-

Forecasting Theories

(Exercises included)

From TETRAPRO, edited by Mr. H. Leijon, ITU

UNION INTERNATIONALE DES TELECOMMUNICATIONS
INTERNATIONAL TELECOMMUNICATION UNION
UNION INTERNACIONAL DE TELECOMUNICACIONES

Forecasting Theories

1. GENERAL ASSUMPTIONS FOR FORECASTING

1.1 Why forecast ?

1.1.1 *Telephone density and economic development*

The telecommunication facilities of a country depend on the economic level in that country. In order to develop further the economy, certain basic telecommunication facilities are necessary. This holds for all countries. The correlation between the number of telephones per inhabitant and the Gross Domestic Product per capita exists, as shown in Figure 1.1/1. Certain deviations occur as a result of differences in the economic structure of different countries. It is easily understood that a country with an economy mainly based on agriculture needs fewer telephones per head than a highly industrialized country.

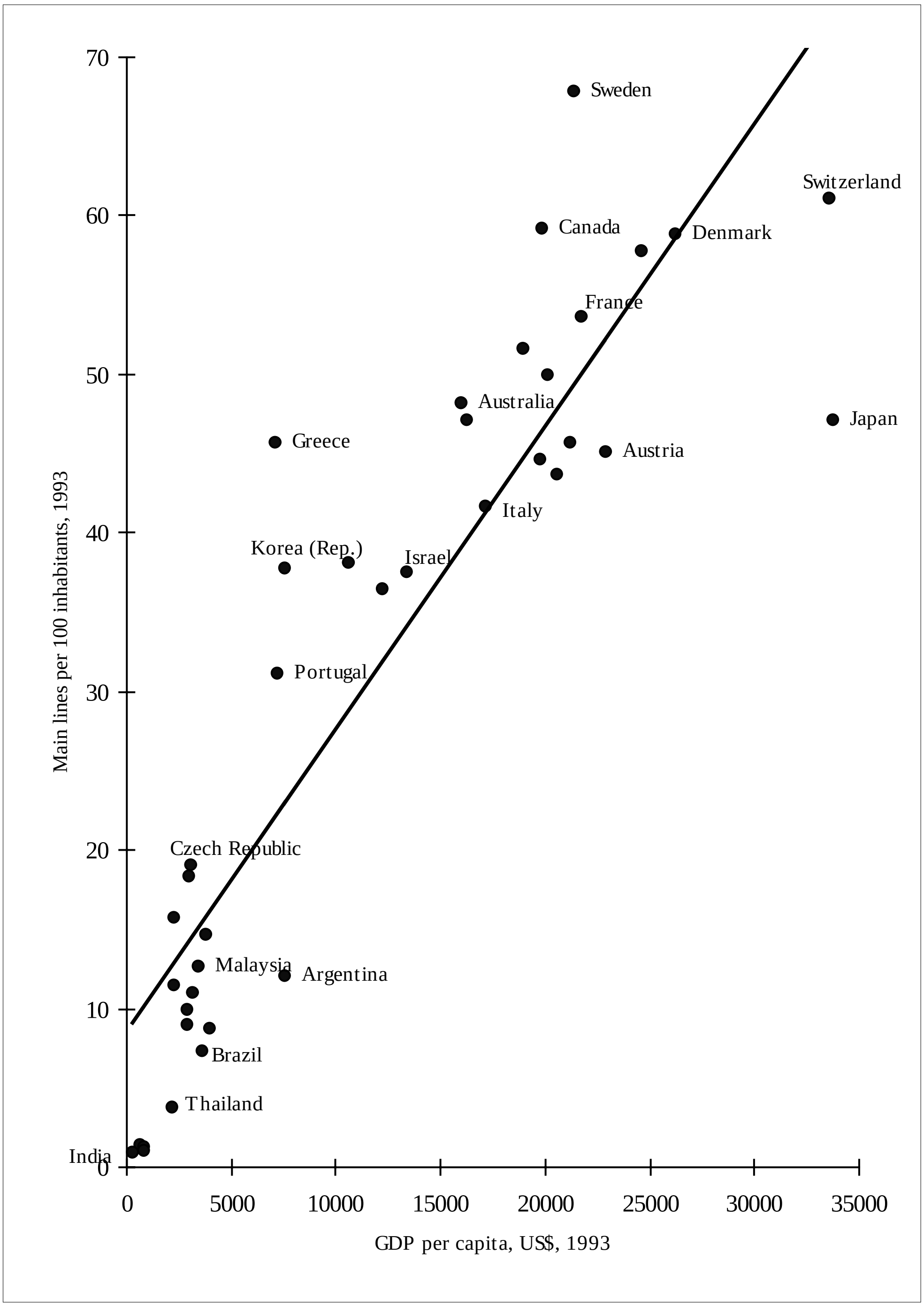


Figure 1.1/1 : Telephone density and Gross Domestic Product (GDP) per capita in some countries, 1993

The demand for telephones depends primarily upon the occupational specialisation of the population. The need for telephones is low in a self-supporting village, but increases as people leave the agricultural sector for other activities (industry, commerce, public services, etc.). Government and private commercial activities require telecommunications to function. The concentration of more and more people in urban areas may also give an increased demand for residential telephones, but this only concerns those persons who have reached a sufficiently hight standard of living. A certain demand for communication with rural areas will also follow from urbanization.. The demand is, of course, also dependent on national preferences and habits, as well as on the telephone tariffs. To avoid the influence of the latter, all investigations and prognosis should be made on the assumption that the telephone tariffs are not prohibitive but only sufficient to support adequate maintenance and sound expansion of the telephone network.

To estimate the future telecommunication demand, the expected general development and especially the economic development must be considered. Published data on the current economic status and information about plans for economic development are useful.

1.1.2 *Purpose of forecasts*

The extension of telephone services requires the provision of telephone apparatus, subscriber line plant, exchange equipment, junction and trunk circuits. However, there will always be a certain time lag between the identification of a need - or rather a future need - and the moment it can be met. The length of time between the identification of and the provision for this need is considerable. To avoid long waiting periods and high congestion, it is desirable for needs to be determined well in advance. This makes it possible to extend the plant at the right moment because the necessary action can be taken at the appropriate time.

A forecast will, therefore, primarily produce accurate estimates of the future demand for telecommunication facilities.

1.1.3 *Forecast, planning and work programme*

Forecasts provide the basis for the plans. The plans are considered by the management which makes its decisions. Following the decisions, a work programme is formulated. The work programme requires detailed planning of what needs to be done until all the equipment becomes operative. There is long-term, medium-term and short-term planning, each of which has its own requirements as concerns the details to be specifided at which point, if the programme's time schedule is to be followed. Each type of planning must use, more or less, detailed forecasts of the required quantities.

We will use the following definitions:

Forecast : A forecast is a prediction of the future demand, generally expressed in quantities.

Plan : A plan is a proposal for action and estimates the cost of carrying out the plan. It may include alternative lines of action.

Decision : The management decides what action shall be taken. It approves a plan which is then converted into a work programme.

Work programme : A work programme is based on an approved plan and defines action to be taken as well as its timing.

Planning A work programme generally requires detailed planning of all action to be taken to realise the work programme. The timing of each action is here of the utmost importance. More or less detailed work plans must be ready at a given time.

The relationship between the above concepts is illustrated in Figure 1.1/2

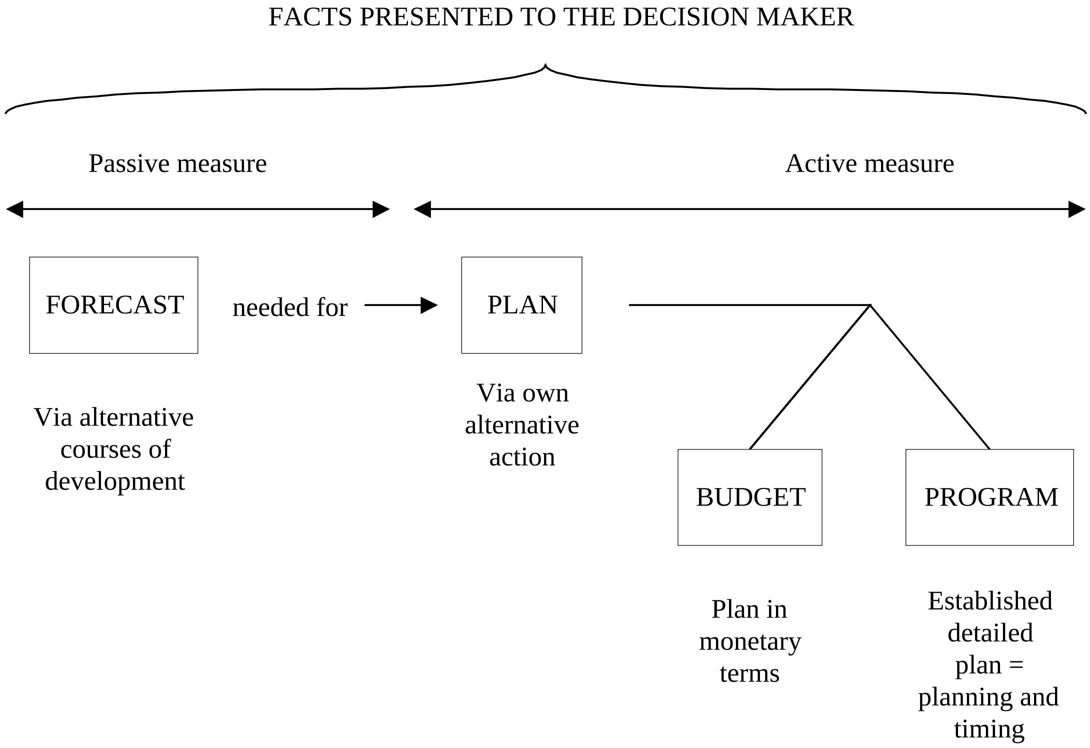


Figure 1.1/2 : Relationship between forecast and plan

Since the work programme and investments are based on forecasts, it is very important for these to be as accurate as possible. It is also desirable that the forecaster defines the degree of uncertainty of the forecast. This enables the planner to know when he can be flexible.

The situation in many developing countries may be such that the demand is far greater than the supply. In such a case, practically all extensions will be used immediately without essentially decreasing the gap between supply and demand. It is very difficult to forecast the demand in such cases since the existing demand cannot easily be measured or estimated. It is still the task of the forecaster to estimate this demand so that the decision-maker will know the effect on the demand/supply situation of his actions. Dealing with such situation is also a matter of what policy the administration will follow. Will it give priority to business subscribers, to development of local exchange areas, or to long distance services, etc.?

1.2 When is a forecast needed ?

The forecast periods depend on the planning periods. They, in turn, depend on the delivery times of the various types of equipment needed. We give a few examples below:

Subscriber apparatus	1-2 years
Exchange switching equipment	3-4 years
Subscriber cables	6-10 years
Ducts in local networks	10-15 years
Buildings	10-20 years
Sites for buildings	up to 50 years

Planning, as well as forecasts, are usually divided into:

Long-term planning/forecast	more than 10-15 years
Medium-term planning/forecast	5-10 years
Short-term planning/forecast	less than 5 years

It follows that the various types of forecast and planning concern different types of equipment dependent on when this equipment has to be used. Clearly, a short-term forecast must be more detailed as it serves the short-term planning in which every detail of the plant has to be specified.

No single forecasting method will be adequate for all time spans. Accordingly, different methods must be used for different periods, and this raises the problem of adjustment when forecasts for overlapping or non-overlapping time periods are derived with different methods since discrepancies may occur between the growth curves.

There is no standard procedure for matching contradictory forecasts. In most cases, judgment based on experience is used and a curve is drawn which bridges the inconsistency between the forecasts. This is shown in Figure 1.3/1.

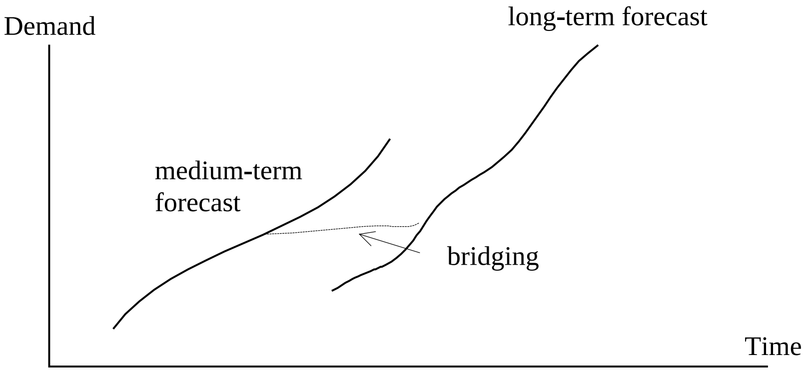


Figure 1.3/1 : Matching contradictory forecasts

1.3 What is to be forecast ?

1.3.1 *Forecasting subjects*

Forecasting concerns four main areas:

Accommodation	Site Buildings
Lines	Subscriber direct exchange lines (subscriber network forecast) Junction network Trunk network
Equipment	Subscriber apparatus Automatic switching equipment Switchboard equipment
Numbering	Exchange numbering schemes (long-term strategy)

Forecasts generally present the expected future number of direct exchange lines (DEL's) and the future traffic, the latter being more or less detailed. Consequently, there are:

- subscriber forecasts;
- traffic forecasts.

1.3.2 *Scope of forecast*

A forecast may concern the total for a given entity, such as a country, a certain part of a country, a metropolitan area, etc. In this case, it is called an *overall forecast* and contains very few details for this entity. It is usually only sufficient for long-term planning purposes. A *detailed forecast* provides detailed data of the expected future development. It is separately made for each part of the country where the division is such that the units become workable.

Only the detailed subscriber forecast specifies the exact location of each subscriber.

A detailed traffic forecast specifies all future traffic within the area concerned. It is generally based on subscriber forecast and on measurement.

1.3.3 *Data needed for planning*

Dependent on the planning need, results of the forecast should be presented in different ways. We give the following examples:

Subscriber distribution within an area

- Exact location of the individual subscribers.
- Number of subscribers in each grid of a grid map.
- Number of subscribers per traffic area in the area concerned.
- Number of subscribers per subscriber category in each traffic area.

Traffic

- Originating and terminating traffic per subscriber.
- Originating and terminating traffic per subscriber in each traffic area.
- Originating and terminating traffic per subscriber for each subscriber category.
- Traffic streams between areas.

We may define:

A = total traffic for a group of subscribers
N = number of subscribers in the group
α = traffic per subscriber
y = call intensity in the group
h = average holding time

Then, $A = y\, h$, but also $A = N\, \alpha$

Traffic A can now be forecasted in different ways:

1. Directly, for example, by trend extrapolation of recorded A values for the past years.
2. Indirectly, by first forecasting either N or α, or alternatively, y and h.

As concerns forecasting the traffic for a group of subscribers, it is generally preferable to make a forecast for each category of subscriber before the total traffic is estimated. The reason for this is that proportions between the different categories may change in the future.

1.4 How is forecasting done ?

1.4.1 *Basic requirements*

There are two main requirements for sound forecasting:

- An adequate supply relevant and accurate information of the past should be available. This will generally consist of records from measurements on existing equipment (connections, calls, traffic, etc.) supplemented by general background information.

