DYNAMIC PROGRAMMING AND OPTIMAL CONTROL FOR VECTOR-VALUED FUNCTIONS: A STATE-OF-THE-ART REVIEW

FOUAD BEN ABDELAZIZ[1], DAVIDE LA TORRE[2,*] AND HOUDA ALAYA[3,4]

Abstract. This paper aims to present a state-of-the-art review of recent development within the areas of dynamic programming and optimal control for vector-valued functions.

1. INTRODUCTION: DYNAMIC PROGRAMMING AND OPTIMAL CONTROL

It is well known that optimization is a key tool in mathematical modeling of real phenomena. Dynamic Programming (DP) and Optimal Control (OC) are concerned with optimization over time in which the evolution of key variables is described by means of deterministic or stochastic dynamical systems, in both discrete or continuous time.

The key idea in DP is that optimization over time can be analyzed through different stages using the optimality principle that states that "from any point on an optimal trajectory the remaining trajectory is optimal for the corresponding problem initiated at that point". The DP approach was introduced by Bellman [7]. The Bellman's principle of optimality breaks a dynamic optimization problem into smaller sub-problems and the Bellman's equation defines the value of a decision problem at a certain point in terms of the payoff from some initial choices and the value of the remaining decision problem that results from those initial choices.

In literature the DP approach has been used to solve OC models. In OC problems one has to determine a control variable for a system that maximizes or minimizes a given objective function. The classical formulation, introduced by Pontryagin *et al.* [38], includes a functional cost that is a function of state and control variables. These variables are linked through a system of differential equations, and the optimal solution can be derived by using the Pontryagin's maximum principle or by solving the Bellman's equation.

When the control variables can be expressed in terms of the state variables and then they can be dropped from the problem formulation, an OC model boils down to a Calculus of Variations (CV) model. CV is a branch

[1] NEOMA Business School, Rouen Campus, Mont Saint-Aignan, France.
[2] SKEMA Business School – Université Côte d'Azur, Sophia Antipolis Campus, Sophia Antipolis, France.
[3] Logistics and Innovation Technology Research Center, Institut des Hautes Etudes de Paris, Paris, France.
[4] Tunis Business School, El Mourouj, Tunisia.
*Corresponding author: `davide.latorre@skema.edu`

of mathematical analysis that uses variations and a CV model consists in finding the maxima or the minima of functionals defined on a set of functions and taking values in the set of real numbers [19].

Recently, various researchers have solved DP and OC problems with multiple objectives and/or stochastic parameters. Diverse applications to different domains such as biology, economics, ecology, finance and management have been proposed (see, for instance [4, 10, 13, 16, 23, 24, 33, 35, 39, 42]).

This paper represents an attempt to review the use of Multiple Objective Programming (MOP) in Dynamic Programming and Optimal Control. It is organized as follows: Section 2 introduces the key concepts in MOP in both deterministic and stochastic contexts. Section 3 presents the case of DP with a single objective. In this section we define the DP process then we propose the equivalent mathematical formulation. Section 4 discusses the DP problem with stochastic parameters. Section 5 addresses the DP problem with multiple objectives in discrete and continuous time. Section 6 provides some examples of application that contain multiple objectives solved through DP. In Section 7 we present multiple objective DP problems with random parameters. Finally, in Sections 8 and 9 both the multiple objective calculus of variations and the multiple objective optimal control models are discussed.

2. Multiple objective programming and multiple objective stochastic programming

The MOP problem is usually proposed in the following form:

$$
\begin{aligned}
&\min F(x) := (f_1(x), f_2(x), \ldots, f_k(x)) \\
&\text{s.t. } x \in D
\end{aligned}
\tag{2.1}
$$

where the index k ($k \geq 2$) indicates the number of objective functions to optimize $f_i : R^n \to R$, $i = 1 \ldots k$, $F(x)$ defines the objective vector, and $x = (x_1, x_2, \ldots, x_n)^T$ is the decision variable vector. D represents the set of the feasible solutions. The MOP solution corresponds to the (Pareto) set of all non-dominated solutions:

- $x' \in S$ is a decision vector dominated by another $x \in S$ if $f_i(x) \geq f_i(x')$ for all $i = 1, \ldots, k$ and $f_j(x) > f_j(x')$ for at least one index j.
- $x' \in S$ is called efficient or Pareto optimal if there does not exist another $x \in S$ where $f_i(x) \geq f_i(x')$ for all $i = 1, \ldots, k$ and $f_j(x) > f_j(x')$ for at least one index j.

Several approaches are proposed in the literature to solve a MOP. Here we cite the most important ones: ϵ-constraint, weighted sum and goal programming.

The ϵ-constraint method was introduced, among others, by Haimes *et al.* [26] to transform a MOP into a single objective problem by minimizing a primary objective function and transforming the remaining objectives into inequality constraints:

$$
\begin{aligned}
&\max f_l(x) \\
&\text{s.t. } f_j(x) \geq \epsilon_j \quad \forall j = 1, \ldots, k; j \neq l \\
&\quad\quad x \in D
\end{aligned}
\tag{2.2}
$$

where ϵ_j is the maximum value of $f_j(x), j \neq l, l \in 1, \ldots, k$.

