

Chapter 2

Maintenance Strategic and Capacity Planning

2.1 Introduction

Planning is one of the major and important functions for effective management. It helps in achieving goals and objectives in the most efficient and effective manner. Planning is usually divided into three levels, depending on the objective and the planning horizon. The levels are as follows:

1. long-range planning (covers a period of five or more years);
2. medium-range planning (covers 1 month to 3 years); and
3. short-range planning (daily, weekly, and monthly plans).

Maintenance strategic planning is a long-range planning and is concerned with the determination of maintenance mission, strategic goals, and objectives. The mission and the strategic goals are derived and aligned with organization's mission and goals. Then, maintenance strategies and programs are developed to achieve a set mission and goals. Maintenance capacity planning is medium- to long-range planning for maintenance involves the determination of the maintenance resources that are needed to perform the maintenance load in order to achieve organizational objectives such as availability, reliability, quality rates and delivery dates.

Maintenance forecasting is an essential activity for planning. Maintenance forecasting comprises the estimation and prediction of the maintenance load. The maintenance load drives the whole maintenance system, and it consists of two major categories. The first category is the scheduled and planned maintenance which is formed of (1) routine and preventive maintenance, (2) scheduled overhauls which involve closure or plant shutdown, (3) corrective maintenance that involves determining the causes of repeated breakdown and substandard performance as a result of design malfunction, and (4) scheduled overhaul, repair, or building of equipment which is not covered under item 2. The second category is emergency or breakdown maintenance load. This category depends primarily on the failure pattern, and it is a major source of uncertainty in the planning process. The sum of the

maintenance load for the two categories is a random variable, and it is the major factor in determining maintenance capacity.

In this chapter, techniques of maintenance load forecasting, strategic planning, and capacity planning are presented. Section 2.2 covers a brief introduction to forecasting. Section 2.3 describes qualitative forecasting techniques. Section 2.4 presents a host of quantitative forecasting models including moving averages, regression analysis, exponential smoothing, and seasonal forecasting. Section 2.5 covers error analysis and measures for testing forecasting models. Section 2.6 outlines the procedure for maintenance load forecasting. Maintenance strategic planning is outlined in Sect. 2.7, and Sect. 2.8 introduces the problem of capacity planning in maintenance. Deterministic techniques for capacity planning are presented in Sect. 2.9. Section 2.10 outlines stochastic techniques for capacity planning. The chapter is summarized in Sect. 2.11.

2.2 Forecasting Preliminaries

Forecasting techniques can be classified into two approaches: qualitative and quantitative. Qualitative forecasting is based on the expert or engineering experience and judgment. Such techniques include historical analogy, surveys, and the Delphi method. Quantitative techniques are based on mathematical models that are derived from the historical data estimates for future trends. These models are either time series-based data such as moving averages and exponential smoothing or structural such as regression models Montgomery and Johnson [7].

A forecasting model is judged by the following criteria: (1) accuracy, (2) simplicity of calculation, (3) data needed for the model and storage requirements, and (4) flexibility. Accuracy is measured by how accurate the model predicts future values and is judged by the difference between the model forecasts and the actual observed values. In general, high accuracy requirements demand a complex relationship and therefore increase the complexity of computation. Flexibility is the ability to adjust to changes in the conditions. In other words, it is a measure of the robustness of the forecasting model. Important considerations in the selection of the forecasting approach are as follows: (1) the purpose of the forecast, (2) the time horizon for the forecast, and (3) the availability of the data for the particular approach. The following are the steps for developing a quantitative forecasting model.

1. Identify the characteristic/item to be forecasted and understand its nature. Define the purpose of the forecast and its time horizon.
2. Screen and validate available data for errors and outliers. Identify the additional data needed and the methodology for collecting it.
3. Use the available data and graphical techniques to hypothesize appropriate models. The model represents a relationship that describes the historical pattern of the data or a relationship between the dependent and independent variables.

4. Use the major part of the data to estimate the parameters of the models. Keep part of it for testing and validating the model. The estimation can be accomplished by an appropriate statistical method such as the least squares or the maximum likelihood.
5. Test and validate the models and select the most appropriate one. Simulation and error analysis are useful tools for testing, validating, and selecting the most appropriate one.
6. Monitor the selected forecasting process and model to detect out-of-control conditions and find opportunities for improving forecasting performance. Improvement can be made by refining parameter estimation or changing the forecasting model. The cycle of the forecasting process is shown in Fig. 2.1

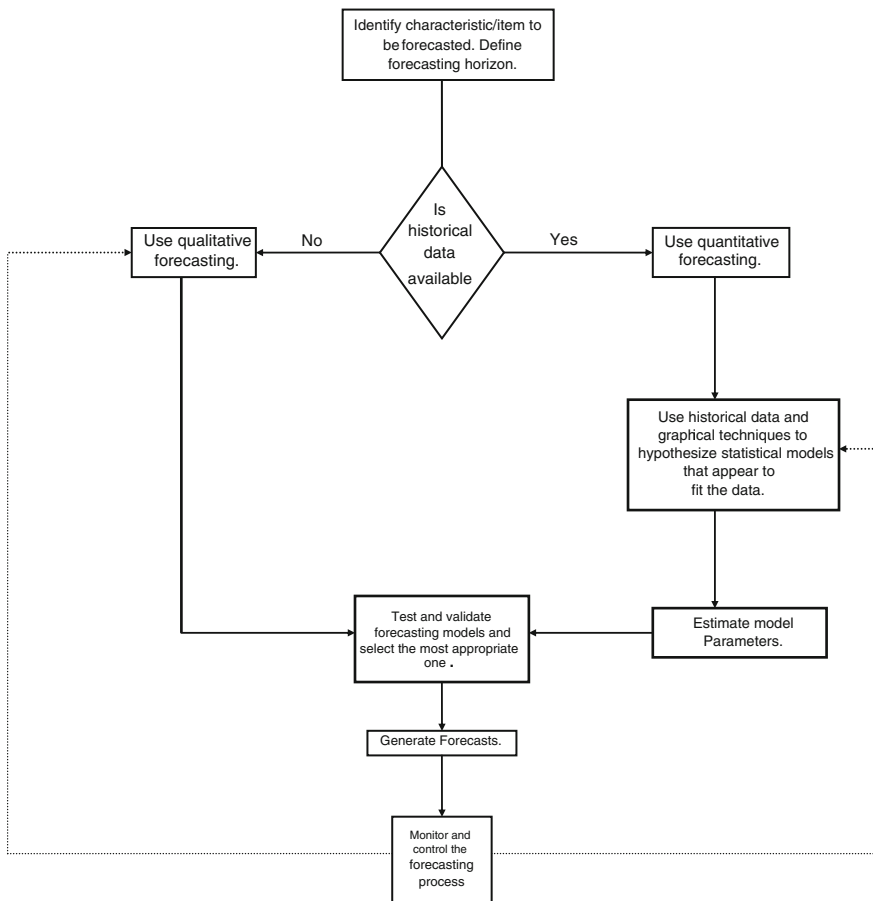


Fig. 2.1 The cycle of the forecasting process

2.3 Qualitative Forecasting Techniques

In the absence of data, the analyst must rely on estimates of experts and their judgment. The role of the analyst in qualitative forecasting is to systematically extract information from the mind of the expert by using structured questionnaires or interviews. He should help the expert or management to quantify their knowledge. Techniques such as cause-and-effect diagrams and the Delphi method can be helpful in identifying relationship among the variables. The analyst should identify which variables influence the forecast and the impact of each one.

After identifying the variables and their impact, the next step is to get an agreement on the magnitude of the variables. Best-case, expected-case, and worst-case scenarios are usually used to estimate the magnitude of the variables. An interactive approach can be used to present arguments to the expert, such as why his estimate differs from the average estimate, and he is asked to revise his/her estimate until a reasonable consensus is reached. When no further reduction in variation about the consensus is possible, the result is used as a forecast.

2.4 Quantitative Forecasting Techniques

In this section, quantitative forecasting techniques are presented. The models presented depend on the availability of the historical data and are usually referred to as time series or structural models. These models either assume future values follow historical trends or that a predictor (independent) variable exists that can provide a model or a functional relationship that predicts the characteristic under study. For example, the age of the equipment can predict the number of maintenance hours required on the equipment. The models presented here include moving averages, regression analysis, exponential smoothing, and seasonal forecasting.

2.4.1 Simple Moving Average

Suppose the characteristic under study is generated by a constant process, plus a random error. An example of this could be the load x_t exerted on an electronic component. Mathematically, this can be represented as

$$x_t = b + \epsilon_t$$

where b is a constant and ϵ_t is a random variable with mean 0 and variance σ_ϵ^2 .

To forecast the future value, we need to estimate the parameter b . Suppose we have the time series observations, $x_1, x_2, \dots, x_n$. If all the observations are assumed to be equally important, i.e., of equal weight, and if we use the least squares method

(see Appendix A, Section A8), then we select a value of b that minimizes the sum of squared error denoted by SS_E .

$$SS_E = \sum_{t=1}^n (x_t - b)^2 \quad (2.1)$$

Equation (2.1) is differentiated with respect to b and equated to zero, and Eq. (2.2) is obtained.

$$\frac{dSS_E}{db} = -2 \sum_{t=1}^n (x_t - b) = 0 \quad (2.2)$$

The estimator of b , $\hat{b}$, is given as

$$\hat{b} = \frac{\sum_{i=1}^n x_i}{n} \quad (2.3)$$

which is just the average of the observations on hand. This method generates the next period's forecast by averaging the actual observations for the last n periods.