- Systematic use of these data in the forecast.

- An educated guess as to future development. This estimate of future development may be an extrapolation of past development, sometimes adjusted to take into acccount available background information. The forecaster needs accurate historical data to improve his forecast.

Consequently, the basis for forecasting is the study of the past. The more we understand and mathematically describe the past development, the better are the chances of making a correct forecast.

It must also be stressed that the degree of uncertainty of forecast data should be given so that those who are going to use this data may take account of these uncertainties.

1.4.2 *Forecast check*

There is always a reason to be critical of the applicability of a forecast. The following questions should be asked:

- Is the forecast relevant (valid) ?

- How accurate is the forecast ?

- Is the outcome credible ?

The relevance of a forecast depends in some cases on the correct use of assumptions for future development.

The accuracy of a forecast may depend on the statistical precision of the historical data and on the extrapolation method used. Whether a forecast is credible or not is often a matter of personal judgment. Here, the human mind may, however, have its limitations when it comes to imagining future growth. Many forecasts have, therefore, been decreased in size since the forecaster could not envisage such growth. The development of many networks has been hampered because of the forecaster's lack of imagination.

When judging the credibility of a forecast, it must be remembered that:

- Forecasts concerning population, economic and industrial development, etc., may be misleading.

- Available telecommunication statistics for past periods may contain errors and may not have always been collected under the same conditions.

- Available data on historical development may only cover a limited period.

- Past and present relationships between studied variables are not necessarily true in the future.

1.4.3 *How do we start ?*

The forecasting process can be divided into the following parts:

Definition of the problem

The purpose of and assumptions for forecasts have to be determined.

Collection of basic data

Various sources for basic data are investigated. The population and economic growth are studied. Results of recent forecasts are essential facts.

Choice of the forecast method

The method has to be chosen with regard to available information and required accuracy, etc.

Analysis and establishment of forecasts

The analysis consists of methodological preparation of the basic data and evaluation of the results obtained.

Documentation

The forecast has to be presented in an easily understandable format. The result should contain alternative forecasts. Besides the most likely forecast, there should be one optimistic and one pessimistic forecast, thus telling the planner where the upper and lower limits may be.

2. FORECASTING METHODS

2.1 Introduction

Almost all forecasting methods assume that the future will in some way resemble the past. This may be interpreted in various ways. For example:

1. For time trend forecasting methods, it is assumed that development will follow a curve which has been fitted to existing historical data.
2. When using explicit relationships between demand and various determining factors, these will remain the same in the future.
3. Comparing various steps of telecommunication development, it is assumed that the less-developed country (or area) will develop to the level of the more developed one.
4. When applying personal (subjective) judgment in the forecast, the future will resemble the person's previous knowledge and experience of past developments.

Although the future often resembles the past in this sense, it is never an exact reproduction. Forecasts must, therefore, never be accepted uncritically. This means that forecasts - as produced by some methods - are frequently modified before being accepted for planning purposes. While such amendments may appear subjective, they are necessary whenever there is reason to believe that the future will show divergences from the past, not accounted for by the forecasting method used. Examples of such situations are tariff changes, radical improvements of the service, etc.

It follows that the assumptions used in forecasting should be questioned, whenever possible changes occur in the environment in which the telephone system operates.

The importance of accurate and detailed data cannot be overemphasized. The quality of a forecast depends directly on the quality of the historical data used. If reliable data are not available, then the forecaster should first establish adequate data collection processes before applying forecasting methods. Many of the methods discussed in this chapter would be difficult, if not impossible, to use without adequate historical data.

2.2 Statistical demand analysis

One can assume that the development of the number of direct exchange lines in a country follows certain patterns. For example, it ought to be dependent on the number of inhabitants, living standard, technical and economic development, etc. If some variables are logically related to the telephone development, those variables may be used to explain the development. In the following, some such variables are described:

Population size - as the population grows, the number of consumers increases and the demand for subscriber lines increases.

Living standard - one measure of the living standard is the Gross Domestic Product per capita. In centrally planned economies, the Net Material Product per capita may be used. Both values should be converted into a fixed price level. When the living standard increases, the demand for subscriber lines also generally increases.

Building activity - as a measure of the change of structure in society due to urbanisation, high living standard, population development, etc., one can use the number of new offices and departments. The building activity can be used to estimate the future demand for subscriber lines.

Price - a high price for using a telephone (i.e. entrance fee, subscription cost and call charges) can be assumed to reduce the demand for subscriber lines.

The choice of explaining variables is dependent on the availability of reliable forecasts for them. Population and economic development are the most common explanatory variables for the number of subscriber lines in a country.

Figure 2.2/1 shows the telephone density and the GDP per capita for Sweden 1900-1965.

On a logarithmic scale, the relationship is approximately linear during the periods 1900-1915 and 1920-1965.

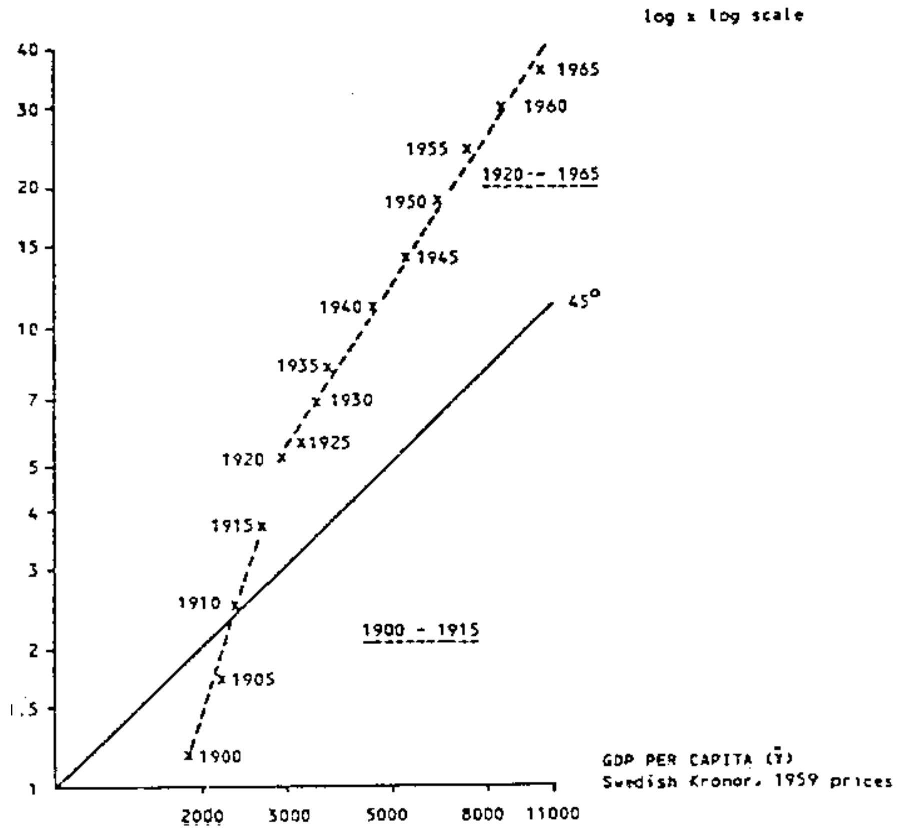


Figure 2.2/1 : Telephone density and GDP per capita in Sweden, 1900-1965

Statistical demand analysis by *econometric modelling* is a technique designed to obtain explicit quantitative relationships between interest and other variables which are believed to influence it.

An econometric model of demand is thus a mathematical representation of customer behaviour, and it is dependent on many more factors than those for which data can be obtained. Some factors, despite having a clear relationship with the forecast variables, are themselves difficult to forecast and other factors may be interdependent so their precise influence is unclear.

There are two main classes of model: *single equation regression* models and *multi-equation regression* models. The former comprises models in which the dependent variable is assumed to be a linear function of several explanatory and independent variables.

The second class of model, the multi-equation regression model, gives us the possibility of accounting simultaneously for the inter-relationship between a set of variables. The weakness of the multi-equation models is that the explanatory variables may be interdependent. The possible feedback between the dependent variable and the explanatory variables is neglected.

2.2.1 *Two-variable regression*

Assume that we want to study the relationship between two variables, x and y. In order to describe this relationship statistically, we need a set of observations of pairs of values, one from each variable:

$$(x_1, y_1), (x_2, y_2), ..., (x_n, y_n)$$

If we plot these quantities, we may find that they show a consistent pattern, and it may therefore be possible to draw a curve that “fits” these points. Two problems then arise:

a) decide what type of curve to use;

b) choose the “best” curve of that type.

The first problem is usually solved simply by looking at a graphical representation of the observed value pairs (x_i, y_i) and preferably by choosing on sight a simple type of curve. Straight lines are very popular in this respect.

The second problem, choosing the optimal values for the parameters of the curve, can only be solved after we have decided on some sort of criterion that informs us about the quality of the curve. The least-squares criterion is mostly used. This is because of its intuitive appeal and its mathematical properties. We will also use it here. According to the least-squares criterion, a curve is good if *S*, the sum of squares of the “distances” from the given points to the curve, is small. The best curve is that for which *S* is the smallest.

The second problem can now be rephrased as: *determine the parameter values that minimize S.*

Suppose that we want to determine a straight line that fits the observations. This means that we assume a linear relationship between *x*, the independent variable, and *y*, the dependent variable. The fact that the observations (x_i, y_i) are not lying on a line is attributed to the influence of an error term.

In short, we assume:

$$y = \alpha + \beta x + \varepsilon \qquad (2.2.1)$$

The parameters α and β are unknown. The error ε is random but has an expected value equal to zero.

Let *a* and *b* be estimates of α, resp. β. The calculated value for *y* for a given value x_i using these estimates:

`

$$\hat{y}_i = a + bx_i \qquad (2.2.2)$$

However, we observed the pair (x_i, y_i), so $y_i - \hat{y}_i$ is the deviation of this particular point from the line. Thus, we have to find estimates *a* and *b*, that minimize:

$$S = \sum_{i=1}^{n} (y_i - \hat{y}_i)^2$$

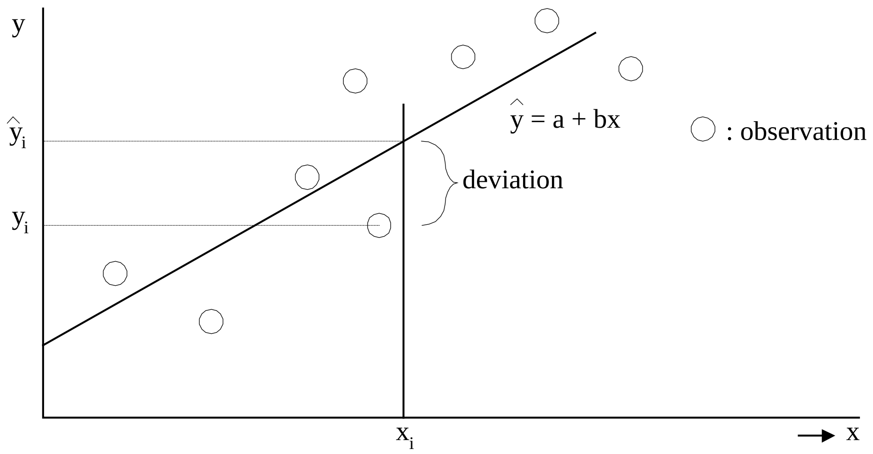


Figure 2.2/1 : The observations and the line

Note that we are only interested in the vertical deviations. This is not surprising as our ultimate goal is to predict the *y*- value when a new *x*- value is given. In general, we have the choice of three possibilities:

V: Vertical distances
H: Horizontal distances
P: Distances perpendicular to the line

The following figure gives the results obtained, starting from the same observations in these three cases.

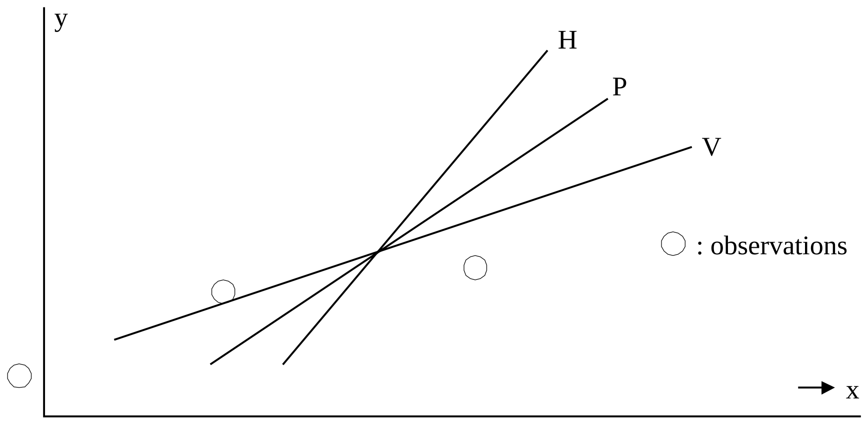


Figure 2.2/2 : Influence of the "distance" chosen

The estimates *a* and *b* follow by solving:

$$\frac{\delta S}{\delta a} = 0 \qquad and \qquad \frac{\delta S}{\delta b} = 0$$

where

$$S = \sum_{i=1}^{n} (y_i - a - bx_i)^2$$

The result is:

$$b = \frac{n\sum_{1}^{n} x_i y_i - \sum_{1}^{n} x_i \sum_{1}^{n} y_i}{n\sum_{1}^{n} x_i^2 - \left(\sum_{1}^{n} x_i\right)^2} \qquad (2.2.3)$$

$$a = y - b\,x \qquad (2.2.4)$$

To judge the quality of the line so produced, we can calculate the coefficient of correlation *r*.

$$r^2 = \frac{\sum(\hat{y}_i - \bar{y})^2}{\sum(y_i - \bar{y})^2} = \frac{Explained \quad variance}{Total \quad variance} \qquad (2.2.5)$$

r has the same sign as *b* and lies in [-1, +1]. $r = 0$ indicates lack of correlation, in other words: a straight line cannot represent the relationship between *y* and *x*.

If *r* deviates from zero, there are difficulties in judging the goodness of fit between the observations and the straight line. Other methods, therefore, have to be applied.

If we are satisfied with the value of *r* obtained, we can use

$$y = a + b\,x$$

to provide point-estimates of *y*, given any value of *x*.

There are, however, other methods to check that the regression line gives a reliable description of how *y* depends on *x*. The first step is to estimate the variance of *b* and check that it lies within reasonable limits. The second step is to check against systematic errors with the Durbin-Watson test.

The first test requires calculation of the quantity

$$T = \frac{b\left(\sum (x - \bar{x})^2\right)^{1/2}}{s} \qquad (2.2.6)$$

where

$$s^2 = \frac{\sum y^2 - a\sum y - b\sum xy}{n-2} \qquad (2.2.7)$$

For accurate checking, it is necessary to use statistical tables for the *t*-distribution, but *T* should be less than -2 or greater than + 2 to make the regression analysis acceptable.

For - 2 < T < 2, the value of *b* cannot be said to be significantly different from zero and we cannot be sure that our straight line with $b \neq 0$ really represents a change of *y* with *x*.

