Updated: 3/5/2024

Production Costing Models

1 Introduction

We have spent time identifying and characterizing costs associated with power plant investment and operations. In production cost analysis, we are interested in costs that vary with production – the *production costs*. These include fuel costs and variable O&M (VOM). For thermal plants, these costs are dominated by fuel costs, but since renewable plants have no fuel costs, they are dominated by VOM.

A production cost program, also referred to as a production cost model (PCM), or a production simulator (PS), is widely used throughout the electric power industry for many purposes. It is used in a stand-alone mode for several applications, as follows:

- Long-term planning: It is a subproblem to evaluate annual production costs within an expansion planning formulation, and in this sense, it is a fundamental part of an expansion planning application. It is also often used to simulate a single future year following identification of a preferred expansion. For example, the Midwest ISO used a production cost program to understand

the effect on energy prices of building HVDC from the Midwest US to the East coast.

- Fuel budgeting: Many companies run PCMs to determine the amount of natural gas and coal they will need to purchase in the coming weeks or months.
- Maintenance: PCMs are run to determine maintenance schedules for generation and transmission.
- Energy interchange: PCMs are run to facilitate negotiations for energy interchange between companies.

There are two essential inputs for any PCM:

1. Data characterizing future load and future renewable generation performance.
2. Data characterizing generation costs, in terms of:
   a. Heat rate curves and
   b. Fuel costs and VOM.

All PCMs require at least the above data. PCMs usually require generator outage rates as well. Specific programs will require additional data depending on their particular design.

The information provided by PCMs includes the annual costs of operating the generation facilities, a cost that is dominated by the fuel costs but also affected by the maintenance costs.

PCMs may also provide reliability indices as well as more time-granular estimates of fuel and maintenance costs, such as monthly, weekly, or hourly, from which it is then possible to obtain annual production costs.

A simplified way to conceptualize a PCM is as an hour-by-hour simulation of the power system over a duration of T hours, where the load is specified hourly, and then for each day or even week

- A unit commitment decision is made, using the security-constrained unit commitment (SCUC, a MIP);

and at each hour

- A dispatch decision is made to obtain the production costs for that hour, using the security-constrained economic dispatch (SCED, an LP)

The total production costs is then the sum of hourly production costs over all hours 1,…,T.

Advanced PCMs do in fact simulate hour-by-hour operation or even at 5-min intervals, with the commitment decision being made a day or a week at a time. As wind & solar generation increases and ancillary service markets become more important, PCM time steps are decreasing from 1-hour to the 5-minute solution frequency of the real-time market.

An important characterizing feature of PCMs is how it makes the unit commitment (UC) and dispatch decisions. The simplest approach makes the UC decision based on merit-ordering (also called priority ordering) such that units with lowest average cost are committed first. Startup costs are added when a unit is started, but in this simplest approach, those costs do not figure into the commitment decision.

The simplest approach for making the dispatch decision is referred to as the block loading principle, where each unit committed is fully loaded before the next unit is committed. The last unit is dispatched at that level necessary to satisfy the load.

Greater sophistication levels may be embedded in PCMs, the most significant of which are:
- Unit commitment and dispatch: A full unit commitment program may be run for certain blocks of intervals at a time, e.g., a week.
- Hydro: Hydro-thermal coordination may be implemented.
- Network representation: The network may be represented using DC flow and branch limits.
- Locational marginal prices: LMPs may be computed (they are the Lagrange multipliers on the nodal power balance equations imposed within the security-constrained economic dispatch (SCED).

- Maintenance schedules: Maintenance schedules may be taken into account.
- Uncertainty: Load uncertainty and generator unavailability may be represented using probabilistic methods. This allows for computation of reliability indices such as loss of load probability (LOLP) and expected unserved energy (EUE).
- Security constraints may be imposed using LODFs.

Below are some slides that Midwest ISO used at one point in time to describe production cost models.

MidwestISO
Energizing the Heartland

What is a Production Cost Model?

- Captures all the costs of operating a fleet of generators
  - Originally developed to manage fuel inventories and budget in the mid 1970's
- Developed into an hourly chronological security constrained unit commitment and economic dispatch simulation
  - Minimize costs while simultaneously adhering to a wide variety of operating constraints.
  - Calculate hourly production costs and location-specific market clearing prices.

4

What Are the Advantages of Production Cost Models?

■ Allows simulation of all the hours in a year, not just peak hour as in power flow models.

■ Allows us to look at the net energy price effects through
- LMP's and its components.
- Production cost.

■ Enables the simulation of the market on a forecast basis

■ Allows us to look at all control areas simultaneously and evaluate the economic impacts of decisions.

5

Disadvantages of Production Cost Models

■ Require significant amounts of data

■ Long processing times

■ New concept for many Stakeholders

■ Require significant benchmarking

■ Time consuming model building process
- Linked to power flow models

■ Do not model reliability to the same extent as power flow

6

2 Probability models

In this section, we briefly describe two kinds of probability models: for the loads, and for the generators. This material is adapted from the "convolution" notes.

1.0 Load duration curves

We have previously presented load duration curves (LDCs). We review them here for completeness, as they are an essential tool for use in basic production costing models we will present in these notes. In addition, we make one key point about them.

Figure 1 illustrates an LDC, and, as we have seen, when we normalize the time axis and then switch the

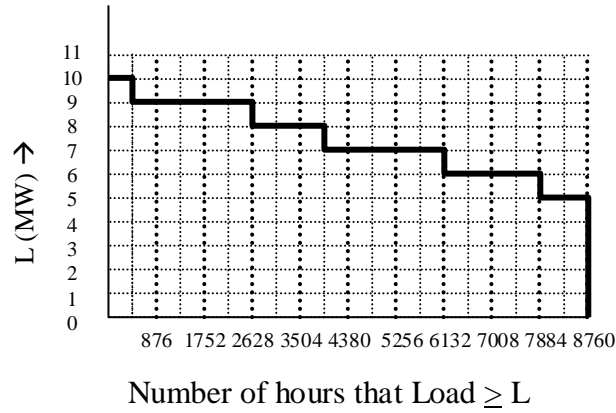time and load axes, we obtain a cumulative distribution function, as shown in Figure 2.

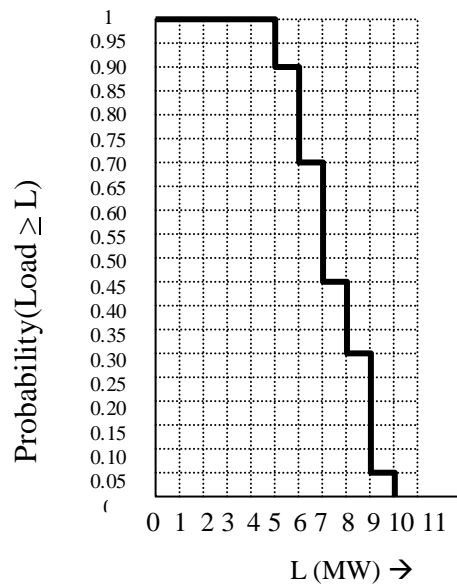**Figure 1: Illustration of LDC**

Figure 2: Cumulative Distribution Function

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. 1 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 the load *always* exceeds. In Figure 3, this value is 5 MW.
- They provide convenient calculation of energy, since energy is just the area under the load duration curve. For example, Figure 3 shows the area corresponding to base load energy consumption, which is 5MWx8760hr=43800 MW-hrs.
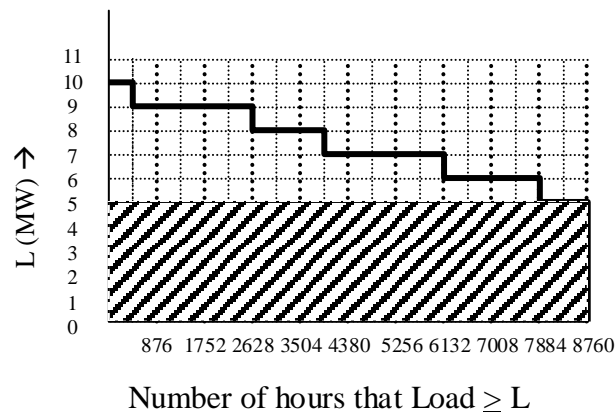

Figure 3: Area corresponding to base load energy consumption

- They allow illustration of generation commitment policies and corresponding yearly unit energy production, as shown in Figure 4, 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

9

gas combined cycle (NGCC) plant #1 are the mid-range plants, and combustion turbine gas plant #1 is a peaker. Assuming units come on-line in a fixed sequence like this is sometimes referred to as *merit-order* or priority-order loading.
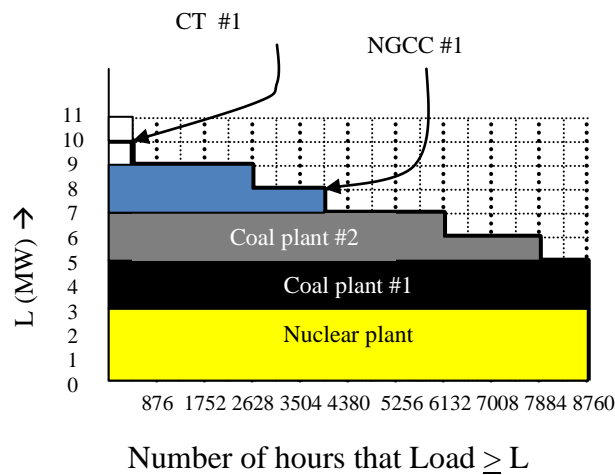
Figure 4: Illustration of Unit commitment policy

Load duration curves can be used in reliability, production costing, and expansion planning programs in computing different reliability indices, as we will see in Sections 5 and 6.

3 Generation probability models

We consider that generators obey a two-state model, i.e., they are either up or down, and we assume that the process by which each generator moves between states is Markov, i.e., the probability distribution of future states depends only on the current state and not on past states, i.e., the process is memoryless.

In this case, it is possible to show that unavailability (or *forced outage rate*, FOR) is the "steady-state" (or long-run) probability of a component not being available and is given by

$$U = p = \frac{\lambda}{\lambda + \mu} \qquad (1)$$

and the availability is the long-run probability of a component being available and is given by

$$A = q = \frac{\mu}{\lambda + \mu} \qquad (2)$$

where λ is the "failure rate" and μ is the "repair rate." Substituting $\lambda=1/MTTF$ and $\mu=1/MTTR$, where *MTTF* is the mean time to failure, and *MTTR* is the mean time to repair, we get that

$$U = p = \frac{MTTR}{MTTF + MTTR} \qquad (3)$$

$$A = q = \frac{MTTF}{MTTF + MTTR} \qquad (4)$$

See www.ee.iastate.edu/~jdm/ee653/U16-inclass.doc for complete derivation of these expressions.

The probability mass function representing the capacity outage corresponding to unit j is then given as $f_{Dj}(d_j)$, expressed as

$$f_{Dj}(d_j) = q_j\delta(d_j) + p_j\delta(d_j - C_j) \qquad (5)$$

and illustrated by Figure 5, where D_j is the random variable corresponding to *the increase in load* for

outage of unit *i*. We will return to this concept in Section 4.2 below.

$f_{Dj}(d_j)$

$A_j = q_j$

$U_j = p_j$

0   C_j Outage load, d_j

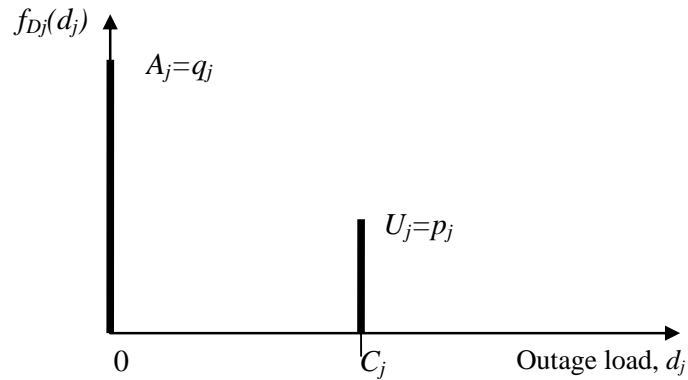**Figure 5: Two state capacity outage model**

4 Reliability calculations

4.1 Preliminary Definitions

Let's characterize the load shape curve (a load duration curve with the axes switched) with $t=g(d)$, as illustrated in Figure 6. It is important to note that the load shape curve characterizes the (forecasted) future time period and is therefore a probabilistic characterization of the demand.
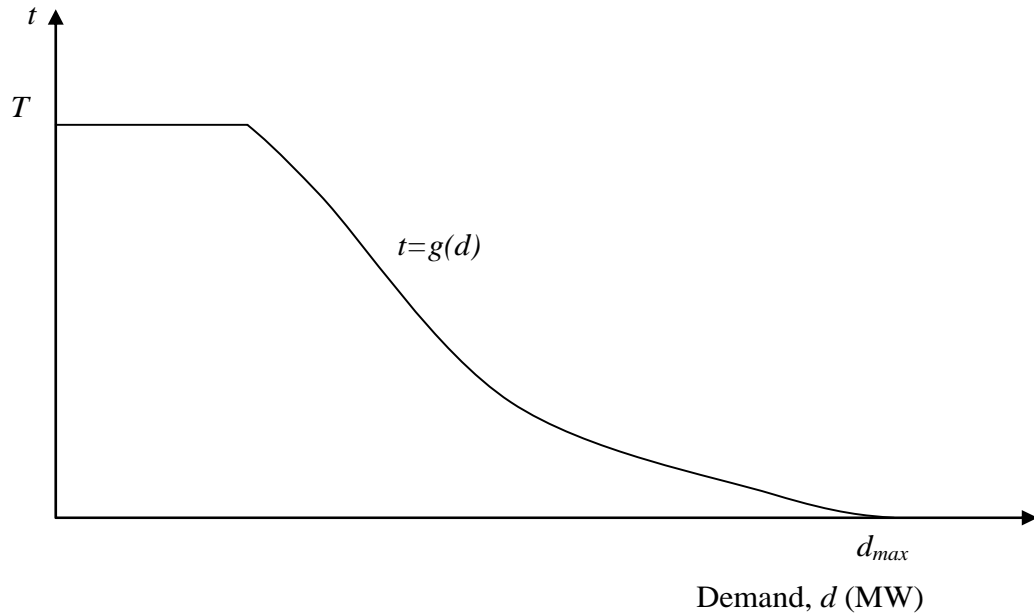
Figure 6: Load shape t=g(d)

Here:

- d is the system load;
- t is the number of time units in the interval T for which the load is greater than d and is most typically given in hours or days;
- $t=g(d)$ expresses functional dependence of t on d;
- T represents, most typically, a year.

