4. NONLINEAR DYNAMICAL SYSTEMS

4.1. Finding Parameters and Initial Conditions

We have recently developed a method to accurately calculate derivatives of a time series data. This ability can be used to identify unknown parameters if the general form of equations is known. For example, in Lorenz's equations,

$$\frac{dx}{dt} = \sigma\,(y - x)$$

$$\frac{dy}{dt} = -xz + \rho x - y$$

$$\frac{dz}{dt} = xy - \beta z$$

if *P(t)* is the observed *x* signal and, *Q, R,* and *S* are the progressively higher derivatives, we can easily show the following relationship.

$$\left[-P.Q + \beta.\frac{(-R-Q)}{P} + \frac{(-P.(S+R)+Q.(R+Q))}{P^2}\right]\gamma +$$

$$+\beta.\left(\rho - 1 - \frac{Q}{P}\right) + \left(\frac{(-P.R+Q^2)}{P^2} - P^2\right) = 0$$

where $\gamma = 1/\sigma$.

Now, we can accurately find Q,R and S only from the knowledge of P. So, all these quantities can be evaluated at a few consecutive points to arrive at a set of simultaneous equations with the parameters as the unknowns. The example below shows a solution by using Mathcad:

$(-3451.6 - 229.20.\beta)\gamma + \beta.(\rho + 22.179) + 281.19 = 0$

$(-4701.7 - 270.11.\beta)\gamma + \beta.(\rho + 25.558) + 408.04 = 0$

$(-7075.5 - 328.24.\beta)\gamma + \beta.(\rho + 30.755) + 647.19 = 0$

$(-6123.6 - 145.39.\beta)\gamma + \beta.(\rho + 20.968) + 563.52 = 0$

$(-12334 - 1421.49.\beta)\gamma + \beta.(\rho + 39.463) + 1174.7 = 0$

The parameters with minimum error obtained are:

Mineer (β, ρ, σ,) = (2.657, 24.837, 0.1)

Once the parameters are known, the initial conditions can also be obtained quite accurately as

$y_0 = 2.5$
$z_0 = 24.087$

These results are quite important in modelling and simulation and in cryptography. It is worth noting that the simplicity of this method is somewhat deceptive. Alternative methods trying to accomplish this via a least square type of methods have proven to be highly complex and time consuming.

(P.G. Vaidya)

4.2 Dynamics of Marine Ecosystem Models

It has been reported in the literature that marine ecosystem models display limit cycles under certain conditions, especially when there is abundance of nutrients. It has also been suggested that introduction of nonlinear mortality and/or self-grazing terms for zooplankton suppresses such limit cycles. These type of oscillations of the models are often believed to be unrealistic. This issue was examined with the help of a class of basic models having the least number of state variables and yet incorporate some of the essential biology. The model was obtained by describing the biomass of all autotrophs as the variable Phytoplankton (*P*), the biomass of heterotrophs as the variable Zooplankton (*Z*), and the mass of nutrients as the variable *N*. In the present study, the growth terms of phytoplankton and zooplankton were modelled using Michaelis-Menten law, mortality of phytoplankton was taken to be linear, mortality of zooplankton was modelled as a sum of linear and quadratic terms and self-grazing of zooplankton was also incorporated. When more than one food type is available for growth, different formulations of growth and grazing terms are possible, like nonswitching, switching etc. Qualitative aspects of the dynamics of the model can be readily understood with the help of a phase plane diagram. Limit cycles arise when the equilibrium point becomes unstable. Stability of the equilibrium point was determined from the sign of the trace of the Jacobian. When the trace is positive, the equilibrium point is unstable and leads to limit cycles in our models. The effect of quadratic mortality term on the trace of the Jacobian at the equilibrium point was studied keeping all the other

