OPTIMAL NEURAL FEEDBACK CONTROL FOR CARBON TAX POLICY GAUGING IN TRANSPORTATION

A. Alessandri,° C. Cervellera,°° and F. Grassia,°°°

° Institute for Studies in Intelligent Systems and Automation,
ISSIA-CNR National Research Council of Italy, Via De Marini 6, 16149 Genova, Italy *
Phone: +39-10-6475611, Fax: +39-10-6475600, Email: angelo@ian.ge.cnr.it

°° Department of Communications, Computer and System Sciences, DIST
University of Genoa, Via Opera Pia 13, 16145 Genova, Italy
Phone: +39-10-3532983, Fax: +39-10-3532948, Email: cristiano.cervellera@dist.unige.it

°°° Institute for Studies in Intelligent Systems and Automation,
ISSIA-CNR National Research Council of Italy, Via De Marini 6, 16149 Genova, Italy
Phone: +39-10-64751, Fax: +39-10-6475600, Email: filippo@ian.ge.cnr.it

Abstract

The effects of carbon emissions have been the objective of an investigation that was based on the model of the nation-wide transportation system with railway, waterway, and roadway. The dynamics of such a complex phenomenon depends on a set of control variables (i.e., the percentage of carbon tax on the fuel cost, the operational cost coverages, and growth rates of the various transportation modes) that can be chosen in a suitable way so as to minimize a given cost function (e.g., carbon emissions, public and private costs, fuel consumption, etc.). This problem has been addressed by searching for a feedback control law that can be approximated by means of the combination of both Dynamic Programming and neural networks. Preliminary simulation results with the afore-mentioned model are presented to demonstrate the effectiveness of the proposed method.

Keywords: carbon taxation, optimal control, feedback, neural networks.

1 Introduction

In recent years, much attention has been devoted to evaluate the impact of carbon abatement policies on economy (see, for example, (Wirl, 1991; Nail *et al.*, 92; Ford, 1994)). A carbon taxation policy provides an emission reduction but affects investment assets, technological change, and economic growth, which, in turn, influence the sources of emissions. Thus, the analysis of the questions is quite difficult and models have been proposed to help in setting economic intervention strategies (see, e.g., (Parry, 1999; Damania, 2000; Aronsson, 2001)).

The problem of determining a carbon tax policy by evaluating its effects on the overall transportation system of a country is the objective of this work. A model of the impact of the so-called green taxes on the transportation market has been proposed in (Piattelli *et al.*, 2002) and represents the starting point for devising a control policy that is aimed at being optimal by minimizing suitable performance indexes.

A computationally efficient design methodology turns out to be an essential tool in decision making, particularly if the dynamic model is complex and affected by exogenous inputs like, for example, transportation demand and cost of fuel. In such cases, a closed-loop strategy is preferable as it allows to account for the capacity of the economy governor to react in the presence of unpredicted, abrupt changes in the values of the above-written inputs.

Unfortunately, most tools for optimization in economy applications allow one to solve only open-loop control problems or are based on linearization procedures that provide linear control strategies (see, for example, (Islam and Craven, 2001; Neck, 1999; Neck, 2001)). A different approach has been followed with respect to the above-mentioned optimization techniques as the optimal policy design is accomplished by solving a functional optimization problem in an approximate way. More precisely, a cost functional has to be minimized with respect to a sequence of decisions described by feedback control functions. For example, in the discrete-time case, the cost is related to the evolution of a system state over a finite horizon of time stages, in the presence of random disturbances of which we suppose to know the probability density functions.

The "classic" algorithm for the solution of such kind of problems is Dynamic Programming (Bellman, 1957; Bertsekas, 2000) (DP in the following). However, it is known that the DP equations can be solved analytically only in simple cases, or under special assumptions on the system and the cost function (typically, linear system and quadratic cost).

For the general case we must look for approximate solutions. This leads to the discretization of the state space, and to the need of approximating the cost-to-go functions for the points that do not belong to such discretization. Unfortunately, the necessity of dealing with grids of samples in the state space causes dimensionality issues that have been faced in literature in different ways, typically by recurring to a simpler model or cost function (see, for example, (Yakowitz, 1982; Archibald *et al.*, 1997)), or by using more "powerful" approximating schemes (see, among others, (Bellman *et al.*, 1963; Johnson *et al.*, 1993)).

In order not to incur an exponential growth of the computational requirements

(commonly known as "curse of dimensionality"), we employ neural networks for their advantageous approximating properties (see (Haykin, 1999) for a detailed description of many neural architectures and theoretical properties) and Montecarlo-like sampling techniques (Bratley *et al.*, 1987) for the various discretizations. In this way it is possible to cope with the high-dimensional context of the carbon-tax model, and actually solve the optimization problem. The experimental results, compared to a fixed unoptimized policy, show the goodness of the method, and the advantages of optimization.

The paper is organized as follows. Section 2 is devoted to the description of the model used to apply a neural DP-based approach for the selection of the control policy. Such method is presented in Section 3. The simulation results are illustrated in Section 4. Finally, the conclusions are drawn in Section 5.

2 A model for the transportation system and carbon taxation

In this section, a brief description of the model proposed in (Piattelli *et al.*, 2002) is presented. The complexity of the system prevents from a complete discussion on the overall model and the interested reader can refer to (Piattelli *et al.*, 2002) for details (see Fig. 1).

The objective of modelling consists in bringing together the knowledge on a system in order to get a deeper understanding of the phenomena and provide support for supervision and/or control. In the case study (Piattelli *et al.*, 2002) regarding the German economy, the complexity of the Transportation System (TS, for short) is somehow tackled by introducing a description of the carbon tax impact on the growth of the transportation market that is shared among railway, waterway, and roadway. Each transportation mode relies on its own network, where trains, ships, and trucks perform, respectively, depending on infrastructure investment, network extension, and taxation policy. For the reader's convenience, the state variables, control and exogenous inputs are shown in Tables 1, 2, and 3, respectively, with the corresponding measurement units, where $1\,DM \triangleq one\,German\,Mark$, $1\,Y \triangleq one\,year$, $1\,Mton \triangleq 10^6 t$, and $1\,KKm \triangleq 10^3 Km$.