Example 1 If the maintenance load in man-hours for the last 6 months is given as

Month	1	2	3	4	5	6
Maintenance load	200	300	200	400	500	600

Find the load forecast for periods 7 and 8 using a 3-month moving average.

The forecasted load for month 7 using $n = 3$ is

$$\hat{x}_7 = \frac{400 + 500 + 600}{3} = 500$$

We did not observe x_7 , so if the load in month 7 is estimated as 500 as calculated above. The forecast for the 8th month is obtained as follows:

$$\hat{x}_8 = \frac{500 + 600 + 500}{3} = 533.33$$

2.4.2 Weighted Moving Average

It is logical to assume that the most recent observations should have more contributions to future forecasts than to more distant observations, especially when the

data are not stable (e.g., there are significant changes). The idea of the weighted moving average is to give each observation a different weight. The forecasting relationship is

$$\bar{x}_{n+1} = \sum_{i=1}^n w_i x_i \quad (2.4)$$

where w_i = weight for the i th actual observation

$$\sum_{i=1}^n w_i = 1 \quad (2.5)$$

The formula given above can be obtained mathematically by minimizing the sum of the weighted square error using the same procedure as in Sect. 5.4.1. The values of w_i can be determined empirically or be estimated based on experience. If the values of w_i are determined based on experience, this method combines qualitative and quantitative forecasting approaches.

Example 2 Assume, for the data used in example 1, that the most recent period should weigh twice as much as the previous months. Find the forecasts for x_7 and x_8 using three-period moving average:

$$\begin{aligned} w_4 + w_5 + w_6 &= 1 \\ w_6 &= 2w_5 = 2w_4 = 2w \\ 4w &= 1 \\ w &= 1/4, \quad w_5 = w_4 = 0.25 \quad w_6 = 0.5 \end{aligned}$$

Therefore,

$$\hat{x}_7 = 0.25(400) + 0.25(500) + 0.5(600) = 525$$

Assuming the load for the seventh month $x_7 = 525$, the forecast for month 8 is given as

$$\hat{x}_8 = 0.25(500) + 0.25(600) + 0.5(525) = 537.5$$

2.4.3 Regression Analysis

If an independent variable exists that can predict a characteristic (dependent variable) and a reasonable correlation exists between the two variables, then a regression model can be used. For example, if the cost of maintenance for this

period, $y(t)$, is a linear function of the number of operational hours in the previous period, $x(t - 1)$, then the model would be

$$y(t) = a + bx(t - 1) + \epsilon_t \quad (2.6)$$

where σ_ϵ^2 .

$y(t)$ desired forecast value for the cost in period t
 $x(t - 1)$ operational hours in period $t - 1$
 a, b parameters to be determined
 ϵ_t a random variable with mean 0 and variance

After estimation, the resulting model that can be used for predicting is of the form:

$$\hat{y}(t) = \hat{a} + \hat{b}x(t - 1)$$

It could be possible that the dependent variable is a function of more than one independent variable. In the case of the maintenance cost, it could be a linear function of operational hours in the previous period, $x(t - 1)$, and the age of the plant, t . Mathematically, this can be expressed as

$$y(t) = a + bx(t - 1) + ct + \epsilon_t \quad (2.7)$$

where a, b , and c are parameters to be determined.

In case of one variable, the parameters a and b can be determined or estimated by finding a good trend line that fits the data points by visual estimation. The parameters a and b are the intercept and the slope of the line, respectively. If more precision is desired, a regression analysis is used. Regression analysis refers to the process of estimating the model parameters using the least squares method. This method fits a line to the observations such that the sum of the square vertical distances from the line is minimized.

The basic equation of a straight line showing a linear trend between an independent variable and a dependent variable $x(t)$ that represents demand for maintenance work is

$$x(t) = a + bt + \epsilon_t \quad (2.8)$$

where a is the intercept and b is the slope that needs to be estimated. The parameter estimation is to determine a and b . The least squares method estimates b and a in terms of $x(t)$ and $t, t = 1, 2, \dots, n$, as follows:

$$\hat{b} = \frac{n \sum_{t=1}^n tx(t) - (\sum_{t=1}^n t)(\sum_{t=1}^n x(t))}{n \sum_{t=1}^n t^2 - (\sum_{t=1}^n t)^2} \quad (2.9)$$

Table 2.1 Data for parameter estimation for the linear regression model

	t	$x(t)$	$tx(t)$	t^2
	1	15	15	1
	2	25	50	4
	3	30	90	9
	4	45	180	16
	5	50	250	25
	6	70	420	36
	7	85	595	49
Summed values	28	320	1600	140

$$\hat{a} = \bar{x} - \hat{b}\bar{t} \quad (2.10)$$

where

$$\bar{x} = \sum_{t=1}^n x(t)/n \quad \text{and} \quad \bar{t} = \sum_{t=1}^n t/n \quad (2.11)$$

The resulting equation that can be used for prediction is

$$x(t) = \hat{a} + \hat{b}t$$

If the model is taken to be constant plus random variation, i.e., $x(t) = a + \epsilon_t$, then the estimate for a is the average as obtained earlier in Sect. 2.4.1.

Regression analysis can easily be generalized to the case of multiple variables and a polynomial relationship between dependent and independent variables.

Example 3 The monthly maintenance load in man-hours is given in the table below. Develop a straight line that best fits the data and can be used to predict future maintenance load.

Month	t	1	2	3	4	5	6	7
Load	$x(t)$	15	25	30	45	50	70	85

The data set and intermediate computation are given in Table 2.1.

The equation of the straight line is

$$x(t) = a + bt$$

The slope of the line is estimated as follows:

$$\hat{b} = \frac{n \sum_{t=1}^7 tx(t) - \left(\sum_{t=1}^7 t \right) \left(\sum_{t=1}^7 x(t) \right)}{n \sum_{t=1}^7 t^2 - \left(\sum_{t=1}^7 t \right)^2} = \frac{7(1600) - (28)(320)}{7(140) - (28)^2} = 11.43$$

$$\hat{a} = \bar{x}(t) - b\bar{t} = \frac{320}{7} - \frac{(11.43)(28)}{7} = -0.005$$

Therefore, $x(t) = -0.005 + 11.43t$.

The above equation can be used for forecasting future load. For example, the load in the ninth month is obtained by substituting 9 in place of t in the equation obtained and is equal to 102.82 man-hours.

2.4.4 Exponential Smoothing

Exponential smoothing is a widely used forecasting method that is simple, efficient, and easy to apply. It assigns weight to observations of previous periods in an inverse proportion to their age. It does this in a very ingenious manner in which only three pieces of data are required to generate the next period's forecast. These are (1) last period's forecast, (2) last period's actual observation, and (3) a smoothing factor, which determines the relative weight given to the recent observation. The basic equation that is the heart of the exponential smoothing is the following:

$$\hat{x}(t) = \alpha x(t-1) + (1-\alpha)\hat{x}(t-1) \quad (2.12)$$

where

- $\hat{x}(t)$ the forecast for period t and all future periods in the case of a constant model
- $x(t-1)$ actual demand at period $t-1$
- $\hat{x}(t-1)$ the forecasted value for $t-1$
- α smoothing constant, $0 < \alpha < 1$.

The exponential smoothing approach can be used to estimate the parameters for a constant model, linear model, and any polynomial functional form. The parameter estimation for the constant and the linear cases is given below:

In the constant case, or zero growth, the process model is given as

$$x(t) = b + \epsilon_t \quad (2.13)$$

where

- b expected demand in any period
- ϵ_t random component having mean 0 and variance σ_ϵ^2 .

At the end of the period $t - 1$, we have the observations $x(1), \dots, x(t - 1)$, from which we need to estimate b and σ_ϵ^2 . This model is referred to as simple exponential smoothing or

$$x(t) = \alpha x(t - 1) + (1 - \alpha) \hat{x}(t - 1) \quad (2.14)$$

where α is determined using experimentation or judgment. A large value of α (closer to 1) indicates a belief that the current observation carries a high weight. In other words, the system has shifted and the most recent observations resemble its behavior. However, small values of α indicate a belief that the past still resembles to a great extent the system.

