1. Introduction

This section details the underlying assumptions of the current model. It will be updated as model development continues. The model is a two stage stochastic model. In the first stage of the model, investments are made in the capacity of power plants. In the second stage of the model, operation decisions are made on how much to run each power plant.

Time periods: 1 Nodes: 1 Technologies: 2 Stages: 2 Stochastic: Not yet

To do:

(1) divide capacity into existing capacity and new capacity
(2) add code to R PTDF file to add a column of zeros for the reference column; remove the set NN
(3) the current iteration of the model parameterizes y as the capacity used instead of hours used

2. Model

Indices and sets

$h = 1, \ldots, H$ index of the operating modes in the load duration curve
$l = 1, \ldots, L$ index of available transmission lines
$n = 1, \ldots, N$ index of the set of buses
$nn = 1, \ldots, NN$ subset of the set n that excludes the reference node
$p = 1, \ldots, P$ index of the available technologies
$t = 1, \ldots, T$ index of the periods in the model
$s = 1, \ldots, S$ index of scenarios

Parameters

a_p = forced outage rate of technology p
b = budget available for investment in new power production technology
c_p = marginal cost of power production technology p in USD per MWh
$d_{h,n,s}$ = demand at node n for operating level h in scenario s
k_p = fixed cost of power production technology p in USD per MW
λ = average capacity in MW of the target Poisson distribution
$l.line_l$ = lower bound on transmission line l capacity
$u.line_l$ = upper bound on transmission line l capacity
$LOLC_C$ = exogenously determined loss of load probability threshold
VLL = value of lost load, the penalty cost for unserved energy in \$ per MWh
$size_p$ = average plant size in MW for power production technology p
$prob_{n,h,s}$ = probability of each scenario (assuming discrete distribution)

Variables

$m_{h,n,s}$ = the import or export of power to or from node n for operating mode h under scenario s
$UE_{h,n,s}$ = unserved energy at node n during operating level h for scenario s
$x_{n,p}$ = number of each type of each power production technology p to build at node n [integer variable]
$y_{h,n,p,s}$ = capacity of power production technology p used at node n during operating mode h under scenario s

$$\underset{x_{n,p}, y_{h,n,p,s}}{\text{minimize}} \quad \sum_{p=1}^{P} k_p \cdot size_p \cdot \sum_{n=1}^{N} x_{n,p} + \sum_{s=1}^{S}\sum_{n=1}^{N}\sum_{h=1}^{H} prob_{h,n,s} \cdot \sum_{p=1}^{P} \Big(c_p \cdot y_{h,n,p,s} + VLL \cdot UE_{h,n,s} \Big)$$

subject to

$$\sum_{p=1}^{P} y_{h,n,p,s} + m_{h,n,s} + UE_{h,n,s} = d_{h,n,s} \ \forall \, h, n, s$$

$$\sum_{n=1}^{N}\sum_{p=1}^{P} x_{n,p} \cdot a_p \cdot \text{size}_p \geq \lambda$$

$$\sum_{h=1}^{H} y_{h,n,p,s} \leq a_p \cdot x_{n,p} \cdot size_p \cdot 8760 \quad \forall \, n, p, s$$

$$\sum_{p=1}^{P} k_p \cdot size_p \cdot \sum_{n=1}^{N} x_{n,p} \leq b$$

$$LOLP \leq LOLP_C$$

$$\sum_{l=1}^{L} m_{h,n,s} = 0$$

$$\sum_{n=1}^{N} PTDF_{l,n} \cdot m_{h,n,s} \leq u.line_l \quad \forall \, l$$

$$-\sum_{n=1}^{N} PTDF_{l,n} \cdot m_{h,n,s} \leq l.line_l \quad \forall \, l$$

$$x_{n,p}, y_{h,n,p,s} \geq 0$$

3. Explanation of the constraints

3.1. Node balance constraints.
The demand at node n during operating level h of scenario s can be met by generators at other nodes in the system and transmitted to the node, generators at the node, or can go unmet.

$$\sum_{p=1}^{P} y_{p,n,h,s} + m_L + UE_{h,n,s} = D_n \quad \forall \, h, n, s \tag{1}$$

3.2. Transmission line capacity.

$$\text{PTDF} \cdot \ldots \leq T_L \quad \forall \, L \tag{2}$$

3.3. Budget constraint.

$$\sum_{p=1}^{P} k_p \cdot size_p \cdot \sum_{n=1}^{N} x_{p,n} \leq b \tag{3}$$

3.4. **Transmission system: Energy balance constraint.** The sum of the flow across the transmission system must equal zero.

$$(4) \qquad \sum_{l=1}^{L} m_l = 0$$

4. METHODS FOR INCORPORATION OF RELIABILITY

4.1. **Average forced outage rate capacity adjustment.** Power production from technology p in each period is constrained to be less than the average forced outage rate of the technology p multiplied by the total installed capacity of technology p at node n.

$$(5) \qquad y_{p,n,h,s} \leq a_p \cdot x_{p,l} \cdot s_p \quad \forall \, h, s$$

4.2. **Cumulants.**

$$(6) \qquad (1 - ls) \cdot \Big(\sum_{n=1}^{N} a_p \cdot size_p \cdot x_{p,n}$$

4.3. **Poisson approximation.** The Poisson approximation method discussed here approaches the problem of identifying an optimal portfolio of generation assets from the perspective of the targeted portfolio of generators, rather than properties of the individual generation units. Each power plant is modeled as having two states: either on or off with the likelihood of each state determined by the forced outage rate (FOR) parameter a_p. The model selects the number of each power production technology to build p at each node n such that a system of constraints governing the operation of the network are met.

We model the likelihood of an individual generation unit being able to turn on as following a Bernoulli distribution. The sum of Bernoulli random variables is a binomial random variable if the probability of success is the same for each Bernoulli R.V. If the probability of success is different amongst the Bernoulli R.V.'s than their sum follows a Poisson Binomial R.V., which is generally characterized as having fatter tails as compared to a Poisson R.V. The Poisson Binomial distribution can be approximated by the Poisson distribution with the quality of the approximation depending on the success probability and number of generation units (according to Le Cam's Theorem).

The Poisson distribution is a single parameter distribution defined by its mean λ. This means that each Poisson distribution, as defined by its λ, value has a unique PDF and CDF. We leverage this information to create a constraint that ensures system reliability is above a selected threshold.

Power system planning involves the determination of an appropriate outage frequency by planners. This is because it becomes prohibitively more expensive to ensure redundancy the higher the probability of success climbs (i.e., the more redundancy built into the system the higher the costs of operating said system). This cost/reliability tradeoff is a central part of any engineered system. In the United States power systems are designed to have system demand exceed system capacity 3 hours a year. With this information, as well as the load duration curve, we can calculate the λ of the target Poisson distribution. We used the R software package to identify the Poisson distribution whereby the percentile of interest matched with the quantity in the load duration curve. It is important to remember that

in the Poisson distribution this is a unique mapping but in the Poisson Binomial distribution, as it is a two parameter distribution, there are an infinite number of mappings.

The constraint to be implemented in the model is that the weighted average available capacity is larger than λ see Equation 7. This constraint leverages information that we necessarily have before we even run the model (i.e., assumptions of future load demand) while also avoiding the need to parameterize reserve capacity into the model. One issue with this approach is the distance between our approximation and the actual Poisson Binomial distribution. I propose that we implement the model and conduct numerical experiments to identify the gap.

$$(7) \qquad \sum_{p=1}^{P} a_p \cdot size_p \cdot \sum_{n=1}^{N} x_{p,n} \geq \lambda$$