Optimal Beaver Population Management Using Reduced Order Distributed Parameter Model and Single Network Adaptive Critics

Radhakant Padhi[1] and **S. N. Balakrishnan**[2]
Department of Mechanical and Aerospace Engineering
University of Missouri - Rolla, MO, 65409, USA

Abstract

Using a distributed parameter model for beaver population that accounts for their spatial and temporal behavior, an optimal control for a desired distribution of the animals is presented. Optimal solutions are obtained through a "single network adaptive critic" (SNAC) neural network architecture.

Keywords: wildlife management, control

1. Introduction

Beavers are small mammal species and have a strong tendency to create nuisances, mainly by building dams on the flowing water thereby creating flooding the low land areas, roads, crop lands etc. However, the same activities sometimes lead to desirable consequences too- like increased vegetation, increased water table etc. However, the same activities sometimes lead to desirable consequences too- like increased vegetation, increased water table etc. Because of this conflicting situation, an optimal management strategy is needed to control their population[McKinstry, McTaggart].With the assumption that the neighboring land owners have a common goal, a distributed parameter model has been proposed in [Bhat]. An optimal harvesting strategy using this model has also been proposed [Bhat, Leinhart].

The main goal of this research is to design an "optimal" beaver harvesting scheme for a region of interest. Solving the associated Hamilton-Jacobi-Bellman (HJB) equation[Bryson] usually demands a very large amount of computations. Werbos proposed an innovative idea to get around this numerical complexity by using an 'Approximate Dynamic Programming (ADP)' formulation that uses two neural networks called adaptive critics. This paper uses a variant of the adaptive critic[Balakrishnan,Werbos] architecture that is named "Single Network Adaptive Critics (SNAC)" using a single network.

[1] Postdoctoral Fellow, Email: padhi@umr.edu
[2] Professor, Email: bala@umr.edu, Tel: (573)341-4675, Fax: (573)341-4607

2. Model and Controller Objective

2.1 Beaver Population Model

Assuming the beaver population distribution to be continuous in a territory, the following distributed parameter model has been developed in the literature for beaver population density [Bhat, Lenhart]:

$$\begin{aligned} \frac{\partial Z}{\partial t} &= \alpha \nabla^2 Z + \left(aZ - bZ^2\right) - PZ && \text{in } \Omega \times \left(o, t_f\right) \\ Z\left(y_1, y_2, t\right) &= 0 && \text{on } \partial\Omega \times \left(o, t_f\right) \\ Z\left(y_1, y_2, 0\right) &= Z_0\left(y_1, y_2\right) && \text{in } \Omega \text{ at } t = 0 \end{aligned} \quad (1)$$

where Z is the beaver population density in heads per square miles (hd/mi^2) and P is the portion of Z to be trapped per year (yr^{-1}), which acts as a control variable. α , a , b are growth parameters of the model (their meanings and values are in Table 1). Note that the term $\left(aZ - bZ^2\right)$ represents the density-dependent annual biological productivity of beavers in the absence of dispersion. Assuming that the spatial domain $\Omega \in \mathbb{R}^2$ and it is a rectangle, $\Omega \triangleq \left\{ y = \left(y_1, y_2\right): \quad y_1 \in \left[0, L_1\right], \quad y_2 \in \left[0, L_2\right] \right\}$, where L_1 and L_2 are the lengths of its sides. $\partial\Omega$ represents the boundary of Ω and time $t \in \left(0, t_f\right)$ and $Z_0\left(y_1, y_2\right)$ represents the initial density distribution. Based on a study for the state of New York, the parameters of the model [Bhat, Lenhart] are given in Table 1.

