1 GEP problem development

1.1 Simple GEP statement

A simple statement of a single-period, single area generation expansion planning (GEP) problem is as follows:

Problem GEP-1:

\[
\begin{align*}
\min & \sum_{j} I_{j} C_{j}^{add} + \sum_{j} F_{j} C_{j} H_{j} P_{j} T \\
\text{subject to} & \quad C_{j} = C_{j}^{existing} + C_{j}^{add} \quad \forall j \\
& \quad \sum_{j} P_{j} = d \\
& \quad 0 \leq P_{j} \leq C_{j} \quad \forall j \\
& \quad C_{j}^{add} \geq 0 \quad \forall j
\end{align*}
\]

(1)

Variable definitions follow:

- \( j \): index over technologies
- \( I_{j} \): investment cost for technology \( j \), $/MW
- \( C_{j}^{add} \): capacity added for technology \( j \), MW
- \( F_{j} C_{j} \): fuel cost for technology \( j \), $/MBTU
- \( H_{j} \): heat-rate for technology \( j \), MBTU/MWhr
- \( P_{j} \): Power generation level for technology \( j \), MW
- \( T \): Planning horizon, hours

In GEP-1, (1) is the objective function which minimizes the investment costs plus the operational costs. There is no need for
accounting for the time value of money here since this is only a single period formulation.

- Equation (2) computes total capacity from existing and added capacity, for each technology $j$.
- Equation (3) expresses power balance for just a single demand level.
- Equation (4) requires power generation level to be within unit capacity.
- Equation (5) imposes non-negativity on the $C_j^{add}$ variables.

The decision-variables for this problem are $C_j^{add}$ and $P_j$.

### 1.2 Reserve constraint

Problem GEP-1 imposes only that capacity be sufficient to meet the demand, via (3) and (4). However, we also need reserves. Therefore we include a reserve constraint via (6), according to Problem GEP-2:

**Problem GEP-2:**

\[
\text{min} \sum_{j} I_j C_j^{add} + \sum_{j} FC_j H_j P_j T
\]

subject to

\[
C_j = C_j^{existing} + C_j^{add} \quad \forall j
\]

\[
\sum_{j} P_j = d
\]

\[
0 \leq P_j \leq C_j \quad \forall j
\]

\[
C_j^{add} \geq 0 \quad \forall j
\]

\[
\sum_{j} C_j(t) \geq (1+r)d
\]

Here, $r$ is the fraction required of capacity over demand.
1.3 Capacity credit

The capacity credit of a resource is used to identify the percentage of the resource’s capacity which should be identified for reliability calculations at peak load. The capacity credit indicates the availability at peak load.

The capacity credit for wind is usually fairly low because wind generation during daytime hours is usually lower. For example, MISO was using 13.3% capacity credit for wind [1].

Capacity credits for solar are generally higher because peak loads usually occur during the daytime when solar insolation is highest.

We modify (6) to account for capacity credit, CC_j, for each technology j.

Problem GEP-3:

\[
\begin{align*}
\min &\sum_j I_j Cap_j^{add} + \sum_j FC_j H_j P_j T \\
\text{subject to} &\quad Cap_j = Cap_j^{existing} + Cap_j^{add} \quad \forall j \\
&\quad \sum_j P_j = d \\
&\quad 0 \leq P_j \leq CC_j Cap_j \quad \forall j \\
&\quad Cap_j^{add} \geq 0 \quad \forall j \\
&\quad \sum_j CC_j Cap_j(t) \geq (1 + r)d
\end{align*}
\]

(1)  
(2)  
(3)  
(4)  
(5)  
(6a)
1.4 Accounting for different demands

The GEP statements defined so far require one to choose a particular demand \( d \). In this case, there are problems encountered, as follows:

- If \( d \) is chosen to be the peak load, then the GEP will build the right amount of capacity but will over-estimate the energy requirements and corresponding generation production.
- If \( d \) is chosen to be the average demand, then the GEP will identify the right energy requirements and corresponding generation production but will underestimate the capacity.

In both cases, the solution will identify the one technology that minimizes the sum of investment plus fuel costs.

An improvement on this formulation is to increase the number of load levels, or load *blocks*, so that we use \( d_s, s=1… \) instead of just \( d \). In doing this, we must also identify the appropriate duration \( h_s \) for each load block.

A tool to use in identifying appropriate load blocks and corresponding durations is the load duration curve (LDC). We describe development of LDCs in the Appendix. Figure 1 illustrates an LDC, where we observe that it provides on the abscissa the number of hours the load is expected to be greater than or equal to the corresponding load given as the ordinate.

In Figure 1, for example, we observe that the load is greater than or equal to 9 MW for 2628 hrs/year, 7 MW for 6132 hrs/year, and 5 MW for 8760 hrs/year (i.e., it is always greater than or equal to 5 MW).
We can understand the LDC from a different perspective if we first normalize the abscissa of Figure 1 by dividing it by 8760, as shown in

**Figure 2: LDC with normalized abscissa**

Now flip the axes of the plot in Figure 2 to show the plot in Figure 3. This plot is a cumulative distribution function (CDF) which provides probabilities that the load will be greater than or equal to the value of the abscissa.
Figure 3: Cumulative Distribution Function

For example, we observe from Figure 3 that
\[ \Pr(\text{Load} \geq 7) \approx 0.45 \]
\[ \Pr(\text{Load} \geq 8) \approx 0.30 \]

We find from the appendix that the actual values are
\[ \Pr(\text{Load} \geq 7) = 0.478 \]
\[ \Pr(\text{Load} \geq 8) = 0.304 \]

We now show how to utilize the LDC in our GEP formulation. Consider the notation associated with the LDC as shown in Figure 4. Only 3 load blocks are shown, but the formulation generalizes to any desired number of load blocks. GEP-1, GEP-2, and GEP-3 can be viewed as special cases where there was only one load block.
The GEP, modified to account for the different number of load blocks, is provided below.

Problem GEP-4:

\[
\begin{align*}
\min & \sum_{j} I_{j} Cap_{j}^{add} + \sum_{j} FC_{j} H_{j} \sum_{s} P_{j,s} h_{s} \\
\text{subject to} & \quad Cap_{j} = Cap_{j}^{existing} + Cap_{j}^{add} \quad \forall j \\
& \quad \sum_{j} P_{j,s} = d_{s} \quad \forall s \\
& \quad 0 \leq P_{j,s} \leq Cap_{j} \quad \forall j, s \\
& \quad Cap_{j}^{add} \geq 0 \quad \forall j \\
& \quad \sum_{j} CC_{j} Cap_{j}(t) \geq (1+r)d_{1}
\end{align*}
\]  

Observe that the following changes were made:
- The energy term \(P_{j,T}\) in (1) was modified in (1a) to the summation over \(s\) of \(P_{j,s} h_{s}\).
• The summation over \( j \) of the operational values \( P_j \) in (3), required to equal \( d \), was modified in (3a) so that such a summation occurs for each load block \( s \).

• The constraints on each operational value \( P_j \) to not exceed capacity in (4) was modified in (4a) so that these constraints are imposed for each load block \( s \).

• The load \( d \) in the reserve constraint (6a) (which only dealt with a single load block) is specified in (6b) to be the peak load \( d_1 \). Only (2) and (5) remain the same.

There is one other issue that deserves some attention.

We generally know when the various load levels \( s=1,2,3 \) occur. For example, the peak condition usually occurs in mid-afternoon, the partial peak condition in the mid-evening, and the off-peak condition in the middle of the night. We may generalize the capacity credit so that it is chosen for each of these time periods. Thus, our capacity credit nomenclature changes from being specific to each technology, \( CC_j \), to being specific to each technology and load block, \( CC_{j,s} \).

