Option D involves the use of computer simulation software to predict facility energy use and savings. For options A, B, and C, standard errors and associated uncertainty have what is referred in mathematics as a “closed form” solution. An equation is said to be a closed form if it solves a given problem in terms of functions and mathematical operations. In cases where such functions are not available, the solution can be obtained through simulation modeling.

In this section, we present three different approaches to conducting simulations to estimate the uncertainty and precision of estimated savings:

* Local One-step-at-a-time (OAT) sampling analysis
* Global OAT screening method
* Bootstrap based analysis

## CASE A- New Construction of a Medium Office

### 1. Situation

This example involves the new construction of a medium office building in San Francisco, CA. Total electric and natural gas use and savings are to be analyzed for energy efficiency measures that reduce the following parameters in the baseline building by a certain percentage:

* Lighting Power Density (LPD)
* Electric Equipment Power Density (EPD)

### 2. Building Modeling and Analysis Tools

For this example, the OpenStudio[[1]](#footnote-1) (OS) tool chain (free and open source software) will be used. Building Energy Models (BEM) created in the OpenStudio format (.osm) can easily be manipulated by OpenStudio Measures,[[2]](#footnote-2) which are a set of programmatic instructions written in the Ruby programming language that makes changes to an energy model. These model changes can be as simple as changing an existing parameter in the model (e.g., change chiller Coefficient of Performance - COP) or as complicated as changing the entire HVAC system. This concept of creating and manipulating energy models through OpenStudio Measures will allow us to turn building characteristics into variables and determine their sensitivity and savings in a way that is consistent, scalable and easily shared with co-workers, clients or the general public as well as allowing for other algorithmic processes such as model calibration and optimization.

The *Medium Office* from the DOE Commercial Prototype Building Models[[3]](#footnote-3) will serve as our example building and the energy model will be programmatically generated using the OpenStudio Prototype Buildings Measure.[[4]](#footnote-4) This has the advantage that the reader can easily reproduce the workflow used to create the model and results shown in this section. The sensitivity analyses presented utilize the OpenStudio Parametric Analysis Tool (PAT)[[5]](#footnote-5) for problem definition and the cloud computing capabilities of OpenStudio Server[[6]](#footnote-6) for algorithm implementation and simulation runs. Both of these tools are free to use, however the actual cloud computing time from Amazon must be purchased from a user specific Amazon cloud computing account (Amazon EC2).

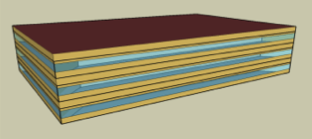
### 3. Energy Model and Variable Characterization

The energy model for this example is created by using the following inputs in the DOE Prototype Building Measure:

* Building Type: Medium Office
* Template: 90.1-2010
* Climate Zone: ASHRAE 169-2006-3C

where ‘Template’ determines the vintage and code requirements for the building. This results in a building that is *typical* of a new construction, medium office building in San Francisco, CA and is depicted in Figure 1.

Figure 1: Medium Office Prototype Building



Energy efficiency measures that reduce the building characteristics of Lighting Power Density LPD and Electric Equipment Power Density (EPD) by 10% will be applied to the building by using OpenStudio Measures that specifically manipulate those values in the energy model. To understand the effect that these parameters have on total building electricity and natural gas usage as well as understanding the sensitivity and uncertainty in the resulting energy savings for these measures, we first need to define valid ranges of the input uncertainty (a minimum and maximum) and a possible distribution type (triangular, uniform, normal, etc.) for each input parameter (LPD and EPD). These input uncertainties can come from many sources such as installation and implementation variations as well as other factors beyond our control. This means that even though we intend to reduce both LPD and EPD by exactly 10% in the building, from a practical standpoint, we won’t see a reduction of exactly 10% but more of a distribution of reductions around that 10% target reduction.

To simulate this effect we need to try and quantify the variation in the input that we will expect to have. For this example, we will consider a 5% variation in the 10% reduction that we wish to apply to the building. This gives us the minimum and maximum values for our distribution and bounds the input parameters at 5% and 15% of the baseline models values of LPD and EPD. We also need to choose a distribution shape to finish quantifying our input uncertainty. We could chose a flat or uniform change from 5% to 15%, a normal or bell curve distribution or one of several others that we think identifies the uncertainty we have about our ability to implement the energy efficiency measures. For this example, we will make the arbitrary choice of using a triangle distribution for the uncertainty that peaks to its maximum at a 10% reduction and tails off at the bounds as depicted in Figures 2-3 with the ranges listed in Table 1. This implies there is a higher probability of actually achieving the 10% reduction and a decreasing probability to achieve reduction values towards the bounds of 5% and 15%.