The cumulative distribution function (cdf) is given by

$$F_D(d) = P(D > d) = \frac{t}{T} = \frac{g(d)}{T} \qquad (6)$$

One may also compute the *total energy E_T* consumed in the period T as the area under the curve, i.e.,

13

$$E_T = \int_0^{d_{max}} g(\lambda)d\lambda$$

(7)

The *average demand* in the period T is obtained from

$$d_{avg} = \frac{1}{T}E_T = \frac{1}{T}\int_0^{d_{max}}g(\lambda)d\lambda = \int_0^{d_{max}}F_D(\lambda)d\lambda$$

(8)

Now let's assume that the planned system generation capacity, i.e., the installed capacity, is C_T, and that $C_T < d_{max}$. This is an undesirable situation, since we will not be able to serve some demands, even when there is no capacity outage! Nonetheless, it serves well to understand the relation of the load duration curve to several useful indices. The situation is illustrated in Figure 7.
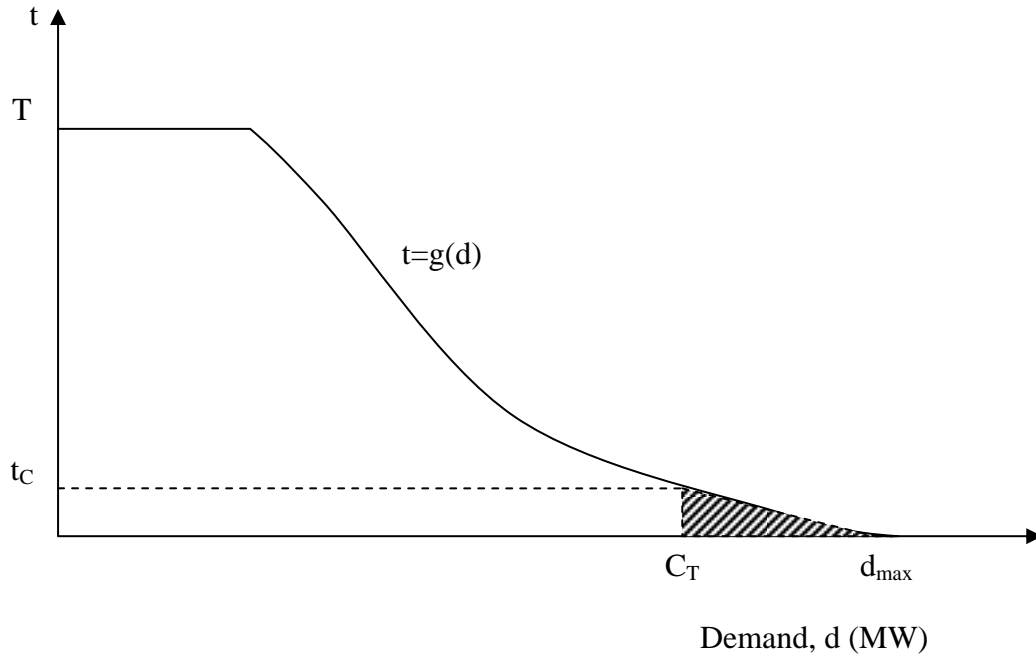

Figure 7: Illustration of Unserved Demand

Then, under the assumption that the given capacity C_T is perfectly reliable, we may express three useful reliability indices:

- Loss of load expectation, *LOLE*: the expected number of time units that load will exceed capacity

$$LOLE = t_{C_T} = g(C_T) \qquad (9)$$

- Loss of load probability, *LOLP*: the probability that the demand will exceed the capacity during *T*:

$$LOLP = P(D \geq C_T) = F_D(C_T) \qquad (10)$$

One may think that, if $d_{max} > C_T$, then *LOLP=1*. However, if $F_D(d)$ is a true probability distribution, then it describes the event $D \geq C_T$ with uncertainty associated with what the load is going to be, i.e., only with a probability. One can take an alternative view, that the load duration curve is certain, which would be the case if we were considering a previous year. In this case, LOLP should be thought of not as a probability but rather as the percentage of time during the interval *T* for which the load exceeds capacity.

It is of interest to reconsider the cdf given by (6), repeated here for convenience:

$$F_D(d) = P(D \geq d) = \frac{t}{T} = \frac{g(d)}{T} \qquad (6)$$

Substituting $d=C_T$, we get:

$$F_D(C_T) = P(D > C_T) = \frac{g(C_T)}{T} \qquad (*)$$

By (9), $g(C_T)=LOLE$; by (10), $P(D \geq C_T)=LOLP$, and so (*) becomes:

$$LOLP = \frac{LOLE}{T} \Rightarrow LOLE = LOLP \times T$$

which expresses that *LOLE* is the expectation of the number of time units within *T* that demand will exceed capacity.

- Expected demand not served, *EDNS*: If the average (or expected) demand is given by (8), then it follows that expected demand not served is (MW):

$$EDNS = \int_{C_T}^{d_{max}} F_D(\lambda)d\lambda \qquad (11)$$

which would be the same area indicated in Figure 7 when the ordinate is normalized to provide $F_D(d)$ instead of *t*. Reference [1] provides a rigorous derivation for (11).

- Expected energy not served, EENS: This is the total amount of time multiplied by the expected demand not served, i.e.,

$$EENS = T \int_{C_T}^{d_{max}} F_D(\lambda)d\lambda = \int_{C_T}^{d_{max}} g(\lambda)d\lambda \qquad (12)$$

which is the area shown in Figure 7.

4.2 Effective load

The notion of effective load is used to account for the unreliability of the generation, and it is essential for understanding the view taken in [1].

The basic idea is that the total system capacity is always C_T, and the effect of capacity outages are accounted for by changing the load model in an appropriate fashion, and then the different indices are computed as given in (10), (11), and (12).

A capacity outage of C_i is therefore modeled as an increase in the demand, not as a decrease in capacity!

We have already defined D as the random variable characterizing the demand. Now we define two more random variables:

- D_j is the random increase in load for outage of unit i.

- D_e is the random load accounting for outage of all units and represents the *effective* load.

Thus, the random variables D, D_e, and D_j are related:

$$D_e = D + \sum_{j=1}^{N} D_j \qquad (13)$$

It is important to realize that, whereas C_j represents the capacity of unit j and is a deterministic value, D_j represents the increase in load corresponding to outage of unit j and is a random variable. The

probability mass function (pmf) for D_j is assumed to be as given in Figure 8, i.e., a two-state model. We denote the pmf for D_j as $f_{Dj}(d_j)$.
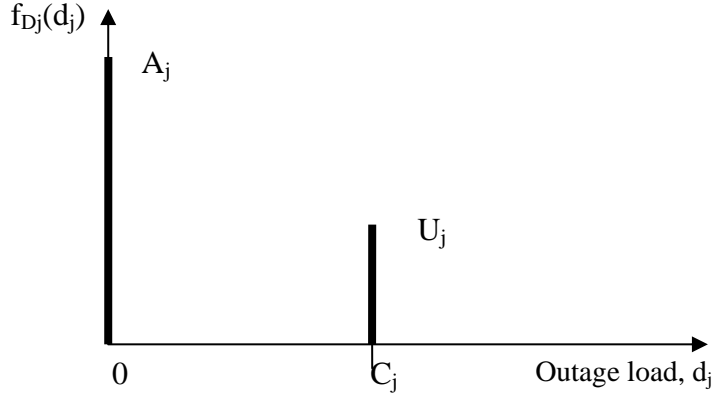
Figure 8: Two state generator outage model

Recall from probability theory that the pdf of the sum of two independent random variables is the convolution of their individual pdfs, that is, for random variables X and Y, with $Z=X+Y$, then

$$f_Z(z) = \int_{\lambda=-\infty}^{\infty} f_X(z-\lambda)f_Y(\lambda)d\lambda$$
(14)

In addition, it is true that the cdf of two random variables can be found by convolving the cdf of one of them with the pdf (or pmf) of the other, that is, for random variables X and Y, with $Z=X+Y$, then

$$F_Z(z) = \int_{\lambda=-\infty}^{\infty} F_X(z-\lambda)f_Y(\lambda)d\lambda$$
(15)

Let's consider the case for only one unit, i.e., from (13),

$$D_e = D + D_j \tag{16}$$

Then, by (15), we have that:

$$F_{D_e}^{(1)}(d_e) = \int_{\lambda=-\infty}^{\infty} F_{D_e}^{(0)}(d_e - \lambda) f_{D_j}(\lambda) d\lambda \tag{17}$$

where the notation $F_D^{(j)}(\cdot)$ indicates the cdf after the jth unit is convolved in. Under this notation, then, (16) becomes

$$D_e^{(j)} = D_e^{(j-1)} + D_j \tag{18}$$

and the general case for (17) is:

$$F_{D_e}^{(j)}(d_e) = \int_{\lambda=-\infty}^{\infty} F_{D_e}^{(j-1)}(d_e - \lambda) f_{D_j}(\lambda) d\lambda \tag{19}$$

which expresses the equivalent load after the jth unit is convolved in.

Since $f_{Dj}(d_j)$ is discrete (a pmf), we rewrite (19) as

$$F_{D_e}^{(j)}(d_e) = \sum_{d_j=-\infty}^{\infty} F_{D_e}^{(j-1)}(d_e - d_j) f_{D_j}(d_j) \tag{20}$$

From an intuitive perspective, (20) is providing the convolution of the load shape $F_D^{(j-1)}(\cdot)$ with the set of impulse functions comprising $f_{Dj}(d_j)$. When using a 2-state model for each generator, $f_{Dj}(d_j)$ is comprised of only 2 impulse functions, one at 0 and one at C_j.

Recalling that the convolution of a function with an impulse function simply shifts and scales that function, (20) can be expressed for the 2-state generator model as:

$$F_{D_e}^{(j)}(d_e) = A_j F_{D_e}^{(j-1)}(d_e) + U_j F_{D_e}^{(j-1)}(d_e - C_j) \quad (21)$$

So the cdf for the effective load following convolution with capacity outage pmf of the jth unit, is the sum of

- the original cdf, scaled by A_j and
- the original cdf, scaled by U_j and right-shifted by C_j.

Example 1:

Figure 9 illustrates the convolution process for a single unit $C_1=4$ MW supplying a system having peak demand $d_{max}=4$ MW, with demand cdf given as in plot (a) based on a total time interval of $T=1$ year.
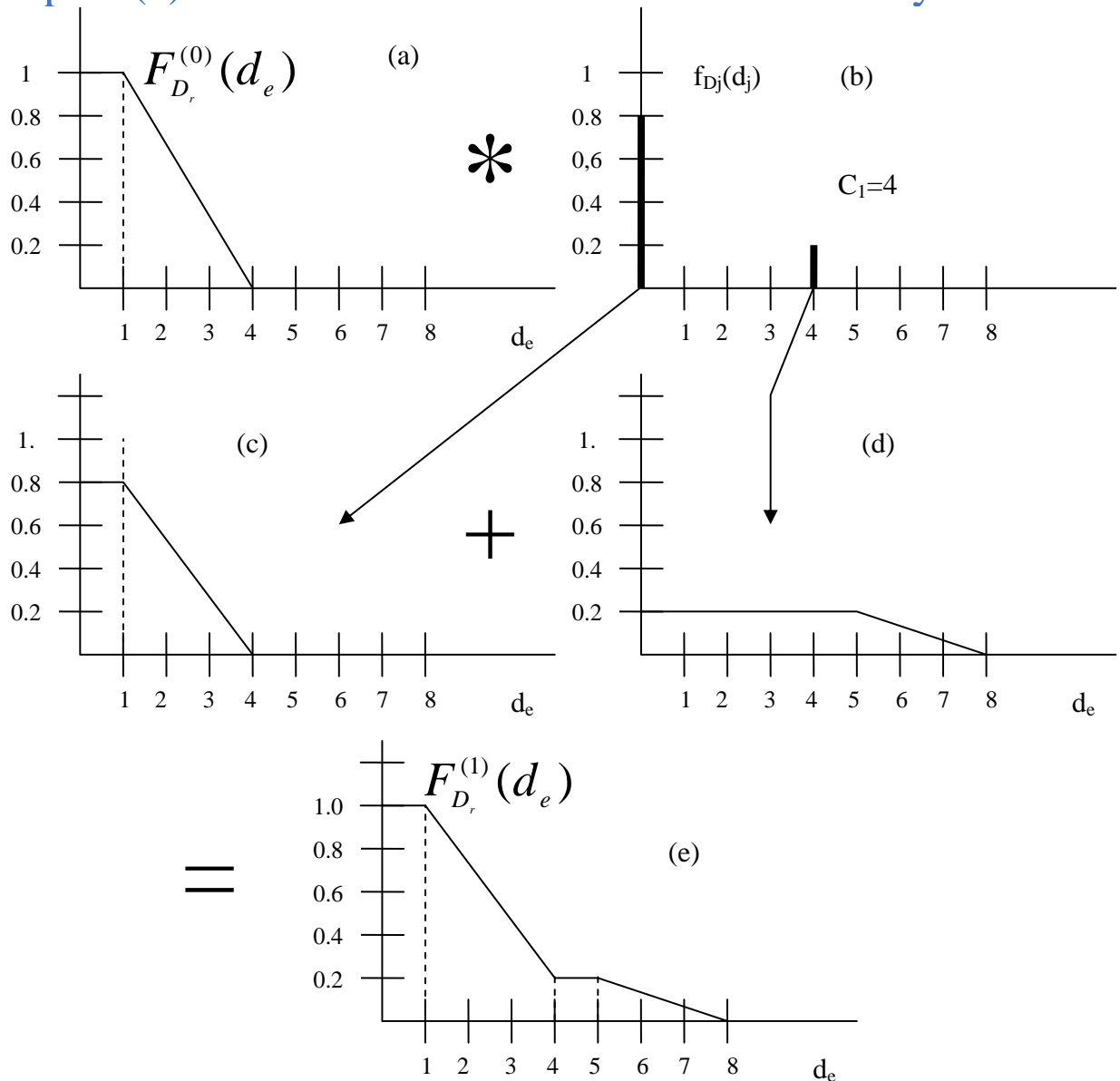
Figure 9: Convolving in the first unit

Plots (c) and (d) represent the intermediate steps of the convolution where the original cdf $F_{D_e}^{(0)}(d_e)$ was scaled by A_1=0.8 and U_1=0.2, respectively, and right-shifted by 0 and C_1=4, respectively. Note the effect of convolution is to spread the original cdf.