If the plotted historical data suggest a linear growth over time, the model is considered a linear model. In that case, the process mean changes linearly with time according to the following equation:

$$x(t) = a + bt + \epsilon_t \quad (2.15)$$

where the expected demand at time t is a linear function of time.

$$E(x(t)|t) = a + bt \quad (2.16)$$

It is known that the lag (the amount by which the forecast deviates from the most recent data value) at the most recent data value in the case of the linear model is

$$\text{lag} = \frac{\beta}{\alpha} \quad \text{slope} = \left(\frac{\beta}{\alpha}\right)b \quad (2.17)$$

If we apply the exponential smoothing again (double smoothing) denoted by

$$\hat{x}(t) = \alpha \hat{x}(t - 1) + (1 - \alpha) \hat{\hat{x}}(t - 1) \quad (2.18)$$

$$\text{lag} = [x(t - 1) - \hat{x}(t - 1)] = [\hat{x}(t - 1) - \hat{\hat{x}}(t - 1)] = \frac{\beta}{\alpha} b \quad (2.19)$$

at each period $t - 1$, the values of a and b are updated as follows:

$$\hat{a}(t - 1) = x(t - 1) - \hat{x}(t - 1) + \text{lag} = 2\hat{x}(t - 1) - \hat{\hat{x}}(t - 1) \quad (2.20)$$

$$\hat{b}(t - 1) = \frac{\alpha}{\beta} [\hat{x}(t - 1) - \hat{\hat{x}}(t - 1)] \quad (2.21)$$

Initial conditions to start the process are

$$\hat{b}(1) = \frac{x(t-1) - x(1)}{N-2}$$

$$\hat{a}(1) = x(1)$$

$$\hat{x}(1) = \hat{a}(1) - \hat{b}(1) \frac{\beta}{\alpha}$$

$$\hat{\hat{x}}(1) = \hat{a}(1) - 2\hat{b}(1) \frac{\beta}{\alpha}$$

Example 4 Consider the data given in example 3 for a linear fit by regression. Use exponential smoothing with a linear growth model and $\alpha = 0.2$. Find $\hat{x}(8)$ and $\hat{\hat{x}}(10)$.

$$\alpha = 0.2$$

$$\beta = 1 - \alpha = 0.8$$

Computing Initial Conditions

$$\hat{a}(1) = x(1) = 15$$

$$\hat{b}(1) = \frac{x(N) - x(1)}{N-1} = \frac{85 - 15}{6} = \frac{70}{6} = 11.67$$

$$\hat{x}(1) = 15 - 11.67 \left(\frac{0.8}{0.2} \right) = -31.68$$

$$\hat{\hat{x}}(1) = 15 - 2(11.67) \left(\frac{0.8}{0.2} \right) = -78.36$$

The first and double exponential smoothing are given in Table 2.2.

Estimates for $\hat{a}(7)$ and $\hat{b}(7)$ can be obtained using Eqs. 2.20, 2.21, respectively.

$$\hat{a}(7) = 2(18) - (-10.67) = 36 + 10.67 = 46.67$$

$$\hat{b}(7) = \frac{0.2}{0.8} [18 - (-10.67)] = \frac{1}{4} [18 + 10.67] = 7.17$$

Table 2.2 The first and double exponential smoothing

t	$x(t)$	$\alpha x(t)$	$\beta \hat{x}(t-1)$	$\hat{x}(t)$	$\alpha \hat{x}(t)$	$\beta \hat{\hat{x}}(t-1)$	$\hat{\hat{x}}(t)$
1	15	—	—	(-31.68)	—	—	(-78.36)
2	25	5.0	-25.34	-20.34	-4.07	-62.69	-66.76
3	30	6	-16.27	-10.27	-2.05	-53.41	-55.46
4	45	9	-8.22	0.78	0.16	-44.37	-44.21
5	50	10	0.62	10.62	2.12	-35.37	-33.25
6	70	14	8.50	22.5	4.5	-26.0	-22.10
7	85	17	19	35	7	-17.68	-10.67

The prediction for any period after L units from 7 is given as

$$\begin{aligned}\hat{x}(7 + L) &= \hat{a}(7) + \hat{b}(7)(L) \\ \hat{x}(8) &= 46.67 + 7.17 = 53.84 \\ \hat{x}(10) &= a(7) + b(7)(3) = 68.18\end{aligned}$$

2.4.5 Seasonal Forecasting

In many cases, the process under study might exhibit seasonality characteristics. For example, demand for electricity is high during summer months in the Middle East, or the rate of absenteeism might be higher at the beginning or at the end of the week. Also, the maintenance load might be higher in certain seasons due to weather and operational conditions. The appropriate period to look for seasonality should be driven by the nature of the operation under study. A quick way to check seasonality and growth is by plotting the historical data.

A set of logical steps to be followed when forecasting a characteristic with combined seasonality and growth is as follows:

1. Plot the data and visually determine clear time series characteristics.
2. Determine the growth model and remove the growth component from the data.
One way to remove the growth component from the data is to determine an average period for each cycle and divide each data value by the average.
3. Determine whether a significant seasonality is present in the data as it appears with the growth component removed (degrowthed). The seasonality index can be computed by averaging the degrowthed data over the seasons (periods exhibiting the similar behavior).
4. Deseasonalize the original data and analyze the growth factor. A plot of the deseasonalized data will reveal the form of the growth component. The deseasonalizing is accomplished by dividing each data by the appropriate seasonal index.
5. Fit the data by some appropriate method, least squares regression, exponential smoothing, etc.
6. A forecast for the future consists of a combination of seasonal and growth trends.

The above steps are demonstrated on the following data. The following data show the number of hours lost due to absenteeism and late arrivals at the maintenance department for a period of 4 weeks.

The data in Table 2.3 exhibit seasonality. It is clear that the data values for Mondays and Fridays are high. The data for Tuesdays, Wednesdays, and Thursdays are relatively low. The daily averages for each week exhibit a growth model (7, 9, 11, 13). To remove the growth (step 2 above), the data for each day in the week are

Table 2.3 Seasonal data for number of hours lost with five seasons (days) across the period

Week	Days					
	Monday	Tuesday	Wednesday	Thursday	Friday	Average
1	9	3	2	5	16	7
2	11	5	5	9	15	9
3	13	6	7	10	19	11
4	16	8	7	12	22	13

Table 2.4 Seasonal data for number of hours lost with five seasons (days) across the period

Week	Days					
	Monday	Tuesday	Wednesday	Thursday	Friday	Average
1	1.29	0.43	0.29	0.71	2.29	5.01
2	1.22	0.55	0.56	1.0	1.67	5.01
3	1.18	0.55	0.64	0.91	1.73	5.01
4	1.23	0.62	0.54	0.92	1.69	5.0
Total	4.92	2.16	2.03	3.54	7.38	–
Average (I_d)	1.23	0.54	0.51	0.89	1.85	5.02

divided by the daily average (e.g., the Monday value of week 1 is found by dividing $\frac{9}{7} = 1.20$. This has been performed in Table 2.4. The seasonality index for each day is obtained by averaging the degrowth data over the 4 weeks. Also, Table 2.4 shows the seasonality index for each day in the last row.

The next step is to deseasonalize the data and determine a model for the growth component. One approach for this is to divide each data point by the seasonal index, and then a growth model should be developed for the daily averages. This is shown in Table 2.5.

The daily averages seem to grow linearly. Using linear regression, the parameter for the growth component is given as

$$\hat{x}(t) = 4.15 + 2.36t \quad (2.22)$$

Table 2.5 Deseasonalized data

Week	Days						
	Monday	Tuesday	Wednesday	Thursday	Friday	Daily total	Average
1	7.32	5.56	3.92	5.62	8.65	31.07	6.21
2	8.94	9.26	9.80	10.11	8.11	46.22	9.24
3	10.57	11.11	13.73	11.24	10.27	56.92	11.38
4	13.01	14.81	13.73	13.48	11.89	66.92	13.38

Table 2.6 Weeks 5 and 6 daily forecast with $x(t) = I_d(4.15 + 2.36t)$

Week	Days					
	Monday	Sunday	Wednesday	Thursday	Friday	Daily average
5	19.62	8.61	8.13	14.20	29.51	15.95
6	22.52	9.89	9.34	16.29	33.87	18.31

Table 2.7 Weeks 5 and 6 daily forecast with $x(t) = I_d(5.0 + 2t)$

Week	Days					
	Monday	Tuesday	Wednesday	Thursday	Friday	Daily average
5	18.45	8.1	7.65	13.35	27.75	15
6	20.91	9.18	8.67	15.13	31.45	17

To obtain the daily forecasts for the 5th and 6th week, use the equation above to forecast the weekly average. Then, multiply the weekly average by the daily seasonal index. For example, the forecast for the 5th week average is

$$x(5) = 4.15 + 2.36(5) = 15.95$$

To obtain the forecast for Monday in the 5th week, this value is multiplied by 1.23, the seasonal index for Monday I_m . It is $(1.23)(15.95) = 19.62$. Table 2.6 presents the daily forecast for the 5th and 6th week.

An alternative approach to estimate the growth component is to total or average the data in each cycle. This approach would lead to the daily average for each week and is shown in Table 2.3. These data show a perfect linear trend. Regression analysis for the four daily averages would yield the model

$$\hat{x}(t) = 5.0 + 2t \quad (2.23)$$

Equation (2.23) predicts the daily averages for the 5th and 6th week as 15 and 17. The daily forecast is obtained by multiplying the daily average for the week by the daily seasonality index I_d . The daily forecasts are shown in Table 2.7.

The forecasts given in Tables 2.6 and 2.7 are both logical and sound. The logical questions are which one is better and which one should the decision maker adopt for planning purposes. Error analysis provides an approach for making the selection.

2.5 Error Analysis

Forecast error analysis provides a valid approach for checking the effectiveness of a forecasting model. It also provides a sound methodology for evaluating and selecting from several forecasting models that are available for a particular situation.

