CHAPTER VIII

REAL TIME FORECASTING OF DISSOLVED OXYGEN LEVEL OF A NON TIDAL RIVER

CHAPTER VIII

REAL TIME FORECASTING OF DISSOLVED OXYGEN LEVEL OF A NON-TIDAL RIVER.

8.0 Abstract :

Recursive least square non-stationary time series analysis techniques of cybernetics has been put to good use in on-line forecasting of daily dissolved oxygen concentrations of a non-tidal stream. A simple dynamic model is obtained. When tested with field data the model is shown to simulate adequately the major variations of D.O. levels observed in the field measurements.

8.1 Introduction :

This work introduces some of the fundamental techniques and principles of control theory broadly classified as cybernetics and indicates how ~~thy~~ they might be put to good use in the real time forecasting of the water quality of a non-tidal stream. This forecasting model would help in the subsequent development of the methodologies for realtime controlling of water quality. For on line (Real time) information processing the basis of the processing mechanism is a recursive least square algorithm for process parameter estimation. Recursive parameter estimation algorithms have the advantages of minimal computational requirements and their ability to tract time varying parameters. In the present work dynamic model for dissolved oxygen in a reach of river has been obtained. The model has been verified by simulation against field measurement collected over a 80-day period from the River Cam in Eastern England [108].

The dissolved oxygen level is a major index of the general state of an aquatic environment. The important engineering decisions in water quality control relate to the determination, regulation and maintenance of the level of D.O. that is consistent with the multiple uses of the natural water bodies. The D.O. concentrations are important aspect of water quality in streams since the temporary fluctuations of D.O. level below critical values can cause long term and possibly irreversible damage to the ecosystem [107].

8.2. Processing of the Statistical Data :

The daily measured data for D.O. concentration of a stream are $P(i)$, $i = 1, 2, \ldots, N1$.

the time instants (days).

The mean value is

$$\bar{P}(N1) = \frac{1}{N1} \sum_{i=1}^{N1} P(i) \qquad \ldots \qquad (8.1)$$

Auto covariance of the data at lag instant k is given by

$$GA(k) = \frac{1}{N1-k} \sum_{i=1}^{N1-k} (P(i) - \bar{P}(N1))(P(i+k) - \bar{P}(N1)) \qquad (8.2)$$

where $k = 0, 1, 2, \ldots, M \quad M < \frac{1}{4} N1$

Normalised co-efficients of covariance is

$$RA(k) = \frac{GA(k)}{GA(o)} \qquad \ldots \qquad (8.3)$$

Estimate of the normalised power density spectra for the data are given as,

$$PS(W_h) = \frac{2}{\pi} \sum_{k=0}^{M} E_k \; RA(k) \; \text{Cos}\, W_h k \qquad (8.4)$$

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

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

E_k has been defined as the weight for window correction,

$$E_k = \begin{cases} 1.0, & 0 < k < M \\ 0.5, & K = 0, M \end{cases} \quad \ldots (8.5)$$

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

$$h = 0, \quad S(w_0) = 0.54 \; PS(w_0) + 0.46 \; PS(w_{h+1})$$

$$0 < h < M; \quad S(w_h) = 0.23 \; PS(w_{h-1}) + 0.54 \; PS(w_h) + 0.23 \; PS(w_{h+1}).$$

$$h = M; \quad S(w_h) = 0.54 \; PS(w_M) + 0.46 \; PS(w_{M-1}) \ldots (8.6)$$

Periodicity in terms of fundamental and its harmonics can be estimated from the power spectral density Vs frequency characteristics.

8.3. Development of the Non-Stationary Process Model :

Non-stationary process with deterministic periodicity can be expressed as

$$P(k) = U(k) + V(k) \quad \ldots (8.7)$$

Where $U(k)$ is the deterministic trend components,

$V(K)$ is the stationary stochastic component,

$k = 1, 2, \ldots,$ are time instances,

Deterministic periodic component can be represented as

$$U(k) = \sum_{l=1}^{L} \left(a_l \cos(l\, 2\pi f_0 k) + b_l \sin(l\, 2\pi f_0 k)\right) \quad \ldots (8.8)$$

where fo is the fundamental cycle per instance of time. Stochastic component V(k) can be expressed in the autoregressive form as,

$$V(k) = \sum_{n=1}^{N} C_{k-n} V(k-n) + \eta_{(k)} \quad \ldots (8.9)$$

where n = 1,2, ..., N, the prehistory interval and $\eta_{(k)}$ is modelled in moving average form

as $$\eta_{(k)} = \sum_{j=1}^{m} C_{k-j} \eta_{(k-j)} + \nu_{(k)} \quad \ldots (8.10)$$

where m is the prehistory interval of the moving average terms and v(k) resembles to a white noise process with zero mean and constant variance.