The weighted sum method [22] consists in assigning a positive weight to each objective and then maximizing the weighted sum of all objective functions:

$$
\begin{aligned}
&\max \sum_{i=1}^{k} \lambda_i f_i(x) \\
&\text{s.t. } x \in D
\end{aligned}
\tag{2.3}
$$

where $\lambda_i > 0$ for all $i = 1, \ldots, k$ are the weight coefficient of the ith objective and $\sum_{i=1}^{k} \lambda_i = 1$.

Goal programming (GP) was introduced by Charnes and Cooper [15]. In this method, the decision maker identifies the goal of each objective and then tries to minimize the deviations from them. The model reads as:

$$
\begin{aligned}
&\min \sum_{i=1}^{k} \gamma_i^- + \gamma_i^+ \\
&\text{s.t.} \quad f_i(x) - \gamma_i^- + \gamma_i^+ = \bar{z}_i \qquad \forall i = 1, \ldots, k \\
&\qquad \gamma_i^-, \gamma_i^+ \geq 0 \qquad \forall i = 1, \ldots, k \\
&\qquad x \in D
\end{aligned}
\tag{2.4}
$$

where γ_i^-, γ_i^+ are the negative and positive deviations of the ith objective function from the corresponding goal $\bar{z}_i$.

In specific cases, the MOP problems are including stochastic parameters. In general, the Multiple Objective Stochastic Programming (MOSP) problem can be written as follows:

$$
\begin{aligned}
&\max f(\omega, x) = (f_1(\omega, x), f_2(\omega, x), \ldots, f_k(\omega, x)) \\
&\text{s.t.} \quad x \in X(\omega) = \{g_j(\omega, x) \leq b_j(\omega), j = 1, \ldots, m\} \\
&\qquad x \in D
\end{aligned}
\tag{2.5}
$$

$\forall \omega \in \Omega$, where D is the set of deterministic constraints and $X(\omega)$ is the set of random constraints. The functions $f_l(\omega, .)$, $g_j(\omega, .)$, and the right hand side of the random constraints b_j depend on the realization of a random observation ω in a probability space (Ω, Ξ, ρ), where Ω is the sample space, Ξ is a σ-algebra of events of Ω, and P is the probability measure defined on Ξ.

When the objective functions and the constraints are linear, the MOSP boils down to:

$$
\begin{aligned}
&\max C(\omega)x \\
&\text{s.t.} \quad A(\omega)x \leq b(\omega) \\
&\qquad x \in D
\end{aligned}
\tag{2.6}
$$

$\forall \omega \in \Omega$, where C is a $(m \times n)$ random matrix, A is a $(k \times n)$ random matrix and b is a $(k \times 1)$ random vector.

To solve the MOSP problem, in [9] the authors proposed two transformations, the stochastic transformation and the multiple objective transformation.

– Stochastic transformation consists in eliminating the randomness of the problem in order to generate an equivalent deterministic problem. So, we deal first with the multiple objective dimension of the MOSP with a chance constrained compromise programming approach [1,8] or a recourse approach [43]. Then in a second stage, we solve the obtained single objective stochastic program with the suitable multiple objective method as weighted sum or a goal programming methods.
– Multiple objective transformation aims to transform the multiple objective functions into a single objective function and then use a stochastic programming method to come with a solution to the problem.

Before adopting one of the previous transformations, we need to eliminate the randomness from the constraints. For this, several transformations are proposed in the literature. The main way to transform random constraints is to consider them as extra stochastic objectives [8,41] or to transform the stochastic constraints with the chance constraint technique [8].

3. Single objective dynamic programming

Dynamic Programming (DP) usually refers to simplifying a decision by breaking it down into a sequence of decision steps over time. According to Kall and Wallace [29] the most important concepts in this method are time horizon (*i.e.* number of stages), state variables, decision variables, transition functions, return function, accumulated return functions and optimal accumulated returns (*i.e.* best return achieved from the present state

until the end of time horizon). DP proceeds as follows:

(1) Take one stage at a time. Start with the last stage.
(2) For each stage calculate the optimal accumulated return and find the optimal decision for all possible states.
(3) Move one step and calculate the returns (immediate returns and returns for all later periods) from that stage until the end of the time horizon.

DP can be formulated as follows [29]:

$$
\begin{aligned}
\max \ & F(r_0(z_0, x_0), \ldots, r_t(z_T, x_T), Q(z_{T+1})) \\
\text{s.t.} \ & z_{t+1} = G_t(z_t, x_t) && t = 0, \ldots, T \\
& A_t(z_t) \leq x_t \leq B_t(z_t) && t = 0, \ldots, T
\end{aligned} \tag{3.1}
$$

where t is the stage, $t = 1, \ldots, T$, T is the time horizon, z_t and x_t are, respectively, the state variable and the decision variable in the stage t. $Q(z_{t+1})$ is the return function for the last stage t. $G_t(z_t, x_t)$ is the transition of the system from the state z_t and the decision taken in the stage t into the state z_{t+1} in the next stage (*i.e.* $z_{t+1} = G_t(z_t, x_t)$). $r_t(z_t, x_t)$ is the immediate return if the system is in the state z_t and the decision x_t is taken. F is the overall objective and x_t is the set of feasible decisions in the stage t. To obtain the optimal accumulated return $f_t^*(z_t)$, the following program must be solved recursively.

$$
\text{Find} \quad f_0^*(z_0) \tag{3.2}
$$

and, recursively, calculate

$$
\begin{aligned}
f_t^*(z_t) &= \max_{A_t(z_t) \leq x_t \leq B_t(z_t)} f_t(z_t, x_t) \\
&= \max_{A_t(z_t) \leq x_t \leq B_t(z_t)} \varphi_t(r_t(z_t, x_t), f_{t+1}^*(z_{t+1})) \quad t = T, \ldots, 0 \\
\text{with} \quad & z_{t+1} = G_t(z_t, x_t) \qquad t = T, \ldots, 0 \\
& f_{T+1}^*(z_{T+1}) = Q(z_{T+1})
\end{aligned} \tag{3.3}
$$

where $f_t(z_t, x_t)$ is the accumulate return function.