Table 1: Input Parameter Ranges for Energy Efficiency Measures

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Variable | | Minimum | Mode | Maximum |
| LPD (% reduction from baseline) | 5 % | | 10 % | 15% |
| EPD (% reduction from baseline) | 5 % | | 10 % | 15 % |

Figure 2: Lighting Power Density Distribution

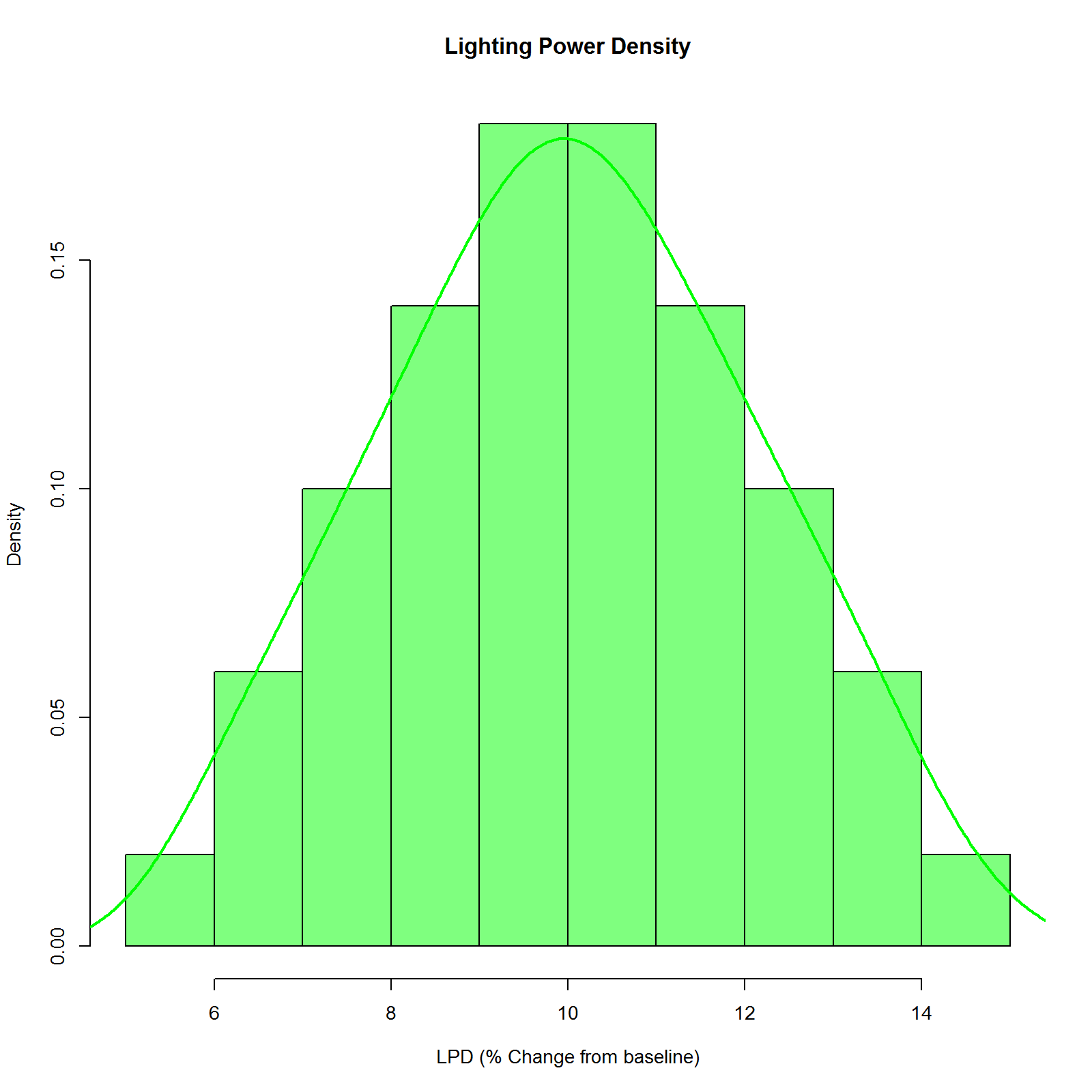
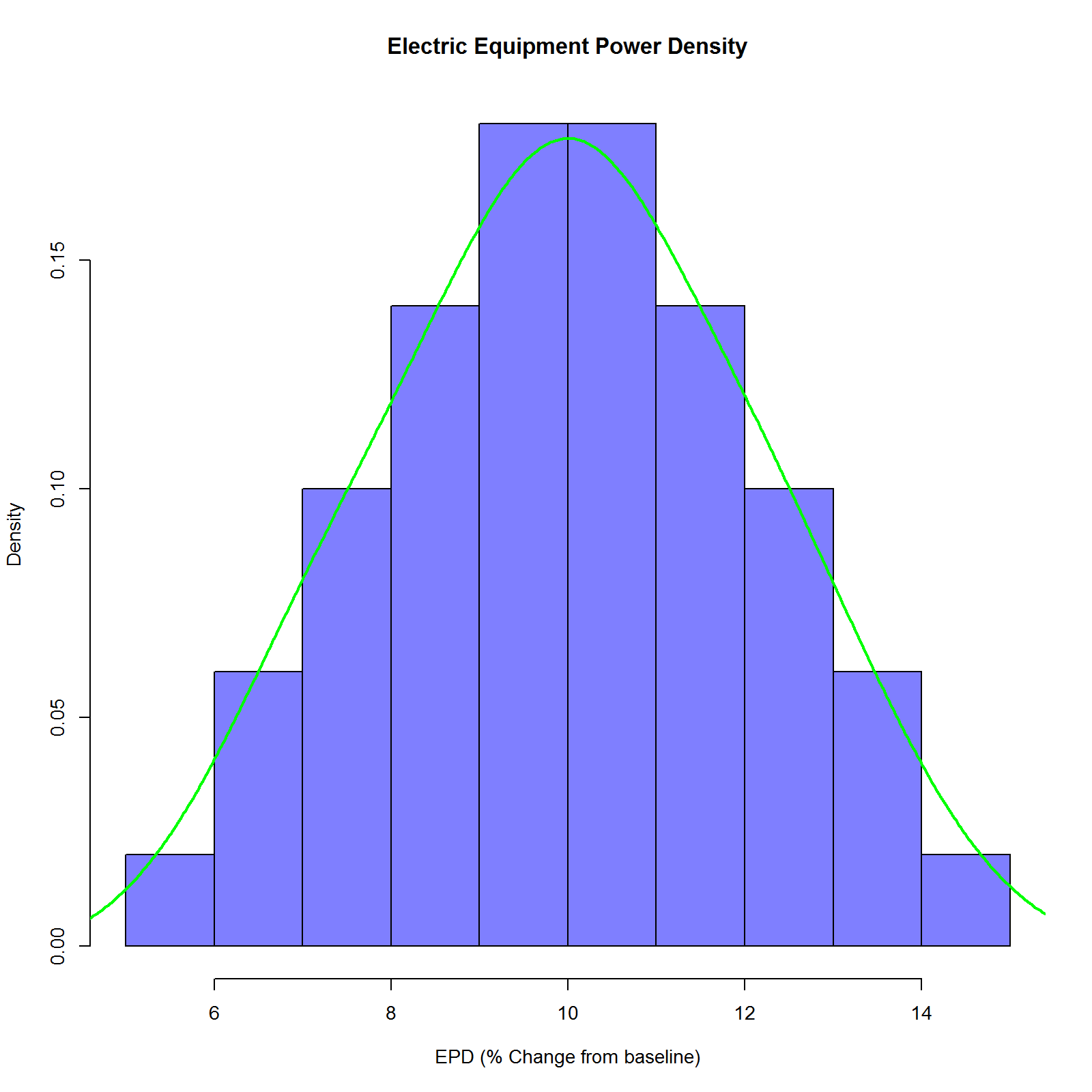