If we want to design confidence intervals for the parameters α and β, or for our predicted *y*-values, we must make the following assumption:

The observations (x_i, y_i) form a random sample from a bivariate-Normal distributed population such that:

$$E\{Y|X = x\} = \alpha + \beta x$$
$$V\{Y| = x\} = \sigma^2 \qquad (2.2.8)$$

From now on, and without loss of generality, we shall also assume that $\bar{x} = \bar{y} = 0$; this means that we choose a new origin so that before starting the calculations, we have to substract $\bar{x}$ from all the x_i -values, and $\bar{y}$ from all y_i -values.

The estimate for β can now be written as

$$b = \frac{\sum x_i y_i}{\sum x_i^2} \qquad (2.2.3a)$$

and

$$a = 0 \qquad (2.2.4a)$$

The variance σ^2 is unknown, but we can estimate it by

$$s_e^2 = \frac{1}{n-2}\sum_{i=1}^{n}(y_i - \hat{y})^2 \qquad (2.2.8)$$

s_e is called the standard error of the estimate. s_e^2 is an unbiased estimator of σ^2.

Moreover,

$$\frac{(n-2)s_e^2}{\sigma^2} = \chi^{\,2}(n-2) \qquad \text{(Chi -square distributed with } n\text{-}2 \text{ degrees of freedom)}$$

$$a \approx N\left(\alpha, \frac{\sigma^2}{n}\right) \ \ and \ \ b \approx N\left(\beta, \frac{\sigma^2}{\sum x_i^2}\right), \ \ so \ \ \frac{a-\alpha}{\sigma}\sqrt{n} \ \ and \ \ \frac{b-\beta}{\sigma}\sqrt{\sum x_i^2}$$

are normal distributed, *N(0,1)*, (with mean = 0 and standard deviation = 1).

It can be shown that *a, b* and s_e are independent of each other, so

$$\frac{a-\alpha}{\sigma} \ \ n \ \ \frac{\sigma^2}{s_e^2} \qquad and \qquad \frac{b-s}{\sigma}\sqrt{\sum x_i^{\,2}}\sqrt{\frac{\sigma^2}{s_e^{\,2}}}$$

are *t*-distributed with *n-2* degrees of freedom. Confidence intervals can be calculated and tests of hypotheses performed. To design a confidence interval for the value of y when $x = x_0$, we must look at the variance of $y_o = bx_o$. Due to our new origin, which is different for each sample, the variance of y_o includes the variance of the origin, V(a), and is equal to:

$$V(y_o) = V(a) + x_o^{\,2}\ V(b) = \sigma^{\,2}\left(\frac{1}{n} + \frac{x_o^{\,2}}{\sum x_i^{\,2}}\right) \qquad (2.2.9)$$

Again, we arrive at a *t(n-2)*-distribution by eliminating the unknown σ^2 because

$$\frac{y_o - \alpha - \beta x_o}{\sigma\sqrt{\frac{1}{n} + \frac{x_o^2}{\sum x_i^2}}} \approx N(0,1)$$

and so

$$\frac{y - \alpha - \beta x_o}{se\sqrt{\frac{1}{n} + \frac{x_o^2}{\sum x_i^2}}} \approx t(n-2)$$

In the above, we have restricted ourselves to a linear relation between *x* and *y*. The same method, however, can be applied in certain cases where a non-linear relationship seems more appropriate.

For this, it is sufficient that the relation can be brought in a linear form by a suitable transformation of the data.

For example, if we assume that the relationship can be expressed by

$$y = \alpha \ \ e^{\beta x} \qquad (2.2.10)$$

then $\lambda n \ \ y = \lambda n\alpha + \beta x$ (2.2.10a)

or $z = \gamma + \beta x$ (2.2.10b)

We determine the estimates *c* and *b* for γ , resp. β, in the usual way and forecast with

$$y = e^z = e^{c+bx} = e^c e^{bx} \qquad (2.2.10c)$$

If transformation of the data to something like a linear relationship is impossible, curvilinear regression can be tried.

The underlying idea is now that the actual relationship between x and y can be represented by

$$y = \beta_o + \beta_1 x + \beta_2 x^2 +, \ldots, + \beta_k x^k \qquad (2.2.11)$$

Our observations (x, y) are the result of random fluctuations ε_I around this line.

Applying the least-squares criterion for estimation of $\beta_o, \beta_1, \ldots, \beta_k$ means again that $\sum \varepsilon_i$ has to be minimized.

Thus, we have tried to find the minimum of

$$\sum_{i=1}^{n} \left(y_i - (b_o + b_1 x_1 + b_2 x_i^2 + \ldots + b_k x_i^k) \right)^2$$

where b_i should estimate β_i .

Differentiating partially to $b_o, b_1, \ldots, b_k$, and equating all the derivatives to zero, gives the linear equation system:

$$\sum y_i = nb_o + b_1 \sum x_i + b_2 \sum x_i^2 +, \ldots, + b_k \sum x_i^k$$

$$\sum x_i y_i = b_o \sum x_i + b_1 \sum x_i^2 + b_2 \sum x_i^3 +, \ldots, + b_k \sum x_i^{k+1}$$

$$\sum x_i^k y_i = b_o \sum x_i^k + b_1 \sum x_i^{k+1} +, \ldots, + b_k \sum x_i^{2k}$$

The solution of this linear equation system with $k+1$ unknown gives the required estimates. It is advisable to limit the number of parameters b_k used.

2.2.2 *Multiple regression*

If we are not satisfied with the value of r, which means that the residual variance (= total variance - explained variance) is too large, we can look at the residues. The residual variance is

$$\left(y_i - \hat{y}_i \right)^2$$

and can have a high value only if at least some of the residues $y_i - \hat{y}_i$ are large, meaning that vertical distances between observations and the line $y = a + bx$ are wide.

If, however, the studied variable y is likely to depend on other variables than x, one can try to reduce the residuals by introducing one or more "explanatory" factors, x_1, x_2, etc. There are two main methods to apply, the *Iterative Method* and the *Direct Method*, which will be explained by two numerical examples.

Iterative Method

Example: Forecasting the number of local calls.

The method shown below is called the iterative method because it works in steps. This is preferable when explanatory factors to be used are not decideed in advance.

It should be noted that although the method involves a certain degree of approximation, the numerical results will be sufficiently accurate for practical cases like the one given here.

Step 1 Select a couple of possible explanatory factors, $x_1, \ldots, x_n$ for which good forecasts are available (number of subscribers, inhabitants, etc.).

Step 2 Calculate the correlation coefficients between the dependent variable y (in the example below the number of local calls) and each of the explanatory factors $x_1, \ldots, x_n$.

Step 3 Select the explanatory factor which has the highest correlation coefficient with y_1.

Step 4 If this factor is presumed to be x_2, make your first attempt with the regression function

$$y = a + b\,x_2 \qquad (2.2.12)$$

and calculate the parameters a and b according to the methods in section 2.2.1.

Step 5 Calculate the residuals

$$r = y - (a + b\,x_2) \qquad (2.2.13)$$

Step 6 Calculate the correlation coefficients between the r_1 :s and everyone of the remaining explanatory factors, in this case $x_1, x_3, \ldots, x_m$

Step 7 Select the remaining explanatory factor which has the highest correlation coefficient with r_1.

Step 8 If this factor is called x_1, make attempt No. 2 by using the regression function

$$r_1 = a_2 + b_2 x_1 \qquad (2.2.14)$$

and calculate the parameters a_2 and b_2, according to the methods in Section 2.2.1.

Step 9 The function from attempt 1, together with the function from attempt 2, now gives

$$\begin{aligned} y_1 &= (a_1 + b_1 x_2) + r_1 = \\ &= (a_1 + b_1 x_2) + (a_2 + b_2 x_1) = \\ &= (a_1 + a_2) + b_2 x_1 + b_1 x_2 \end{aligned} \qquad (2.2.15)$$

which is an improvement compared with the regression function from the first attempt.

Step 10 Calculate the residuals

$$r_2 = y_1 - [(a_1 + a_2) + b_2 x_1 + b_1 x_2] \qquad (2.2.16)$$

Step 11 Calculate the correlation coefficients between the r_2:s and each of the remaining explanatory factors, in this case $x_3, \ldots, x_n$

Step 12 Go on with attempt 3 and continue in the same way as long as the calculated residuals are not small enough.

Perfect fitting will not mean that all residuals are zero. This state will only be reached when the number of parameters is the same as the number of observation in the statistical material.

Numerical Calculations

Financial year	Local calls y_1	Long distance calls Manual service	Automatic service	Total y_2
1958/59	2305	110	125	235
1959/60	2423	103	165	268
1960/61	2605	95	220	315
1961/62	2852	84	267	351
1962/63	3040	78	322	400
1963/64	3376	73	372	445
1964/65	3642	55	445	500
1965/66	3700	38	530	568
1966/67	3884	21	609	630
1967/68	4059	13	650	663

Table 2.2/1 Telephone Calls (million)

Step 1: Select possible explanatory factors

Year 1:st Jan.	Population x_1	Telephone Subscribers man. exch.	aut. exch.	Total x_2	Telephone Apparatus man. ex-exch.	aut. ex-exch.	Total x_3
1959	7 436	352	1 712	2 064	408	2 118	2 526
1960	7 471	311	1 839	2 150	359	2 278	2 637
1961	7 499	271	1 973	2 244	313	2 448	2 761
1962	7 542	207	2 143	2 350	236	2 668	2 904
1963	7 581	174	2 294	2 468	198	2 856	3 054
1964	7 627	144	2 454	2 598	163	3 060	3 223
1965	7 695	101	2 622	2 723	114	3 273	3 387
1966	7 773	75	2 785	2 860	86	3 487	3 573
1967	7 844	41	2 956	2 997	46	3 711	3 757
1968	7 894	27	3 095	3 122	32	3 903	3 935

Table 2.2/2 Explanatory Factors (thousands)

1:st Jan.	Population x_1	Subscribers Total x_2	Apparatus Total x_3
1969	7 964	3 214	4 111
1970	8 042	3 335	4 288
1971	8 120	3 453	4 458
1972	8 201	3 566	4 626
1973	8 281	3 672	4 793
1974	8 360	3 776	4 957
1975	8 430	3 879	5 120

Table 2.2/3 Available Forecasts for Explanatory Factors (thousands)

Step 2: Calculate the correlation coefficients. Note that the basic figures given in Tables 2.2/1-2 are rounded off in the following calculations.

Calculation of the correlation coefficients between the number of local calls y and the explanatory factors x_1, x_2 and x_3 gave the results shown. Data is taken from Tables 2.2/1-2.

y_1 10^9	x_1 10^6	x_2 10^6	x_3 10^6
2.31	7.44	2.06	2.53
2.42	7.47	2.15	2.64
2.60	7.50	2.24	2.76
2.85	7.54	2.35	2.90
3.04	7.58	2.47	3.05
3.38	7.63	2.60	3.22
3.64	7.70	2.72	3.39
3.70	7.77	2.86	3.57
3.88	7.84	3.00	3.76
4.06	7.89	3.12	3.94
31.88	76.36	25.57	31.76

Table 2.2/4 Given Explanatory Factors

The following sums are calculated:

$$\sum y_i = 31.88 \qquad \sum y_i^2 = 105.22$$

$$\sum x_1 = 76.36 \qquad \sum x_2 = 25.57 \qquad \sum x_3 = 31.76$$

$$\sum x_1^2 = 583.31 \qquad \sum x_2^2 = 66.58 \qquad \sum x_3^2 = 102.97$$

$$\sum x_1 y_1 = 244.32 \qquad \sum x_2 y_1 = 83.57 \qquad \sum x_3 y_1 = 103.96$$

The correlation coefficient is calculated from:

$$R(y_1 x_i) = \frac{D_i}{\sqrt{G_i H_i}}$$

where

$$D_i = n\sum x_i y_1 - \sum x_i \sum y_1$$

$$G_i = n\sum y_1^2 - \left(\sum y_1\right)^2$$

$$H_i = n\sum x_i^2 - \left(\sum x_i\right)^2$$

which gives

$$R(x_1 y_1) = 0.979$$
$$R(x_2 y_1) = 0.991$$
$$R(x_3 y_1) = 0.989$$

Step 3: Select the explanatory factor to be used first.

In this first attempt, we choose the explanatory factor with the highest correlation coefficient with y_1. Here, it was the number of subscribers = x_2.

Step 4: Make attempt No. 1 with the regression function

$$y_1 = a_1 + b_1 x_2$$

Formulae for calculation of *a* and *b* in Section 2.2 can be written as follows:

$$a_1 n + b_1 \sum x_2 = \sum y_1$$
$$a_1 \sum x_2 + b_1 \sum x_{2i}^2 = \sum x_2 y_1$$

or with the numerical values:

$$10.00\ a_1 + 25.57\ b_1 = 31.88$$

$$25.57\ a_1 + 66.58\ b_1 = 83.57$$

The solution is:

$$a_1 = -1.196$$

$$b_1 = 1.715$$

Step 5: Calculate the residual r_1.

The function arrived at in Step 4 (attempt 1) gives the following fitted values of y_1:

Table 2.2/5 Fitted values of *y* from Step 4

Financial year		Fitted values - 1.13 + 1.677 x =	Comparison with the given values of *y*	
			Table 2.2/1	residual
	x_2	y_1	y	r_1
1958/59	2.06	2.34	2.31	- 0.03
1959/60	2.15	2.49	2.42	- 0.07
1960/61	2.24	2.64	2.60	- 0.04
1961/62	2.35	2.83	2.85	+ 0.02
1962/63	2.47	3.04	3.04	+ 0.00
1963/64	2.60	3.26	3.38	+ 0.12
1964/65	2.72	3.47	3.64	+ 0.17
1965/66	2.86	3.71	3.70	+ 0.01
1966/67	3.00	3.95	3.88	- 0.07
1967/68	3.12	4.15	4.06	- 0.09

Step 6: Calculate the correlation coefficients with r_1.

Before our second attempt, we calculate the correlation coefficients between the residuals from attempt 1 and the remaining explanatory factors x_1 and x_3.

Table 2.2/6 Calculation of correlation coefficients.

Residuals from attempt No. 1 r_1	x_1 from Table 2.2/2	x_3 from Table 2.2/2	Calculated sums: Variable	Calculated sums: Value
- 0.03	7.44	2.53	$\sum r_1^2$	= 0.0642
- 0.07	7.47	2.64		
- 0.04	7.50	2.76	$\sum x_1^2$	= 583.31
- 0.02	7.54	2.90		
- 0.00	7.58	3.05	$\sum x_3^2$	= 102.97
+ 0.12	7.63	3.22		
+ 0.17	7.70	3.39	$\sum r_1 x_1$	= - 0.0073
- 0.01	7.77	3.57		
- 0.07	7.84	3.76	$\sum r_1 x_3$	= - 0.1199
- 0.09	7.89	3.94		
Total	76.36	31.76		

The correlation coefficients were:

$R\ (r_1, x_1) = -\ 0.029$

$R\ (r_1, x_3) = -\ 0.019$

This indicates a very low correlation.