Table 1. Beaver Population Model Parameters

Symbol	Meaning	Units	Value
a	Maximum rate of net recruitment	yr^{-1}	0.335
b	Density dependence of beaver stock	$mi^2 \; hds^{-1}$	0.2066315
α	Diffusion coefficient	$mi^2 \; yr^{-1}$	725.27

It is clear from Eq.(1) that the growth (or decay) of the population density is a dynamic process that depends on the parameters of the model. Reasons for the diffusion term in the model include migration of two-year olds to set up new colonies, migration of the entire colonies for better food availability etc. Similarly, decay terms represent their natural demise, being eaten by predators, diseases due to contaminated water (their habitat is always close to water resources) etc.

2.2 Controller Objective

The objective of this controller in this study is to trap the beavers throughout the territory in an optimal way that leads to a desired distribution $Z^*(y)$ in the long run.

2.2.1 Choice of the Desired Distribution

The territory considered in this paper is a forest land and the desired distributions wanted by a wildlife manager $Z^*(y)$ is restricted to satisfy the following conditions:

(i) $Z^* \geq 0$ in $\Omega \setminus \partial\Omega$, $Z^* < 0$ is meaningless

(ii) $Z^* = 0$ on $\partial\Omega$

(iii) Z^* is continuous and smooth (i.e. $\nabla^2 Z^*$ continuous)

(iv) $\nabla^2 Z^* / Z^*$ is finite for $\Omega \cup \partial\Omega$

Condition (ii) is imposed because the boundary of the forest land usually consists of human habitation. However, the conditions $Z^* = 0$ and $\left(\nabla^2 Z^* / Z^*\right)$ being finite are in conflict. Hence, condition 1 is restricted to $Z^* > 0$ in $\Omega \setminus \partial\Omega$ and an approximation for condition (ii) that $Z^* \to 0^+$ on $\partial\Omega$ is introduced. One such approximation is

$$Z^*(y_1, y_2) = f(y_1)\, f(y_2) \tag{2}$$

$$f(y_1) = A \frac{1}{\sigma_1 \sqrt{2\pi}} \quad e^{-\frac{1}{2}\left(\frac{y_1 - \mu_1}{\sigma_1}\right)^2} \tag{3a}$$

$$f(y_2) = B \frac{1}{\sigma_1 \sqrt{2\pi}} \quad e^{-\frac{1}{2}\left(\frac{y_2 - \mu_2}{\sigma_2}\right)^2} \tag{3b}$$

Means of the distributions are selected as $\mu_1 = L_1/2, \mu_2 = L_2/2,\ 3\sigma_1 = L_1/2,\quad 3\sigma_2 = L_2/2$; For continuity in both y_1 and y_2 dimensions, a condition $f_1(\mu_1) = f_2(\mu_2)$ is imposed. This leads to $B = A(\sigma_2/\sigma_1)$ and $Z^*(y_1, y_2)$ becomes

$$Z^*(y_1, y_2) = \frac{A^2}{2\pi\sigma_1^2} \quad e^{-\frac{1}{2}\left[\left(\frac{y_1 - \mu_1}{\sigma_1}\right)^2 + \left(\frac{y_2 - \mu_2}{\sigma_2}\right)^2\right]} \tag{4}$$

Selection of this distribution is found to be good enough for this application, even though it does not satisfy the boundary condition of the model (see Eq.(1)) in a strict sense. For a particular selection of the parameter A, the total number of species in a rectangular territory having sides L_1 and L_2 for such a selection of Z^* can be computed as follows.

$$\begin{aligned} N^* &= \int_0^{L_1}\int_0^{L_2} Z^*(y_1, y_2)\, dy_2 dy_1 \\ &= \left(A\int_{\mu_1 - 3\sigma_1}^{\mu_1 + 3\sigma_1} f_1(y_1) dy_1\right)\left(B\int_{\mu_2 - 3\sigma_2}^{\mu_2 + 3\sigma_2} f(y_2) dy_2\right) \\ &= (0.9973)^2 A^2 (\sigma_2/\sigma_1) = (0.9973)^2 A^2 (L_2/L_1) \end{aligned} \tag{5}$$

