CHAPTER V

MEDIUM TERM AND LONG TERM PREDICTION MODELS OF ANNUAL INSTALLED PLANT CAPACITY AND CONSUMPTION OF ELECTRICAL ENERGY BY COMPUTER-AIDED SELF-ORGANISATION OF MATHEMATICAL MODELS

CHAPTER V

MEDIUM TERM AND LONG TERM PREDICTION MODELS OF ANNUAL INSTALLED PLANT CAPACITY AND CONSUMPTION OF ELECTRICAL ENERGY BY COMPUTER-AIDED SELF-ORGANISATION OF MATHEMATICAL MODELS

5.0 Introduction

This work relating to the short term and long term forecasting models of annual installed plant capacity and consumption of electrical energy of India has been divided into four parts. In the first part an annual model of installed capacity of electrical energy for medium term (6 to 7 years) prediction has been obtained. Different types of polynomials of increasing complexity have been tested. The polynomial which gives minimum of a selection criterion has been found. In the second part assuming an annual growth rate of installed capacity of 8 % a long term prediction model of installed plant capacity of electrical energy has been obtained.

In the third part assuming an annual growth rate of electrical energy consumption of 8 % a long term prediction model of annual energy consumption has been obtained. The annual plant load factor has been found to have a distinct preiodicity.

In the fourth part a polynomial model of annual load factor has been obtained incorporating periodic terms.

With the theory of self-organisation commonly known as group method of data handling it has been possible to formulate mathematical models for complex processes with prediction optimisation.

The concept of self-organisation can be illustrated as follows :

When the model complexity gradually increases the computer finds by shifting the different models, the minimum of a selection criterion which the computer has been asked to look for. Thus the computer indicates to the operator the model of optimum complexity.

5.1 Medium-term prediction model of annual installed plant capacity of electrical energy

We have annual installed plant capacity data from 1961 - 1981. Power density spectra versus cycle per annum characteristic [31] of the data does not show any harmonicity in the process. So it is obvious that the process does not contain any sinusoidal harmonic parts. The correlation co-efficients versus shift of instances of time (equation 5.2.2) show that the current year installed plant capacity is strongly correlated with the past three years installed plant capacity.

Thus the process is assumed to have a finite difference form structure of the following nature.

$$Y_{k+1} = f\ (Y_k,\ Y_{k-1},\ Y_{k-2},\ t_k\ ) \qquad (5.1.1)$$

We write,

Y_{k+1} = y, installed plant capacity of electrical energy for the k+1 - th year.

Y_k = x_1, installed plant capacity of electrical energy for the k - th year.

Y_{k-1} = x_2, installed plant capacity of electrical energy for the k-1 - th year.

Y_{k-2} = x_3, installed plant capacity of electrical energy for the k-2 - th year.

$t_k = x_4$, time instant for the k-th year

$$\text{so, } y = f\ (x_1,\ x_2,\ x_3,\ x_4\ ) \qquad (5.1.2)$$

The function $f(.)$ is sought in the class of quadratic polynomials on the basis of a Table of polynomial of gradually increasing complexity of four varieblee as shown in Table 5.1.1, with the help of the theory of the self-organisation of combinatorial group method of data handling algorithm. The model of optimum complexity is selected on the basis of the minimum of the integral square error criterion.

TABLE 5.1.1

GRADUALLY INCREASING COMPLEXITY OF POLYNOMIALS FOR FOUR VARIABLES

General form of polynomial :

$$y = a_0 + a_1X_1 + a_2X_2 + a_3X_1X_2 + a_4X_3 + a_5X_1X_3 + a_6X_2X_3 + a_7X_4 + a_8X_1X_4 + a_9X_2X_4 + a_{10}X_3X_4$$

X_1	X_2	X_1X_2	X_3	X_1X_3	X_3X_4
$y_1 = a_0 + a_1X_1$	$y_2 = a_0 + a_2X_2$	$y_4 = a_0 + a_3X_1X_2$	$y_8 = a_0 + a_4X_3$	$y_{16} = a_0 + a_5X_1X_3$	$y_{512} = a_0 + a_{10}X_3X_4$
	$y_3 = a_0 + a_1X_1 + a_2X_2$	$y_5 = a_0 + a_1X_1 + a_3X_1X_2$	$y_9 = a_0 + a_1X_1 + a_4X_3$	$y_{17} = a_0 + \cdots$	$y_{513} = a_0 + a_1X_1 + a_{10}X_3X_4$
		$y_6 = a_0 + a_2X_2 + a_3X_1X_2$	...	...	...
		$y_7 = a_0 + a_1X_1 + a_2X_2 + a_3X_1X_2$	...	...	...
			$y_{15} = a_0 + a_1X_1 + a_2X_2 + a_3X_1X_2 + a_4X_3$	...	$y_{1023} = a_0 + a_1X_1 + a_2X_2 + a_3X_1X_2 + a_4X_3 + a_5X_1X_3 + a_6X_2X_3 + \cdots + a_{10}X_3X_4$ *

* General rule of total number of combination is $2^n - 1$.
Where n is the number of terms containing the variables in the general form of polynomial.