Plot (d) may raise some question since it appears that the constant part of the original cdf has been extended too far to the left. The reason for this apparent discrepancy is that all of the original cdf, in plot (a), was not shown. The complete cdf is illustrated in Figure 10 below, which shows clearly that $F_{D_e}^{(0)}(d_e)=1$ for $d_e<0$, reflecting the fact that $P(D_e>d_e)=1$ for $d_e<0$.
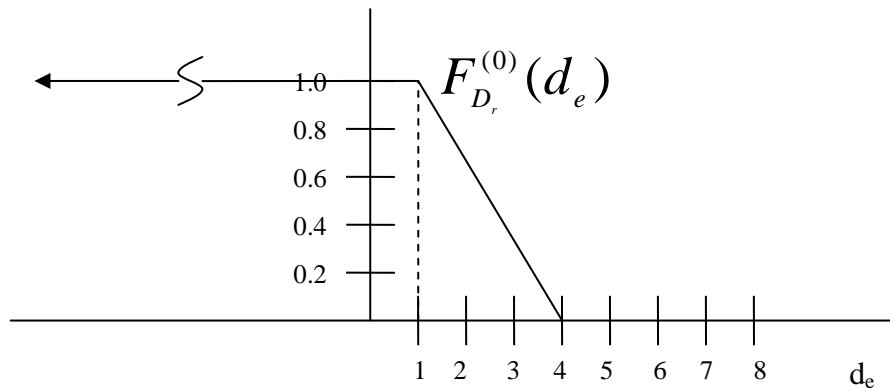
Figure 10: Complete cdf including values for $d_e<0$

Let's consider that the "first" unit we just convolved in is actually the only unit. If that unit were perfectly reliable, then, because C_1=4 and d_{max}=4, our system would never have loss of load. This would be the situation if we applied the ideas of Figure 7 to Figure

9, plot (a). However, plot (e) tells a different story. Figure 11 applies the ideas of Figure 7 to plot (e) to show how the cdf on the *equivalent load* indicates that, for a total capacity of $C_T=4$, we do in fact have some chance of losing load.
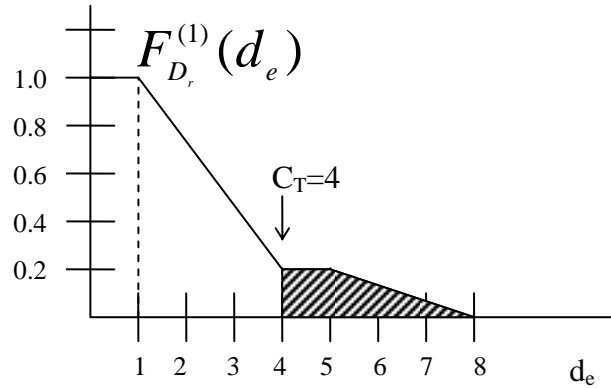
Figure 11: Illustration of loss of load region

The desired indices are obtained from (10),(11), (12):

$$LOLE = t_{C_T} = g_e(C_T) = T \times F_{D_r}(C_T = 4) = 1 \times 0.2 = 0.2\,years$$

A LOLE of 0.2 years is 73 days, a very poor reliability level that reflects the fact we have only a single unit with a high FOR=0.2.

The LOLP is given by:

$$LOLP = P(D_e \geq C_T) = F_{De}(C_T) = 0.2$$

and the EDNS is given by:

$$EDNS = \int_{C_T}^{d_{e,max}} F_{De}(\lambda)d\lambda$$

23

which is just the shaded area in Fig. U19.23, most easily computed using the basic geometry of the figure, according to:

$$0.2(1) + \frac{1}{2}(3)(0.2) = 0.5MW$$

The EENS is given by

$$EENS = T \int_{C_T}^{d_{e,max}} F_{De}(\lambda)d\lambda = \int_{C_T}^{d_{e,max}} g_e(\lambda)d\lambda$$

or *T×EDNS*=1(0.5)=0.5MW-years,

or 8760(0.5)=4380MWhrs.

Example 2: This example is from [2].
A set of generation data is provided in Table 1.

Table 1

Unit No.	Unit name	Rated capacity (MW(e))	Forced outage (%)	Type of fuel	Variable cost ($/MW·h)
1	NUC1	200	20	Nuclear	6.5
2	NUC2	200	20	Nuclear	6.5
3	COAL1	200	10	Coal	27.0
4	COAL2	200	10	Coal	27.0
5	OIL1	100	10	Oil	58.1
6	OIL2	100	10	Oil	58.1
7	OIL3	100	10	Oil	58.1
8	OIL4	100	10	Oil	58.1
9	CT1	100	5	Distillate oil	113.2
System capacity		1300			

The 4th column provides the forced outage rate, which we have denoted by *U*. The two-state

generator outage model for each unit is obtained from this value, together with the rated capacity, as illustrated in Figure 12, for unit 1. Notice that the units are ordered from least variable cost to highest variable cost.
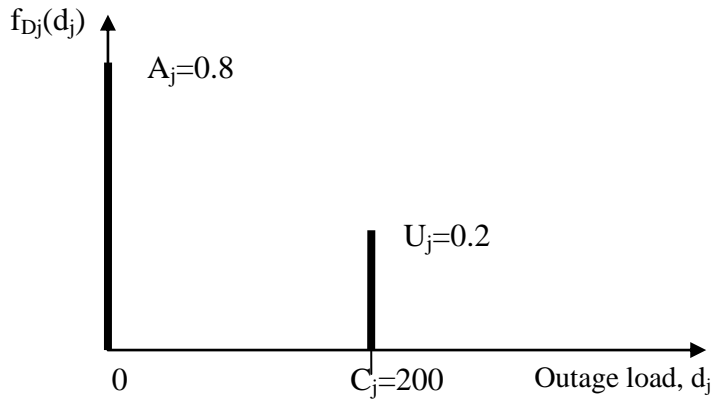
Figure 12: Two-state outage model for Unit 1

Load duration data is provided in Table 2, and the CDF is plotted in Figure 13.

Table 2

Load MW(e)	Fraction of time load exceeds given load
0	1.00
100	1.00
200	1.00
300	1.00
400 (minimum load)	1.00
500	0.80
600	0.40
700	0.20
800	0.10
900	0.05
1000 (peak load)	0.00
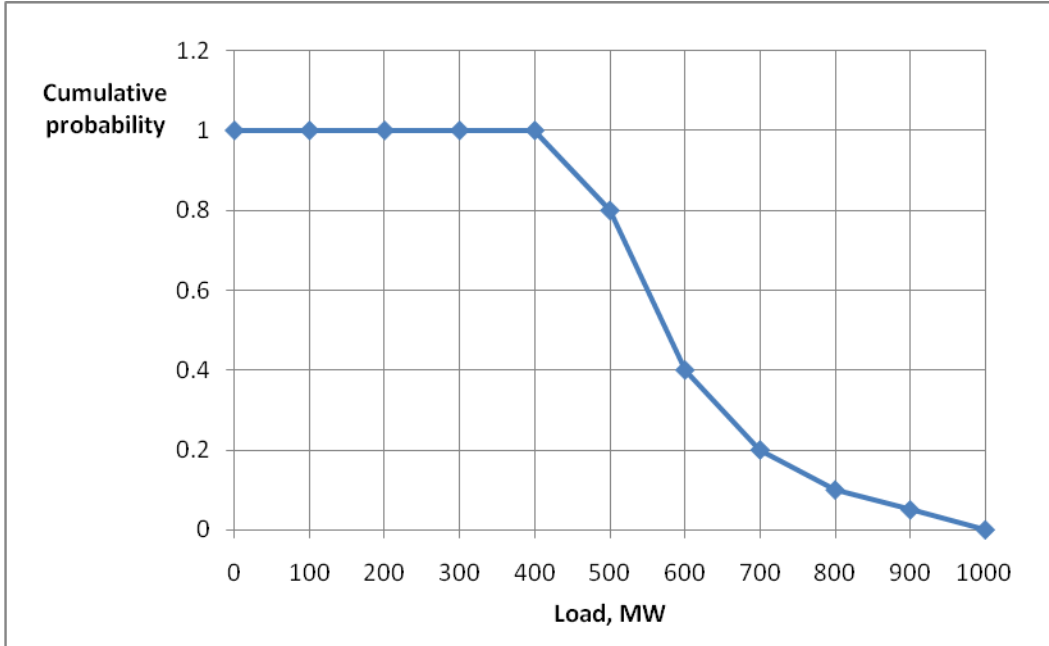
Figure 13: CDF for Example 2

We now deploy (21), repeated here for convenience,

$$F_{D_e}^{(j)}(d_e) = A_j F_{D_e}^{(j-1)}(d_e) + U_j F_{D_e}^{(j-1)}(d_e - C_j) \qquad (21)$$

to convolve in the unit outage models with the load duration curve of Figure 13. The procedure is carried out in an Excel spread sheet, and the result is provided in Figure 14. In Figure 14, we have shown

- Original load duration curve, F0;
- Load duration curve with unit 1 convolved in, F1.
- Load duration curve with all units convolved in, F9

We could, of course, show the load duration curves for any number of units convolved in, but this would be a cluttered plot.
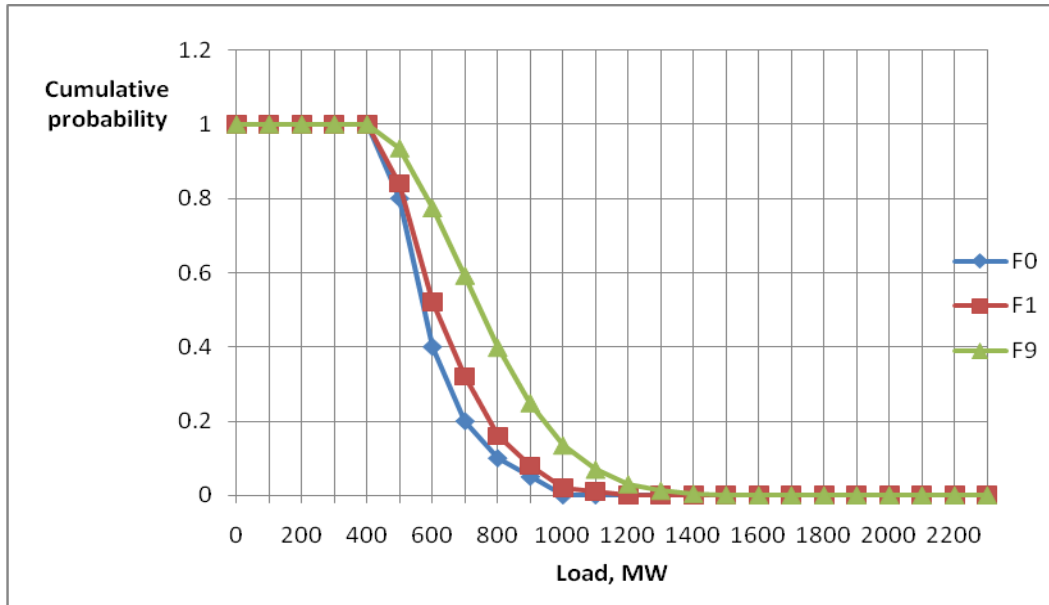
Figure 14

We also show, in Table 3, the results of the calculations performed to obtain the series of load duration curves (LDC) F0-F9. Notice the following:

- Each LDC is a column F0-F9
- The first column, in MW, is the load.
  - It begins at -200 to facilitate the convolution for the largest unit, which is a 200 MW unit.
  - Although it extends to 2300 MW, the largest actual load is 1000 MW; the extension is to obtain the equivalent load corresponding to a 1000 MW load with 1300 MW of failable generation.
- The entries in the table show the % time the load exceeds the given value.
- LOLP is, for a particular column, the % time the load exceeds the total capacity corresponding to that column, and is underlined.

For example, one observes that *LOLP=1* if we only have units 1 (F1, $C_T=200$) or only units 1 and 2 (F2, $C_T=400$). This is because the capacity would never be enough to satisfy the load, at any time. And *LOLP=0.6544* if we have only units 1, 2, and 3 (F3, $C_T=600$). This is because we would be able to supply the load for some of the time with this capacity. And *LOLP=0.012299* if we have all units (F9, $C_T=1300$), which is non-0 (in spite of the fact that $C_T>1000$) because units can fail.

