



Uncertainty Quantification for Capacity Expansion Planning

Paul Diaz, Ryan King, Devon Sigler, Wesley Cole,
and Wesley Jones

National Renewable Energy Laboratory

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

This report is available at no cost from the National Renewable Energy Laboratory (NREL) at www.nrel.gov/publications.

Contract No. DE-AC36-08GO28308

**Technical Report
NREL/TP-2C00-76708
June 2020**



Uncertainty Quantification for Capacity Expansion Planning

Paul Diaz, Ryan King, Devon Sigler, Wesley Cole,
and Wesley Jones

National Renewable Energy Laboratory

Suggested Citation

Diaz, Paul, Ryan King, Devon Sigler, Wesley Cole, and Wesley Jones. 2020. *Uncertainty Quantification for Capacity Expansion Planning* Golden, CO: National Renewable Energy Laboratory. NREL/TP-2C00-76708. <https://www.nrel.gov/docs/fy20osti/76708.pdf>

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

This report is available at no cost from the National Renewable Energy Laboratory (NREL) at www.nrel.gov/publications.

Contract No. DE-AC36-08GO28308

Technical Report
NREL/TP-2C00-76708
June 2020

National Renewable Energy Laboratory
15013 Denver West Parkway
Golden, CO 80401
303-275-3000 • www.nrel.gov

NOTICE

This work was authored in part by the National Renewable Energy Laboratory, operated by Alliance for Sustainable Energy, LLC, for the U.S. Department of Energy (DOE) under Contract No. DE-AC36-08GO28308. Funding provided by U.S. Department of Energy Office of Energy Efficiency and Renewable Energy Wind Energy Technologies Office. The views expressed herein do not necessarily represent the views of the DOE or the U.S. Government. This research was performed using computational resources sponsored by the Department of Energy's Office of Energy Efficiency and Renewable Energy and located at the National Renewable Energy Laboratory.

This report is available at no cost from the National Renewable Energy Laboratory (NREL) at www.nrel.gov/publications.

U.S. Department of Energy (DOE) reports produced after 1991 and a growing number of pre-1991 documents are available free via www.osti.gov.

Cover Photo by Lee Jay Fingersh: NREL 35721.

NREL prints on paper that contains recycled content.

Executive Summary

This report quantifies the uncertainty in output decisions from a Capacity Expansion Planning (CEP) model. The need to understand how uncertainties within CEP models and modeling assumptions affect Quantities of Interest (QoIs) such as expansion and operating costs, as well as expansion decisions remains an ongoing challenge in scientific research and industrial operations. This area of research is particularly important for models which seek to capture how large networks will evolve and operate under increased sources of variable generation, i.e., higher penetration of renewable technologies such as solar and wind generators. Uncertainty quantification (UQ) of CEP models which estimate expansion costs and decisions, and *production cost models* which estimate operating costs and dispatch decisions, is a key focus of research at NREL. The Regional Energy Deployment System (ReEDS) represents a state-of-the-art CEP model and considers a range of possible grid evolutions in an attempt to identify key drivers, ramifications, and decisions which contribute to better informed investment and policy decisions. However, research to quantify how uncertainties and model assumptions, such as unit commitment (UC), within ReEDS may be affecting its outputs remains challenging due to the size and complexity of the model (Cohen et al. 2019; Murphy et al. 2019, Appendix A).

The Scalable Power-System Economic Expansion Dispatch (SPEED) model is similar to ReEDS and is used in this work as a stochastic approach to study CEP and production cost modeling (PCM). Simulations yield expansion decisions in the form of either natural gas or wind generators, of different sizes, at different locations, over a ten year period, using data for a hypothetical electric grid overlapping parts of California, Nevada, and Arizona. Relevant QoIs are identified as 1) expansion cost, 2) operations cost, 3) maximum installed gas capacity, and 4) maximum installed wind capacity, resulting from any given SPEED simulation. The key uncertain model input parameters we study are capacity reserve margin for expansion, cost of loss of load, cost of excess load, natural gas price, wind installation cost, and transmission capacity. Simulations are performed utilizing the modified Institute of Electrical and Electronics Engineers' (IEEE) Reliability Test System (RTS) provided by the Grid Modernization Lab Consortium (GMLC). The resulting data is post-processed, and two different approaches are considered to quantify the uncertainty in the model including sparse Polynomial Chaos Expansions (PCEs) and Active Subspace analysis. Five global sensitivity metrics, which explicitly quantify uncertainty by measuring an input parameter's influence on the variance of the QoIs and provide a measure of explainable uncertainty, are reported. PCE surrogate models were constructed and exploited to generate rich posterior distributions of the SPEED model output QoIs. To better understand the uncertainty associated with associated UC, we conducted two independent numerical experiments holding all modeling conditions equal except for the integer modeling assumption regarding the dispatch decision variables, which has a significant impact on the model complexity. Our results support the following conclusions:

- The global sensitivity metrics indicate that all four QoIs are relatively insensitive to the input parameters cost of loss of load, cost of excess load, and cost of natural gas, while they are sensitive to reserve capacity margin, cost of wind, and transmission capacity but in different ways. See Sections 4.1 and 4.2.
- The global sensitivity metrics indicate that all four QoIs are influenced by cost of loss of load, cost of excess load, and cost of natural gas only through their interactions with other terms, if at all.
- Among the 6 uncertain inputs considered, expansion cost in this model is driven primarily by cost of wind and secondarily by reserve capacity margin.
- Among the 6 uncertain inputs considered, operation cost in this model is driven primarily by cost of wind, secondarily by transmission availability, and slightly by reserve capacity margin.
- Among the 6 uncertain inputs considered, maximum installed gas capacity in this model is driven primarily by reserve capacity margin, secondarily by cost of wind, and slightly by transmission availability.
- Among the 6 uncertain inputs considered, maximum installed wind capacity in this model is driven primarily by the cost of wind and secondarily by reserve capacity margin.
- When considering relaxed vs. binary UC, the posterior distributions of the expansions cost, operations cost, and maximum installed gas capacity predicted by the PC surrogates are similar, while the distribution of maximum installed wind capacity differs significantly. See Section 4.3.

- The mean values of each QoI, approximated by the PCE surrogates, are similar for both relaxed and binary UC. However, we report larger standard deviations in the distributions of expansion cost, operations cost, and max installed wind capacity for relaxed UC compared to the binary UC model assumption. In contrast, the standard deviation for max installed gas capacity was smaller for relaxed UC compared to binary.
- The distribution of max installed wind for relaxed UC is bimodal and right-skewed, while the distribution for binary UC shows no apparent skew or multiple modes, potentially indicating that by relaxing the UC dispatch decision variables, CEP models could be significantly underestimating the amount of installed wind capacity required to satisfy design constraints, e.g., transmission guidelines or resource adequacy constraints such as reserve capacity margin.
- Prior work has demonstrated that relaxing binary UC variables in PCMs alone does not necessarily reduce computational burden, contrary to expectations larger binary decision trees improved UC model resolution, and that results are likely solver dependent (Alemany, Kasprzyk, and Magnago 2018). Our results show significant reduction in the computational burden of the SPEED model by relaxing binary UC variables, see Section 4.3.

The implementation of the SPEED model in this study was relatively narrow both in the geographic location of the RTMS-GMLC data set within the U.S., but also in the greater scheme of CEP research. This study did not consider modeling the expansion of battery storage technologies or other variable generating technologies such as solar PV and hydro due to the limited scope of analysis. In order to meet the growing demand for the UQ and prediction capabilities of CEP and PCM models we highlight the importance of continuing to study state-of-the-art methods in stochastic programming, data driven modeling, and UQ specifically to improve or better inform both CEP and PCM frameworks. As we discuss in Section 5, a collaborative research effort between modelers, uncertainty analysts, and domain scientists is necessary to better understand the fundamental shift predicted in the future from traditional thermal generation to a more diverse U.S. electric grid that is more dependent on variable generation technologies and renewables (Cole et al. 2020; Murphy et al. 2019). This research effort is particularly important regarding efforts to scale up existing modeling approaches to higher temporal or spatial fidelities, as identifying non-sensitive model parameters can greatly reduce model complexity and allow scientists and engineers to study other meaningful uncertainties or physical processes within a model.

Acknowledgments

We want to thank all of the folks who either directly or indirectly contributed to this work through conversations, meetings, and feedback. Special thanks Karen Stengel (NREL) and Andrew Glaws (NREL) for their continual feedback and input over the course of this project, to Caroline Draxl (NREL) for inviting us to present these findings at the National Wind Technology Center (NWTC), and to Prof. Alireza Doostan (CU Boulder) for his expert feedback. We also thank Max Brown for providing sources and documentation on the ReEDS model, and more generally the entire ReEDs modeling and analysis team for the foundation of research they have provided in this area. This work was authored in part by the National Renewable Energy Laboratory, operated by Alliance for Sustainable Energy, LLC, for the U.S. Department of Energy (DOE) under Contract No. DE-AC36-08GO28308, and in part by the University of Colorado, Boulder, Department of Aerospace Engineering Sciences for the U.S. Department of Education under award No. P200A180014.

Acronym List

CC	Combined Cycle
CEP	Capacity Expansion Planning
CT	Combustion Turbine
DOE	U.S. Department of Energy
DGSM	Derivative-based Global Sensitivity Metric
EIA	U.S. Energy Information Administration
GMLC	Grid Modernization Laboratory Consortium
HPC	High Performance Computer
IEEE	Institute of Electrical and Electronics Engineers
LP	Linear Program
LCOE	Levelized Cost of Energy
MILP	Mixed-Integer Linear Program
MMBTU	Million British Thermal Unit
MWh	Mega-Watt Hour
NREL	National Renewable Energy Laboratory
NWTC	National Wind Technology Center
PC	Polynomial Chaos
PV	Photovoltaic
QoI	Quantity of Interest
ReEDS	Regional Energy Deployment System
RTS	Reliability Test System
UQ	Uncertainty Quantification
UC	Unit Commitment
SPEED	Scalable Power-System Economic Expansion Dispatch

Table of Contents

1	Introduction	2
1.1	CEP with the Regional Energy Deployment System	2
1.2	Stochastic Programming for Capacity Expansion Planning	4
1.3	Contributions of this research	8
2	The RTS-GMLC Data Set	8
3	Methods for Uncertainty Quantification	10
3.1	Polynomial Chaos Expansions	10
3.2	Active Subspaces	14
4	Results and Interpretation	16
4.1	Relaxed UC Global Sensitivity Analysis	18
4.2	Binary UC Global Sensitivity Analysis	23
4.3	Exploiting the PC surrogates	26
5	Conclusions and Future Work	28

List of Figures

Figure 1.	An example output of the ReEDS model.	3
Figure 2.	A schematic of the ReEDS model structure.	4
Figure 3.	A stereotypical L-curve.	13
Figure 4.	SPEED model QoIs with relaxed and binary UC.	17
Figure 5.	Global sensitivity analysis for the SPEED model expansion cost with relaxed UC.	19
Figure 6.	Global sensitivity analysis for the SPEED model operations cost with relaxed UC.	20
Figure 7.	Global sensitivity analysis for the SPEED model installed gas capacity with relaxed UC.	21
Figure 8.	Global sensitivity analysis for the SPEED model installed wind capacity with relaxed UC.	22
Figure 9.	Global sensitivity analysis for the SPEED model expansion cost with binary UC.	23
Figure 10.	Global sensitivity analysis for the SPEED model operations cost with binary UC.	24
Figure 11.	Global sensitivity analysis for the SPEED model installed gas capacity with binary UC.	25
Figure 12.	Global sensitivity analysis for the SPEED model installed wind capacity with binary UC.	26
Figure 13.	SPEED model QoIs approximated by the PC surrogate models.	27

List of Tables

Table 1.	SPEED Model Parameters	5
Table 2.	SPEED Model Uncertain Input Parameters	9
Table 3.	SPEED Model Quantities of Interest	9
Table 4.	Single index, multi-index, and tensored polynomials for $(d, p) = (3, 2)$.	11
Table 5.	Four standard types of orthogonal polynomials according to the Wiener-Askey Polynomial Scheme.	11

Table 6. The validated point-wise relative residual fit, measures of complexity, and QoI statistics of the PCE surrogate models for relaxed UC.	18
Table 7. The validated point-wise relative residual fit, measures of complexity, and QoI statistics of the PCE surrogate models for binary UC.	18

1 Introduction

Imagine that you are in charge of planning operations as an electrical service provider. You would like to expand your ability to meet electricity demand by building new types of generation technologies, but you are not sure what comprises the best portfolio of generators. Moreover, there is uncertainty in the market, for example in the form of the total amount of generating capacity required on the electric grid, regulatory considerations, or in the installation and operational costs of the available generating technologies. The problem you are faced with is referred to as a Capacity Expansion Planning (CEP) problem, and within the context of Uncertainty Quantification (UQ), it is classified as decision making under uncertainty.

CEP is a resource allocation problem in which decisions are made and represented as a binary variable. Generally speaking, there are two types of decision variables associated with *expansion* and *dispatch* decisions, and for each decision there are costs associated with installing new generating capacity and dispatching generators. Within energy studies, CEP often refers to the planning of expansion decisions, which represent different types, sizes, and locations of electricity generating sources to meet demand. Industry professionals, scientists, researchers, and engineers use CEP in order to make decisions regarding what technologies, e.g., wind, solar, natural gas, nuclear, etc., to build in order to meet demand and satisfy regulatory or reliability conditions. In contrast, Production Cost Modeling (PCM) involves simulating the operation of an electric grid and focuses on dispatch decisions. These types of problems are important for designing reliable and robust electric grids, and their use can have significant and lasting real world impacts. Often, computational models are formed of an electric grid and an optimization problem is solved which seeks to minimize the costs of expanding the system.

In this work, we numerically study an electric grid whose expansions costs are on the order of hundreds of millions of dollars (\$ U.S.D.) and operations costs are on the order of tens of millions of dollars. It is of interest to industry professionals, the National Renewable Energy Lab (NREL), and more broadly the Department of Energy (DoE), to understand how uncertainties in the inputs of a CEP model may affect the resulting overall build decisions. It is also of interest to understand how modeling assumptions within common CEP models may be affecting the resulting overall build decisions, in particular the expansion decisions associated with renewable energy technologies such as wind turbines and solar photovoltaic (PV) generators. Further complicating matters, CEP problems may become more difficult in the future due to the fact that many renewable energy sources are variable generators which depend on environmental factors that impact wind and solar availability. The variable nature of these technologies has associated uncertainty that is not present in thermal generating technologies, for example natural gas, coal, and nuclear. Given this variable uncertainty in generating technologies, a future shift from traditional thermal technologies to renewables may cause the overall uncertainty within CEP models to grow in the future. This observation motivates our work, and highlights that studying the uncertainty will remain necessary for building reliable and efficient electric grids.

1.1 CEP with the Regional Energy Deployment System

This work focuses significantly on electrical CEP and PCM and is closely related to NREL’s Regional Energy Deployment System (ReEDS) model. While the research in this report does not directly involve simulation with the ReEDS model, we mention it to serve as an introduction to state-of-the-art capacity expansion planning. ReEDS is NREL’s flagship model for long-term power sector analysis and enables research regarding clean energy policy, renewable energy integration, technology development, and issues related to future generation and transmission infrastructure. Specifically, ReEDS models the contiguous United States electric power sector and relies on system-wide cost optimization in order to estimate the types and locations of future generation and transmission capacity (Eurek et al. 2016).

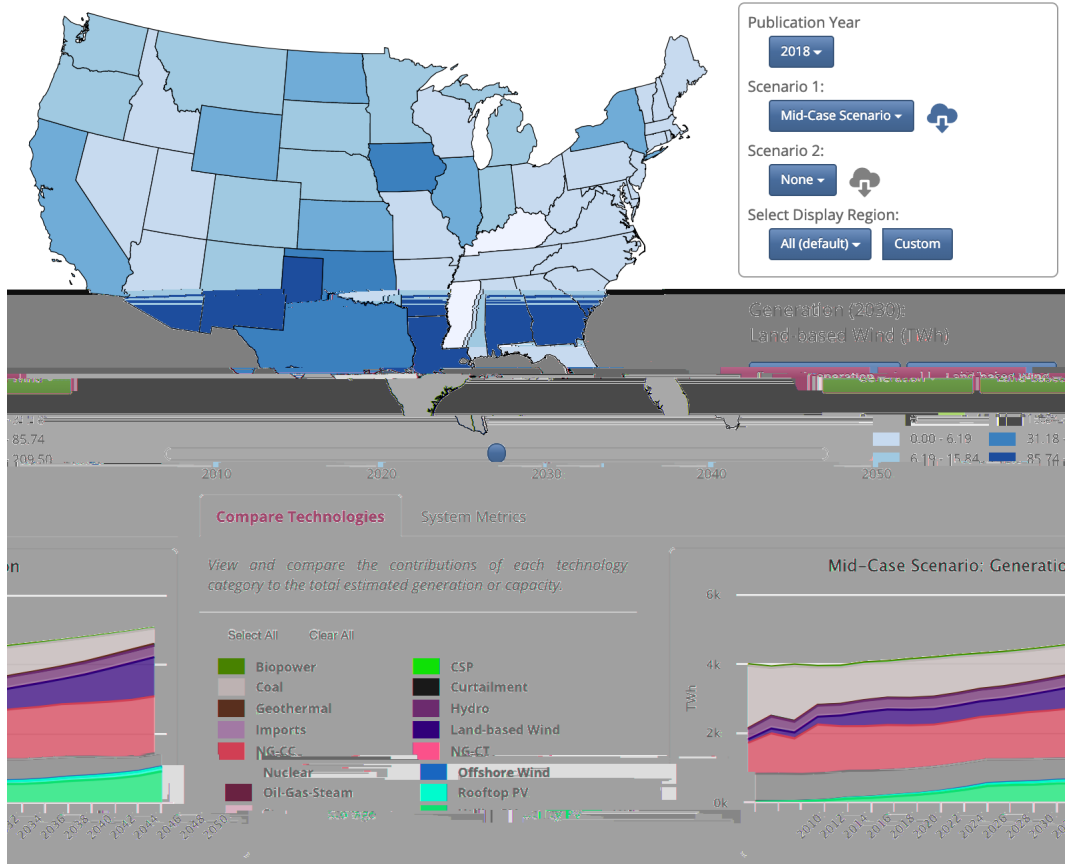


Figure 1. An example output of the ReEDS model (Cohen et al. 2019).