The integral square error is defined as

$$I.S.E. = \frac{\sum_{i=1}^{N} (Y_{tab}(i) - Y_{d.m}(i))^2}{\sum_{i=1}^{N} (Y_{tab}(i))^2} \qquad (5.1.3)$$

where $Y_{tab}(i)$, $i = 1,2,\ldots, N$ years are the tabulated values of the variables in the interpolation region and $Y_{d.m}(i)$ are the values of the variable obtained from the model. The time instances t_k is taken as $1,2,\ldots,$ $k = 1961, 1962,\ldots$ and so on. The models from the Table 5.1.1 comprising of four variables are tested for all the data points.

The model of annual installed plant capacity of electrical energy in MW for India is obtained as

$$\begin{aligned} y = {} & 5584.157096 + 0.599659\ x_1 \\ & -0.330972\ x_2 - 0.185849\ x_3 \\ & +770.911284\ x_4 + 1.05645165\ E\text{-}05\ x_1 x_2 \\ & -8.57186948\ E\text{-}09\ x_1 x_3 \end{aligned} \qquad (5.1.4)$$

In finite difference form,

$$\begin{aligned} Y_{k+1} = {} & 5584.157096 + 0.59965\ Y_k \\ & - 0.330972\ Y_{k-1} - 0.185849\ Y_{k-2} \\ & + 770.911284\ t_k + 1.05645165\ E\text{-}05\ Y_k Y_{k-1} \\ & - 8.57186948\ E\text{-}09\ Y_k Y_{k-2} \end{aligned} \qquad (5.1.5)$$

The corresponding minimum value for the integral square error and the mean error are 9.366641 E - 04 and 8.6845 E - 06 respectively.

The errors between the observed and the modelled values are found to be almost uncorrelated for $k \neq j$. Fig. 5.1.1a and 5.1.1b shows the observed and errors between the observed and the modelled values respectively. Fig. 5.1.1a has been extrapolated for seven years of prediction of the installed plant capacity of electrical energy i.e., upto 1988.

5.2 Long term prediction model of installed plant capacity

During the period from 1970 - 80 the installed plant capacity of electrical energy has grown at an average annual rate of 7.2 %. Considering the deleterious impact of power shortage on the productive sectors, both industry and agriculture, of the economy the Planning Commission of the Govt. of India has suggested an average annual growth rate of 11.3 per cent during the Sixth Five Year Plan period (1980 - 85). The growth rate has been suggested on the assumption that a distinct improvement in the working of power plants and strict adherence to the working schedules of power projects. Over the years the trend has been a