V(k) can also be expressed as,

$$V(k-n) = P(k-n) - U(k-n) \quad \ldots (8.11)$$

Hence equation (8.9) can be written as

$$V(k) = \sum_{n=1}^{N} C_{k-n} P_{(k-n)} - \sum_{n=1}^{N} C_{k-n} U_{(k-n)} + \eta_{(k)} \quad \ldots (8.12)$$

It follows from equations (8.7) and (8.9)

$$P(k) = \sum_{n=1}^{N} C_{k-n} P_{(k-n)} + \sum_{n=0}^{N} \sum_{l=0}^{L} [d^{l}_{k-n} \cos (l 2\pi fo (k-n))$$

$$+ g^{l}_{k-n} \sin (l 2\pi fo (k-n))] + \eta(k) \quad \ldots (8.13)$$

In matrix form the process P(k) can be expressed as

$$P(k) = a^{T} (k-1) Z(k-1) + \eta(k) \quad \ldots (8.14)$$

where $a^T(k-1) = [C_{k-1}, \ldots, C_{k-n}, \sum_{n=0}^{N} (d^o_{k-n}), \ldots, g^t_{k-w}]$ ... (8.15)

And

$$Z(k-1) = [P(k-1), \ldots, P(k-n), 1, \ldots, \cos i(2\pi fo\ (k-n)), \ldots, \sin i(2\pi fo\ (k-n))]^T \quad \ldots \quad (8.16)$$

$\widehat{P}(k)$, the estimate of P(k) can be very clearly written as

$$\widehat{P}(k) = a^T\ (k-1)\ Z\ (k-1) \quad \ldots \quad (8.17)$$

8.4 Development of the Real Time Recursive Least Square Prediction Algorithm :

The co-efficient vector a can be estimated by minimizing the quadratic function $J_k(a)$,

$$J_k(a) = \sum_{j=1}^{k} (P(j) - a^T Z(j-1))^2 + (a-a(o))^T S^{-1}(o) \ (a-a(o)) \quad (8.18)$$

where a(o) is the available a priori estimate of the co-efficient vector 'a' and S(o) is the positive definite weighting matrix of the order mxm, where

$$m = (N+1)\ (2L+1) \quad \ldots \quad (8.19)$$

and N is the prehistory internal

and L is the number of harmonics including the fundamental.

For minimization,

$$\frac{\partial J_k(a)}{\partial a} = -2 \sum_{j=1}^{k} Z(j-1)(P(j)-A^T Z(j-1)) + 2S^{-1}\ (o)\ (a - a(o)) \quad \ldots \quad (8.20)$$

Since we are seeking for a minima,

$$\frac{\partial J_k(a)}{\partial a} = 0 \qquad \ldots \quad (8.21)$$

It follows from equation (8.21)

$$\sum_{j=1}^{k} Z(j-1)P(j) + S^{-1}(0)\, a(o) = \sum_{j=1}^{k} Z(j-1)Z^T(j-1)a + S^{-1}(o)a \qquad (8.22)$$

$$\text{Let } S^{-1}(k) = \sum_{j=1}^{k} Z(j-1)Z^T(j-1) + S^{-1}(o) \qquad \ldots \quad (8.23)$$

$$\text{and } \quad d(k) = \sum_{j=1}^{k} Z(j-1)\, P(j) + S^{-1}\, a(o) \qquad \ldots \quad (8.24)$$

Denoting the estimate of the co-efficient vector 'a' as a(k) at time instant k,

$$S^{-1}(k)\;\; a(k) = d(k) \qquad \ldots \quad (8.25)$$

$$a(k) = S(k)\;\; d(k)$$

From equations (8.23) and (8.24) the following recursive equations are obtained,

$$S^{-1}(k+1) = S^{-1}(k) + Z(k)\, Z^T(k) \qquad \ldots \quad (8.26)$$

$$d(k+1) = d(k) + Z(k)\, P(K+1) \qquad \ldots \quad (8.27)$$

By matrix inversion lemma the recursive least square parameter estimation algorithms are

$$a(k+1) = a(k) + S(k+1)Z(k) \left[P(k+1) - a^T(k)\, Z(k) \right] \quad \ldots \quad (8.28a)$$

$$S(k+1) = S(k) - S(h)Z(k)Z^T(k)S(k) \left[1+Z^T(k)S(k)Z(k) \right] \ldots \quad (8.28b)$$