Table 3

	F0	F1	F2	F3	F4	F5	F6	F7	F8	F9
		0.2	0.2	0.1	0.1	0.1	0.1	0.1	0.1	0.05
		0.8	0.8	0.9	0.9	0.9	0.9	0.9	0.9	0.95
		200	200	200	200	100	100	100	100	100
Load (MW)	Fraction of time load exceeds given load									
-200	1	1	1	1	1	1	1	1	1	1
-100	1	1	1	1	1	1	1	1	1	1
0	1	1	1	1	1	1	1	1	1	1
100	1	1	1	1	1	1	1	1	1	1
200	1	1	1	1	1	1	1	1	1	1
300	1	1	1	1	1	1	1	1	1	1
400	1	1	1	1	1	1	1	1	1	1
500	0.8	0.84	0.872	0.8848	0.89632	0.906688	0.916019	0.924417	0.931976	0.935377
600	0.4	0.52	0.616	0.6544	0.68896	0.709696	0.729395	0.748058	0.765694	0.774008
700	0.2	0.32	0.424	0.4688	0.5104	0.528256	0.5464	0.5647	0.583035	0.592168
800	0.1	0.16	0.232	0.2704	0.3088	0.32896	0.34889	0.368641	0.388247	0.397986
900	0.05	0.08	0.128	0.1576	0.18872	0.200728	0.213551	0.227085	0.241241	0.248591
1000	0	0.02	0.048	0.0664	0.0868	0.096992	0.107366	0.117984	0.128894	0.134512
1100	0	0.01	0.024	0.0344	0.04672	0.050728	0.055354	0.060556	0.066298	0.069428
1200	0	0	0.004	0.0084	0.0142	0.017452	0.02078	0.024237	0.027869	0.02979
1300	0	0	0.002	0.0042	0.00722	0.007918	0.008871	0.010062	0.01148	0.012299
1400	0	0	0	0.0004	0.0012	0.001802	0.002414	0.003059	0.00376	0.004146
1500	0	0	0	0.0002	0.0006	0.00066	0.000774	0.000938	0.00115	0.001281
1600	0	0	0	0	0.00004	0.000096	0.000152	0.000215	0.000287	0.00033
1700	0	0	0	0	0.00002	0.000022	2.94E-05	4.17E-05	5.9E-05	7.04E-05
1800	0	0	0	0	0	0.000002	0.000004	6.54E-06	1.01E-05	1.25E-05
1900	0	0	0	0	0	0	2E-07	5.8E-07	1.18E-06	1.62E-06
2000	0	0	0	0	0	0	0	2E-08	7.6E-08	1.31E-07
2100	0	0	0	0	0	0	0	0	2E-09	5.7E-09
2200	0	0	0	0	0	0	0	0	0	1E-10
2300	0	0	0	0	0	0	0	0	0	0

5 A basic production cost model

The most basic production cost model obtains production costs of thermal units over a period of time, say 1 year, by building upon the equivalent load duration curve described in Section 4.2.

To describe this, we will assume that generator variable cost, in \$/MWhr, for unit j operating at P_j over a time interval t, is expressed by

$$C_j(E_j)=b_jE_j$$

where $E_j=P_jt$ is the energy produced by the unit during the hour and b_j is the unit's _average_ variable costs of producing the energy. This is a simple generator cost model which lumps both fuel and non-fuel variable costs into the parameter b_j, and assumes all variable costs are proportional to energy production.

The production cost model begins by assuming the existence of a loading (or merit) order, which is how the units are expected to be called upon to meet the demand facing the system. We assume for simplicity that each unit consists of a single "block" of capacity equal to the maximum capacity. It is possible, and more accurate, to divide each unit into multiple capacity blocks, but there is no conceptual difference to the approach when doing so.

Table 1, listed previously in Example 2, provides the variable cost for each unit in the appropriate loading order. This table is repeated here for convenience.

Table 1

Unit No.	Unit name	Rated capacity (MW(e))	Forced outage (%)	Type of fuel	Variable cost ($/MW·h)
1	NUC1	200	20	Nuclear	6.5
2	NUC2	200	20	Nuclear	6.5
3	COAL1	200	10	Coal	27.0
4	COAL2	200	10	Coal	27.0
5	OIL1	100	10	Oil	58.1
6	OIL2	100	10	Oil	58.1
7	OIL3	100	10	Oil	58.1
8	OIL4	100	10	Oil	58.1
9	CT1	100	5	Distillate oil	113.2
System capacity		1300			

The criterion for determining loading order is clearly economic. Sometimes it is necessary to alter the economic loading order to account for must-run units or spinning reserve requirements. We will not consider these issues in the discussion that follows.

To motivate the approach, we introduce the concept of a unit's *observed energy demand* as the energy demand "seen" by a unit just before it is committed in the loading order. In considering a unit's observed energy demand, it will be the case that all higher-priority units will have been committed. One may also think of the observed energy demand as *the energy the system is requesting from the unit.*

30

If all higher-priority units are perfectly reliable (A_j=1), then the observed energy demand seen by the next unit would be the area under the LDC from the total committed capacity to a capacity equal to the total committed plus that of the next unit.

However, all higher-priority units are not perfectly reliable, i.e., they may fail according to the forced outage rate U_j. This means we must account for their stochastic behavior over time. This can be done in a straight-forward fashion by using the equivalent load duration curve developed for the last unit committed.

In the notation of (21), pg. 19, unit j sees an energy demand characterized by $F_{D_e}^{(j-1)}(d_e)$. Thus, the energy provided by unit j is proportional to the area under $F_{D_e}^{(j-1)}(d_e)$ from x_{j-1} to x_j, where

- x_{j-1} is the summed capacity over all previously committed units and
- x_j is the summed capacity over all previously committed units and unit j.

But unit j is only going to be available A_j% of the time. Also, since $F_{D_e}^{(j-1)}(d_e)$ is a probability function, we must multiply it by T, resulting in the following expression for energy provided by unit j [3]:

$$E_j = T \times A_j \int_{x_{j-1}}^{x_j} F_{D_e}^{(j-1)}(\lambda) d\lambda \qquad (22)$$

where

$$x_j = \sum_{i=1}^{j} C_i \, , \quad x_{j-1} = \sum_{i=1}^{j-1} C_i \qquad (23)$$

Referring back to Example 2, we describe the computations for the first three entries. This description is adapted from [2].

For unit 1, the original load duration curve F0 is used, as forced outages of any units in the system do not affect unit 1's observed load. The energy requested by the system from unit 1, is the area under $F_{D_e}^{(0)}(d_e)$ over the range of 0 to 200 MW (unit 1 's position in the loading order) times the number of hours in the period (8760) times unit 1's availability A_1. The area under $F_{D_e}^{(0)}(d_e)$ from 0 to 200, illustrated in Figure 15 below, is 200.

Figure 15

Therefore,

$$E_1 = 8760 \times 0.8 \times 200 = 1{,}401{,}600 \text{ MWhrs}$$

For unit 2, the load duration curve F1 is used, as forced outage of unit 1 will affect unit 2's observed load. The energy requested by the system from unit 2, is the area under $F_{D_e}^{(1)}(d_e)$ over the range of 200 to 400 MW (unit 2 's position in the loading order) times the number of hours in the period (8760) times unit 2's availability A_2. The area under $F_{D_e}^{(1)}(d_e)$ from 200 to 400, illustrated in Figure 16 below, is 200.
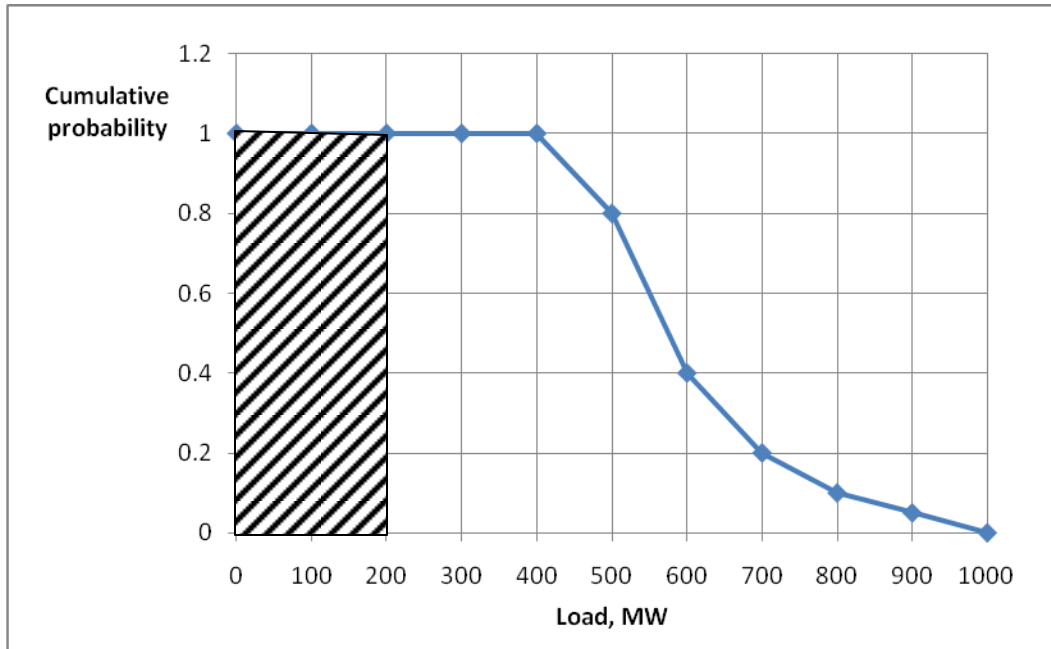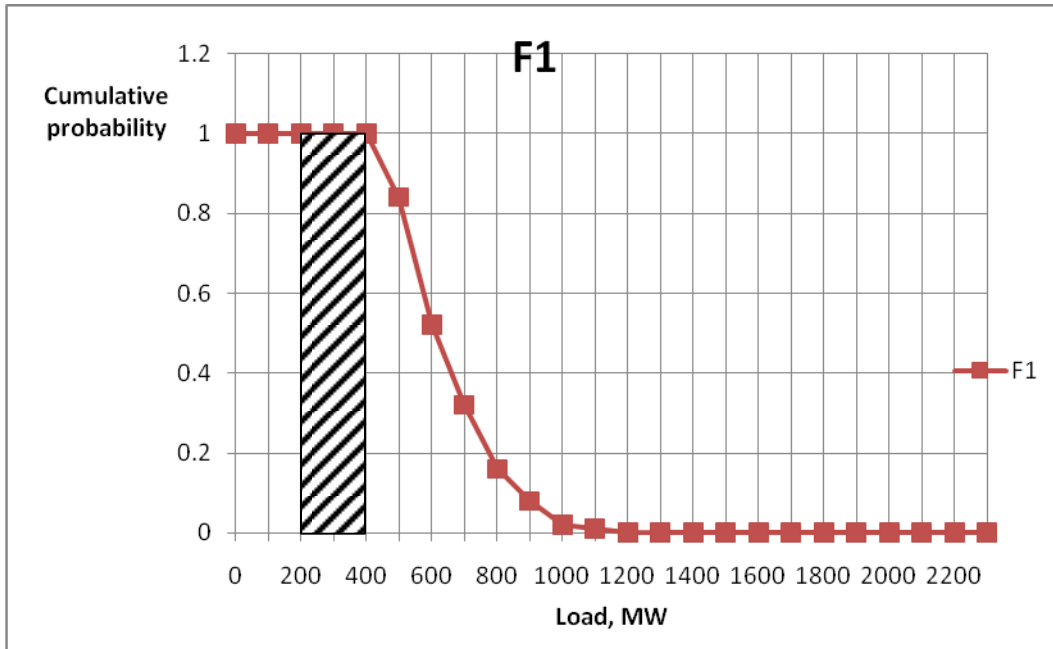
Figure 16

Therefore,

$$E_2 = 8760 \times 0.8 \times 200 = 1,401,600 \text{ MWhrs}$$

For unit 3, the load duration curve F2 is used, as forced outage of units 1 and 2 will affect unit 3's observed load. The energy requested by the system from unit 3, is the area under $F_{D_e}^{(2)}(d_e)$ over the range of 400 to 600 MW (unit 3 's position in the loading order) times the number of hours in the period (8760) times unit 3's availability A_3. The area under $F_{D_e}^{(2)}(d_e)$ from 400 to 600, illustrated in Figure 17, is calculated below Figure 17. The coordinates on Figure 17 are obtained from Table 3, repeated on the next page for convenience.

Figure 17

The area, indicated in Figure 17, is obtained as two applications of a trapezoidal area (1/2)(h)(a+b), as

$$\underbrace{\frac{1}{2}(100)(1+.872)}_{LeftPortion} + \underbrace{\frac{1}{2}(100)(.872+.616)}_{RightPortion}$$