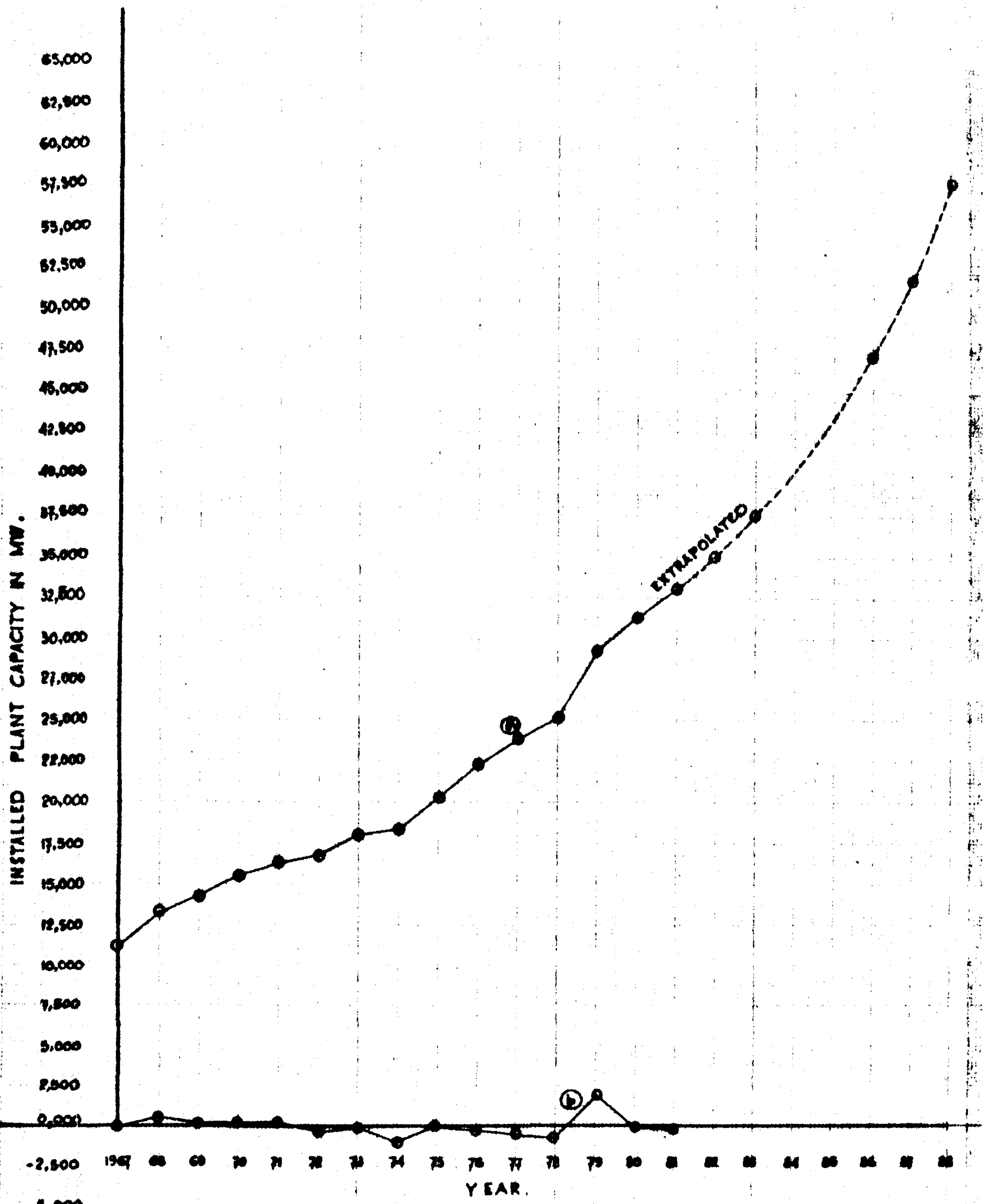


FIG 5.1.1a & b ANNUAL INSTALLED PLANT CAPACITY EXTRAPOLATED FROM 1982-88 IN MW AND ERRORS BETWEEN THE OBSERVED AND MODELLED VALUES, RESPECTIVELY.

mounting negligence and an appalling backlogs. The plant load factor has been moving in the vicinity of 45 per cent. It is now suggested that an annual growth rate of installed plant capacity of 8 per cent would be a realistic estimate. Because of the interactions of different deep lying feedback paths it may be recognised that this growth rate is a complicated process and is likely to change slowly rather than quickly.

On the basis of an annual growth rate of 8 per cent the installed plant capacity of electrical energy has been extrapolated upto 4 times the 1981 figures.

A mathematical model in the form of a finite difference equation has been postulated as

$$Y_{k+1} = f\,(Y_k, Y_{k-1}, Y_{k-2}, Y_{k-3}, t_k) \qquad (5.2.1)$$

The four arguments Y_k, Y_{k-1}, Y_{k-2}, Y_{k-3} are selected because of their strong correlation on Y_{k+1}. The correlation co-efficient at a shift instance λ is defined as

$$\phi_{YY}(\lambda) = \frac{\sum_{i=1}^{N-\lambda} \left[\left(Y_i - \frac{1}{N}\sum_{k=1}^{N} Y_k\right)\left(Y_{i+\lambda} - \frac{1}{N}\sum_{k=1}^{N} Y_k\right) \right]}{\left[\sum_{i=1}^{N-\lambda} \left(Y_i - \frac{1}{N}\sum_{k=1}^{N} Y_k\right)^2 \sum_{j=1+\lambda}^{N} \left(Y_j - \frac{1}{N}\sum_{k=1}^{N} Y_k\right)^2 \right]^{\frac{1}{2}}} \qquad \ldots (5.2.2)$$

Polynomials of gradually increasing complexity of five variables are shown in Table 5.2.1.

The model of optimum complexity which gives minimum of integral square criterion is found to be

$$
\begin{aligned}
Y_{k+1} = & - 232.764083 + 0.998565\, Y_k \\
& + 0.167739\, Y_{k-1} - 0.020582\, Y_{k-2} \\
& + 0.086204\, Y_{k-3} - 171.471615\, t_k \\
& - 6.719787\, E - 07\, Y_k\, Y_{k-1} \\
& + 3.512649\, E - 11\, Y_k\, Y_{k-2}
\end{aligned}
\qquad (5.2.3)
$$

The corresponding values of the integral square error and the mean error are 1.284107 E - 04 and 1.512884 E - 05.

The errors are found to be almost uncorrelated for $k \neq j$.

Fig. 5.2.1a and 5.2.1b shows the extrapolated annual installed plant capacity upto 1998 and errors of modelling respectively.

