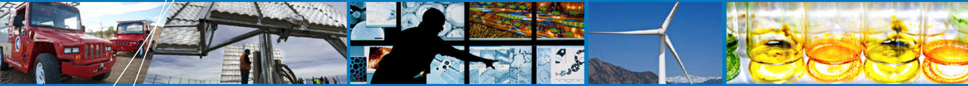
Scenario Generation for Two-Stage Stochastic Economic Dispatch

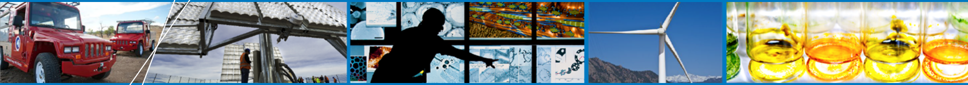


I. Satkauskas, M. Reynolds, D. Sigler, J. Maack, W. Jones

FERC Technical Conference, June 27, 2019

NREL is a national laboratory of the U.S. Department of Energy, Office of Energy Efficiency and Renewable Energy, operated by the Alliance for Sustainable Energy, LLC.

Table of Contents



Problem Formulation

Exascale Computing Project (ECP)

The Exascale Computing Project (ECP) was established to deliver exascale-ready applications and solutions that address currently intractable problems of strategic importance and national interest.

ECP Applications Target National Problems:

- National security, energy security (e.g. ExaWind)
- Scientific discovery, earth system, healthcare
- Economic security
  - ExaSGD: Reliable and efficient planning of the power grid

EXASCALE COMPUTING PROJECT

https://www.exascaleproject.org

Two-stage linear stochastic programming problem

$$\begin{aligned} \min_{\boldsymbol{x}} \quad & \boldsymbol{c}^T \boldsymbol{x} + \mathbb{E}_{\boldsymbol{\xi}}\left[L\left(\boldsymbol{x}, \boldsymbol{\xi}\right)\right] \\ \text{s. t.} \quad & A\boldsymbol{x} \leq \boldsymbol{b} \end{aligned}$$

where $L(\boldsymbol{x}, \boldsymbol{\xi})$ is the optimal value of the second-stage problem,

$$\begin{aligned} \min_{\boldsymbol{y}} \quad & \boldsymbol{q}_{\boldsymbol{\xi}}^T \boldsymbol{y} \\ \text{s. t.} \quad & T_{\boldsymbol{\xi}} \boldsymbol{x} + W_{\boldsymbol{\xi}} \boldsymbol{y} \leq \boldsymbol{h}_{\boldsymbol{\xi}}, \ \boldsymbol{y} \geq 0 \end{aligned}$$

and

- $\boldsymbol{x}$ - first stage decision vector (e.g. thermal dispatch)
- $\boldsymbol{y}$ - second stage decision vector (e.g. wind dispatch, wind spilled, load shed)
- $\boldsymbol{\xi}$ - uncertain data (e.g deviation in wind power from forecast).

Cost functions

First stage:

- thermal generation costs plus expectation of second stage costs

$$\min_{\boldsymbol{x}} \quad \sum_{g \in G} c_g x_g + \mathbb{E}_{\boldsymbol{\xi}} \left[L\left(\boldsymbol{x}, \boldsymbol{\xi}\right)\right]$$

Second stage:

- Wind generation, spilling wind, and slack variable costs (i.e. overload and loss-of-load) at buses $i \in \mathcal{I}$ and wind plants $w \in W$

$$\min_{\boldsymbol{y}^{\pm}, \boldsymbol{\omega}, \boldsymbol{\omega}^{spl}} \quad \sum_{w \in W} \left(c_w \omega_w + c_w^{spl} \omega_w^{spl} \right) \quad + \sum_{i \in \mathcal{I}} \left(c_i^{+} y_i^{+} + c_i^{-} y_i^{-} \right)$$

Constraints

First stage:

- constraints on output of thermal generators $g \in G$

$$x_g^{min} \leq x_g \leq x_g^{max} \quad \forall\, g$$

- ramping constraints on thermal generators

$$-R_g^{down} \leq x_g - l_g \leq R_g^{up} \quad \forall\, g$$

Second stage:

- slack variable (loss-of-load and overload) constraints on buses

$$0 \leq y_i^{\pm} \quad \forall\, i \in \mathcal{I}$$

- wind power constraints

$$\begin{aligned} 0 \leq \omega_w \leq \omega_w^{fcst} + \xi_w \quad \forall\, w \\ \omega_w^{spl} = \left(\omega_w^{fcst} + \xi_w\right) - \left(\omega_w\right) \quad \forall\, w \end{aligned}$$

Constraints continued

- power balance constraints at every node i

$$y_i^+ - y_i^- + \sum_{w \in W_i} \omega_w + \sum_{g \in G_i} x_g = \sum_{q \in D_i} d_q - \sum_{\substack{e=(j,i)\\(j,i)\in\mathcal{E}}} f_e + \sum_{\substack{e=(i,j)\\(i,j)\in\mathcal{E}}} f_e$$

- line flow constraints

$$\underline{F}_e \leq f_e \leq \overline{F}_e \quad \forall\, e \in \mathcal{E}$$

- power flow physics, e.g. ACOPF and DCOPF (below)

$$B_e\left(\theta_i - \theta_j\right) - f_e = 0 \quad \forall\, e \in \mathcal{E}$$

Solving the SAA: matrix structure

$$
\begin{array}{lllllll}
\min\limits_{\boldsymbol{x},\boldsymbol{y}_1,\dots,\boldsymbol{y}_N} & \multicolumn{6}{l}{\left(\boldsymbol{c}^{\mathsf{T}}\boldsymbol{x}\right) + \frac{1}{N}\sum\limits_{s=1}^{N}\left(\boldsymbol{c}_{\xi_s}^{\mathsf{T}}\boldsymbol{y}_s\right)} \\
\text{such that} & A\boldsymbol{x} & & & & & = \boldsymbol{h} \\
& T_{\xi_1}\boldsymbol{x} & + & W_{\xi_1}\boldsymbol{y}_1 & & & = \boldsymbol{h}_1 \\
& T_{\xi_2}\boldsymbol{x} & + & & W_{\xi_2}\boldsymbol{y}_2 & & = \boldsymbol{h}_2 \\
& T_{\xi_3}\boldsymbol{x} & + & & \ddots & & = \vdots \\
& T_{\xi_N}\boldsymbol{x} & + & & & W_{\xi_N}\boldsymbol{y}_N & = \boldsymbol{h}_N \\
& \multicolumn{6}{c}{\boldsymbol{x} \geq 0,\ \boldsymbol{y}_1 \geq 0,\ \boldsymbol{y}_2 \geq 0,\ \dots,\ \boldsymbol{y}_N \geq 0\,.}
\end{array}
$$

There are specialized algorithms for solving these optimization problems:

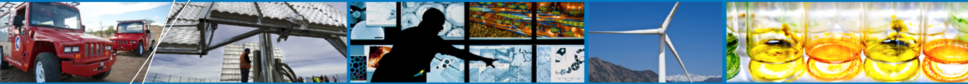
- Schur complement approaches to solving updates in interior point methods, e.g. PIPS[1].
- Progressive hedging algorithm, (e.g. http://www.pyomo.org/)

---

[1] https://github.com/Argonne-National-Laboratory/PIPS

Uncertainty

- In this problem formulation the uncertainty is included by taking $\mathbb{E}_{\boldsymbol{\xi}}\left[L\left(\boldsymbol{x},\boldsymbol{\xi}\right)\right]$, that we approximate by $\frac{1}{N}\sum_{s=1}^{N} L\left(\boldsymbol{x},\boldsymbol{\xi_s}\right)$
- Since we are solving economic dispatch problem for 5-minute periods, $\boldsymbol{\xi}$ represents power deviation from persistence.
- $\boldsymbol{\xi} = (\xi_1, \xi_2, \ldots, \xi_m)$, where ξ_i is power deviation at wind farm i.
- We use high-fidelity data to produce realistic wind scenarios that respect physics and spatio-temporal relations.