4. Single objective stochastic dynamic programming

Stochastic DP differs from the deterministic one in the state transition equation. When events in the future are uncertain, the state does not evaluate deterministically. The procedure of the stochastic DP approach is as follows [29]:

$$
\text{Find} \quad f_0^*(z_0) \tag{4.1}
$$

recursively, calculate

$$
\begin{aligned}
f_t^*(z_t) &= \min_{A_t(z_t) \leq x_t \leq B_t(z_t)} f_t(z_t, x_t) \\
&= \min_{A_t(z_t) \leq x_t \leq B_t(z_t)} E_{\widetilde{\xi}_t} \{\varphi_t(r_t(z_t, x_t, \widetilde{\xi}_t), f_{t+1}^*(z_{t+1}))\} \quad t = 0, \ldots, T \\
\text{with} \quad & \\
z_{t+1} &= G_t(z_t, x_t, \xi_t) \quad t = 0, \ldots, T \\
f_{T+1}^*(z_{T+1}) &= Q(z_{T+1})
\end{aligned} \tag{4.2}
$$

where $\widetilde{\xi}_t$ is the random vector for stage t, $r_t(z_t, x_t, \widetilde{\xi}_t)$ is the return function and $z_{t+1} = G_t(z_t, x_t, \xi_t)$ is the transition function.

To solve a stochastic DP, we can use a backward recursion or a forward recursion algorithm.

– Backward recursion: the aim of this algorithm is to provide a sequence of alternating states and decision from an initial state to a final state.
– Forward recursion: it is an opposite procedure to the backward recursion. In this method, we start with a final state and work backwards. This algorithm can be used to solve different problems and specially for the problems in which the final state is unknowns. In this case it is the only feasible approach.

5. MULTIPLE OBJECTIVE DYNAMIC PROGRAMMING (MODP)

Recently, the DP is developed to solve problems with multiple objective functions. In general, two properties are considered in the objective function witch are, separability and monotonicity properties [7]. Those properties can ease the solution of MODP.

– Separability:
$$
\begin{aligned}
&F(x_0, \ldots, x_N; u_1, \ldots, u_N) \\
&\quad = Q_N(x_N, u_N; q_{N-1}(x_0, \ldots, x_{N-1}; u_1, \ldots, u_{N-1})) \\
&q_n(x_0, \ldots, x_n; u_1, \ldots, u_n) \\
&\quad = Q_n(x_n, u_n; q_{n-1}(x_0, \ldots, x_{n-1}; u_1, \ldots, u_{n-1})) \\
&(n = 1, \ldots, N - 1), q_0 = q_0(x_0)
\end{aligned}
\tag{5.1}
$$

where F, Q_j, q_j are real scalar-valued functions. $x = x_1, \ldots, x_N$ is the set of state variables and $u = u_1, \ldots, u_N$ is the set of control variables.
– Monotonicity:
$$
\begin{aligned}
&Q_n(x_n, u_n; w) \leq Q_n(x_n, u_n; w') \\
&\text{for every pair } w, \ w' \in E^1 \text{ with } w \leq w'.
\end{aligned}
\tag{5.2}
$$

MODP models can be generally divided into discrete or continuous.

5.1. MODP in discrete time

In this section we will present different MODP models in discrete time. In [2,3] the authors proposed to solve the following formulation:

$$
\begin{aligned}
&\min F_k(f_{k1}(x_1), \ldots, f_{kN}(x_N)) \quad k = 1, 2, \ldots, K, \quad K \geq 2 \\
&\text{s.t. } G_m(g_{m1}(x_1), \ldots, g_{mN}(x_N)) \leq 0 \qquad m = 1, \ldots, M \\
&\qquad x_n \in X_n, \qquad\qquad\qquad\qquad\qquad\qquad n = 1, \ldots, N
\end{aligned}
\tag{5.3}
$$

where X_n, $n = 1, \ldots, N$ is a subset of R^{t_n}, and x_n is a t_n dimensional vector. The objective functions F_k, $k = 1, \ldots, K$, the constraints $G_m, m = 1, \ldots, M$, are functions of class C^1 on R^N taking values in R and $f_{kn}, g_{mn}, k = 1, \ldots, K, m = 1, \ldots, M, n = 1, \ldots, N$ are real valued functions defined on X_n. To determine the efficient set, the authors proposed the ϵ-constraint approach for dealing with non-convex models. Their approach consists in retaining one objective function as primary and transform the other $k - 1$ objective functions as constraints. Therefore, the proposed model becomes equivalent to the following $P(\epsilon)$ formulation:

$$
\begin{aligned}
&\min F_1(f_{11}(x_1), \ldots, f_{1N}(x_N)) \\
&\text{s.t. } F_k(f_{k1}(x_1), \ldots, f_{kN}(x_N)) \leq \epsilon_k, \quad k = 2, \ldots, K \\
&\qquad G_m(g_{m1}(x_1), \ldots, g_{mN}(x_N)) \leq 0, \quad m = 1, \ldots, M \\
&\qquad x_n \in X_n, \qquad\qquad\qquad\qquad\qquad n = 1, \ldots, N
\end{aligned}
\tag{5.4}
$$

where $\epsilon = (\epsilon_2, \ldots, \epsilon_k)$ is a $k - 1$ dimensional vector. We note that, for each fixed ϵ, at least one optimal solution of the problem $P(\epsilon)$ is an efficient solution of the defined problem. To generate the entire set of efficient solutions, in [3] the authors decomposed the parametric space by using the constrained approach. Moreover, if the proposed formulation is separable and the separation is monotone, the problem $P(\epsilon)$ can be decomposed into sub-problems by DP in the sense of [40]. In this case, real-valued functions $B_1^n(\epsilon, s)$ are defined for each $n = 1, \ldots, N, \epsilon = (\epsilon_2, \ldots, \epsilon_k), s = (s_1, \ldots, s_M)$ as follows:

$$
\begin{aligned}
&B_1^n(\epsilon, s) = \min \ F_1^n(f_{11}(x_1), \ldots, f_{1n}(x_n)) \\
&\text{s.t.} \\
&F_k^n(f_{k1}(x_1), \ldots, f_{kn}(x_n)) \leq \epsilon_k, \qquad k = 2, \ldots, K \\
&G_m^n(g_{m1}(x_1), \ldots, g_{mn}(x_n)) \leq s_m; \qquad m = 1, \ldots, M \\
&x_1 \in X_1, \ldots, x_n \in X_n.
\end{aligned}
\tag{5.5}
$$

The recursive relations:

$$
\begin{aligned}
&B_1^n(\epsilon_2^n,\dots,\epsilon_k^n,s_1^n,\dots,s_M^n) = \\
&\min_{x_n \in X_n} \Phi_1^n(f_{1n}(x_n), B_1^{n-1}[\epsilon_k^{n-1}(\epsilon_k^n, f_{kn}(x_n)), s_m^{n-1}(s_m^n, g_{mn}(x_n))]) \\
&\text{s.t.} \\
&f_{kn}(x_n) \le \epsilon_k^n, \qquad k = 2,\dots,K \\
&g_{mn}(x_n) \le s_m^n, \quad m = 1,\dots,M
\end{aligned}
\tag{5.6}
$$

for $n = 1,\dots,N$ with $B_k^o = 0, k = 1,\dots,K$ and assuming the monotonicity of ψ_m^n, let s_m^{n-1} be defined by:

$$
s_m^{n-1}(s_m^n, g_{mn}(x_n)) = \sup\{\xi \in R | \psi_m^n(\xi, g_{mn}(x_n)) \le s_m^n\}, \quad m = 1,\dots,M.
$$

The authors also proposed an algorithm for decomposing the parametric space into multiple objective nonlinear programming problems which are solved using a DP approach.

In [30] a multiple objective optimal stopping problem is solved. The authors proposed to stop the selection process of the Bilateral Optimal Selection Problem (BOSP) if either DM decides to stop. A utility function for each $\mathrm{DM}_p(p = 1,2)$ is considered. This function takes into account the position of the opponent for the selected offer. A game payoff of each DM is defined. The authors presented a dynamic formulation of the BOSP where each DM's utility is expressed as in the above payoff matrix. The optimality equation for DM_p, $p = 1,2$, at stage i, $i = 1,\dots,n-1$, is the following:

$$
Ev^p(i, k_p, k_{p'}) = \begin{cases} \max\{Ev_s^p(i, 1, k_{p'}), Ev_c^p(i)\} & \text{if } k_p = 1 \text{ or } k_{p'} = 1 \\ Ev_c^p(i) & \text{elsewhere} \end{cases}
\tag{5.7}
$$

where $Ev^p(i, 1, k_{p'})$ is the maximum expected utility of DM_p by following the optimal strategy when the ith offer, with relative ranks k_p and $k_{p'}$ is examined, $Ev_c^p(i)$ is the expected utility of DM_p when postponing the decision to the next stage and $Ev_s^p(i, 1, k_{p'})$ is the expected utility of DM_p when stopping at stage i and selecting the current offer with relative ranks k_p and $k_{p'}$. In [30] the authors developed the game formulation of the problem through the backward recursive equations of each DM. Then, they studied the problem for a specific case of the proposed utility and they showed how each DM should foresee his opponent's individual decision. The numerical investigation of the problem when varying the parameters of the proposed utility showed the rational decisions that can be taken by each DM and their expected rewards before beginning the selection process. Compared with the case of a single DM, the experimental results illustrate the great loss of expected utility when considering two DMs.