5.3 Long term prediction model of annual electrical energy consumption

On the basis of an annual growth rate of 8 per cent the electrical energy consumption has been extrapolated to 5 times its 1981 consumption figures. It has been observed

TABLE 5.2.1

GRADUALLY INCREASING COMPLEXITY OF POLYNOMIALS OF FIVE VARIABLES

General form of polynomial :

$y = a_0 + a_1x_1 + a_2x_2 + a_3x_1x_2 + a_4x_3 + a_5x_1x_3 + a_6x_2x_3 + a_7x_4 + a_8x_1x_4 + a_9x_2x_4 + a_{10}x_3x_4 + a_{11}x_5 + a_{12}x_1x_5 + a_{13}x_2x_5 + a_{14}x_3x_5 + a_{15}x_4x_5$

x_1	x_2	x_1x_2	x_3	x_1x_3	...	x_3x_4	...	x_4x_5
$y_1 = a_0 + a_1x_1$	$y_2 = a_0 + a_2x_2$	$y_4 = a_0 + a_3x_1x_2$	$y_8 = a_0 + a_4x_3$	$y_{16} = a_0 + a_5x_1x_3$	...	$y_{512} = a_0 + a_{10}x_3x_4$	...	....
	$y_3 = a_0 + a_1x_1 + a_2x_2$	$y_5 = a_0 + a_2x_2 + a_3x_1x_2$	$y_9 = a_0 + a_1x_1 + a_4x_3$	$y_{17} = a_0 +$		$y_{513} = a_0 + a_1x_1 + a_{10}x_3x_4$		
		$y_6 = a_0 + a_2x_2 + a_3x_1x_2$	⋮			⋮		⋮
		$y_7 = a_0 + a_1x_1 + a_2x_2 + a_3x_1x_2$	$y_{15} = a_0 + a_1x_1 + a_2x_2 + a_3x_1x_2 + a_4x_3$			$y_{1023} = a_0 + a_1x_1 + a_2x_2 + a_3x_1x_2 + a_4x_3 + a_5x_1x_3 + \dots + a_{10}x_3x_4$		$y_{32767} = a_0 + a_1x_1 + a_2x_2 + a_3x_1x_2 + \dots + a_{15}x_4x_5$

The General rule of total number of combination is $2^n - 1$ where n is the total number of terms containing the variables in the general form of the polynomial.