If we do not do this, we will tend to obtain unrealistic values of generation levels \( P_{j,s} \), e.g., we may obtain significantly high values of wind output during load block \( s=1 \) (which corresponds to the daytime period), or worse yet, significant solar output during load block \( s=3 \) (which corresponds to the night-time period).

We will still utilize the capacity credit for the peak period, \( s=1 \), for the reserve constraint.
These changes lead us to Formulation GEP-5. Formulation GEP-5:

\[
\begin{align*}
\min & \sum_j {I_j}Cap_j^{add} + \sum_j FC_j H_j \sum_s P_{j,s} h_s \\
\text{subject to} & \\
Cap_j &= Cap_j^{existing} + Cap_j^{add} \quad \forall j \\
\sum_j P_{j,s} &= d_s \quad \forall s \\
0 &\leq P_{j,s} \leq CC_{j,s} Cap_j \quad \forall j, s \\
Cap_j^{add} &\geq 0 \quad \forall j \\
\sum_j CC_{j,1} Cap_j(t) &\geq (1+r)d_1
\end{align*}
\]

Observe that the following changes were made:

- The upper bound $Cap_j$ on the operational variables $P_{j,s}$ in (4a) are multiplied by $CC_{j,s}$ in (4b).
- The capacity credit $CC_j$ in the reserve constraint (6b) is specified to be the capacity credit corresponding to the peak period, $CC_1$, in (6c).

One observes that each load block is likely to differ in terms of which technology is least cost. This can result in solutions providing a combination of technologies in order to achieve capacity benefits from technologies with low investment costs (e.g., combustion turbines) and energy benefits from technologies with low fuel costs (e.g., wind and coal).
1.5 Accounting for annual energy production

The last change made (CC) accounts for the tendency of each technology to operate within certain ranges during certain time periods (where time periods correspond to load blocks).

Another feature which is as yet unaccounted for is the tendency of each technology to produce over a time frame (e.g., a year) a certain fraction of the energy it would produce if it continuously operated at its capacity during that time frame. This feature is typically captured via the capacity factor, denoted here by $\text{CF}_j$.

We account for this tendency in Formulation GEP-6.

Formulation GEP-6:

$$\min \sum_j I_j \text{Cap}_j^{\text{add}} + \sum_j FC_j H_j \sum_s P_{j,s} h_s$$

subject to

$$\text{Cap}_j = \text{Cap}_j^{\text{existing}} + \text{Cap}_j^{\text{add}} \quad \forall j$$

$$\sum_j P_{j,s} = d_s \quad \forall s$$

$$0 \leq P_{j,s} \leq CC_{j,s} \text{Cap}_j \quad \forall j, s$$

$$\text{Cap}_j^{\text{add}} \geq 0 \quad \forall j$$

$$\sum_j CC_{j,s} \text{Cap}_j(t) \geq (1 + r)d_1$$

$$\sum_s P_{j,s} h_s \leq \text{CF}_j \text{Cap}_j \sum_s h_s \quad \forall j$$

The only change made to GEP-6, relative to GEP5, is that (7) has been added to the formulation.
1.6 Representing multiple areas

So far, we have formulated the GEP for only a single area. Doing so is typical when a single company is performing generation expansion, particularly if the company is not transmission limited. However, when solving GEP under transmission limitations, it is important to represent the GEP in terms of multiple areas.

This means we will also be interested to represent transmission. A preliminary step to representing transmission is to first state the multiarea GEP without transmission. To do this, we introduce the subscript \( i \) to denote a particular area, in Formulation GEP-7a.

Formulation GEP-7a:

\[
\begin{align*}
\min & \quad \sum_i \sum_j I_{i,j} Cap_{i,j}^{add} + \sum_i \sum_j FC_{i,j} H_j \sum_{s} P_{i,j,s} h_s \\
\text{subject to} & \\
& \quad Cap_{i,j} = Cap_{i,j}^{existing} + Cap_{i,j}^{add} \quad \forall i, j \\
& \quad \sum_i \sum_j P_{i,j,s} = \sum_i d_{i,s} \quad \forall s \\
& \quad 0 \leq P_{i,j,s} \leq CC_{i,j,s} Cap_{i,j} \quad \forall i, j, s \\
& \quad Cap_{i,j}^{add} \geq 0 \quad \forall i, j \\
& \quad \sum_i \sum_j CC_{i,j,l} Cap_{i,j}(t) \geq (1 + r) \sum_i d_{i,l} \\
& \quad \sum_s P_{i,j,s} h_s \leq CF_{i,j} Cap_{i,j} \sum_s h_s \quad \forall i, j
\end{align*}
\]

Observe that the following changes were made:
The subscript \( i \) was added to all variables and parameters, with the only exceptions being
- \( r \): the reserve requirement, which is considered to be imposed across the entire interconnection;
- \( h_s \): the duration for each load block.

Investment costs and operational costs in (1b) were summed across all areas \( i \).

Equation (2a) is written for each technology \( j \), as in (2), and also for each area \( i \).

Equation (3b) is summed across all areas \( i \) on both the left-hand-side (to obtain total generation for load block \( s \)) and on the right-hand-side (to obtain the total load for load block \( s \)).

Equation (4c) allows the capacity credit for each technology \( j \) and load block \( s \) to also be specified for each area \( i \).

Equation (5a) specifies a \( C_{i,j}^{\text{add}} \) variable for each area \( i \).

Equation (6d) sums across all areas \( i \) on both the left-hand-side (to obtain the total capacity of the interconnection) as well as on the right-hand-side (to obtain the interconnection’s reserve requirement).

Equation (7a) allows the technology capacity factor constraint (to limit total energy over the time period) to be area-specific.

The implication of (3b) and (6d) are that transmission capacity is unlimited, i.e.,
- (3b) says the total interconnection load must be supplied by the total interconnection generation;
- (6d) says the reserve requirement is interconnection-wide (and not area-specific).

Formulation GEP-7a differs from solving GEP-6 for the entire interconnection in that in GEP-7a, parameters are geographically
differentiated, whereas in GEP-6, there is only one set of parameters.

An alternative formulation would assume that there is no transmission between the areas, as in Formulation 7b.

Formulation GEP-7b:
\[
\begin{align*}
\min & \sum_{i} \sum_{j} I_{i,j} \text{Cap}_{i,j}^{\text{add}} + \sum_{i} \sum_{j} F_{i,j} H_{j} \sum_{s} P_{i,j,s} h_{s} \\
\text{subject to} & \quad \text{Investment Costs} \quad \text{Operational Costs}
\end{align*}
\]

subject to
\[
\begin{align*}
\text{Cap}_{i,j} &= \text{Cap}_{i,j}^{\text{existing}} + \text{Cap}_{i,j}^{\text{add}} \quad \forall i, j \\
\sum_{j} P_{i,j,s} &= d_{i,s} \quad \forall i, s \\
0 &\leq P_{i,j,s} \leq C_{i,j,s} \text{Cap}_{i,j} \quad \forall i, j, s \\
\text{Cap}_{i,j}^{\text{add}} &\geq 0 \quad \forall i, j \\
\sum_{j} C_{i,j,1} \text{Cap}_{i,j}(t) &\geq (1 + r_{i})d_{i,1} \quad \forall i \\
\sum_{s} P_{i,j,s} h_{s} &\leq C_{F,i,j} \text{Cap}_{i,j} \sum_{s} h_{s} \quad \forall i, j
\end{align*}
\]

Observe that the following changes were made:
- Equation (3c) is written for each area, thus, each area must satisfy its own demand.
- Equation (6e) is also written for each area, thus, each area must satisfy its own reserve requirement. Notice also that each area’s reserve requirement \( r_{i} \) is specific to area \( i \).