Qualitatively, ReEDS identifies a cost-optimal mix of technologies which satisfy the regional power demand, subject to grid reliability (or reserve) requirements, technology resource constraints, and policy constraints. A cost minimization problem is solved for 21 two-year periods from 2010 to 2050. The main outputs of the ReEDS model include the location of generator capacity and annual generation from each technology, storage capacity expansion, transmission capacity expansion, total electric sector costs, electricity price, fuel demand and prices, and carbon dioxide (CO₂) emissions. A graphical result of the ReEDS model is shown in Figure 1, this image was generated using the ReEDS standard scenario viewer publicly available at <https://openei.org/apps/reeds/>.

Quantitatively, ReEDS is a recursive-dynamic model that for each scenario (or year) $s \in S$, solves a linear program of the form

$$\text{minimize } \tilde{\mathbf{c}}^T \tilde{\mathbf{x}}_s \quad \text{subject to } \tilde{\mathbf{x}}_s \in \mathcal{Q}_s, \quad (1)$$

where $\tilde{\mathbf{x}}_s \in \mathbb{R}^n$ is a decision vector, $\tilde{\mathbf{c}} \in \mathbb{R}^n$ is a cost coefficient vector, and the requirement $\tilde{\mathbf{x}}_s \in \mathcal{Q}_s$ expresses the problem constraints ensuring $\tilde{\mathbf{x}}_s$ is a feasible solution. A schematic of the model structure is shown in Figure 2. ReEDS solves 21 = |S| individual, but interacting, linear programming optimization problems. Each optimization problem minimizes both capital and operating costs for the U.S. electric grid subject to a number of constraints for a two-year period. These constraints fall into the following categories: load constraints, planning reserve constraints, operating reserve constraints, transmission constraints, resource constraints, emissions constraints, and policy constraints.

For each year of the ReEDS model, 17 time-slices are used to represent the 8760 hours of the year. Each of the four seasons is modeled by a representative day of four time-slices: overnight, morning, afternoon, and evening. The 17th

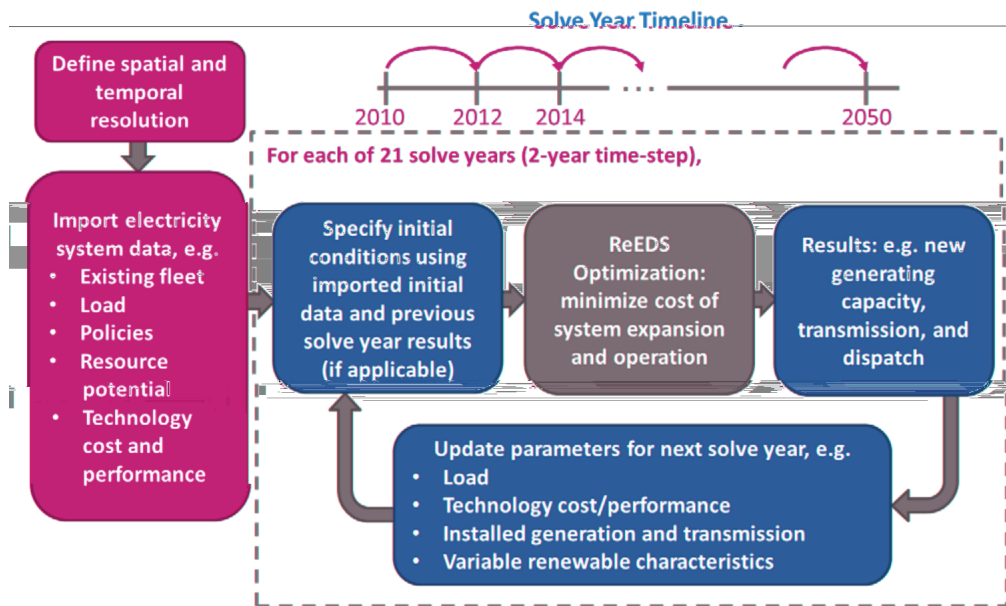


Figure 2. A schematic of the ReEDS model structure (Cohen et al. 2019).

time-slice represents a summer “superpeak” representing the top 40 hours of summer load (Eurek et al. 2016; Cohen et al. 2019). The modeled load is averaged over all hours represented by each time-slice. While this time schedule allows for seasonal and diurnal variations in demand, wind, and solar profiles, it cannot address shorter timescale challenges associated with unit commitment (UC)¹ and economic dispatch, particularly under circumstances with high penetration of variable generation (e.g. wind and solar). However, efforts have been made to inform the dispatch order associated with the 17 timeslices by building hourly modules that define ReEDS parametric constraints, and the resulting decisions were post-processed with PLEXOS PCM to investigate unserved load and reserve violations, which are indicators of grid reliability and resource adequacy (Frew et al. 2019).

Efforts have also been made to study uncertainty in the ReEDS model, but due to the complexity and scale of the model uncertainty analysis is typically performed through scenario analysis (Cole et al. 2018). The standard scenarios² used to study the ReEDS model are a relatively small collection of forward-looking projections of the U.S. power section which are used to identify trends between inputs and expansion decisions (outputs). The standard scenarios consider input factors like electricity demand growth, fuel prices, financing assumptions, model foresight, expansion costs for generating technologies, resource and system constraints, and existing fleet retirements (Cole et al. 2020). While scenario analysis provides some insights for such a complex model, efforts to quantify uncertainty globally by varying all factors of uncertainty *simultaneously* have received comparably less attention. One approach that allows for more freedom in modeling uncertainty in these types of problems involves stochastic programming.

1.2 Stochastic Programming for Capacity Expansion Planning

The *Scalable Power-System Economic Expansion Dispatch* (SPEED) model is a similar model to ReEDS which may be used to study CEP and production cost modeling. This model has several advantages compared to ReEDS including a simple scheme for data storage and transfer, a streamlined approach for new scenario generation, a well-suited parallel computing implementation compatible with the Eagle HPC at NREL, and significantly improved

¹Unit commitment (UC) refers to the scheduling of generating units such that total operating cost is minimized.

²A standard ReEDs scenario should not to be confused with s . The scenario s represents a year’s worth of data, and it provides time-series data for input parameters such as wind availability, solar availability, and electricity demand as well as regulatory or reliability constraints associated with factors like emissions, transmission, and safety margins. These input parameters are used to make decisions by solving the linear program (1). The effect of changing from one standard scenario to the next can be interpreted as using different sets S .

flexibility of time series data sampling. The last advantage is of particular interest. Because the SPEED model enables a greater degree of fidelity in time compared to the ReEDS model, it enables the study of smaller time scales associated with UC and economic dispatch scenarios with high penetration of renewables.

With perfect knowledge of the future a decision maker would simply make decisions according to $\tilde{\mathbf{x}}_s^*$ that is an optimal solution to (1) for the scenario s which they knew would be realized. Realistically, decision makers must decide on capacity expansion plans without *a priori* knowledge which scenario will be ultimately realized. This fact motivates the use of an optimization model which possess a mechanism for dealing with such uncertainty.

Let $\mathbb{P}[s]$ denote the probability of the occurrence of each scenario $s \in S$. Such probabilities enable a model which takes into account prior knowledge of the distribution of individual scenarios, or in the case of CEP to weight the relative importance of particular scenarios based on problem specific knowledge. Under this probabilistic framework, decisions can be made according to the stochastic optimization problem

$$\text{minimize } \tilde{\mathbf{c}}^T \tilde{\mathbf{x}} + \sum_{s \in S} \mathbb{P}[s] (\mathbf{f}_s^T \mathbf{y}_s) \quad \text{subject to } (\tilde{\mathbf{x}}, \mathbf{y}_s) \in \mathcal{Q} \quad \forall s \in S, \quad (2)$$

where the decision vector $\tilde{\mathbf{x}}$ (with $\tilde{\mathbf{x}}_s = \tilde{\mathbf{x}}, \forall s \in S$) does not allow decisions to depend on the scenario s through *non-anticipativity* constraints (Watson, Woodruff, and Strip 2008). In this problem, \mathbf{y}_s represents scenario-specific decision associated with costs \mathbf{f}_s , which are determined by $\tilde{\mathbf{x}}$ and a particular scenario $s \in S$. Below is a complete definition of the SPEED model parameters and stochastic optimization problem associated with CEP.

Table 1. SPEED Model Parameters

G	Set of thermal generator types which exist on the system	E^{cap}	Minimum capacity expansion required for the system
G_{new}	Set of thermal generator types which can be added to the system	$n_g^{therm,min}$	Minimum expansion number of thermal generators of type g
R	Set of renewable generator types	$n_r^{renew,min}$	Minimum expansion number of renewable generators of type r
R_{new}	Set of renewable generator types which can be added to the system	$n_g^{therm,max}$	Maximum expansion number of thermal generators of type g
T	Set of time steps	$n_r^{renew,max}$	Maximum expansion number of renewable generators of type r
D	Set of load buses	p_g^{min}	Min output of thermal generator type g
S	Set of dispatch scenarios	p_g^{max}	Max output of thermal generator type g
K	Set of lines	p_r^{max}	Max output of renewable generator type r
$G_{[q]}$	Set of thermal generators at bus q	$\gamma_{r,t}^s$	Percentage of renewable capacity available for renewable generator type r at time t in scenario s
$G_{new[q]}$	Set of new thermal generators at bus q	R_g^{up}	Max ramp up for thermal generator type g
$R_{[q]}$	Set of renewable generators at bus q	R_g^{down}	Max ramp down for thermal generator type g
$R_{new[q]}$	Set of new renewable generators at bus q	$d_{q,t}^s$	Load q at time t in scenario s
n_g	Expansion decision for the number of thermal generators of type g	R^{cap}	Capacity reserve factor
n_r	Expansion decision for the number of renewable generators of type r	R_g^{min}	Minimum reserve amount that must be provided for thermal generator type g if providing reserves
$p_{g,t}^s$	Dispatch decision for thermal generator type g at time t under scenario s	R_g^{sys}	Minimum operating reserve amount for the system
$p_{r,t}^s$	Dispatch decision for renewable generator type r at time t under scenario s	I_g	Binary indicator parameter which determines if a generator type g provides operating reserves

$v_{q,t}^s$	Voltage value for bus q at time t under scenario s	I_g^{dis}	Parameter which determines the percent of maxgen a generator type g provides to the system capacity
$N_{g,t}^s$	Number of expanded generators being utilized of type g at time t generating in scenario s	A	Incidence matrix for the network (Buses \times Lines)
$SU_{g,t}^s$	Number of generators started up of type g at time t generating in scenario s	B	Susceptance matrix
$SD_{g,t}^s$	Number of generators shut down of type g at time t generating in scenario s	$v_{q,t}$	voltage at bus q
$y_{g,t}^s$	Amount of reserves generator type g at time t providing in scenario s	F.S. vars	n_g for $g \in G \cap G_{new}$ and n_r for $r \in R \cap R_{new}$
$L_{q,t}^s$	Amount of load lost at bus q at time t in scenario s	S.S. vars	$p_{g,t}^s, N_{g,t}^s, y_{g,t}^s, SD_{g,t}^s$ for $g \in G \cap G_{new}$ $p_{r,t}^s$ for $r \in R \cap R_{new}$ $L_{q,t}^s, OL_{q,t}^s$ for $q \in D$
$OL_{q,t}^s$	Amount of excess load at bus q at time t in scenario s	λ_s	Weight of scenario s
c^{loss}	Cost of loss of load	c^{exce}	Cost of excess load
c_g	Cost of expansion for thermal generator type g	c_r	Cost of expansion for renewable generator type r
c_g^{dis}	Cost of dispatch for thermal generator type g	c_r^{dis}	Cost of dispatch for renewable generator type r
c_g^{su}	Cost of start up for thermal generator type g	c_g^{sd}	Cost of shut down for thermal generator type g
n_g^{old}	Existing number of thermal generators of type g	n_r^{old}	Existing number of renewable generators of type r

The stochastic optimization problem is defined as follows.

$$\begin{aligned}
& \text{minimize } \sum_{g \in G_{new}} c_g p_g^{max} n_g + \sum_{r \in R_{new}} c_r p_r^{max} n_r + \dots \\
& \sum_{s \in S} \lambda_s \sum_{t \in T} \sum_{g \in G \cup G_{new}} \left(c_g^{dis} p_{g,t}^s + c_g^{su} SU_{g,t}^s + c_g^{sd} SD_{g,t}^s \right) + \dots \\
& \sum_{s \in S} \lambda_s \sum_{t \in T} \sum_{r \in R \cup R_{new}} c_r^{dis} p_{r,t}^s + \dots \\
& \sum_{s \in S} \lambda_s \sum_{t \in T} \left(c^{exce} OL_t^s + c^{loss} L_t^s \right)
\end{aligned}$$

(continued)

$$\begin{aligned}
\text{subject to } (1+R^{cap})E^{cap} &\leq \sum_{g \in G} p_g^{max} n_g^{old} I_g^{dis} + \sum_{g \in G_{new}} p_g^{max} n_g I_g^{dis} + \sum_{r \in R} p_r^{max} a_x n_r^{old} I_r^{dis} + \sum_{r \in R_{new}} p_r^{max} n_r I_r^{dis} \\
n_g^{therm,min} &\leq n_g \leq n_g^{therm,max} && \forall g \in G_{new} \\
n_r^{renew,min} &\leq n_r \leq n_r^{renew,max} && \forall r \in R_{new} \\
0 &\leq N_{g,t}^s \leq n_g^{old} && \forall s,t,g \in G_{new} \setminus G \\
0 &\leq N_{g,t}^s \leq n_g && \forall s,t,g \in G \setminus G_{new} \\
0 &\leq N_{g,t}^s \leq n_g + n_g^{old} && \forall s,t,g \in G \cap G_{new} \\
p_g^{min} N_{g,t}^s &\leq p_{g,t}^s && \forall s,t,g \\
p_{g,t}^s + y_{g,t}^s &\leq p_g^{max} N_{g,t}^s && \forall s,t,g \\
R_g^{min} N_{g,t}^s (I_g) &\leq y_{g,t}^s \leq (p_g^{max} - p_g^{min}) N_{g,t}^s (I_g) && \forall s,t,g \\
N_{g,t-1}^s - N_{g,t}^s + SU_{g,t}^s - SD_{g,t}^s &= 0 \text{ given } N_{g,0}^s && \forall s,t,g \\
R^{sys} &\leq \sum_{g \in G \cup G_{new}} y_{g,t}^s && \forall s,t \\
0 &\leq p_{r,t}^s \leq \gamma_{r,t}^s p_r^{max} n_r^{old} && \forall s,t,r \in R \setminus R_{new} \setminus G \\
0 &\leq p_{r,t}^s \leq \gamma_{r,t}^s p_r^{max} n_r && \forall s,t,r \in R_{new} \setminus R \\
0 &\leq p_{r,t}^s \leq \gamma_{r,t}^s p_r^{max} (n_r + n_r^{old}) && \forall s,t,r \in R \cap R_{new} \\
R_g^{down} &\leq p_{g,t}^s - p_{g,t-1}^s \leq R_g^{up} \text{ given } p_{g,0}^s && \forall s,t,g \\
f_l^{max} &\leq f_{l,t}^s \leq f_l^{max} && \forall s,t,l \\
v_{q,t}^{min} &\leq v_{q,t}^s \leq v_{q,t}^{max} && \forall s,t,q \\
f_{l,t} &= B_{q,q'} (v_{q,t}^s - v_{q',t}^s) \quad (l := \text{line from } q \text{ to } q') && \forall s,t,l \\
\sum_{g \in G_{[q]} \cup G_{new[q]}} p_{g,t}^s + \sum_{r \in R_{[q]} \cup R_{new[q]}} p_{r,t}^s + \sum_{l \in K} A_{q,l} f_{l,t}^s &= d_{q,t}^s + OL_{q,t}^s - L_{q,t}^s && \forall s,t,q \\
N_{g,t}^s, SU_{g,t}^s, SD_{g,t}^s &\in \mathbb{N} && \forall s,t,g
\end{aligned}$$

Generally, UC involves solving optimization problems where dispatch decisions represent which generators are switched on to meet the electricity demand. In the SPEED model, UC is determined by the dispatch decision variables $p_{g,t}^s$ and $p_{r,t}^s$ which are typically modeled as integers. However, it is of interest to see how relaxing this integer assumption to a continuous variable would impact the expansion and operating costs as well as the expansion decisions.

The mathematical significance of this relaxation, is that it changes the dispatch problem from solving a mixed-integer linear program (MILP) to a linear program (LP) which can greatly reduce the complexity of the problem and allow other physically meaningful aspects of CEP to be modeled. However, it should be mentioned prior work has demonstrated that relaxing binary UC variables in PCMs alone does not necessarily reduce computational burden, contrary to expectations larger binary decision trees improved UC model resolution, and results are likely solver dependent (Alemany, Kasprzyk, and Magnago 2018). If the MILP can be relaxed to an LP, then it might be possible to include a reactive (or alternating current) optimal power flow model in SPEED, as opposed to the simpler but less accurate direct current optimal power flow. Relaxing the UC problem is possible with the SPEED model. In this work, we study this relaxation by conducting two similar experiments produced both with and without this UC integer relaxation. The goal of this experiment is to see if this model assumption significantly impacts the quantities of interest, and to investigate possible changes in computational burden due to differing model complexities.

1.3 Contributions of this research

The need to understand how uncertainties within CEP models and modeling assumptions affect quantities of interest such as expansion and operating costs, as well as expansion decisions is substantial. This area of research is particularly important for models which seek to capture how large networks will evolve and operate under increased sources of variable generation, i.e., higher penetration of renewable technologies such as solar and wind power. Uncertainty quantification of CEP models which estimate expansion costs and decisions and *production cost models* which estimate operating costs and dispatch decisions is a key focus of research at NREL. ReEDS considers a range of possible grid evolutions in an attempt to identify key drivers, ramifications, and decisions which contribute to better informed investment and policy decisions. However, research to quantify how uncertainties and model assumptions, such as UC, within ReEDS may be affecting its outputs *globally* remains challenging due to size and complexity of the model (Murphy et al. 2019, Appendix A).

This report focuses specifically on uncertainty quantification of the SPEED model. Mathematically, the SPEED model is a generalization of the ReEDS linear optimization problem (1). It allows for higher fidelity time-series data to simulate CEP and production cost modeling compared to ReEDS. This ability to sample at a higher fidelity in time enables us to study the smaller time scale features of the model associated with UC, where ReEDS is limited.