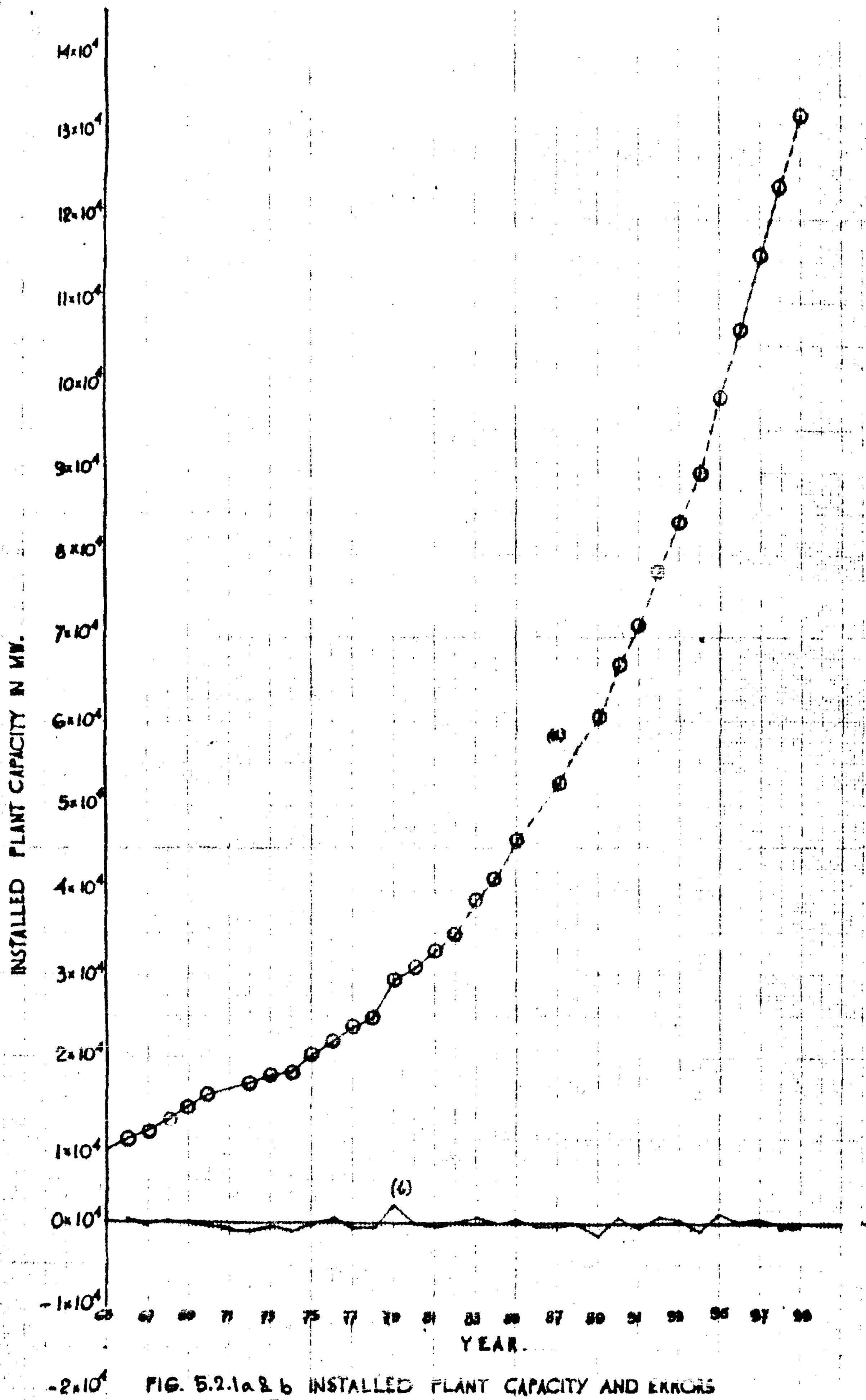


FIG. 5.2.1a & b INSTALLED PLANT CAPACITY AND ERRORS BETWEEN THE OBSERVED AND THE MODELLED VALUES.

there will be a significant improvement of the quality of life of the Indian people if the energy consumption is increased to five times its 1981 figures. All reasonable solutions to alleviate the multitudes of complex problems involving population growth, economics, energy, development, transport and communication involve sharp increases both in the amount of energy consumed and in the efficiency of their use.

A long term energy consumption model has been postulated in the form of a finite difference equation as stated below :

$$Y_{k+1} = f\,(\,Y_k,\ Y_{k-1},\ Y_{k-2},\ Y_{k-3},\ t_k\,) \qquad (5.3.1)$$

The arguments are selected on the basis of the correlation with the output Y_{k+1}. The polynomial model of optimum complex has been obtained as

$$\begin{aligned} Y_{k+1} = & -62.374577 + 1.115110\,Y_k \\ & + 0.068099\,Y_{k-1} - 0.114979\,Y_{k-2} \\ & + 0.028869\,Y_{k-3} - 128.665721\,t_k \\ & - 3.701969\,E-08\,Y_k\,Y_{k-1} \\ & + 4.507208\,E-12\,Y_k\,Y_{k-2} \end{aligned} \qquad (5.3.2)$$

The corresponding integral square error and mean error are 7.162363 E - 06 and 3.409563 E - 02.

The errors are found to be uncorrelated for $k \neq j$. Fig. 5.3.1a and 5.3.1b show the extrapolated annual electrical energy consumption and the errors of modelling.

5.4 Polynomial model with periodicity terms for annual plant load factor

The purpose of this part of the work is to obtain a polynomial model for annual plant load factor. The observed data are processed according to the method stated below.

The annual measured data for load factor are

$P(i),\ i = 1,2,\ \ldots,\ N_1$

i being the instant in years.