Note that the cost of fossil fuel results from the sum of industrial price and taxation, which depends on the country policy. The total cost of fuel determines the fares of the various transportation modes. Moreover, for the sake of compactness, let the fare cost functions be defined as follows. For the railway, the fare cost of transportation is given by

$$F_R(x_1, u_1) = [(C_f + x_1)\,\alpha_R + C_{R0}]\,(1 - u_1)$$

where C_f is fuel net cost in DM/m^3, $\alpha_R = 1.7 \cdot 10^5\,m^3/(Mton\,KKm)$ is the mean railway fuel consumption rate, $C_{R0} = 50 \cdot 10^6\,DM/(Mton\,KKm)$ is the amount of fixed cost for railway transportation, and u_1 is the public coverage of the operational cost for railway ($u_1 \in [0, 1]$).

The fare cost for waterway transportation is

$$F_W(x_1, u_2) = [(C_f + x_1)\,\alpha_W + C_{W0}]\,(1 - u_2)$$

x_1	carbon tax	DM/m^3
x_2	transportation demand	$Mton\,KKm/Y$
x_3	railway network extension	KKm
x_4	railway transportation capacity	$Mton/Y$
x_5	number of train deliveries	
x_6	railway transportation amount	$Mton\,KKm/Y$
x_7	inland waterway network extension	KKm
x_8	inland waterway transportation capacity	$Mton/Y$
x_9	number of inland ship deliveries	
x_{10}	inland waterway transportation amount	$Mton\,KKm/Y$
x_{11}	roadway network extension	KKm
x_{12}	roadway transportation capacity	$Mton/Y$
x_{13}	number of truck deliveries	
x_{14}	roadway transportation amount	$Mton\,KKm/Y$

Table 1: State variables with measurement units.

u_1	operational cost coverage for railway	
u_2	operational cost coverage for waterway	
u_3	operational cost coverage for roadway	
u_4	yearly carbon tax rate	1/Y
u_5	yearly railway growth rate	1/Y
u_6	yearly waterway growth rate	1/Y
u_7	yearly roadway growth rate	1/Y

Table 2: Input variables with measurement units.

C_f	fuel cost	DM/m^3
r	yearly transportation demand growth rate	1/Y

Table 3: Exogenus inputs with measurement units.
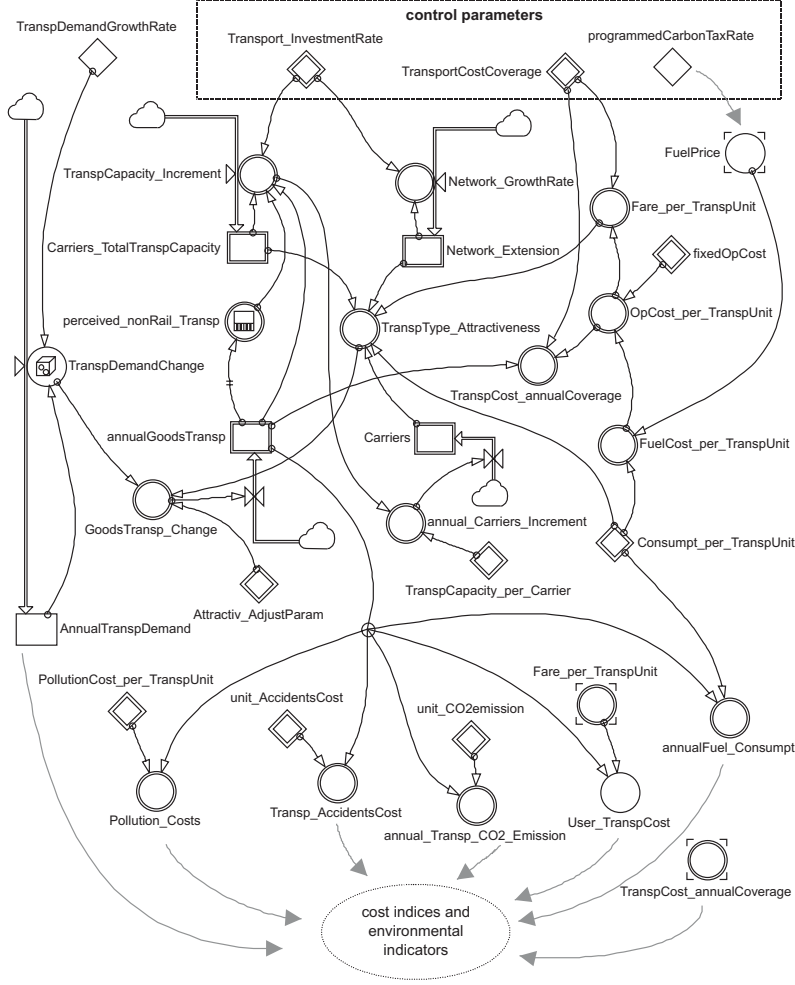
Figure 1: Flow diagram of the model.

where $\alpha_W = 1.3 \cdot 10^5 \, m^3/(Mton \, KKm)$ is the mean waterway fuel consumption rate, $C_{W0} = 1.79 \cdot 10^8 \, DM/(Mton \, KKm)$ is the amount of fixed cost for waterway transportation, and u_2 is the public coverage of the operational cost for waterway ($u_2 \in [0,1]$).

For roadway transportation, the fare function has been taken equal to

$$F_T(x_1, u_3) = [(C_f + x_1) \, \alpha_T + C_{T0}] \, (1 - u_3)$$

where $\alpha_T = 4.1 \cdot 10^5 \, m^3/(Mton \, KKm)$ is the mean roadway fuel consumption rate, $C_{W0} = 2.49 \cdot 10^9 \, DM/(Mton \, KKm)$ is the amount of fixed cost for roadway transportation, and u_3 is the public coverage of the operational cost for roadway ($u_3 \in [0,1]$).

2.1 Carbon tax

The carbon tax is accounted for by the simple equation

$$\dot{x}_1 = x_1 \, u_4 \tag{1}$$

where u_4 is the yearly percentage rate of taxation.

2.2 Transportation demand

The transportation demand dynamics is described by

$$\dot{x}_2 = x_2\, r \tag{2}$$

where r is the yearly percentage of growth rate. r is usually affected by a noise. For example, the trend of growth may be 2.7 % a year with random fluctuations in the range between -3 % and +3 %.