$$= 93.6 + 74.4 = 168$$

Therefore,

$$E_3 = 8760 \times 0.9 \times 168 = 1,324,512 \text{MWhrs}$$

Table 3

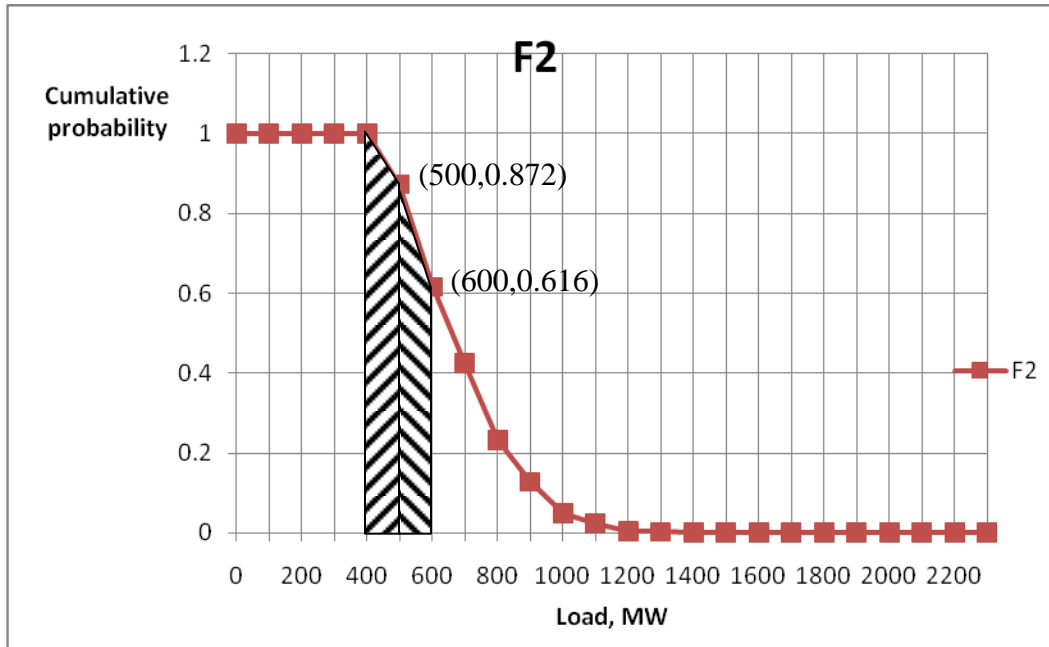| Load (MW)	F0	F1	F2	F3	F4	F5	F6	F7	F8	F9
		0.2	0.2	0.1	0.1	0.1	0.1	0.1	0.1	0.05
		0.8	0.8	0.9	0.9	0.9	0.9	0.9	0.9	0.95
		200	200	200	200	100	100	100	100	100
	Fraction of time load exceeds given load									
-200	1	1	1	1	1	1	1	1	1	1
-100	1	1	1	1	1	1	1	1	1	1
0	1	1	1	1	1	1	1	1	1	1
100	1	1	1	1	1	1	1	1	1	1
200	1	1	1	1	1	1	1	1	1	1
300	1	1	1	1	1	1	1	1	1	1
400	1	1	1	1	1	1	1	1	1	1
500	0.8	0.84	0.872	0.8848	0.89632	0.906688	0.916019	0.924417	0.931976	0.935377
600	0.4	0.52	0.616	0.6544	0.68896	0.709696	0.729395	0.748058	0.765694	0.774008
700	0.2	0.32	0.424	0.4688	0.5104	0.528256	0.5464	0.5647	0.583035	0.592168
800	0.1	0.16	0.232	0.2704	0.3088	0.32896	0.34889	0.368641	0.388247	0.397986
900	0.05	0.08	0.128	0.1576	0.18872	0.200728	0.213551	0.227085	0.241241	0.248591
1000	0	0.02	0.048	0.0664	0.0868	0.096992	0.107366	0.117984	0.128894	0.134512
1100	0	0.01	0.024	0.0344	0.04672	0.050728	0.055354	0.060556	0.066298	0.069428
1200	0	0	0.004	0.0084	0.0142	0.017452	0.02078	0.024237	0.027869	0.02979
1300	0	0	0.002	0.0042	0.00722	0.007918	0.008871	0.010062	0.01148	0.012299
1400	0	0	0	0.0004	0.0012	0.001802	0.002414	0.003059	0.00376	0.004146
1500	0	0	0	0.0002	0.0006	0.00066	0.000774	0.000938	0.00115	0.001281
1600	0	0	0	0	0.00004	0.000096	0.000152	0.000215	0.000287	0.00033
1700	0	0	0	0	0.00002	0.000022	2.94E-05	4.17E-05	5.9E-05	7.04E-05
1800	0	0	0	0	0	0.000002	0.000004	6.54E-06	1.01E-05	1.25E-05
1900	0	0	0	0	0	0	2E-07	5.8E-07	1.18E-06	1.62E-06
2000	0	0	0	0	0	0	0	2E-08	7.6E-08	1.31E-07
2100	0	0	0	0	0	0	0	0	2E-09	5.7E-09
2200	0	0	0	0	0	0	0	0	0	1E-10
2300	0	0	0	0	0	0	0	0	0	0

Continuing in this way, we obtain the energy produced by all units. This information, together with the average variable costs from Table 1, and the resulting energy cost, is provided in Table 4 below.

Table 4

Unit	MW-hrs	Avg. Variable Costs, $/MWhr	Costs, $
1	1,401,600	6.5	9,110,400
2	1,401,600	6.5	9,110,400
3	1,324,500	27.0	35,761,500
4	734,200	27.0	19,823,400
5	196,100	58.1	11,393,410
6	117,400	58.1	6,820,940
7	64,100	58.1	3,724,210
8	33,400	58.1	1,940,540
9	16,400	113.2	1,856,480
Total	5,289,300		99,541,280

It is interesting to note that the total energy supplied, $E_{S,9,1300}$=5,289,300 MWhrs, is less than what one obtained when the original load duration curve is integrated. This integration can be done by applying our trapezoidal approach to curve F0 in Table 7. Doing so results in $E_{D,0,1300}$=5,299,800 MWhrs. The difference is

$E_{D,0,1300}$-$E_{S,9,1300}$=5,299,800-5,289,300=10,500 MWhr

What is this difference of 10,500 MWhrs?

➔To answer this question, we first describe notation.

- $E_{S,j,C}$ is the energy supplied by the generation, as computed by our method up to and including unit j, for a total capacity of C. Here, the C gen capacity is failable.

- $E_{D,j,C}$ is the energy demanded by the load as computed by integrating from 0 to C the effective LDC after convolving in unit j. Here, the C gen capacity is perfectly reliable.

Now, we let's consider:

- The total area under the original curve F0, integrated from 0 to 1000 (the peak load), is 5,299,800 MWhrs, as shown in Figure 18. This is the amount of energy provided to the *actual* load if it were supplied by perfectly reliable generation having capacity of 1000 MW. We can denote this as $E_{D,0,1000}$, or, as indicated above, it is equivalent to denote it as $E_{D,0,1300}$, since there is no area under the curve from 1000 to 1300.
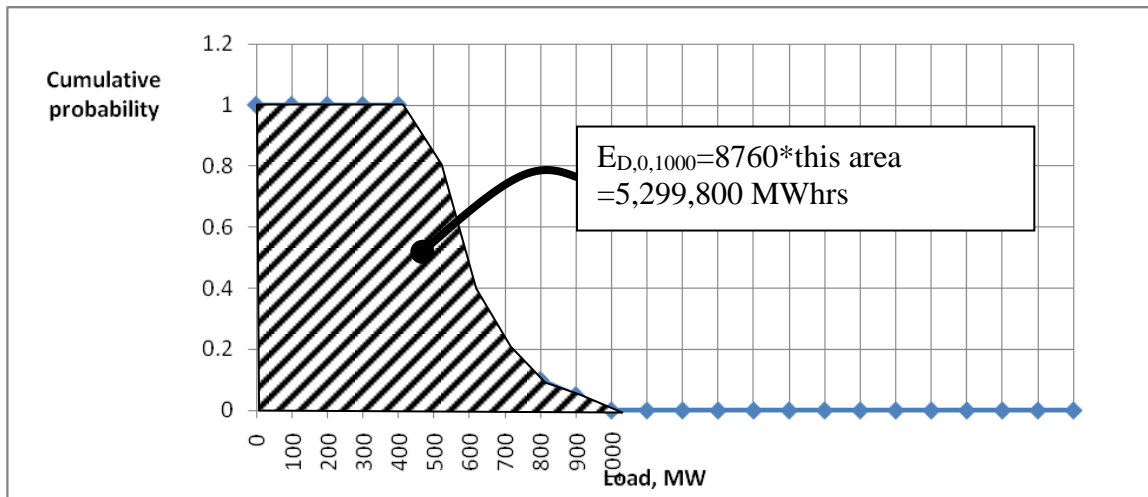

Figure 18

- The area under the final curve, F9, integrated from 0 to 1300 MW (total generation capacity) is $E_{D,9,1300}$=6,734,696 MWhrs, as shown in Figure 19. This is the amount of energy provided to the *effective* load if it were supplied by perfectly reliable generation having capacity of 1300 MW.
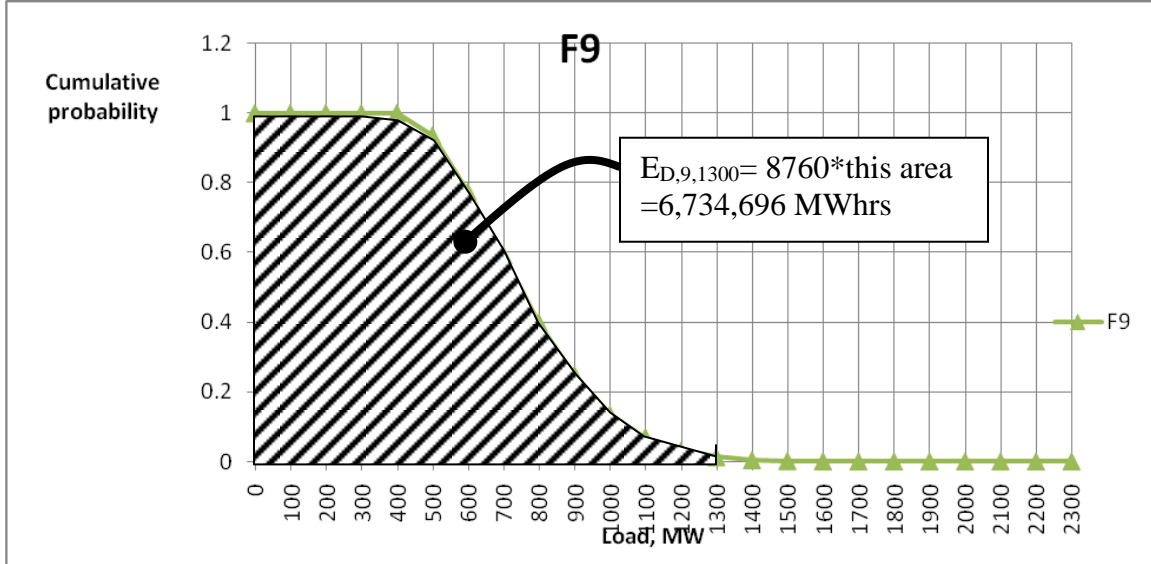
Figure 19

The energy represented by the area of Figure 19, which is the energy provided to the *effective* load if it were supplied by perfectly reliable generation having capacity of 1300 MW, is greater than the energy provided by the actual 1300 MW, that is

$$E_{D,9,1300} > E_{S,9,1300}$$

because $E_{D,9,1300}$ includes load required to be served when the generators are outaged, and this portion was explicitly removed from the calculation of Table 4 ($E_{S,9,1300}$). One observes this by considering a system with only a single unit. Recalling the general formula (22) for obtaining *actual energy supplied by a unit* per the method of Table 4:

$$E_j = T \times A_j \int_{x_{j-1}}^{x_j} F_{D_e}^{(j-1)}(\lambda) d\lambda$$

(22)

and applying this to the one-unit system, we get:

$$E_{S,1,C_1} = E_1 = T \times A_1 \int_0^{C_1} F_{D_e}^{(0)}(\lambda)d\lambda \tag{24}$$

In contrast, the energy $E_{D,1,C1}$ obtained when we integrate the effective load duration curve (accounting for only the one unit) is

$$E_{D,1,C_1} = T \int_0^{C_1} F_{D_e}^{(1)}(\lambda)d\lambda \tag{25}$$

Recalling the convolution formula (21),

$$F_{D_e}^{(j)}(d_e) = A_j F_{D_e}^{(j-1)}(d_e) + U_j F_{D_e}^{(j-1)}(d_e - C_j) \tag{21}$$

and for the one-unit case, we get

$$F_{D_e}^{(1)}(d_e) = A_1 F_{D_e}^{(0)}(d_e) + U_1 F_{D_e}^{(0)}(d_e - C_1) \tag{26}$$

Substituting (26) into (25) results in

$$E_{D,1,C_1} = T \int_0^{C_1} A_1 F_{D_e}^{(0)}(\lambda) + U_1 F_{D_e}^{(0)}(\lambda - C_1)d\lambda \tag{27}$$

Breaking up the integral gives

$$E_{D,1,C_1} = T \int_0^{C_1} A_1 F_{D_e}^{(0)}(\lambda)d\lambda + T \int_0^{C_1} U_1 F_{D_e}^{(0)}(\lambda - C_1)d\lambda$$

$$= T \times A_1 \int_0^{C_1} F_{D_e}^{(0)}(\lambda)d\lambda + T \times U_1 \int_0^{C_1} F_{D_e}^{(0)}(\lambda - C_1)d\lambda \tag{28}$$

Comparing (28) with (24), repeated here for convenience:

$$E_{S,1,C_1} = E_1 = T \times A_1 \int_0^{C_1} F_{D_e}^{(0)}(\lambda)d\lambda \tag{24}$$

we observe the expressions are the same except for the presence of the second integration in (28). This proves that $E_{D,1,C1} > E_{S,1,C1}$, i.e.,

effective energy demanded > energy served by generation

- Now consider computing the energy consumed by the total effective load as represented by Figure 20 (note that in this figure, the curve should go to zero at Load=2300 but does not due to limitations of the drawing facility used).
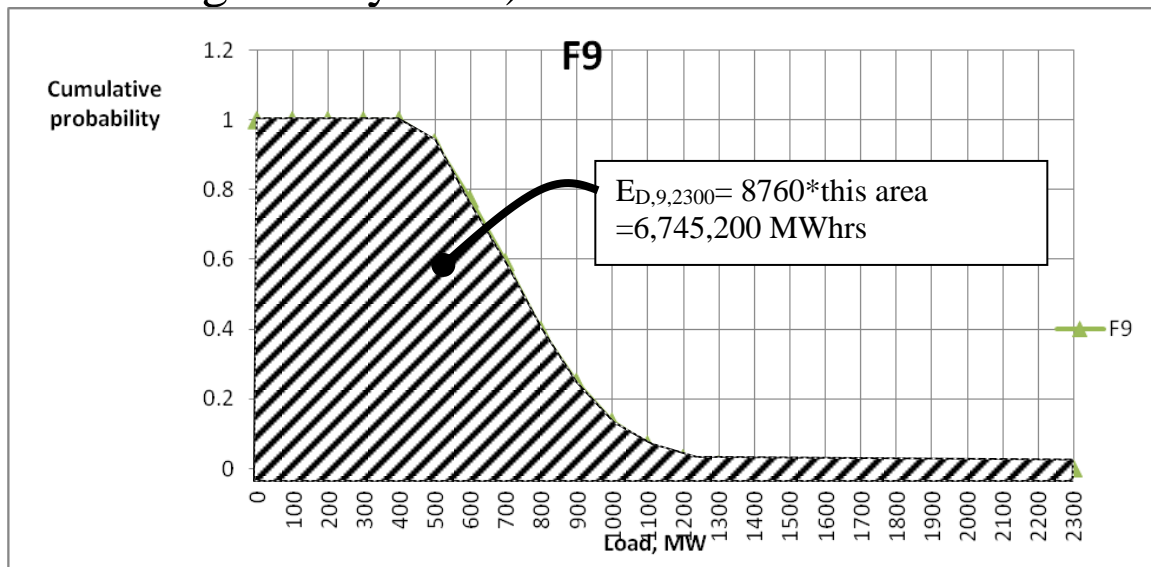

Figure 20

Using the trapezoidal method to compute this area results in $E_{D,9,2300}$=6745200 MWhrs, which is the energy provided to the *effective* load if it were supplied by perfectly reliable generation having capacity of 2300 MW. This would leave zero energy unserved.