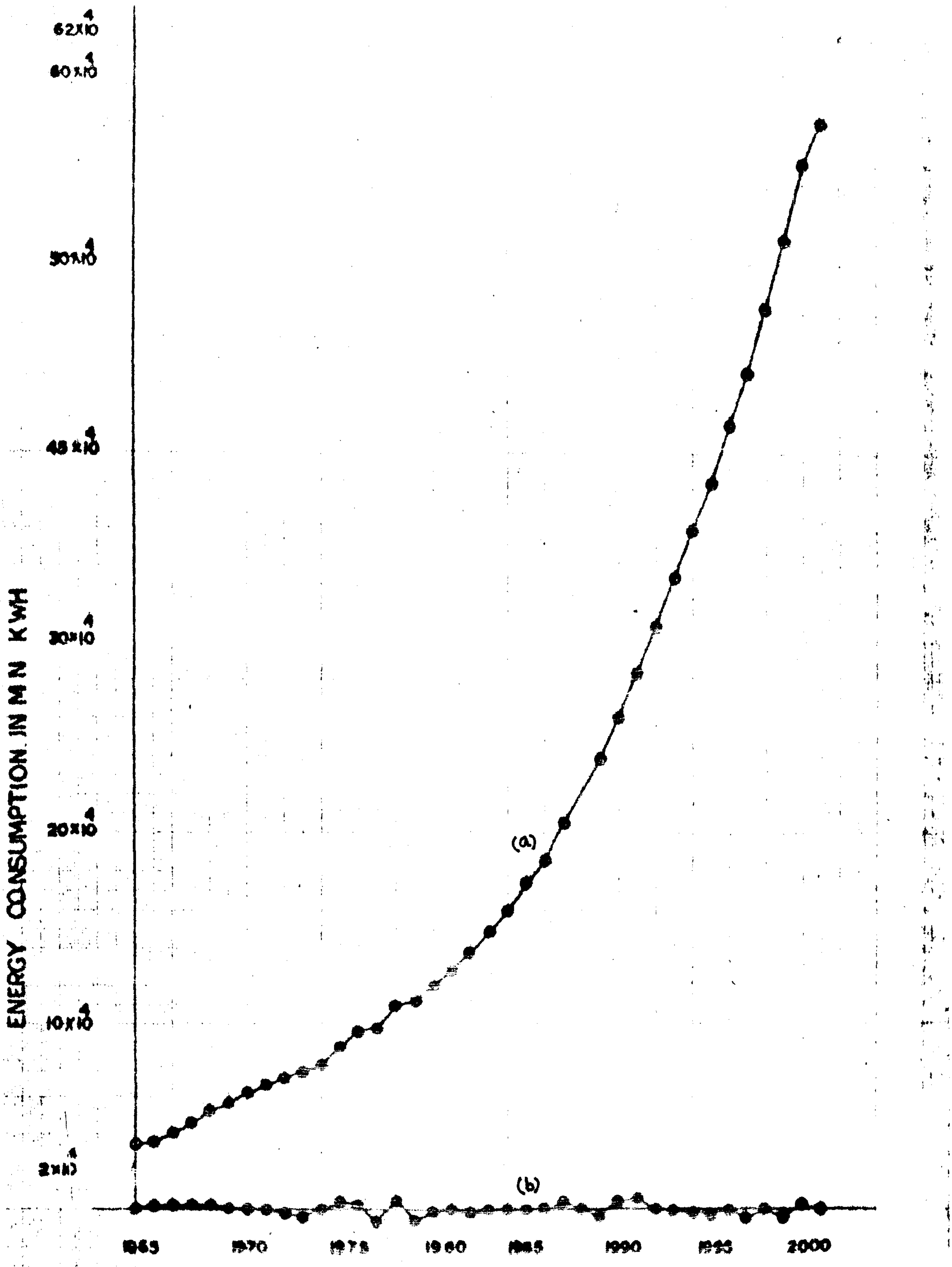
The mean value is given by

$$\bar{P}(N_1) = \frac{1}{N_1} \sum_{i=1}^{N_1} P(i) \qquad (5.4.1)$$

Autocovariance of the data at lag instant k is given by

$$CA(k) = \frac{1}{N_1 - k} \sum_{i=1}^{N_1-k} \left[P(i) - \bar{P}(N_1) \right] \left[P(i+k) - \bar{P}(N_1) \right] \qquad \ldots (5.4.2)$$

where $k = 0,\ 1,2,\ \ldots,\ M$ and $M < \frac{1}{4} N_1$.



ANNUAL ENERGY CONSUMPITION EXTRA POLATED FROM 1981-2001 AND ERRORS, BETWEEN THE OBSERVED AND MODELLED VALUES RESPECTIVELY. FIG 5.3.1a & 5.3.1b

The normalised co-efficients of covariance are given by

$$RA(k) = \frac{GA(k)}{GA(o)} \qquad (5.4.3)$$

The estimate of the normalised power density spectra for the data are given by

$$PS(W_h) = \frac{2}{\pi} \sum_{k=0}^{M} E_k \, RA(k) \cos W_h k \qquad (5.4.4)$$

where $W_h = 2\pi f_h, \; f_h = \frac{h}{2M}$; $0 \leqslant f_h \leqslant 0.5$

$$h = 0, 1, 2, \ldots, M$$

E_k has been defined as the weight for window correction and may be taken as

$$E_k = 1.0 \quad \text{for} \quad 0 < k < M$$
$$0.5 \quad \text{for} \quad k = 0, M \qquad (5.4.5)$$

These raw estimate of power spectral density are smoothed by using Hamming Window to obtain the final estimates of the power spectrum. The smoothed estimates of the ordinates of the power spectrum are

$$S(W_h) = 0.54\, PS(W_h) + 0.46\, PS(W_1)\ ;\ \text{for}\ \ h = 0$$

$$S(W_h) = 0.23\, PS(W_{h-1}) + 0.54\, PS(W_h) + 0.23\, PS(W_{h+1})\ ;\ \text{for}\ \ 0 < h < M$$

$$S(W_h) = 0.54\, PS(W_h) + 0.46\, PS(W_{h-1})\ ;\ \text{for}\ \ h = M$$

... (5.4.6)

The periodicity in terms of fundamental and its harmonics can be estimated from the power spectral density - frequency characteristics as shown in Fig. 5.4.1. The process has been found to have a 10 yearly cycle (i.e., 0.1 cycle per annum). Only one lag instant of the annual plant load factor has been found to be strongly correlated with current instant of the annual plant load factor.