2.3 Railway transportation

The model of railway TS is composed of three different subsystems, i.e., (i) network extension (in KKm), (ii) the number of carrier deliveries, and (iii) capacity (in $Mton$). On the basis of (Piattelli *et al.*, 2002), railway capacity and network grow according to the scheduled growth rate. Thus, the dynamics of railway network extension and capacity are given by

$$\dot{x}_3 = x_3\, u_5 \tag{3}$$

$$\dot{x}_4 = x_4\, u_5 \tag{4}$$

respectively, where u_5 is the yearly percentage of growth rate in railway transportation.

The variation in the number of train deliveries is represented by means of

$$\dot{x}_5 = \frac{x_4}{C_1} \tag{5}$$

where $C_1 \triangleq 360\, t$ is the mean transportation capacity of a train.

The behavior of the transportation amount may be somehow related to the its own "appeal" that depends mainly on the transportation demand but also on other factors like, for example, extension of the network, capacity, number of carrier deliveries, fuel cost, and transportation fare. Such a modeling paradigm in general case studies has been proposed in (Senge, 1975). In our case, each of the aforementioned variables provides a contribution to the attractiveness of the transportation mode that may be expressed by the following dynamic equation (see, for details, (Piattelli *et al.*, 2002)):

$$\dot{x}_6 = C_2\, x_2 \left\{ \frac{2\, x_3}{x_3 + x_7 + x_{11}} + \frac{x_4}{x_4 + x_8 + x_{12}} + \frac{x_5}{x_5 + x_9 + x_{13}} \right.$$
$$\left. + \frac{\alpha_R}{\alpha_R + \alpha_W + \alpha_T} + \frac{1}{\left[\dfrac{1}{F_R(x_1, u_1)} + \dfrac{1}{F_W(x_1, u_2)} + \dfrac{1}{F_T(x_1, u_3)} \right] F_R(x_1, u_1)} \right\} \tag{6}$$

where $C_2 \triangleq 0.87$ is an adjustment coefficient. Roughly speaking, the transportation amount grows as much as demand increases. Moreover, such variations depend on the characteristics of the transportation mode, e.g., railway is as more attractive as the network is larger, the capacity is bigger, and the number of train is more

numerous (we assume that it is proportional to the resulting load deliveries). As outlined in (Piattelli *et al.*, 2002), also the rate of fuel consumption reflects the attractiveness of the transportation mode, in the sense that consumers prefer to use the high-consuming transportation means. However, a TS is more appealing if the fare is lower. It is worth noting that the network is made more important than the other contributions are by means of the selection of a parameter equal to 2 in (6) (see (Piattelli *et al.*, 2002)).

2.4 Waterway transportation

The model of the inland waterway TS is composed of four dynamic equation concerning network, carriers, capacity, and transported quantity. The network changes are modelled by using

$$\dot{x}_7 = x_7 \, u_6 \tag{7}$$

where u_6 is the yearly percentage of waterway growth rate.

The dynamics of the transportation capacity is given by

$$\dot{x}_8 = C_4 \, x_8 \, x_{10} \tag{8}$$

where $C_4 = 0.5/KKm$ is a constant parameter. In other words, the transportation capacity grows as more as the waterway transportation amount increases. This relationship is quite reasonable as, unlike what cast for the railway model, the waterway TS is not directly controlled by the Government and its evolution depends on the market (see, for details, (Piattelli *et al.*, 2002)).

The variation in the number of ship deliveries is represented by means of

$$\dot{x}_9 = \frac{x_8}{C_3} \tag{9}$$

where $C_3 \triangleq 980 \, t$ is the mean transportation capacity of a ship for inland navigation.

The attractiveness of waterway TS may be expressed like in (6), i.e.,

$$\dot{x}_{10} = C_5 \, x_2 \left\{ \frac{2 \, x_7}{x_3 + x_7 + x_{11}} + \frac{x_8}{x_4 + x_8 + x_{12}} + \frac{x_9}{x_5 + x_9 + x_{13}} \right.$$
$$\left. + \frac{\alpha_W}{\alpha_R + \alpha_W + \alpha_T} + \frac{1}{\left[\dfrac{1}{F_R(x_1, u_1)} + \dfrac{1}{F_W(x_1, u_2)} + \dfrac{1}{F_T(x_1, u_3)} \right] F_W(x_1, u_2)} \right\} \tag{10}$$

where $C_5 \triangleq 0.94$ is an adjustment coefficient. The same comments on (6) in Section 2.3 apply to (10).

2.5 Roadway transportation

The dynamics of network, capacity, carrier deliveries, and transportation amount for the roadway TS is given by a set of four equations similar to those of the waterway TS.

The dynamics of the roadway network is

$$\dot{x}_{11} = x_{11}\, u_7 \tag{11}$$

where u_7 is the percentage of roadway growth rate. The dynamics of the transportation capacity resembles to that of the waterway one, i.e.,

$$\dot{x}_{12} = C_7\, x_{12}\, x_{14} \tag{12}$$

where $C_7 = 1.0/KKm$ is a constant parameter. Like for the waterway TS dynamics, the roadway transportation capacity grows as much more as the corresponding transportation amount increases and depends on the transportation market.

The evolution in the number of truck deliveries may be modelled as follows:

$$\dot{x}_{13} = \frac{x_{12}}{C_6} \tag{13}$$

where $C_6 \triangleq 3.8\, t$ is the mean transportation capacity of a truck.

The attractiveness of this TS influences the transportation amount like in (6) and (10), i.e.,

$$\dot{x}_{14} = C_8\, x_2 \left\{ \frac{2\, x_{11}}{x_3 + x_7 + x_{11}} + \frac{x_{12}}{x_4 + x_8 + x_{12}} + \frac{x_{13}}{x_5 + x_9 + x_{13}} + \right.$$
$$\left. \frac{\alpha_T}{\alpha_T + \alpha_W + \alpha_T} + \frac{1}{\left[\dfrac{1}{F_R(x_1, u_1)} + \dfrac{1}{F_W(x_1, u_2)} + \dfrac{1}{F_T(x_1, u_3)}\right] F_T(x_1, u_3)} \right\}$$

$$\tag{15}$$

where $C_8 \triangleq 1.074$ is an adjustment coefficient.