- The difference between

41

- $E_{D,9,2300}$, the energy provided to the effective load if it were supplied by 2300 MW of perfectly reliable generation and
- $E_{D,9,1300}$, the energy provided to the effective load if it were supplied by 1300 MW of perfectly reliable generation

is given by:

$E_{D,9,2300} - E_{D,9,1300} = 6,745,200 - 6,734,696 = 10,504$ MWhr

This is the expected energy not served (EENS), sometimes called the expected unserved energy (EUE).

We observe, then, that we can obtain EENS in two different approaches.

1. $E_{D,0,1300} - E_{S,9,1300} = 5,299,800 - 5,289,300 = 10,500$ MWhrs
2. $E_{D,9,2300} - E_{D,9,1300} = 6,745,200 - 6,734,696 = 10,504$ MWhrs

We remind the reader of the notation:

- $E_{S,j,C}$ is the energy supplied by the generation, as computed by our method up to and including unit j, for a total capacity of C. Here, the C gen capacity is failable.
- $E_{D,j,C}$ is the energy demanded by the load as computed by integrating from 0 to C the effective LDC after convolving in unit j. Here, the C gen capacity is considered to be perfectly reliable.

Approach 1 may be computationally more convenient for production costing because $E_{D,0,1300}$ comes from the original LDC, and $E_{s,9,C}$ is easily obtained as the summation of all the energy values (via Table 4).

Approach 2 may be more convenient conceptually as it is simply the area under the final effective load curve from total capacity (I call it C_T) to infinity.

6 Industry-grade commercial PCMs

In the previous notes, we reviewed a relatively simple production cost model (PCM). This PCM required two basic kinds of input data:
- Annual load duration curve
- Unit data:
    o Capacity
    o Forced outage rate
    o Variable costs

It then computes load duration curves for *effective* load (which accounts for the unreliability of the generators supplying that load) through a convolution process and provides the following information:
- Reliability indices: LOLP, LOLE, EDNS, EENS (EUE)
- Annual energy produced by each unit
- Annual production costs for each unit
- Total system production costs

Another approach to PCMs is to simulate each hour of the year. This allows more rigorous models and more refined results, which comes with a significant computational cost. There are several industry-grade models available today, as summarized in Table 5. Reference [4] summarizes these products; additional information is found in [5, 6].

Table 5: Summary of PCMs

Name	Owner
Promod	ABB
Plexos	Energy Exemplar
SERVM	Astrape Consulting
Gridview	ABB
GE-MAPS	GE
UPLAN	LCG
GTMax	Argonne Nat. Lab
SCOPE	Nexant
GenTrader	PCI
RIM	PNNL

Promod and Plexos are two such models heavily used today. I will describe the conceptual approach to such PCMs. They generally consist of the following loops:

1. <u>Annual loop</u>: Most PCMs have only one annual loop, i.e., the annual simulation is deterministic. But it is conceivable to make multiple runs through a particular year, each time selecting various variables based on probability distributions for those variables. Such an approach is referred to as a Monte Carlo approach, and it requires many loops in order to "converge" with respect to the average annual production costs. It can be quite computational[1].

---

[1] One may reduce computational cost for this loop by selecting "representative" time intervals, e.g., "representative weeks." The idea here is that although a "complete" year-long simulation requires simulation of 52 consecutive weeks, many weeks will be very similar. Thus, the 52 weeks may become only 50 if weeks 24, 25, and 26 are judged similar and only 45 if weeks 10, 11, 12, 33, 34, and 35 are judged similar, with weeks 25 and 12 chosen as the "representative weeks" for their respective groups. By adjusting the "similarity criteria" one may reduce the number of simulated weeks as much as desired,

2. <u>UC loop</u>: The program must have a way for deciding, in each hour, which units are committed. A UC program is typically implemented within the PCM on a weekly, 48, or 24-hour basis. The latter may be preferred by some because it is consistent with the fact that most electricity market structures include a day-ahead market using the security-constrained-unit-commitment (SCUC).

3. <u>Hourly loop</u>: A security-constrained-economic-dispatch (SCED) is implemented to dispatch committed units. In addition, it is within the hourly loop that reliability indices are computed. There are two broad ways of doing this:

    i. Generation adequacy assessment

    ii. Multiarea adequacy assessment.

We will study multiarea adequacy assessment in the next set of notes. In this document, we have focused on generation adequacy assessment, and methods used in PCM are similar, with the exception that the load is deterministic during the hour and so is represented by a single number. Thus, the only randomness is in regards to the status of committed generators and whether they are in-service or out-of-service due to a forced

---

although there will be a reduction in simulation accuracy as the number of representative weeks decreases. There has been much research on developing similarity criteria using machine-learning based grouping methods such as K-means. Key to maintaining good fidelity is the choice of features used to form the groups, e.g., PCM input data such as load levels together with unit or line maintenance schedules. PCM output data such as line flows and LMPs can also be useful.

outage. The are two broad methods for performing generation adequacy assessment:

- Analytic: Because the load is deterministic (represented by a single number), the capacity outage table approach is utilized (instead of the effective load duration approach).
- Monte Carlo: In this method, the status of each committed generator is identified via random draw of a number between $0\rightarrow1$. If a number between 0 and the probability of the unit being down (e.g., $0\rightarrow0.03$) is chosen, the unit is outaged. If a number between probability of unit being down & 1 is chosen (e.g., $0.03\rightarrow1$), the unit is up.

Comment: In hourly production cost models, it is important to use outage replacement rate (ORR) as the probability of the unit being down, rather than the forced outage rate (FOR). The ORR is the probability that the unit will go down in the next time interval given it is up at the beginning of the time interval. It is shown in U20 (Section U20.4) to be given by λT where λ is the failure rate (failures per hour) and T is the time interval of interest, usually 1 hour.

7 A reported model

Models are reported in [7, 8] which capture some of the above attributes. Figure 22 shows a PCM flow chart from [7], and Figure 22 shows a PCM flow chart from [8].

46

There are two loops in both of these flow charts. These loops are:

- Inner loop: hour-by-hour simulation for one day (SCED);
- Middle loop: day-by-day simulation for one year (SCUC);

The second flow chart also has an outer loop that enables multiple years to be simulated, but PCMs are usually done for only a single year.

In Figure 21, the first loop includes a hydro-thermal coordination step, important to do for hydro-rich systems.

In Figure 22, the flow chart to the right provides the evaluation for each hour of adequacy indices by randomly selecting unit outages (only one selection per hour). With respect to this evaluation, observe note 1 in the figure, which indicates a variety of adequacy evaluations may be computed within this flow chart to evaluate the hourly resource adequacy.

There is an excellent treatment of PCM fundamentals at this URL, but you have to be an IEEE PES member to get it.

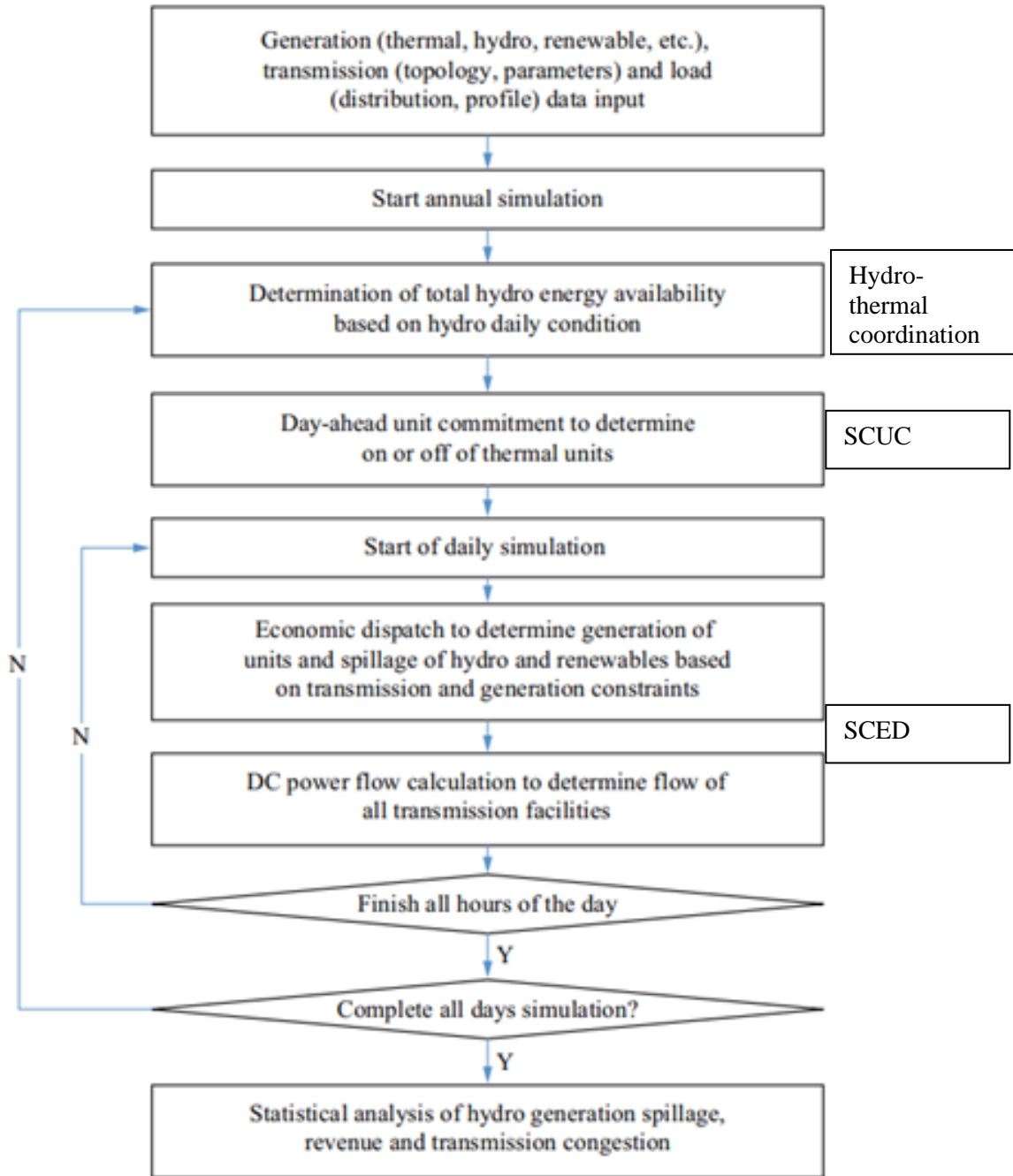https://resourcecenter.ieee-pes.org/education/tutorials/pes_ed_tut_pbr_081021_sld

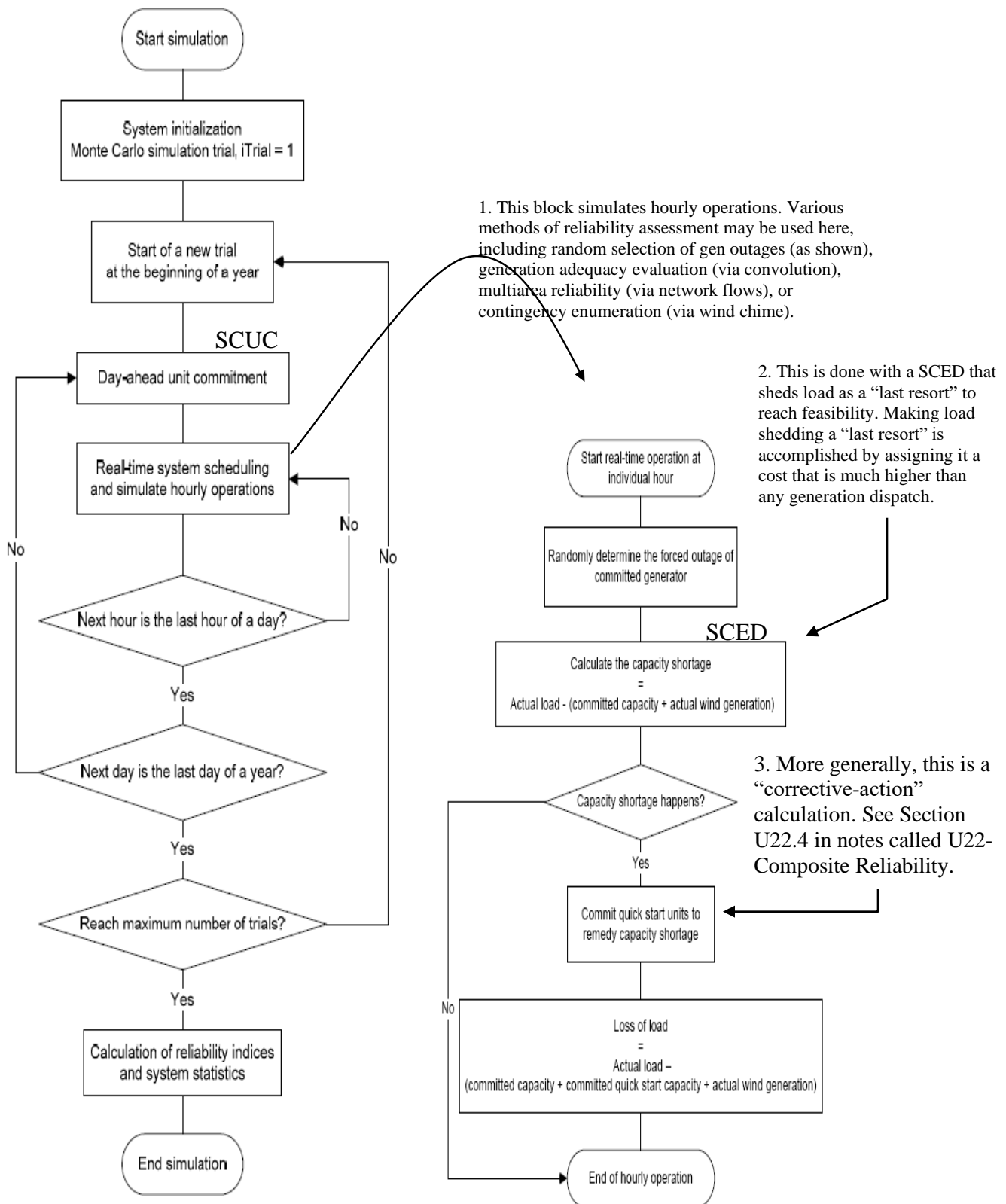**Figure 21: Flow chart for typical PCM**