Since there is no transmission between the areas, this problem is equivalent to solving GEP-6 for each of the areas.
1.7 Representing transmission, but without limits

The intermediate formulation between Formulations GEP-7a and GEP-7b is to represent capacitated (i.e., limited) transmission. We will do this in the next subsection, but first, we represent transmission without imposing transmission limits. Here, each area i is considered to be a network node. We extend this from GEP-7b and not GEP-7a in order to take advantage of GEP-7b’s node-specific power balance equations (3c).

Formulation GEP-8:

\[
\min \sum_i \sum_j I_{i,j} \text{Cap}_{i,j}^{\text{add}} + \sum_i \sum_j FC_{i,j} H_j \sum_s P_{i,j,s} h_s
\]

subject to

\[
\text{Cap}_{i,j} = \text{Cap}_{i,j}^{\text{existing}} + \text{Cap}_{i,j}^{\text{add}} \quad \forall i, j \quad (2a)
\]

\[
\sum_j P_{i,j,s} - \sum_k b_l S_{i,l} S_{l,k} \theta_{k,s} = d_{i,s} \quad \forall i, s \quad (3d)
\]

\[
0 \leq P_{i,j,s} \leq CC_{i,j,s} \text{Cap}_{i,j} \quad \forall i, j, s \quad (4c)
\]

\[
\text{Cap}_{i,j}^{\text{add}} \geq 0 \quad \forall i, j \quad (5a)
\]

\[
\sum_i \sum_j CC_{i,j,l} \text{Cap}_{i,j}(t) \geq (1 + r) \sum_i d_{i,l} \quad (6d)
\]

\[
\sum_s P_{i,j,s} h_s \leq CF_{i,j} \text{Cap}_{i,j} \sum_s h_s \quad \forall i, j \quad (7a)
\]

This formulation is no different Formulation GEP-7a in terms of solution, except that it is more computational (due to its inclusion of the linearized power flow equations) and that it provides flows between areas for each of the load blocks.
The only difference between GEP-8 and GEP-7a is the presence of (3d) instead of (3b). The nomenclature here needs definition, as follows:

- $b_l$ is the negative of the susceptance of branch $l$ (therefore +)
- $S_{l,i}$ is element $(l,i)$ of the node-arc incidence matrix.
- $S_{l,k}$ is element $(l,k)$ of the node-arc incidence matrix.
- $\theta_{k,s}$ is the angle of node (or area) $k$ under loading block $s$.

A word about the node-arc incidence matrix is in order here. It is also called the adjacency matrix, or the connection matrix. This matrix is well known in any discipline that has reason to structure its problems using a network of nodes and “arcs” (or branches or edges). Any type of transportation engineering is typical of such a discipline.

The node-arc incidence matrix contains a number of rows equal to the number of arcs and a number of columns equal to the number of nodes.

Element $(l,i)$ of $A$ is 1 if the $l$th branch begins at node $i$, -1 if the $l$th branch terminates at node $i$, and 0 otherwise.

A branch is said to “begin” at node $i$ if the power flowing across branch $l$ is defined positive for a direction from node $i$ to the other node.

A branch is said to “terminate” at node $i$ if the power flowing across branch $l$ is defined positive for a direction to node $l$ from the other node. Figure 5 is an illustration of a node-arc incidence matrix for a particular network.
The admittance for all branches is \( y = g + jb = -j10 \), so that the susceptance is \( b = -10 \) (implying it is reactive, not capacitive).

\[
A = \begin{bmatrix}
1 & 0 & 0 & -1 \\
1 & -1 & 0 & 0 \\
0 & 1 & -1 & 0 \\
0 & 0 & -1 & 1 \\
1 & 0 & -1 & 0
\end{bmatrix}
\]

Reconsider (3d):
\[
\sum_{i,j,s} P_{i,j,s} - \sum_{l,k} b_l S_{i,l} S_{l,k} \theta_{k,s} = d_{i,s} \quad \forall i, s
\]  
(3d)

In this equation:

- \( b_l \) as the negative susceptance is a positive number;
For each branch l, there will be one product term $S_{l,i} S_{l,k}$ equal to 1 (when $k=i$) and one product term $S_{l,i} S_{l,k}$ equal to -1 (when $k$ is a terminating node for a branch connected to node i); all other product terms $S_{l,i} S_{l,k}$ will be zero.

**ASIDE:** We illustrate the above relation. Dropping the subscripts $s$ (for LDC blocks) and $j$ (for technologies), we have

$$ P_i - \sum_l \sum_k b_l S_{l,i} S_{l,k} \theta_k = d_i $$

We write the expression using the above system for node 1:

$$ P_1 - \sum_l \sum_k b_l S_{l,1} S_{l,k} \theta_k = d_1 $$

We observe that there are 5 branches, so that $l=1,\ldots,5$, and there are 4 nodes, so that $k=1,\ldots,4$. Therefore:

$$ P_1 - \sum_{l=1}^{5} \sum_{k=1}^{4} b_l S_{l,1} S_{l,k} \theta_k = d_1 $$

Expanding the inner summations results in

$$ P_1 - \sum_{l=1}^{5} b_l S_{l,1} (S_{l,1} \theta_1 + S_{l,2} \theta_2 + S_{l,3} \theta_3 + S_{l,4} \theta_4) = d_1 $$

Now expand the outer summation to obtain:

$$ P_1 - b_1 S_{1,1} (S_{1,1} \theta_1 + S_{1,2} \theta_2 + S_{1,3} \theta_3 + S_{1,4} \theta_4) $$

$$ - b_2 S_{2,1} (S_{2,1} \theta_1 + S_{2,2} \theta_2 + S_{2,3} \theta_3 + S_{2,4} \theta_4) $$

$$ - b_3 S_{3,1} (S_{3,1} \theta_1 + S_{3,2} \theta_2 + S_{3,3} \theta_3 + S_{3,4} \theta_4) $$

$$ - b_4 S_{4,1} (S_{4,1} \theta_1 + S_{4,2} \theta_2 + S_{4,3} \theta_3 + S_{4,4} \theta_4) $$

$$ - b_5 S_{5,1} (S_{5,1} \theta_1 + S_{5,2} \theta_2 + S_{5,3} \theta_3 + S_{5,4} \theta_4) $$

$$ = d_1 $$

Recalling the node-arc incidence matrix below, we may fill in the values for $S_{l,k}$.
\[ A = \begin{bmatrix} 1 & 0 & 0 & -1 \\ 1 & -1 & 0 & 0 \\ 0 & 1 & -1 & 0 \\ 0 & 0 & -1 & 1 \\ 1 & 0 & -1 & 0 \end{bmatrix} \]

\[
P_1 - b_1(\theta_1 - \theta_4) \\
- b_2(\theta_1 - \theta_2) \\
- 0 \\
- 0 \\
- b_5(\theta_1 - \theta_3) \\
= d_1
\]

This may be re-written as
\[
P_1 - d_1 = b_1(\theta_1 - \theta_4) + b_2(\theta_1 - \theta_2) + b_5(\theta_1 - \theta_3)
\]
from which it is easy to see the correspondence to the one-line diagram, repeated below for convenience.
1.8 Representing transmission, with limits