Step 7: Select the next explanatory factor.

The highest correlation coefficient is the one between r_1 and x_1, so our second attempt will be to use the explanatory factor x_1.

Step 8: Make attempt 2 with the function

$r_1 = a_2 + b_2\, x_1$

Parameters a_1 and b_2 are given from the following equations:

$$a_2 n + b_2 \sum x_1 = \sum r_1$$
$$a_2 \sum x_1 + b_2 \sum x_1^2 = \sum r_1 x_1$$

or $10.00\ a_2 + 76.36\ b_2 = -\ 0.04$

$76.36\ a_2 + 583.31\ b_2 = -\ 0.3089$

The solution is then:

$a_2 = 0.18 \quad 0.113$

$b_2 = -\ 0.023 - 0.0154$ (cor - 0.028822)

Consequently $\quad r_1 = 0.113 - 0.0154\, x_1$

Step 9: The results from attempts 1 and 2 above-mentioned will now give the following regression function for local calls:

$y_1 = (-1.196 + 1.715\, x_2) \quad + \quad r_1 \quad =$

↑ ↑

(from attempt 1) (from attempt 2)

↓ ↓

$= (-1.196 + 1.715\, x_2) \quad + \quad (0.113 - 0.0154\, x_1)$

consequently

$y_1 = -1.083 - 0.0154\, x_1 + 1.715\, x_2$

This result is compared with the given y_1 -values in Table 2.2/1.

Step 10: Calculate the residuals r_2

Financial year	Fitted values for local calls $-\ 1.083 - 0.0154\ x_1 + 1.715\ x_2 = y_1$	Given values of *y*	
		Table 2.2/1 x_2	Residual r_2
1958/59	2.34	2.31	- 0.03
1959/60	2.49	2.42	- 0.07
1960/61	2.64	2.60	- 0.04
1961/62	2.83	2.85	+ 0.02
1962/63	3.04	3.04	0.00
1963/64	3.26	3.38	+ 0.12
1964/65	3.46	3.64	+ 0.18
1965/66	3.70	3.70	0.00
1966/67	3.94	3.88	- 0.06
1967/68	4.15	4.06	- 0.09

From this regression function, $y\ (x_1,\ x_2)$, we find the following future number of local calls (in million):

	x_1	x_2	$y\ (x_1 x_2)$	$y(x_2)$
1968/69	7.96	3.21	4300	4310
1969/70	8.04	3.34	4520	4530
1970/71	8.12	3.45	4710	4720
1971/72	8.20	3.57	4910	4930
1972/73	8.28	3.67	5080	5100
1973/74	8.36	3.78	5270	5290
1974/75	8.43	3.88	5440	5460

The result is given in Figure 2.2/3

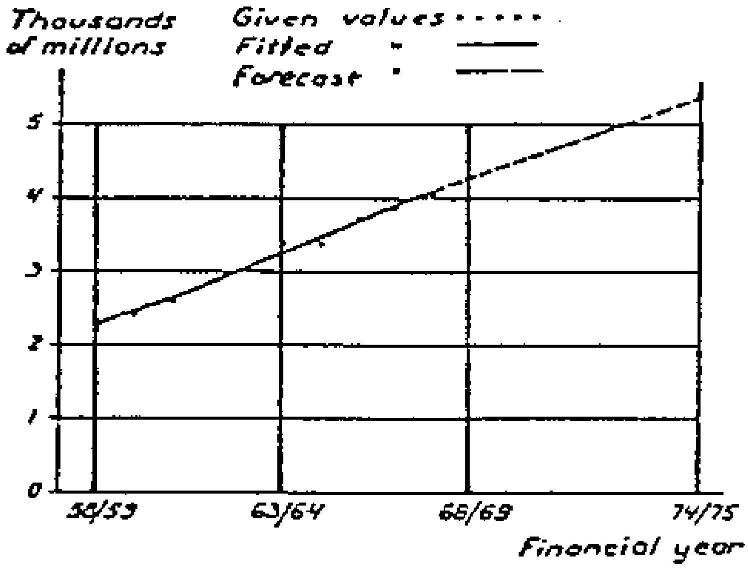


Figure 2.2/3

From the residuals r_2 follows that attempt 2 gave practically no difference as compared with attempt 1, $y(x_2)$ above. This followed already from the rather small correlation coefficient $R(r_1, x_1)$.

We should, therefore, stop at step 8 and use the regression function.

$y_1 = -1.083 - 0.0154\, x_1 + 1.715\, x_2$

Direct Method

For this method, we assume from the start that the variable we want to forecast depends on a number of explanatory variables.

Example: *Forecasting the number of long-distance calls.*

The number of long-distance calls, y_2, generally increases heavily by the introduction of automatic service (STD). It was frequently observed that introduction of STD between two areas caused a doubling of the number of calls.

Therefore, it seems appropriate to use the explanatory factor "*degree of automixation*" (denoted by x_4), to forecast the number of long-distance calls. This degree is defined as the ratio between the number of long-distance calls and all long-distance calls.

It also seems likely that the number of telephone apparatus (denoted x_3) could be a useable explanatory factor. After these considerations, we select a priori the two explanatory factors:

- number of telephone apparatus $= x_3$
- degree of automization $= x_4$

It is presumed that we need no further explanatory factors, so we can apply the "*direction method*" for calculation of the parameters in the regression function:

$$y_2 = a + b_3 x_3 + b_4 x_4 \qquad (2.2.17)$$

Financial year	Long-distance calls in thousand million y_2	Telephone apparatus in million x_3	Degree of automazation x_4
1958/59	0.235	2.53	0.53
1959/60	0.268	2.64	0.62
1960/61	0.315	2.76	0.70
1961/62	0.351	2.90	0.76
1962/63	0.400	3.05	0.81
1963/64	0.445	3.22	0.84
1964/65	0.500	3.39	0.89
1965/66	0.568	3.57	0.93
1966/67	0.630	3.76	0.97
1967/68	0.663	3.94	0.98
Total	4.375	31.76	8.03

Table 2.2./8 Long-distance calls and explanatory factors

Financial year		Telephone apparatus in million x_3	Degree of automazation x_4
1968/69		4.11	0.99
1969/70		4.29	0.99
1970/71		4.46	1.00
1971/72		4.63	1.00
1972/73		4.79	1.00
1973/74		4.96	1.00
1974/75		5.12	1.00

Table 2.2/9 Given forecasts for the explanatory factors:

From the figures for the period 1958-1968, we calculate the following totals:

$$\sum x_3^2 = 102.97 \qquad \sum x_3 x_4 = 26.13$$
$$\sum x_4^2 = 6.65 \qquad \sum y_2 x_3 = 14.55$$
$$\sum y_2 x_4 = 3.\ 71$$

The parameters *a,* b_3 and b_4 are estimated from the equation system:

$$a\,n + b_3 \sum x_3 + b_4 \sum x_4 = \sum y_2$$
$$a \sum x_3 + b_3 \sum x_3^2 + b_4 \sum x_3 x_4 = \sum y_2 x_3$$
$$a \sum x_4 + b_3 \sum x_3 x_4 + b_4 \sum x_4^2 = \sum y_2 x_4$$

or

$$10.00\ a + 31.76 b_3 + 8.03 b_4 = 4.38$$
$$31.76\ a + 102.97 b_3 + 26.13 b_4 = 14.55$$
$$8.03\ a + 26.13 b_3 + 6.65 b_4 = 3.71$$

The solution of the equation system is:

$a = -0.481$
$b_3 = 0.222$
$b_4 = 0.267$

Consequently, the regression function will be:

$y_2 = -0.481 + 0.222\, x_3 + 0.267\, x_4$

The result is compared with the given y_2 -values in the following table:

Financial year	y_2 = -0.481 + 0.222 x_3 + 0.267 x_4	Comparison with given values of y_2	
		from Table 2.2/8	Residual *r*
1958/59	0.223	0.235	+0.012
1959/60	0.271	0.268	-0.003
1960/61	0.319	0.315	-0.004
1961/62	0.366	0.351	-0.015
1962/63	0.412	0.400	-0.012
1963/64	0.458	0.445	-0.013
1964/65	0.510	0.500	-0.010
1965/66	0.560	0.568	+0.008
1966/67	0.613	0.630	+0.017
1967/68	0.656	0.663	+0.007

Table 2.2/10 Fitted values of long-distance calls *(y_2)*

Extrapolating these figures into the future, we find the following forecast for the number of long-distance calls in thousands of millions:

1968/69	0.695
1969/70	0.735
1970/71	0.776
1971/72	0.814
1972/73	0.887
1974/75	0.923

The result is given in Figure 2.2/4.

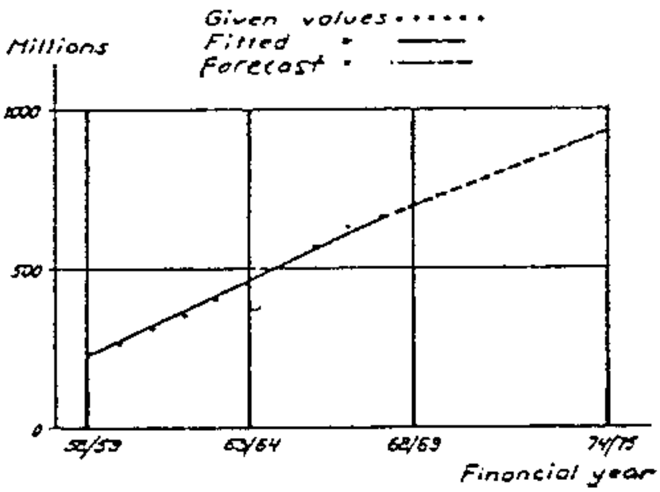


Figure 2.2/4

2.3 Trend methods

A time series is a set of observations of a quantity as a function of time. The observations may follow a certain pattern and the forecaster wants to find out how the observed quantity will develop in the future, i.e. to estimate the trend.

This section will deal with possible curves for development, such as linear trend, exponential trend and trends towards saturation level. The section on time series deals with methods to explain an observed variation and growth pattern as a result of four factors: trend, cyclical variations, seasonal variations and random variations.

If correct historical data are available, the trend methods can be used to forecast future devleopment on one important assumption: the same dependence between the variables will exist in the future. This is the same assumption as is made for all forecasts based on extrapolation of curves describing the past.

2.3.1 *Time series analysis*‘

Time series analysis is the name for a number of methods required to find an explanation for the fluctuations in the studied values (y_i).

Usually, the values of *y* are thought to be the result of four factors:

T. The trend T is the fundamental growth process (this growth can of course be negative or zero).

C. The cyclical variations C are sinus-like quantities with a very low frequency. The business cycles, seven fat years followed by seven lean years, are an example of this.

S. The seasonal variations S also oscillate, but with a higher frequency, and are generally more pronounced than the cyclical variations. Daily, weekly or yearly patterns are usually incorporated in S.

I. The irregular or random movements (I) are due to chance events such as strikes, earthquakes, changes in tariff, etc., and may be regarded as residual, unexplained by T, C and S.

In Figure 2.3/1, the effects of T, C and S are shown.

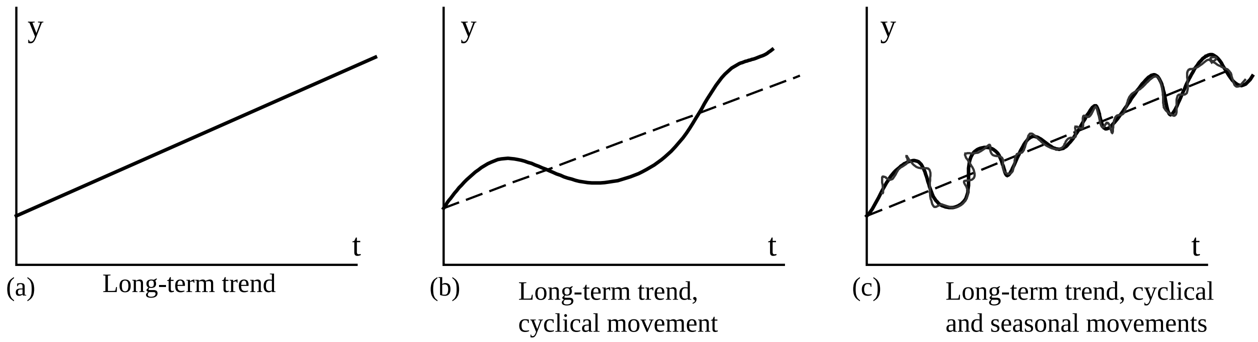


Figure 2.3/1
Influence of cyclical and seasonal variations on the trend, in an idealized time series

The first three faccctors are studied separately. Fact I cannot be expressed as a function of time and acts as a residual variation.

We can express the combined effect of the four factors as

$$Y = T + S + C + I \qquad (2.3.1)$$

or by multiplication, expressed by the formula

$$Y = T \times S \times C \times I \qquad (2.3.2)$$

Combined addition and multiplication is also possible, e.g.

$$Y = T + S \times C \times I \qquad (2.3.3)$$

It is usually sufficient for long-term forecasting to aggregate data into yearly figures and to attempt to fit a trend line through them.

A variety of mathematical curves can be fitted for analysing time series data ranging from freehand extrapolation of a graph of the series, fitting a simple straight line, fitting a straight line to transformed data series, to the use of a polynomial curves and on to still more complicated methods.

Linear trend

Any straight line giving the relationship between two variables can be expressed as:

$$y = a + bt \qquad (2.3.4)$$

When a and b have their values specified, the line is uniquely defined and the value of y can be calculated for any value of t. A number of pairs of recorded values of y and t are available for the forecaster. These values do not in general lie exactly along a straight line. His task is to estimate a and b. Assuming that n records are available, the values of a and b calculated under the least squares criterion are obtained by solving the following pair of simultaneous equations:

$$\sum y = na + b\sum t \qquad (2.3.5)$$

$$\sum t\,y = a\sum t + b\sum t^2 \qquad (2.3.6)$$

where n = number of pairs of values (t, y);

Method

Calculate the following values:

$$\sum ty = t_1 y_1 + t_2 y_2 +,\ldots,+ t_n y_n$$
$$\sum t^2 = t_1^2 + t_2^2 +,\ldots,+ t_n^2$$
$$\sum t = t_1 + t_2 +,\ldots,+ t_n$$
$$\sum y = y_1 + y_2 +,\ldots, y_n$$

n = number of pairs = number of t-values = number of y-values

Here: $\bar{t} = \frac{\sum t}{n}$ $\qquad$ $\bar{y} = \frac{\sum y}{n}$

Finally, calculate a and b for the equation system (2.3.5) and (2.3.6):

$$b = \frac{\sum ty - n\bar{t}\,\bar{y}}{\sum t^2 - n\bar{t}^2} \qquad (2.3.7)$$

and

$$a = \bar{y} - b\,\bar{t} \qquad (2.3.8)$$

For each value t, the y-value is given from:

$$y = a + b\,.\,t$$

which is the equation for the straight trend line we were seeking.

Exponential trend

Exponential growth can be expressed with the equation

$$y = a \cdot e^{b \cdot t} \qquad (2.3.9)$$

where y and t are the variables, and a and b are the parameters of the curve, and e = base for natural logarithms.