2.6   Cost function for performance evaluation

Different cost functions have been considered to assess the above-described (see, for details, (Piattelli *et al.*, 2002)). Such cost functions for a generic time interval from 0 to $T > 0$ are defined as follows:

$J_1 \triangleq \int_0^T x_{14}\, dt$: roadway transportation amount (in $Mton\, KKm$);

$J_2 \triangleq \int_0^T \left(\beta_{ER}\, x_6 + \beta_{EW}\, x_{10} + \beta_{ET}\, x_{14}\right) dt$: CO_2 emission (in m^3), where $\beta_{ER} = 48.1\, m^3/(Mton\, KKm)$, $\beta_{EW} = 33.4\, m^3/(Mton\, KKm)$, and $\beta_{ET} = 164.0\, m^3/(Mton\, KKm)$ are the CO_2 emission rates for railway, waterway, and roadway transportation mode, respectively;

$J_3 \triangleq \int_0^T \left(\alpha_R\, x_6 + \alpha_W\, x_{10} + \alpha_T\, x_{14}\right) dt$: fuel consumption (in m^3);
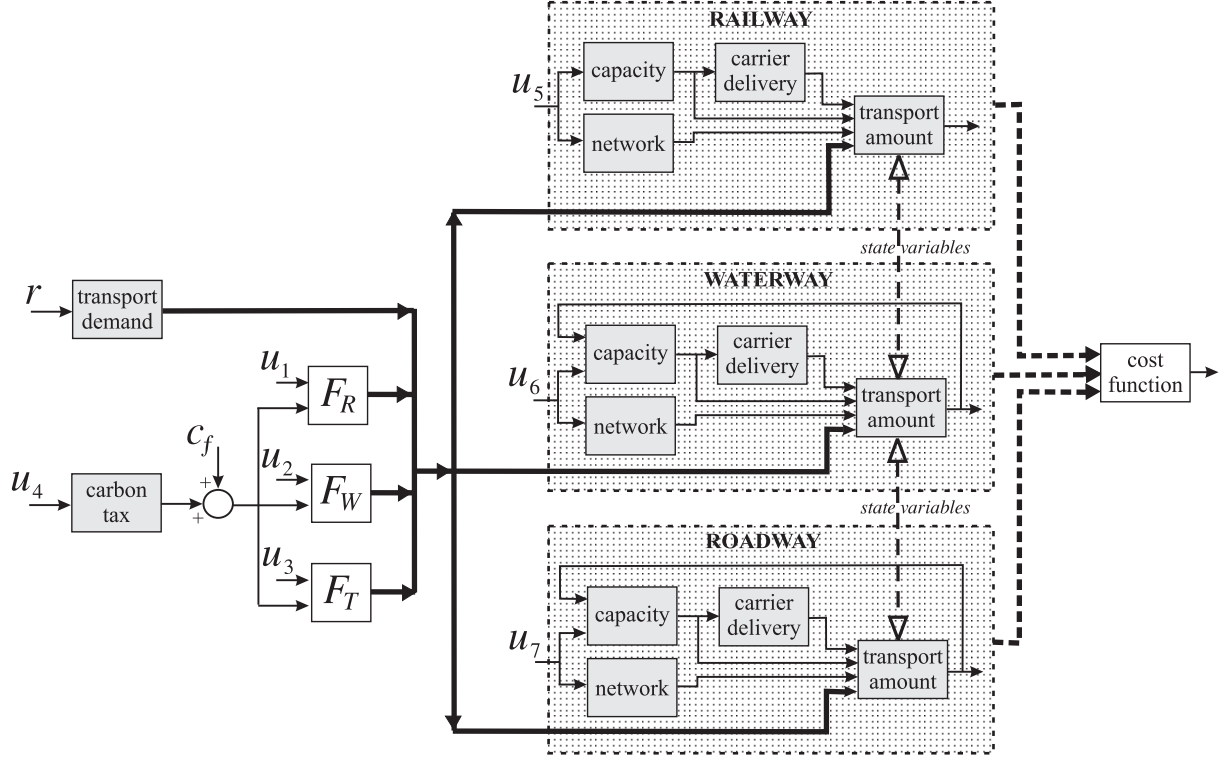
Figure 2: Block diagram of the model (grey and white boxes correspond to dynamic and algebraic mappings, respectively).

$J_4 \triangleq \int_0^T \left(\gamma_{PR}\, x_6 + \gamma_{PW}\, x_{10} + \gamma_{PT}\, x_{14} \right) dt$: pollution cost (in DM), where $\gamma_{PR} = 11.5 \cdot 10^6\, DM/(Mton\,KKm)$, $\gamma_{PW} = 3.5 \cdot 10^6 DM/(Mton\,KKm)$, and $\gamma_{PT} = 50.1\, DM/(Mton\,KKm)$ are the pollution cost rates for railway, waterway, and roadway transportation mode, respectively;

$$J_5 \triangleq \int_0^T \Big[\left(\gamma_{PR} + \left((C_f + x_1)\, \alpha_R + C_{R0} \right) u_1 - \alpha_R x_1 \right) x_6$$
$$+ \left(\gamma_{PW} + \left((C_f + x_1)\, \alpha_W + C_{W0} \right) u_2 - \alpha_W x_1 \right) x_{10}$$
$$+ \left(\gamma_{PT} + \left((C_f + x_1)\, \alpha_T + C_{T0} \right) u_3 - \alpha_T x_1 \right) x_{14} \Big] dt$$: carbon tax net public cost (in DM);

$J_6 \triangleq \int_0^T \left[F_R(x_1, u_1)\, x_6 + F_W(x_1, u_2)\, x_{10} + F_T(x_1, u_3)\, x_{14} \right] dt$: user cost (DM);

$J_7 \triangleq J_4 + J_6$: pollution and user cost (DM);

$J_8 \triangleq J_5 + J_6$: total (public and user) cost (in DM);

$J_9 \triangleq J_5 + \int_0^T \left(\gamma_{IR}\, x_6 + \gamma_{IR}\, x_{10} + \gamma_{IT}\, x_{14} \right) dt$: cost for transportation incidents (in DM), where $\gamma_{IR} = 1.2 \cdot 10^6\, DM/(Mton\,KKm)$, $\gamma_{IW} = 0.01 \cdot 10^6 DM/(Mton\,KKm)$, and $\gamma_{IT} = 17.8 \cdot 10^6\, DM/(Mton\,KKm)$ are the incident cost rates for railway, waterway, and roadway transportation mode, respectively.