2.2.2 Feed-forward Controller

Let P^* be the associated control with Z^* so that Z^* remains at steady state. Then from Eq.(1), it is clear that Z^* and P^* should satisfy the following equation:

$$\alpha\, \nabla^2 Z^* + Z^*\left(a - bZ^* - P*\right) = 0 \tag{6}$$

which leads to

$$P^* = \left(a - bZ^*\right) + \alpha\left(\frac{\nabla^2 Z^*}{Z^*}\right) \tag{7}$$

Note that the conditions (iii) and (iv) imposed on Z^* in Subsection 2.2.1 makes P^* well-behaved. By using Eqs.(4) and (7) P^* can be written as:

$$\begin{aligned} P^*(y_1, y_2) = \left(a - bZ^*\right) + \left[\left(\frac{y_1 - \mu_1}{\sigma_1^2}\right)^2 - \frac{1}{\sigma_1^2}\right]\ldots \\ + \left[\left(\frac{y_2 - \mu_2}{\sigma_2^2}\right)^2 - \frac{1}{\sigma_2^2}\right] \end{aligned} \tag{8}$$

One may observe from Eq.(8) that the steady state control P^* is a function of A *via* the steady state Z^*. However since $Z^* \to 0$ at the boundary, P^* is not a function of A at the boundary.

2.2.3 Deviation Dynamics and Cost Function

With the availability of the desired final values for state Z^* and control P^*, $Z \triangleq Z^* + x$ and $P \triangleq P^* + u$, where x and u are deviations in state and control respectively. Then it follows from Eq.(1) that

$$\frac{\partial x}{\partial t} = \alpha\, \nabla^2 x + \left(a - P^* - 2bZ^*\right)x - Z^* u + f(x, u) \tag{9}$$

where

$$\begin{aligned} & f(x,u) \triangleq \left(-bx^2 - xu\right) \\ & x(y_1, y_2, t) = 0 \ \text{ on } \partial\Omega \\ & x(y_1, y_2, 0) = Z_0(y_1, y_2) - Z^*(y_1, y_2) \ \text{ on } \Omega \end{aligned} \tag{10}$$

The goal of the controller design now is to cancel the deviation terms x and u throughout the domain. This can be achieved by finding a controller that minimizes

$$J = \frac{1}{2}\int_{0}^{t_f\to\infty}\int_0^{L_1}\int_0^{L_2}\left(qx^2+ru^2\right)dy_2dy_1\,dt \quad (11)$$

where $q \ge 0$ and $r > 0$ are the weights on state and control respectively.

3. Reduced Order Model Development

3.1 Basis Function Design Based on Proper Orthogonal Decomposition (POD)

Proper Orthogonal Decomposition (POD) is a technique of finding an optimal set of basis functions, which spans an ensemble of data optimally in an average sense. Let $\{U_i(y):\ 1 \le i \le N,\ y \in \Omega\}$ be an ensemble of data, consisting of set of N *snapshot* solutions (observations), of some physical process over the domain Ω at arbitrary instants of time. It is an effort to find all possible basis functions Φ, each of which provides a local maximum for the following figure of merit

$$I = \frac{1}{N\langle\Phi,\Phi\rangle}\sum_{i=1}^{N}\left|\langle U_i,\Phi\rangle\right|^2 \quad (12)$$

The problem is reduced to finding eigenvalues and eigenvectors where the normalized orthogonal eigenvectors are given by $W^i = \left[\,w_1^i \quad w_2^i \quad \cdots \quad w_N^i\,\right]^T$.

The N basis functions can be written as

$\Phi i = \sum_{i=1}^{N} w_i^i\, U_i(y)$.

The eigenspectrum can then be truncated judiciously such that $\sum_{j=1}^{\tilde N}\sigma_j \approx \sum_{j=1}^{N}\sigma_j$, where the truncated system has $\tilde N \le N$ eigenvalues and eigenvectors.