Formulation GEP-9:
\[
\min \sum_{i} \sum_{j} I_{i,j} \text{Cap}_{i,j}^{\text{add}} + \sum_{i} \sum_{j} \text{FC}_{i,j} \text{H}_{j} \sum_{s} P_{i,j,s} h_{s} \quad \text{(1b)}
\]

subject to
\[
\text{Cap}_{i,j} = \text{Cap}_{i,j}^{\text{existing}} + \text{Cap}_{i,j}^{\text{add}} \quad \forall i, j \quad \text{(2a)}
\]
\[
\sum_{j} P_{i,j,s} - \sum_{k} b_{l} S_{i,l} S_{l,k} \theta_{k,s} = d_{i,s} \quad \forall i, s \quad \text{(3d)}
\]
\[
0 \leq P_{i,j,s} \leq \text{CC}_{i,j,s} \text{Cap}_{i,j} \quad \forall i, j, s \quad \text{(4c)}
\]
\[
\text{Cap}_{i,j}^{\text{add}} \geq 0 \quad \forall i, j \quad \text{(5a)}
\]

Need to state here that all quantities must be in per-unit.
\[ \sum_{i} \sum_{j} CC_{i,j} Cap_{i,j}(t) \geq (1 + r) \sum_{i} d_{i,1} \]  
(6d)

\[ \sum_{s} P_{i,j,s} h_{s} \leq CF_{i,j} Cap_{i,j} \sum_{s} h_{s} \quad \forall i, j \]  
(7a)

\[ b_{l} \left| \sum_{i} S_{i,l} \theta_{i,s} \right| \leq F_{l}^{\text{max}} \quad \forall l, s \]  
(8)

Formulation GEP-9 is exactly like Formulation GEP-8 except for the addition of (8) which limits the flows on the branches.

### 1.9 Including adequacy constraints

Adequacy evaluation is the evaluation of the extent to which the available generation and transmission is sufficient to supply the load considering steady-state limitations on components (generators and circuits) and their potential for failure. The multi-area reliability analysis we recently learned (using max-flow and decomposition) was adequacy evaluation.

We augment Formulation GEP-9 with a constraint on an adequacy index. For example, we may constrain loss-of-load probability (LOLP) according to the following:

Formulation GEP-10:

\[
\begin{align*}
\min & \quad \sum_{i} \sum_{j} I_{i,j} Cap_{i,j}^{\text{add}} + \sum_{i} \sum_{j} FC_{i,j} H_{j} \sum_{s} P_{i,j,s} h_{s} \\
\text{subject to} & \\
Cap_{i,j} &= Cap_{i,j}^{\text{existing}} + Cap_{i,j}^{\text{add}} \quad \forall i, j \\
\sum_{j} P_{i,j,s} - \sum_{l} \sum_{k} b_{l} S_{i,l} S_{l,k} \theta_{k,s} &= d_{i,s} \quad \forall i, s \\
0 \leq P_{i,j,s} &\leq CC_{i,j,s} Cap_{i,j} \quad \forall i, j, s
\end{align*}
\]  
(1b) 
(2a) 
(3d) 
(4c)
\[ Cap_{i,j}^{add} \geq 0 \quad \forall i, j \]  
(5a)

\[ \sum_i \sum_j CC_{i,j,l} Cap_{i,j}(t) \geq (1 + r) \sum_i d_{i,l} \]  
(6d)

\[ \sum_s P_{i,j,s} h_s \leq CF_{i,j} Cap_{i,j} \sum_s h_s \quad \forall i, j \]  
(7a)

\[ b_l \left| \sum_i S_{l,i} \theta_{i,s} \right| \leq F_{l}^{\text{max}} \quad \forall l, s \]  
(8)

\[ \text{LOLP} \leq \text{LOLP}_0 \]  
(9)

where LOLP₀ is a chosen maximum level of LOLP.

The problem with Formulation 10 is that computation of LOLP is intensive, as we have observed in our work with multi-area reliability analysis. There has been work to include adequacy constraints within the GEP problem [2, 3], and it is of interest to study the various approaches to determine the additional computation necessary. One approach that is perhaps most straightforward is to exclude the adequacy constraint within the optimization problem and instead evaluate the adequacy of the system after obtaining a solution to the optimization problem; if the adequacy level is not acceptable, then adjustments are made (e.g., by increasing reserve requirements), and the optimization is performed again. Such an approach, illustrated in Figure 6, is amendable to optimization methods which employ decomposition (not the same as the decomposition we learned for multiarea reliability analysis). One well-known method is called Benders decomposition; we will study it later.
1.10 Multiperiod formulation

The problem we have described up until this point has retained the single-period simplification. However, a realistic version of the GEP is necessarily multi-period, to account for the variation in time when new facilities come on-line, the variation in time when existing facilities are retired, demand growth, and the need to meet demand and reserve requirements but without overbuilding.

A key issue for multi-period planning is what is referred to as end effects. End effects refer to the difficulty of appropriately representing the influence of investment costs and operational costs at the end of the planning period. There are two problems to address:

1. Remnant investment value: The retirement year for some facilities occurs after the final year. For these facilities, there
is a value to the facility because it has remaining life, i.e., some of the investment amount paid has not yet been depreciated.

2. Remnant operational cost: Because the simulation must be truncated at a particular final year, operational costs after that final year are not included since those years are not simulated. Both of the above effects tend to bias decisions in favor of alternatives that have low investment costs, since ignoring the remnant investment value and the remnant operational cost means the optimization sees only the investment cost, and because it is minimizing costs, it chooses the investments with the lowest investment costs.

These issues are addressed by doing two things:

1. **Extend the simulation time beyond the planning horizon:** A multiperiod formulation with extended planning horizon is based on the observation that end effects increase their influence as the final year gets closer. For example, if the planning horizon is 20 years, end effects have more influence in years 5-20 than they do in years 1-5. This observation leads to a very natural solution: extend the final year well-beyond the planning horizon. Therefore if the planning horizon is 20 years, we may run the optimization problem to 50 years; but the decisions for years 20-50 will be ignored. A general guide for how far to extend the simulation beyond the planning horizon is that the final year of the simulation should exceed the final year of the planning horizon by the lifetime of the facility with the longest life.

2. **Model salvage values to facilities:** Salvage value is the net sum to be realized from the disposal of an asset (net of
disposal costs) at the time of its replacement or resale, or at
the end of the study period [6, pg. 134].

A revised formulation of the GEP is provided in Formulation
GEP-11, which accounts for extended horizon and modeling of
salvage values. There are 5 salient changes to our formulation:

- Inclusion of the discount factor $\zeta$ in the objective function.
- Provision of subscript $t$ - this subscript $t$ goes from $t=1,\ldots,T$,
  where $T$ is chosen to reflect the extended time; it affects all
  equations in the formulation.
- Inclusion of the salvage value in objective function.
- Inclusion of retirement values in the update equation (2b-i).
- Addition of (2b-ii), (9), (10), and (11) and introduction of
  parameters Life and AvgAge.