Data

WIND Toolkit[2]: domains

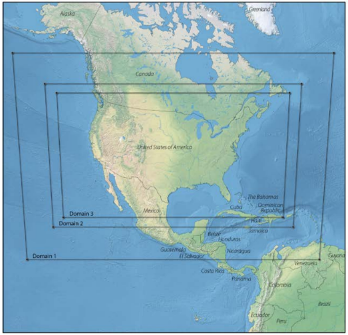


Figure: WRF simulation domains. (from "Overview and Meteorological Validation of the Wind Integration National Dataset Toolkit")

Wind Integration National Dataset (WIND) Toolkit

- Data sets: 2-TB, 50-TB, 0.5 PB
- 2 km x 2 km grid with 20 m vertical resolution.
- 5 min time resolution: pressure, wind speed, direction, humidity, temperature, and density.
- Techno-Economic data set: 5 min time series at 120 000 wind sites.

---

[2] https://www.nrel.gov/grid/wind-toolkit.html

WIND Toolkit: power curves

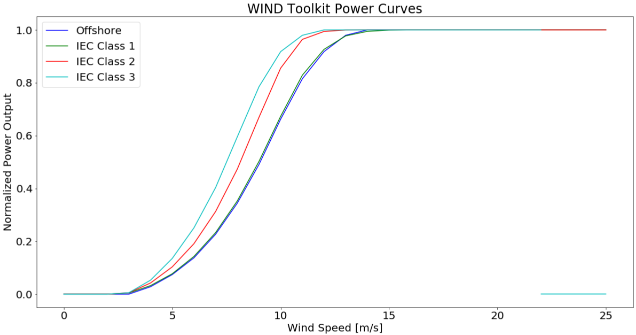


Figure: Power curves used to convert modeled wind speed to power.

- Power curves were chosen according to the estimated long-term wind conditions at each site.
- Each turbine is assumed to have a rated power of 2 MW (at 100 m)

TAMU[3] 200 grid (Illinois)

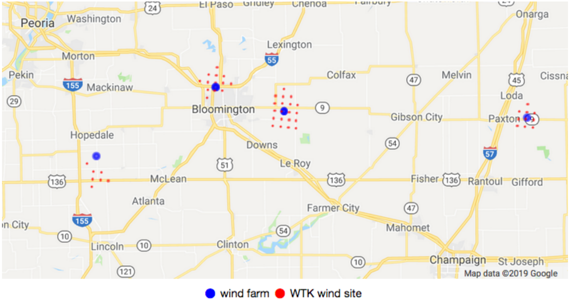


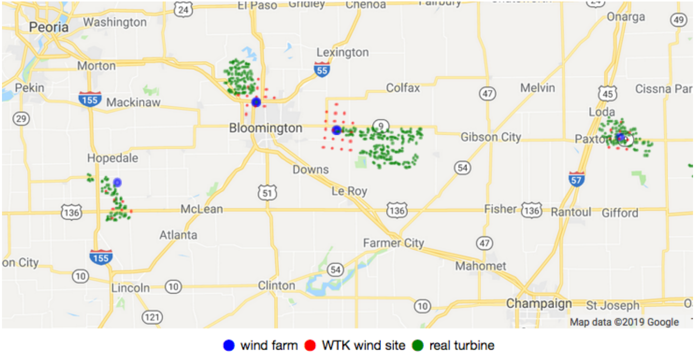
Figure: TAMU 200 bus grid.

- Wind capacity 760 MW (19.42%).
- 6 wind farms, 50 wind sites (NREL Wind Toolkit).