To quantify the uncertainty we build low-dimensional representations of specific quantities of interest which enable global sensitivity analysis. Specifically, we use sparse polynomial expansions to build surrogate models of specific Quantities of Interest (QoIs) using simulations of the SPEED model run on the Eagle HPC, surrogate models are validated to ensure their accuracy and avoid over-fitting potentially noisy or corrupted data. These two approaches are chosen because they may be used to compute *global sensitivity metrics* of the SPEED model.

Our results in Sections 4 show that by relaxing the binary UC dispatch decision variables, CEP models could be significantly underestimating the amount of installed wind capacity required to satisfy design constraints, e.g., transmission guidelines or safety constraints such as reserve capacity margin. While prior work has demonstrated that relaxing binary UC variables in PCMs alone does not necessarily reduce computational burden, contrary to expectations larger binary decision trees improved UC model resolution, and overall results are likely solver dependent (Alemany, Kasprzyk, and Magnago 2018). Our results show significant reduction in the computational burden of the SPEED model with relative increases in uncertainty (variability in QoI estimates) by relaxing binary UC decision variables in QoI estimates. See Section 4.3.

2 The RTS-GMLC Data Set

The SPEED model uses two stage stochastic optimization to make generation capacity expansion decisions for a power grid. For this work, the SPEED model parameters in Table 1 are defined according to a modified version of the Reliability Test System by the Grid Modernization Laboratory Consortium (RTS-GMLC). The RTS-GMLC is based upon the 1979 and 1996 Institute of Electrical and Electronics Engineers (IEEE) Reliability Test Systems (RTS-96) (Grigg et al. 1999). The RTS-GMLC features several key changes from the RTS-96, for more information on this system see <https://github.com/GridMod/RTS-GMLC/>. Overall, RTS-GMLC is a data set used for production cost modeling of a hypothetical electric grid operating over the geographic region overlapping Los Angeles, Las Vegas, and Western Arizona.

For this work the SPEED model was configured to compute an optimal capacity expansion plan for the modified RTS-GMLC 10 years in the future. An annual load growth of 3% a year was assumed and a typical pattern of retirement for generators was used, which was informed from the U.S. Energy Information Administration (EIA) report (Jell and Bowman 2018). Additionally, all hydro-electric and photovoltaic (PV) generation were removed from the system. This decision was made due to the limited scope of this project. The generator expansions were restricted to be typical sized wind power plants, combined cycle (CC) gas generators, and combustion turbine (CT) gas generators at specified buses. In summary the model aimed to build out the optimal combination of wind and gas generation for the projected load on the system in 10 years.

Part of the aim of this research is to identify key input parameters and QoIs (or output) that are relevant for capacity expansion decisions. For the scope of this work, we assume the uncertain model inputs include 1) the capacity

Parameter ξ_i	Description	Units	Range
R^{cap}	Capacity Reserve Margin	%	[10, 20]
c^{loss}	Loss of load cost	\$/MWh	[4000, 15000]
c^{exce}	System excess load cost	\$/MWh	[100, 1000]
c^{ng}	Natural gas price	\$/MMBTU	[1, 10]
c^{wind}	Wind Install Cost	%	[70, 105]
t^{cap}	Transmission Capacity	%	[50, 150]

Table 2. SPEED Model Uncertain Input Parameters

QoIs u_j	Description	Units
c^{exp}	System Expansion Cost	\$
c^{oper}	System Operation Cost	\$/year
max_{wind}	Installed Wind Capacity	MW
max_{gas}	Installed Gas Capacity	MW

Table 3. SPEED Model Quantities of Interest

reserve factor R^{cap} , 2) the cost of loss of load c^{loss} , 3) the cost of excess load c^{exce} , 4) the price of natural gas³ c^{ng} , 5) the cost of wind installation c^{wind} , and 6) the transmission capacity t^{cap} . See Table 2. The cost of loss of load c^{loss} represents the dollar amount which is lost due to a generator falling short of its available capacity, a nominal value of this cost is \$10,000/ MWh (Megawatt-hour). The capacity reserve factor R^{cap} is a percentage of the required capacity which is added to the total expansion requirement to prevent a supply shortage, typically this is modeled as 15% of the required expansion capacity following the North American Electric Reliability Corporation (NERC) recommendation (Reimers, Cole, and Frew 2019). The cost of system overload c^{exce} is typically modeled as an order of magnitude smaller than c^{loss} and represents the dollar amount associated with a system overload, there is not a well establish nominal value for this parameter making it an important uncertainty of the model. Such overloads can damage existing infrastructure and might occur when variable generation peaks, for example when both high-wind and solar availability occur at the same time. The natural gas price c^{ng} is informed by the Henry Hub Natural Gas Spot Price history provided by the EIA. The cost of wind installation c^{wind} is informed by the 2018 land-based wind annual technology baseline levelized cost of energy (LCOE) projections (Vimmerstedt et al. 2018), and the 2017 Wind Technologies Market Report (Wiser and Bolinger 2017). The transmission capacity t^{cap} is an artificial parameter, not listed in Table 1, which we introduce for the purpose of uncertainty quantification. Transmission capacity is modeled by considering a percentage t^{cap} of the transmission capacity defined by the RTS-GMLC data set. For example, $t^{cap} = 100\%$ represents the exact total amount of transmission capacity for RTS-GMLC data set, while a value of $t^{cap} = 50\%$ represents a grid which has 50% the transmission capacity between all nodes, compared to RTS-GMLC. Traditionally, the term *transmission capacity* may refer to the amount of power in watts which can be sent over a transmission line within acceptable line loss limits, our definition of t^{cap} is a percentage which scales this value for every line in the RTS-GMLC data set.

The QoIs for this study are aggregate outputs of the SPEED model. We define the QoIs as 1) system expansion cost c^{exp} , 2) system operation cost c^{oper} , 3) installed gas capacity max_{gas} , 4) and installed wind capacity max_{wind} . See Table 3. Mathematically, the expansion and operations costs are associated with the first and second terms of $\tilde{\mathbf{c}}^T \tilde{\mathbf{x}} + \sum_{s \in S} \mathbb{P}[s] (\mathbf{f}_s^T \mathbf{y}_s)$, the objection function of the stochastic optimization problem (2).

³This quantity is modeled through the speed model parameter c_g^{dis} .

3 Methods for Uncertainty Quantification

In this section we briefly summarize two distinct approaches for uncertainty quantification that are used to study the SPEED model. These methods are chosen specifically because they admit the calculation of *global sensitivity metrics*, which explicitly quantify uncertainty in the model's variance. Global sensitivity metrics are often used for sensitivity analysis to quantify each uncertain input parameter's influence or importance on a specified QoI. The term *global* means that the metrics access the importance of each variable over a range of parameters compared to *local* which often refers to measuring the model's response to small perturbations around a nominal parameter value.

3.1 Polynomial Chaos Expansions

One common approach for UQ of problems with random inputs involves expanding a QoI, with finite variance, in a multivariate polynomial basis using so-called *Polynomial Chaos (PC) expansions*. Consider our *uncertain model inputs*

$$\boldsymbol{\xi} = [\xi_1, \dots, \xi_d]^T, \quad (3)$$

and an orthogonal polynomial basis $\{\psi_1, \dots, \psi_P\}$ such that

$$\mathbb{E}[\psi_i(\boldsymbol{\xi})\psi_j(\boldsymbol{\xi})] = \int_{\Omega} \psi_i(\boldsymbol{\xi})\psi_j(\boldsymbol{\xi})f(\boldsymbol{\xi})d\boldsymbol{\xi} = \langle \psi_i(\boldsymbol{\xi}), \psi_j(\boldsymbol{\xi}) \rangle = \delta_{ij}\gamma_i \quad (4)$$

where $\Omega \subseteq \mathbb{R}^d$ is the input parameter space, $f(\boldsymbol{\xi})$ is the product of densities corresponding to each uncertain input, and $\gamma_i = \mathbb{E}[\psi_i^2]$ is a product of univariate normalization constants. For a random process $u(t, x, \boldsymbol{\xi}) : [0, T] \times \mathcal{D} \times \Omega \rightarrow \mathbb{R}$ a PC expansion can be written as

$$u(\boldsymbol{\Xi}) = \sum_{k=0}^{\infty} c_k \psi_k(\boldsymbol{\Xi}). \quad (5)$$

We truncate the expansion in (5) for computation, i.e.,

$$u(\boldsymbol{\Xi}) = \sum_{k=1}^P c_k \psi_k(\boldsymbol{\Xi}) + \varepsilon(\boldsymbol{\Xi}) \approx \sum_{k=1}^P c_k \psi_k(\boldsymbol{\Xi}), \quad (6)$$

where $P = \frac{(p+d)!}{p!d!}$ is the the number of basis functions, p is the highest polynomial order allowed, sometimes referred to as *total order*, d is the number of uncertain inputs, sometimes referred to as the *stochastic dimension*, and $\mathbf{c} = [c_1, \dots, c_P]^T$ is a vector of coefficients. In words, the expansion in (6) is the projection of u onto the space $\mathbb{P}_p(\boldsymbol{\xi})$ which is the space of all polynomials of $\boldsymbol{\xi} \in \mathbb{R}^d$ of up to degree p . Note as we explain in Section 2 the random processes or QoI's of the SPEED model are aggregated and defined over a specified spatial region for a specific point in time, for these reasons we suppress the explicit dependence of u on x and t for the remainder of this work. It has been observed for a variety of physics based science and engineering problems that often, many of the coefficients c_k are negligible and thus $u(\boldsymbol{\Xi})$ admits a sparse representation of the form

$$u(\boldsymbol{\xi}) \approx \sum_{k \in C} c_k \psi_k(\boldsymbol{\xi}), \quad (7)$$

where the index set C has few elements, say $s = |C| \ll P$, and we say that our QoI is *approximately sparse* in the polynomial basis. Sparse PCEs have been used within the context of UQ for a variety of problems (Doostan and Owhadi 2011; Blatman and Sudret 2011; Mathelin and Gallivan 2012; Jones, Parrish, and Doostan 2015; Sargsyan et al. 2014; Yan, Guo, and Xiu 2012; Yang and Karniadakis 2013; Peng, Hampton, and Doostan 2014; Schiavazzi, Doostan, and Iaccarino 2014; West IV and Hosder 2014; Diaz, Doostan, and Hampton 2018; Jakeman, Eldred, and Sargsyan 2015; Hampton and Doostan 2015; Bouchot et al. 2015; Peng, Hampton, and Doostan 2016; Chkifa et al. 2016; Winokur et al. 2016; Yang et al. 2016; Adcock 2017; Jakeman, Narayan, and Zhou 2017).

The multivariate polynomial basis functions $\psi_k(\boldsymbol{\xi})$ are the result of a tensor product where each basis function can be expressed as a product of 1-dimensional polynomials of order n , $\varphi_n(\xi_i)$ for $i = 1, \dots, d$. To better illustrate this we highlight that (6) can be written using multi-indices, which are tuples of integers $\mathbf{k}' \in \mathbb{N}^d$, where the components of

\mathbf{k}' indicate the order of the 1-dimensional polynomials which when multiplied together make up a single multivariate basis function. We can re-write (6) as

$$u^P(\boldsymbol{\xi}) = \sum_{|\mathbf{k}'|=0}^P c_{\mathbf{k}'} \psi_{\mathbf{k}'}(\boldsymbol{\xi}) = \sum_{k=1}^P c_k \psi_k(\boldsymbol{\xi}), \quad (8)$$

which is a useful representation to describe the global sensitivity metrics PC expansions provide. Table 4 shows the single index, multi-index, and tensored polynomials for a PC expansion with $(d, p) = (3, 2)$, we encourage those who are interested to read Smith 2013, Section 10.1 for more foundational information on PCEs. There are many types

k	order $ \mathbf{k}' $	Multi-index	Polynomial $\psi_k(\boldsymbol{\xi})$
1	0	(0, 0, 0)	$\varphi_0 = 1$
2	1	(1, 0, 0)	$\varphi_1(\xi_1) \varphi_0(\xi_2) \varphi_0(\xi_3)$
3		(0, 1, 0)	$\varphi_0(\xi_1) \varphi_1(\xi_2) \varphi_0(\xi_3)$
4		(0, 0, 1)	$\varphi_0(\xi_1) \varphi_0(\xi_2) \varphi_1(\xi_3)$
5	$p = 2$	(2, 0, 0)	$\varphi_2(\xi_1) \varphi_0(\xi_2) \varphi_0(\xi_3)$
6		(1, 1, 0)	$\varphi_1(\xi_1) \varphi_1(\xi_2) \varphi_0(\xi_3)$
7		(1, 0, 1)	$\varphi_1(\xi_1) \varphi_0(\xi_2) \varphi_1(\xi_3)$
8		(0, 2, 0)	$\varphi_0(\xi_1) \varphi_2(\xi_2) \varphi_0(\xi_3)$
9		(0, 1, 1)	$\varphi_0(\xi_1) \varphi_1(\xi_2) \varphi_1(\xi_3)$
$P = 10$		(0, 0, 2)	$\varphi_0(\xi_1) \varphi_0(\xi_2) \varphi_2(\xi_3)$

Table 4. Single index, multi-index, and tensored polynomials for $(d, p) = (3, 2)$.

of orthogonal polynomials, some standard distributions and their corresponding orthogonal polynomials referred to as the Wiener-Askey Polynomial Scheme are shown in Table 5, in this work we use Legendre polynomials (Xiu and Karniadakis 2002). As a practical note, if a scientist or engineer is working with data $\boldsymbol{\xi}$ which follows an unknown distribution $f(\boldsymbol{\xi})$, numerical methods exist, typically based on Gram-Schmidt Orthonormalization, to construct an appropriate basis.

$f(\xi_i)$	Polynomial Type	Support
Gaussian	Hermite	$(-\infty, \infty)$
Uniform	Legendre	$[a, b]$
Beta	Jacobi	$[a, b]$
Gamma	Laguerre	$(0, \infty)$

Table 5. Four standard types of orthogonal polynomials according to the Wiener-Askey Polynomial Scheme.

To construct a polynomial expansion for a given QoI we perform two main steps. For the first main step, we draw $i = 1, \dots, N$ independent samples from $f(\boldsymbol{\xi})$, which we denote $\boldsymbol{\xi}_i$, then evaluate the computational model which yields corresponding values of the QoI $u(\boldsymbol{\xi}_i)$. The polynomial coefficients \mathbf{c} are approximated using the experimental design consisting of samples $\{\boldsymbol{\xi}_i\}_{i=1}^N$ and their corresponding QoIs $\{u(\boldsymbol{\xi}_i)\}_{i=1}^N$, which are related by the linear system $\mathbf{u} \approx \boldsymbol{\Psi}\mathbf{c}$, where

$$\boldsymbol{\Psi}(i, j) := \psi_j(\boldsymbol{\xi}_i) \text{ and } \mathbf{u} := [u(\boldsymbol{\xi}_1), \dots, u(\boldsymbol{\xi}_N)]^T, \quad (9)$$

so that our QoI is expanded into a polynomial basis $\boldsymbol{\Psi}$.

In some studies, it may be of interest to investigate how best to construct the experimental design $\{\boldsymbol{\xi}_i\}_{i=1}^N$ with the intent of building a surrogate model prior to computing or sampling an expensive QoI, this type of research is commonly referred to as Optimal Design of Experiments (Pukelsheim 2006). Generally speaking, in the context of PCE-based surrogate modeling this type deterministic sampling, and even adaptive sampling schemes, have

been shown to greatly benefit problems of high polynomial order p and relatively low stochastic dimension d , i.e., problems such that $p > d$ (Fajraoui, Marelli, and Sudret 2017; Diaz, Doostan, and Hampton 2018). In this work, $(d, p) = (6, 6)$ and constructing our sparse PCE surrogates can be considered a relatively high-dimensional problem, therefore, as the prior research suggests, deviating from traditional Monte Carlo sampling would be expected to have limit benefits and was deemed out of the scope of this particular study.

Our motivation for choosing a Legendre polynomial dictionary is rooted in the fact that the Legendre polynomial basis is orthogonal with respect to the uniform probability density function over $[-1, 1]^d$. The orthogonality of a given basis is intimately tied the *mutual coherence* of Ψ , given by

$$\mu(\Psi) = \max_{k < l} \frac{|\langle \Psi(:, k), \Psi(:, l) \rangle|}{\|\Psi(:, k)\|_2 \|\Psi(:, l)\|_2} = \max_{i \neq j} |\tilde{\mathbf{G}}(i, j)| \in [\underline{\mu}(\Psi), 1] \quad (10)$$

where $\tilde{\mathbf{G}} = \tilde{\Psi}^T \tilde{\Psi}$ is called the Gram matrix of $\tilde{\Psi} = \Psi \mathbf{S}$ and \mathbf{S} is a diagonal scaling matrix ensuring each column of $\tilde{\Psi}$ has unit length, and the value $\underline{\mu}(\Psi) = \sqrt{\frac{P-N}{N(P-1)}}$ defines the Welch bound (Hong and Zhu 2018).

The mutual coherence is closely related to the maximum pairwise correlation between any two basis functions of a given basis, i.e., the worst-case coherence between any two columns. In terms of compressed sensing, it is generally advantageous to have an *incoherent* basis which is a basis with small mutual coherence. If a solution of interest has an exact representation in an appropriate polynomial basis and the QoI values are exact, then any K -sparse solution can be exactly recovered by compressed sensing algorithms provided that

$$K < \frac{1}{2} \left[1 + \frac{1}{\mu(\Psi)} \right], \quad (11)$$

where K is the number of non-zero entries in the solution vector \mathbf{c} (Hong and Zhu 2018; Elad 2010). Moreover, if a vector $\mathbf{c} \in \mathbb{R}^P$ has K non-zero entries then we define the ℓ_0 pseudo-norm as $\|\xi\|_0 = K$. Intuitively, if two columns of Ψ are closely correlated it is difficult to distinguish whether the nonlinear model features expressed by the data may be attributed to one basis function or the other (Candes et al. 2011). If the data ξ_i are uniformly distributed over $[-1, 1]$ then the Legendre polynomial dictionary has precisely the incoherence properties which provide optimal compressed sensing results, and certain reconstruction guarantees exist (Schaeffer, Tran, and Ward 2018). For analysis, it is often useful to deal with Legendre polynomials because they are bounded over $[-1, 1]$. From (11), we highlight that the convergence of our PCE surrogates depends explicitly on K and $\mu(\Psi)$, and implicitly on N and P . Employing compressed sensing can allow for stable and convergent solutions when the number of samples in the experimental design is limited, i.e., $N < P$, due to the high computational cost of simulating or collecting values of the QoI. However, because the exact coefficients \mathbf{c} and therefore K is not known *a priori*, it is not possible to know precisely the number of samples N required for a specific QoI surrogate model, which highlights the importance of validation. This validation process is part of the second step for constructing a PCE surrogate model.