3.2 Reduced Order Model: Galerkin Projection

After obtaining the basis functions, x and u are expanded as follows

$$x(t,y) = \sum_{j=1}^{\tilde N}\hat x_j(t)\Phi_j(y_1,y_2) \quad (13)$$

$$u(t,y) = \sum_{j=1}^{\tilde N}\hat u_j(t)\Phi_j(y_1,y_2) \quad (14)$$

Note that:

- The principle of Galerkin projection [Holmes] is used after substituting Eqs.13-14 in Eqs.(9-10) to obtain the following reduced-order finite-dimensional model for the deviation dynamics.

$$\dot{\hat X} = \hat A\hat X + \hat B\hat U + \hat F\left(\hat X,\hat U\right) \quad (15)$$

where $\hat X \triangleq \left[\hat x_1 \ldots \hat x_{\tilde N}\right]^T$, $\hat U \triangleq \left[\hat u_1 \ldots \hat u_{\tilde N}\right]^T$ and

$$\begin{aligned}
\hat A_{nj} &= -\int_0^{L_1}\int_0^{L_2}\left[\nabla\Phi_n \bullet \nabla\Phi_j + \left(2bZ^* + P^*\right)\Phi_n\Phi_j\right]dy_2dy_1 + a f_{n_j}\\
\hat B_{nj} &= -\int_0^{L_1}\int_0^{L_2} Z^*\Phi_n\Phi_j dy_2dy_1\\
\hat F(\hat x,\hat u) &= \int_0^{L_1}\int_0^{L_2} f(x,u)\Phi_n dy_2dy_1 = -\int_0^{L_1}\int_0^{L_2}\left(bx^2 + xu\right)dy_2dy_1
\end{aligned} \quad (16)$$

Similarly substitutions for x and u from Eqs.(13-14) in the expression for the cost function Eq.(11) results in

$$J = \frac{1}{2}\int_0^{\infty}\left(\hat X^T\hat Q\hat X + \hat U^T\hat R\hat U\right)dt \quad (17)$$

where $\hat Q = qI_{\tilde N}$ and $\hat R = rI_{\tilde N}$

3.4 Snapshot Solution Generation

The spatial domain $\Omega \setminus \partial\Omega$ is discretized denoting $m_1 = 1,\ldots,M_1$ as the node points along y_1 and $m_2 = 1,\ldots,M_2$ as the node points along y_2. Then for $m_1 = 2,\ldots,(M_1-1)$ and $m_2 = 2,\ldots,(M_2-1)$ the following ordinary differential equations can be written [Gupta].

$$\begin{aligned}
\dot x_{m_1,m_2} = \alpha &\left[\begin{array}{l}\frac{1}{\Delta y_1^2}\left(x_{m_{1+1},m_2} - 2x_{m_1,m_2} + x_{m_{1-1},m_2}\right) + \ldots\\ \frac{1}{\Delta y_2^2}\left(x_{m_1,m_{2+1}} - 2x_{m_1,m_2} + x_{m_1,m_{2-1}}\right)\end{array}\right]\ldots\\
&+\left(a - 2bZ^*_{m_1,m_2} - P^*_{m_1,m_2}\right)x_{m_1,m_2} - Z^*_{m_1,m_2}U_{m_1,m_2}\ldots\\
&-bx^*_{m_1,m_2} - x_{m_1,m_2}U_{m_1,m_2}
\end{aligned} \quad (18)$$

It is also observed that $x_{m_1,m_2} = 0$ for either $m_1 = 1, M_1$ or $m_2 = 1, M_2$ for all time t (because of the boundary conditions). By defining

$X \triangleq \left[\left[x_{2,2}\ldots x_{M_1-1,2}\right]^T \vdots \cdots \vdots \left[x_{2,M_2-1}\ldots x_{M_1-1,M_2-1}\right]^T\right]$