---

[3] https://electricgrids.engr.tamu.edu/electric-grid-test-cases/

TAMU 200 grid

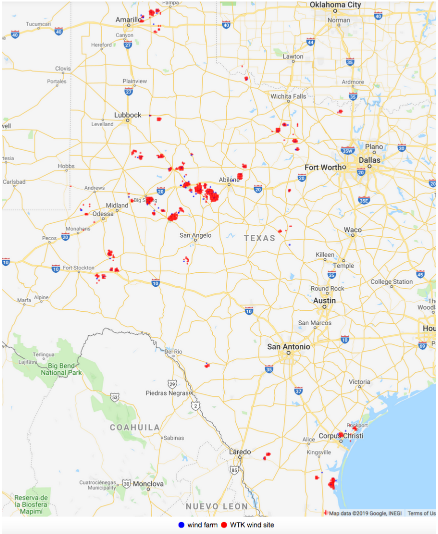


- Wind capacity 760 MW (19.42%).
- 6 wind farms, 50 wind sites (NREL Wind Toolkit).
- 437 real turbines (USWTDB[4]).

---

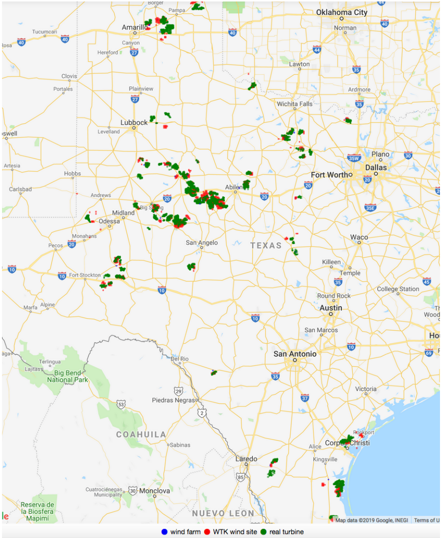
[4] https://eerscmap.usgs.gov/uswtdb/

TAMU 2000 grid (Texas)



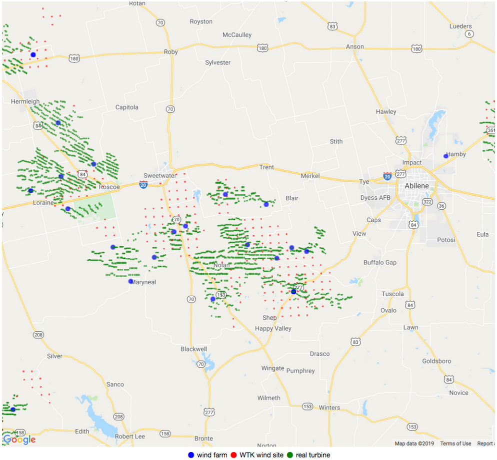
- 2000 buses
- Wind capacity 12574 MW (12.56%).
- 87 wind farms, 859 wind sites (NREL Wind Toolkit).

TAMU 2000 grid



- 2000 buses
- Wind capacity 12574 MW (12.56%).
- 87 wind farms, 859 wind sites (NREL Wind Toolkit).
- 7312 real turbines (USWTDB).

TAMU 2000 grid



- 2000 buses
- Wind capacity 12574 MW (12.56%).
- 87 wind farms, 859 wind sites (NREL Wind Toolkit).
- 7312 real turbines (USWTDB).

Techno-Economic data set

- 7 years (2007-2013) of power output data at 5-min resolution.
- One year we leave as “actuals” for experiments; 6 years are used to generate scenarios $\boldsymbol{\xi}$.
- For each wind farm on the grid we aggregate WIND Toolkit wind sites up to farm’s capacity.
- Additionally, we split our data set based on total wind power in the network: low, medium, high.