Figure 3: Electric Equipment Power Density Distribution



The actual OpenStudio Measures used in this example, along with the code used in this analysis can be downloaded from (XXX). The OpenStudio PAT allows for triangular, normal, log-normal and uniform distribution types. Ranges and distribution types for each variable are typically defined by project constraints or best practices. It should be noted that while the plotted distributions in Figures 2-3 do not appear to be triangles, that is just an artifact of plotting them using histogram and curve fit functions in R.

Next, we explore 3 methods to get an understanding of how the uncertainty in the inputs LPD and EPD translates to uncertainty and precision of estimated energy use and savings. All 3 methods involve different methods of discretizing the uncertainty distributions of the input variables and running building energy simulations at those discretized points.

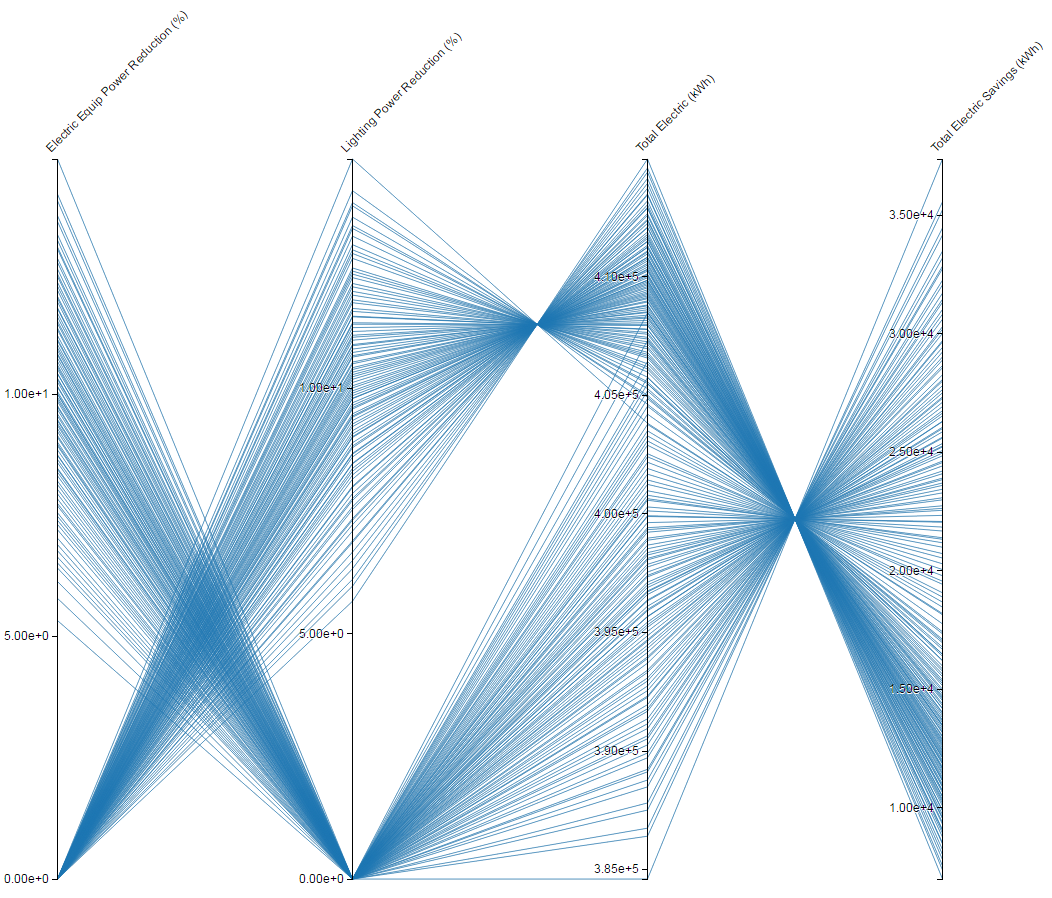
### 4. Local One-step-at-a-time (OAT) sampling analysis

Latin Hypercube Sampling[[7]](#footnote-7) (LHS) attempts to distribute sample points evenly over the sample space and can be applied to multiple variables or dimensions. The method essentially divides the distribution to be sampled into areas of equal probability and places an equal number of points into each area. This has the effect of giving a more even sampling across a distribution with fewer points compared to a purely random sample.