$U \triangleq \left[\left[u_{2,2}\ldots u_{M_1-1,2}\right]^T \vdots \cdots \vdots \left[u_{2,M_2-1}\ldots u_{M_1-1,M_2-1}\right]^T\right]$,

the following finite-dimensional approximated system dynamics can be written as

$$\dot X = AX + BU + f\left(X,U\right) \quad (19)$$

where matrices A, B and the function $f\left(X,U\right)$ are appropriately defined (we have omitted the detailed expressions for brevity). Next the cost function was also approximated using this discretized system in the form

$$J = \frac{1}{2}\int_0^\infty \left(X^T QX + U^T RU \right) dt \qquad (20)$$

where $Q = (1/2)\, q\, (\Delta y_1 \Delta y_2)\, I_{(M_1-2)(M_2-2)}$ and $R = (1/2)\, r\, (\Delta y_1 \Delta y_2)\, I_{(M_1-2)(M_2-2)}$.

4. Single Network Adaptive Critics (SNAC)

4.1 Optimality Conditions

The necessary conditions of optimality for a lumped system driven by the system dynamics in Eq.(15) and cost function in Eq.(17) is

$$\hat{U} = -\hat{R}^{-1}\left[\hat{B}^T + \frac{\partial \hat{F}}{\partial \hat{U}}\left(\hat{X} \right) \right]^T \lambda \qquad (21)$$

where the nk^{th} element of $\left[\partial \hat{F} / \partial \hat{U} \right]$ matrix is given by

$$\frac{\partial \hat{F}_n}{\partial \hat{u}_k} = -\int_0^{L_1}\int_0^{L_2} x\, \Phi_k \Phi_n dy_2 dy_1 \qquad (22)$$

The costate equation is given by

$$\dot{\lambda} = -\frac{\partial H}{\partial \hat{X}} = -\left(\hat{Q}\hat{X} + \left[\hat{A} + \frac{\partial \hat{F}}{\partial \hat{X}} \right]^T \lambda \right) \qquad (23)$$

where the nk^{th} element of $\left[\partial \hat{F} / \partial \hat{X} \right]$ matrix is given by

$$\frac{\partial \hat{F}_n}{\partial \hat{x}_k} = -\int_0^{L_1}\int_0^{L_2} (2bx + u)\, \Phi_k \Phi_n\, dy_2 dy_1 \qquad (24)$$

Eq.(15), Eq.(21) and Eq.(23) need to be solved simultaneously, along with the boundary conditions for optimal control with $\hat{X}(0)$ is known and $\lambda\left(t_f \to \infty\right) = 0$.

4.2 Neural Network Synthesis Process

4.2.1 Sate Generation for Training

Let $\hat{X}_{\max}$ denote the vector of maximum values for $\hat{X}_k$ and $\hat{X}_{\min}$ the vector for minimum values. Then fixing a positive constant $0 \le c_i \le 1$, the states $\hat{X}_k \in c_i\left[\hat{X}_{\min}, \hat{X}_{\max} \right]$ are selected. Let $S_i = \left\{ \hat{X}_k : \hat{X}_k \in c_i\left[\hat{X}_{\min}, \hat{X}_{\max} \right] \right\}$. Then for $c_1 \le c_2 \le c_3 \le \ldots$, $S_1 \subseteq S_2 \subseteq S_3 \subseteq \ldots$. Hence, for some $i = I$, $c_I = 1$ and S_I will include the domain of interest for initial conditions. At the beginning a small value for the constant c_1 is fixed and the networks is trained with these states, randomly generated within S_1. Once the critic networks converge for this set, a higher value of c_i s are picked and the network training is continued until the set S_i includes domain of interest for the initial conditions.

4.2.2 Training Procedure

The SNAC training algorithm is described in Figure 1.