TAMU 200 grid

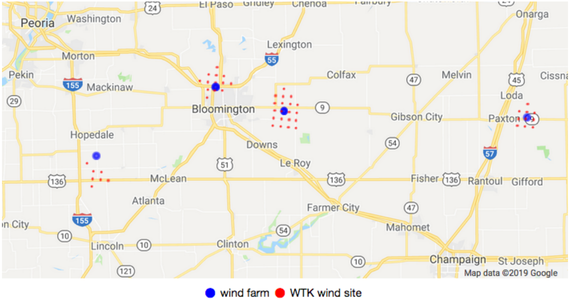


Figure: TAMU 200 bus grid.

- Power (760 MW) conditioning splits data into 3 sets:

1. $low < 11MW$
2. $11MW < medium < 717MW$
3. $high > 717MW$

Distribution of deviations: TAMU 200

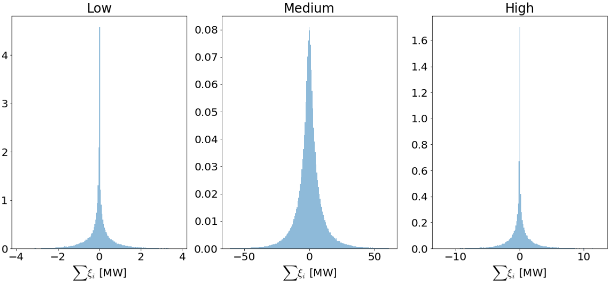
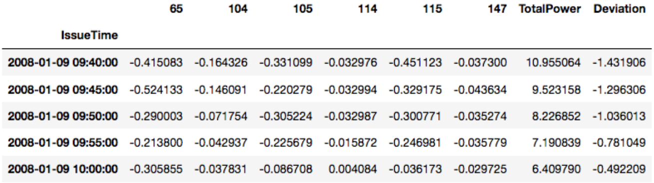


Figure: Power conditioned distributions of deviations

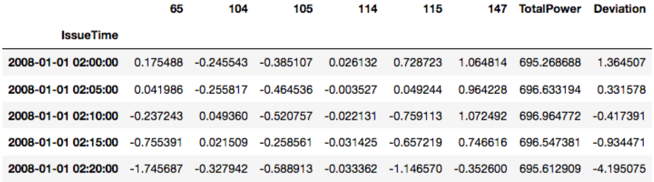
.

Examples of scenario tables for TAMU 200 grid



	65	104	105	114	115	147	TotalPower	Deviation
IssueTime								
2008-01-09 09:40:00	-0.415083	-0.164326	-0.331099	-0.032976	-0.451123	-0.037300	10.955064	-1.431906
2008-01-09 09:45:00	-0.524133	-0.146091	-0.220279	-0.032994	-0.329175	-0.043634	9.523158	-1.296306
2008-01-09 09:50:00	-0.290003	-0.071754	-0.305224	-0.032987	-0.300771	-0.035274	8.226852	-1.036013
2008-01-09 09:55:00	-0.213800	-0.042937	-0.225679	-0.015872	-0.246981	-0.035779	7.190839	-0.781049
2008-01-09 10:00:00	-0.305855	-0.037831	-0.086708	0.004084	-0.036173	-0.029725	6.409790	-0.492209

Table: Low power



	65	104	105	114	115	147	TotalPower	Deviation
IssueTime								
2008-01-01 02:00:00	0.175488	-0.245543	-0.385107	0.026132	0.728723	1.064814	695.268688	1.364507
2008-01-01 02:05:00	0.041986	-0.255817	-0.464536	-0.003527	0.049244	0.964228	696.633194	0.331578
2008-01-01 02:10:00	-0.237243	0.049360	-0.520757	-0.022131	-0.759113	1.072492	696.964772	-0.417391
2008-01-01 02:15:00	-0.755391	0.021509	-0.258561	-0.031425	-0.657219	0.746616	696.547381	-0.934471
2008-01-01 02:20:00	-1.745687	-0.327942	-0.588913	-0.033362	-1.146570	-0.352600	695.612909	-4.195075