Figure 22: Flow chart for typical PCM

The flow chart contains the following text elements:

- Start simulation
- System initialization / Monte Carlo simulation trial, iTrial = 1
- Start of a new trial / at the beginning of a year
- SCUC
- Day-ahead unit commitment
- Real-time system scheduling / and simulate hourly operations
- No
- Next hour is the last hour of a day?
- Yes
- No
- Next day is the last day of a year?
- Yes
- No
- Reach maximum number of trials?
- Yes
- Calculation of reliability indices / and system statistics
- End simulation

- Start real-time operation at / individual hour
- Randomly determine the forced outage of / committed generator
- SCED
- Calculate the capacity shortage / = / Actual load - (committed capacity + actual wind generation)
- Capacity shortage happens?
- Yes
- No
- Commit quick start units to / remedy capacity shortage
- Loss of load / = / Actual load − / (committed capacity + committed quick start capacity + actual wind generation)
- End of hourly operation

1. This block simulates hourly operations. Various methods of reliability assessment may be used here, including random selection of gen outages (as shown), generation adequacy evaluation (via convolution), multiarea reliability (via network flows), or contingency enumeration (via wind chime).

2. This is done with a SCED that sheds load as a "last resort" to reach feasibility. Making load shedding a "last resort" is accomplished by assigning it a cost that is much higher than any generation dispatch.

3. More generally, this is a "corrective-action" calculation. See Section U22.4 in notes called U22-Composite Reliability.

One more comment is worth making here. Because power system flexibility demand and provision depend on sub-hourly actions, a high-fidelity treatment of operational flexibility within a PCM (of interest because wind and solar growth significantly raise the importance of operational flexibility) can benefit from employing a temporal granularity consistent with this need. To this end, there has been recent interest to implement PCMs using

- one hour time steps for commitment decisions (for which decision variables are integers) and
- five-minute time steps for economic dispatch decisions (for which decision variables are continuous).

This kind of design has been called a subordinated time structure [9]. It is illustrated in Figure 23.
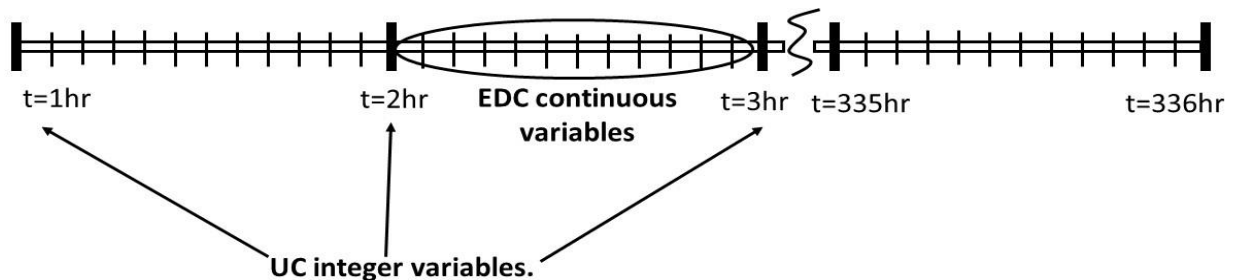

Figure 23: Subordinated time structure

The new feature of the subordinated time structure is that there are two layers of decision variables: the integer variables associated with hourly commitment decisions, and the continuous values associated with the 5-min dispatch. Thus, there are two solutions, one with a MILP across the second loop time frame (a day or week), and the other having 12 LP solutions within each hour. An initial approach decouples the two, but there may be benefit from solving them together, as recently investigated in [9].
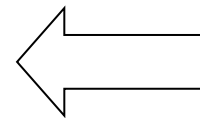
8 MISO's use of a PCM

Below are a few more slides that characterize how MISO utilizes production costing. Although these slides are dated, most points in them are applicable.
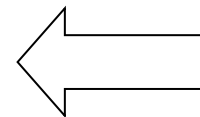
MidwestIS◉
Energizing the Heartland

Background

■ PROMOD is a Production Cost Model developed by Ventyx (Formerly known as NewEnergy Associates, A Siemens Company).

Promod is now marketed by Hitachi-ABB.

■ Detailed generator portfolio modeling, with both region zonal price and nodal LMP forecasting and transmission analysis including marginal losses
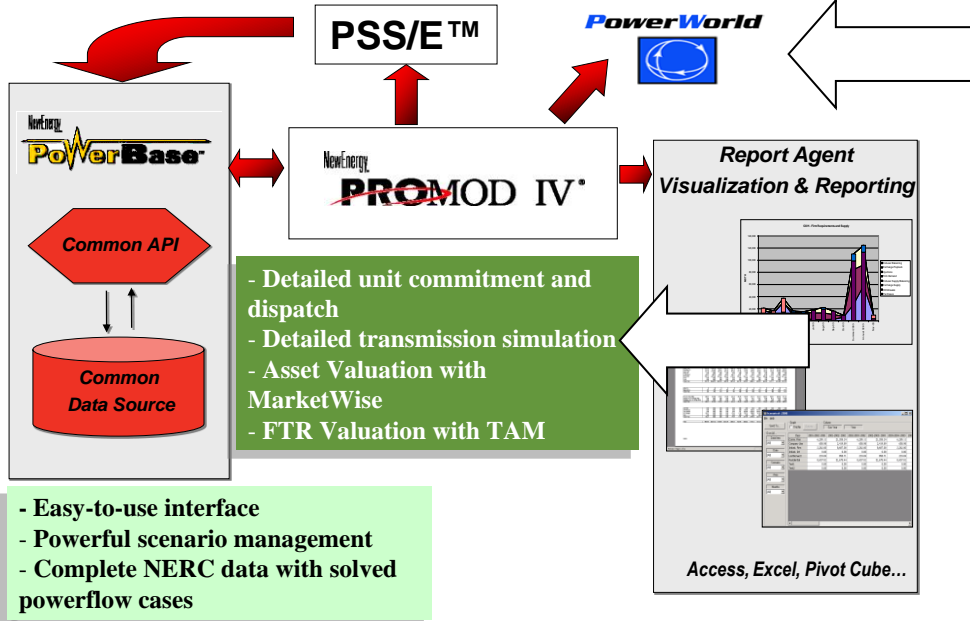
30

How PROMOD Works - PROMOD Structure

MidwestIS◉
Energizing the Heartland

PSS/E™

PowerWorld

NewEnergy
PowerBase

Common API

Common Data Source

NewEnergy
PROMOD IV

Report Agent
Visualization & Reporting

- Detailed unit commitment and dispatch
- Detailed transmission simulation
- Asset Valuation with MarketWise
- FTR Valuation with TAM

Access, Excel, Pivot Cube...

- Easy-to-use interface
- Powerful scenario management
- Complete NERC data with solved powerflow cases

31

How PROMOD Works –
Input and Output of PROMOD

- Generation Data: heat rate, different costs, etc.
- Demand & Energy
- Fuel Forecasts: Gas, Coal, Oil
- Environmental Costs: Sox, Nox, Mercury
- Power Flow Case
- Monitored Flowgates
- Other Information: reserve requirement, market territory, etc.

PROMOD

- Hourly LMP of buses and hubs, include energy, loss and congestion components.
- Hourly unit generation and production cost
- Hourly binding constraints and shadow prices
- Hourly line flows
- Hourly company purchase/sale
- Environmental emissions.
- Fuel consumptions.
- etc.

32

Magnitude of the Challenge

Real System Dimensions –

MTEP 08 PROMOD Cases

- Footprint: East interconnection excluding FRCC
- Generators: ~ 4,700
- Buses: ~ 47,500
- Branches: ~ 60,000
- Monitored Lines: ~ 1,500
- Contingencies: ~ 500
- Run Time: 60-90 Hrs (for one year 8760 hours)

33

Data in PowerBase

- Generation

- Demand & Energy

- Transmission Network Data

- Fuel Forecasts
  - Coal, Uranium, Gas, Coal, Oil

- Environmental Effluent and Costs
  - CO2, Sox, Nox, Mercury

35

PROMOD input files

- PFF file
  - Main input file, includes units, fuels, environmental and transmission data, pool configuration, reserve requirement, run option switches, etc.

- Load data file
  - Hourly load profiles for each company for a selected study period.
  - Based on the 8760 hour load shape and each year's peak load and annual energy for each company defined in PowerBase.

- Gen Outage Library and automatic maintenance schedule
 - Same outage library and maintenance schedule used by all cases
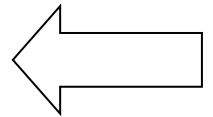
37

PROMOD input files

■ Event files

- Define the monitored line/contingency pairs which are the transmission constraints

- Combine MISO and NERC Book of Flowgates

- Modify existing events or add new events according to member's comments.

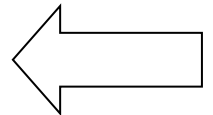- Create new events which have the potential of overflow using PAT tool
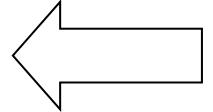
38

PROMOD Assumptions

■ Study Footprint

- East interconnection excluding Florida

- Hourly fixed transactions modeled to include the influence of external areas to the study footprint

  ❖ SETRANS sale to Florida

39

PROMOD Assumptions (Cont')

- Pool Definition

  - a group of companies in which all its generators are dispatched together to meet its loads.

  - Hurdle rates are defined between pools to allow the energy exchange between pools.

  - Hurdle rates are based on the filed transmission through-and-out rates, plus a market inefficiency adder.

 - In current MISO cases, 11 pools are defined: MISO, PJM, TVA, MRO, East Canada, SPP, IMO, MHEB, ISONE,NYISO,SERC
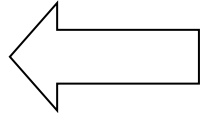
40

PROMOD Assumptions (Cont')

- Loss Calculation

  - Option1: Load is equal to actual load plus loss. Loss and LMP loss component are not calculated.

  - Option 2: Load is equal to actual load plus loss. Loss is not calculated while LMP loss component is calculated using an approximation method – Single Pass Loss Calculation.

  - Option 3: Load is equal to actual load. Loss and LMP loss component are calculated – Multi Pass Loss Calculation. Run time is 4 times of Option 2.
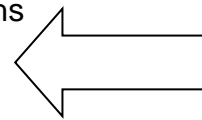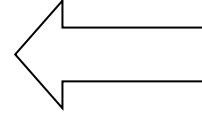
  Option 2 is used in MISO PROMOD cases.
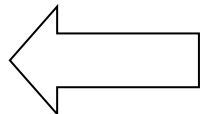
41

PROMOD Assumptions (Cont')

- Wind Units – fixed load modifier transactions

  - Set at a same capacity factor for every hour (~ 33%);

  - Set different capacity factors for different months (15% for summer months, and 20% for winter months);

  - Set hourly profile for each unit to capture geographical diversity.

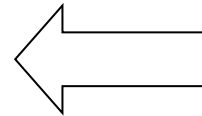- Smelter Loads modeled as transactions

42

PROMOD Output

- LMPs (include the energy, loss and congestion components):

  - ❖ Hourly LMP of selected buses, defined hubs.

  - ❖ Hourly Load Weighted and Gen Weighted LMP of defined zones.

- Constraints:

  - ❖ Hourly shadow price;

  - ❖ Number of hours at Pmax, total shadow price at Pmax;

  - ❖ Number of hours at Pmin, total shadow price at Pmin;

44

PROMOD Output (Cont')

■ Generators:

❖ Hourly generation

❖ Hourly production cost (sum of fuel, variable O&M, environmental cost)

❖ Hourly fuel consumption, BTU consumption

❖ Hours on line, hours of startup, hours at margin, Hours profitable.

❖ Monthly variable O&M cost, fuel cost, emission, and emission cost.
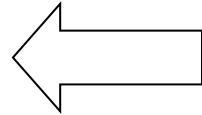
45

PROMOD Output (Cont')

■ Fuel:

❖ Hourly fuel consumption.

■ Power Flow:

❖ Hourly flow for selected lines, interfaces, and DC lines.

❖ Monthly transmission losses (only for marginal loss calculation option)

■ Company:

❖ Hourly purchase/sale.

❖ Hourly dump and emergency energy.

46

Economic Benefit

- To capture the economic benefit of transmission upgrade: run two PROMOD cases, one with transmission upgrade, one without. For each case, calculate (for each region):

  - Load Cost = Load LMP * Load

  - Adjusted Production Cost = Production Cost + Import * Load Weighted LMP (or) - Export *Gen Weighted LMP

The adjustment to each region's production cost is to account for
- imports (evaluated at LWLMP since it is load-dominated region, added);
- exports (evaluated at GWLMP since it is gen-dominated region, subtracted)

- Economic Benefit:

  - Load Cost Saving: Load Cost difference between two cases;

  - Adjusted Production Cost Saving: Adjusted Production Cost difference between two cases

  - RECB II Benefit = sum over all regions (30%* Load Cost Saving + 70%*Adjusted Production Cost Saving)

RECB: Regional expansion criteria & benefits. Today, there is an RECB work grp "the forum for stakeholders to discuss existing or proposed criteria & cost allocation policies for regional and interregional cost shared transmission projects.