Formulation GEP-11:

$$
\min \sum_{t} \zeta^{t-1} \left\{ \sum_{i} \sum_{j} I_{i,j,t} \text{Cap}_{i,j,t}^{\text{add}} \right\} - \zeta^{T-1} \left\{ \sum_{i} \sum_{j} SV_{i,j,T} \text{Cap}_{i,j,T} \right\} \\
+ \sum_{t} \zeta^{t-1} \left\{ \sum_{i} \sum_{j} FC_{i,j} H_{j} \sum_{s} P_{i,j,s} h_{s} \right\}
$$

subject to
\[
\begin{align*}
\text{Cap}_{i,j,t} &= \text{Cap}_{i,j,t-1} + \text{Cap}_{i,j,t}^{\text{add}} - \text{Cap}_{i,j,t}^{\text{ret}} \quad \forall i, j, t \\
\text{Cap}_{i,j,0} &= \text{Cap}_{i,j}^{\text{exist}} \quad \forall i, j \\
\text{Cap}_{i,j,t}^{\text{ret}} &= \text{Cap}_{i,j,t}^{\text{ret,exist}} + \text{Cap}_{t-Life_j}^{\text{Add}} \\
\sum_j P_{i,j,s,t} - \sum_l \sum_k b_l S_{l,i} S_{l,k} \theta_{k,s,t} &= d_{i,s,t} \quad \forall i, s, t \\
0 \leq P_{i,j,s,t} &\leq CC_{i,j,s} \text{Cap}_{i,j,t} \quad \forall i, j, s, t \\
\text{Cap}_{i,j,t}^{add} &\geq 0 \quad \forall i, j, t \\
\sum_i \sum_j CC_{i,j,1} \text{Cap}_{i,j,t} (t) &\geq (1 + r) \sum_i d_{i,1,t} \quad \forall t \\
\sum_s P_{i,j,s,t} h_s &\leq CF_{i,j} \text{Cap}_{i,j,t} \sum_s h_s \quad \forall i, j, t \\
b_l \left| \sum_i S_{l,i} \theta_{i,s,t} \right| &\leq F_l^{\text{max}} \quad \forall l, s, t \\
\text{AvgAge}_{i,j,t} &= \frac{\text{AvgAge}_{i,j,t-1} + \text{Life}_j \times \text{Cap}_{i,j,t}^{\text{ret}} + 1 \times \text{Cap}_{i,j,t}^{\text{add}}}{\text{Cap}_{i,j,t-1} - \text{Cap}_{i,j,t}^{\text{ret}} + \text{Cap}_{i,j,t}^{\text{add}}} \quad \forall i, j, t \\
\text{SV}_{i,j,t} &= \frac{\text{Life}_j - \text{AvgAge}_{i,j,t}}{\text{Life}_j} \times I_{i,j,t} \quad \forall i, j, t
\end{align*}
\]

We describe the new and significantly modified equations in what follows:
- Equation (1c): There are three changes to describe within the objective function:
  - The discount factor is given by \( \zeta = 1/(1+i) \) where \( i \) is the discount rate. Thus we have that
  \[
  \zeta^t = \frac{1}{(1+i)^t}
  \]
We assume the investments made in year 1 are already present value, and so it is not until year 2 that we need to discount to present worth; therefore we utilize $\zeta^{t-1}$ as the discount factor.

- The investment cost $I_{i,j,t}$, in $$/MW, is subscripted by time $t$ to reflect the fact that it can change over time due to inflation and/or technology maturation.
- The salvage value is subtracted from the investment costs to accommodate the remaining value of the capacity at the simulation end, $T$. The salvage value $SV_{i,j,t}$ is in units of $$/MWhr and when multiplied by $Cap_{i,j,t}$, gives $$. 

**Equation (10):** This equation computes $SV_{i,j,t}$, the salvage value of technology $j$ in region $i$ at time $t$ as a fraction of the investment cost $I_{i,j,t}$ of technology $j$ in region $i$ at time $t$. There are two issues to consider here:

- Investment cost: If there is no inflation or technology maturation, then the investment cost is constant throughout the entire simulation time $t=1,\ldots,T$. But if investment cost varies, we utilize the investment cost at time $t$ because it better reflects the buying power of money at time $t$ which better reflects the salvage value.
- The fraction utilized to compute $SV_{i,j,t}$ is the ratio of the average remaining life for technology $j$ in region $i$ at time $t$ to the expected life of technology $j$. If there is no remaining life ($AvgLife=Life$), then salvage value is 0.

It is appropriate at this point to make the following observation in regards to all GEP modeling done within these notes: we are modeling aggregated capacity, and not capacity of individual units. This means that we may make capacity investments (and retirements) of any magnitude. In reality, capacity is built (and
retired) in increments of discrete values. We make this general point at this time in order to point out that our valuation of salvage value is also based on aggregated capacity and not discrete units. This point also justifies our estimation of the average life, AvgLife, a key parameter used to compute the salvage value.

• Equation (9): This equation, repeated here for convenience, computes average age of technology \( j \) in region \( i \) and time \( t \).

\[
\text{AvgAge}_{i,j,t} = \frac{\left(\text{AvgAge}_{i,j,t-1} + 1\right) \times \text{Cap}_{i,j,t-1} - \text{Life}_j \times \text{Cap}^{\text{ret}}_{i,j,t} + 1 \times \text{Cap}^{\text{add}}_{i,j,t}}{\text{Cap}_{i,j,t-1} - \text{Cap}^{\text{ret}}_{i,j,t} + \text{Cap}^{\text{add}}_{i,j,t}} \quad \forall i, j, t
\]

This calculation provides a \textit{weighted} average age, where the weights are the capacities. The equation indicates the following: the average age at time \( t \) depends on

- the \([\text{average age at time } t-1] + 1 \text{ yr}] \times \text{capacity at time } t-1$
- less the influence of retiring units at time \( t \) (which all have an age of \( \text{Life}_j \))
- plus the influence of added units at time \( t \) (which are assumed to have an age of 1 year)

• Equations (2b-i, 2b-ii, and 2b-iii) are described as follows:
  - Equation (2b-i), repeated below for convenience, is the capacity update equation and has been modified to account for retirements via the term \( \text{Cap}^{\text{ret}}_{i,j,t} \).

\[
\text{Cap}_{i,j,t} = \text{Cap}_{i,j,t-1} + \text{Cap}^{\text{add}}_{i,j,t} - \text{Cap}^{\text{ret}}_{i,j,t} \quad \forall i, j, t
\]  

(2b-i)
  - Equation (2b-ii), repeated below for convenience, initiates capacities at time \( t=0 \) to the capacities that existed at the beginning of the simulation, \( \text{Cap}^{\text{exist}}_{i,j} \).

\[
\text{Cap}_{i,j,0} = \text{Cap}^{\text{exist}}_{i,j} \quad \forall i, j
\]

(2b-ii)
  - Equation (2b-iii), repeated below for convenience, computes retirements at time \( t \) as the retirements at time \( t \) of units that existed at the beginning of the simulation
(assumed to be input data) plus retirements from capacity that has been added since the beginning of the simulation but has exceeded its life, which is $Cap_{i,j,t}^{Add}$. 

$$Cap_{i,j,t}^{ret} = Cap_{i,j,t}^{ret, exist} + Cap_{i,j,t}^{Add} - Life_j$$

(2b-iii)

We emphasize that the only decision variables in our problem formulation are the capacity-add variables, i.e., $Cap_{i,j,t}^{Add}$ and the operational variables $P_{i,j,s,t}$. The retirement variables $Cap_{i,j,t}^{ret}$ are not decision variables.