Table: Medium power

Distribution of deviations: TAMU 200

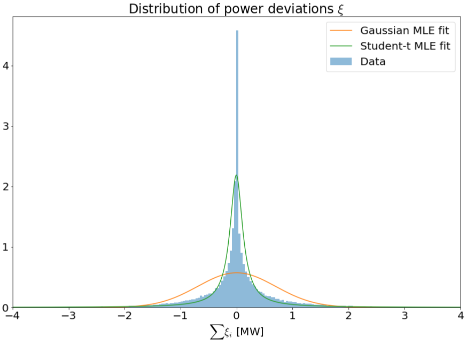


Figure: Low power deviations

Distribution of deviations: TAMU 2000

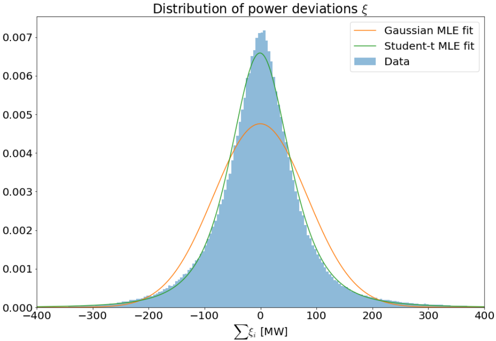
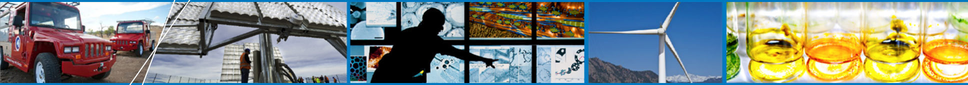


Figure: Medium power deviations

.



Importance Sampling

Basic idea

Importance sampling, is based on the following idea:

$$\mathbb{E}_p\left[L\left(\xi\right)\right] = \int L\left(\xi\right)p\left(\xi\right)d\xi = \int L\left(\xi\right)\frac{p\left(\xi\right)}{q\left(\xi\right)}q(\xi)d\xi = \mathbb{E}_q\left[L\left(\xi\right)R\left(\xi\right)\right],$$

where $R\left(\xi\right) = p\left(\xi\right)/q\left(\xi\right)$, $p\left(\xi\right)$ is called nominal and $q\left(\xi\right)$ is called importance distribution.

Basic idea

Low fidelity approach: surrogate model for $L(\boldsymbol{x}, \boldsymbol{\xi})$ using loss-of-load and costs of spilling wind

$$\tilde{L}(\boldsymbol{\xi}) = \begin{cases} c^{-} \sum_w \xi_w & \sum_w \xi_w > 0 \\ c^{+} \sum_w \xi_w & \sum_w \xi_w \leq 0 \end{cases}$$

where c^{-} is cost of spilling wind, and c^{+} is the cost of loss-of-load.

Nominal and importance distribution

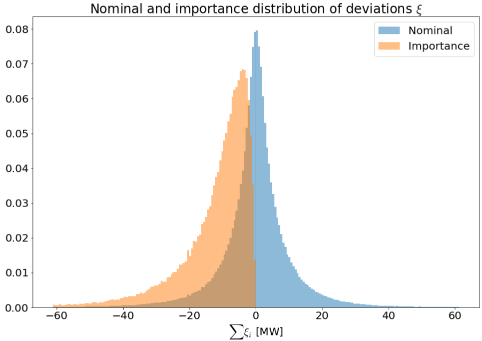
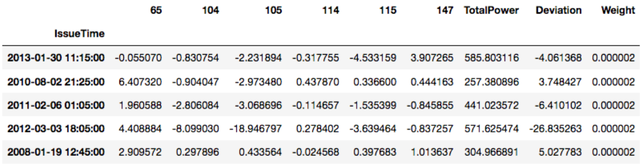


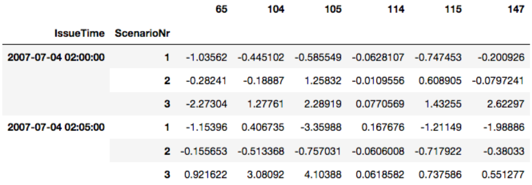
Figure: Histograms of deviations (TAMU 200 grid, medium power)