J_1	roadway transportation amount	$Mton\,KKm$
J_2	CO_2 emission	m^3
J_3	fuel consumption	m^3
J_4	pollution cost	DM
J_5	carbon tax net public cost	DM
J_6	user cost	DM
J_7	pollution and user cost	DM
J_8	total (public and user) cost	DM
J_9	cost for transportation incidents	DM

Table 4: Cost functions with measurement units.

Table 4 summarizes the above-defined cost functions.

2.7 Discretization

A pictorial representation of the overall model is shown in Fig. 2. Equations (1)–(14) have been discretized by using a simple Euler's method with sample time ΔT equal to Y/τ, where $\tau \in \mathbb{N}$, $\tau \geq 1$, the number of samples in which the unit time Y is divided. The complete model includes 16 auxiliary state variables $\chi_i(t)$, $i = 1, 2, \ldots, 16$. Thus, we will refer to a state vector $\underline{x}_k \triangleq \mathrm{col}\,\{\mathrm{col}\,[x_i(k\,\Delta T)\,,\,i=1,2,\ldots,14]\,,\,\mathrm{col}\,[\chi_i(k\,\Delta T)\,,\,i=1,2,\ldots,16]\,\}$, input vector $\underline{u}_k \triangleq \mathrm{col}\,[u_i(k\,\Delta T)\,,\,i=1,2,\ldots,7]$, and disturbance vector $\underline{\xi}_k \triangleq \mathrm{col}\,[r(k\,\Delta T), c_f(k\,\Delta T)]$.

In the following, for the sake of compactness, let $h_p(\underline{x}, \underline{u})$ be the integrand of the corresponding cost function J_p, $p = 1, 2, \ldots, 9$, i.e., $J_p \triangleq \int_0^T h_p(\underline{x}, \underline{u})\,dt$. Moreover, it is useful to re-state the system in terms of years of evolution. Specifically, since a year corresponds to the evolution of the system for τ temporal stages, we can define $\underline{\tilde{x}}_t \triangleq \underline{x}_{\tau t}$, as the new state vector.

If we suppose that the input vector and the disturbance vectors remain the same during a whole year, we can write the new state equation as

$$\underline{\tilde{x}}_{t+1} = \tilde{f}(\underline{\tilde{x}}_t, \underline{\tilde{u}}_t, \underline{\tilde{\xi}}_t)$$

where $\underline{\tilde{u}}_t$ and $\underline{\tilde{\xi}}_t$ are the input vector and the disturbance vector of the t-th year, and $\tilde{f}$ keeps track of the evolution of f during the τ stages that cover one year. Thus, in order to minimize the cost indexes during the whole year, we can define the new functions

$$\tilde{h}_t^{(p)}(\underline{\tilde{x}}_t, \underline{\tilde{u}}_t, \underline{\tilde{\xi}}_t) = \sum_{j=0}^{\tau} h^{(p)}(\underline{x}_{\tau\,t+j}, \underline{\tilde{u}}_t, \underline{\tilde{\xi}}_t) \quad , \quad p = 1, 2, \ldots, 9$$

Therefore, the cost function for the whole horizon can be written, for $p = 1, \ldots, 9$, as

$$\tilde{J}^{(p)} = \sum_{t=0}^{T-1} \tilde{h}_t^{(p)}(\underline{\tilde{x}}_t, \underline{\tilde{u}}_t, \underline{\tilde{\xi}}_t) \quad .$$

3 A Dynamic Programming scheme for optimization

As we have defined the state vector, the input vector and the disturbance vector, we can state the optimization problem in the following way:

Problem P

Find the optimal control law $\underline{\mu}^\circ = \mathrm{col}(\underline{\mu}_0^\circ, \ldots, \underline{\mu}_{T-1}^\circ)$ that minimizes

$$F^{(p)}(\underline{\mu}) = \mathop{\mathrm{E}}_{\underline{\tilde{\xi}}} \tilde{J}^{(p)}(\underline{\tilde{x}}_0, \underline{\mu}, \underline{\tilde{\xi}})$$

where $\underline{\tilde{x}}_0$ is a given initial state and $\underline{\tilde{\xi}} \triangleq \mathrm{col}(\underline{\tilde{\xi}}_0, \ldots, \underline{\tilde{\xi}}_{T-1})$, subject to the constraints

$$\underline{\tilde{u}}_t = \underline{\mu}_t(\underline{\tilde{x}}_t) \in U_t, \quad t = 0, \ldots, T-1$$

and

$$\underline{\tilde{x}}_{t+1} = \tilde{f}_t(\underline{\tilde{x}}_t, \underline{\tilde{u}}_t, \underline{\tilde{\xi}}_t), \quad t = 0, \ldots, T-1 \quad .$$

□

We define the *feasible space* $\hat{X}_t$ for the t-th state space as the set of all points that can be actually reached from year $t-1$ by applying all the possible input vectors $\underline{\tilde{u}}_{t-1} \in U_t$ and all the possible random vectors $\underline{\tilde{\xi}}_{t-1} \in D_t$:

$$\hat{X}_t \triangleq \left\{ \underline{\tilde{x}}_t = \tilde{f}(\underline{\tilde{x}}_{t-1}, \underline{\tilde{u}}_{t-1}, \underline{\tilde{\xi}}_{t-1}), \underline{\tilde{x}}_{t-1} \in \hat{X}_{t-1}, \underline{\tilde{u}}_{t-1} \in U_{t-1}, \underline{\tilde{\xi}}_{t-1} \in D_{t-1} \right\}$$

$$\hat{X}_0 \triangleq \{\underline{\tilde{x}}_0\}$$

As already said, the well known Dynamic Programming equations are the main method of solution for problems like Problem P. Anyway, as we are not under "LQ" hypotheses, we have to solve them numerically.