Using the OpenStudio Analysis Workflow we will sample the two variables, one at a time (OAT), while holding the other constant (at their default values) resulting in an OAT sampling method. This method is also described as a local method since we are holding all the non-sampled variables to their default values. Thus, while we are getting some sense of the effect of the perturbed variable, this perturbation is only with respect to the same locally fixed values for the other variables. The details on how to accomplish this can be found in the user guide for the OpenStudio Analysis Spreadsheet[[8]](#footnote-8) and the actual spreadsheet or PAT version used in this example can be found here.

Choosing the “correct” number of samples to use in the LHS method is very problem dependent. For example, a smooth linear problem would require very few samples to understand the variation in the output, while a highly nonlinear problem would require “lots” of points to characterize the nonlinear behavior of the problem. It is typical under these conditions to do a grid refinement study where a small number of points are chosen and used, and then twice as many points are used, and then twice again, this process repeating until the variation or characterization of the outputs becomes “good enough” for us to have confidence that we have chosen the “right” number of points. For illustration purposes, the number of samples in this example for each variable was chosen to be 100. This resulted in a total of 200 simulations to run (100 points x 2 variables = 200 total runs) and the solution space is depicted in the parallel coordinates plot in Figure 5. The first two coordinates are the variables: LPD and EPD. The next two coordinates are the dependent variables Total Electric Use (kWh) and Total Electric Savings (kWh) where savings is computed by subtracting the simulated value from the baseline models Total Electric Use. As can be seen in the parallel coordinate plots, this is an OAT method since half the simulations (there are only 2 variables) have each variable being held to its baseline value of 0% reduction.

Figure 5: LHS OAT Solution Space for Total Electric Use and Savings (kWh) due to changes in LPD and EPD



This visual inspection can be followed up by overlaying the resulting histograms of Total Electric Use and Savings variation due to the variation in the variables which is depicted in Figure 6-7. In Figure 6, the vertical line is the baseline model Total Electric Use and the sensitivity of each variable on Total Electric Use can be determined by looking at the support or base width of each colored distribution. For example, the green distribution is the Total Electric Use variation resulting from the distribution of LPD given in Figure 2. The blue distribution is the Total Electric Use variation resulting from the distribution of EPD in Figure 3. From these results we see that EPD results in greater Electric Savings while the uncertainty in that savings is much larger than the LPD case since the width of the distribution is much larger. It should be noted that the height of the distributions are such that the area under the distribution equals one so that each is a probability distribution which makes direct comparison easier.