Example scenarios: nominal



	65	104	105	114	115	147	TotalPower	Deviation	Weight
IssueTime									
2013-01-30 11:15:00	-0.055070	-0.830754	-2.231894	-0.317755	-4.533159	3.907265	585.803116	-4.061368	0.000002
2010-08-02 21:25:00	6.407320	-0.904047	-2.973480	0.437870	0.336600	0.444163	257.380896	3.748427	0.000002
2011-02-06 01:05:00	1.960588	-2.806084	-3.068696	-0.114657	-1.535399	-0.845855	441.023572	-6.410102	0.000002
2012-03-03 18:05:00	4.408884	-8.099030	-18.946797	0.278402	-3.639464	-0.837257	571.625474	-26.835263	0.000002
2008-01-19 12:45:00	2.909572	0.297896	0.433564	-0.024568	0.397683	1.013637	304.966891	5.027783	0.000002

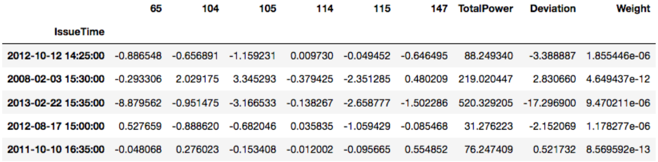
Figure: TAMU 200 medium power scenarios



		65	104	105	114	115	147
IssueTime	**ScenarioNr**						
2007-07-04 02:00:00	**1**	-1.03562	-0.445102	-0.585549	-0.0628107	-0.747453	-0.200926
	2	-0.28241	-0.18887	1.25832	-0.0109556	0.608905	-0.0797241
	3	-2.27304	1.27761	2.28919	0.0770569	1.43255	2.62297
2007-07-04 02:05:00	**1**	-1.15396	0.406735	-3.35988	0.167676	-1.21149	-1.98886
	2	-0.155653	-0.513368	-0.757031	-0.0606008	-0.717922	-0.38033
	3	0.921622	3.08092	4.10388	0.0618582	0.737586	0.551277

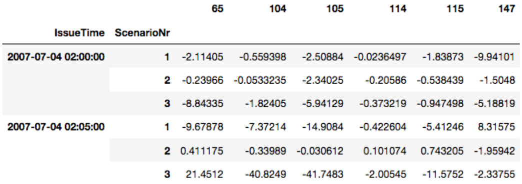
Figure: TAMU 200 scenrios drawn from nominal distribution

Example scenarios: importance



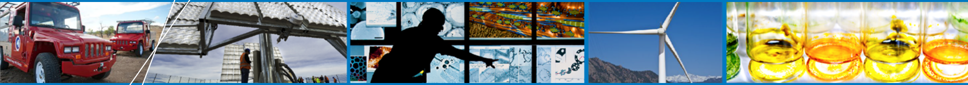
IssueTime	65	104	105	114	115	147	TotalPower	Deviation	Weight
2012-10-12 14:25:00	-0.886548	-0.656891	-1.159231	0.009730	-0.049452	-0.646495	88.249340	-3.388887	1.855446e-06
2008-02-03 15:30:00	-0.293306	2.029175	3.345293	-0.379425	-2.351285	0.480209	219.020447	2.830660	4.649437e-12
2013-02-22 15:35:00	-8.879562	-0.951475	-3.166533	-0.138267	-2.658777	-1.502286	520.329205	-17.296900	9.470211e-06
2012-08-17 15:00:00	0.527659	-0.888620	-0.682046	0.035835	-1.059429	-0.085468	31.276223	-2.152069	1.178277e-06
2011-10-10 16:35:00	-0.048068	0.276023	-0.153408	-0.012002	-0.095665	0.554852	76.247409	0.521732	8.569592e-13