The forecast error at period t is the difference between the actual data value $x(t)$ and the forecasted value for it

$$e(t) = x(t) - \hat{x}(t) \quad (2.24)$$

The sum of the errors

$$\sum_{t=1}^N e(t) = \sum_{t=1}^N [x(t) - \hat{x}(t)] \quad (2.25)$$

is not a valid measure of effectiveness of a forecasting model, but it is a measure of bias. The sum of the errors should approach zero if the model is fitted using the least squares method. The sum of the errors has the problem that large positive errors, $e(t)$, can offset large negative errors. To eliminate this problem, we take either absolute errors or squared errors. The following error measures are commonly used for error analysis and evaluation of forecasting models.

1. Mean absolute deviation (MAD)

$$\text{MAD} = \frac{\sum_{t=1}^N |x(t) - \hat{x}(t)|}{N} \quad (2.26)$$

2. Mean-squared error (MSE)

$$\text{MSE} = \frac{\sum_{t=1}^N (x(t) - \hat{x}(t))^2}{N} \quad (2.27)$$

3. Mean absolute percent error (MAPE):

$$\text{MAPE} = \frac{100}{N} \sum_{t=1}^N \left[\left| \frac{x(t) - \hat{x}(t)}{x(t)} \right| \right] \quad (2.28)$$

4. Mean-squared percent error (MSPE)

$$\text{MSPE} = \frac{100}{N} \sum_{t=1}^N \left[\left(\frac{x(t) - \hat{x}(t)}{x(t)} \right)^2 \right] \quad (2.29)$$

One of the above measures can be calculated for all the available models, and the one with the minimum value is selected. This approach can be applied easily for more on some of the models presented in Sect. 2.4 see [3, 7].

2.6 Forecasting Maintenance Work

Prior to performing capacity planning or designing a new maintenance organization, it is essential to have some forecast of the expected maintenance load. The load comprises the following:

1. *Emergency maintenance workload.* This can be forecasted using actual historical workloads and the appropriate techniques of forecasting and/or management experience. This component of the load is random and can be minimized by having a well-designed planned maintenance.
2. *Preventive maintenance workload.* This can be forecasted using actual historical records coupled with newly developed preventive maintenance programs. This should include routine inspection and lubrications.
3. *Deferred corrective maintenance.* This can be forecasted based on historical records and future plans.
4. A forecast for overhaul/removed items and fabrication. This can be estimated from historical records coupled with future plans for improvements.
5. *Shutdown, turnarounds, and design modifications.* This can be forecasted from actual historical records and the future maintenance schedule.

The forecasting of the maintenance load for a new plant is more difficult and must rely on similar plants' experience, benchmarking, management experience, and manufacturers' information (Table 2.8).

Once a plant is in operation, errors in forecasting and job standards may lead to a backlog. An alternative approach is to do forecasting by examining the maintenance backlog. The usefulness of calculating a backlog for planned work is seen when load forecasting is done for the coming week's or month's work. Table 2.8 illustrates this.

The table shows a year-to-date average emergency work level in mechanical trades of 400 h and in electrical, 80 h. This week PM requirements are 1200 h for mechanical and 600 h for electrical. There are an average number of hours required for the minor unplanned work that is done with non-emergency, yet high priority, work orders. Some work was begun last week but was not finished at week's end; it will be completed this week. The total planned priority 3 work in backlog awaiting scheduling—that is work that has been planned and the materials required are

Table 2.8 Weekly backlog report in hours

	Mechanical	Electrical
YTD average "emergency"	400	80
This week's PM level	1200	600
YTD average minor unplanned	600	400
This week's carryover work	800	40
Total planned priority 3 work backlog	12,000	900
Total shutdown backlog	10,000	3000

available—is 12,000 h for the mechanical and 900 h for electrical. Finally, there is a total of 13,000 h of mechanical and electrical work awaiting the next plant shutdown.

This represents the total backlog that is relevant for scheduling the maintenance work for the upcoming week.

2.7 Maintenance Strategic Planning

This section addresses maintenance planning at the strategic level which falls under the long-range planning. In the past, maintenance is not considered as a strategic unit within the organization, but lately this has changed and researchers and maintenance decision makers have brought forward the strategic role of maintenance. Maintenance strategic planning is the process of ensuring the alliance of maintenance mission, goals, objectives, and programs with the organization's mission and objectives. Recent work of Tsang [9], Murthy et al. [8], and Al-Turki [1] has addressed and identified several important issues that are essential in deciding maintenance strategic plan. These issues include maintenance mode of delivery, organization, methodology, and support. The following ten steps are suggested to put to work the frameworks suggested in the recent literature:

- Top management develops organization mission, strategic goals, and objectives.
- Maintenance management (MM) reviews the organization mission and objectives and identifies the role of maintenance in achieving them.
- MM analyzes the current internal and external situations related to the maintenance function using the methodology for determining strengths, weaknesses, opportunities, and threats known as SWOT analysis. SWOT leads to the identification of the strategic issues that need to be addressed. The strategic issues are expected to include mode of delivery, maintenance organization, methodology, manning, and training.
- MM formulates the identified role as a mission statement for maintenance and develops strategic goals and objectives to address the strategic issues.
- Align and prioritize maintenance strategic goals, objectives, and programs with maintenance mission.
- MM checks the alignment of maintenance strategic goals and objectives with organization set mission and strategic goals and objectives.
- MM develops maintenance programs to achieve set strategic goals and objectives.
- Develop a set of quantitative performance measures for the maintenance strategic goals and objectives.
- Evaluate periodically the progress toward achieving the goals and the objectives using the performance measures and identify the gap between the actual and the desired situation.
- Identify the root causes of the gap and implement effective corrective actions.

2.7.1 Maintenance Strategic Issues

In this section, several strategic issues in maintenance that are identified in the literature are discussed. The first issue is the delivery mode. This issue deals with whether maintenance is outsourced, conducted in-house or a combination of both. Each option has its pros and cons. The selection of the right option should be made in light of the set maintenance strategic goals and objectives at the same time to minimize risks and threats to maintenance systems' efficiency and effectiveness. The second maintenance strategic issue is maintenance organization as identified by Tsang [10]. Several options are available in this issue. The options include functional, process, or network. Each of the previous organization options can be crossed with centralized, decentralized, or a cascade option. The third issue is the work structure and control. Many options are available for the work structure depending on the reporting and supervision structure. The fourth strategic maintenance issue is the selection of the support system that includes information system, training, and performance management and reward system. Each element has to be carefully selected to support the overall objective of the organization.

2.8 Maintenance Capacity Planning

Maintenance capacity planning determines the optimal level of resources (crafts, skills, spares, inventory equipment, and tools) to meet the forecasted maintenance load that consists of future load forecast plus the maintenance backlog. An essential element of capacity planning is the determination of the skills of craftsmen, the exact number of various types of craftsmen, the healthy level of backlog, overtime capacity, and contract maintenance. The optimal allocation of the maintenance resources to meet a random and varying workload is a complex and challenging problem. Capacity planning techniques play very important roles in handling this complex problem. The steps involved in capacity planning can be summarized as follows:

1. Determine the total maintenance load.
2. Estimate the required spares and material to meet the load.
3. Determine equipment and tools that are necessary for all types of maintenance work.
4. Determine the skills and the number of crafts from each skill. A special attention should be given to multi-skill crafts.
5. Sometimes in highly computerized equipment, there may be a need to provide a certain specialty that may require special plans.

In maintenance capacity planning, a major issue is to determine the optimal mix of skills of crafts from the available sources to the organization. The usual available sources are regular and overtime in-house crews and contract maintenance. The best mix using these sources is determined using cost and availability measures.

Multi-skilling provides capacity planners with options and alternatives that enrich and generate a wide range of options for determining the optimal capacity. Crafts with multi-skills can be utilized in more than one type of maintenance work. This usually improves manpower utilization.

Techniques for capacity planning can be divided into two major categories: deterministic and stochastic. The deterministic approach assumes that the forecasted maintenance load, standard times, and other random variables are fixed constants. A plan that minimizes cost or maximizes availability depending on the organization objectives that may include reliability, availability, and cost is then determined. In this chapter, two deterministic techniques will be presented, namely:

1. Heuristic tableau method, and
2. linear programming.

The stochastic approach models the maintenance load, standard times, job arrival times, and other variables as random variables with certain probability distributions and uses standard statistical techniques to identify these distributions. Then, a stochastic model is utilized to determine the optimal capacity that meets the maintenance load. Two stochastic techniques will be presented in this chapter, namely:

1. Queuing models and
2. stochastic discrete event simulation

Next, a brief description for these techniques is provided.

2.9 Deterministic Approaches for Capacity Planning

In Sects. 2.9.1 and 2.9.2, two deterministic techniques for capacity planning are presented. These are the heuristic tableau method and linear programming.