Taking the logarithm of the whole equation:

$$\lambda n \;\; y = \lambda n \;\; a + b \cdot t \qquad (2.3.10)$$

Introduce

$z = \lambda n \;\; y$ and $c = \lambda n \;\; a$

The equation can then be written

$$z = c + b \,.\, t \qquad (2.3.11)$$

We see that this equation is a straight line.

<u>Method</u>

Take the logarithm of all y-values and call these values z.

Then calculate:

$$\sum tz \qquad \sum t^2 \qquad \sum t \qquad \sum z$$

$$\bar{t} = \frac{\sum t}{n} \qquad \bar{z} = \frac{\sum z}{n} \qquad n \;\; \bar{t}^2$$

as before !

Then, we have:

$$b = \frac{\sum t \;\; z - n\bar{t}\,\bar{z}}{\sum t^2 - n\bar{t}^2} \qquad (2.3.12)$$

and

$$c = \bar{z} - b \;\; t$$

$$z = c + b \;\; t \qquad (2.3.13)$$

For each t, the value y-value is now calculated as

y = anti-logarithm of z (2.3.14)

<u>Checking the regression model</u>

A. Testing the significance of the time parameters

This regression approach depends on the assumption that the quantity under consideration (either a measured traffic or a calling rate) is correlated with time. This may or may not be true, the values may be quite independent of time. The formulae given above will, however, always produce a numerical value for b. It is, therefore, necessary to test if this value is arrived at by pure chance and if its true mean should be zero. This can be performed by calculating the value

$$H = \frac{b\sum(t-t)^2}{s} \qquad \text{where} \qquad s^2 = \frac{\sum y^2 - a\sum y - b\sum ty}{n-2} \qquad (2.3.15)$$

Calculate s^2 first, then calculate s = square root of s^2.

The notations above are valid for the straight line. For the exponential case, calculate s^2 as

$$s^2 = \frac{\sum z^2 - c\sum z - b\sum t \;\; z}{(n-2)} \qquad (2.3.16)$$

For accurate checking, it is necessary to use statistical tables, but as an approximate guide, *H* should be greater than 2 or less than - 2 for the linear description to be valid. If values in the range - 2 to + 2 are obtained, then the value of *b* is not significantly different from zero; meaning that the traffic quantity does not change systematically with time and the analytical description is inappropriate. The best forecast for future values of *y* is then no change at all.

B. *Examining systematic errors*

Even if the parameter *b* passes the above test, the model may be invalid due to the existence of systematic errors in the way the “best“ line is fitted to the data points. The following procedure should be used to check for systematic form of errors:

Denote the values on the fitted line corresponding to the recorded values of *y* by $\hat{y}$, and further denote the first readings by $y_1, \hat{y}_1$, the second by $y_2, \hat{y}_2$, and so on.

Calculate

$$w = \sum_{i=1}^{n-1} (y_i - \hat{y}_i)(y_{i+1} - \hat{y}_{i+1}) \qquad (2.3.17)$$

and $$v = \sum_{i=1}^{n} (y_i - \hat{y}_i)^2 \qquad (2.3.18)$$

Calculate the value $2 - 2\frac{w}{v}$

The statistic, known as the *Durbin-Watson statistic*, should lie between 1.5 and 2.5 and preferably between 1.7 and 2.3. If the value is outside this range, the data should be examined for indications of systematic deviation from the fitted curve. Common causes are either wrong curve shape (see Figure 2.3/2, Graph 3) or discontinuity in the data (see the same figure, Graph 4). If the *Durbin-Watson statistic* gives an unsatisfactory test result, suitable adjustments must be made to the data and the analysis be repeated. The forecasts may otherwise be subject to serious error.

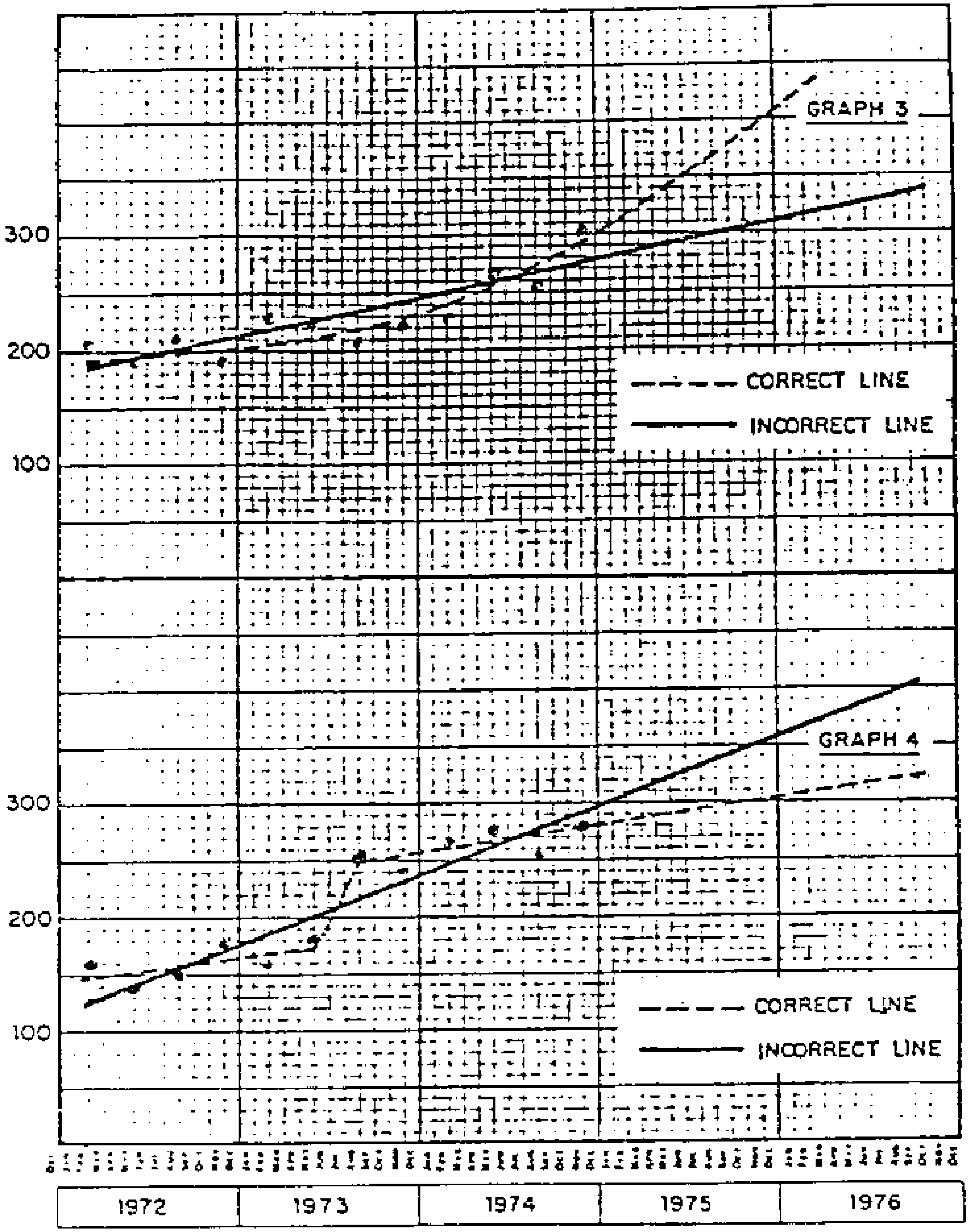

GRAPH 3
300
200
100
CORRECT LINE
INCORRECT LINE
GRAPH 4
300
200
100
CORRECT LINE
INCORRECT LINE
1972
1973
1974
1975
1976

Figure 2.3/2

Forecasting future levels

Assuming that the analytical description of the past successfully passes the tests described above, the next stage is to provide the required forecasts.

All that is required is to calculate the value of y from

$$y = a + bt$$

(If an exponential growth curve is fitted to the historical data, the actual forecast level is obtained by taking the natural anti-logarithm of the value of y.)

A general idea of the precision of the forecast can be obtained by designing a confidence interval for the forecast value, within which it can be stated with a *pre-determined degree of probability that the future value will be*.

For instance, if the forecast for a future time t_0 is y_o, then there is approximately a 95 % probability that the achieved value of y will be within the range.

$$y_o \pm 2\sqrt{u}$$

where $$u = s^2\left[1 + \frac{1}{n} + \frac{(t_o - \bar{t})^2}{\sum (t - \bar{t})^2}\right] \qquad (2.3.19)$$

and s^2 is calculated from (2.3.15).

It will be noticed (see Figure 2.3/3) that the further the forecast is made into the future, the wider the confidence interval, and accordingly the more uncertain the forecast.

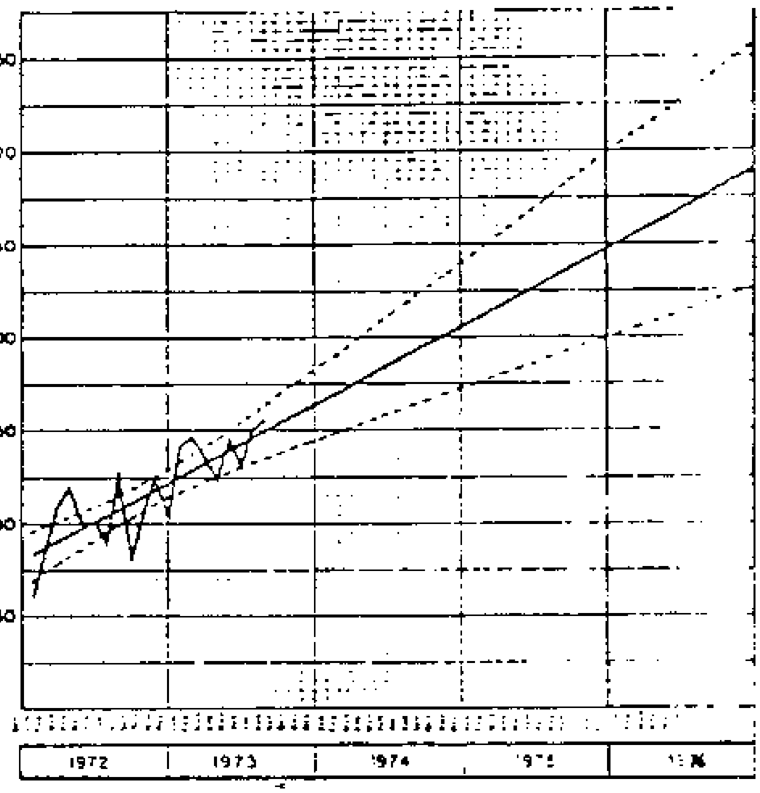


Figure 2.3/3 : Confidence interval (95 %) for forecasted linear development

If the quantities plotted on a graph show a consistent pattern over time, it is possible to draw a curve of best fit. Future values are then forecast by extrapolation of this curve.

The first problem is to find the equation for the curve which best describes the relationship between quantity and time.

For many types of technical equipment, the history of its usage can be described in three phases. In the first phase, we observe an accelerated usage which in the second phase changes to a linear increase over time. In the third phase, the usage slows down and eventually reaches saturation level. Examples of such phenomena may be the number of households with radios, television sets and telephones.

For other cases, the saturation level may become zero. Example: Use of horse-drawn cabs in London, use of the telegram service, use of conventional sailing ships for transporting goods, use of stone axes in battle, etc.

We will not deal with this latter type of development curve.

For a quantity that in the future may reach a saturation level, it may be possible to describe the whole history with one single mathematical expression. It may however be simpler and more accurate to describe each phase separately.

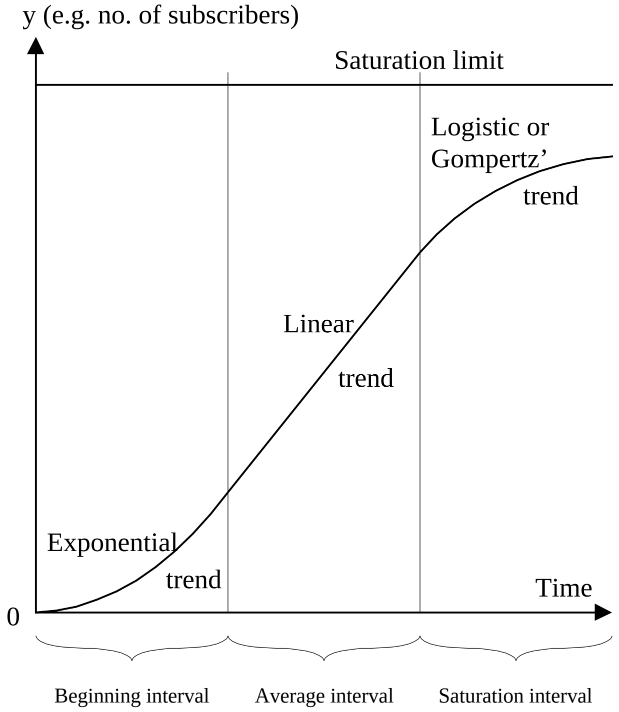


Figure 2.3/4
Development over time for a telecommunication service

Curves describing the development trend over time are often called growth curves, even if the "growth" is sometimes really a decrease in quantity. Here are some common types of trend curves:

Linear: $y = a + b\,t$ (2.3.20)

Parabolic: $y = a + b\,t + c\,t^2$ (2.3.21)

Exponential: $y = a\,e^{bt}$ (2.3.22)

Gompertz: $y = e^{a-br(t)}$ (2.3.23)

<u>Notations</u>

t = point of time (independent variable)

a, b, c, r = parameters to be calculated from historical data

y = item to be forecast (dependent variable)

e = the basis for natural logarithm

Some simple numerical examples are on trends:

1) Linear trend $y = a + b\,t$

The formula has the unknown parameters *a* and *b* which should be calculated from given historical data. To calculate two parameters, we need two equations.

These two equations might be obtained from two points in a diagram through which the straight line should pass.