The second main step involves evaluating the polynomial basis to build the matrix Ψ , then solving a sparse regression problem. Two such problems are detailed in this section.

In the study of compressed sensing, a sparse approximation of \mathbf{c} can be obtained by solving the optimization problem

$$\hat{\mathbf{c}} = \underset{\mathbf{c}}{\operatorname{argmin}} \|\mathbf{c}\|_0 \quad \text{subject to } \|\mathbf{u} - \Psi \mathbf{c}\|_2 \leq \delta, \quad (12)$$

where $\|\mathbf{c}\|_0 = \{\text{the number of indices } k \text{ such that } c_k \neq 0\}$ indicates the sparsity \mathbf{c} . In (12), δ is a tolerance of solution inaccuracy due to the truncation of the expansion. While, the problem (12) is NP-hard to solve, approximate solutions may be obtained in polynomial time using a variety of greedy algorithms including *orthogonal matching pursuit (OMP)* (Tropp and Gilbert 2005; Tropp and Gilbert 2007; Needell and Vershynin 2010; Davenport and Wakin 2010), *compressive sampling matching pursuit (CoSaMP)* (Needell and Tropp 2009; Pal and Mengshoel 2016), and *subspace pursuit (SP)* (Dai and Milenkovic 2009). A convex relaxation of (12) can also be solved via ℓ_1 -minimization (Candès and Wakin 2008; Donoho 2006). Specifically,

$$\hat{\mathbf{c}} = \underset{\mathbf{c}}{\operatorname{argmin}} \|\mathbf{c}\|_1 \quad \text{subject to } \|\mathbf{u} - \Psi \mathbf{c}\|_2 \leq \delta, \quad (13)$$

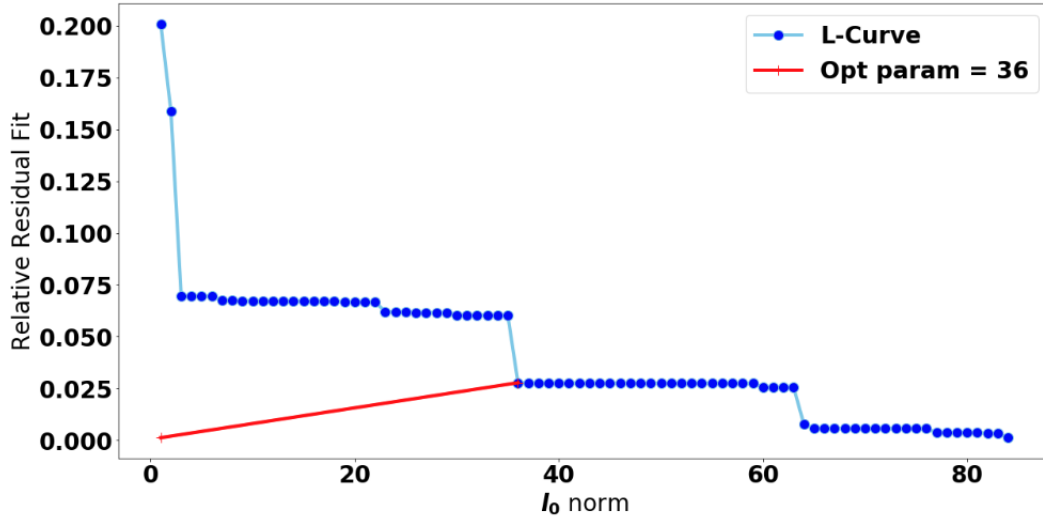


Figure 3. A stereotypical L-curve for the ℓ_0 optimization problem (12).

is referred to as Basis Pursuit Denoising (BPDN). The key advantage of an approximation via compressed sensing is that, if the QoI is approximately sparse, stable and convergent approximations of \mathbf{c} can be obtained using $N < P$ random samples of $u(\mathbf{E})$, as long as Ψ satisfies certain conditions (Candès and Wakin 2008; Donoho 2006; Doostan and Owhadi 2011; Rauhut and Ward 2012; Hampton and Doostan 2015; Adcock 2017). To solve the BPDN problem we use the `SPGL1` Matlab package (Berg and Friedlander 2007).

We solve the basis pursuit denoising problem (13) via the `SPGL1` Matlab package (Berg and Friedlander 2007) with Robust Projection Matrix Optimization to improve the accuracy following (Hong and Zhu 2018). L-curve optimization is used to validate an optimal value of δ which balances the residual fit $\|\mathbf{u} - \Psi\mathbf{c}\|_2$ with the sparsity (or model complexity) $\|\mathbf{c}\|_1$ so that the PC model is not overfit.

L-curve optimization generally speaking involves solving (13) for a variety of tolerance parameters $\{\delta_k\}_{k=1}^{N_\delta}$ and approximating the optimal value according to the *L-curve criterion*. Because $\|\mathbf{c}\|_1$ and $\|\mathbf{c}\|_0$ are typically decreasing functions of δ and $\|\mathbf{u} - \Psi\hat{\mathbf{c}}\|_2^2$ is a strictly increasing function of δ , when plotted on a log-log scale the curve of optimal values of $\|\mathbf{u} - \Psi\hat{\mathbf{c}}\|_2^2$ vs. $\|\mathbf{c}\|_1$ or $\|\mathbf{c}\|_0$ often take on a stereotypical “L” shape in linear problems called an *L-curve* (Aster, Borchers, and Thurber 2018; Hansen 1992). The L-curve criterion is a way to pick δ according to the solution closest to the corner of the L-curve which balances the trade off between the residual fit and the sparsity of the PC model which is a measure of the surrogate complexity. Another common approach for selecting a tolerance parameter δ is cross-validation (Diaz, Doostan, and Hampton 2018; Doostan and Owhadi 2011). The L-curve method is preferable to other approaches because it does not involve further partitioning of small data sets into training and test data, like for example, cross-validation. Figure 3 depicts a stereotypical L-curve and was generated using the `python PCE` tutorial available at https://github.nrel.gov/aces/CEP_UQ. Each point on the L-curve corresponds to a different value of δ and the optimal solution depicted in Figure 3 was such that $\|\hat{\mathbf{c}}\|_0 = 36$.

PC expansions offer explicit formulas for valuable statistical information. The mean and variance of the QoI are given by

$$\mu_u = c_1, \quad \text{and} \quad \text{Var}[u] = \sigma_u^2 = \sum_{k=2}^P c_k^2. \quad (14)$$

Recall that estimating the mean and variance via (14) exploits compressed sensing, because we are solving the BPDN optimization problem defined by (13). PCEs via compressed sensing have been shown to produce more accurate estimates of the true QoI mean and variance than traditional Monte Carlo estimators, e.g., sample mean and standard deviations of the data, for high-dimensional parametric operator equations (Rauhut and Schwab 2017).

More generally stochastic spectral methods which PCEs belong to, are able to exploit smoothness and converge at rates significantly faster than the well known $1/\sqrt{N}$ convergence rate associated with Monte Carlo methods, for moderate parameter dimensions (Smith 2013, Section 10).

Global sensitivity metrics, which explicitly quantify uncertainty, are often computed from *total Sobol' indices* defined by

$$\tau_i = \frac{1}{\text{Var}[u]} \sum_{\mathbf{k}' \in \mathcal{S}_i^+} c_{\mathbf{k}'}^2 = \frac{\mathbb{E}[\text{Var}[u|\boldsymbol{\xi}_{\sim i}]]}{\text{Var}[u]}, \quad (15)$$

where $\mathcal{S}_i^+ := \{\mathbf{k}' \in \mathbb{N}_o^d : 0 \leq |\mathbf{k}'| \leq d, k'_i \neq 0, \text{ and } \boldsymbol{\xi}_{\sim i} = [\xi_1, \dots, \xi_{i-1}, \xi_{i+1}, \dots, \xi_d]\}$. Moreover, we can also compute *first order Sobol' indices* defined by

$$\tau_{1,i} = \frac{1}{\text{Var}[u]} \sum_{\mathbf{k}' \in \mathcal{S}_{1,i}^+} c_{\mathbf{k}'}^2 = \frac{\mathbb{E}[\text{Var}[u|\xi_i]]}{\text{Var}[u]}, \quad (16)$$

where $\mathcal{S}_{1,i}^+ := \{\mathbf{k}' \in \mathbb{N}_o^d : 0 \leq |\mathbf{k}'| \leq d, k'_i \neq 0 \text{ and } k'_j = 0 \text{ for } j \neq i\}$. In words, the total Sobol' index is the expectation of the conditional variance of the QoI given the values of all but the i th input, normalized by the total variance. And the first order Sobol' index is the expectation of the conditional variance of the QoI given the value of the i th input, normalized by the total variance. When $\tau_i = \tau_{1,i}$, this means that globally the i th parameter does not interact with other uncertainty inputs to affect the variance of the QoI. When $\tau_{1,i} = 0$ and $\tau_i \neq 0$, the i th parameter globally affects the variance of the QoI *only* through interactions with other parameters. And when $\tau_i \approx \tau_{1,i} = 0$, the i th input parameter globally doesn't affect the variance either independently or through interactions, and we say the the QoI is *globally insensitive* to that parameter.

3.2 Active Subspaces

Active subspaces describe important directions within a mathematical model's input parameter space/domain. They provide a powerful data-driven approach which may be exploited to a number of ends including uncertainty quantification, data-visualization, optimization, approximation/surrogate modeling, and parameter space dimension reduction. Recall that, $u(t, x, \boldsymbol{\xi}) : [0, T] \times \mathcal{D} \times \Omega \rightarrow \mathbb{R}$. Let's assume that our QoI is for a fixed values of t and x and drop the explicit dependence. Further, we assume that:

- u is differentiable (not necessary in practice);
- $u, \frac{\partial u}{\partial \xi_i} \in L^2(\Omega)$, i.e., the QoI and its partial derivatives are square integrable over the input parameter space Ω ;
- $\boldsymbol{\xi} \in \Omega = [-1, 1]^d$, the input parameters have been shifted and scaled to be within a d -dimensional hypercube;
- and $f(\boldsymbol{\xi}) = 2^{-d}$ for $\boldsymbol{\xi} \in [-1, 1]^d$, and zero elsewhere, i.e., each of our d parameters obey a uniform probability density.

Under these assumptions the statistics, the mathematical expected value of the QoI and its variance are defined by $\mathbb{E}[u] = \int u f d\boldsymbol{\xi}$ and $\text{Var}[u] = \int (u - \mathbb{E}[u])^2 f d\boldsymbol{\xi}$. We may describe a special symmetric positive semidefinite $d \times d$ matrix

$$\mathbf{C} = \int \nabla u \nabla u^T f d\boldsymbol{\xi} = \mathbf{W} \boldsymbol{\Lambda} \mathbf{W}^T, \quad (17)$$

which admits an eigen-decomposition, where $\mathbf{W} = [\mathbf{w}_1, \dots, \mathbf{w}_d]$ is an orthogonal matrix of eigenvectors and $\boldsymbol{\Lambda} = \text{diag}(\lambda_1, \dots, \lambda_d)$ is a diagonal matrix of eigenvalues such that $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_d \geq 0$. One reason the matrix \mathbf{C} is special is that the formula which defines its diagonal components

$$v_i = \int \left(\frac{\partial u}{\partial \xi_i} \right)^2 f d\boldsymbol{\xi} \quad (18)$$

is exactly the *derivative-based global sensitivity metric (DGSM)* originally described by (Kucherenko et al. 2009). Because $f(\boldsymbol{\xi})$ represents a probability density function, $\nabla u(\boldsymbol{\xi})$ is a random vector, and \mathbf{C} represents its uncentered covariance matrix.

In order to explicitly quantify uncertainty via active subspace-based global sensitivity metrics we need to focus on the right hand side of (17). Generally speaking, the eigenvectors of a matrix identify important directions in the column span of that matrix. Moreover, the eigenpairs of the matrix \mathbf{C} satisfy

$$\lambda_i = \int (\nabla u^T \mathbf{w}_i)^2 f d\boldsymbol{\xi}. \quad (19)$$

In words, the i th eigenvalue is the averaged, squared, inner-product of the QoI's gradient and the i th eigenvector. The expression (19) provides insight as to why this method may be used for global sensitivity analysis and parameter space dimension reduction. Consider if $\lambda_i = 0$, then u is constant along the direction \mathbf{w}_i , i.e., then the i th parameter doesn't affect the QoI *globally* which is to say when averaged over the entire input parameter space Ω .

Suppose there is a gap, often an order of magnitude is sufficient, between eigenvalues $\lambda_n > \lambda_{n+1}$ for $n < d$, then we partition the eigenpairs into two sets

$$\mathbf{\Lambda} = \begin{bmatrix} \mathbf{\Lambda}_1 & \\ & \mathbf{\Lambda}_2 \end{bmatrix}, \quad \text{and} \quad \mathbf{W} = [\mathbf{W}_1 \quad \mathbf{W}_2], \quad (20)$$

where $\mathbf{\Lambda}_1 = \mathbf{\Lambda}(:, 1:n) :=$ the first n columns of $\mathbf{\Lambda}$, $\mathbf{\Lambda}_2 = \mathbf{\Lambda}(:, n+1:d)$, $\mathbf{W}_1 = \mathbf{W}(:, 1:n)$, and $\mathbf{W}_2 = \mathbf{W}(:, n+1:d)$. The eigenvectors \mathbf{W} describe a rotation of \mathbb{R}^d and consequently the domain of the QoI $\Omega \subset \mathbb{R}^d$. The subspace $\text{span}(\mathbf{W}_1) \subset \mathbb{R}^d$ defines an n -dimensional *active subspace* of u . On average, perturbing the inputs in the first set of coordinates affects the QoI more than in the second set of coordinates. Because of this relationship we define the following

$$\mathbf{y} = \mathbf{W}_1^T \boldsymbol{\xi} \in \mathbb{R}^n, \quad \text{and} \quad \mathbf{z} = \mathbf{W}_2^T \boldsymbol{\xi} \in \mathbb{R}^{d-n}, \quad (21)$$

where \mathbf{y} are referred to the *active variables* and \mathbf{z} the *inactive variables*. Opposed to the active subspace, the $\text{span}(\mathbf{W}_2)$ is referred to as the *inactive subspace*. Both the active and inactive variables are a linear combinations of the uncertain model inputs, i.e., a weighted sum of the components of $\boldsymbol{\xi}$. For example, $y_1 = \mathbf{w}_1^T \boldsymbol{\xi} = \sum_{i=1}^d w_{1,i} \xi_i$. Moreover because $|\mathbf{y}|, |\mathbf{z}| < d$, the matrices \mathbf{W}_1 and \mathbf{W}_2 define rotations of $\mathbb{R}^d \supset \Omega$ which map to a lower-dimension. Because \mathbf{W}_1 maps the input parameters to a dimension lower than d by rotating \mathbb{R}^d such that the most important features of the model are expressed, active subspaces make it possible to easily visualize the high-dimension model response surface of the QoI as a function of the active variables \mathbf{y} , provided $n \leq 3$.

We can check a model for active subspaces by first approximating the gradients, assuming most models do not provide gradient calculations. In this work, we approximate gradients via local linear gradients as in (Hristache et al. 2001).⁴ The following Monte Carlo method for estimating \mathbf{C} was proposed and analyzed in (Constantine and Gleich 2014). For $i = 1, \dots, N$, draw $\boldsymbol{\xi}_i$ according to $f(\boldsymbol{\xi})$. Compute $\nabla u_i = \nabla u(\boldsymbol{\xi}_i)$. Then estimate

$$\mathbf{C} \approx \hat{\mathbf{C}} = \frac{1}{N} \sum_{i=1}^N \nabla u_i \nabla u_i^T = \hat{\mathbf{W}} \hat{\mathbf{\Lambda}} \hat{\mathbf{W}}^T. \quad (22)$$

The *active subspace error* is defined as

$$e_{as} = \|\mathbf{W}_1 \mathbf{W}_1^T - \hat{\mathbf{W}}_1 \hat{\mathbf{W}}_1^T\| = \|\mathbf{W}_1 \hat{\mathbf{W}}_1^T\|, \quad (23)$$

which measures error between the approximated active subspace and the true active subspace of the QoI. Mathematically, e_{as} has an upper bound which is proportional to the ratio $\lambda_1/(\lambda_n - \lambda_{n+1})$, and there are estimates which inform

⁴python 2.0 codes for active subspace analysis which implement local linear gradient approximation are available at https://github.com/paulcon/active_subspaces (Constantine et al. 2016), python 3.0 codes are available at https://github.com/mathematicalmichael/active_subspaces.

how many samples are required to achieve specific probabilistic guarantees on e_{as} Constantine 2015, Lemma 3.9. Practically speaking, the bound means that if the gap between eigenvalues is large then the active subspace error will be small, provided we have enough samples. In this work we bootstrap to estimate the subspace error by resampling with 10,000 replicates as in Constantine 2015, Algorithm 3.2.

The active subspace provides two different global sensitivity metrics. The first is given by the *first eigenvector* of the matrix (22), specifically

$$\mathbf{w}_1 \tag{24}$$

identifies the most important direction the input parameter space if $\lambda_1 > \lambda_2$. The relative importance of \mathbf{w}_2 is measured by the gap in λ_1 and λ_2 . The components of \mathbf{w}_1 may be used as global sensitivity metrics. The relative importance of each parameter is indicated by the magnitude of the components of \mathbf{w}_1 . Each eigenvector \mathbf{w}_i is unique up to sign, consider that the span defined by \mathbf{w}_i is also defined by $-\mathbf{w}_i$. The sign of components of \mathbf{w}_1 can be identified by physical reasoning about the model or by computing the regression coefficients of a linear model, see Constantine and Diaz 2017, Section 2.3. The sign indicates whether, on average, small perturbations of a parameter positively or negatively affect the QoI. The second metric is the *activity score* which is defined as

$$\alpha_i = \alpha_i(n) = \sum_{j=1}^n \lambda_j w_{i,j}^2. \tag{25}$$

The activity score and the total Sobol' index are related according to the inequality

$$\tau_i \leq \frac{1}{4\pi^2 \text{Var}[u]} (\alpha_i(n) + \lambda_{n+1}), \tag{26}$$

and it has been demonstrated that the global sensitivity metrics τ_i , v_i , $w_{1,i}$, and $\alpha_i(n)$ can provide consistent results for high-dimensional, physics-based models (Constantine and Diaz 2017). Additionally, the activity score and the DGSM are related by

$$\alpha_i(n) \leq v_i, \quad \text{and} \quad \alpha_i(n = d) = v_i, \text{ as in (18)}. \tag{27}$$

In words, the activity score is a truncation of the DGSM. This relationship means that when $\alpha_i(n) \approx v_i$ and $n < d$ so that the model admits an active subspace, the contributions to the QoI's derivatives, associated with the inactive variables, are globally negligible. Mathematically, this scenario occurs when the non-zero entries of $\mathbf{\Lambda}_2$ are sufficiently small relative to $\mathbf{\Lambda}_1$. Generally speaking, these sensitivity metrics are expected to provide similar rankings of importance for model parameters, although it is possible to produce a model where they differ, e.g., a model that does not admit an active subspace.