2.9.1 Heuristic Tableau Method

In maintenance capacity planning, the heuristic tableau method derives intuitively appealing plans to determine a feasible craft mix, based on sound principles and guidelines. The tableau is used to evaluate the cost of each alternative, and the plan with minimum cost is selected. Sound principles and guidelines include, providing sufficient in-house crafts for high priority work, a reasonable ratio of overtime work to regular time work and a fixed level of a healthy backlog. The following discussion presents the reasoning that may be adopted for determining the required crafts and skills to meet the forecasted mechanical maintenance load using this approach. The same approach can be repeated for other types of maintenance load such as electrical or instrument. As an example, the mechanical workload is first

classified into two grades. The skill the work requires and its impact on the facility determines the grade. In this example, it is classified as grade one and grade two mechanical workload, although other classifications based on skill and priority are possible. Prior to presenting the plan in Table 2.9, the following notation is defined as follows:

FM_t	Total forecasted mechanical load
B_{t-1}	Mechanical workload backlog from period $t - 1$
TM_t	Total mechanical workload, for period t . $TM_t = FM_t + B_{t-1}$
$TM_{i,t}$	Total mechanical workload of grade i in period t , $i = 1, 2$
$RM_{i,t}$	Regular in-house capacity for mechanical workload of grade i in period t
$OM_{i,t}$	Overtime capacity for mechanical workload of grade i in period t
$CM_{i,t}$	Contract capacity for mechanical workload of grade i in period t

The plan is derived based on the following commonsense principles and guidelines:

1. All priority work is met by regular in-house crafts as much as possible.
2. If it is not possible to satisfy priority one work by regular in-house crafts use overtime.
3. No backlog is allowed for grade 1 work.
4. The manning level must be determined based on the average maintenance load plus a healthy backlog from grade 2 work.
5. The priority two work is met with overtime or contract maintenance.
6. The overtime capacity is at most 25 % of the regular in-house capacity.
7. In the example, the maximum for the backlog is 100 man-hours. If the backlog exceeds this limit, subcontracting is utilized, and it is assumed that subcontracting can provide as much capacity as needed.

If the periods in Table 2.9 are taken to be weeks and the guidelines (1–6) are taken into consideration, a flexible plan for meeting the maintenance load from the available sources is shown in Table 2.9. Column two contains the forecasted mechanical maintenance load, and column three has the expected backlog from the previous period. Column four contains the total load. Columns five and six present the two grades of maintenance work. The sum of these columns gives column four. Columns seven and eight present the amount of work that will be met by regular in-house maintenance crafts. Columns nine and ten contain the work that is expected to be performed by overtime in-house maintenance. Columns eleven and twelve present the amount of contract maintenance for both grades of mechanical work. The hours in all the columns are standard hours.

It should be noted that this is a target plan; the realized workload might be different from the forecasted load, and an adjustment to the plan at the execution stage might be needed. Different plans can be generated based on the same guidelines, and the one with minimum cost is selected. The required number of employees can be based on the number of standard hours in the regular in-house column, after making an allowance for the expected productivity of the trades.

Table 2.9 Sample data for heuristic tableau capacity planning

Col. #	2	3	4	5	6	7	8	9	10	11	12
<i>Period t</i>	FM_t	B_{t-1}	TM_t	$TM_{1,t}$	$TM_{2,t}$	$RM_{1,t}$	$RM_{2,t}$	$OM_{1,t}$	$OM_{2,t}$	$CM_{1,t}$	$CM_{2,t}$
1	150	100	250	150	100	150	10	-	40	-	-
2	200	50	250	150	100	150	10	-	40	-	-
3	250	50	300	150	150	150	10	-	40	-	-
4	180	100	280	180	100	160	-	20	20	-	-
5	200	80	280	200	80	160	-	40	-	-	-
6	150	80	230	120	110	120	40	-	40	-	-
7	200	30	230	150	80	150	10	-	40	-	-
8	250	30	280	150	130	150	10	-	40	-	-
9	200	80	280	150	130	150	10	-	40	-	-
10	300	80	380	200	180	160	-	40	-	10	110
11	250	70	320	210	110	160	-	40	-	10	-
12	200	60	260	200	60	160	-	40	-	-	-

The above heuristic tableau approach can be formalized by using a tableau similar to the one used in production planning with some modifications. In this tableau, the sources of manpower supply are shown on the left side of the tableau and the capacity of each source on the right-hand side on the same row. The maintenance load for each period is shown in the columns of the tableau. The cost of meeting the load from a particular source is shown in the corner of each cell in the table.

In production planning products, from the current period can be kept in inventory to satisfy future periods demand, however, this is not the case in maintenance planning. In maintenance, unfinished work is backlogged and performed in future periods at an added cost to the system. It is also possible to divide the maintenance load in each period by skill or priority, and the same method can be applied. Table 2.10 shows a tableau for a three-period maintenance plan for one kind of maintenance work which is the mechanical workload. The same table can be replicated for other types of work. The notations on the tableau are as follows:

- C_r hourly cost of mechanical trade on regular time
- C_o hourly cost of mechanical trade on overtime
- C_s hourly cost of subcontracting
- B_t backlog in man-hours at the beginning of period t
- CR_t capacity of in-house regular time in period t
- CO_t capacity of in-house overtime in period t
- CS_t capacity of subcontracting in period t
- FM_t forecasted maintenance load in period t

Table 2.10 Data for capacity allocation problem

PERIODS	Period Sources	1	2	3	Capacity
1	Regular Time	C_r	∞	∞	CR_1
	Overtime	C_o	∞	∞	CO_1
	Subcontract	C_s	∞	∞	CS_1
2	Regular Time	$C_r + \pi$	C_r	∞	CR_2
	Overtime	$C_o + \pi$	C_o	∞	CO_2
	Subcontract	$C_s + \pi$	C_s	∞	CS_2
3	Regular Time	$C_r + 2\pi$	$C_r + \pi$	C_r	CR_3
	Overtime	$C_o + 2\pi$	$C_o + \pi$	C_o	CO_3
	Subcontract	$C_s + 2\pi$	$C_s + \pi$	C_s	CS_3
Maintenance load		M_1	M_2	M_3	

If a work is backlogged for r period and performed at period $r + 1$ with regular in-house manpower, it will have a cost of $C_r + r\pi$ per hour. Table 2.10 shows the data needed for a three-period planning horizon. The table shows the costs, capacities, and maintenance load. The symbol ∞ in the cost cell means that a work cannot be done in this period. For example, work that came in period 2 cannot be done in period 1.

A simple least cost heuristic method can be used to compute the allocation of the maintenance load to different sources of manpower supply. The method starts with the least cost cell and satisfies the load as much as possible and then moves to the next least cost until all the load is met. It is highly likely to attain near-optimal solutions with the least cost method. Next, an example is given to demonstrate the tableau approach.

Assume that we have three time periods with maintenance loads 400, 300, and 500, respectively. The in-house regular capacity is 200, 350, and 300 for periods 1, 2, and 3 respectively. The overtime is at most 25 % of the in-house capacity. Subcontracting is abundant, and technically, there is no limit on this source. The cost of performing one in-house man-hour is taken to be 1 unit, overtime man-hour costs 50 % more than this regular time, i.e., 1.5 units, and the subcontracting costs 2 units. Backlogging of one man-hour costs $\pi = 0.3$. The capacity for subcontracting can be taken a large number in this example, and for purpose of demonstration, it is taken to be 500 man-hours. The data and the solution for this situation are shown in Table 2.11.

Table 2.11 Data and solution of the example

Periods	Sources	Period			Capacity
		1	2	3	
1	Regular Time	1 200	8	∞	200
	Overtime	1.5 40	∞	∞	40
	Subcontract	2 30	∞	∞	500
2	Regular Time	1.3 50	1 300	∞	350
	Overtime	1.8 80	1.5	∞	80
	Subcontract	2.3	2	∞	500
3	Regular Time	1.6	1.3	1 300	300
	Overtime	2.1	1.8	1.5 75	75
	Subcontract	2.6	2.3	2 125	500
Maintenance load		400	300	500	

Applying the least cost heuristic, the plan in Table 2.11 is obtained. The plan calls for meeting the maintenance load by 200 h of regular in-house, 40 h of overtime, and 30 h of subcontracting in the first period. In addition to 50 h of regular time, and 80 h of overtime in the second period. The number of hours backlogged from the first period to the second period is 130. The second-period load is met by 300 h of regular in-house maintenance in the same period. The third-period load is met by 300 h of regular in-house, 75 h of overtime, and 125 h of subcontracting. The total cost (TC) of the plan is sum of the costs in each cell, which is obtained by multiplying the hours in the cell by the cost of the hour.

$$\begin{aligned} \text{TC} &= 1 * 2000 + 1.5 * 40 + 2 * 30 + 1.3 * 50 \\ &\quad + 1.8 * 80 + 1 * 300 + 1 * 300 + 1.5 * 75 + 2 * 125 = 1491.5 \text{ units.} \end{aligned}$$

2.9.2 Linear and Integer Programming for Maintenance Capacity Planning

Linear programming is a mathematical model that optimizes a linear function subject to linear inequalities. The linear programming model determines the optimal values of decision variables that optimize a given objective such as minimizing cost or maximizing profit. Decision variables are elements under the control of the decision maker, and their values determine the solution of the model. The objective function in linear programming is the criteria with which feasible solutions are evaluated. A feasible solution is a solution that satisfies all the constraints in the system. A constraint is a condition on the system that must be satisfied.