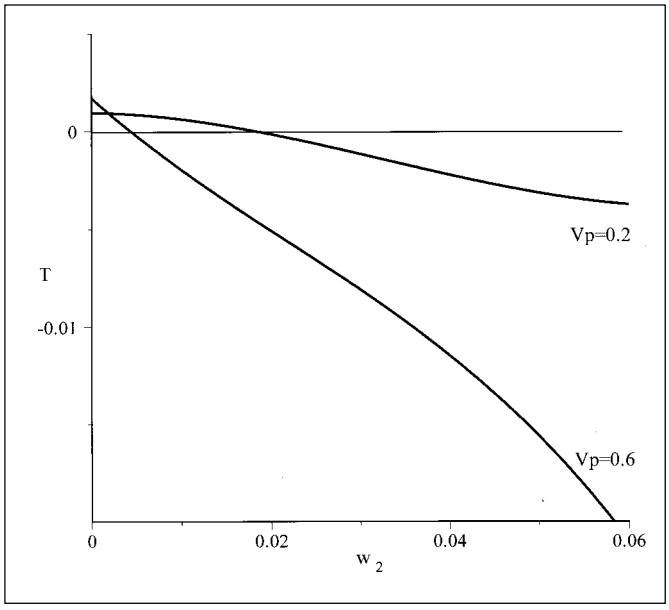

Fig. 4.2.1. Effect of nonlinear mortality coefficient (W_2) on the trace of the Jacobian (T) for two values of asymptotic growth rate (Vp) of phytoplankton.

model parameters fixed and it was found that the value of the trace decreased when the quadratic mortality term was increased leading to stability of the equilibrium point *(Fig.4.2.1).* Similarly, when self-grazing of zooplankton was introduced, limit cycles were suppressed. It was shown using phase plane analysis that even a rather small nonlinear term describing nonlinearity in mortality or self-grazing of zooplankton can effectively alter the linear stability conditions of this class of ecosystem models and thereby suppress the limit cycles.

(M.K. Sharada)

4.3 Neural Network Forecasts of All-India Monthly and Summer Monsoon Rainfall

In view of the critical importance of knowledge of rainfall patterns in areas like agricultural planning, accurate and long range forecast of rainfall is of vital importance. Given the present status of prediction algorithms, this requires persistent effort and perhaps a multi-pronged strategy using different techniques and methodologies. One such promising forecasting tool is neural networks (NN), which we have been exploring and developing now for several years. In particular, a generalized NN, termed cognitive network (CN), was developed and tested for hindcast skill in long range forecast of all India summer monsoon rainfall (ISMR).

Subsequent to the development of CN, we have followed use-and-probe strategy with NN forecasts. In particular, we believe that, while many of the shortcomings of NN in general and, CN in particular, need to be addressed at conceptual and theoretical level, the capabilities and the scope of applicability of CN as a forecast tool need to be explored alongside. These two efforts can, and should, feedback upon each other, leading to better understanding and further development.

With this philosophy in the background, we have been generating experimental forecasts of ISMR for the last four years. It is noteworthy that all the three forecasts for 1996, 1997 and 1998 were generated well ahead of the monsoon season, and were found to be of good quality. The CN forecast for 1998 ISMR was, for example, 945 mm (or 107% of long term mean) while the observed value was about 106%. The success for 1998 assumes further significance since it was made in 1996 two years in advance. These successes, along with the hindcast skill achieved during our investigation, suggest it to be worthwhile to pursue this approach. We have been thus encouraged to investigate the forecast skill of CN for all India monthly rainfall (AIMR). *Table 4.3.1* shows the performance of CN in AIMR forecast from seventy four hindcast experiments. We also record in *Table 4.3.2* the CN forecasts of ISMR for 1999 and 2000.

These forecasts should be interpreted as ensemble forecasts where, in the present case, an ensemble is formed by generating forecasts through a number of CN configurations with comparable hindcast skill. The inputs to the networks consist of past rainfall data of appropriate time scale (e.g. past ISMR data and past AIMR data for the corresponding forecasts). Error back propagation algorithm was employed and, for ISMR forecast, 50 points of a 123 year data set was used to train the CN.