In [44] the authors considered the following multiple objective optimal control problems subjected to nonlinear control dynamics with unknown disturbance:

$$
(P) \quad \min_U f(X,U) = (f^1(X,U), f^2(X,U),\dots,f^M(X,U))^T \\
\text{s.t.} \ x_{t+1} = f(x_t, u_t, \xi_t)
\tag{5.8}
$$

where $U = (u_0^T, u_1^T,\dots)^T$ denotes the control set, $X = (x_0^T, x_1^T,\dots)^T$ denotes the state set, the state $x_t \in R^n$, the control $u_t \in R^m$ and the bounded disturbance $\xi_t \in R^p, t = 0,1,\dots$. In the problem (P), there are M performance indices with all the elements in the performance index vector $f^i(X,U) \ge 0$, expressed by the following form:

$$
f^1(X,U) = \sum_{t=0}^{\infty} y_t^i, \quad i = 1,\dots,M
\tag{5.9}
$$

where $y_t^i = \Phi^i(x_t, u_t) \ge 0, i = 2,\dots,M$ is the utility function. To solve the multiple objective optimal control problem (P), the aim is to find an optimal control sequence $U^* = ((u_0^*)^T, (u_1^*)^T,\dots)^T$ so as to make the vector

valued performance index minimized which makes the system achieve optimality. The optimal state trajectory is formulated as $X^* = ((x_0^*)^T, (x_1^*)^T, \ldots)^T$. Let $U^* \in EZ$, where

$$EZ = \{U | U \text{ is an optimal solution to } (P)\}$$

and let X^* be the corresponding state trajectory for (P). Then for any t, the control sequence $(u_{t+1}^*, u_{t+2}^*, \ldots)$ of U^* constitutes an optimal solution of the following problem:

$$(P_1(t)) \quad \min_{u_t} \; y_t + f(X_{t+1}^*, U_{t+1}^*)$$
$$\text{s.t.} \; x_{t+1} = f(x_t, u_t, \xi_t) \tag{5.10}$$

where $y_t = (y_t^1, y_t^2, \ldots, y_t^m)^T$, and $X_{t+1} = (x_{t+1}^T, x_{t+2}^T, \ldots)^T$ and $U_{t+1} = (u_{t+1}^T, u_{t+2}^T, \ldots)^T, t = 0, 1, \ldots \infty$.
   The problem $(P_1(t))$ can be stated as the following minimum problem:

$$\min_U f(X, U) = \left(\Phi^i(x_j, u_j)_{i,j}\right) = \left(y_j^i\right)_{i,j} = (y_j)_j. \tag{5.11}$$

In [44] the 2-norm of the vector-valued index is minimized. According to the assumption $y_t \geq 0$, the above model can be transformed into the following problem:

$$(P_2) \quad \min_U \; V = (y_0)^T y_0 + (y_1)^T y_1 + \ldots$$
$$\text{s.t.} \; x_{t+1} = f(x_t, u_t, \xi_t), t = 0, 1, \ldots. \tag{5.12}$$

Define the utility function as:

$$(x_t, u_t) = (y_t^1)^T y_t^1 + \ldots + (y_t^M)^T y_t^M = \sum_{i=1}^{M} l^i(x_t, u_t) = \sum_{i=1}^{M} (y_t^i)^T y_t^i.$$

So the performance index function can be expressed as:

$$\sum_U V = \sum_{t=0}^{\infty} l(x_t, u_t). \tag{5.13}$$

Note that $x_t \in R^m, u_t \in R^n, \xi_t \in R^p$.

   In [31] the basic problem is defined as follows:

$$\max \; F(x, u) \tag{5.14}$$

subject to state variables $x = (x_0, \ldots, x_N)$ and control variables $u = (u_1, \ldots, u_N)$ under restrictions of the type:

$$x_{n-1} = g_n(x_n, u_n), u_n \in U_n(x_n), \quad n = 1, \ldots, N \tag{5.15}$$

and the initial condition $x_n = \alpha$. Each feasible pair $\langle x, u \rangle$ of the problem is said to be a process. The $F(x, u)$ fulfills the properties of separability and monotonicity. Then, the problem can be decomposed into the following multistage system of recurrence relations:

$$f_n(x_n) = \max_{u_n \in U_n(x_n)} Q_n(x_n, u_n; f_{n-1}(g_n(x_n, u_n))), \quad n = 2, \ldots, N$$
$$f_1(x_1) = \max_{u_1 \in U_1(x_1)} Q_1(x_1, u_1; q_0(g_1(x_1, u_1))) \tag{5.16}$$

where $f_n(x_n)$ is the Bellman function and the optimal return of the corresponding $n-$stage process by the initial value x_n. Another problem is then proposed:

$$\max F(u, x)$$
$$\text{s.t.} \quad x = (x_0, \ldots, x_N) \text{ and } u = (u_1, \ldots, u_n) \text{ satisfying}$$
$$x_{n-1} = g_n(x_n, u_n), u_n \in U_n(x_n), n = 1, \ldots, N, x_n = \alpha. \tag{5.17}$$

The range W of F is a subset of E^L. This formulation aims to find the whole set W^* of maximal (efficient) values of W in the sense of Pareto optimality. In respect of all admissible n-stage process by the initial value x_n. Instead of the properties of separability and monotony, in [31] assumed the following modified property.

Modified Separability:

$$
\begin{aligned}
F(x_0, \ldots, x_N; u_1, \ldots, u_N) &= Q_N(x_N, u_N; q_{N-1}(x_0, \ldots, x_{N-1}; u_1, \ldots, u_{N-1})) \\
q_n(x_0, \ldots, x_n; u_1, \ldots, u_n) &= Q_n(x_n, u_n; q_{n-1}(x_0, \ldots, x_{n-1}; u_1, \ldots, u_{n-1})) \\
&\quad n = 1, \ldots, N-1, q_0 = q_0(x_0)
\end{aligned}
\tag{5.18}
$$

by real vector values functions F, Q_j, q_j as mappings in E^L.