4 Results and Interpretation

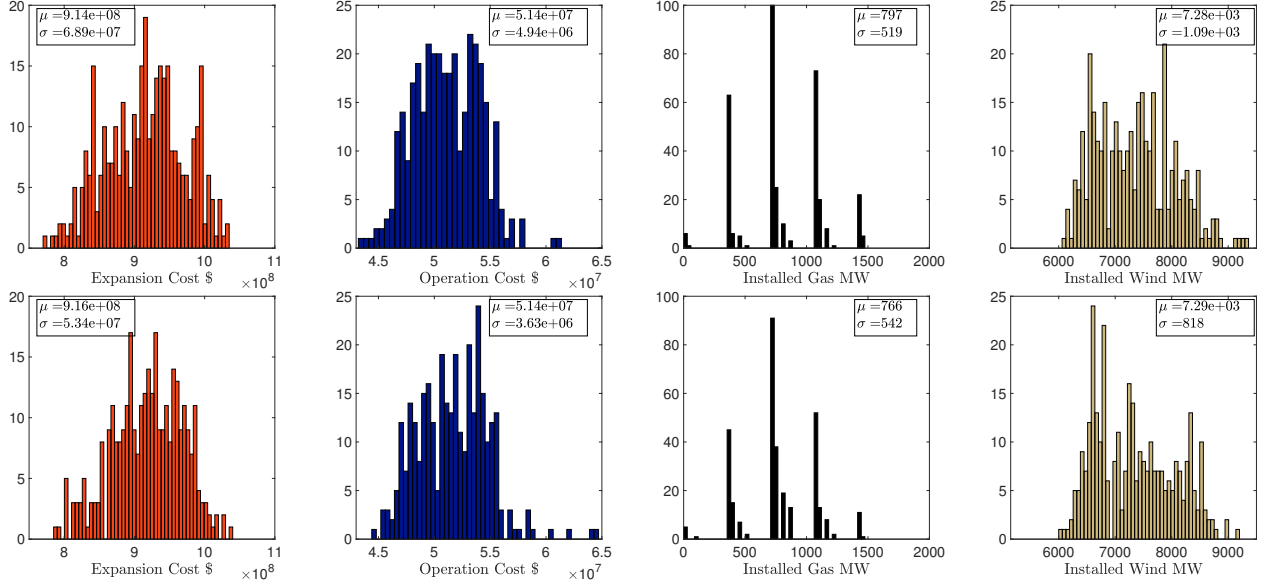


Figure 4. Histograms of the four SPEED model QoIs with *relaxed* (top) and *binary* UC (bottom). The reported values for the mean and standard deviation are estimates of the true posterior distribution of the QoI computed from the PC coefficients according to (14) and are considerably more accurate estimates compared to the sample average and standard deviation.

In this Section, we provide the results obtained from simulating the SPEED model for hundreds of randomly sampled uncertain input parameter values both with and without binary UC. The stochastic optimization problem (2) is solved using the Gurobi optimizer (Optimization 2014) to within a mixed-integer programming gap of 0.5%. The model is built in the scalable Pyomo PySP framework which enables the model to be solved using the progressive hedging algorithm across multiple nodes and cores on an HPC system (Watson, Woodruff, and Hart 2012; Hart et al. 2017).⁵ Figure 4 shows histograms of the four SPEED model QoIs with relaxed and binary UC along with values for the mean and standard deviation that are estimates of the true posterior distribution of the QoI computed from the PC coefficients according to (14). For each QoI we report:

- the five global sensitivity metrics⁶ τ_i , $\tau_{1,i}$, v_i , $\alpha_i(n)$, and $w_{1,i}$;
- eigenvalues of the matrix \mathbf{C} with 95% confidence intervals via bootstrap estimates;
- 1 and 2D *sufficient summary* plots showing the data $\{\xi_i\}_{i=1}^N$ as a function of the active variables \mathbf{y} ;
- the active subspace error e_{as} with 95% confidence intervals via bootstrap estimates;
- and the results of the L-curve optimization, i.e., the two measures of surrogate complexity $\|\hat{\mathbf{c}}\|_0$ and $\|\hat{\mathbf{c}}\|_1$, and the validated point-wise relative residual fit error of the PC surrogate model given by

$$e_{pc} = \frac{\|\mathbf{u} - \Psi\hat{\mathbf{c}}\|_2}{\|\mathbf{u}\|_2}. \quad (28)$$

Samples ξ_i were drawn randomly according to $f(\xi)$ which defines a uniform probability density function over the hyper cube $[-1, 1]^d$, parameters are shifted and scaled to be within the ranges defined in Table 2. Sample sizes for

⁵The SPEED model is available at <https://github.nrel.gov/aces/SPEED> and a python PCE tutorial with codes for post processing the data is available at https://github.nrel.gov/aces/CEP_UQ. The SPEED model input and output data and Matlab code used for the post processing and generation of these results is publicly available at <https://github.com/PaulMDiaz/>.

⁶The reported values of the activity score and DGSM are obtained by normalizing the metrics α and \mathbf{v} by dividing their true value by $\|\alpha\|_2$ and $\|\mathbf{v}\|_2$, respectively. This normalization is not necessary for the remaining metrics which are defined to be in the interval $[0, 1]$.

QoI u	c^{exp}	c^{oper}	max^{gas}	max^{wind}
e_{PC}	0.0202	0.0230	0.0085	0.0273
$\ \hat{\mathbf{c}}\ _0$	116	48	373	124
$\ \hat{\mathbf{c}}\ _1$	1.2051×10^9	6.1586×10^7	9765.8	1.0819×10^4
μ_u	9.14×10^8	5.14×10^7	797	7.28×10^3
σ_u	6.89×10^7	4.94×10^6	519	1.09×10^3

Table 6. The validated point-wise relative residual fit, measures of complexity, and QoI statistics of the PCE surrogate models for relaxed UC.

QoI u	c^{exp}	c^{oper}	max^{gas}	max^{wind}
e_{PC}	0.0299	0.0242	0.0509	0.0420
$\ \hat{\mathbf{c}}\ _0$	17	29	271	21
$\ \hat{\mathbf{c}}\ _1$	1.0287×10^9	6.0367×10^7	5720.4	8.5495×10^3
μ_u	9.16×10^8	5.14×10^7	766	7.28×10^3
σ_u	5.34×10^7	3.63×10^6	542	818

Table 7. The validated point-wise relative residual fit, measures of complexity, and QoI statistics of the PCE surrogate models for binary UC.

each model assumption are $N = 349$ for relaxed UC and $N = 329$ for binary UC. PC surrogates are constructed using Legendre polynomials up to 6th order where $(d, p) = (6, 6)$ for a total of $P = 924$ total basis functions. This experimental configuration results in matrices Ψ which are wide and correspond to under-determined linear systems with noise⁷, which occurs because we have limited samples of the SPEED model, we are considering 6 uncertain input parameters, and we are using polynomials of up to order 6; recall that $P = \binom{p+d}{p}$. Without reducing the total order p of the PCE surrogates, which would reduce the overall accuracy, or by generating more QoI samples by running SPEED simulations, we are forced to consider an under-determined system of the form $\mathbf{u} \approx \Psi \mathbf{c}$. Therefore, to construct PCE surrogates we exploit compressed sensing to solve (12) or (13). Tables 6 and 7 show the validated point-wise relative residual fit, measures of complexity, and QoI statistics for the PCE surrogate models for binary and relaxed UC.

4.1 Relaxed UC Global Sensitivity Analysis

Figure 5 shows the global sensitivity and active subspace analysis for the expansion cost c^{exp} with relaxed UC. The five global sensitivity metrics indicate consistent rankings of global sensitivities. The expansion cost is globally sensitive to reserve capacity margin and the cost of wind. In comparison, the expansion cost is globally insensitive to cost of loss of load, cost of excess load, cost of natural gas, and transmission capacity but may be locally sensitive to these inputs. The Sobol' indices indicate that cost of loss of load, cost of excess load, cost of natural gas, and transmission capacity globally affect the expansion cost *only* through interactions with other inputs, and that reserve capacity margin and cost of wind affect expansion cost through *both* interactions with other inputs and independently. The relative gap between the first ($n = 1$) and second eigenvalue indicates the model admits a 1D active subspace where the first active variable $y_1 = \mathbf{w}_1^T \boldsymbol{\xi} \approx 0.51R^{cap} + 0.86c^{wind}$. The 1D sufficient summary depicts a strong linear correlation between the QoI and y_1 , and is supported by the active subspace error e_{as} which increases as a function of n .

Figure 6 shows the global sensitivity and active subspace analysis for the operation cost c^{oper} with relaxed UC. The five global sensitivity metrics indicate consistent rankings of global sensitivities. The operation cost is globally sensitive to reserve capacity margin, cost of wind, and transmission availability. In comparison, the operation cost is

⁷Although the SPEED model is deterministic in the sense the same inputs always yield the same outputs, the mixed-integer programming gap acts as a tolerance of solution accuracy meaning that the SPEED model solutions differ from the true optimal solution slightly. This affect can be interpreted as noise in the simulated QoI values.

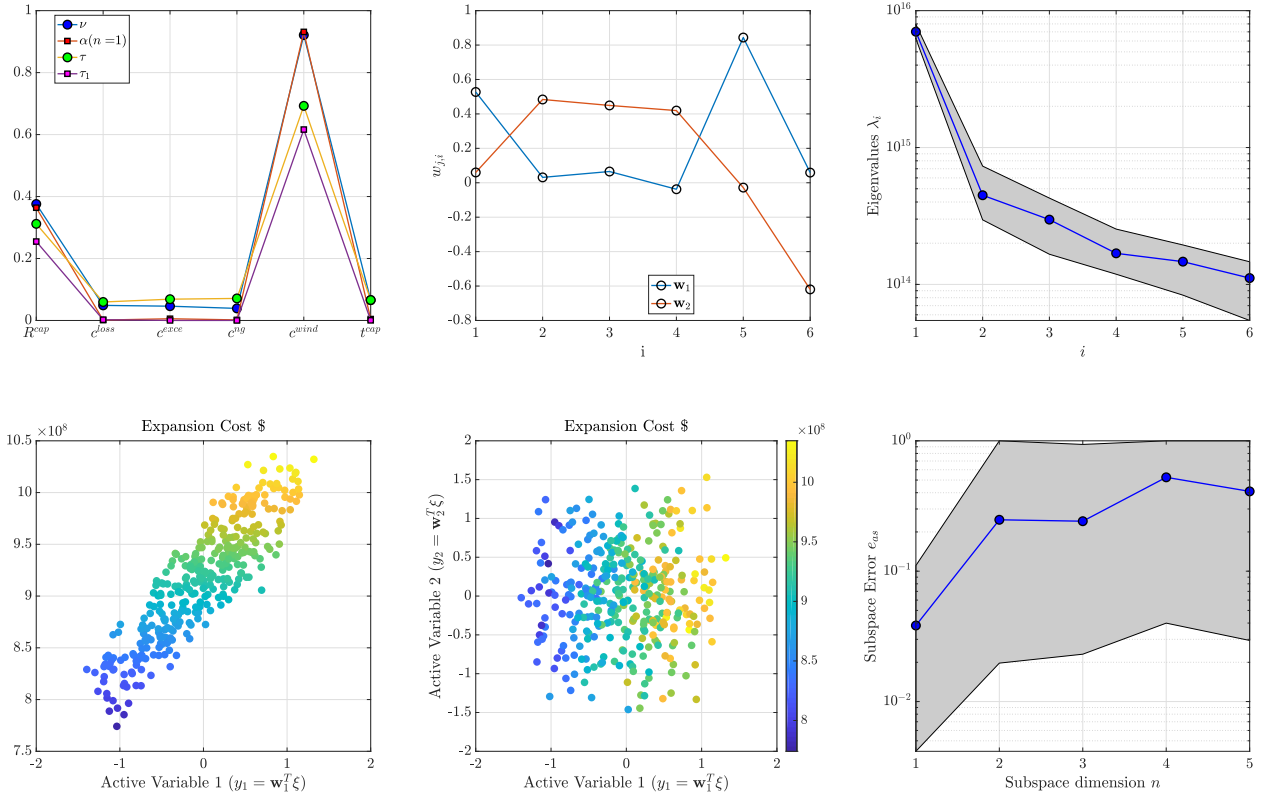


Figure 5. Sensitivity and active subspace analysis for the SPEED model expansion cost c^{exp} with relaxed UC. The top row depicts global sensitivity metrics (left), active subspace eigenvectors (middle), and eigenvalues (right). The bottom row depicts 1D (left) and 2D (middle) sufficient summary plots, and the active subspace error (right).

globally insensitive to cost of loss of load, cost of excess load, and cost of natural gas but may be locally sensitive to these inputs. The Sobol' indices indicate that cost of loss of load, cost of excess load, and cost of natural gas globally affect the operation cost *only* through interactions with other inputs, and that reserve capacity margin, cost of wind, and transmission availability affect operation cost *both* through interactions with other inputs and independently. The relative gap between the first ($n = 1$) and second eigenvalue indicates the model admits a 1D active subspace where the first active variable $y_1 = \mathbf{w}_1^T \boldsymbol{\xi} \approx 0.26R^{cap} - 0.81c^{wind} + 0.53t^{cap}$. The 1D sufficient summary depicts a strong linear correlation between the operation cost and y_1 , and is supported by the active subspace error e_{as} which increases as a function of n , however the confidence intervals suggest the model might admit a 2D subspace.

Figure 7 shows the global sensitivity and active subspace analysis for the maximum installed gas capacity max^{gas} with relaxed UC. Four of the five global sensitivity metrics indicate consistent rankings of global sensitivities with the total Sobol' index showing small differences. Max installed gas is globally sensitive to reserve capacity margin, cost of wind, and transmission capacity. In comparison, the maximum installed gas is globally insensitive to cost of loss of load, cost of excess load, and the cost of natural gas but may be locally sensitive to these inputs. The Sobol' indices indicate that cost of loss of load, cost of excess load, and cost of natural gas globally affect max installed wind *only* through interactions with other inputs, and reserve capacity margin, cost of wind, and transmission capacity affect maximum installed gas through *both* interactions with other inputs and independently. Notice that τ_i is the only metric which ranks cost of loss of load, cost of excess load, and cost of natural gas as globally sensitive, this is a known fault of the total Sobol' indices which may give inaccurate estimates for small sensitivities particularly if the QoI is non-smooth or discontinuous (Myshetskaya et al. 2008; Constantine and Diaz 2017; Diaz 2016). Approx-

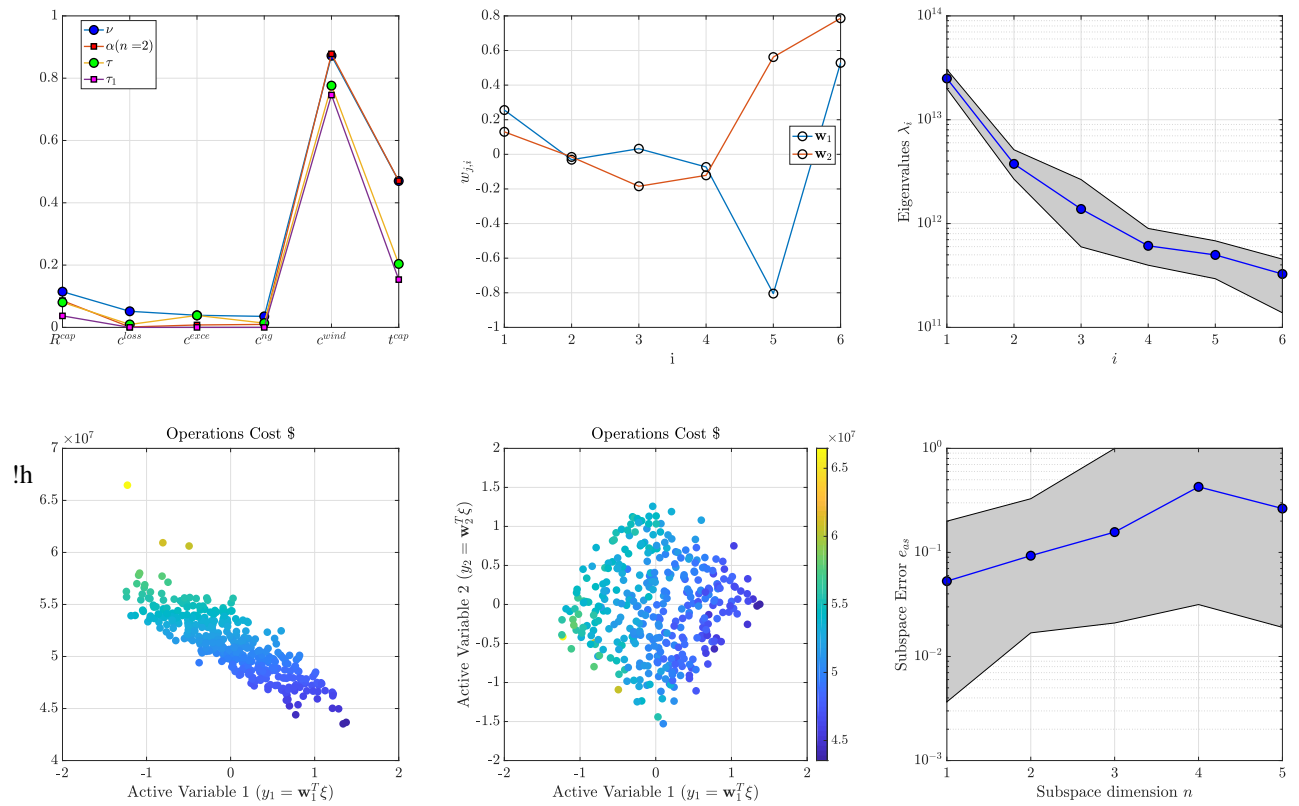


Figure 6. Sensitivity and active subspace analysis for the SPEED model operations cost c^{oper} with relaxed UC. The top row depicts global sensitivity metrics (left), active subspace eigenvectors (middle), and eigenvalues (right). The bottom row depicts 1D (left) and 2D (middle) sufficient summary plots, and the active subspace error (right).