Figure 6: Total Electric Use (kWh) Variation due to Variables

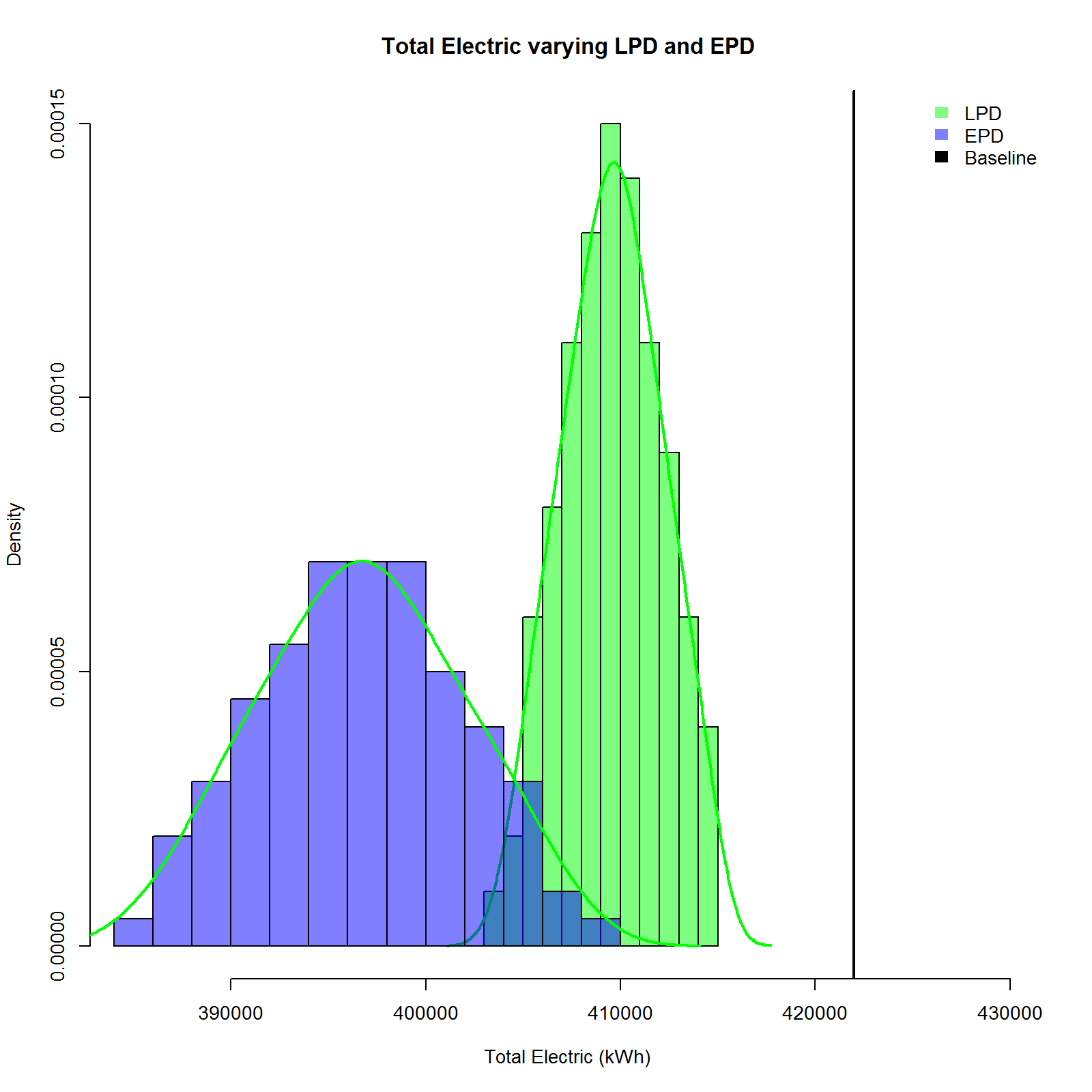
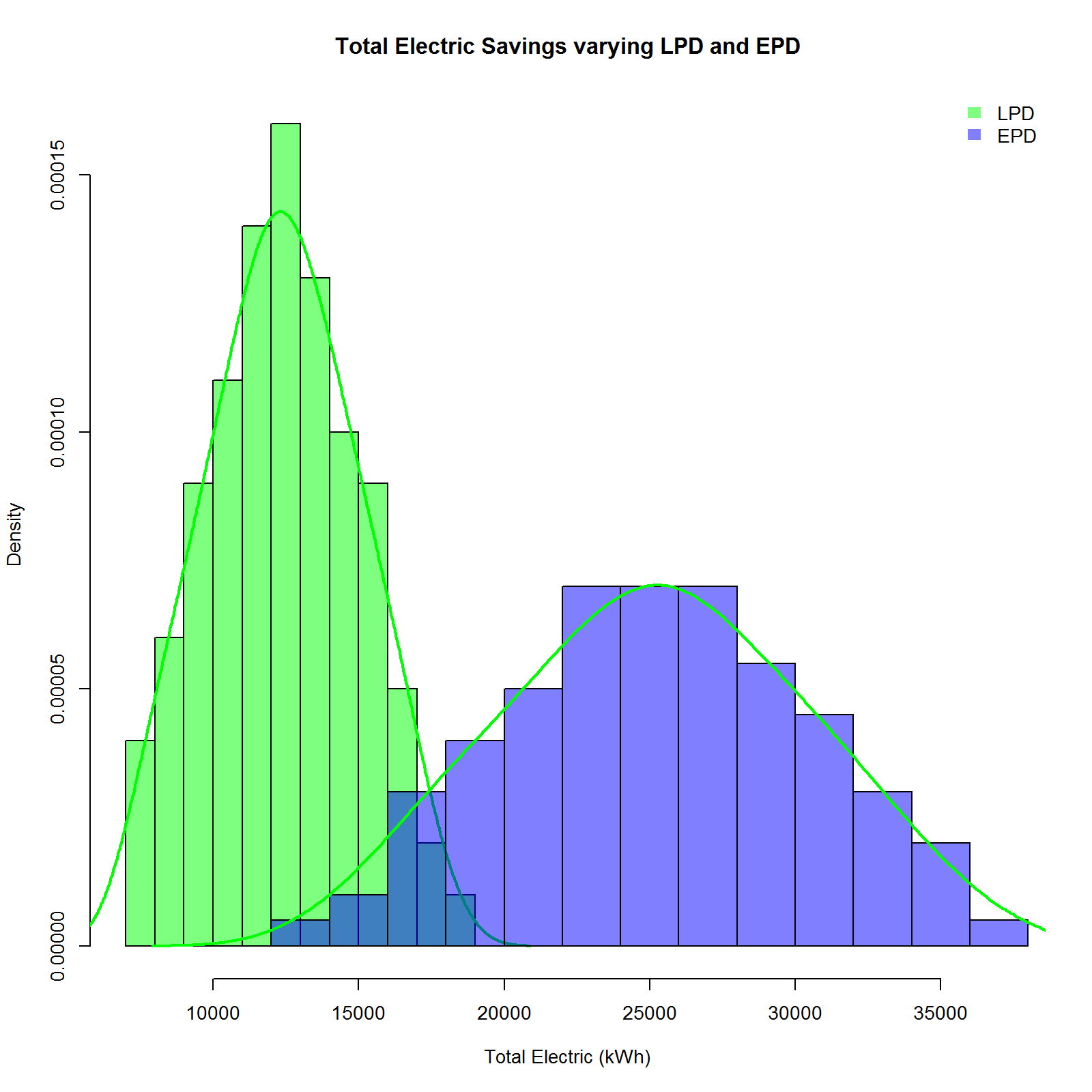