The following parameters were used to evaluate the performance of hindcast skill.

1. Success rate:

$$\mu = \frac{n - n^1}{n} \; x \; 100$$

where n is the total number of predictions and n^1 is the number of predictions out of phase. Thus for $n^1=0$, the success rate is 100%, while for $n=n^1$ the success rate is 0%.

Table 4.3.1
Performance of CN in AIMR forecast

Evaluation Parameters	Month											
	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
<e> (mm)	9.9	9.8	7.7	7.3	16.2	54	65	53	44	37	24	9.7
μ (%)	50	48	36	60	43	46	48	43	48	38	45	46
γ_1	0.8	0.9	1	1	1	1	1	1	1	1	0.8	1
γ_2	0.7	0.9	1	1	1	0.7	0.6	0.9	0.8	0.8	0.8	1
Bias	-2.5	-1	2.6	1.7	5	19.2	6.1	13.3	10.7	-1.2	-8.1	1.3

Table 4.3.2
The CN forecasts of ISMR for 1999 and 2000

Year	Ensemble mean (mm)	Ensemble SD (mm)
1999	861	43
2000	793	30

2. Average absolute error:

$$< e > = \frac{1}{N} \sum_{i=1}^{N} e_i$$

with

$$e_i = \left| X_p^i - X_t^i \right|$$

where X_p is the predicted value and X_t is the target value.

The other evaluation parameters are :

$$\gamma_1 = \frac{M_p}{M_o}, \qquad \gamma_2 = \frac{\sigma_p}{\sigma_o}$$

and

Bias = M_o - M_p

where M_p and M_o are the mean of the predicted and observed values and s_p and s_o are the standard deviations (SD) of the predicted and observed values respectively.

The ensemble standard deviations in this case provide a measure of the uncertainty in the forecasts. A similar methodology was adopted for generating the forecasts for 1997 and 1998. Thus, according to these forecasts, the year 2000 is likely to be a deficit monsoon year. However, the forecast for 2000 shall be updated should our forecast for 1999 turn out to be appreciably different from the observed value (no updating was necessary for 1998 as our forecast for 1997 turned out to be close to the observed value).

(K. Rameshan and P. Goswami)

4.4 Computational Biology

4.4.1 DNA sequence analysis of complete bacterial genomes

The sequences of all the bacterial genomes sequanced so far were analysed to understand the distribution of repetitive DNA sequences and, in particular, the polypurine/polypyrimidine repeats. These repeats are of interest since they have been shown to influence gene expression in *cis* and have also been found at regulatory regions of genes. Consistent with our earlier results on the yeast genome, a majority of the bacterial genomes also showed asymmetry in the distribution of polypurines and polypyrimidines between the coding and non-coding strands. In addition, the knowledge that this asymmetry can show a wide range of variation was an important outcome of this analysis. This has also helped strengthen the hypothesis that repetitive sequences in DNA can have important biological roles. A brief summary of these results is shown in *Table 4.4.1.*

(Sowmya Raghavan)

Table 4.4.1
Distribution of polypurine (polyR) and pyrimidine (polyY) repeats in bacterial genomes

Organism*	Total polyR and polyY	polyR in coding regions	poly Y in coding regions	Strand bias[1]	polyR in non-coding regions	polyY in non-coding regions
M.genitalium (580)	69	40	15	2.7	5	9
M.pneumoniae (816)	80	46	9	5.1	13	12
B.burgdorferri (911)	167	127	32	3.9	4	4
C.trachomatis (1042)	374	137	160	0.85	38	39
R.prowazekeii (1111)	71	52	6	8.7	8	5
T.pallidum (1138)	251	91	125	0.73	18	17
A.aeolicus (1551)	1351	1075	152	7.07	59	65
M.jannaschii (1665)	735	696	8	87	13	18
H.pylori (1667)	315	232	35	6.1	30	18
P.horikoshii (1738)	992	840	140	6.0	5	7
M.thermoautotrophicum (1751)	315	250	37	6.7	15	13
H.influenzae (1830)	115	57	21	2.7	17	20
A.fulgidus (2178)	1005	867	93	9.1	25	20
*Synechocystis sp.(*3573)	502	253	143	1.76	62	44
B.subtilis (4215)	856	543	83	6.8	111	119
M.tuberculosis (4412)	21	8	4	2	6	3
E.coli (4639)	198	81	38	2.1	37	42

[1] Strand bias refers to the ratio, polyR /polyY, in coding regions.
[*] Size in kilobases is given in brackets.