In the case of maintenance capacity planning (MCP), the decision variables could be the number of hours from different skills and trades made available for maintenance capacity through regular in-house, overtime, or contract maintenance. The objective could be to maximize resource utilization or minimize total cost. An example of a constraint is the ratio of overtime hours to regular in-house hours should not exceed certain percentage. In this section, linear programming is introduced via a simple hypothetical example.

Suppose it is forecasted that the mechanical maintenance load is 100 man-hours divided into two grades: 60 h of grade 1 and 40 h of grade 2. Grade classification is based on the skill the work requires and the impact the work has on the facility. The capacity in man-hours for this work can be provided from two skills of workers, i.e., skill 1 and skill 2. Both skills of workers can perform the two types of mechanical work but with different productivity. The productivity of skill one worker is 0.75 and 0.8 for grade 1 and grade 2 mechanical works, respectively. The productivity of skill two worker is 0.5 for grade 1 and 0.7 for grade 2 works. The cost of one man-hour of skill 1 worker is 30 dollars, and the cost of one man-hour of skill 2 worker is 20 dollars. The objective is to determine the number of

man-hours from each skill level workers assigned to perform the maintenance load (work) at a minimum cost.

Prior to stating the model constraints and objective, let x_{ij} be the number of man-hours of skill i assigned to perform the mechanical work of grade j , $i = 1, 2, j = 1, 2$. The first constraint is to meet the load for grade 1 mechanical work. This is stated mathematically as

$$0.75x_{11} + 0.5x_{21} \geq 60 \quad (2.30)$$

The second constraint is to meet the grade 2 mechanical work, mathematically stated as

$$0.8x_{12} + 0.7x_{22} \geq 40 \quad (2.31)$$

The non-negativity restrictions state that the man-hours obtained from each skill of workers should be nonnegative, mathematically stated as

$$x_{11}, x_{12}, x_{21}, x_{22} \geq 0$$

The objective function is to minimize the sum of the cost of meeting the required load from the two skill levels. The linear programming model for determining the capacity for the mechanical work in the example above is

$$\text{Min } 30(x_{11} + x_{12}) + 20(x_{21} + x_{22}) \quad (2.32)$$

Subject to

$$0.75x_{11} + 0.5x_{21} \geq 60$$

$$0.8x_{12} + 0.7x_{22} \geq 40$$

$$x_{11}, x_{12}, x_{21}, x_{22} \geq 0$$

Solving the above model will determine the allocation of man-hours from different skills to perform the two grades of mechanical work at minimum cost. The solution of the linear program could be fractional values. If we want to obtain the solution in whole hours, an additional restriction must be imposed. That restriction requires the variables to be integers. Adding this restriction transforms the linear program to an integer program. In many real situations rounding the linear programming, solution to the nearest hour would be a good solution.

The above formulation can be employed for maintenance capacity planning (MCP) with minor generalization. The formulation will be given for planning the capacity for the mechanical work. The objective is to determine the number of hours from different skill levels from the mechanical trade made available from different sources to perform different grades of mechanical work. The available sources include regular in-house trades, overtime, and contract maintenance. The same model can be used for determining the number of hours needed from other

trades such as electrical, instrument, and civil. The following notations are necessary for stating the linear programming model: of a mechanical trade of skill i , from source j when performing mechanical work of grade k

x_{ijkt}	Number of man-hours from the mechanical trade of skill i ($i = 1, 2, \dots, I$) from source j ($j = 1, 2, \dots, J$) made available to perform the mechanical work of grade k ($k = 1, 2, \dots, K$) in period t , $t = 1, 2, \dots, T$
C_{ij}	Hourly cost of a mechanical trade of skill i from source j
P_{ijk}	Productivity
F_{kt}	Forecasted mechanical load of grade k in period t
B_{kt}	Backlog of work of grade k work in period t
UB_k	Upper limit for a healthy backlog for grade k work
LB_k	Lower limit for a healthy backlog for grade k work
U_{ijt}	Upper limit on the availability of skill i mechanical trade from source j in period t
r_{kt}	Cost of backlogging one man-hour of grade k work in period t

The linear programming model for determining the required number of mechanical man-hours of different skills, and sources made available to perform all grades of the mechanical work, consists of an objective function and a set of constraints. The objective function is to minimize manpower and backlog cost. The constraints include the work balance constraints, a reasonable ratio of in-house man-hours and overtime to limits on manpower availability. The model is stated as follows:

$$\text{Min } \sum_i \sum_j C_{ij} \left(\sum_k \sum_t x_{ijkt} \right) + \sum_k \sum_t r_{kt} B_{kt} \quad (2.33)$$

Subject to:

Work balance constraints,

$$\sum_i \sum_j P_{ijk} x_{ijkt} + B_{kt} = F_{kt} + B_{k,t-1} \quad (2.34)$$

Limit on overtime man-hours in terms of in-house man-hours (taken as 25 %)

$$x_{i2kt} - 0.25x_{i1kt} \leq 0 \quad (2.35)$$

Limits on man-hours availability,

$$x_{ijkt} \leq U_{ijt} \quad (2.36)$$

Lower and upper limits on backlog on different work grades,

$$LB_k \leq B_k \leq UB_k$$

$i = 1, 2, \dots, I$ (number of skills), $j = 1, 2, \dots, J$ (number of sources), $k = 1, 2, \dots, k$ (number of grades), $t = 1 \dots T$ (time period in the planning horizon).

The above model can be solved by any linear programming algorithm such as the simplex-based code in the package Linear Interactive and Discrete Optimizer (LINDO) or the one in the package International Mathematical Software Library (IMSL) or the Optimization Software Library (OSL).

The output of the linear programming model will determine the optimal number of man-hours of different skill levels from a specific trade (mechanical) made available to perform the maintenance work. Based on the number of man-hours from source 1 (in-house), the manning level will be determined.

The linear programming formulation assumed that the parameters of MCP problem are fixed constants, which in reality is not true. In order to examine the sensitivity of the solution obtained, the standard linear programming sensitivity analysis must be conducted. Sensitivity analysis addresses issues, such as what will happen to the solution if the demand changed or a new cheaper resource became available. Also, the complete linear programming solution that includes the dual variables provides information about which resource to increase its level if the demand for maintenance work increased. The interested reader is referred to Taha [9] for more detail on the subject of dual variables and sensitivity analysis.

If we require the crafts and skill levels to be specified in terms of employees to be hired in a maintenance department, we either round up the linear programming solution to the nearest number of employees or reformulate the problem as an integer programming model. This is done as follows:

Let employee of skill i , from source j in period t

n_{ijkt}	Number of employees from the mechanical trade of skill i , from source j made available to perform the mechanical load of grade k in period t
$NR \times S_j$	Number of hours worked per period for an employee from source j (S_j is a constant, and it is usually taken to be 1 if an employee is full time in-house, and between 0 and 1 if an employee is part time or overtime. NH is the number of hours worked by a regular in-house employee per period)
U_{ijt}	Upper bound on the availability
C_{ijt}	Salary of an employee with skill i , from source j in period t

Assuming that the periods are in months and $NH = 160$, the integer programming model for determining the number of employees from the mechanical trade of different skills to perform the mechanical load is stated as follows:

$$\text{Min } \sum_i \sum_j \sum_k \sum_t C_{ijt} n_{ijkt} + \sum_k \sum_t r_{kt} B_{kt} \quad (2.37)$$

Subject to:

$$\sum_i \sum_j (160S_j) P_{ijk} n_{ijkt} + B_{Rt} = F_{k,t} + B_{k,t-1} \quad (2.38)$$

$$n_{i2k} - 0.25n_{i1k} \leq 0 \quad (2.39)$$

$$n_{ijkt} \leq U_{ijt} \quad (2.40)$$

$$LB_R \leq B_{k,t} \leq UB_k \quad (2.41)$$

$$n_{ijk}, \text{ integer}$$

The packages mentioned in this section have the capabilities of solving the integer programming model.

2.10 Stochastic Techniques for Capacity Planning

Queuing models and stochastic simulation are two important techniques for capacity planning. Queuing models address the situation where customers arrive at a service facility, perhaps wait in a queue, and then are served by a server and thus leave the facility. In maintenance capacity planning, the customers may take the form of a maintenance job arriving at the maintenance planning and scheduling unit, that job is then planned and routed to a workshop for repair (which represents the service facility). The results from queuing theory allow us to evaluate the performance of such systems under different configurations. Queuing theory has been used to determine maintenance staffing. However, when queuing models are used for capacity planning, the measures of performance are obtained for the system under steady-state conditions which do not represent the transient system behavior which is more close to day-to-day operations.

Stochastic simulation offers a viable alternative when the decision maker is interested in a transient situation or when the system under consideration is complex (which is usually the case in maintenance systems). In stochastic simulation, the maintenance system is represented on the computer and well-designed experiments (scenarios) are used for system performance evaluation. The experiment described for capacity planning of a maintenance system is an allocation of maintenance crews and skills under certain maintenance policy and procedures. The performance measures can be cost, utilization of resources, and availability of critical and major equipment.

In this section, these two stochastic approaches for capacity planning are presented.