Periodicity L may be estimated from the power density spectra versus cycle per time instance. Initial value of the prehistory internal N is estimated from the correlation function. The correlation function is defined as

$$\phi_{pp}(\lambda) = \frac{\sum_{i=1}^{N1-\lambda} (P(i)-\bar{P}(N1))(P(i+\lambda)-\bar{P}(N1))}{\left[\sum_{i=1}^{N1-\lambda} (P(i)-\bar{P}(N1))^2 \sum_{j=1+\lambda}^{N1} (P(j)-\bar{P}(N))^2 \right]^{\frac{1}{2}}} \quad \ldots (8.29)$$

where λ is the shift of instances of time. For estimation of the parameters of the moving average components of equation (8.10) recursive algorithms (8.28a) and (8.28b) have been followed. The algorithms have been initialised with $k = o$; $S(o) = I$ (unit matrix), $a(o) = 0$ and $P(j) = 0$, for $j = u, -1, -2, \ldots$

The non-stationary model of optimum complexity is obtained by comparing the criterion of integral square error between the predicted values and the observed values in the interpolation region for the different values of the prehistory interval N.

Criterion for the integral square error is defined as

$$\sigma_P^2 = \frac{\sum_{i=1}^{N1} (P(i) - \hat{P}(i))^2}{\sum_{i=1}^{N1} (P(i))^2} \quad \ldots (8.30)$$

8.5 Illustration :

Figure 8.5.1a shows the D.O. level at down stream end of 4.7 Km stretch of the River Cam in Eastern England, collected over an 80 day period during summer of 1972. Figure 8.5.2 shows the power density spectra versus cycle per day characteristic. It is evident that the fundamental component of the periodicity has 0.03125 cycle per day and the harmonics are negligible. From the correlation coefficient versus shift of time as shown in figure 8.5.3 it is obvious the process follows a maximum prehistory interval of 5 time instances (days).

Minimum value of the integral square error was obtained as 0.0102826 with prehistory interval $N = 2$. The innovation process $v(k)$ has been identified to have a mean of 0.077871 with variance 1.06121 for moving average terms in corporating $M = 2$. Figure 8.5.1b shows the error of prediction.

8.6. Conclusion :

A simple dynamic model is presented for the description of D.O. level in a non-tidal stream. When tested with field data the model is shown to simulate adequately the major variations observed in the field.