In order to do so we discretize each feasible space in the most uniform way by L points, which form the following sets

$$X_{tL} = \left\{ \underline{\tilde{x}}_t^{(l)} \in \hat{X}_t : l = 1 \ldots, L \right\}, \quad t = 1 \ldots, T-1$$

For notational convenience, in the following we will omit the superscript p. Let us now write, for stage $T-1$:

$$J_{T-1}^\circ(\underline{\tilde{x}}_{T-1}^{(l)}) = \min_{\underline{\tilde{u}}_{T-1} \in U_{T-1}} \mathop{\mathrm{E}}_{\underline{\tilde{\xi}}_{T-1}} \left\{ \tilde{h}_{T-1}(\underline{\tilde{x}}_{T-1}^{(l)}, \underline{\tilde{u}}_{T-1}, \underline{\tilde{\xi}}_{T-1}) \right.$$

$$\left. + \underline{\tilde{h}}_T[\tilde{f}(\underline{\tilde{x}}_{T-1}^{(l)}, \underline{\tilde{u}}_{T-1}, \underline{\tilde{\xi}}_{T-1})] \right\}, \quad \underline{\tilde{x}}_{T-1}^{(l)} \in X_{T-1,L}$$

Once we have obtained the L pairs $[\underline{\tilde{x}}_{T-1}^{(l)}, J_{T-1}^\circ(\underline{\tilde{x}}_{T-1}^{(l)})]$, we can approximate the cost-to-go function J_{T-1}° for the whole $X_{T-1,L}$ by means of a neural network having the structure $\hat{J}_{T-1}(\underline{\tilde{x}}_{T-1}, \underline{w}_{T-1})$, where $\underline{w}_{T-1} \in \mathbb{R}^K$ is the vector of the parameters of the network.

Specifically, we obtain the optimal parameter vector $\underline{w}^{\circ}_{T-1}$ by minimizing the empirical risk in this way

$$\underline{w}^{\circ}_{T-1} = \arg\min_{\underline{w}_{T-1}} \sum_{l=1}^{L} \left[J^{\circ}_{T-1}(\underline{\tilde{x}}^{(l)}_{T-1}) - \hat{J}^{\circ}_{T-1}(\underline{\tilde{x}}^{(l)}_{T-1}, \underline{w}_{T-1}) \right]^2$$

By the means of the newly obtained $\hat{J}_{T-1}(\underline{\tilde{x}}_{T-1}, \underline{w}^{\circ}_{T-1})$, we can solve the DP equations for stage $T-2$, i.e., we compute

$$\bar{J}^{\circ}_{T-2}(\underline{\tilde{x}}^{(l)}_{T-2}) = \min_{\underline{\tilde{u}}_{T-2} \in U_{T-2}} \operatorname*{E}_{\underline{\tilde{\xi}}_{T-2}} \left\{ \tilde{h}_{T-2}(\underline{\tilde{x}}^{(l)}_{T-2}, \underline{\tilde{u}}_{T-2}, \underline{\tilde{\xi}}_{T-2}) \right.$$

$$\left. + \hat{J}_{T-1}[\underline{\tilde{f}}(\underline{\tilde{x}}^{(l)}_{T-2}, \underline{\tilde{u}}_{T-2}, \underline{\tilde{\xi}}_{T-2}), \underline{w}^{\circ}_{T-1}] \right\}, \quad \underline{\tilde{x}}^{(l)}_{T-2} \in X_{T-2,L}$$

This time $\bar{J}^{\circ}_{T-2}$ is an approximation of the real J°_{T-2}, due to the fact that we use the neural network $\hat{J}_{T-1}$ in the DP equation, instead of the real J°_{T-1} (which is unknown).

Once again, we can approximate $\bar{J}^{\circ}_{T-2}$ by a new neural network $\hat{J}_{T-2}(\underline{\tilde{x}}_{T-2}, \underline{w}_{T-2})$. The optimal parameter vector $\underline{w}^{\circ}_{T-2}$ is given by

$$\underline{w}^{\circ}_{T-2} = \arg\min_{\underline{w}_{T-2}} \sum_{l=1}^{L} \left[\bar{J}^{\circ}_{T-2}(\underline{\tilde{x}}^{(l)}_{T-2}) - \hat{J}^{\circ}_{T-2}(\underline{\tilde{x}}^{(l)}_{T-2}, \underline{w}_{T-2}) \right]^2 \quad .$$

If this procedure is repeated for $t = T-3, \dots, 1$, we obtain "off-line" all the neural network approximations $\hat{J}_t(\underline{\tilde{x}}_t, \underline{w}_t)$ that can be used for the "on-line" minimization of the cost. Thus, in general, at a given year t and for a given state $\underline{\hat{x}}_t$, the approximate optimal control $\underline{\hat{u}}^{\circ}_t = \underline{\hat{\mu}}^{\circ}_t(\underline{\hat{x}}_t)$ is given by

$$\underline{\hat{u}}^{\circ}_t = \arg\min_{\underline{\tilde{u}}_t \in U_t} \operatorname*{E}_{\underline{\tilde{\xi}}_t} \left\{ \tilde{h}_t(\underline{\hat{x}}_t, \underline{\tilde{u}}_t, \underline{\tilde{\xi}}_t) + \hat{J}_{t+1}[\underline{\tilde{f}}(\underline{\hat{x}}_t, \underline{\tilde{u}}_t, \underline{\tilde{\xi}}_t), \underline{w}^{\circ}_t] \right\}$$

and the new state vector is computed by

$$\underline{\hat{x}}_{t+1} = \underline{\tilde{f}}(\underline{\hat{x}}_t, \underline{\hat{u}}^{\circ}_t, \underline{\hat{\xi}}_t)$$

where $\underline{\hat{\xi}}_t$ is the actual random disturbance acting on the system "on-line".

The discretization issue deserves a comment. In fact, the key for the success of the neural DP method is that the "training sets" $X_{t,L}$ for the various networks must grow "moderately" with the dimension d of the state vector.

It is well known that the "classic" choice of uniform discretization of each component of the state vector in the same number of levels does not satisfy this requirement, as it leads to an exponential growth of L with the dimension d ("curse of dimensionality").

Montecarlo methods can be applied in order to generate sequences of points with L that does not depend "structurally" on d. In fact it is possible to randomly extract points with uniform probability, in order to have sequences of any desired length L, for each dimension d. Of course, we still must expect some dependency of L on d, since in general it is likely that higher-dimensional functions require bigger training sets in order to be approximated with the same level of accuracy (see, for a discussion, (Cervellera, 2001)).

4   Numerical results

The optimization method has been tested in simulations corresponding to the evolution of the model for eight years ($T = 8$). The chosen cost index to be minimized is J_8, as it is one of the most representative and important. The kind of approximators employed is feedforward one-hidden layered neural networks with sigmoidal activation function, i.e.,

$$\hat{J}_t(\underline{\tilde{x}}_t, w_t) = \sum_{i=1}^{\nu} c_{it} \sigma(\underline{\tilde{x}}^\top \alpha_{it} + \beta_{it})$$

where σ is the hyperbolic tangent.