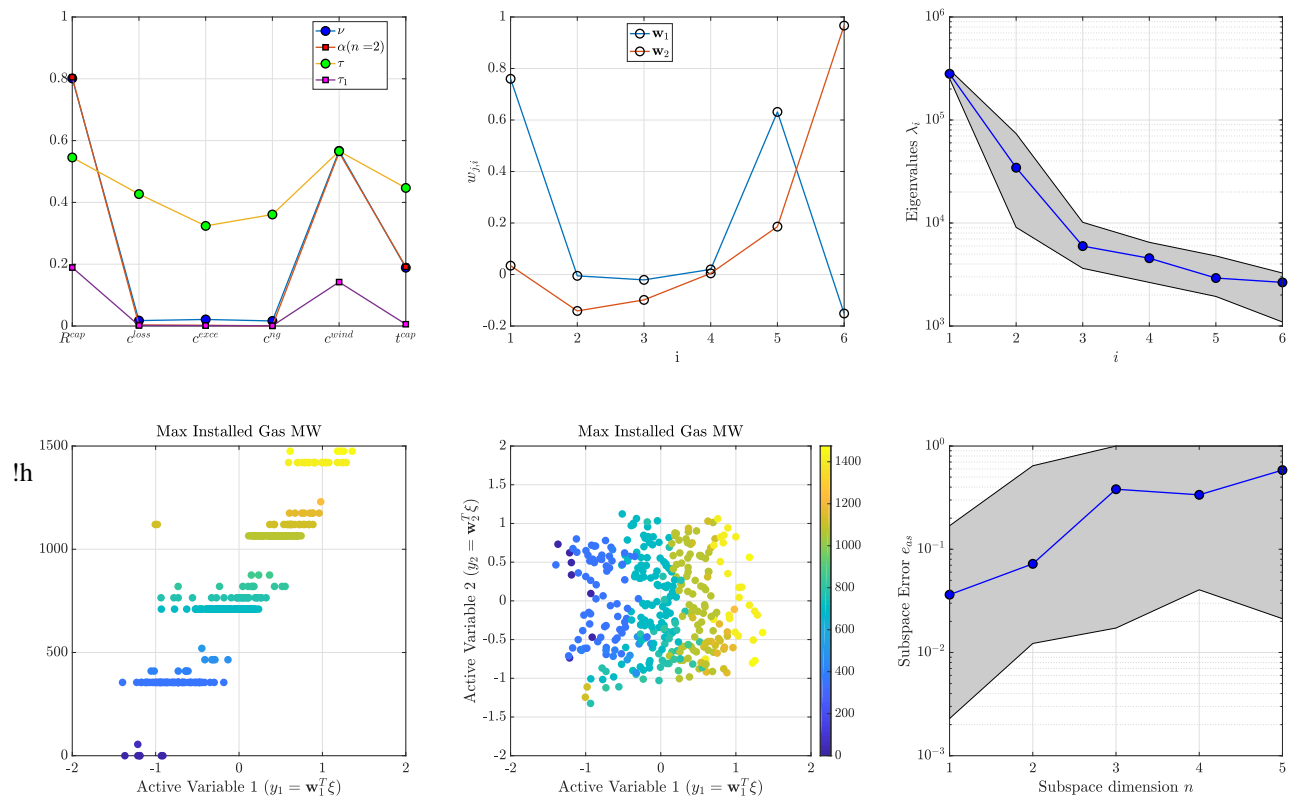


Figure 7. Sensitivity and active subspace analysis for for the SPEED model maximum installed gas capacity max^{gas} with relaxed UC. The top row depicts global sensitivity metrics (left), active subspace eigenvectors (middle), and eigenvalues (right). The bottom row depicts 1D (left) and 2D (middle) sufficient summary plots, and the active subspace error (right).

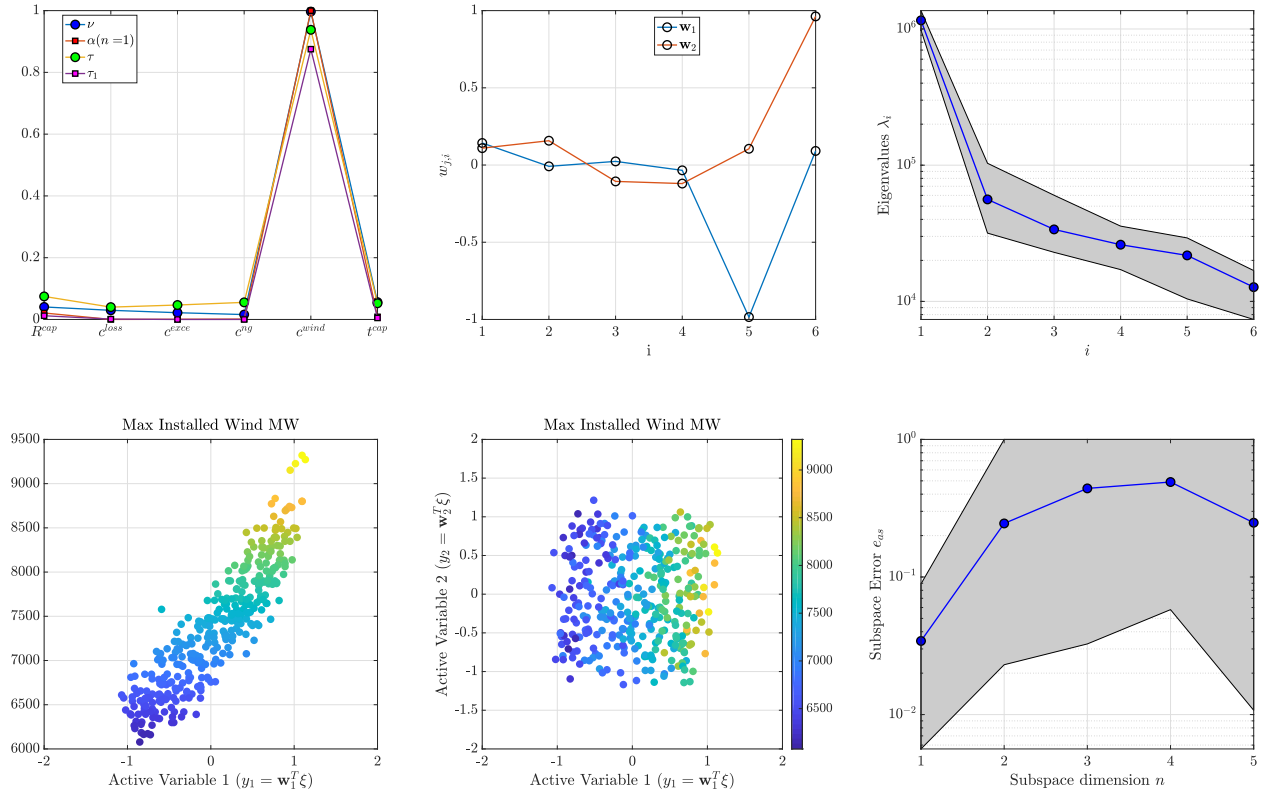


Figure 8. Sensitivity and active subspace analysis for the SPEED model maximum installed wind capacity max^{wind} with relaxed UC. The top row depicts global sensitivity metrics (left), active subspace eigenvectors (middle), and eigenvalues (right). The bottom row depicts 1D (left) and 2D (middle) sufficient summary plots, and the active subspace error (right).

imaging non-smooth or discontinuous functions is a known disadvantage to polynomial approximation (Crestaux, Le Maître, and Martinez 2009, Section 5.3). The relative gap between the first ($n = 1$) and second eigenvalue suggests the model admits a 1D active subspace where the first active variable $y_1 = \mathbf{w}_1^T \xi \approx 0.73R^{cap} + 0.67c^{wind} - 0.13t^{cap}$, however the gap between the second and third eigenvalue is comparable. The 1D sufficient summary depicts a linear correlation between the maximum installed gas capacity and y_1 and is supported by the active subspace error e_{as} which increases as a function of n , however the confidence intervals suggest the model might admit a 2D subspace.

Figure 8 shows the global sensitivity and active subspace analysis for the maximum installed wind capacity max^{wind} with relaxed UC. The five global sensitivity metrics indicate consistent rankings of global sensitivities. Max installed wind is globally sensitive to cost of wind and reserve capacity margin. In comparison, the QoI is globally insensitive to cost of loss of load, cost of excess load, cost of natural gas, and transmission capacity but may be locally sensitive to these inputs. The Sobol' indices indicate that cost of loss of load, cost of excess load, cost of natural gas, and transmission capacity globally affect max installed wind *only* through interactions with other inputs, and cost of wind and reserve capacity margin affect maximum installed wind through *both* interactions with other inputs and independently. The relative gap between the first ($n = 1$) and second eigenvalue indicates the model admits a 1D active subspace where the first active variable $y_1 = \mathbf{w}_1^T \xi \approx 0.13R^{cap} - 0.99c^{wind}$. The 1D sufficient summary depicts a strong linear correlation between the QoI and y_1 . The 1D active subspace is further supported by the active subspace error e_{as} which increases as a function of n .

Overall, the global sensitivity metrics indicate that all four QoIs are relatively insensitive to the input parameters cost of loss of load, cost of excess load, and cost of natural gas, while they are sensitive to reserve capacity margin,

cost of wind, and transmission availability but in different ways. Specifically: expansion cost is driven largely by cost of wind and reserve capacity margin; operation cost is driven by cost of wind, transmission availability, and reserve capacity margin; maximum installed gas capacity is driven by reserve capacity margin, cost of wind, and transmission availability; and maximum installed wind capacity is driven by cost of wind and reserve capacity margin.

4.2 Binary UC Global Sensitivity Analysis

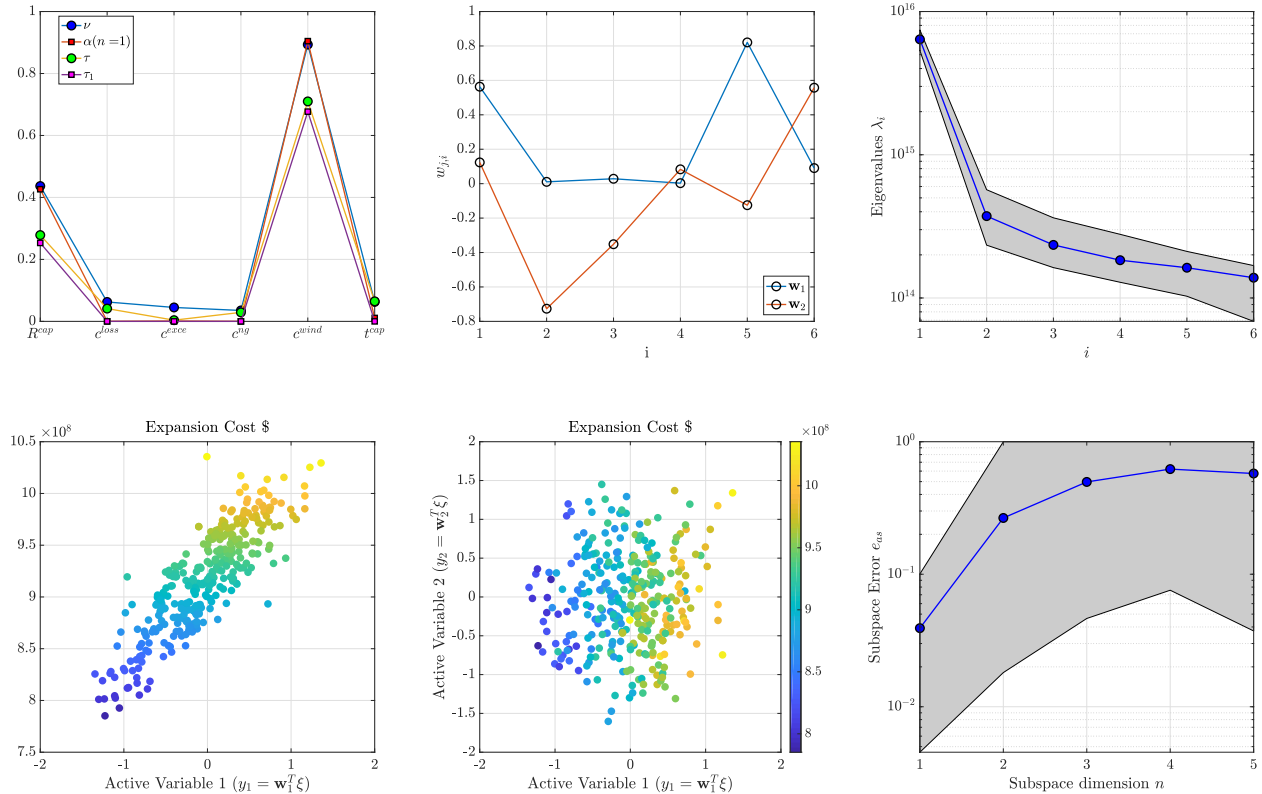


Figure 9. Sensitivity and active subspace analysis for the SPEED model expansion cost c^{exp} with binary UC. The top row depicts global sensitivity metrics (left), active subspace eigenvectors (middle), and eigenvalues (right). The bottom row depicts 1D (left) and 2D (middle) sufficient summary plots, and the active subspace error (right).

Figure 9 shows the global sensitivity and active subspace analysis for the expansion cost c^{exp} with binary UC. The five global sensitivity metrics indicate consistent rankings of global sensitivities, and the relative ranking among sensitivities is similar to the relaxed UC results shown in Figure 5. The expansion cost is globally sensitive to reserve capacity margin and the cost of wind. In comparison, the expansion cost is globally insensitive to cost of loss of load, cost of excess load, cost of natural gas, and transmission capacity but may be locally sensitive to these inputs. The Sobol' indices indicate that cost of loss of load, cost of excess load, cost of natural gas, and transmission capacity globally affect the expansion cost *only* through interactions with other inputs, and that reserve capacity margin and cost of wind affect expansion cost through *both* interactions with other inputs and independently. The relative gap between the first ($n = 1$) and second eigenvalue indicates the model admits a 1D active subspace where the first active variable $y_1 = \mathbf{w}_1^T \xi \approx 0.56R^{cap} + 0.82c^{wind}$. The 1D sufficient summary depicts a strong linear correlation between the QoI and y_1 , and is supported by the active subspace error ϵ_{as} which increases as a function of n .

Figure 10 shows the global sensitivity and active subspace analysis for the operation cost c^{oper} with binary UC. The five global sensitivity metrics indicate consistent rankings of global sensitivities, and the relative ranking among

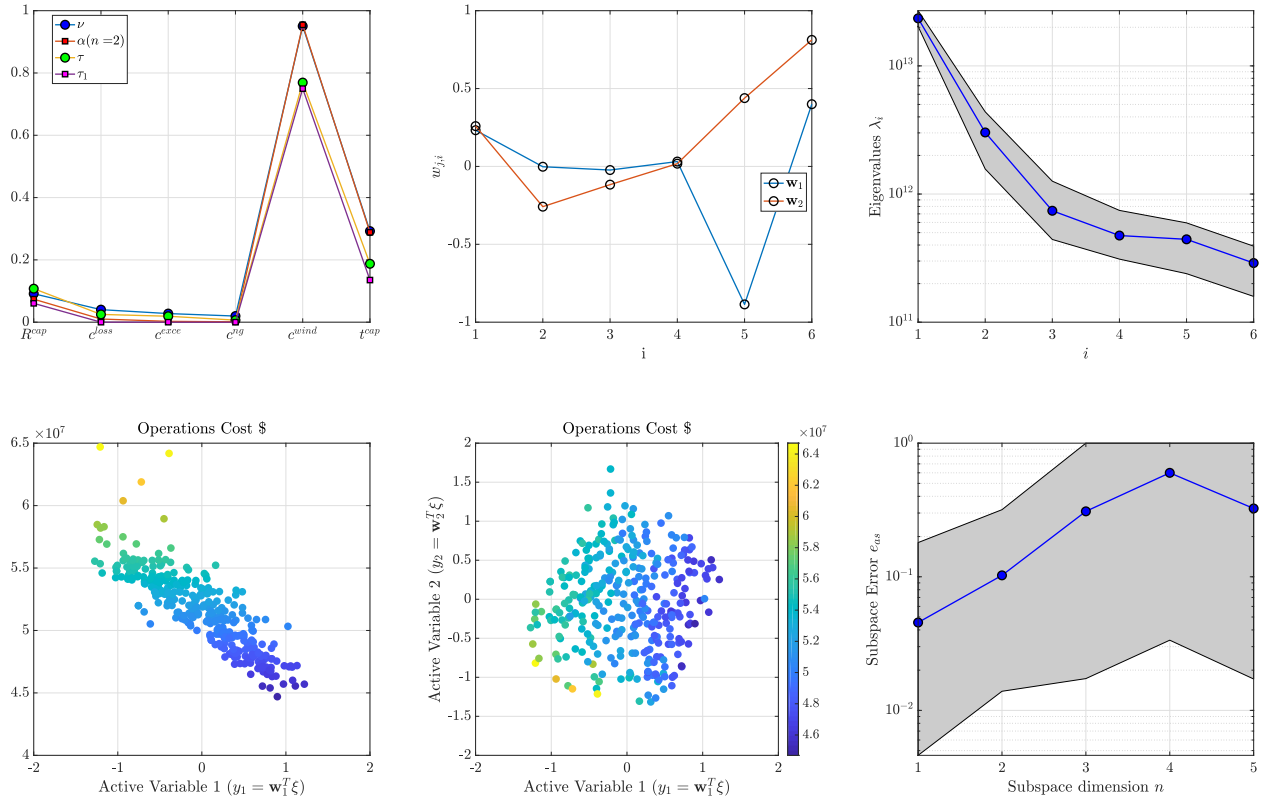


Figure 10. Sensitivity and active subspace analysis for the SPEED model operations cost c^{oper} with binary UC. The top row depicts global sensitivity metrics (left), active subspace eigenvectors (middle), and eigenvalues (right). The bottom row depicts 1D (left) and 2D (middle) sufficient summary plots, and the active subspace error (right).

sensitivities is similar to the relaxed UC results shown in Figure 6. The operation cost is globally sensitive to reserve capacity margin, cost of wind, and transmission availability. In comparison, the operation cost is globally insensitive to cost of loss of load, cost of excess load, and cost of natural gas but may be locally sensitive to these inputs. The Sobol' indices indicate that cost of loss of load, cost of excess load, and cost of natural gas globally affect the operation cost *only* through interactions with other inputs, and that reserve capacity margin, cost of wind, and transmission availability affect operation cost *both* through interactions with other inputs and independently. The relative gap between the first ($n = 1$) and second eigenvalue indicates the model admits a 1D active subspace where the first active variable $y_1 = \mathbf{w}_1^T \xi \approx 0.25R^{cap} - 0.87c^{wind} + 0.43t^{cap}$. The 1D sufficient summary depicts a strong linear correlation between the operation cost and y_1 , and is supported by the active subspace error e_{as} which increases as a function of n , however the confidence intervals suggest the model might admit a 2D subspace.