4.4.2 Analysis of Mycobacterium tuberculosis genome to identify novel gene functions

Apart from the analysis of bacterial genomes as a whole, efforts have also been made to subject the genome of *M. tuberculosis* to intensive sequence analysis. Since the report of the complete genome sequence in the literature, no new functions have been reported. We have analysed all the open reading frames of unknown function using the similarity search algorithm BLAST. This exercise revealed that at least ten new functions can be discerned. Since some of these functions are enzymes, the possibility of using these as potential drug targets emerges. The process of validating these new functions from a structural point of view, using the technique of homology modelling is now in progress. *(Sowmya Raghavan)*

4.5 Non-linear Dynamical Applications to Land-Atmosphere Interactions

The important role played by ground hydrology in the dynamics of monsoonal precipitation has been emphasized by a number of studies. While appropriate representation of ground hydrological processes is imperative for improved range and quality of rainfall forecast, its modelling is complicated by the interplay of a number of stochastic and non-linear processes. To gain insight into the behavior and hence modelling of ground hydrological processes, we have considered a statistical dynamical model representing the surface hydrology of large continental regions. The model represents rainfall rate E_p, moisture infiltration function $\phi(s)$ and evapotranspiration rate $E(s)$ as functions of moisture s. The evapotranspiration rate and infiltration rate are parameterized as

$$E(s) = E_p\, s^c$$

and

$$\phi(s) = 1 - \varepsilon\, s^r$$

where c, ε and r are nonnegative constants.

Non-linear and stochastic analysis of the soil moisture balance equation yields two fixed points akin to floods and droughts as reported in the literature. Stochastic