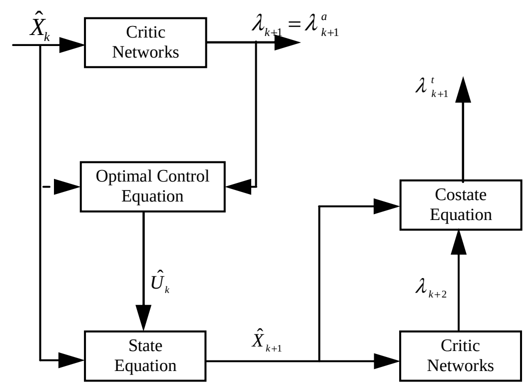

Figure 1: Schematic of SNAC synthesis

4.2.3 Network Structure

Since $\tilde{N} = 5$ for the beaver problem, five critic networks are used with X_k as inputs and a component of the vector λ_{k+1} as the output. For the wildlife management problem the network architecture is $\pi_{5,8,1}$ where $\pi_{5,8,1}$ means five neurons in the input layer, eight neurons in the hidden layer and one neuron in the output layer. For activation functions, a *tangent sigmoid* function for the input and hidden layers and a *linear* function for the output layer are used. Simulation results indicate the network choices were adequate.

5. Numerical Results

5.1 Selection of Numerical Values

The values of parameters used in the numerical experiments in this study are the same as used in [Bhat, Leinhart]. A spatial domain having $L_1 = 62.75$ miles and $L_2 = 112.95$ miles was selected and the grid parameters are $\Delta y_1 = \Delta y_2 = 12.55$ miles. The time step $\Delta t = (7/365)\, yr$ (one week), means that the control solution (rate of beavers to be harvested) is updated every one week. For the costhence N^*) in Eq.(4), random values were used for

$A \in [A_{\min}, A_{\max}]$ where $A_{\min} = 0.4\sqrt{2\pi}\ \sigma_1$ and $A_{\min} = 0.5\sqrt{2\pi}\ \sigma_1$ in the simulation.

5.2 Analysis of Results

The main goal in this study is to drive $Z \to Z^*$ and in the process drive the control $P \to P^*$ for any initial condition in the chosen domain of interest. A lot of initial profiles were used and simulation studies carried out where these conditions were met. However, since it is impossible to include a number of simulation results (due to space constraints), results for one representative random case is presented in Figures 2-9.

Figures 2 and 3 depict the steady state (or target) state and control profiles respectively. In other words, starting from any initial condition that has been accounted for training the networks, the sate and control should converge to these profiles with time.

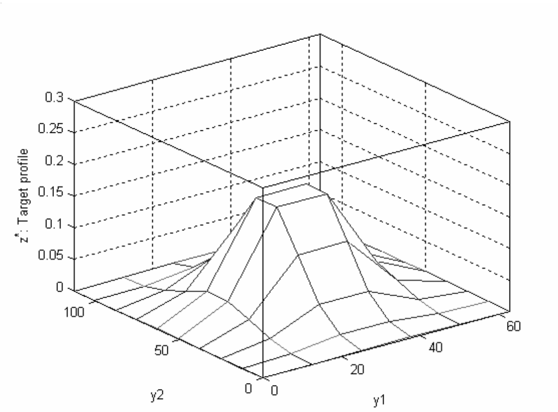

Figure 2: Target profile for state

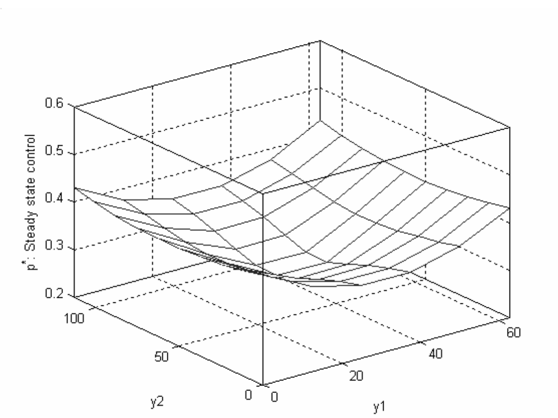

Figure 3: Target profile for control

The randomly chosen initial condition (for time $t = 0$) for this simulation and the associated control computed are shown in Figures 4 and 5 respectively.