Figure 11 shows the global sensitivity and active subspace analysis for the maximum installed gas capacity max^{gas} with binary UC. Four of the five global sensitivity metrics indicate consistent rankings of global sensitivities with the total Sobol' index showing small differences. The relative ranking among sensitivities is similar to the relaxed UC results shown in Figure 7. Max installed gas is globally sensitive to reserve capacity margin, cost of wind, and transmission capacity. In comparison, the maximum installed gas is globally insensitive to cost of loss of load, cost of excess load, and the cost of natural gas but may be locally sensitive to these inputs. The Sobol' indices indicate that cost of loss of load, cost of excess load, and cost of natural gas globally affect max installed wind *only* through interactions with other inputs, and reserve capacity margin, cost of wind, and transmission capacity affect maximum installed gas through *both* interactions with other inputs and independently. As in Figure 7, τ_i is the only metric

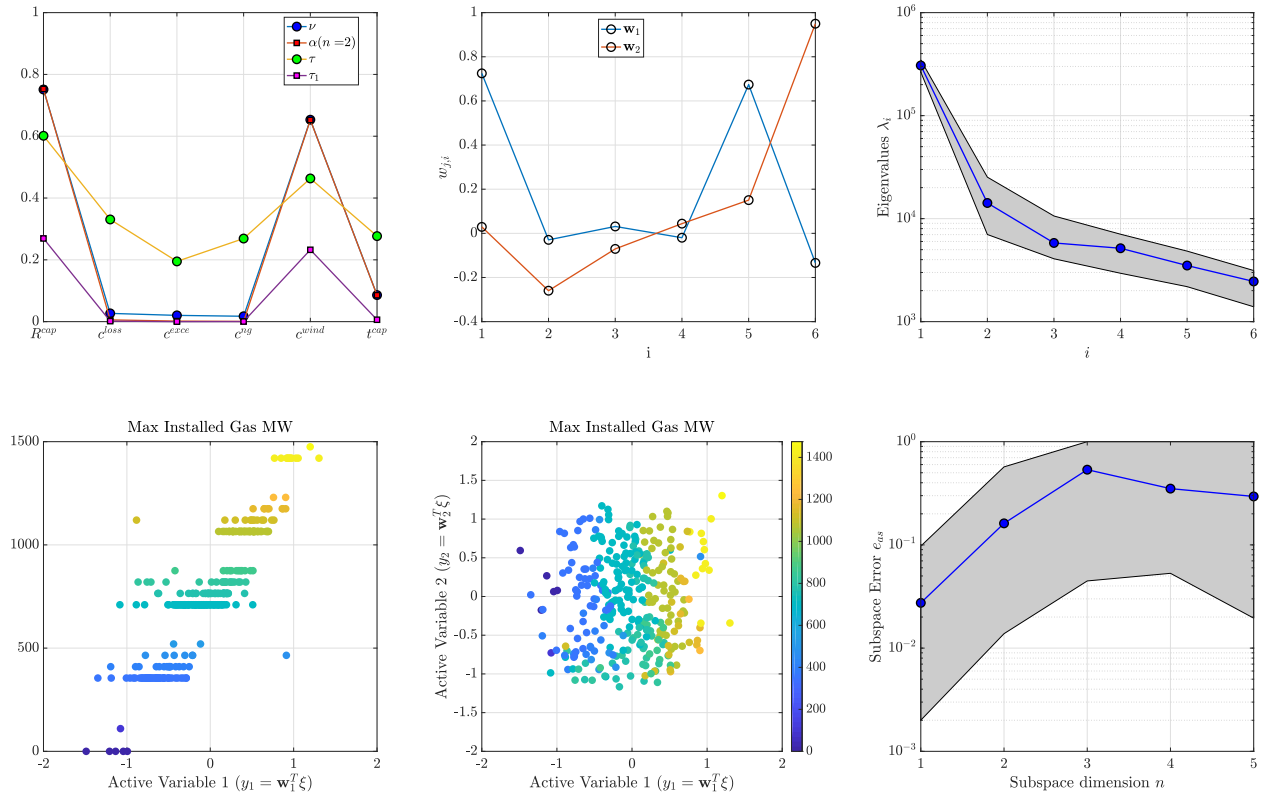


Figure 11. Sensitivity and active subspace analysis for for the SPEED model maximum installed gas capacity max^{gas} with binary UC. The top row depicts global sensitivity metrics (left), active subspace eigenvectors (middle), and eigenvalues (right). The bottom row depicts 1D (left) and 2D (middle) sufficient summary plots, and the active subspace error (right).

which ranks cost of loss of load, cost of excess load, and cost of natural gas as globally sensitive, this is a known fault of the total Sobol' indices which may give inaccurate estimates for small sensitivities particularly if the QoI is non-smooth or discontinuous (Myshetskaya et al. 2008; Constantine and Diaz 2017; Diaz 2016). Similar to the results shown in Section 4.1, the PC error, shown in Table 6, is large for this QoI because the response surface is discontinuous and not smooth. The relative gap between the first ($n = 1$) and second eigenvalue suggests the model admits a 1D active subspace where the first active variable $y_1 = \mathbf{w}_1^T \boldsymbol{\xi} \approx 0.73R^{cap} + 0.66c^{wind} - 0.15t^{cap}$, however the gap between the second and third eigenvalue is comparable. The 1D sufficient summary depicts a linear correlation between the maximum installed gas capacity and y_1 and is supported by the active subspace error e_{as} which increases as a function of n , however the confidence intervals suggest the model might admit a 2D subspace.

Figure 12 shows the global sensitivity and active subspace analysis for the maximum installed wind capacity max^{wind} with binary UC. The five global sensitivity metrics indicate consistent rankings of global sensitivities, and the relative ranking among sensitivities are similar to the relaxed UC results shown in Figure 8. Max installed wind is globally sensitive to cost of wind and reserve capacity margin. In comparison, the QoI is globally insensitive to cost of loss of load, cost of excess load, cost of natural gas, and transmission capacity but may be locally sensitive to these inputs. The Sobol' indices indicate that cost of loss of load, cost of excess load, cost of natural gas, and transmission capacity globally affect max installed wind *only* through interactions with other inputs, and cost of wind and reserve capacity margin affect maximum installed wind through *both* interactions with other inputs and independently. The relative gap between the first ($n = 1$) and second eigenvalue indicates the model admits a 1D active subspace where the first active variable $y_1 = \mathbf{w}_1^T \boldsymbol{\xi} \approx 0.11R^{cap} - 1c^{wind}$. The 1D sufficient summary depicts a strong linear correlation

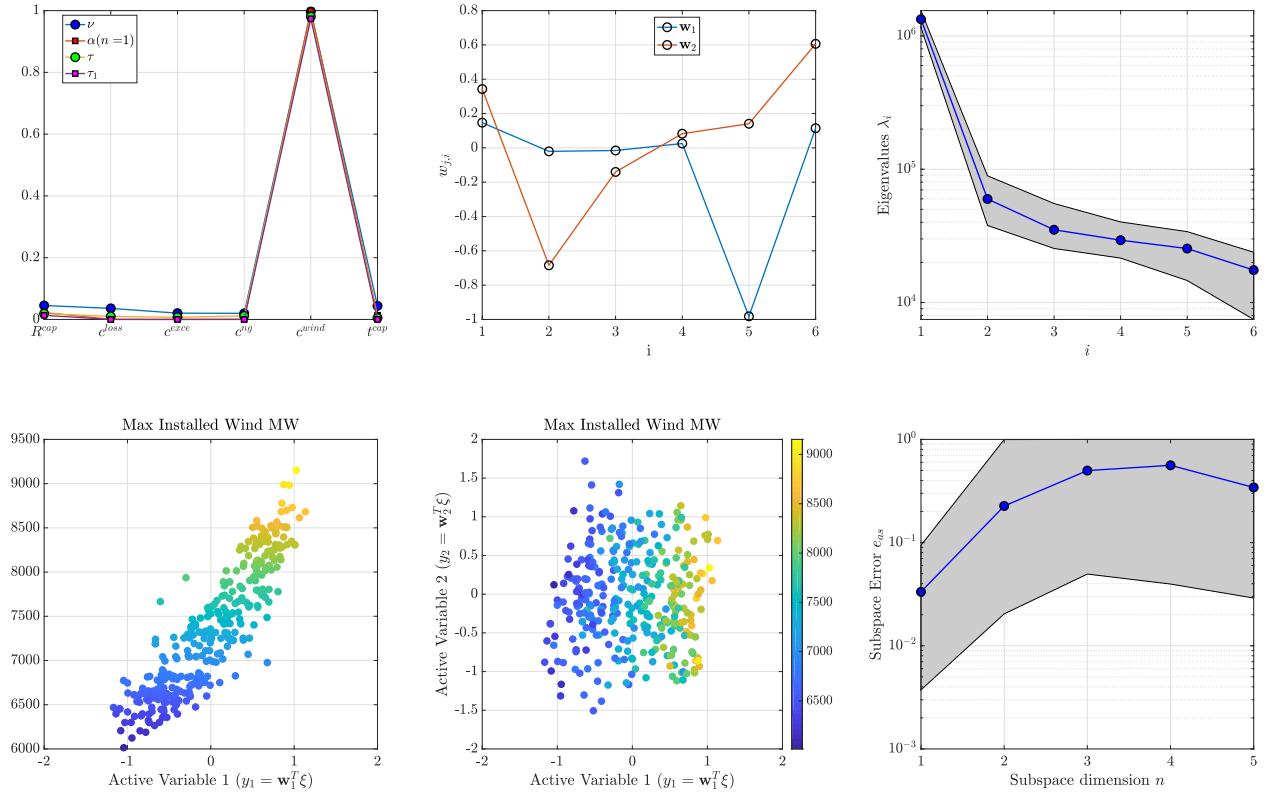


Figure 12. Sensitivity and active subspace analysis for the SPEED model maximum installed wind capacity max^{wind} with binary UC. The top row depicts global sensitivity metrics (left), active subspace eigenvectors (middle), and eigenvalues (right). The bottom row depicts 1D (left) and 2D (middle) sufficient summary plots, and the active subspace error (right).

between the QoI and y_1 . The 1D active subspace is further supported by the active subspace error e_{as} which increases as a function of n .

In summary, these results are remarkably similar to those considering the relaxed UC assumption shown in Section 4.1. This result was not necessarily expected as it was unclear what, if any, affect the UC modeling assumption would have on the model sensitivities. As in Section 4.1, the global sensitivity metrics indicate that all four QoIs are relatively insensitive to the input parameters cost of loss of load, cost of excess load, and cost of natural gas, while they are sensitive to reserve capacity margin, cost of wind, and transmission availability but in different ways. Specifically: expansion cost is driven largely by cost of wind and reserve capacity margin; operation cost is driven by cost of wind, transmission availability, and reserve capacity margin; maximum installed gas capacity is driven by reserve capacity margin, cost of wind, and transmission availability; and maximum installed wind capacity is driven by cost of wind and reserve capacity margin.

4.3 Exploiting the PC surrogates

So far we have only used the PC surrogate models to compute the Sobol' global sensitivity metrics according to (15) and (16), this is a good use for PC expansions but they provide more information. A PC surrogate model is made up of essentially two main components. The first, the polynomial basis, was known *a priori*, it is asymptotically orthogonal with respect to the distribution of our uncertain input parameters $f(\xi)$. The second, the vector of coefficients \hat{c} , was unknown and had to be approximated by randomly sampling $f(\xi)$ to produce an experimental design $\{\xi_i\}_{i=1}^N$, then feeding said design through the SPEED model to generate random values of the QoI. The coefficient vector \hat{c}

is expensive to compute initially, not because (12) or (13) is computationally expensive to solve, but because obtaining random values of the QoIs $\{u(\xi_i)\}_{i=1}^N$ requires running the SPEED model N times. When modeling relaxed UC running SPEED many times is feasible, on the NREL Eagle HPC, $N = 349$ samples took ~ 24 hours to complete. However, when modeling binary UC running SPEED many times is infeasible, e.g., on the NREL Eagle HPC, $N = 323$ samples took $\sim 10^3$ hours to complete. Fortunately, having computed approximate values of \hat{c} , and knowing the appropriate polynomial basis and probability density function $f(\xi)$ provides a computationally expeditious way to approximate values of the QoI without running the SPEED model. This method is computationally advantageous because estimating any single value of the QoI can be accomplished via the inner product of two P -dimensional vectors (6). Additionally, knowing the validated relative residual fit error e_{PC} informs us as to how much, on average, point-wise estimates of the QoI may differ from their true value. This surrogate modeling framework allows us to *predict* values of the QoI which are within our specified input parameter domain Ω , without the computational expense required to run the SPEED model.

We demonstrate how PC surrogates can be exploited to generate rich *posterior*⁸ distributions of the QoI, an additional benefit beyond providing global sensitivity analysis. To accomplish this we first draw a large experimental design $\{\xi_i\}_{i=1}^{10000}$ according to $f(\xi)$, then for each design point we compute an approximate value of the QoI via (6). This process can be repeated for any set of design points which define a subset of Ω , however the practitioner should be warned that samples which take on extreme values of sensitive model parameters are expected to have higher variance and assumed to be less accurate estimates, i.e., the further away the distribution of experimental design samples is from $f(\xi)$ the larger the prediction error. Figure 13 show histograms of the four SPEED model QoIs approximated by the PC surrogate models with relaxed and binary UC.

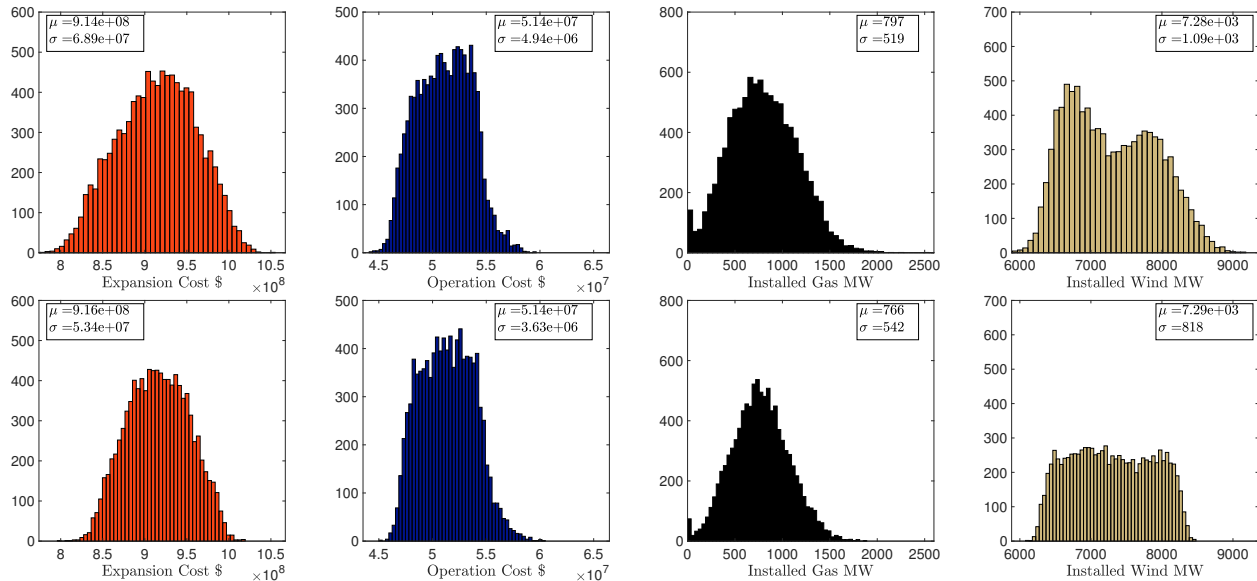


Figure 13. Histograms of the four SPEED model QoIs approximated by the PC surrogate models with *relaxed* (top) and *binary* UC (bottom).

Qualitatively, the distributions of expansion cost, operation cost, and maximum installed gas capacity appear similar, while the distribution of max installed wind capacity differs significantly. Specifically, the distribution of max installed wind for relaxed UC appears bimodal and right-skewed, while the distribution for binary UC shows no apparent skew or multiple modes. Quantitatively, the means of each distribution are similar for both relaxed and binary

⁸We use the term *posterior* loosely in the sense that, assuming independent identically distributed, unbiased, normal errors, maximum likelihood estimation is the same as ordinary least squares estimation. Moreover, the maximum a posteriori estimate is equivalent to the maximum likelihood estimate given a uniform prior distribution of the inputs (Smith 2013).

UC, however, we see noticeably larger standard deviations in both the distribution of expansion cost and max installed wind for relaxed UC compared to binary UC. Figures 8 and 12 clearly show that the UC modeling assumption does not change the global sensitivities of max installed wind capacity. However, only by exploiting the PC surrogate model to visualize rich posterior distributions can we see the significant difference in the overall distributions of the QoI, when considering relaxed vs. binary UC. It is also important to highlight that both PC surrogates for max installed wind capacity have validated errors on the order of 10^{-2} , which gives us measure of the degree of confidence we have in comparing these results. The relative point-wise validation errors of each PCE surrogate are reported in Tables 6 and 7.