Direct use of the above model results in a nonlinearity. How to express the GEP-11 formulation (with retirements) as an LP? Formulation GEP-11 has the following term in the objective function.

$$-\zeta^T \left\{ \sum_i \sum_j SV_{i,j,t} Cap_{i,j,t} \right\}$$

where

$$Cap_{i,j,t} = Cap_{i,j,t-1}^{add} + Cap_{i,j,t}^{add} - Cap_{i,j,t}^{ret} \quad \forall i, j, t$$

(2b-i)

$$\text{AvgAge}_{i,j,t} = \frac{\text{AvgAge}_{i,j,t-1} + 1 \times Cap_{i,j,t-1}^{ret} \times Cap_{i,j,t}^{add} + Cap_{i,j,t}^{add}}{Cap_{i,j,t-1}^{ret} + Cap_{i,j,t}^{add}} \quad \forall i, j, t$$

(9)

$$SV_{i,j,t} = \frac{Life_j - \text{AvgAge}_{i,j,t} \times I_{i,j,t}}{Life_j} \quad \forall i, j, t$$

(10)

One brief note: The term $SV_{i,j,t}$ depends on investment cost ($$/\text{MW}) at time $t$ – the rationale here is that the investment cost at time $t$ better reflects the buying power of money at time $t$, which better reflects the salvage value (if there is no inflation or technology maturation, then this number is constant throughout
the planning horizon). The basic term in the objective function for the salvage value for a given technology is:

\[ SV_{i,j,T} \cdot \text{Cap}_{i,j,T} = \left( \frac{L_i - \text{AvgAge}_{i,j,T}}{L_i} \right) \times I_{i,j,T} \times \text{Cap}_{i,j,T} \]

\[ = \left( 1 - \frac{\text{AvgAge}_{i,j,T}}{L_i} \right) \times I_{i,j,T} \times \text{Cap}_{i,j,T} \]

\[ = \left( \text{Cap}_{i,j,T} - \frac{\text{AvgAge}_{i,j,T} \times \text{Cap}_{i,j,T}}{L_i} \right) \times I_{i,j,T} \]

(\#)

Now the term \( \text{AvgAge}_{i,j,T} \times \text{Cap}_{i,j,T} \) needs to be simplified using the expression for \( \text{vgAge}_{i,j,T} \). It can be seen from (2b-i) and (9) that the denominator for \( \text{AvgAge}_{i,j,T} \) is \( \text{Cap}_{i,j,T} \). That is,

\[ \text{AvgAge}_{i,j,T} = \frac{(\text{AvgAge}_{i,j,T-1} + 1) \times \text{Cap}_{i,j,T-1} - L_i \times \text{Cap}_{i,j,T-1}^{\text{ret}} + 1 \times \text{Cap}_{i,j,T}^{\text{add}}}{\text{Cap}_{i,j,T-1} - \text{Cap}_{i,j,T}^{\text{ret}} + \text{Cap}_{i,j,T}^{\text{add}}} \]

(\*)

And from (2b-i) we see that

\[ \text{Cap}_{i,j,T} = \text{Cap}_{i,j,T-1} + \text{Cap}_{i,j,T}^{\text{add}} - \text{Cap}_{i,j,T}^{\text{ret}} \]

(**)

Substituting (**) into (*), we obtain

\[ \text{AvgAge}_{i,j,T} = \frac{(\text{AvgAge}_{i,j,T-1} + 1) \times \text{Cap}_{i,j,T-1} - L_i \times \text{Cap}_{i,j,T-1}^{\text{ret}} + 1 \times \text{Cap}_{i,j,T}^{\text{add}}}{\text{Cap}_{i,j,T}} \]

Thus,

\[ \text{AvgAge}_{i,j,T} \times \text{Cap}_{i,j,T} = (\text{AvgAge}_{i,j,T-1} + 1) \times \text{Cap}_{i,j,T-1} - L_i \times \text{Cap}_{i,j,T-1}^{\text{ret}} + 1 \times \text{Cap}_{i,j,T}^{\text{add}} \]

Distributing the term \( \text{Cap}_{i,j,T} \) on the right-hand-side,

\[ \text{AvgAge}_{i,j,T} \times \text{Cap}_{i,j,T} = \text{AvgAge}_{i,j,T-1} \times \text{Cap}_{i,j,T-1} + 1 \times \text{Cap}_{i,j,T-1} \]

\[ - L_i \times \text{Cap}_{i,j,T-1}^{\text{ret}} + 1 \times \text{Cap}_{i,j,T}^{\text{add}} \]

The right hand side of the above equation contains \( \text{AvgAge}_{i,j,T-1} \times \text{Cap}_{i,j,T-1} \), and this is rewritten as
\(\text{AvgAge}_{i,j,T-1} \times \text{Cap}_{i,j,T-1}\)

\(= \text{AvgAge}_{i,j,T-2} \times \text{Cap}_{i,j,T-2} + 1 \cdot \text{Cap}_{i,j,T-2} - \text{Life}_j \cdot \text{Cap}_{i,j,T-1}^{\text{ret}} + 1 \cdot \text{Cap}_{i,j,T-1}^{\text{add}}\)

Continuing like this for all time periods, we get T equations with the last equation given below.

\(\text{AvgAge}_{i,j,1} \times \text{Cap}_{i,j,1}\)

\(= \text{AvgAge}_{i,j,0} \times \text{Cap}_{i,j,0} + 1 \cdot \text{Cap}_{i,j,0} - \text{Life}_j \cdot \text{Cap}_{i,j,1}^{\text{ret}} + 1 \cdot \text{Cap}_{i,j,1}^{\text{add}}\)

Adding all the equations, we observe that the terms \(\text{AvgAge}_{i,j,T-1} \times \text{Cap}_{i,j,T-1}\), \(\text{AvgAge}_{i,j,T-2} \times \text{Cap}_{i,j,T-2}\), etc get cancelled on both sides and the remaining terms give the following equation:

\(\text{AvgAge}_{i,j,T} \times \text{Cap}_{i,j,T} = \text{AvgAge}_{i,j,0} \times \text{Cap}_{i,j,0} + \sum_{t=1}^{T} (1 \cdot \text{Cap}_{i,j,t-1} - \text{Life}_j \cdot \text{Cap}_{i,j,t}^{\text{ret}} + 1 \cdot \text{Cap}_{i,j,t}^{\text{add}})\)

\((***)\)

This is a linear expression in all the variables.
Recall eq. (\#) which reflects what we need in the objective function:

\(\text{SV}_{i,j,T} \cdot \text{Cap}_{i,j,T} = \left(\text{Cap}_{i,j,T} - \frac{\text{AvgAge}_{i,j,T} \times \text{Cap}_{i,j,T}}{\text{Life}_j}\right) \times I_{i,j,T}\)

\((\#)\)

Substitute \((***)\) into \((\#)\) to obtain:

\(\zeta^{T-1} \left(\text{Cap}_{i,j,T} - \frac{\text{AvgAge}_{i,j,0} \times \text{Cap}_{i,j,0} + \sum_{t=1}^{T} (1 \cdot \text{Cap}_{i,j,t-1} - \text{Life}_j \cdot \text{Cap}_{i,j,t}^{\text{ret}} + 1 \cdot \text{Cap}_{i,j,t}^{\text{add}})}{\text{Life}_j}\right) \times I_{i,j,T}\)

Thus the constraints (9) & (10) can be removed if we express the objective function like this.
\[
\min \sum_{t} \zeta_{t-1} \left\{ \sum_{i} \sum_{j} I_{i,j,t} \text{Cap}_{i,j,t}^{\text{add}} \right\} \\
- \zeta_{T-1} \left\{ \sum_{i} \sum_{j} \text{SV}_{i,j,t} \text{Cap}_{i,j,T} \right\} \\
+ \sum_{t} \zeta_{t-1} \left\{ \sum_{i} \sum_{j} \text{FC}_{i,j} H_{j} \sum_{s} P_{i,j,s} h_{s} \right\}
\]

\[
= \min \sum_{t} \zeta_{t-1} \left\{ \sum_{i} \sum_{j} I_{i,j,t} \text{Cap}_{i,j,t}^{\text{add}} \right\} \\
- \zeta_{T-1} \left\{ \sum_{i} \sum_{j} \text{Cap}_{i,j,T} - \frac{\text{AvgAge}_{i,j,0} \times \text{Cap}_{i,j,0} + \sum_{t=1}^{T} (\text{Cap}_{i,j,t-1} - \text{Life}_{j} \times \text{Cap}_{i,j,t}^{\text{ret}} + \text{Cap}_{i,j,t}^{\text{add}})}{\text{Life}_{j}} \right\} \\
+ \sum_{t} \zeta_{t-1} \left\{ \sum_{i} \sum_{j} \text{FC}_{i,j} H_{j} \sum_{s} P_{i,j,s} h_{s} \right\}
\]

General comments about multi-period planning models follow. It is possible to find publications that relate to each of the below.