Example:
5 Bus Power Network

49

5 Bus Power Network (Original) – PROMOD result

Gen: 150 MW
LMP: 15$/MWH
Prod. Cost: 2,250$

Load: 200 MW
LMP: -15$/MWH
Load Cost: -3,000$

This LMP is neg. because of the binding transmission constraint, if we increase load by 1 MW, it is supplied by 3 MW at bus 1 (+$45) & -2 MW at bus 5 (-$60).

Line is binding

50 MW

Region 2:
Export: 150 MWH
Gen Weighted LMP =30$/MWH
Load Cost = 22,500$
Adjusted Production Cost = 13,500$ - 150MWH*30$/MWH =9,000$

GWLMP=(GenProdCost)/Gen=13500/450=$30/MWh

LWLMP=(Load cost)/Load
=(-3000+4500)/(100+200)
=$5/MWh

Region 1:
Import: 150 MWH
Load Weighted LMP = (-3,000+4,500)/(100+200) =5$/MWH
Load Cost = -3,000 + 4,500 = 1,500$
Adjusted Production Cost = 2,250$+150MWH*5$/MWH =3,000$

{PrC+Imp*LWLMP, for importing region
AdjPrC={
    {PrC-Exp*GWLMP, for exporting region
and since region 2 is exporting region,
    =13,500-150*30=$9000

Load: 100 MW
LMP: 45$/MWH
Load Cost: 4,500$

Load: 300 MW
LMP: 75$/MWH
Load Cost: 22,500$

Gen: 450 MW
LMP: 30$/MWH
Prod. Cost: 13,500$

50

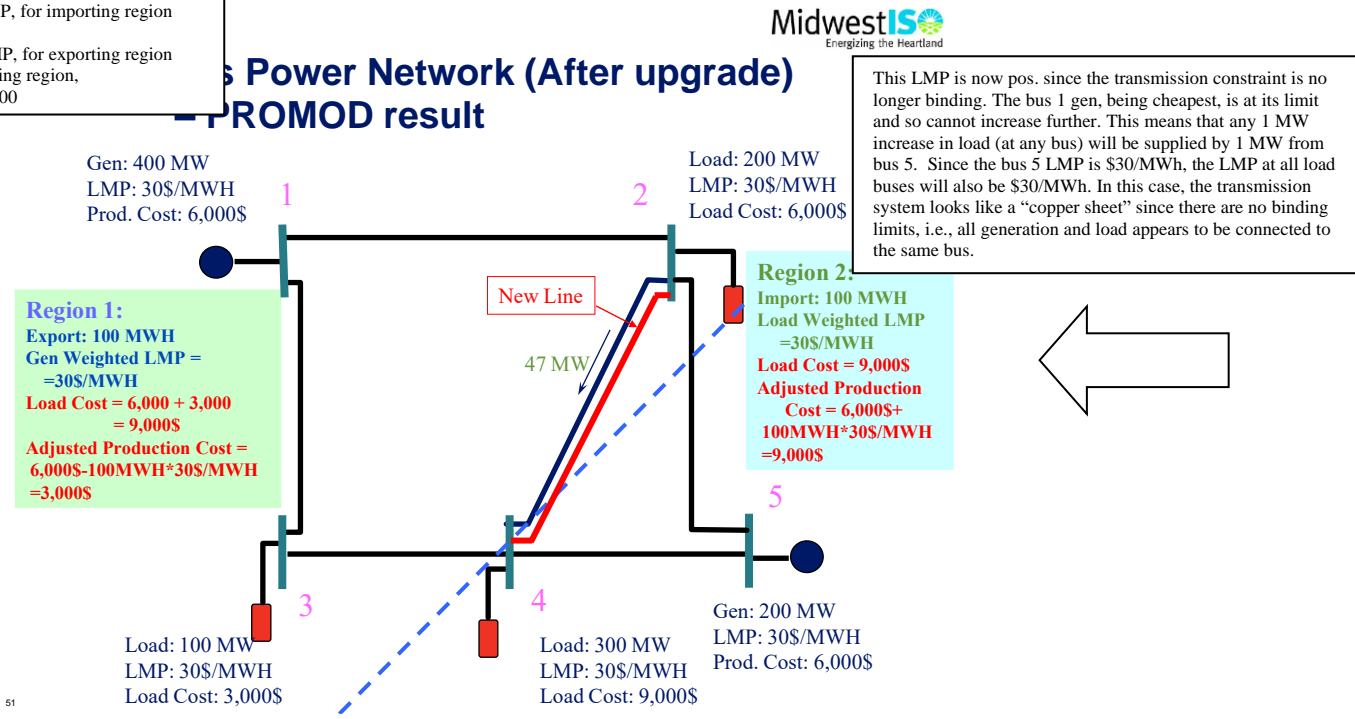{PrC+Imp*LWLMP, for importing region
AdjPrC={
    {PrC-Exp*GWLMP, for exporting region
and since region 1 is importing region,
    =2250+150*5=$9000

Bus Power Network (After upgrade) – PROMOD result

This LMP is now pos. since the transmission constraint is no longer binding. The bus 1 gen, being cheapest, is at its limit and so cannot increase further. This means that any 1 MW increase in load (at any bus) will be supplied by 1 MW from bus 5. Since the bus 5 LMP is $30/MWh, the LMP at all load buses will also be $30/MWh. In this case, the transmission system looks like a "copper sheet" since there are no binding limits, i.e., all generation and load appears to be connected to the same bus.

Gen: 400 MW
LMP: 30$/MWH
Prod. Cost: 6,000$

Load: 200 MW
LMP: 30$/MWH
Load Cost: 6,000$

New Line

47 MW

Region 2:
Import: 100 MWH
Load Weighted LMP =30$/MWH
Load Cost = 9,000$
Adjusted Production Cost = 6,000$+ 100MWH*30$/MWH =9,000$

Region 1:
Export: 100 MWH
Gen Weighted LMP = =30$/MWH
Load Cost = 6,000 + 3,000 = 9,000$
Adjusted Production Cost = 6,000$-100MWH*30$/MWH =3,000$

Load: 100 MW
LMP: 30$/MWH
Load Cost: 3,000$

Load: 300 MW
LMP: 30$/MWH
Load Cost: 9,000$

Gen: 200 MW
LMP: 30$/MWH
Prod. Cost: 6,000$

51

5 Bus Power Network
– New Transmission RECB II Benefit

		Original Case	Case with New Line	Saving

For both regions, the Adj Pr Cost stay the same after the new line is built because, in each region, the adjustment exactly compensates for the change in production cost:
Region 1: prod costs increase but it is exporting➔adjustment is subtracted.
Region 2: prod costs decrease but it is importing➔adjustment is added.

Region 1

| Load Cost | $1,500 | $9,000 | $-7,500 |
| Adjusted Production Cost | $3,000 | $3,000 | $0 |

The load in region 1 is hurt by the new line (the LMP for the largest load has increased significantly).

Region 2

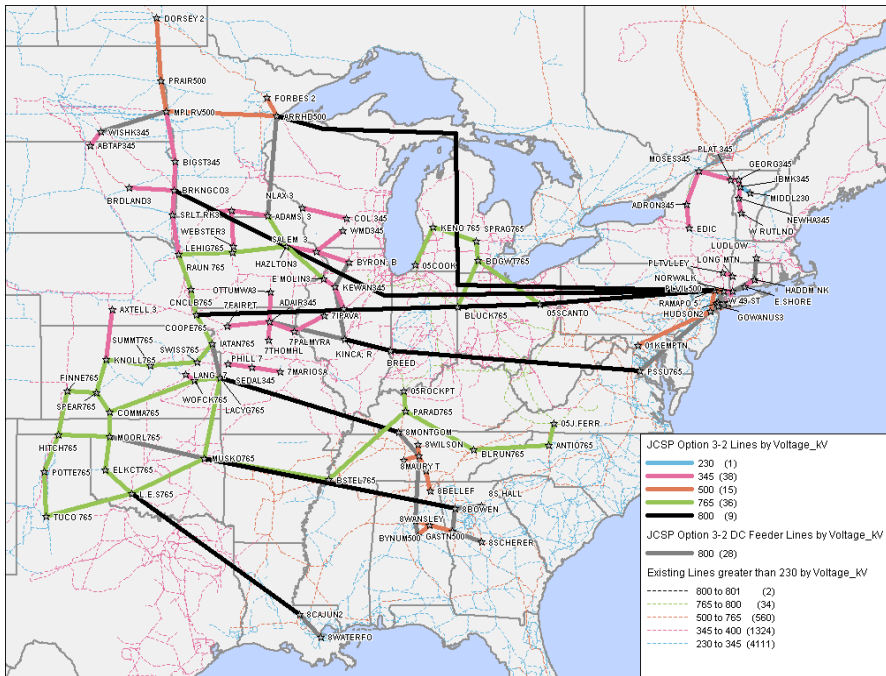| Load Cost | $22,500 | $9,000 | $13,500 |
| Adjusted Production Cost | $9,000 | $9,000 | $0 |

The load in region 2 is helped by the new line (the LMP for its only load has decreased significantly).

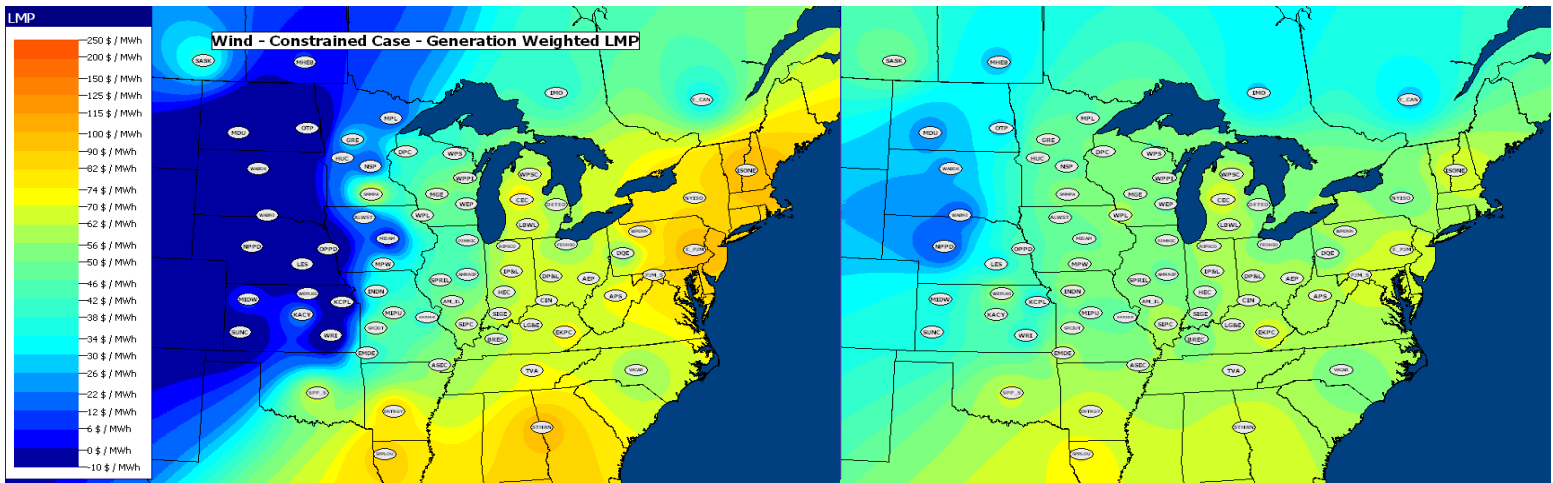**RECB II Benefit = 70% * 0 +30% * (-7,500+13,500) = $1,800**

52

There is more on RECB II Benefit at [10].

Joint Coordinated System Plan Overlay

20% Wind Energy Scenario

Without and With the JCSP

Overlay

Annual Generator Location Marginal Prices

References

[1] R. Sullivan, "Power System Planning," McGraw Hill, 1977.

[2] International Atomic Energy Agency, "Expansion planning for electrical generation systems," Technical Reports Series No 241, 1984.

[3] X. Wang and J. McDonald, "Modern Power System Planning," McGraw-Hill, 1994.

[4] D. Anderson, N. Samaan, T. Nguyen, and M. Kintner-Meyer, "North America Modeling Compendium and Analysis," March, 2016. [Online]. Available: www.energy.gov/sites/prod/files/2017/01/f34/North%20America%20Modeling%20Compendium%20and%20Analysis.pdf.

[5] "Final Report of the GTMax Model Review Panel: Report of a Workshop held August 31-Sept 1, 2011, in Flagstaff, Arizona," Sept 4, 2012. [Online]. Available: http://gcdamp.com/images_gcdamp_com/7/7e/130114_GTMax_Model_Review_Workshop_-_Final_reports_of_the_review_panel_and_workshop_participants.pdf.

[6] A. Gaikwad, S. Agarwal, K. Carden, N. Wintermantel, S. Meliopoulos, and M. Kumbale, "Case Studies on Risk Assessment for Transmission and Other Resource Planning," prepared for Eastern Interconnection States' Planning Council and National Association of Regulatory Utility Commissioners (NARUC), Jan., 2015. [Online]. Available: https://pubs.naruc.org/pub.cfm?id=536DCE1C-2354-D714-5175-E568355752DD.

[7] C. Liang, J. Meng, C. Chen, and Y. Zhou, "A production-cost-simulation-based method optimal planning of the grid interconnection between countries with rich hydro energy," Global Energy Interconnection, Vol. 3, Num 1, Feb 2020 (23-29), DOI: 10.1016/j.gloei.2020.03.002.

[8] S. Wang and M. Baran, "Reliability assessment of power systems with wind power generation," Proceedings of the IEEE Power and Energy Society General Meeting, 2010.

[9] J. Dillon, "System operation and technology value in future energy grids: a methodological framework," Ph. D. dissertation, University College, Dublin, Ireland, 2020.

[10] https://apps.psc.wi.gov/ERF/ERFview/viewdoc.aspx?docid=160002.