Figure 7: Total Electric Savings (kWh) Variation due to Variables

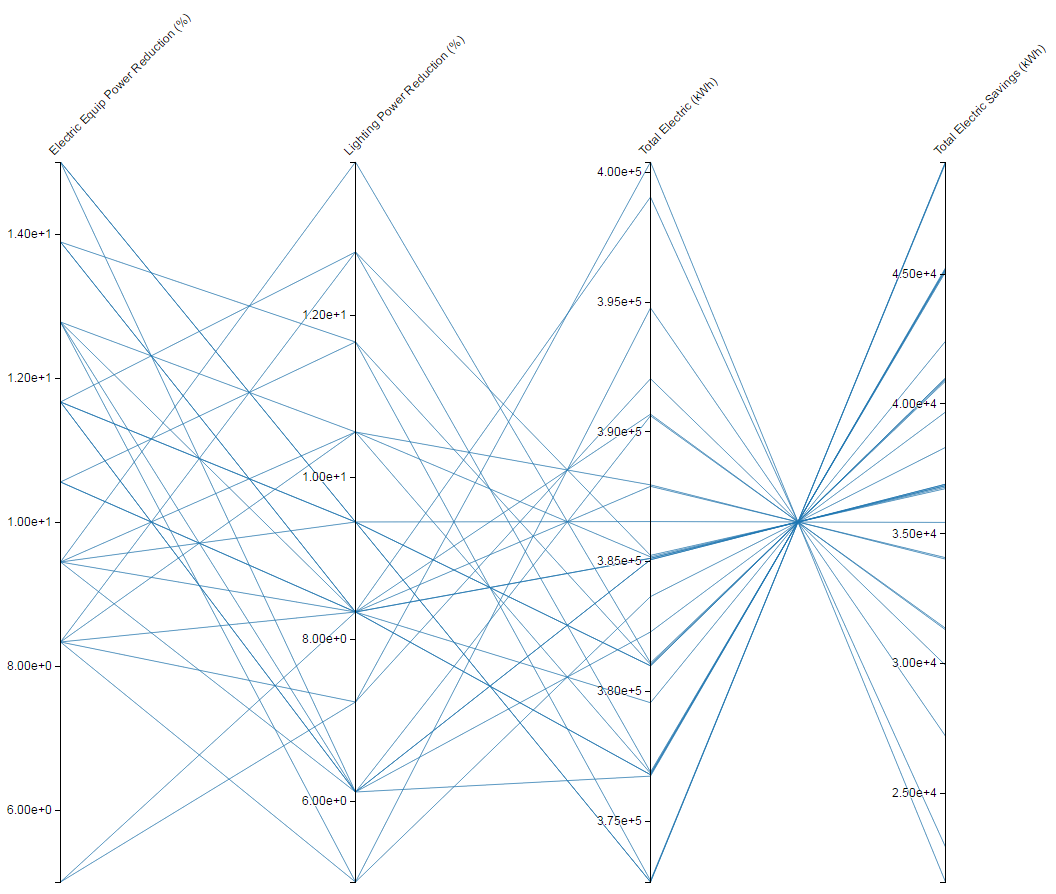


### 5. Global OAT screening method

To gain a more qualitative assessment of the sensitivity of the output variable to each input variable, we will use the Morris OAT Method.[[9]](#footnote-9) Morris’s elementary effects screening method identifies the few important factors or elementary effects (EE) at a cost of r × (p + 1) simulations (where *p* is the number of factors (in this example there are 2 factors or variables) and *r* is the number of EE computed per factor). The EE for each factor or variable is essentially a first order numerical derivative of the output variable of interest (in this example it is Total Electric Use and Savings) with respect to the factor. Thus the EE calculates the finite-difference of the output variable changes when the input factor is varied at specific points in the parameter space; however the step size is typically chosen to be much larger than one would chose for an approximation to a derivative. This calculation is repeated *r* times at different locations in the parameter space to get a global sense of how sensitive each factor or input variable is with respect to the outputs. The Morris Method implementation used in the OpenStudio Server comes from the R package *Sensitivity,* which includes improvements to the original method defined by Morris, specifically a space-filling optimization of the design[[10]](#footnote-10) and a simplex-based design.[[11]](#footnote-11)

For illustrative purposes, the *r* value for this example was set to 10 resulting in a total number of 30 simulations. The resulting sampling pattern of the solution space is plotted in Figure 8. Since *r* was set to 10 for this example, there will be 10 calculations of the EE for the variables, with each one coming from a *different initial condition* in the solution space (computed by various algorithmic settings) depicted in Figure 8, thus making the method more global in nature. Typically *r* is chosen between 4-10, depending on the number of variables and the computational cost of the model.