- It is possible to model discrete capacities, and there are two ways:
  - Impose constraints that require capacity variables to lie between tight bounds, e.g.,
    \[
    \text{Cap}_{\text{Low},j} \leq \text{Cap}_{i,j,t}^{\text{Add}} \leq \text{Cap}_{\text{High},j} \quad \forall j
    \]
- Convert the problem to a mixed integer program by defining decision variables to be integers $c_{i,j,t}$ which are deployed as multiplying coefficients of $Cap_{i,j,t}^{add}$ in the above formulations. Other changes may be needed as well.

- Fixed and variable O&M: In what we have done so far, the only costs have been investment costs and fuel costs. However, generation also has fixed and variable O&M costs. In any of our formulations, we may include these costs in the objective function, representing the fixed O&M proportion to the capacity and the variable O&M proportional to the energy produced. We do so to single-period Formulations GEP-9 to produce Formulation GEP-9-OM.

Formulation GEP-9-OM:

$$\min \sum{\sum{I_{i,j}Cap_{i,j}^{add}}} + \sum{\sum{OM_{j}^{fixed} Cap_{i,j}}}$$

subject to

$$Cap_{i,j} = Cap_{i,j}^{existing} + Cap_{i,j}^{add} \quad \forall i, j \quad (2a)$$

$$\sum{P_{i,j,s}} - \sum{\sum{l}{b_lS_{l,i}S_{l,k}\theta_{k,s}}} = d_{i,s} \quad \forall i, s \quad (3d)$$

$$0 \leq P_{i,j,s} \leq CC_{i,j,s} Cap_{i,j} \quad \forall i, j, s \quad (4c)$$

$$Cap_{i,j}^{add} \geq 0 \quad \forall i, j \quad (5a)$$

$$\sum{\sum{l}{CC_{i,j,l}Cap_{i,j}(t)}} \geq (1 + r)\sum{d_{i,l}} \quad (6d)$$

$$\sum{s}{P_{i,j,s}h_s} \leq CF_{i,j} Cap_{i,j} \sum{s}{h_s} \quad \forall i, j \quad (7a)$$
\[ b_l \left| \sum_i S_{l,i} \theta_{i,s} \right| \leq F_{l}^{\max} \quad \forall l, s \]  

(8)

We also do so to multi-period Formulations GEP-11 to produce Formulation GEP-11-OM:

**Formulation GEP-11-OM:**

\[
\begin{align*}
\min \sum_t \xi^{-1} \left( \sum_i \sum_j I_{i,j,t} \text{Cap}_{i,j,t}^{add} - \xi^{T-1} \left( \sum_i \sum_j SV_{i,j,T} \text{Cap}_{i,j,T} \right) + \sum_i \sum_j \text{OM}_{j}^{fixed} \text{Cap}_{i,j,t} \right) \\
+ \sum_t \xi^{-1} \left( \sum_i \sum_j FC_{i,j} H_j \sum_s P_{i,j,s} h_s \right) + \sum_t \xi^{-1} \sum_i \sum_j \text{OM}_{j}^{var} \sum_s P_{i,j,s,t} h_s \\
\end{align*}
\]

subject to

\[ Cap_{i,j,t} = Cap_{i,j,t-1} + Cap_{i,j,t}^{add} - Cap_{i,j,t}^{ret} \quad \forall i, j, t \]  

(2b-i)

\[ Cap_{i,j,0} = Cap_{i,j}^{exist} \quad \forall i, j \]  

(2b-ii)

\[ Cap_{i,j,t}^{ret} = Cap_{i,j,t}^{ret}^{exist} + Cap_{i}^{Add} \]  

(2b-iii)

\[ \sum_j P_{i,j,s,t} - \sum_k b_l S_{l,i} S_{l,k} \theta_{k,s,t} = d_{i,s,t} \quad \forall i, s, t \]  

(3e)

\[ 0 \leq P_{i,j,s,t} \leq CC_{i,j,s} Cap_{i,j,t} \quad \forall i, j, s, t \]  

(4c)

\[ Cap_{i,j,t}^{add} \geq 0 \quad \forall i, j, t \]  

(5a)

\[ \sum_j CC_{i,j,1} Cap_{i,j,t} (t) \geq (1+r) \sum_i d_{i,1,t} \quad \forall t \]  

(6d)

\[ \sum_s P_{i,j,s,t} h_s \leq CF_{i,j} Cap_{i,j,t} \sum h_s \quad \forall i, j, t \]  

(7a)

\[ b_l \left| \sum_i S_{l,i} \theta_{i,s,t} \right| \leq F_{l}^{\max} \quad \forall l, s, t \]  

(8)
\[ \text{AvgAge}_{i,j,t} = \frac{(\text{AvgAge}_{i,j,t-1} + 1) \times \text{Cap}_{i,j,t-1} - \text{Life}_j \times \text{Cap}_{i,j,t}^{\text{ret}} + 1 \times \text{Cap}_{i,j,t}^{\text{add}}}{\text{Cap}_{i,j,t-1} - \text{Cap}_{i,j,t}^{\text{ret}} + \text{Cap}_{i,j,t}^{\text{add}}} \quad \forall i, j, t \quad (9) \]

\[ \text{SV}_{i,j,t} = \frac{\text{Life}_j - \text{AvgAge}_{i,j,t}}{\text{Life}_j} \times I_{i,j,t} \quad \forall i, j, t \quad (10) \]

- It is possible to provide that the retirements become decisions within the optimization. But to do this, there must be potential for retirement to be more economic than plant retention. When the Clean-Air Act Amendments were made in 1990, this problem was a very good problem to solve because generation owners had to decide, for coal plants, whether to retire them, convert to a new fuel, install scrubbers, or purchase allowances. It is also a good problem to solve in today’s environment because of the recent EPA actions regarding (a) mercury and CO₂. Owners of coal plants must decide whether to retrofit existing units with
- There have been two other ways suggested to handle end-effects, as indicated in [4, 5]. One is called the primal equilibrium method and the other is called the dual equilibrium method. The basic idea of the primal equilibrium method is that
  - The costs of the objective function are written as infinite series in time, from \( t=1, \ldots, \infty \);
  - A relation is imposed on all decision variables \( x \), which requires that, beyond a certain time \( T \), the value in one time period relates to the value in another time period, according to
    \[ x(T+t) = \lambda \times x(T+t-1) \]
    where \( \lambda \) is the annual demand growth. Doing so means the infinite series beyond \( T \) can be approximated with a single term, so that the objective function can be modeled as
before, for \( t=1,\ldots,T \) and just a single term for the infinite horizon after that.

The basic idea of the dual equilibrium is similar to that of the primal equilibrium, except the equilibrium condition is imposed on the dual variables (prices) instead of the primal variables (capacity additions).

A comparative study was performed on a 5-node representation of the US generation portfolio to observe the influence of end effects using different methods, including (a) truncation to 40 years; (b) extended to 120 years; (c) include salvage values; (d) dual equilibrium. Results are illustrated below.