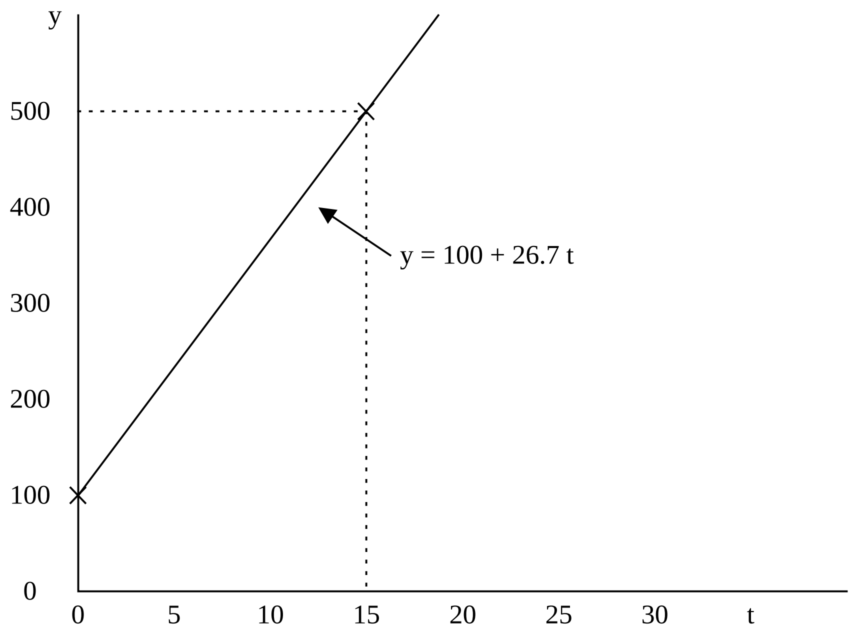


Figure 2.3/5
Linear trend, estimated from 2 points

Assume the following points: $t = 0$ $t = 15$
$y = 100$ $y = 500$

Insertion of the 2 points in the equation for the straight line (2.3.20) gives:

$100 = a + b \quad 0$

$500 = a + b\;15$

which gives $a = 100$ and $b = \frac{500 - 100}{15} = 26.7$

This gives the trend:

$y = 100 + 26.7\,t$

2) *Exponential trend* $y = a\,e^{bt}$

Assume the same points as in the previous case:

$t = 0 \quad t = 15$

$y = 100 \quad y = 500$

The two required equations will be

$100 = a\,e^{b0}$

$500 = a\,e^{b15}$ So: $a = 100$ and b is calculated from:

$500 = 100\,e^{b15}$

$5 = e^{b15}$

ln5 = 1.609 = 15b which gives $b = 0.1073$

The trend is: $y = 100\,e^{0.1073t}$

The curve is plotted in the diagram below:

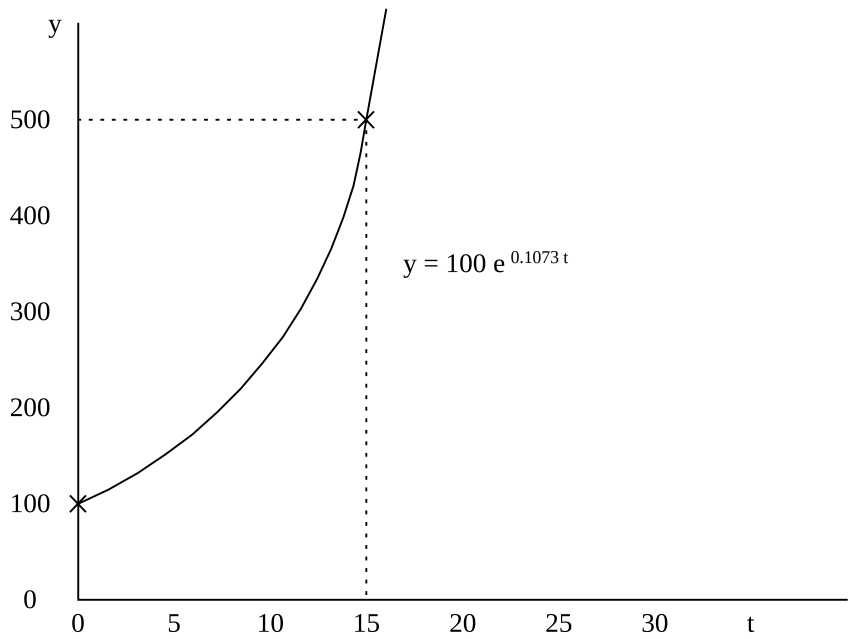


Figure 2.3/6
Exponential trend determined from 2 points

3) *Gompertz' trend* $y = e^{a-br^t}$

In this case, we have three parameters *a, b* and *r*, and the calculation requires three equations for determination of the parameters. We use the same two points as before:

$t = 0 \quad t = 15$

$y = 100 \quad y = 500$

A third equation can be obtained by assuming a ***saturation*** value in infinity, i.e. the point:

$t = \infty \quad y = 3000$ (i.e. the value 3000 will be reached after an infinite number of years)

The parameters are calculated from:

$100 = e^{a-b\ \ r^{0}}$ $500 = e^{a-b\ \ r^{15}}$

The third equation gives since $r < 1$ and t = ∞

$3000 = e^{a}$ therefore $a = 1n3000 = 8.006$

If $a = 8.006$ is inserted in the first equation,

$4.605 = 8.006 - b$; which gives $b = 3.401$

Then r is calculated from the second equation for

$a = 8.006$ and $b = 3.401$

$6.215 = 8.006 - 3.401\ r^{15}$

or $r = \left(\frac{8.006 - 6.215}{3.401}\right)^{1/15}$ $r = 0.958$

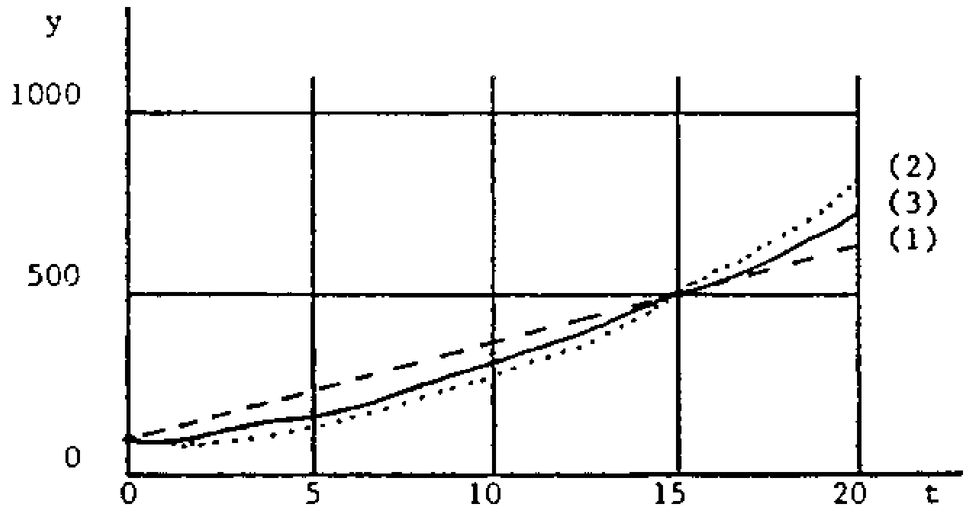


Figure 2.3/7 : Numerical examples on trends

1) Linear trend: $y = 100 + 26.\ 7t$

2) Exponential trend: $y = 100e^{0.1073t}$

3) Gompertz' trend: $y = e^{8.006 - 3.401(0.958)^{t}}$
(saturation value = 3000)

Other growth curves:

Consider

$$y = M - ae^{bt} \tag{2.3.24}$$

There are four configurations of this curve according to the signs of a and b. These are illustrated below.

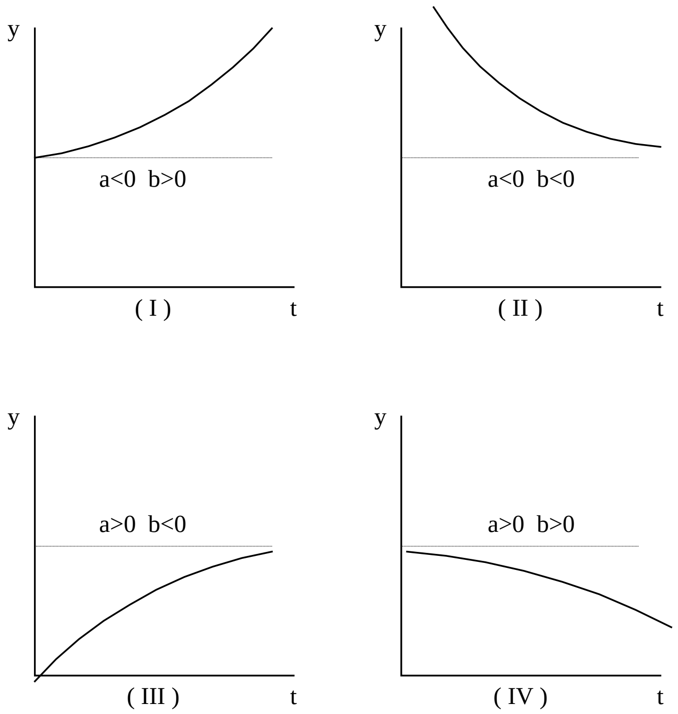


Figure 2.3/8 : Modified exponential curves

When forecasting demand for telecommunications, only (ii) and (iii) are applicable for long term forecasting, of which (ii) is less frequently used. A decline towards a minimum as shown in (ii) might be used for such variables as the number of manual calls per subscriber or the number of telegrams per capita. However, few telecommunication services are characterized by steady decline. Thus (iii) is most applicable. Of course, the implied approach to a saturation level as shown in (iii) may cover a period stretching over several decades. It may possibly be applied to forecast telephone density in countries with very high penetration, but is is inapplicable where telephone development is at an earlier stage. An alternative to the Gompertz' curve is the logistic curve, as described below.

<u>Logistic curve</u>

The simple logistic curve is given by the formula:

$$y = F + \frac{M - F}{1 + ae^{-bt}} \qquad (2.3.25)$$

The curve is *S*-shaped with a floor value *F*, and a maximum *M*. The curve increases to its maximum over time for $a > o, b > o.$

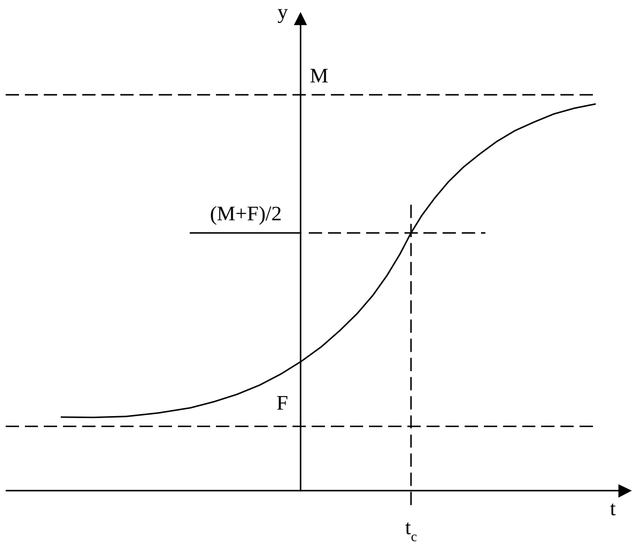


Figure 2.3/9 : Logistic curve

The curve has the following properties:

i. The growth increases steadily to a peak and then decreases.

ii. The point of inflexion at which the change in the absolute growth rate changes from an increase to a decrease occurs at a level midway between floor and ceiling, at time $t = t_c$, where t_c is given by 1/b 1n (1/a).

Comments

The purpose of a trend analysis may not always be to find out how the ***average*** for all traffic values changes with time. For dimensioning purposes, we are particularily interested in how the upper tail of the annual traffic distribution changes with time in order to obtain a representative value for the busy hour traffic. This busy hour traffic then determines the number of circuits and switches required in the future. To carry out a meaningful time series extrapolation, it is therefore essential that the data used really refers to values that are representative for our purpose. A set of historical data based on paid minutes per month, for example, would certainly not be worth the effort of any serious time series analysis. It is so because the relationship between the busy hour traffic and the monthly paid minutes is not generally established by experience but has to be calculated as a product of several inaccurate factors.

Another case where a time series analysis will not work is when the operational conditions have changed. For example, if the service is suddenly remarkably improved, the traffic ususally changes drastically. Such service improvement jumps may mean an increase in traffic of 100 % or more. To estimate the magnitude of such a sudden increase in traffic, information is required of the service performance before and after the change, how much traffic was suppressed before the change, etc. Such data are not generally available or are rather hard to collect.

2.4 *Individual judgment*

The mathematically simplest forecasting method is individual judgment. The forecast is based on experience and collected information. No systematic analysis is made.

However, the quality of a forecast is directly dependent on the quality of the basic data. Unless reliable basic data are available, the forecaster must establish an adequate data collection process before starting his forecasting work, or rather, to secure more reliable data for his future work.

A typical example of forecasts based on individual judgment is estimation of the future number of subscribers in a given area. These forecasts are generally related to existing and future buildings. Each type of buildings has its typical telephone density. The short-term forecast may concern existing buildings, buildings under construction and planned ones. The medium- and long-term forecasts may concern different types of areas as well, where areas with office buildings will require more telephones than residential areas. The forecaster may, as an aid, have tables showing the average number of subscribers in different types of buildings and areas. When carrying out this type of investigation, correct information from city planning authorities is essential, as well as past experience.

An example of the average number of subscribers for different types of buildings is shown in Table 2.4/1. The table should naturally be adjusted to agree with the situation in the country or towns concerned. For medium- and long-term forecasting, Table 2.4/2 shows typical values for different types of built-up area. The figures should also be adjusted here to agree with the country concerned. The figures given in the tables need to be updated when the telephone density increases.

Type of buildings	Number of subscribers
Official buildings and offices, banks, insurance companies, large hotels, clubs, large restaurants, hospitals, large commercial houses	investigate
Small hotels, restaurants, food shops, boarding houses	1 -2
Chemists, doctors, lawyers, etc.	1 - 1.5
Shops	0.5 - 1
Large factories	investigate
Small factories, workshops	0.5 - 1.5
Cinemas, petrol stations	1 - 2
Private houses of highest class	1
Private houses of lower class	0.3 - 0.5
Terraced houses	0.3 / apartment
Blocks of flats, highest class	0.5 - 1 / apartment
Blocks of flats, lower class	0.2 / apartment

Table 2.4/1 **Assessment of number of subscribers per different kind of buildings.**This example is valid for a developing country with 4-8 subscribers per 100 inhabitants.

Type of built-up area		Subscribers per hectare
A.	Slum	0.25
B.	Parks, gardens, etc.	0.50
C.	Old private houses with large gardens	1.00
D.	Poor workers’ residential districts	1.50
E.	Better class workers’ residential districts	2.00
F.	Modern private houses with large gardens	3.00
G.	Modern workers’ residential districts	4.00
H.	Industrial areas	5.00
I.	Modern private houses with small gardens	7.00
J.	Non-detached houses of older type	8.00
K.	Area consisting of working class dwellings and small workshops	10.0
L.	Modern non-detached houses	13.0
M.	Non-detached 1-2 storey residential buildings and small shops	18.0
N.	Blocks of flats up to 4 storeys	25.0
O.	Blocks of flats and shops of up to 4 storeys	28.0
P.	Business centre in residential area	30.0
Q.	Blocks of flats of more than 4 storeys	40.0
R.	Office buildings up to 3 storeys	80.0
S.	Office buidings of 4-6 storeys	150
T.	Office buildings of more than 6 storeys	250

Table 2.4/2 **Assessment of number of subscribers per hectare in built-up areas of different kinds.**

This example is valid for a developing country with 4-8 subscribers per 100 subscribers.

2.5 Other methods

The following is a survey of some other methods used for forecasting.

Analytical comparisons

This method relies on data from other, preferably comparable areas where the development is more advanced. The forecast is then made by comparison and assuming that the development will be the same.

This method is used when the series of adequate historical data is too short to permit reliable trend analysis. This is frequently the case in developing countries. It is then assumed that the development will follow the same pattern as in the more developed country or area.