Modified Monotonicity:

$$
Q_n(x_n, u_n; w) \leq Q_n(x_n, u_n; w') \text{ for every pair } w, \ w' \in E^L \text{ with } w \leq w'.
\tag{5.19}
$$

The system of recurrence set relations is:

$$
\begin{aligned}
f_n(x_n) &= \max_{u_n \in U_n(x_n)} Q_n(x_n, u_n; f_{n-1}(g_n(x_n, u_n))) \quad n = 2, \ldots, N \\
f_1(x_1) &= \max_{u_1 \in U_1(x_1)} Q_1(x_1, u_1; q_0(g_1(x_1, u_1))).
\end{aligned}
\tag{5.20}
$$

5.2. MODP in continuous time

In [14] the authors proposed a schema by which a typical multiple objective capacity expansion problem can be solved. In the capacity expansion problem, N projects available for construction in T time periods and two objectives are considered. The problem formulation is considered of the following form:

$$
\begin{aligned}
&\min \ (f_1(x,y), f_2(x,y)) \\
&\text{s.t. } \ x \in X, y \in Y
\end{aligned}
\tag{5.21}
$$

where $x = (x_{11}, \ldots, x_{it}, \ldots, x_{NT})$, x_{it} is equal to 1 if project i is constructed in time period t and 0 otherwise. $y = (y_{11}, \ldots, y_{it}, \ldots, y_{NT})$, y_{it} is the utilization of project i in period t. X is a feasible set of x and Y is a feasible set of y. In [14] the authors proposed to convert, first, the problem into an ϵ-constraint problem by taking the first objective as the primary objective function.

$$
\begin{aligned}
P(\epsilon) : &\min \ f_1(x,y) \\
&\text{s.t. } \ f_2(x,y) \leq \epsilon \\
&\quad\quad x \in X, y \in Y.
\end{aligned}
\tag{5.22}
$$

Then, two steps are considered:

– solve the above problem for various values of ϵ,
– obtain trade off and all relevant information from the first step, then interact with decision maker to choose the final solution.

The authors suggested to carry out the first step. To do this, the objective that is turned as a constraint is expressible as some function of similar indicators in each time period t.

$$
f_2(x,y) = \hat{f}_2(f_2^1, \ldots, f_2^t, \ldots, f_2^T)
\tag{5.23}
$$

where f_2^t is the corresponding index of f_2 in period t. Another assumption is then suggested:

$$
f_2(x,y) = \hat{f}_2(f_2^1, \ldots, f_2^t, \ldots, f_2^T) \leq \epsilon.
\tag{5.24}
$$

Chankong *et al.* [14] assumed that they can find $F^1(\epsilon)$, $F^2(\epsilon)$, ..., $F^T(\epsilon)$ such that (35) is equivalent to the following set of inequalities:

$$f_2^t(x,y) \leq F^t(\epsilon), \quad t = 1, \ldots, T \tag{5.25}$$

in the sens that if $X_1 = \{(x,y)|f_2(x,y) \leq \epsilon\}$ and $X_2 = \{(x,y)|f_2^t(x,y) \leq F^T(\epsilon) \ t = 1, \ldots, T\}$ then $X_1 = X2$. Thus, $P(\epsilon)$ can be transformed into:

$$P_1(\epsilon) : \min_{} f_1(x,y)$$
$$\text{s.t.} \quad f_2^t(x,y) \leq F^t(\epsilon), \quad t = 1, \ldots, T \tag{5.26}$$
$$x \in X, y \in Y.$$

An equivalent DP problem is defined as follows:

$$f_n^{(k_n(t))} = \min_{i,t_n}(C_{it_n} + f_{n-1}^{k_{n-1}(t_n)}(t_n) + h^{k_{n-1}(t_n)\theta i}(t_n,t)) \tag{5.27}$$

where C_{it} is the capital and annual cost of constructing project i in time period t. k_n is the set of n or fewer constructed projects. $h^{k_{n-1}(t_n)}$ is equal to the minimum present value of the variable costs satisfying the demand.

To find $h(.)$, the following formulation should be solved:

$$P_2(\epsilon) : h^{k_{n-1}(t_n)\theta_t}(t_n,t) = \min_{y_{i\tau}} \sum \tau = t_n{}^t \sum_{j \in k_{n-1}(t_n)^{\theta i}} b_{j\tau}y_{j\tau}$$
$$\text{s.t.} \quad \sum_{j \in k_{n-1}(t_n)^{\theta i}} y_{j\tau} \geq q(t)$$
$$0 \leq y_{j\tau} \leq Q_j \tag{5.28}$$
$$f_2^\tau(y; k_{n-1}(t_n)\theta i) \leq F^\tau(\epsilon)$$
$$j = 1, \ldots, N, \quad \tau = t_n, \ldots, t.$$

The authors assumed that the function $F^t(\epsilon)$ is a continuously differentiable function of ϵ for each $t = 1, \ldots, T$. For some given ϵ^0, let y^0 solve the following problem:

$$P_3(\epsilon) : \min_{y \in Y} f_1(x^0, y)$$
$$\text{s.t.} \quad f_2^t(x^0, y) \leq F^t(\epsilon^0) \quad t = 1, \ldots, T \tag{5.29}$$

with the following properties:

– y^0 satisfies the regularity conditions for P_3
– second-order sufficiency condition is satisfied at y^0
– there are no degenerate constraints at y^0. Let $\lambda_{12}^t = \sum_{t=1}^{T} \frac{dF^t(\epsilon^0)}{d\epsilon}$, the optimal Kuhn-Tucker multiplier. Therefore, the trade offs between f_1 and f_q, $q = 2, \ldots, n$ with $\lambda_{1q}^t > 0$ are given by:

$$-\lambda_{1q} = -\sum_{t=1}^{T} \lambda_{1q}^t \frac{dF_q^t(\epsilon_q)}{d\epsilon_q}. \tag{5.30}$$

6. Solving MOP through dynamic programming

Some researchers used the DP to solve various practical applications of MOP. In [40] the authors illustrated a MODP model with the independent operation of Shasta Reservoir in California. Three objectives are considered: maximizing the cumulative dump energy generated above the level of firm energy, minimizing the cumulative evaporation or loss of the resource and maximizing the firm energy. The problem was formulated as follows:

$$f_j(S_j, V_j) = \max q_j E_j(q_j, S_j, V_j, \text{FE}_j) + f_{j-1}(S_{j-1}, V_{j-1})$$
$$S_{j-1} = T_j^s(q_j, S_j, V_j)$$
$$V_{j-1} = T_j^v(q_j, S_j, V_j) \tag{6.1}$$
$$q_j \geq q_{\min j}; S_{\min} \leq S_j \leq S_{\max}$$

where $f_j()$ is long-range returns (dump energy accumulated through stage j), $E_j()$ is short-range returns (stage j dump energy), FE_j is a firm energy required in time period j, S_j is the volume of storage at stage j, V_j is a cumulative evaporation through stage j and q_j is a release volume during stage j. MODP generates the entire non inferior set of solutions (convex or otherwise) as well as the explicit trade-off between objectives at all objective levels. MODP is therefore an applicable technique that can further assist the water resource planner.

Dabia *et al.* [17] proposed a time dependent DP approach to solve a multiple objective vehicle routing problem in which two objectives are defined: minimize the tour's execution time and maximize the total demand fulfilled. The authors presented two time assumptions. The first one suggests that each customer i having an opening time t_i^l can only be visited before customers with later opening times and the second one, called FIFO assumption, guarantees that, for every link, a later departure will not result in earlier arrival. Dabia *et al.* [17] formulated an approximate DP based on the non-dominance principle and on an additional assumption. This assumption means that fast increases in travel time are not allowed and it is defined as follows: for every real number $1 \le \alpha \le 2$ and every two customers i and j, it holds that for every time t, $tt_{ij}(\alpha t) \le \alpha tt_{ij}(t)$. $tt_{ij}(t)$ is the travel time from customer i to j when leaving at time t.

Jacquin *et al.* [28] used the DP to solve a scheduling problem called multiple objective unit commitment problem (UCP). Two objectives are defined, which are, minimize the production cost and minimize the gas emissions. In this work a generalization of DYNAMOP (DP based metaheuristic for optimization problems) method is proposed.

7. MULTIPLE OBJECTIVE STOCHASTIC DYNAMIC PROGRAMMING

Various applications of MODP considered random parameters. These problems can be captured within the framework of Multiple Objective Stochastic DP (MOSDP). The MOSDP is expressed in most of the cases by a discrete time modelisation. In [46] the authors proposed a Dynamic Cell Formation Problem (DCFP). This problem has a dynamic and stochastic nature. The dynamic aspect is due to varying product mix and product demand volumes. The authors developed a bi-objective stochastic model for:

– Minimizing the total cost of machine procurement, machine relocation, inter-cell moves, overtime utilization, working hiring/laying-off, and worker moves between cells.

$$
\begin{aligned}
\min Z_1 = & \sum_{t=1}^{T} \sum_{m=1}^{M} \left(\mu_m K_{mt}^+ - \pi_m K_{mt}^- \right) \\
& + \sum_{t=1}^{T} \sum_{c=1}^{C} \sum_{m=1}^{M} \left(\gamma_m^+ . N_{mct}^+ + \gamma_m^- . N_{mct}^- \right) \\
& + \frac{1}{2} \sum_{t=1}^{T} \sum_{p=1}^{P} \Phi_p \left[\frac{D_{pt}}{B_p} \right] . \sum_{c=1}^{C} \sum_{j=1}^{J_p} \left| \sum_{m=1}^{M} X_{(j+1)pmct} - \sum_{m=1}^{M} X_{jpmct} \right| \\
& + \sum_{t=1}^{T} OC_t . \sum_{m=1}^{M} \tau_{mt} \\
& + \sum_{t=1}^{T} (W_t^+ m_t + W_t^- n_t) + \sum_{c=1}^{C} \sigma_c . \sum_{t=1}^{T} W_{ct}^+
\end{aligned}
\tag{7.1}
$$