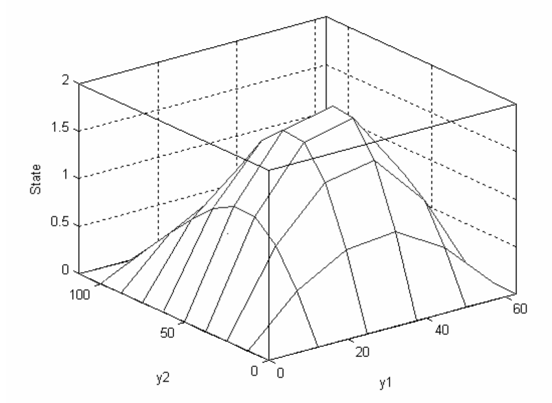

Figure 4: Initial condition for state

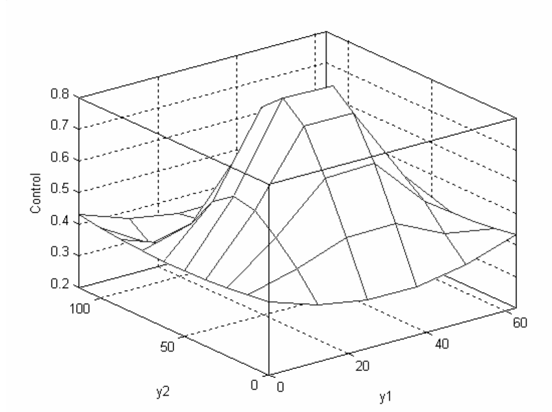

Figure 5: Control for the initial condition

Since it is impossible to show a three dimensional surface plot as it evolves, we have included the state and control at different time instants. At $t = 6$ months the state and control are as in Figures 6 and 7 respectively.

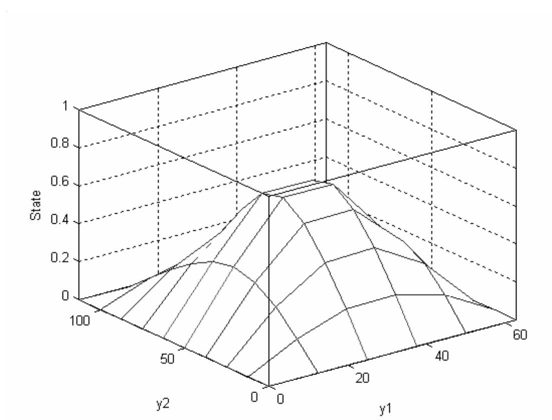

Figure 6: State at $t = 6$ months

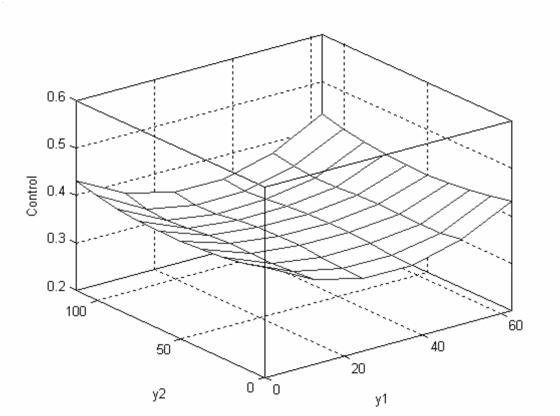

Figure 7: Control at $t = 6$ months

To further illustrate the way the state and control develop towards their steady state, we have included the time histories of the lumped parameter states and controls for the state and control deviations (see Section 3.2) in Figures 8 and 9 respectively.