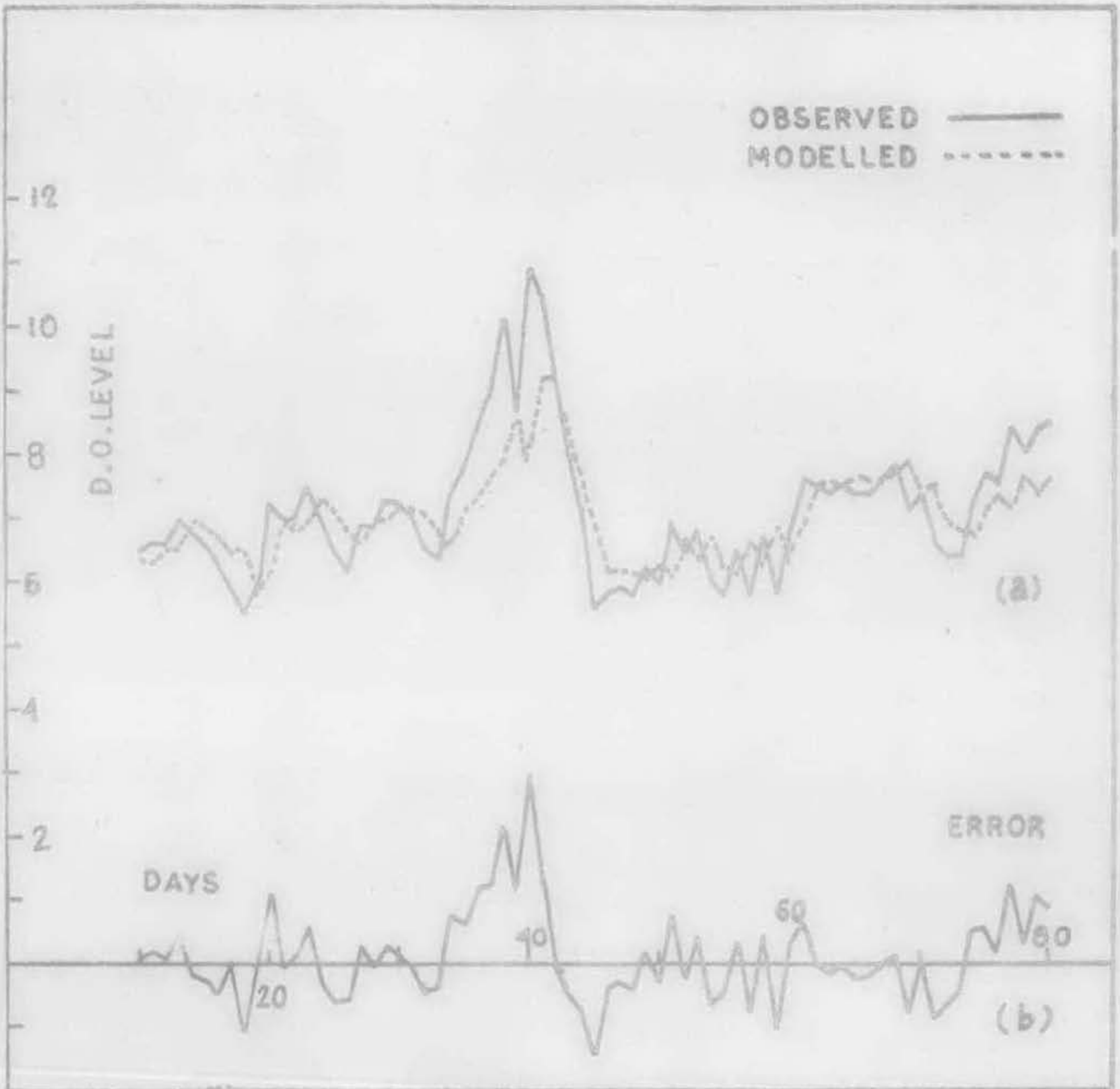
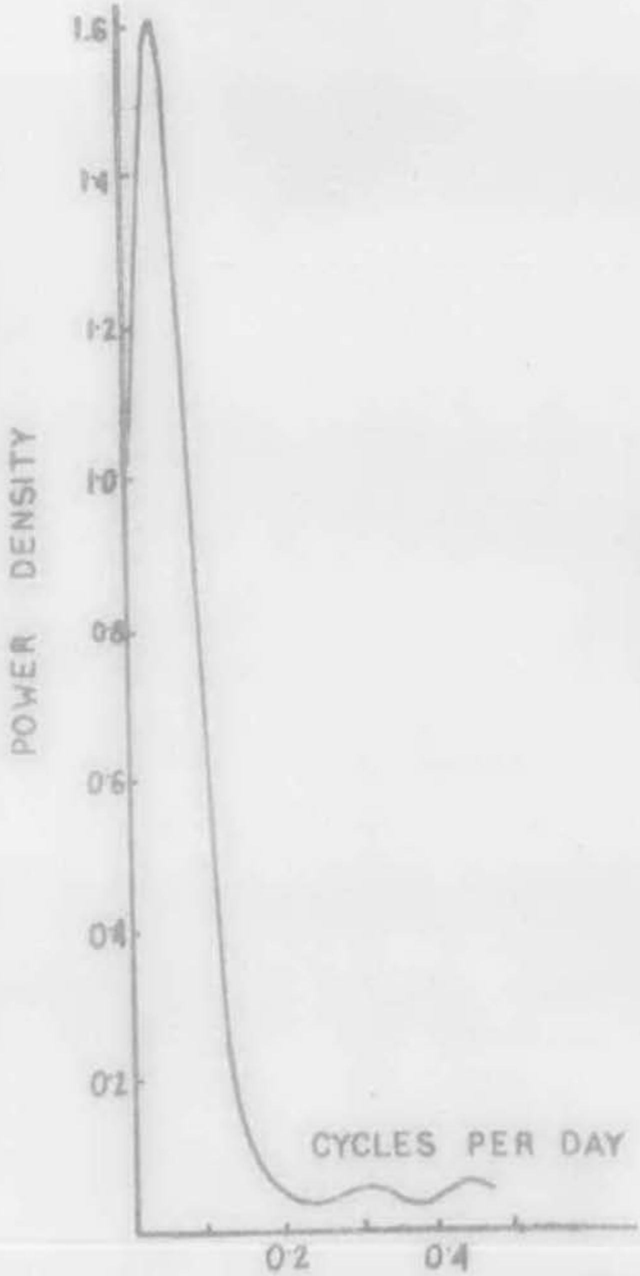