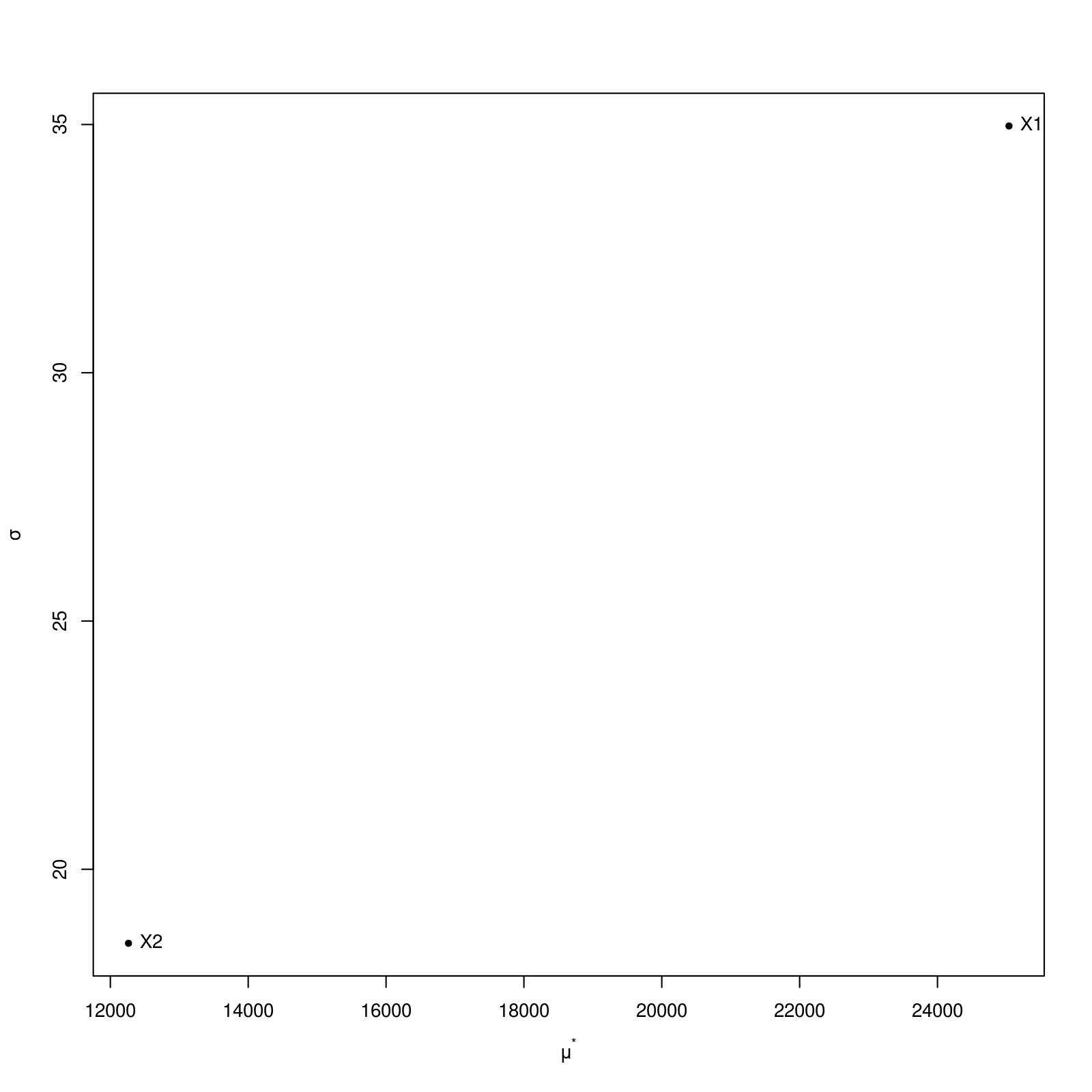
Figure 8: Total Electric Use and Savings (kWh) Sampling pattern for Morris Method



The main output of the Morris Method is the mean and standard deviation of the EEs for each variable. In this example, the 10 EEs are averaged and the standard deviation is computed and denoted µ and σ, respectively. It is typical to take the absolute value of the EEs before averaging to eliminate any possible canceling of values if the EE results are both positive and negative. This is the typical output value for the Morris Method and is denoted µ\* along with the value of σ which characterizes the variation of the EEs. A low value of σ typically means the EE value is relatively consistent and that there is negligible interaction with other possible factors or variables. A higher σ value typically means the EE value is not consistent and there could be nonlinear effects at play and/or there are possible interactive effects with other factors or variables.

The µ\* and σ values for our example are plotted in Figure 9. These results are consistent with the LHS approach where EPD has the larger Energy Savings yet has a greater uncertainty than LPD, however the results are computed with far less simulations than the LHS case and are more global in nature.

Figure 9: µ\* vs σ. X1 is EPD and X2 is LPD

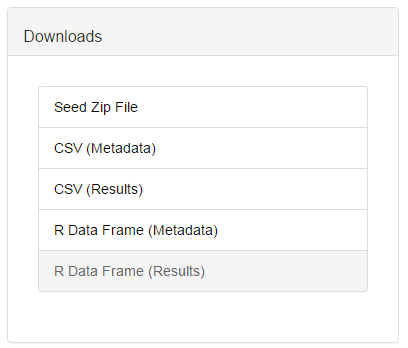


### 6. Bootstrap based analysis

In this section, we attempt to compute the precision of the estimated Total Electric Use and Savings due to both the energy efficiency measures we are applying to the building.