2.10.1 Queuing Models

A queuing model can be described as follows: If the customers arrive at a facility, they join a working queue (line). A server chooses a customer according to a certain discipline from the queue and serves him. Upon service completion, the customer leaves the service facility. A queuing model can represent many situations in real life. A relevant example is the arrival of maintenance jobs to a workshop. Other examples include the arrival of customers to a bank and the arrival of telephone calls to a telephone number. The major components of a queuing system are the customers and the servers. The interaction of the customer with the server is gauged by the time the server spends serving the customer. The customers drive the system, and we are interested in their arrival pattern or in their interarrival time, which is usually modeled as a random variable. The service time is also modeled as a random variable with a certain probability distribution. Other important elements in a queuing model are: (1) the service discipline, which refers to the manner in which customers are selected for service, e.g., it could be first-come first-serve (FCFS) or some other certain priority rule; (2) the design of the facility, which refers to the number of servers and design of the queues (parallel, series, tandem, network); (3) the queue size (finite, infinite); (4) the size of the source of arrivals (finite, infinite); and (5) human behavior (jockey, balk). In summary, the elements of a queuing model are the following:

- (1) arrival distribution,
- (2) service time distribution,
- (3) design of service facility,
- (4) service discipline,
- (5) customer population, and
- (6) human behavior.

Specifying the above factors will result in a specific queuing model. Some queuing models have steady-state results that give the expected number of customers in the system, L_s , the expected queue length, L_q , the expected waiting time in the system, W_s , and the expected waiting time in the queue, W_q . The above results can be used to evaluate a queuing system.

A notation that summarizes the main characteristic of a parallel queuing model is $(a/b/c)$, $(d/e/f)$ where

- a Arrival distribution
- b Service time distribution
- c Number of parallel servers

- d* Service discipline
- e* Maximum number allowed in the system
- f* Size of customer population

One of the simplest queuing models is the one denoted by (M/M/C) (GD/∞/∞), in which there are C servers and interarrival and service times are exponential. If all the servers have exponential service time distribution with parameter μ and if we let $\rho = \lambda/\mu$, then we have the following steady-state results:

$$P_0 = \left\{ \sum_{n=0}^{c-1} \frac{\rho^n}{n!} + \frac{\rho^c}{c!(1 - \rho/c)} \right\}^{-1} \quad (2.42)$$

where P_0 is the probability of zero customers in the queue.

$$L_q = \frac{\rho^{c+1}}{(c-1)!(c-\rho)^2} P_0 \quad (2.43)$$

$$L_s = L_q + \rho \quad (2.44)$$

$$W_q = \frac{L_q}{\lambda} \quad (2.55)$$

$$W_s = W_q + \frac{1}{\mu} \quad (2.46)$$

If the above model is used as an approximation for a maintenance department, in which we want to determine the optimal number of repairmen (servers). The steady-state results can be used to evaluate machines' availability and utilization.

Another queuing model known as the machine servicing model is one in which we have R repairmen servicing a total of K machines, and, because, a broken machine cannot generate a breakdown while in service, the arrival population is finite. The number of repairmen R is less than K . The objective is to find a value R that minimizes the total expected costs that consist of the cost of failure and cost of service. This type of model is denoted by (M/M/R):(GD/K/K), $R < K$.

The model is a special case of the general queuing model. If λ is the rate of breakdown per machine, then if there are n broken machines, the arrival rate from the system is given as:

$$\lambda_n = \begin{cases} (K-n)\lambda & 0 \leq n \leq K \\ = 0 & n \geq K \end{cases} \quad (2.47)$$

and the service rate is as follows:

$$\mu_n \begin{cases} n\mu & 0 \leq n \leq R \\ R\mu & R \leq n \leq K \\ 0 & n > K \end{cases} \quad (2.48)$$

P_n is the probability of n machines in the system. The steady-state results for the system are derived as follows:

$$P_n = \begin{cases} \binom{K}{n} \rho^n P_0 & 0 \leq n \leq R \\ \binom{K}{n} \frac{n! \rho^n}{R! R^{n-R}} & R \leq n \leq K \end{cases} \quad (2.49)$$

$$P_0 = \sum_{n=0}^R \binom{K}{n} \rho^n + \sum_{n=R+1}^K \binom{K}{n} \frac{n! \rho^n}{R! R^{n-R}} \quad (2.50)$$

$$L_q = \sum_{n=R+1}^K (n - R) P_n \quad (2.51)$$

$$L_s = L_q + (R - \bar{R}) = L_q + \frac{\lambda_{eff}}{\mu}$$

$$\bar{R} = \text{expected number of idle repairmen} = \sum_{n=0}^R (\{R - n\}) P_n$$

$$\lambda_{eff} = \mu = (R - \bar{R}) = \lambda(K - L_s) \quad (2.52)$$

If $R = 1$, the model yields results for a system with single server. If a maintenance system can be represented by such a queuing model, its effectiveness can be evaluated using steady-state measures.

2.10.2 Stochastic Simulation

Maintenance systems have several characteristics that make capacity planning a rather complex problem. These characteristics are as follows:

- Maintenance as a function interacts with other technical and engineering functions in a complex fashion.
- The maintenance factors are highly dependent on each other.
- Maintenance as a function has many uncertain elements. These elements include demand for maintenance, time of arrival of job requests, content, time to complete a job, tools, equipment, and spare parts availability.

- The complexity of the maintenance capacity planning suggests that simulation is one of the most desirable approaches for modeling it.

Stochastic simulation is the process of representing a system on the computer and then employing well-designed experiments (scenarios), to evaluate the system performance. Using this process, systems can be analyzed, planned, and designed. Law and Kelton [6] provide ten major steps for conducting a typical simulation study. In this section, these ten steps are summarized in eight steps and the relationships among them are outlined:

1. *Purpose of simulation*: The first step toward a successful simulation study is to state precisely the purpose of the study. Simulation has been used in maintenance systems for the following purposes: to determine the optimal crew size and staffing, to evaluate the effect of maintenance policies on production systems, to design and plan maintenance operations, and to determine the shutdown time periods.
2. *Simulation models*: The conceptual model used in building the computer simulation study will affect the simulation accuracy and efficiency. The simulation model should contain only the necessary information that captures the essence of the system under study.
3. *Model assumptions*: The assumptions of a simulation model will affect the realism of the simulation results. They also may affect the way results are interpreted. Therefore, each assumption should be reviewed carefully before putting it into effect. Availability of manpower, equipment, job standards, and spare parts are some of the assumptions used in maintenance systems.
4. *Data Accuracy*: Accurate data and their distributions are very essential for a reliable simulation model. To simulate a maintenance system, the distribution of equipment failures and repair times must be identified using sound statistical methodology.
5. *Simulation languages and computers*: One of the major jobs in building a simulation model is to convert the conceptual model into an actual computer simulation program. There are over 100 simulation software programs currently available for a variety of computers. Computer languages such as GPSS, SLAM II, SIMAN, and SIMSCRIPT II.5 are generally used in simulation. Several other criteria also have been used in practice to classify the simulation software.
6. *Program verification and model validation*: Verification is testing and checking the computer code to show that it performs as intended. Validation is to ensure that the model's assumptions are realistic and correct, and the simulation model fairly represents the behavior of the modeled system. Even though this step is fairly tedious and time-consuming, it is the most important step in simulation studies.
7. *Design of experiment*: Another significant element in any simulation study is the design of the experiment. This comprises the following:
 - (a) selecting experimental factors,
 - (b) selecting measures of performance,

- (c) determining the initial conditions,
 - (d) determining the steady-state conditions,
 - (e) determining the length of a simulation run,
 - (f) determining the number of replications, and
 - (g) applying variance reduction techniques
8. *Output analysis:* In any simulation study, it pays very well to spend time on output analysis. To check for the true estimate, test, validate, and decide on the output results from your simulation, statistical techniques that ensure reliable estimates for system performance must be used. These include deciding on the length of the simulation run, the number of runs, and confidence intervals for estimated measures of performance.

A machine servicing model with general input distributions and general servicing distributions can be used to model maintenance operations and capacity planning. Depending on the maintenance activities and the maintenance load forecasted, the system can be divided into types of jobs requiring different crafts. Craft types include mechanical (ME), electrical (EE), instrumental (INS), civil (CI), and janitorial (JAN). The historical data can be used to develop a probability distribution that represents the arrival of maintenance jobs and the time to complete repairs for a job. The machine servicing model can be modified so that instead of considering physical machines, each type of work generated by the system may be considered as being generated by a hypothetical machine. Each type of work (machine) requires a specific craft or skills. The number of hypothetical machines is determined by the types of work generated by the system. Such a representation can be used within a simulation model to determine the size of crews needed. Figure 2.2 is a schematic representation of a model that can be simulated to determine the staffing levels in a maintenance department.