Table 2.5/1 shows growth in telephone penetration over the period 1960-1975, expressed in the number of telephones per 100 inhabitants for certain countries. It also shows during which period the United Kingdom developed at the same rate and how many years it took there to go from the penetration of the country concerned in 1960 to that in 1975.

Country	Telephones per 100 inhabitants		Corresponding UK penetration	
	1960	1975	Dates	Period
Argentina	5.99	9.41	12 years	1937-1949
Chile	2.43	4.26	8 years	1923-1931
Costa Rica	1.34	5.02	25 years	1910-1935
Cyprus	2.95	10.65	25 years	1925-1950
Fiji	1.71	4.53	18 years	1914-1932
Greece	2.30	20.71	45 years	1922-1967
Ireland	5.13	12.78	20 years	1935-1955
Japan	5.21	37.88	41 years	1935-1976
Mexico	1.46	4.37	20 years	1911-1931
Spain	5.47	12.46	18 years	1936-1954

Table 2.5/1 Number of telephones per 100 inhabitants for certain countries.

Table 2.5/1 shows that there is no clear connection between the time it took for a particular country to develop from the 1960 value to the value of 1975, and the time that was needed in the United Kingdom to reach the same degree of development. It is therefore not recommended to compare countries with a too large difference in development. Table 2.5/2 gives the same information as the previous table but for a set of countries with approximately the same economic development.

Country	Telephones per 100 inhabitants		Corresponding dates for UK	
	1960	1975		
Australia	20.88	37.49	9 years	1969-1976
Denmark	22.17	42.48	10 years	1968-1978
Finland	12.89	35.78	20 years	1955-1975
Germany (F.R.)	9.98	30.25	24 years	1949-1973
Iceland	21.42	40.41	10 years	1967-1977
Netherlands	13.15	34.41	19 years	1955-1974
Norway	19.50	33.90	8 years	1966-1974

Table 2.5/2 Telephone penetration 1960-1975 for some countries with approximately the same economic development.

This table shows a better agreement with the development in the United Kingdom than the previous one.

We can now conclude the following:

1. Short-term projection of a country's telephone density from a given level, by comparison with the growth of some more advanced country being at the same level earlier, cannot always be recommended since special circumstances may cause development not to follow the same trend. Such circumstances are, e.g., economic recession or expansion, supply limitations, different tariff policies, etc.

2. For long-term projections, it is necessary to choose a country which is at least 15 years more advanced in order to have a series long enough for the possible future growth.

3. If such a comparison is carried through, however, we assume that a comparison between one country's performance over the last 15 years and another country's over the previous 15 years is valid. This is unlikely because of the rapid changes in the world economy and because of disruptive events such as wars and major recessions which are likely to distort the comparison. Data confirms this.

Further Methods:

The eye-ball method

The method is simply that the forecaster inspects the plotted data in order to get an idea of some possible trend or relationship between the variables and then manually tries to draw a corresponding curve.

Moving average

The new value is calculated as the mean of a number of observed values. The purpose of this approach is to reduce irregularities caused by, for instance, seasonal variations.

Exponential smoothing

Similar to the moving average method but weight is given to the observations in such a manner that more recent observations get more weight.

Market research

This method involves a study of the market by asking questions to potential customers.

Combined methods (including structural analysis)

A combination of different methods such as market research, growth curves, projection and econometric models.

3. FORECASTING FOR EXCHANGE-PLANNING

For exchange-planning, the demand should be forecast for all particular items for each exchange separately. This includes the following, per category of subscriber:

- the number of direct subscriber lines (DEL)
- originating traffic
- terminating traffic
- local traffic
- long-distance traffic
- international traffic
- special services

Once these forecasts have been made, corresponding figures from different exchanges can be added to produce results for greater areas or for the country as a whole. This is known as forecasting "from the botton up".

When a forecast for a an area shall be made, all information available about future changes in this area must be collected. The relevant information obtained in this way can make it possible to foresee discontinuities in the development of the demand. On the other hand, fluctuations of the quantities of interest are relatively great when we are looking at a small area. The number of reliable historic records is also usually very limited.

These two factors combine to make the small area forecast less reliable. Adding the results to obtain the corresponding figures for a wider area can lead to grave errors. To guard against this, the "from the *bottom up*" forecast is compared with a "from the *top down*" forecast.

Starting from the forecast for the country as a whole or for a large area, one can of course arrive at forecasts for smaller regions by splitting up the amounts forecasted. Usually, more extensive and reliable data are available for the forecasts of large areas. Therefore, the "from the *top down*" forecast is generally regarded as more reliable than the "from the *bottom up*" forecast.

There are two ways in which agreement between the two forecasts can be reached:

1. By changing the division of the total amount over the various areas.

2. By reducing the local forecasts by a given factor, either uniformly or after some system.

Clearly, this procedure of matching the two forecasts is rather subjective. It is also clear that an area which has submitted a high forecast will have to reduce its figures unless a satisfactory reason for this high forecast can be given. Such a reason may be a known development plan for this area, which motivates a sudden increase in the demand. The large-area forecast does not always take into account such local peculiarities.

4. FORECASTING FOR NETWORK PLANNING

4.1 *Introduction*

The planning of a telephone network is based on estimates of future traffic need.A *long-term forecast* is necessary for the development plan to ensure coordinated extension over a period of 15-25 years. This development plan has to be brought up to date approximately every 2-4 years.

Within the development plan, *short-term forecasts* are needed to provide the basic data for planning the actual extension steps. They should contain estimates of the traffic needs for four to six years ahead to cater for delivery times, etc. The short-term forecast should be brought up to date every year.

The traffic needs are expressed in erlang. To forecast these needs, the traffic is estimated between each pair of exchanges, generally separately for each direction. The traffic within each exchange area is also estimated.

4.2 *Traffic matrix*

To specify the traffic needs in a region with *n* exchanges, n^2 traffic values are required. A standard way to specify these traffics is to present them in a matrix, the so-called traffic matrix.

to from	1	i	j	n	SO
1	A(11)			A(1n)	O(1)
i		A(ii)	A(ij)		O(i)
j		A(ji)	A(jj)		O(j)
n	A(n1)			A(nn)	O(n)
ST	T(1)	T(i)	T(j)	T(n)	A(11)

Here:

A(ij) is the traffic from *i* to *j*;

A(ji) is the traffic from *j* to *i;*

A(ii) is the local traffic in exchange *i*;

O(i) is the sum of all traffic originating in *i*;

T(j) is the sum of all traffic terminating in *j*.

Adding all the row-totals O(i), i.e. the entries in column SO (sum originating traffic) give the total traffic A.. The same result is obtained by adding all the column-totals T(j), i.e. the entries in the row ST (sum terminating traffic). In short:

$$\sum_i O(i) = \sum_j T(j) = A..$$

As long as no confusion arises, the symbol *A(i, j)* can be used. However, it will frequently be necessary to distinguish between the present traffic from *i* to *j A(i, j/0)* and the estimated traffic at some future date *t*: *A(i, j/t)*. Then, of course,

$$O(i/t) = \sum_j A(ij/t) \qquad \text{and} \qquad T(j/t) = \sum_i A(ij/t)$$

4.3 *Point-to-point forecast*

Various methods exist to estimate *A(i, j/t)* based on the expected growth of the number of subscribers in areas *(i)* and *(j),* expected changes in traffic per subscriber, tec. The traffic matrix can then be completed by adding the entries row-wise to obtain the O(i/t) and column-wise to obtain the T(j/t).

For estimation of the future point-to-point traffics in a network, one usually bases the calculations on forecasted growth of subscriber lines and the present traffic matrix. Different formulae are used, of which the most common ones will be given below. It cannot be stated that one formula is more accurate than the other. Only feedback from future records may indicate which formula is the best for a particular case. When this has been found out, there is however no guarantee that this will be so for ever.

Estimation of total traffic

Taking into account that different categories of subscribers initiate different amounts of traffic, it may sometimes be possible to estimate a future traffic from:

$$A(t) = N_1(t) \cdot \alpha_1 + N_2(t) \cdot \alpha_2 + -- \qquad (4.3.1)$$

where $N_1(t)$, $N_2(t)$, etc., are the forecasted number of subscribers of category 1, 2, etc., and α_1, α_2, etc., are the traffic per subscriber of category 1, 2, etc.

If it is not possible to separate the subscribers into categories with different traffic, the future traffic may simply be estimated as:

$$A(t) = A(0) \frac{N(t)}{N(0)} \qquad (4.3.2)$$

where *N(t)* and *N(0)* are the number of subscribers at times *t* and zero.

Estimation of point-to-point traffic

For estimation of the traffic from one exchange to another, various formulae may be applied. The main idea is to take into account the increase of subscribers in the two exchanges and to apply certain weight factors to these growths.

$$A_{ij}(t) = A_{ij}(0) \frac{W_i G_i + W_j G_j}{W_i + W_j} \qquad (4.3.3)$$

where W_i and W_j are the weights and G_i is the growth of subscribers in exchange *i,* and G_j in exchange *j*.

$$G_i = \frac{N_i(t)}{N_i(0)} \qquad G_j = \frac{N_j(t)}{N_j(0)} \qquad (4.3.4)$$

Different methods exist for W_i and W_j .

Rapp's Formula 1

$$W_i = N_i(t) \qquad W_j = N_j(t) \qquad (4.3.5)$$

The assumption here is that the traffic per subscriber from exchange *i* to exchange *j* is proportional to the number of subscribers in exchange *j*.

Rapp's Formula 2

$$W_i = N_i(t)^2 \qquad W_j = N_j(t)^2 \qquad (4.3.6)$$

This formula assumes that the change of originated and terminated traffic per subscriber is as small as possible.

Australian Telecom Formula

$$W_i = \frac{N_i(0) + N_i(t)}{2} \qquad W_j = \frac{N_j(o) + N_j(t)}{2} \qquad (4.3.7)$$

This formula is a modification of Rapp's Formula 1.

A fourth formula is derived as from the following assumption: the traffic per one subscriber in exchange *i* to all subscribers in exchange *j* is constant.

$$\frac{A_{ij}(t)}{N_i(t) \cdot N_j(t)} = \frac{A_{ij}(0)}{N_i(0) \cdot N_j(0)} \qquad (4.3.8)$$
$$A_{ij}(t) = A_{ij}(0) \cdot G_i \cdot G_j$$

These four formulae can be further adjusted by introducing *N* in the weighting factors.

Gravity Model

The traffic between two exchanges can be expressed as:

$$A_{ij} = K(d_{ij}) \cdot N_i \cdot N_j \qquad (4.3.9)$$

where $K(d_{ij})$ = community of interest factor.

It has been found that this factor, to a certain degree, depends on distances between the exchanges d_{ij} We can, therefore, write $K(d_{ij})$ as follows:

$$K(d_{ij}) = e^{-\gamma d_{ij}} \qquad (4.3.10)$$

or $$K(d_{ij}) = d_{ij}^{-g} \qquad (4.3.11)$$

The value for the parameters γ or *g* can be calculated from a known traffic matrix. It may be necessary to adjust the expression for A_{ij} for pairs of exchanges with special relations to each other; for example, a big factory in one part of a country and the head office in another part.

The formulae given here are all deducted under certain assumptions. One assumption is that the traffic per subscriber remains constant, another is that a subscriber has the same interest to call every other subscriber. It is, however, found in developed countries that new subscribers may have less traffic than the old ones. This is partly taken into account in the Australian formula.

It is in principle impossible to say that one formula is more reliable than the other. The forecaster should try different formulae and try to find out which one gives the most credible result for his particular case.

4.4 *Kruithof's double factor method*

Kruithof's method enables us to estimate the future individual traffic values, $A(i, j)$ in a traffic matrix. The values, at present, are assumed to be known and so is the future row and column sums.

The procedure is to adjust the individual $A(i, j)$ so as to agree with the new row and column sums, i.e.

$A(i, j)$ is changed to $A(i,j)\frac{S_1}{S_o}$

where S_o is the present sum and S_1 the new sum for the individual row or column.

If we start with adjusting the $A(i, j)$, with regard to the new row sums, S_j, these sums will agree, but the column sums will not. Next step is then to adjust the found values of $A(i, j)$ to agree with the column sums. This makes the row sums disagree, so the next step is to adjust the new $A(i, j)$ to agree withthe row sums. The procedure is continued until sufficient accuracy for both column and row sums has been reached. The iteration is rather fast and gives in general a satisfying result after about three corrections.

The procedure is better understood from the numerical example given below.

The method is applicable when the change of proportions between the individual $A(i, j)$ is not expected to change much and in cases where it is not possible to forecast the individual $A(i, j)$ in other ways.

Example of the use of Kruithof's Double Factor Method

Consider a telephone network with two exchanges:

Given:

1. The present traffic interests $A_{ij}(0)$

i j	1	2	sum
1	10	20	30
2	30	40	70
sum	40	60	100

(4.4.4)

2. Forecast of the future total originating and terminating traffic per exchange:

$A_{i\cdot}(t)$ and $A_{\cdot j}(t)$:

i j	1	2	sum
1			45
2	?		105
sum	50	100	150

(4.4.5)

Problem:

Estimate the traffic values $A(i, j/t)$ with Kruithof's method.

Solution:

Iteration 1: Row multiplication.
$A_{i\cdot}$ is distributed as given by present traffic interest.

Result:

i j	1	2	sum
1	15	30	45
2	45	60	105
sum	60	90	150

$$A_{ij}(1) = \frac{A_{ij}(0)}{A_{i\cdot}(0)} A_{i\cdot}(t) \qquad (4.4.6)$$

After row multiplication, the sums of columns differ from the forecast. Next iteration will be column multiplication.

Iteration 2: Column multiplication.
$A_{\cdot j}(2)$ is distributed as given by iteration 1.

Result:

i j	1	2	sum
1	12.5	33.33	45.83
2	37.5	66.67	104.17
sum	50	100	150

$$A_{ij}(2) = \frac{A_{ij}(1)}{A_{\cdot j}(0)} A_{\cdot j}(t) \qquad (4.4.7)$$

After column multiplication, the sums of rows differ from the forecasted values. Next iteration will be row multiplication.

Iteration 3: Row multiplication.
$A_{i\cdot}$ is distributed as given by iteration 2.

Result:

i j	1	2	sum
1	12.27	32.73	45
2	37.80	67.20	105
sum	50.07	99.93	150

$$A_{ij}(3) = \frac{A_{ij}(2)}{A_{i\cdot}(0)} A_{i\cdot}(t) \qquad (4.4.8)$$

Iteration 4: Column multiplication.
$A_{\cdot j}$ is distributed as given by iteration 3.

Result:

i j	1	2	sum
1	12.25	32.75	45
2	37.75	67.25	105

sum	50	100	150

$$A_{ij}(4) = \frac{A_{ij}(3)}{A_{.j}(0)} A_{.j}(t) \qquad (4.4.9)$$

After 4 iterations, the sums of rows and columns are equal to the forecasted values. We can now write:

$$A_{ij}(t) = A_{ij}(4) \qquad (4.4.10)$$

5. CONCLUSIONS

This chapter has been concerned mainly with the mathematics behind the forecasting methods. This will make it possible for the reader to be able to assist in the forecasting work, especially as concerns preliminary treatments of historical data.