The number of neural units ν used in the tests is 20 for each stage t, and the number of points used for the discretization of the various feasible sets is $L = 2000$. Such points were extracted randomly by using a uniform distribution.

The growth of the transportation demand r has been taken constant at every time stage and equal to 2.275% per year. A random perturbation of the fuel price c_f has been considered, by assuming that the price can change every new year by a value in a range of $\pm 50\%$ of the price of the first year. Such perturbation are modeled as random variables with uniform distribution. In order to approximate the expected value in the DP equations, such random variables have been discretized in 10 equispaced values, on which the cost-to-go is averaged and then the minimum computed.

In order to test the goodness of the "on-line" closed-loop solutions, three different sequences of disturbances have been chosen, starting from the same initial point $\underline{\tilde{x}}_0$, and the cost given by the DP method has been compared to the cost corresponding to the real policies applied in 1992, here kept constant for the 8 years. $\underline{\tilde{x}}_0$ has been chosen according to the available data on the German economy for the year 1992.

Four different "on-line" disturbance sequences have been used, in order to show the advantages of the closed-loop control system. Such sequences are displayed in Table 6. For such random inputs, Fig.s 3-5 show the comparison between the closed-loop neural policies and the 1992 "real," fixed policies.

Note that the neural policy produces a lower cost than the 1992 fixed policy in all the simulation runs. For a better understanding of the results, Table 6 shows a comparison of the various final cost values at the end of the 8-th year.

The comparison of Fig.s 3 and 4 suggests that the oscillating behavior of the fuel price may be tackled by pursuing the expansion of the waterway TS that is obtained by a higher cost coverage and an increase of the network growth (see the plots of u_3 and u_6, respectively). Note that the carbon tax rate u_4 remains constant and equal to 0.05 (i.e., 5.0 % increment of taxation per year), which appears in accordance with the conclusions of (Piattelli *et al.*, 2002). Of course, the results change with the selection of a different cost function.

$\underline{\tilde{\xi}}^{(1)}$	col(0,0,0,0,0,0,0,0)
$\underline{\tilde{\xi}}^{(2)}$	col(.5,-.5,.5,-.5,.5,-5,.5,-.5)
$\underline{\tilde{\xi}}^{(3)}$	col(0.45,-0.27,0.11,-0.01,0.39,0.26,-0.04,-0.48)

Table 5: "On-line" random sequences.

Random Sequence	1992 Policy	Closed-loop Policy	Difference (saving)
$\underline{\tilde{\xi}}^{(1)}$	$6.6850 \cdot 10^{11}$ DM	$6.6634 \cdot 10^{11}$ DM	$2.16 \cdot 10^{9}$ DM
$\underline{\tilde{\xi}}^{(2)}$	$6.4462 \cdot 10^{11}$ DM	$6.4258 \cdot 10^{11}$ DM	$2.04 \cdot 10^{9}$ DM
$\underline{\tilde{\xi}}^{(3)}$	$6.4551 \cdot 10^{11}$ DM	$6.4314 \cdot 10^{11}$ DM	$2.37 \cdot 10^{9}$ DM

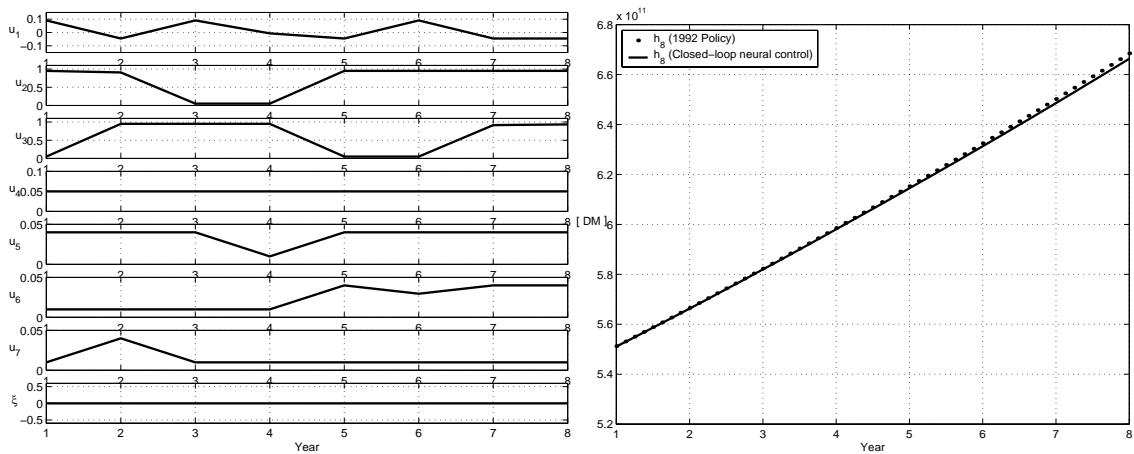Table 6: Comparison of the final costs (i.e., for the 8-th year).

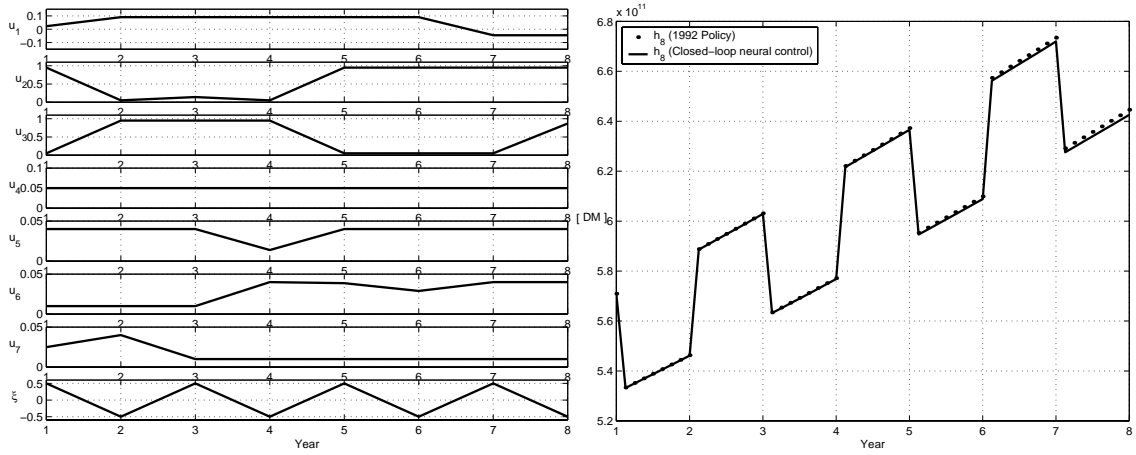Figure 3: Neural control actions, disturbance, (on the left) and cost function h_8 (on the right) for $\underline{\tilde{\xi}}^{(1)}$.