FIG 8.5.1 OBSERVED D.O. LEVEL AND ERROR OF PREDICTION



POWER DENSITY SPECTRA OF D.O. LEVEL

FIG 8.5.2

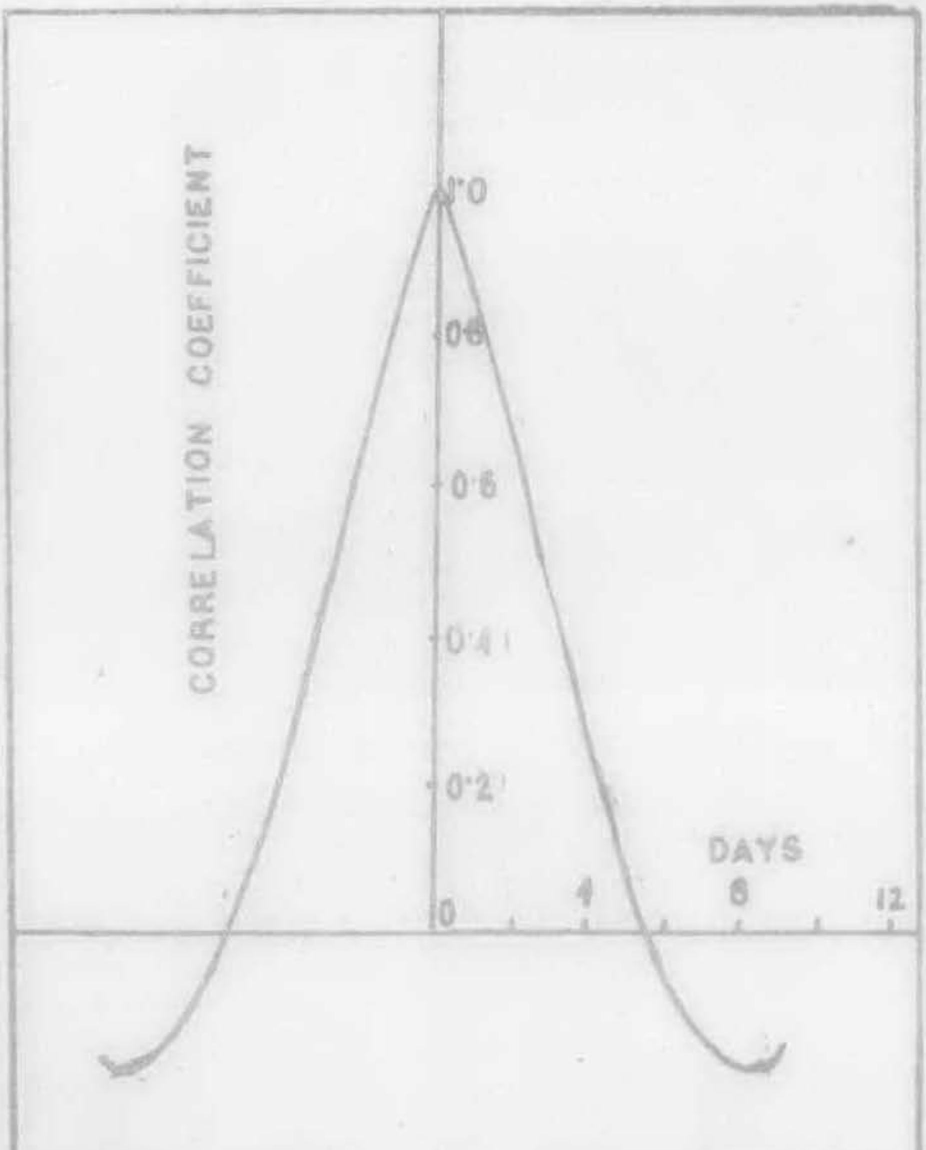


FIG 6.53 CORRELATION COEFFICIENT. D.O. LEVEL