5 Conclusions and Future Work

The aim of this work was to quantify the uncertainty within the SPEED model, and to provide insights into CEP at large. The relevant model output QoIs considered were expansion cost, operations cost, maximum installed gas capacity, and maximum installed wind capacity, however other aggregate outputs of a CEP model could also be considered. The uncertain model input parameters considered were the capacity reserve margin for expansion, cost of loss of load, cost of excess load, natural gas price, wind installation cost, and transmission capacity. Simulations are performed on the RTS-GMLC data set and solved with the SPEED model, a stochastic programming framework. The resulting data was post-processed, and sparse PC expansions, and active subspace analysis were employed to quantify the uncertainty. Five global sensitivity metrics, which quantify uncertainty by measuring an input parameter's influence on the variance of the QoIs are reported. PCE surrogate models were constructed and exploited to generate rich posterior distributions of the SPEED model output QoIs. To better understand the uncertainty associated with associated UC we conducted two independent numerical experiments holding all modeling conditions equal except for the integer modeling assumption regarding the dispatch decision variables, which has a significant impact on the model complexity. Our results support the following conclusions:

- The global sensitivity metrics indicate that all four QoIs are relatively insensitive to the input parameters cost of loss of load, cost of excess load, and cost of natural gas, while they are sensitive to reserve capacity margin, cost of wind, and transmission capacity but in different ways, see Sections 4.1 and 4.2.
- The global sensitivity metrics indicate that all four QoIs are influenced by cost of loss of load, cost of excess load, and cost of natural gas only through their interactions with other terms, if at all.
- Among the 6 uncertain inputs considered, expansion cost in this model is driven primarily by cost of wind and secondarily by reserve capacity margin.
- Among the 6 uncertain inputs considered, operation cost in this model is driven by primarily cost of wind, secondarily by transmission availability, and slightly by reserve capacity margin.
- Among the 6 uncertain inputs considered, maximum installed gas capacity in this model is driven primarily by reserve capacity margin, secondarily by cost of wind, and slightly by transmission availability.
- Among the 6 uncertain inputs considered, maximum installed wind capacity in this model is driven primarily by the cost of wind and secondarily by reserve capacity margin.
- When considering relaxed vs. binary UC, the posterior distributions of the expansions cost, operations cost, and maximum installed gas capacity predicted by the PC surrogates are similar, while the distribution of maximum installed wind capacity differs significantly.
- The mean value of each QoI's distribution are similar for both relaxed and binary UC, however we report larger standard deviations in the distributions of expansion cost, operations cost, and max installed wind capacity for relaxed UC compared to the binary UC model assumption. In contrast, the standard deviation for max installed gas capacity was smaller for relaxed UC compared to binary.
- The distribution of max installed wind for relaxed UC is bimodal and right-skewed, while the distribution for binary UC shows no apparent skew or multiple modes, potentially indicating that by relaxing the UC dispatch decision variables, CEP models could be significantly underestimating the amount of installed wind

capacity required to satisfy design constraints, e.g., transmission guidelines or safety constraints such as reserve capacity margin.

- Prior work has demonstrated that relaxing binary UC variables in PCMs alone does not necessarily reduce computational burden, contrary to expectations larger binary decision trees improved UC model resolution, and that results are likely solver dependent (Alemany, Kasprzyk, and Magnago 2018). Our results show significant reduction in the computational burden of the SPEED model by relaxing binary UC variables, see Section 4.3.

A natural question to ask given these results is, why are the driving sources of uncertainty considered in the SPEED model the reserve capacity margin, cost of wind installation, and transmission capacity? One possible explanation is that these uncertainties are associated either exclusively with expansion decisions in the case of reserve capacity margin and cost of wind installation, or partially in the case of transmission capacity, whereas the other relatively insensitive inputs are associated with dispatch decisions. This explanation would suggest that uncertainties associated with the first term in (2) are more significant on build decisions than uncertainties associated with the second term. A second possible explanation might follow the argument that the expansion decisions were more sensitive to the cost of wind installation than they were to the cost of natural gas, for instance, because wind generation is variable, and that uncertainties associated with variable generation will have greater significance. The reality is that the results of this work do not provide enough information to answer this question and a follow up study is warranted. Beyond this question, the implementation of the SPEED model in this study was relatively narrow both in the geographic location of the RTMS-GMLC data set, but also in the greater scheme of CEP and PCM research. Specifically, this study did not consider modeling the expansion of battery storage technologies or other variable generating technologies such as solar PV and hydro which are considered in the ReEDS model. In order to meet the growing demand for the UQ and prediction capabilities of CEP and PCM models we highlight the importance of continuing to study state-of-the-art methods in stochastic programming, data driven modeling, and UQ specifically to improve or better inform both the SPEED and ReEDS modeling framework, e.g. perhaps the ReEDS standard scenarios. This collaborative research effort is necessary to better understand the fundamental shift predicted in the future from traditional thermal generation to a more diverse electric grid in the U.S, and is particularly important regarding efforts to scale up existing modeling approaches to higher temporal or spatial fidelities.

In the future, we hope to further our efforts to quantify uncertainty within the SPEED model by investigating multiple *time fidelities*. As was previously mentioned, the SPEED model has the capability to substantially increase the fidelity of time-series data, or scenarios, compared to ReEDS. For example, the SPEED model is able to utilize data for all 365 days of a year, whereas ReEDS uses 17 scenarios (4 representative days) to model an entire year. Of particular interest, is whether or not increasing levels of fidelity significantly impact the model sensitivities or expansion decisions. Following this study we propose investigating this issue by comparing different surrogate models corresponding to different time-series sampling fidelities, similar to the comparison in this work between binary vs. relaxed UC. We also seek to allow the SPEED model to build different types of generating technologies such as solar PV or battery storage technologies, such additions or modifications to the model would prompt future global sensitivity analysis and model validation studies.

We highlight the need for future work to explore other potential surrogate modeling methods so that practitioners have a variety to choose from when presented with data that is disagreeable to one particular method or another. For other data-sets or models, problems using sparse PCEs can arise when the input samples obey potentially unknown, discrete, or correlated distributions. One potential solution to this problem might be changing from a continuous polynomial basis to a multiwavlet basis through the use of Piece-Wise Polynomial Chaos Expansions which provide an inherent multi-resolution modeling framework. Multiwavlet basis expansions have been used to study hyperbolic partial differential equations which can exhibit discontinuous shock-wave phenomenon (Pettersson, Iaccarino, and Nordstrom 2015).

References

- Adcock, Ben. 2017. “Infinite-Dimensional ℓ_1 Minimization and Function Approximation from Pointwise Data”. *Constructive Approximation* 45 (3): 345–390.
- Aleman, Juan, Leszek Kasprzyk, and Fernando Magnago. 2018. “Effects of binary variables in mixed integer linear programming based unit commitment in large-scale electricity markets”. *Electric Power Systems Research* 160:429–438.
- Aster, Richard C, Brian Borchers, and Clifford H Thurber. 2018. *Parameter estimation and inverse problems*. Elsevier.
- Berg, Ewout van den, and Michael P Friedlander. 2007. *SPGL1: A solver for large-scale sparse reconstruction*.
- Blatman, Géraud, and Bruno Sudret. 2011. “Adaptive sparse polynomial chaos expansion based on least angle regression”. *Journal of Computational Physics* 230 (6): 2345–2367.
- Bouchot, Jean-Luc, et al. 2015. “Compressed sensing Petrov-Galerkin approximations for parametric PDEs”. In *Sampling Theory and Applications (SampTA), 2015 International Conference on*, 528–532. IEEE.
- Candès, Emmanuel J, and Michael B Wakin. 2008. “An introduction to compressive sampling”. *IEEE signal processing magazine* 25 (2): 21–30.
- Candes, Emmanuel J, et al. 2011. “Compressed sensing with coherent and redundant dictionaries”. *Applied and Computational Harmonic Analysis* 31 (1): 59–73.
- Chkifa, Abdellah, et al. 2016. “Polynomial approximation via compressed sensing of high-dimensional functions on lower sets”. *arXiv preprint arXiv:1602.05823*.
- Cohen, Stuart M, et al. 2019. *Regional Energy Deployment System (ReEDS) Model Documentation: Version 2018*. Tech. rep. National Renewable Energy Lab.(NREL), Golden, CO (United States).
- Cole, Wesley J, et al. 2018. *2018 Standard Scenarios Report: A US Electricity Sector Outlook*. Tech. rep. National Renewable Energy Lab.(NREL), Golden, CO (United States).
- Cole, Wesley J, et al. 2020. *2019 Standard Scenarios Report: A US Electric Sector Outlook*. Tech. rep. National Renewable Energy Lab.(NREL), Golden, CO (United States).
- Constantine, Paul, and David Gleich. 2014. “Computing active subspaces with Monte Carlo”. *arXiv preprint arXiv:1408.0545*.
- Constantine, Paul, et al. 2016. “Python active-subspaces utility library”. *Journal of Open Source Software* 1 (5): 79.
- Constantine, Paul G. 2015. *Active subspaces: Emerging ideas for dimension reduction in parameter studies*. Vol. 2. SIAM.
- Constantine, Paul G, and Paul Diaz. 2017. “Global sensitivity metrics from active subspaces”. *Reliability Engineering & System Safety* 162:1–13.
- Cresta, Thierry, Olivier Le Maître, and Jean-Marc Martinez. 2009. “Polynomial chaos expansion for sensitivity analysis”. *Reliability Engineering & System Safety* 94 (7): 1161–1172.
- Dai, Wei, and Olga Milenkovic. 2009. “Subspace pursuit for compressive sensing signal reconstruction”. *IEEE Transactions on Information Theory* 55 (5): 2230–2249.
- Davenport, Mark A, and Michael B Wakin. 2010. “Analysis of orthogonal matching pursuit using the restricted isometry property”. *IEEE Transactions on Information Theory* 56 (9): 4395–4401.
- Diaz, Paul, Alireza Doostan, and Jerrad Hampton. 2018. “Sparse polynomial chaos expansions via compressed sensing and D-optimal design”. *Computer Methods in Applied Mechanics and Engineering* 336:640–666.
- Diaz, Paul Marcus. 2016. *Global sensitivity metrics from active subspaces with applications*. MA thesis, Colorado School of Mines. Arthur Lakes Library.
- Donoho, David L. 2006. “Compressed sensing”. *IEEE Transactions on information theory* 52 (4): 1289–1306.
- Doostan, Alireza, and Houman Owhadi. 2011. “A non-adapted sparse approximation of PDEs with stochastic inputs”. *Journal of Computational Physics* 230 (8): 3015–3034.
- Elad, Michael. 2010. “From Exact to Approximate Solutions”. In *Sparse and Redundant Representations*, 79–109. Springer.
- Eurek, Kelly, et al. 2016. *Regional energy deployment system (reeds) model documentation: Version 2016*. Tech. rep. National Renewable Energy Lab.(NREL), Golden, CO (United States).
- Fajraoui, Noura, Stefano Marelli, and Bruno Sudret. 2017. “Sequential design of experiment for sparse polynomial chaos expansions”. *SIAM/ASA Journal on Uncertainty Quantification* 5 (1): 1061–1085.

- Frew, Bethany, et al. 2019. “Sunny with a chance of curtailment: Operating the US grid with very high levels of solar photovoltaics”. *iScience* 21:436–447.
- Grigg, Cliff, et al. 1999. “The IEEE reliability test system-1996. A report prepared by the reliability test system task force of the application of probability methods subcommittee”. *IEEE Transactions on power systems* 14 (3): 1010–1020.
- Hampton, Jerrad, and Alireza Doostan. 2015. “Coherence motivated sampling and convergence analysis of least squares polynomial Chaos regression”. *Computer Methods in Applied Mechanics and Engineering* 290:73–97.
- Hansen, Per Christian. 1992. “Analysis of discrete ill-posed problems by means of the L-curve”. *SIAM review* 34 (4): 561–580.
- Hart, William E, et al. 2017. *Pyomo-optimization modeling in python*. Vol. 67. Springer.
- Hong, Tao, and Zhihui Zhu. 2018. “An efficient method for robust projection matrix design”. *Signal Processing* 143:200–210.
- Hristache, Marian, et al. 2001. “Structure adaptive approach for dimension reduction”. *The Annals of Statistics* 29 (6): 1537–1566.
- Jakeman, John D, Michael S Eldred, and Khachik Sargsyan. 2015. “Enhancing ℓ_1 -minimization estimates of polynomial chaos expansions using basis selection”. *Journal of Computational Physics* 289:18–34.
- Jakeman, John D, Akil Narayan, and Tao Zhou. 2017. “A generalized sampling and preconditioning scheme for sparse approximation of polynomial chaos expansions”. *SIAM Journal on Scientific Computing* 39 (3): A1114–A1144.
- Jell, Scott, and Michelle Bowman. 2018. “Almost all power plants that retired in the past decade were powered by fossil fuels”. *Washington, DC: US Energy Information Administration*.
- Jones, Brandon A, Nathan Parrish, and Alireza Doostan. 2015. “Postmaneuver collision probability estimation using sparse polynomial chaos expansions”. *Journal of Guidance, Control, and Dynamics* 38 (8): 1425–1437.
- Kucherenko, S, et al. 2009. “Derivative based global sensitivity measures and their link with global sensitivity indices”. *Mathematics and Computers in Simulation* 79 (10): 3009–3017.
- Mathelin, L, and KA Gallivan. 2012. “A compressed sensing approach for partial differential equations with random input data”. *Communications in computational physics* 12 (04): 919–954.
- Murphy, Caitlin, et al. 2019. *Electrification Futures Study: Scenarios of Power System Evolution and Infrastructure Development for the United States*. Tech. rep. National Renewable Energy Lab.(NREL), Golden, CO (United States).
- Myshetskaya, EE, et al. 2008. “Monte Carlo estimators for small sensitivity indices”. *Monte Carlo Methods and Applications mcma* 13 (5-6): 455–465.
- Needell, Deanna, and Joel A Tropp. 2009. “CoSaMP: Iterative signal recovery from incomplete and inaccurate samples”. *Applied and Computational Harmonic Analysis* 26 (3): 301–321.
- Needell, Deanna, and Roman Vershynin. 2010. “Signal recovery from incomplete and inaccurate measurements via regularized orthogonal matching pursuit”. *IEEE Journal of selected topics in signal processing* 4 (2): 310–316.
- Optimization, Gurobi. 2014. *Inc., “Gurobi optimizer reference manual,” 2015*.
- Pal, Dipan K, and Ole J Mengshoel. 2016. “Stochastic CoSaMP: Randomizing Greedy Pursuit for Sparse Signal Recovery”. In *Joint European Conference on Machine Learning and Knowledge Discovery in Databases*, 761–776. Springer.
- Peng, Ji, Jerrad Hampton, and Alireza Doostan. 2014. “A weighted ℓ_1 -minimization approach for sparse polynomial chaos expansions”. *Journal of Computational Physics* 267:92–111.
- . 2016. “On polynomial chaos expansion via gradient-enhanced ℓ_1 -minimization”. *Journal of Computational Physics* 310:440–458.
- Pettersson, M Per, Gianluca Iaccarino, and Jan Nordstrom. 2015. “Polynomial chaos methods for hyperbolic partial differential equations”. *Springer Math Eng* 10 (1007): 978–973.
- Pukelsheim, Friedrich. 2006. *Optimal design of experiments*. SIAM.
- Rauhut, Holger, and Christoph Schwab. 2017. “Compressive sensing Petrov-Galerkin approximation of high-dimensional parametric operator equations”. *Mathematics of Computation* 86 (304): 661–700.
- Rauhut, Holger, and Rachel Ward. 2012. “Sparse Legendre expansions via ℓ_1 -minimization”. *Journal of approximation theory* 164 (5): 517–533.

- Reimers, Andrew, Wesley Cole, and Bethany Frew. 2019. “The impact of planning reserve margins in long-term planning models of the electricity sector”. *Energy policy* 125:1–8.
- Sargsyan, Khachik, et al. 2014. “Dimensionality reduction for complex models via Bayesian compressive sensing”. *International Journal for Uncertainty Quantification* 4 (1).
- Schaeffer, Hayden, Giang Tran, and Rachel Ward. 2018. “Extracting sparse high-dimensional dynamics from limited data”. *SIAM Journal on Applied Mathematics* 78 (6): 3279–3295.
- Schiavazzi, Daniele, Alireza Doostan, and Gianluca Iaccarino. 2014. “Sparse multiresolution regression for uncertainty propagation”. *International Journal for Uncertainty Quantification* 4 (4).
- Smith, Ralph C. 2013. *Uncertainty quantification: theory, implementation, and applications*. Vol. 12. Siam.
- Tropp, J, and Anna C Gilbert. 2005. *Signal recovery from partial information via orthogonal matching pursuit*.
- Tropp, Joel A, and Anna C Gilbert. 2007. “Signal recovery from random measurements via orthogonal matching pursuit”. *IEEE Transactions on information theory* 53 (12): 4655–4666.
- Vimmerstedt, Laura J, et al. 2018. *2018 Annual Technology Baseline (ATB)*. Tech. rep. National Renewable Energy Lab.(NREL), Golden, CO (United States).
- Watson, Jean-Paul, David L Woodruff, and William E Hart. 2012. “PySP: modeling and solving stochastic programs in Python”. *Mathematical Programming Computation* 4 (2): 109–149.
- Watson, Jean-Paul, David L Woodruff, and David R Strip. 2008. “Progressive hedging innovations for a class of stochastic resource allocation problems”.
- West IV, Thomas K, and Serhat Hosder. 2014. “Uncertainty quantification of hypersonic reentry flows with sparse sampling and stochastic expansions”. *Journal of Spacecraft and Rockets* 52 (1): 120–133.
- Winokur, Justin, et al. 2016. “Sparse pseudo spectral projection methods with directional adaptation for uncertainty quantification”. *Journal of Scientific Computing* 68 (2): 596–623.
- Wiser, Ryan H, and Mark Bolinger. 2017. “2017 Wind Technologies Market Report”.
- Xiu, Dongbin, and George Em Karniadakis. 2002. “The Wiener–Askey polynomial chaos for stochastic differential equations”. *SIAM journal on scientific computing* 24 (2): 619–644.
- Yan, Liang, Ling Guo, and Dongbin Xiu. 2012. “STOCHASTIC COLLOCATION ALGORITHMS USING ℓ_1 -minimization”. *International Journal for Uncertainty Quantification* 2 (3).
- Yang, Xiu, and George Em Karniadakis. 2013. “Reweighted ℓ_1 minimization method for stochastic elliptic differential equations”. *Journal of Computational Physics* 248:87–108.
- Yang, Xiu, et al. 2016. “Enhancing sparsity of Hermite polynomial expansions by iterative rotations”. *Journal of Computational Physics* 307:94–109.