Consequently the functional model of the annual plant load factor in functional form is given by

$$P(k) = f(P(k-1),\ \sin(2\pi f_o k),\ \cos(2\pi f_o k),\ \sin(2\ f_o(k-1)),\ \cos(2\pi f_o(k-1))) + \eta(k)$$

... (5.4.7)

Let $P(k) = y, P(k-1) = x_1,\ \sin(2\pi f_o k) = x_2,$

$\cos 2\pi f_o k = x_3,\ \sin 2\pi f_o(k-1) = x_4,$

$\cos 2\pi f_o(k-1) = x_5$ and $\eta(k) = \xi$

Equation (5.4.7) becomes

$$y = f(x_1, x_2, x_3, x_4, x_5) + \xi \qquad (5.4.8)$$

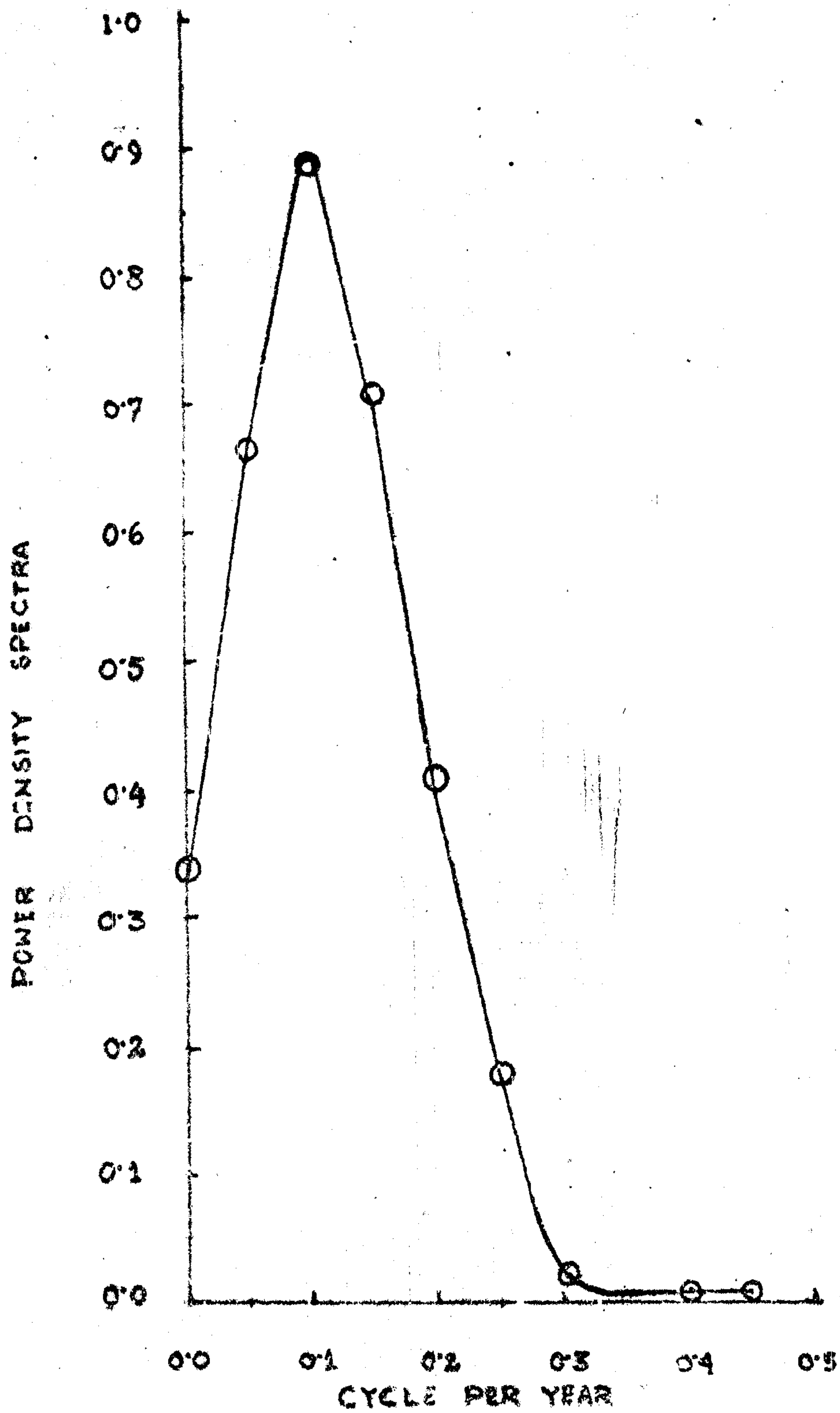


FIG-5.4.1 POWER DENSITY SPECTRA VS CYCLE PER ANNUM

The estimate of y as $\hat{y}$ is defined as

$$\hat{y} = f\,(x_1, x_2, x_3, x_4, x_5) \qquad (5.4.9)$$

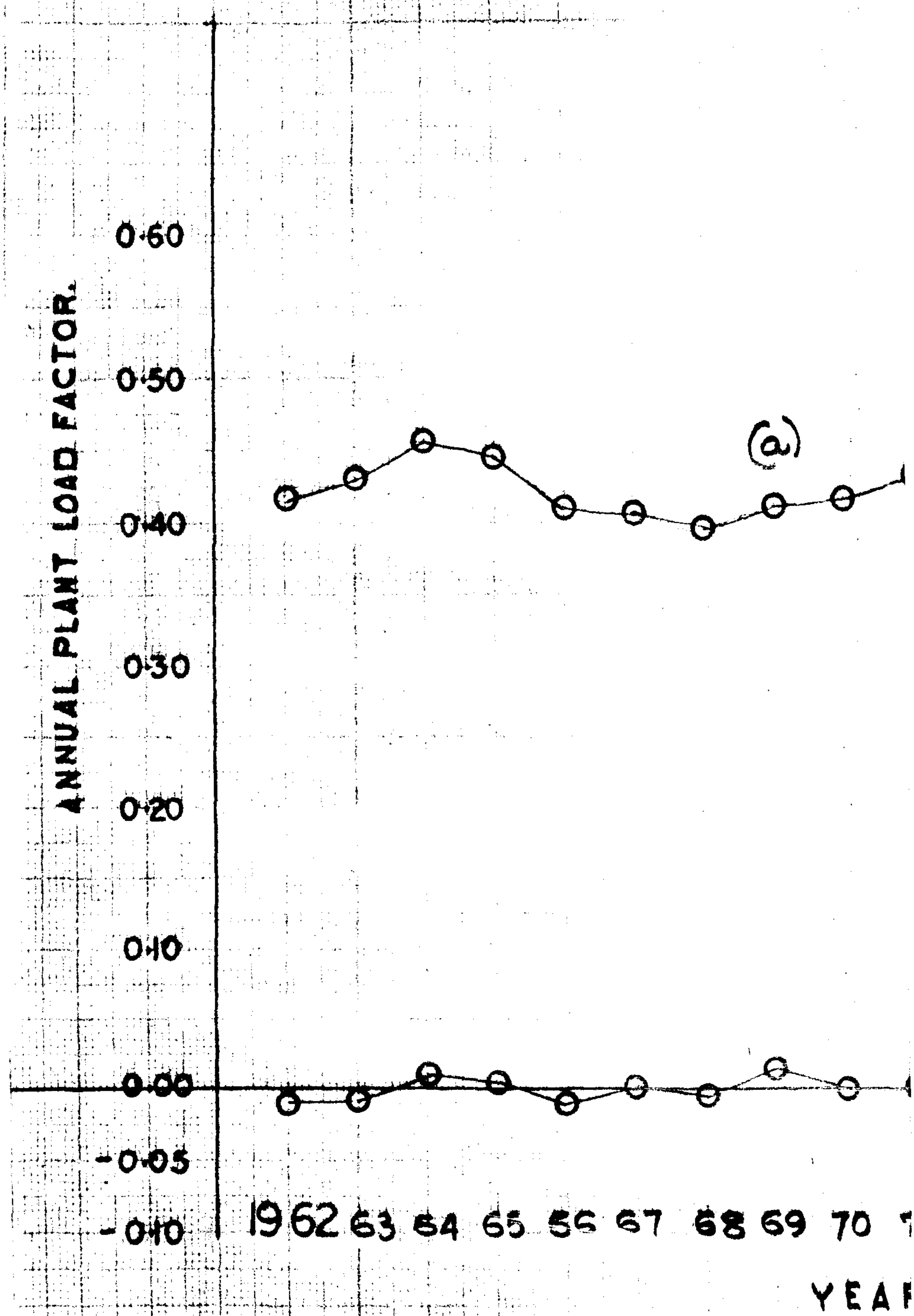
The function $f(\,.\,)$ in equation (5.4.9) is sought in the class of quadratic polynomials on the basis of a Table of polynomials of gradually increasing complexity of five variables as shown in Table 5.2.1 with the theory of self-organisation of mathematical models. The model of optimum complexity is selected on the basis of minimum integral square error criterion as defined in equation (5.1.3).