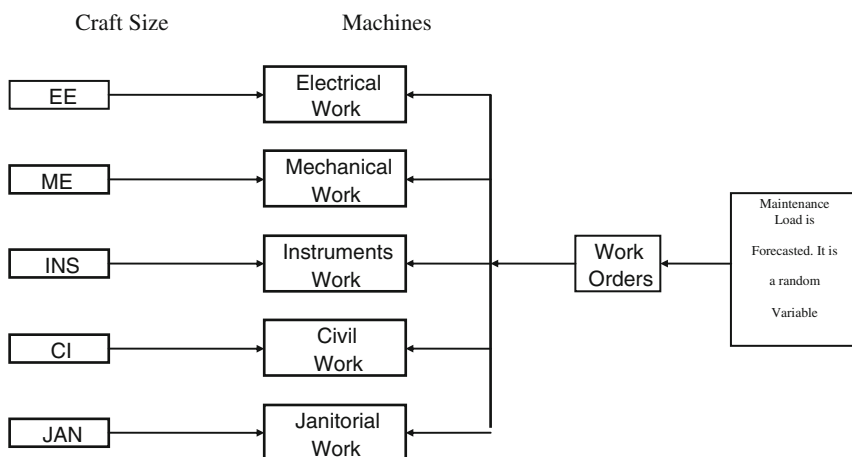


Fig. 2.2 Representation of maintenance capacity planning in a machine servicing framework

In order to implement a simulation model for capacity planning, the following practical steps are needed Duffuaa and Raouf [5].

1. A detailed study of the organization/plant maintenance requirements to determine the types of maintenance crafts and crews required, the types and criticality of equipment repaired, the failure mechanism for each equipment, and the effect of a failure on production or service provided by the organization.
2. Forecast the maintenance workload and divide it according to priority.
3. Outline the existing work order system and define the logic of work assignments (system discipline).
4. Set up the relevant machine servicing model after determining the failure rate of each machine, the service rate, and the cost of each machine being out of service.
5. Develop the simulation software and verify and validate the model.
6. After the model is validated, perform production runs and, on the basis of all measures of performance, find the optimal staffing levels.

2.11 Contract Maintenance

Contracting out maintenance activities is common throughout all industrial and public sectors. The premise is that certain aspects of maintenance can be done as effectively and at a lower cost than in-house resources can accomplish it. A key question is whether contracting maintenance builds or diminishes the competitive advantage of the organization. It is difficult to see how, for example, contracting out custodial services would harm the competitive advantage of a business, unless somehow in-house personnel could do it more effectively and significantly cheaper and with less management attention than a firm specializing in facility cleaning. On the other hand, if haul truck maintenance at a remote mine site is contracted out to a local garage and body shop, it is possible that the mining operation could suffer, regardless of the competitive rates.

In the petrochemical and oil refining sector in North America, it is not uncommon to see the majority of the execution of maintenance contracted to a firm specializing in that business, whereas the reliability engineering and maintenance management is often kept in-house. Major shutdown and overhaul maintenance requires the contracting out of a large segment of the shutdown work backlog, because there is usually a short, finite time period to accomplish all the work and not enough capacity within the organization to accomplish it.

Contracting maintenance during peak periods is more effective in most working environments when finite projects or tasks can be estimated and outsourced as a work package, as one would normally do for a capital expansion or modification project. This allows for a separate contracted crew to work together outside the

day-to-day organizational structure of the plant or facility. The work package would specify the contractor's trade skills and numbers and contain any required instructions. Contracting individual trades and integrating them with the in-house staff can lead to inefficiencies and conflicts, as these contracted staff may not know the routines, equipment, or working procedures and rules, thereby reducing the productivity of the entire crew. Further, particularly in a unionized environment, contracting out routine work without extensive bilateral communication on the issue of job security can often lead to labor relations' difficulties and some resentments that may lead to strikes or slowdowns.

2.12 Summary

This chapter presented the essential techniques for maintenance forecasting and strategic and capacity planning. Forecasting is a prerequisite for capacity planning, and capacity planning is a major element in maintenance organization and determines its ability to perform its mission in an effective way. The techniques covered in the forecasting sections of this chapter are moving averages, linear regression, exponential smoothing, and seasonal forecasting. Steps for developing a strategic plan for maintenance are outlined. The methods presented for capacity planning are simple heuristic tableau form, linear programming, queuing theory, and stochastic simulation. The application of the stochastic techniques requires some maturity on the part of the analyst and is introduced only briefly here (for more detail on stochastic techniques see [1–4]).

Exercises

1. The forecasting of the maintenance load that is generated by a newly developed sophisticated equipment does not seem to be amenable to time series forecasting. Why? Suggest a procedure to predict the maintenance load for the next two years for this equipment.
2. In the absence of data, you may have to resort to qualitative forecasting. Explain how would you perform qualitative forecasting and how to validate it?
3. When is it possible to use a control chart to control the qualitative forecasting process? What type of control charts will you use? Explain how.
4. Prove that using the least squares method to estimate the slope, b , and the intercept, a , of the single variable linear fit, $x(t) = a + bt$ results in estimates for b and a as follows:

$$\hat{b} = \frac{n \sum_{t=1}^n tx(t) - (\sum_{t=1}^n t)(\sum_{t=1}^n x(t))}{n \sum_{t=1}^n t^2 - (\sum_{t=1}^n t)^2}$$
$$\hat{a} = \bar{x}(t) - b\bar{t}$$

5. Use the least squares method to estimate the parameters of the quadratic model

$$x(t) = a + bt + bt^2$$

6. Derive the simultaneous equations a and b for the model $x(t) = ab^t$ using the least squares method.
7. Given the following data:

t	1	2	3	4	5	6
$x(t)$	5	4	7	8	7	10

- Predict $x(t)$ for periods 7 and 8 using five-period moving average and linear regression model. Which resulting prediction will you recommend for use in this case and why?
8. Show quantitatively that the estimate you obtain from the basic exponential smoothing equation is a function of all previous observations; however, the most recent ones are heavily weighted than the far-distant ones.
9. Use a linear model and exponential smoothing to predict the values of $x(t)$ for periods 7, 8, and 10 for the data in Exercise 7.
10. Prove that if a set of data have a noise-free ramp with slope b , the first exponential smoothing will lag the true value by $\left(\frac{\beta}{\alpha}\right)b$ where $\beta = 1 - \alpha$
11. How would you determine α optimally in the exponential smoothing model?
12. Consider the following quarterly data for the maintenance load in man-hours for the last 5 years.

Year	Quarter			
	I	II	III	IV
1	215	120	150	100
2	250	175	75	150
3	300	250	75	165
4	350	275	100	200
5	400	300	125	225

- (a) Determine an appropriate seasonal index for each quarter.
- (b) Deseasonalize the data and fit it to an appropriate growth model.
- (c) Predict the quarter values for the 6th and 7th years.

13. Discuss the advantages and disadvantages of the approaches given in this chapter for capacity planning.
14. Suppose you were asked to develop a capacity plan for a maintenance department. Which approach will you select from the approaches given in this chapter? and why?
15. For the data in Table 2.9, evaluate the cost of the plan given in the table if an in-house regular hour is 10\$, overtime hour is 15\$, and a contract man-hour costs 25. In addition, every backlog man-hour delayed from period t to $t + 1$ costs 5. Develop an alternative competitive plan in which you can have instead of 5 skilled employees have a mixture of skill and unskilled. The hour of unskilled worker is 8\$.
16. Develop a linear programming model for the data given in Table 2.9. Assume that the trades in problem 15 are available.
17. Explain how would you use the machine servicing model for determining optimal mix of crafts and skills to meet the maintenance load.
18. What are the disadvantages of using linear programming for capacity planning.
19. Locate a factory near your area and study its operations and maintenance.
 - (a) Forecast their maintenance load.
 - (b) Use the structured tableau method to determine their maintenance capacity in terms of staff only.
 - (c) Use linear and integer programming to plan their maintenance capacity.
20. Apply stochastic simulation to plan the maintenance capacity for the factory in problem 19. Is there a difference between the approaches in problems 19 and 20. Why do you expect such a difference? Which is more appropriate for this case and why?

References

1. Al-Turki OM (2011) Strategic maintenance management. *J Qualit Maint Eng* 18(2):87–94
2. Andijani AA, Duffuaa SO (1996) A critical evaluation of simulation study in maintenance. In: *Proceedings of the 1st international conference on industrial engineering, theory and applications*, Houston, Texas, Dec 1996
3. Bedworth DD, Bailey JE (1987) *Integrated production systems and contro*, 2nd edn. Wiley, New York
4. Duffuaa SO, Andijani AA (1996) Stochastic simulation: an effective tool for modeling maintenance systems. A working paper, Department of Systems Engineering, KFUPM, Dhahran 31261, Saudi Arabia
5. Duffuaa SO, Raouf A (1995) A simulation model for determining maintenance staffing in an industrial environment. *Simulation* 59:293–299
6. Law AM, Kelton WD (1993) *Simulation modeling and analysis*. McGraw Hill, New York
7. Montgomery DC, Johnson LA (1976) *Forecasting and time series analysis*. McGraw Hill, New York

8. Murthy DN, Atrens P, Eccleston JA (2002) Strategic maintenance management. *J Qualit Maint Eng* 8(4):287–305
9. Taha H (1992) *Operations research: an introduction*, 5th edn. Macmillan Publishing Company, New York
10. Tsang AHC (1998) A strategic approach to managing maintenance performance. *J Qualit Maint Eng* 4(2):87–94

Planning and Control of Maintenance Systems

Modelling and Analysis

Duffuaa, S.; Raouf, A.

2015, XXIII, 348 p. 80 illus., 6 illus. in color., Hardcover

ISBN: 978-3-319-19802-6