Besides explaining when and why forecasts are needed, this chapter contained some frequently used methods of describing historical data in mathematical formulae. It should be emphasized again that the *treatment of historical data is not forecasting*. It is only treatment of historical data. Forecasting, on the other hand, is the art of guessing future development. Only in certain cases for short-term purposes will extrapolation of curves, obtained from historical data, provide reliable forecasts. In most cases, it must be judged if the future conditions and assumptions are the same as for the past. Adjustments of extrapolated values may be necessary for various reasons. Improved service performance may double the traffic within a month or two. New activities in society (e.g. industries) may cause drastic changes in traffic dispersion, etc.

Forecasting generally concerns the *demand* for subscriptions and traffic caused by various telecommunication services. Since demand quantities are difficult or impossible to measure, the historical data given to the forecaster are generally measured quantities, such as carried traffic. Therefore, the forecaster should first of all transfer the given data into demand data. If they concern a country and a service where the demand is well satisfied, the demand may be estimated with acceptable accuracy.

If the demand is far greater than the supply, it is very difficult to estimate the real demand. The forecaster has to take this difficulty into account as long as conditions do not permit a balance between supply and demand.

Another difficulty in using historical data is that they may not be sufficiently reliable to permit an analysis. One reason for this is that the data have been collected on different principles, for different purposes and at different times. It may even happen that there are no historical data available at all. Forecasting must then be based on other indirect indications, which does not generally improve the reliability of the forecast.

6. REFERENCES

BEAR, D.	Some Theories on Telephone Traffic Distribution; A critical survey (ITC-paper, Stockholm, 1973)
LEIJON, K.H.	Introduction to Practical Teletraffic Engineering. Section F. Forecasting (ITU) REM/72/038, Benghazi, 1976)
CCITT (GAS 5)	Methods Used in Long-term Forecasting of Domestic Telecommunication Demand and Required Resources (1980)
LONNQUIST, I.	Theory and Methods for Traffic Forecasts (TELE, Stockholm, 1968)
LONNQUIST, I.	Forecasting. Teletraffic Engineering Seminar, Istanbul, 1980 (ITU)
LONNSTROM, M.M. MARKLUND, F. MOO, I.	A Telephone Development Project, Stockholm, 1967
NYGAARD, J.	An Integrated Telephone Traffic Forecating Tool for Use in Planning, I. (Teleteknik, Vol. XIV, 1970, No. 1)
RAPP, Y.	Some Economic Aspects on the Long-term Planning of Telephone Networks (Stockholm, 1969)
TURNER, W.M.	Forecasting. ITU/SIDA Seminar, New Delhi, 1975. SEM/ASIA/75

Note: This chapter has been edited from a manuscript prepared by K.H. Leijon, L.M. Ericsson, Sweden.

7. EXERCISES

FORECASTING TECHNIQUES

An object is to be forecasted. The historical development was as follows:

Year (at the end of)	Time Scale (t)	Stock (y)	Increase Absolute	%
1968	1	583		
1969	2	615	32	5.5
1970	3	646	31	5.0
1971	4	697	51	7.9
1972	5	738	41	5.9
1973	6	802	64	8.7
1974	7	844	42	5.2

Make a forecast for the expected stock 5 and 10 years ahead, in four different ways:

a. Assuming an unchanged and constant *absolute* increase per year, based on the historical average increase.

b Assuming an unchanged and constant *percentage* increase per year, based on the total increase over the past six years.

c. By adapting a linear trend line to the historical data.

d. By adapting an exponential trend to the historical data.

Plot the historical data in a diagram and show in the same diagrams the two trend lines c) and d) expanded to year 16 (6 + 10), and also the two other forecasts a) and b).

Compare the outcome from the forecasts.

Check the trend lines as concerns statistical reliability.
(Correlation coefficient, T - test and Durbin - Watson Test)

Discuss the credibility of the forecasts.

2. Calculate the overall connection forecast for a local telephone network for 1983 (31.3.83)

Given:

GROWTH CHART OF DEMANDS IN A LOCAL TELEPHONE NETWORK

As on	Equipped capacity	Working connect	Waiting list	Total demand
31/3/72	11800	9825	1901	11,726
31/3/73	12000	11114	1781	12,925
31/3/74	12100	11458	3555	15,013
31/3/75	12500	11530	5106	16,636
30/9/75	12500	11653	5662	17,315
31/3/76	12800	12275	2587	14,862
31/3/77	14150	13683	1624	15,307
31/3/78	15500	14437	1893	16,330

3. Forecast the total originating traffic for December 1986 in a local telephone exchange.

Given:

a. Traffic records for January 1979 - December 1981 (Table A)

b. Calling rate records for January 1979 - December 1981 (Table B)

c. Connection forecast for December 1986 = 2969
(1170 subscribers in January 1979 and 1600 in December 1981)

A. TOTAL ORIGINATING TRAFFIC

MONTH	1979	1980	1981
JANUARY	38.6	39.4	45.6
FEBRUARY	37.9	43.7	46.2
MARCH	42.1	48.7	47.2
APRIL	40.6	43.8	46.2
MAY	40.1	40.2	45.6
JUNE	38.1	42.6	48.5
JULY	37.7	41.1	44.4
AUGUST	39.9	44.2	47.4
SEPTEMBER	40.4	41.0	49.1
OCTOBER	40.7	43.8	48.7
NOVEMBER	40.8	41.8	45.0
DECEMBER	42.2	49.5	49.5

B. OVERALL ORIGINATING CALLING RATES (Erlang/subscriber)

MONTH	1979	1980	1981
JANUARY	.033	.031	.033
FEBRUARY	.033	.034	.033
MARCH	.036	.038	.034
APRIL	.035	.033	.033
MAY	.034	.030	.032
JUNE	.032	.032	.035
JULY	.031	.031	.032
AUGUST	.033	.032	.034
SEPTEMBER	.034	.030	.035
OCTOBER	.034	.032	.032
NOVEMBER	.033	.030	.028
DECEMBER	.034	.036	.031

4. Consider a telephone network with two exchanges.

Given:

a. The present traffic interest $[A_{ij}(0)]$

j / i	1	2	0_2 (o)
*	10	20	30
2	30	40	70
T_j (o)	40	60	100

b. Forecasted values on the future total originating and terminating traffics per exchange [0; (t) T; (t)]

j / i	1	2	o_i (E)
1	?	?	45
2	?	?	105
T_j (t)	50	100	150

Estimate the traffic values $A_{ij}^{(t)}$ by using Kruithof's method.

5. Calculate the traffic interest between two exchanges in a local area.

Given:

a. A local area is divided into *traffic areas* No. 1, 2, 3, 4, ... The future traffics between all *traffic areas* are forecasted.

b. The local area is to be divided into exchange areas. The *exchange areas* do not coincide with the *traffic areas*.

c. Try to calculate the expected future traffic from exchange *A* to exchange *B*.

d. We have the following information:

Exchange *A* will get

- 5000 subscriber lines from traffic area I, which totally has 10,000 subscriber lines.

- 8000 subscriber lines from traffic area II, which totally has 12,000 subscriber lines.

Exchange *B* will get

- 9000 subscriber lines from traffic area III, which is all of the subscriber lines that are there.

- 2000 subscriber lines from traffic area IV, which totally has 6000 subscriber lines.

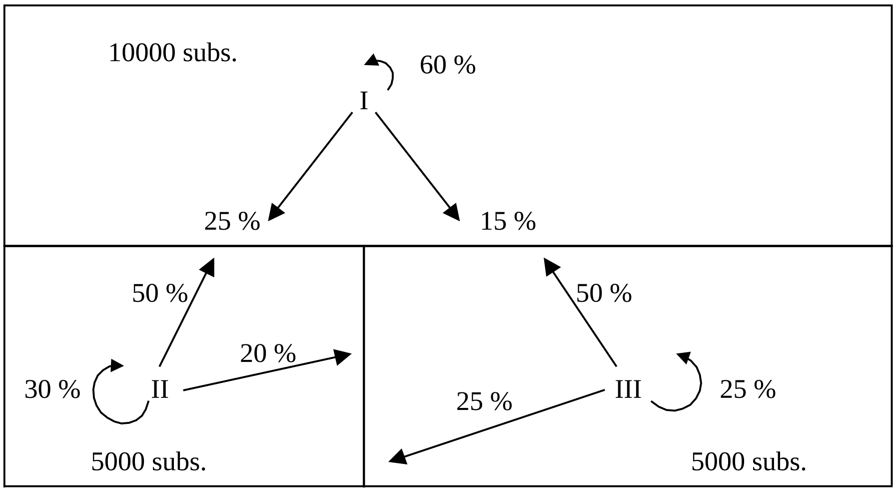
e. We also know from the forecast the total expected traffic from each one, to each one of the involved traffic areas, as following:

From traffic area No.	To traffic area No.	Total expected traffic
I	III	100 erl.
I	IV	90 erl.
II	III	105 erl.
II	IV	95 erl.

6. Estimation of traffic interests between exchanges: The objective of the example is to illustrate how a traffic matrix for exchanges can be computed from a given traffic matrix for *traffic areas*.

Given:

An area consisting of three traffic areas: I, II, III:



Traffic Area 1

No. of subscribers		10,000
Total originated traffic/sub	0.06 erl.	
Distribution of this traffic	60 %	I
25 %	II	
15 %	III	

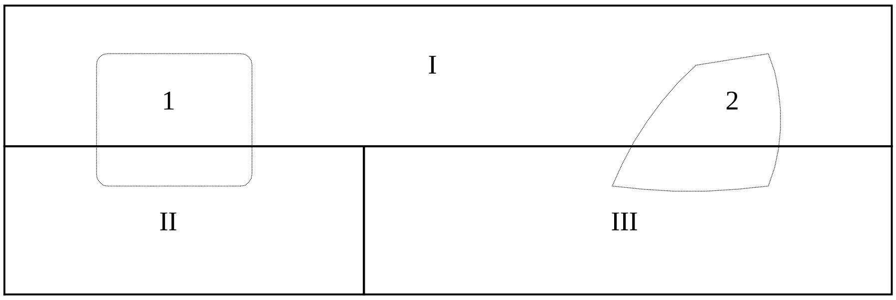
Traffic Area II

No. of subscribers		5,000
Total originated traffic/sub	0.05 erl.	
Distribution of this traffic	50 %	I
30 %	II	
20 %	III	

Traffic Area III

No. of subscribers		5,000
Total originated traffic/sub	0.04 erl.	
Distribution of this traffic	50 %	I
25 %	II	
25 %	III	

The area is served by a number of exchanges 1, 2, ... etc.



Exchange 1

Total No. of subscribers 8,000

No. of subscribers belonging to traffic area I 5,000

No. of subscribers belonging to traffic area II 3,000

Exchange 2

Total No. of subscribers 6,000

No. of subscribers belonging to traffic area I 4,000

No. of subscribers belonging to traffic area III 2,000

Task:

Calculate the expected total traffic flow from exchange 1 to exchange 2.

Hint:

Start with calculation of traffic from *one* subscriber in traffic area I to *one* subscriber in traffic area II, etc.

7. Present traffic matrix $A_{ij}^{(0)}$ has been estimated:

To From	Exchange No. j 1	 2	 3	 Sum
1	25	30	45	100
2	35	55	110	200
3	60	85	155	300
Sum	120	170	310	600

The number of main lines per exchange in year *t* has been forecasted:

Exchange No.	N_i (0)	N_i(t)
1	2000	3000
2	3500	3500
3	6800	7500

The main lines have not been classified into different categories since the proportion of different subscribers is expected to be the same in the future.

The total originating and terminating traffic per exchange is therefore forecasted from the formulae:

$$A_{i.}^{(t)} = N_i^{(t)} \frac{A_i^{(0)}}{N_i^{(0)}}$$

$$A_{.j}^{(t)} = N_j^{(t)} \frac{A_{.j}^{(0)}}{N_j^{(0)}}$$

Exchange No.	$A_i^{(t)}$	$A_j^{(t)}$
1	150.0	180.0
2	200.0	170.0
3	331.9	341.9
Sum	681.9	691.9

Since the sum of $A_{i.}^{(t)}$ and $A_{.j}^{(\ )}$ differ, we can use the mean value of these sums as an estimate of $A_{..}^{(t)}$ and adjust the $A_{i.}^{(t)}$ and $A_{.j}^{(\ )}$. This will give:

Exchange No.	$A_i^{(t)}$	$A_{.j}^{(t)}$
1	151.1	178.7
2	201.5	168.8
3	334.3	339.4
Sum	686.9	686.9

a. Now draw a traffic matrix for the future year *t* and fill in the total traffics calculated above.

b. Calculate the different point-to-point traffics by using the weighted growth methods, from known data about present and future number of main lines and the present point-to-point traffics. The type of weights is your own choice.

c. Fill in these values in a new traffic matrix and calculate the sums of columns and rows. You will see that these sums do not agree with the values in the matrix you first drew.

d. If we consider the traffic per main line to be constant during the forecast period, then the obtained point-to-point traffics must be corrected so that their sums agree with the forecasted total traffics.

If time permits it, do that correction by using Kruithof's Double Factor method.

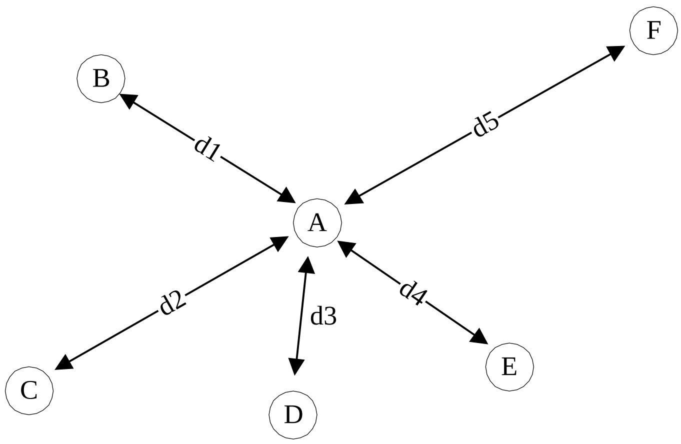
8. Suppose that a busy hour trunk traffic forecast for exchange "A" of 75 erlangs is to be distributed over 5 routes to exchanges "B", "C", "D", "E" and "F", and for which no traffic data is available. Suppose also that the connection forecasts for the exchanges are:

A	10 000	(c_o)
B	5 000	(c_1)
C	7 000	(c_2)
D	4 000	(c_3)
E	2 000	(c_4)
F	10 000	(c_5)

And that the distances from "A" are:

B	20 miles (d_1)
C	30 miles (d_2)
D	10 miles (d_3)
E	10 miles (d_4)
F	50 miles (d_5)

The situation is illustrated in the diagram below.



Estimate the distribution of the total trunk traffic over the five different routes *A* to *B*, *A* to *C*, ..., etc.

Hint:

The traffic forecast for a route can be obtained from

$$t_i = \frac{\frac{c_o c_i}{d_i^2}}{\sum_{j=1}^{n} \frac{c_o c_j}{d_j^2}}$$

where *T* is the total traffic.