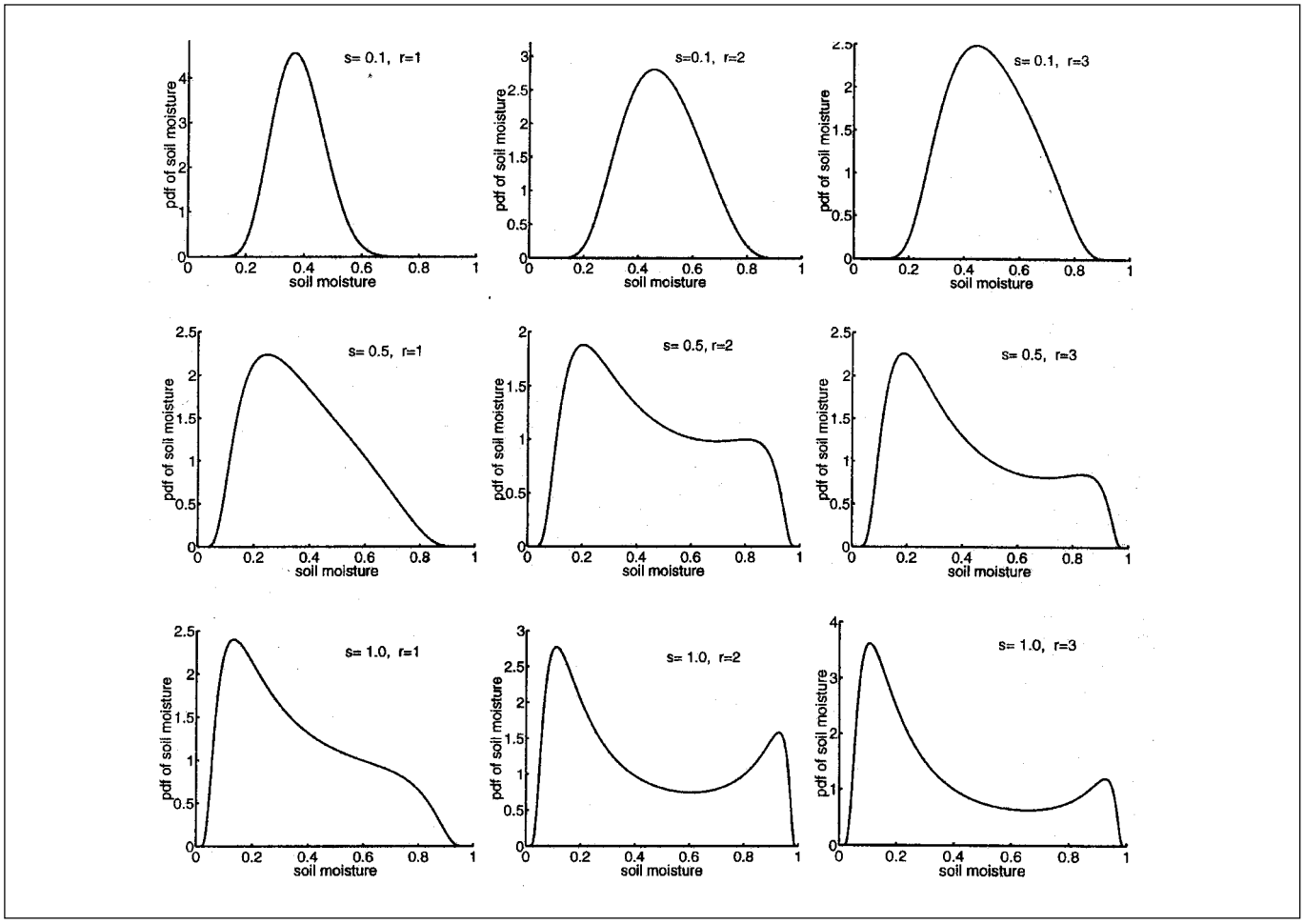

Fig. 4.5.1. Behaviour of the soil moisture for different infiltration functions and different intensity of environmental fluctuations (s) as indicated in the panels.

fluctuations in the forcing of the soil moisture balance equation lead to noise induced transitions from one stable mode to another. This is the hydrologic situation whereby a region may experience prolonged periods of drought and then there is an abrupt change to above average conditions until another dry spell occurs. Such persistent behaviour will have two stable modes: dry and wet. During each run there are stochastic fluctuations about the local mean. The regional hydrology may shift from one mode to another by one of these stochastic fluctuations. This occurs if the forcing is large enough to free the regional hydrologic balance away from one stable mode and lock it onto another.

We have estimated sensitivity of the dry/wet states due to variation in infiltration function parameter *r*. *Fig. 4.5.1* shows the probability density function (PDF) of soil moisture for different infiltration parameters *r* and for different intensities of environmental fluctuations ω. For ω = *0.1*, the PDF shows only one stable mode for all ϕ. This shows that when the intensity of fluctuation is low, the system coincides with deterministic case. When *w* is increased to 0.5, for *r=1*, PDF has only one mode which is shifted to left, i.e. towards dry mode. The case *r=2* shows two stable modes, dry and wet. This shows that when the intensity of fluctuations is more, noise induced transitions occur. The system experiences two modes for *r=3* also. When the intensity is increased to 1.0, the three cases of *r=1,2* and 3 show two stable modes. The wet mode for *r=2* and 3 are more prominent than the case for *r=1*. It is also observed that, as the intensity of fluctuations is increased, the dry mode is shifted towards left and wet mode is shifted towards right. It appears necessary, therefore, to include appropriate representation of (short term) stochastic processes in the modelling of ground hydrological processes in models of climate.

(R. Bhagyalakshmi, P. Goswami and R.N. Singh)