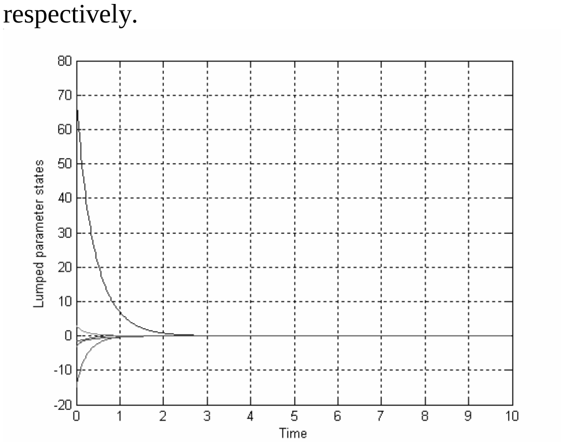

Figure 8: Lumped parameter state histories for the state deviation

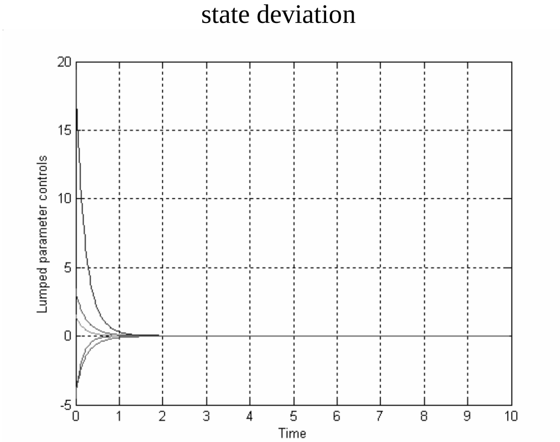

Figure 9: Lumped parameter control histories for the control deviation

From Figures 8 and 9 it is clear that the state and control converge to their respective targets in about two years time and it stays there afterwards. We observed this in all of the large number of simulations we have carried out. The results in Figures 2 through 9 clearly indicate that the control design achieved it objective.

6. Conclusions

The optimal harvesting technique presented for managing the beaver population leads to a healthy desired distribution. Hence this strategy will not invoke much of opposition from the animal conservationist and may be a great tool for a wildlife manager.

Acknowledgement

Dr. M. G. Bhat, of Florida International University, Miami, FL, was quite helpful. This research was supported by NSF grants 0201076 and 0324428.

References

1. Balakrishnan S. N. and Biega V., Adaptive-Critic Based Neural Networks for Aircraft Optimal Control, *J. of Guid., Cont.& Dynamics, Vol. 19, 1996, 893-898.*
2. Bhat M. G., Controlling Wildlife Damage by Diffusing Beaver Population: A Bioeconomic Application of the Distributed Parameter Control Model, *Ph.D. Dissertation, The University of Tennessee, Knoxville, 1991.*
3. Bryson A. E. and Ho Y. C., Applied Optimal Control, *London: Taylor and Francis, 1975.*
4. Holmes P., Lumley J. L. and Berkooz G., Turbulence, Coherent Structures, Dynamical Systems and Symmetry, *Cambridge University Press, 1996, 87-154.*
5. Lenhart S. M. and Bhat M. G., Application of Distributed Parameter Control Model in Wildlife Damage Control Management, *Math. Models and Meth. in Applied Sciences, Vol. 4, 1992, 423-439.*
6. McKinstry M. C. and Anderson S. H., Attitude of Private and Public managers in Wyoming, USA, Towards Beaver, *Environmental Management, Vol. 23, 1999, 95-101.*
7. McTaggart S. T. and Nelson T. A., Composition and Demographics of Beaver (Castor Canadensis) Colonies in Central Illinois, *American Midland Naturalist, Vol. 150, 2003, 139-150.*
8. Werbos P. J., Approximate dynamic programming for real-time control and neural modeling. In White D.A., & Sofge D.A (Eds.), *Handbook of Intelligent Control, Multiscience Press, 1992.*