The model of optimum complexity is obtained as

$$\begin{aligned} P(k) = {} & 0.215817 + 0.502090\, P(k-1) \\ & + 17.065098 \sin 0.2\pi k \\ & - 16.103889 \cos 0.2\pi k \\ & - 23.161123 \sin 0.2\pi(k-1) \\ & + 2.964121 \cos 0.2\pi(k-1) \\ & - 17.073073\, P(k-1) \sin 0.2\pi k \\ & + 12.529761\, P(k-1) \cos 0.2\pi k \\ & + 20.968465\, P(k-1) \sin 0.2\pi(k-1) \\ & - 4.75\text{E-}12\, P(k-1) \cos 0.2\pi(k-1) \end{aligned} \qquad \ldots (5.4.10)$$

The corresponding integral square error and mean error are 9.242190 E-04 and 1.808587 E-03. $\eta(\,.\,)$ is found to be almost uncorrelated for $k \neq j$ and with variance at 0.999692.

Fig. 5.4.2a and 5.4.2b show the observed values of the plant load factor and the errors between the observed and the modelled values respectively [90]. The software developed in BASIC language is given the Appendix as A3.1 and A3.2.



ANNUAL PLANT LOAD FACTOR

AND MODELLED VALUES RESPEC

FIG 5.4.2a & 5.4.2