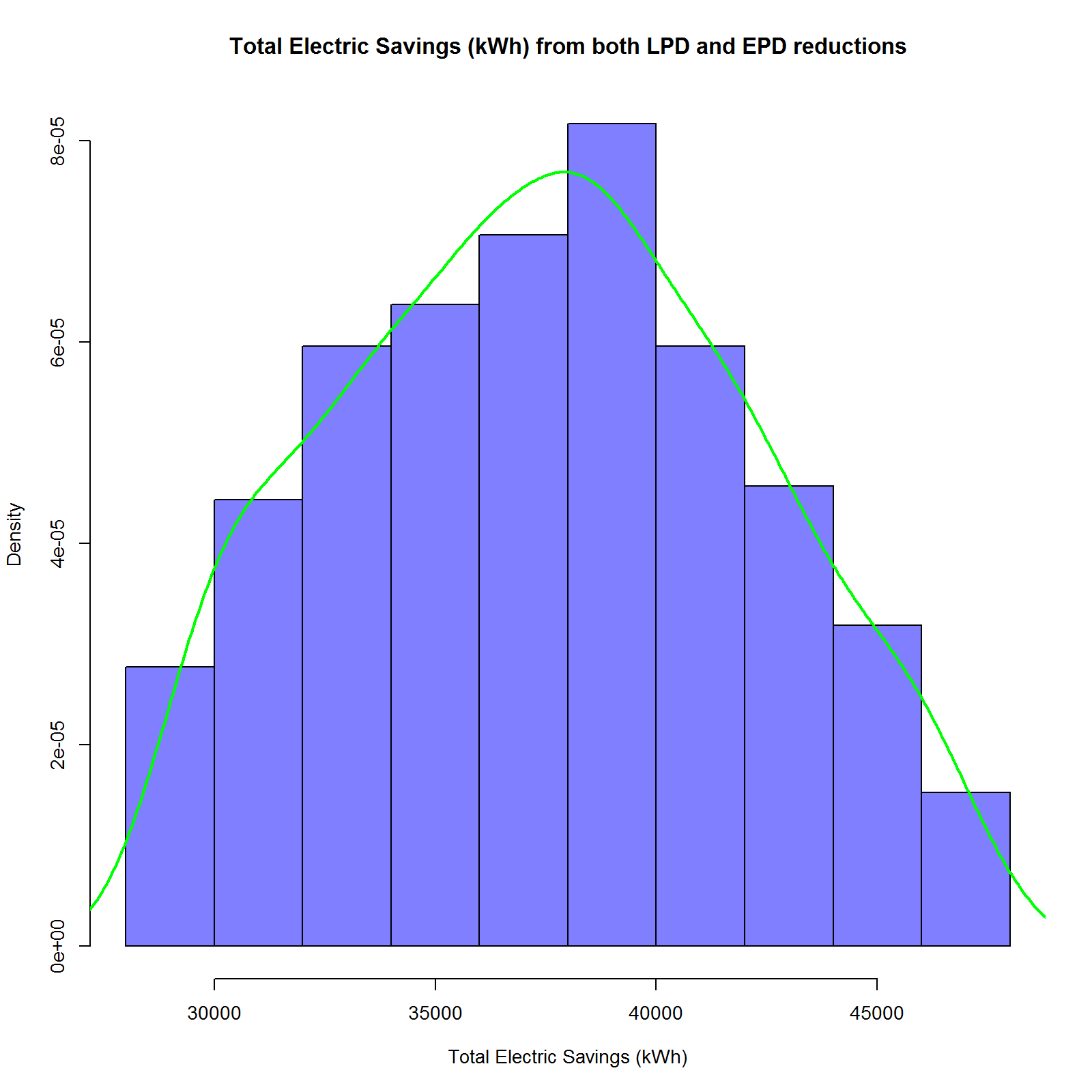
To create the data necessary to examine such a model, we will again resort to using LHS, however this time we will **not** do a OAT method but instead sample **all** the Input Variables at the **same** time. For illustrative purposes, we will use 400 as the sample size with the resulting solution space being much denser than the OAT sampling depicted in Figure 5. After running the 400 samples using OS Server, an R data frame can easily be obtained by clicking the *download R data frame (Results)* button on OS Server as depicted in Figure 10.

Figure 10: Download R dataframe from OS Server



To do the bootstrapping analysis, we first sort the resulting Electric Savings values from 400 simulation runs from low to high values and pick the 21st – 381st values as the range of savings. The resulting Total Electric Savings from both LPD and EPD measures are plotted in Figure 11.

Figure 11: Total Electric Savings (kWh) from both LPD and EPD measures



Next, the mean value for that limited set is then computed as well as the measured distance from the mean to the 21st value of Electric Savings, which is called the absolute savings. The relative precision is then the absolute savings divided by the mean. The results for this example are listed in Table 2.

**Table 2: Bootstrap Results**

|  |  |  |
| --- | --- | --- |
| Absolute Energy Savings | | **8460 kWh** |
| Relative Precision | **22 %** | |

1. https://www.OpenStudio.net [↑](#footnote-ref-1)
2. http://nrel.github.io/OpenStudio-user-documentation/getting\_started/about\_measures/ [↑](#footnote-ref-2)
3. https://www.energycodes.gov/commercial-prototype-building-models [↑](#footnote-ref-3)
4. https://github.com/NREL/openstudio-standards [↑](#footnote-ref-4)
5. http://nrel.github.io/OpenStudio-user-documentation/reference/parametric\_analysis\_tool\_2 [↑](#footnote-ref-5)
6. https://github.com/NREL/OpenStudio-server [↑](#footnote-ref-6)
7. M. Mckay, R. Beckman and W. Conover, “A comparison of three methods for selecting values of input variables in the analysis of output from a computer code,” Technometrics, vol. 21, no. 2, pp. 239-245, May. 1979 [↑](#footnote-ref-7)
8. https://github.com/NREL/OpenStudio-analysis-spreadsheet/tree/develop/documentation [↑](#footnote-ref-8)
9. Morris, M.D. (1991). "Factorial Sampling Plans for Preliminary Computational Experiments" (PDF). Technometrics 33: 161–174. [↑](#footnote-ref-9)
10. F. Campolongo, J. Cariboni and A. Saltelli, 2007, An effective screening design for sensitivity,

    Environmental Modelling \& Software, 22, 1509–1518. [↑](#footnote-ref-10)
11. G. Pujol, 2009, Simplex-based screening designs for estimating metamodels, Reliability Engineering and System Safety 94, 1156–1160. [↑](#footnote-ref-11)