Figure 4: Neural control actions, disturbance, (on the left) and cost function h_8 (on the right) for $\underline{\tilde{\xi}}^{(2)}$.
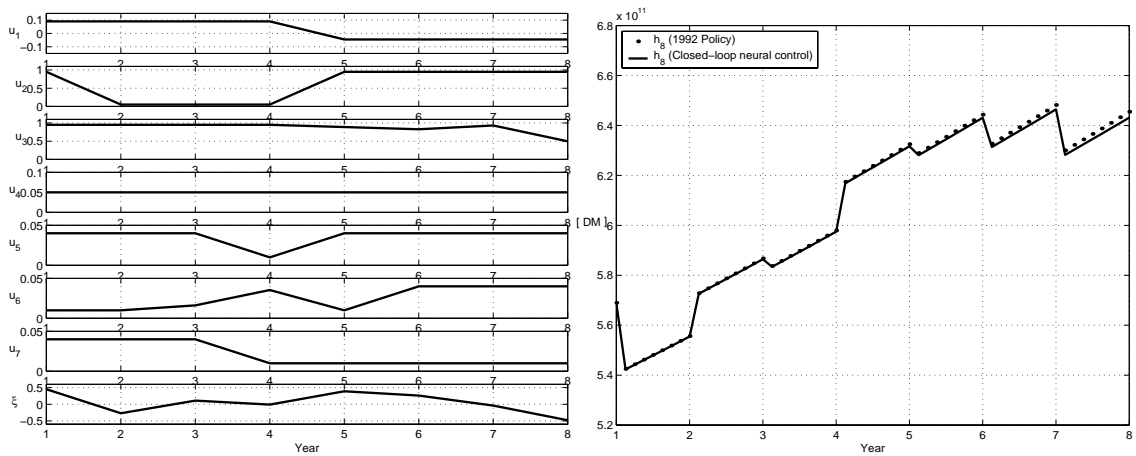

Figure 5: Neural control actions, disturbance, (on the left) and cost function h_8 (on the right) for $\underline{\tilde{\xi}}^{(3)}$.

5 Conclusions

An approach to the design of optimal feedback control laws has been presented. This method relies on the combination of both dynamic programming and neural networks. Neural networks are used as approximators of the solution of the DP problem associated with a general discrete-time nonlinear system.

The generality of such approach has allowed us to address the problem of finding closed-loop control policies that minimize a given cost function for a complex model of a nation-wide transportation system, with railway, waterway, and roadway. It is worth noting that the model is composed of 30 state variables. Preliminary simulation results confirm the effectiveness of the method and its potential as decision support tool for policy making.

References

Archibald, T.W., K.I.M. McKinnon and L.C. Thomas (1997). An aggregate stochastic dynamic programming model of multireservoir systems. *Water Resour. Res.* **33**, 333–340.

Aronsson, T. (2001). Green taxes and uncertain timing of technological change. *Resource and Energy Economics* **23**, 21–36.

Bellman, R. (1957). *Dynamic Programming*. Princeton University Press. Princeton.

Bellman, R., R. Kalaba and B. Kotkin (1963). Polynomial approximation - a new computational technique in dynamic programming. *Math. Comp.* **17**, 155–161.

Bertsekas, D. (2000). *Dynamic Programming and Optimal Control (2nd Edition)*. Vol. I. Athena Scientific. Belmont.

Bratley, P., B.L. Fox and L.E. Schrage (1987). *A guide to simulation (2nd Edition)*. Springer. New York.

Cervellera, C. (2001). Approximating Networks for the Solution of T-Stage Stochastic Optimal Control Problems. PhD thesis. University of Genoa.

Damania, R. (2000). Financial structure and the effectiveness of pollution control in an oligopolistic industry. *Resource and Energy Economics* **22**, 21–36.

Ford, A. (1994). Simulating the controllability of feebates. *System Dynamics Review* **11**(1), 3–29.

Haykin, S. (1999). *Neural Networks: A Comprehensive Foundation (2nd Edition)*. Prentice Hall. New Jersey.

Islam, S. M.N. and B. D. Craven (2001). Computation of non-linear continuous optimal growth models: experiments with optimal control algorithms and computer program. *Economic Modelling* **18**, 551–568.

Johnson, S.A., J.R. Stedinger, C.Shoemaker, Y.Li and J.A. Tejada-Guibert (1993). Numerical solution of continuous-state dynamic programs using linear and spline interpolation. *Oper. Res.* **41**, 484–500.

Nail, R. F., S. Belanger, A. Klinger and E. Petersen (92). An analysis of the cost effectiveness of the U.S. energy policies to mitigate global warming. *System Dynamics Review* **8**(2), 111–128.

Neck, R. (1999). Dynamic games of fiscal and monetary policies for Austria. In: *Annals of Operations Research*. Vol. 88. pp. 233–249. Kluwer.

Neck, R. (2001). Modeling and control of national and international economies. In: *Proc. American Control Conference*. Arlington, VA. pp. 1053–1067.

Parry, I. W.H. (1999). Environmental taxes and quotas in the presence of distorting taxes in factor markets. *Resource and Energy Economics* **19**, 203–220.

Piattelli, M. L., M. A. Cuneo, N. P. Bianchi and G. Soncin (2002). The control of goods transportation growth by modal share re-planning: the role of a carbon tax. *System Dynamics Review* **18**, 47–69.

Senge, P. M. (1975). Multiplicative formulations in urban dynamics. In: *Readings in Urban Dynamics*. Vol. 2. pp. 116–131. Wright-Allen Press, Cambridge.

Wirl, F. (1991). Dynamics of coommodity taxation: an example of an energy tax. *System Dynamics Review* **7**(2), 145–158.

Yakowitz, S. (1982). Dynamic programming applications in water resources. *Water Resour. Res.* **18**, 673–696.