Figure: TAMU 200 medium power table



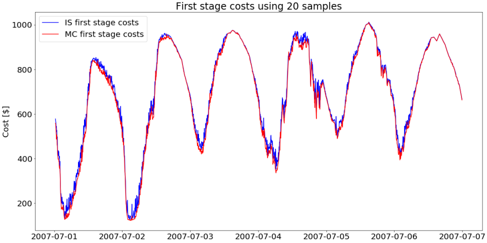
IssueTime	ScenarioNr	65	104	105	114	115	147
2007-07-04 02:00:00	**1**	-2.11405	-0.559398	-2.50884	-0.0236497	-1.83873	-9.94101
	2	-0.23966	-0.0533235	-2.34025	-0.20586	-0.538439	-1.5048
	3	-8.84335	-1.82405	-5.94129	-0.373219	-0.947498	-5.18819
2007-07-04 02:05:00	**1**	-9.67878	-7.37214	-14.9084	-0.422604	-5.41246	8.31575
	2	0.411175	-0.33989	-0.030612	0.101074	0.743205	-1.95942
	3	21.4512	-40.8249	-41.7483	-2.00545	-11.5752	-2.33755

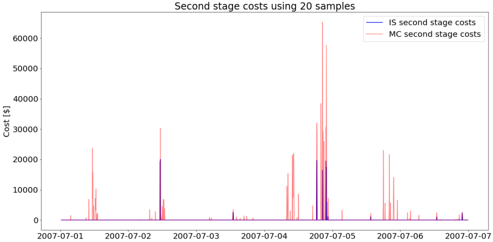
Figure: TAMU 200 scenrios drawn from importance distribution



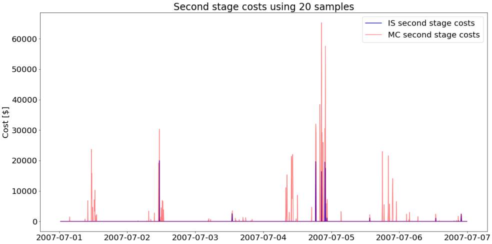
Experimental Results

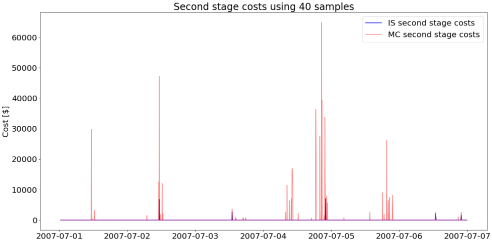
One week economic dispatch for TAMU 200 grid



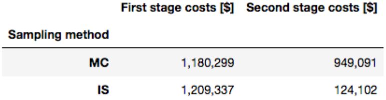


One week economic dispatch for TAMU 200 grid



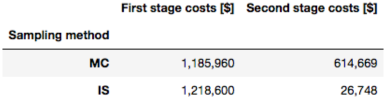


One week economic dispatch for TAMU 200 grid



Sampling method	First stage costs [$]	Second stage costs [$]
MC	1,180,299	949,091
IS	1,209,337	124,102

Figure: 20 samples



Sampling method	First stage costs [$]	Second stage costs [$]
MC	1,185,960	614,669
IS	1,218,600	26,748

Figure: 40 samples

Future work

- High fidelity surrogate model for $L(\boldsymbol{x}, \boldsymbol{\xi})$.
- Extend to multi-period scenarios.
- More realistic physics, e.g. ACOPF.
- Extending toolset to data sets other than WIND toolkit.

Thank you. Questions?

- Exascale Project https://www.exascaleproject.org/project/exasgd-optimizing-stochastic-grid-dynamics-at-exascale/
- WIND Toolkit https://www.nrel.gov/grid/wind-toolkit.html
- TAMU Sythetic Grids https://electricgrids.engr.tamu.edu/electric-grid-test-cases/
- USWTDB https://eerscmap.usgs.gov/uswtdb/