![Figure 7](image)

Figure 7: (a) Truncated; (b) Extended; (c) Salvage value; (d) Dual equilibrium

We take the extended 120 year simulation (Plot (b)) to be the basis of comparison. Comparing Plot (b) to Plot (a), it is clear influence of end effects is seen in the results of the truncated model (Plot (a)), where there are no investments in solar units which have high investment. Both salvage value and dual equilibrium methods reduce the end effects. While the dual equilibrium method seems to do well during the end periods, it has impacted the investments trend in the early periods, which the salvage value method has preserved.
Some commercially available expansion planning tools include EGEAS, GEM, STRATEGIST, and PLEXOS.

The New Zealand Electricity Commission developed GEM as a long range generation capacity planning model. It is formulated as a mixed integer programming problem (MIP). The computer code is written using the GAMS optimization software and the model is solved with CPLEX.

The EPRI product EGEAS is a modular production cost and generation expansion software package which employs dynamic programming algorithm to form candidate portfolios from identified alternatives meeting a capacity planning constraint. It also has modules which accommodate demand-side management options and facilitate development of environmental compliance plans. Some of the key functions of EGEAS are asset retirement evaluation, emission evaluation from new plants, and Scenario analysis for various generation options. EGEAS is widely used by many utilities and regulators.

The Ventyx product STRATEGIST is composed of multiple application modules incorporating all aspects of utility planning and operations. This includes forecasted load modeling, marketing and conservation programs,
production cost calculations including the dispatch of energy resources, benefit-cost (B/C) ratios calculation for different alternatives, capital project modeling, financial and rate impacts evaluation, and analysis of long-range rate strategy and the implications of utility plans on customer classes.

PLEXOS, from Energy Exemplar, is a versatile software system that finds optimal combinations of new generation units, unit retirements and transmission system upgrades on a least-cost basis over a long-term planning horizon. PLEXOS in itself does not incorporate the optimization engine but rather produces optimization code that can be read by an external solver such as CPLEX or MOSEK.

We provide an overview of EGEAS, adapted from [6].

The EGEAS computer model was developed by researchers at MIT under funding from the Electric Power Research Institute (EPRI). EGEAS can be run in both the expansion optimization and the production simulation modes. Uncertainty analysis, based on automatic sensitivity analysis and data collapsing via description of function estimation, is also available. A complete description of the model can be found in [7].
The production simulation option consists of production cost/reliability evaluation for a specified generating system configuration during one or more years. Probabilistic production cost/reliability simulation is performed using a load duration curve based model. Customer load and generating unit availability are modeled as random variables to reflect demand fluctuations and generation forced outages. Two algorithmic implementations are available: an analytic representation of the load duration curve (cumulants) and a piecewise linear numerical representation.

The below illustrates the MISO planning process. Step 1 is based on use of EGEAS.
Appendix – Load Duration Curves

A critical issue for planning is to identify the total load level for which to plan. One extremely useful tool for doing this is the so-called load duration curve, which is formed as follows. Consider that we have obtained, either through historical data or through forecasting, a plot of the load vs. time for a period T, as shown in Fig. A1 below.

Fig. A1: Load curve (load vs. time)

Of course, the data characterizing Fig. A1 will be discrete, as illustrated in Fig. A2.

Fig. A2: Discretized Load Curve
We now divide the load range into intervals, as shown in Fig. A3.

![Fig. A3: Load range divided into intervals](image)

This provides the ability to form a histogram by counting the number of time intervals contained in each load range. In this example, we assume that loads in Fig. A3 at the lower end of the range are “in” the range. The histogram for Fig. A3 is shown in Fig. A4.

![Fig. A4: Histogram](image)

Figure A4 may be converted to a probability mass function, pmf, (which is the discrete version of the probability density function, pdf) by dividing each count by the total number of time intervals, which is 23. The resulting plot is shown in Fig. A5.
Like any pmf, the summation of all probability values should be 1, which we see by the following sum:

$$0.087 + 0.217 + 0.217 + 0.174 + 0.261 + 0.043 = 0.999$$

(It is not exactly 1.0 because there is some rounding error). The probability mass function provides us with the ability to compute the probability of the load being within a range according to:

$$\Pr(\text{Load within Range}) = \sum_{L \in \text{Range}} \Pr(\text{Load} = L) \quad (2)$$

We may use the probability mass function to obtain the cumulative distribution function (CDF) as:

$$\Pr(\text{Load} \geq \text{Value}) = \sum_{L \geq \text{Value}} \Pr(\text{Load} = L) \quad (3)$$

From Fig. A5, we obtain:

$$\Pr(\text{Load} \geq 1) = \sum_{L \geq 1} \Pr(\text{Load} = L) = 1.0$$
$$\Pr(\text{Load} \geq 2) = \sum_{L \geq 2} \Pr(\text{Load} = L) = 1.0$$
$$\Pr(\text{Load} \geq 3) = \sum_{L \geq 3} \Pr(\text{Load} = L) = 1.0$$
\[
\Pr(\text{Load} \geq 4) = \sum_{L \geq 2} \Pr(\text{Load} = L) = 1.0
\]
\[
\Pr(\text{Load} \geq 5) = \sum_{L \geq 5} \Pr(\text{Load} = L) = 0.217 + 0.217 + 0.174 + 0.261 + 0.043 = 0.912
\]
\[
\Pr(\text{Load} \geq 6) = \sum_{L \geq 6} \Pr(\text{Load} = L) = 0.217 + 0.174 + 0.261 + 0.043 = 0.695
\]
\[
\Pr(\text{Load} \geq 7) = \sum_{L \geq 6} \Pr(\text{Load} = L) = 0.174 + 0.261 + 0.043 = 0.478
\]
\[
\Pr(\text{Load} \geq 8) = \sum_{L \geq 8} \Pr(\text{Load} = L) = 0.261 + 0.043 = 0.304
\]
\[
\Pr(\text{Load} \geq 9) = \sum_{L \geq 9} \Pr(\text{Load} = L) = 0.043
\]
\[
\Pr(\text{Load} \geq 10) = \sum_{L \geq 10} \Pr(\text{Load} = L) = 0
\]
Plotting these values vs. the load results in the CDF of Fig. A6.
Fig. A6: Cumulative distribution function

The plot of Fig. A6 is often shown with the load on the vertical axis, as given in Fig. A7.

If the horizontal axis of Fig. A7 is scaled by the time duration of the interval over which the original load data was taken, T, we obtain the load duration curve. This curve provides the number of time intervals that the load equals, or exceeds, a given load level. For example, if the original load data had been taken over a year, then the load duration curve would show the number of hours out of that year for which the load could be expected to equal or exceed a given load level, as shown in Fig. A8.
Load duration curves are useful in a number of ways.

- They provide guidance for judging different alternative plans. For example, one plan may be satisfactory for loading levels of 90% of peak and less. One sees from Fig. A8 that such a plan would be unsatisfactory for 438 hours per year (or 5% of the time).
- They identify the base load. This is the value that the load always exceeds. In Fig. A8, this value is 5 MW.
- They provide convenient calculation of energy, since energy is just the area under the load duration curve. For example, Fig. A9 shows the area corresponding to the base load energy consumption, which is $5\text{MW} \times 8760\text{hr} = 43800 \text{ MW-hrs}$. 

Fig. A8: Load duration curve
Fig. A9: Area corresponding to base load energy consumption

- They allow illustration of generation commitment policies and corresponding yearly unit energy production, as shown in Fig A10, where we see that the nuclear plant and coal plant #1 are base loaded plants, supplying 26280 MWhrs and 17520 MWhrs, respectively. Coal plant #2 and natural gas combined cycle (NGCC) plant #1 are the mid-range plants, and combustion turbine gas plant #1 is a peaker.
Load duration curves are also used in reliability, production costing, and expansion planning programs in computing different reliability indices.