where c is an index for manufacturing cells ($c = 1, \ldots, C$), m is an index for machine types ($m = 1, \ldots, M$), p is an index for part types ($p = 1, \ldots, P$), t is an index for time periods ($t = 1, \ldots, T$), j is an index for operations belonging to part type p ($j = 1, \ldots, J_p$). C is the maximum number of allowed cells, M, P, T and J_p correspond to the numbers of machine types, part types, periods and operations required for part type p, respectively. D_{pt} is the demand for part type p in period t. Φ_p represents the inter-cell movement

cost per unit of part type p batch. B_p is the batch size of product type p. π_m is the selling cost of machine type m. μ_m is purchasing cost of machine type m. γ_m^+ is installing cost of machine type m. γ_m^- is removing cost of machine type m. m_t is hiring cost in period t. n_t is laying off cost in period t. OC_t is worker cost per unit of overtime. σ_c is cost of moving worker to cell c. X_{jpmct} takes 1, if operation j of part type p is performed on machine m in cell c in period t; 0, otherwise. N_{mct}^+ is the number of machine type m added to cell c in period t. N_{mct}^- is the number of machine type m removed from cell c in period t. K_{mt}^+ is the number of machine type m purchased at the beginning of period t. K_{mt}^- is the number of machine type m sold at the beginning of period t. W_t^+ is the number of workers hired in period t. W_t^- is the number of workers laid off in period t. W_{ct}^+ is the number of workers added to cell c in period t and τ_{mt} is overtime utilization on machine m in period t.

– Maximizing labor utilization of the cellular manufacturing system.

$$\max Z_2 = \sum_{c=1}^{C} \sum_{t=1}^{T} \mathrm{MHU}_{ct} \tag{7.2}$$

where MHU_{ct} is the minimum worker utilization of cell c in period t.

Zohrevand *et al.* [46] defined several constraints such as, time capacities of machines, limitations of overtime and cell sizes, minimum worker utilization in the planning periods, numbers of assigned workers to cells and numbers of workers in the periods. Some parameters are changing during different planning periods, including the part demands, worker hiring/ laying off costs, and worker overtime costs. In addition to the defined parameters and variables, some constraints ensure dynamic reconfiguration of part assignment, cells, and the total CMS. The authors adopted two transformations approaches; the approach of Lai and Hwang [32] and the weighted sum approach. The first approach tackled the fuzzy parameters and the second approach transformed the proposed model into single-objective model. A hybrid tabu search - genetic algorithm is developed to obtain desired solutions.

8. Multiple objective calculus of variations

In [5] the authors established weak, strong, and converse duality results for weakly efficient solutions in vector or multiple objective variational problems. These results extend and improve recent papers. The authors considered Kuhn-Tucker optimality conditions, weighting variational problems, and some classes of generalized convex functions, recently introduced, which are extended in this work. The authors considered the following model:

$$\min \int_a^b f(t, x, \dot{x}) \mathrm{d}t = \left(\int_a^b f_1(t, x, \dot{x}) \mathrm{d}t, \ldots, \int_a^b f_n(t, x, \dot{x}) \mathrm{d}t \right) \tag{8.1}$$

subject to

$$x(a) = \alpha, x(b) = \beta, \tag{8.2}$$

$$g(t, x, \dot{x}) \leq 0, \text{a.e.} \ \ t \in [a, b]. \tag{8.3}$$

In [25] the authors proved necessary optimality conditions in the form of a generalized Euler equation for an infinite dimensional dynamic maximization program posed as a vector problem of calculus of variations with fixed end points. Under concavity of the integrand with respect to the space variables they showed that these conditions are also sufficient. They also used these results to study a vector-valued extension of the Ramsey economic growth model in Banach spaces with partial order given by a closed convex cone and provided several concrete economic examples. They considered problems involving objective functionals taking the form:

$$J(x) = \int_a^b f(t, x(t), x'(t)) \, \mathrm{d}t \tag{8.4}$$

where $f : [a, b] \times X \times X \to Y$ is a continuous function, $(X, \| \cdot \|_X)$, $Y = (\| \cdot \|_Y)$ are Banach spaces, and $P \subset Y$ is a closed convex cone with nonempty interior.

9. MULTIPLE OBJECTIVE OPTIMAL CONTROL

Several approaches are proposed in the literature to solve a Multiple Objective Optimal Control Problem (MOOCP). Gambier and Jipp [20] classified these approaches as follows:

– Parameter-optimized PID controllers: this approach aims to find parameters for the digital PID controller and to define the performance index for the time discrete case [27].
– Eigenvalue assignment as a MOP problem: the basic principle of this approach is to find a feedback matrix K for an input matrix $B \in \Re^{n*l}$ and a system matrix $A \in \Re^{n*n}$ such that the eigenvalues of the closed loop system matrix $A^* = (A - BK)$ lie in the position defined by the designer [11] [36].
– Multiple objective robust control: this approach is adapted to performance specifications while provides closed -loop stability in the presence of disturbances [6].
– Multiple objective predictive control: this approach compromises a family of controllers signed by the same design philosophy and similar characteristics [12].
– Supervisory switching control: in this approach, the supervisor decides on the basis of available information like states, input and output which of the available controllers should be activated . Then, he establishes the corresponding reference signal. The final result is the switching function acquired from the switching function and the corresponding reference signal [34].
– Multi-loop control systems with multiple controllers: in this approach the controllers are assumed to be players of a non-zero sum cooperative game [21].

Logist *et al.* [37] established a computationally efficient frame for a robust MOOCP yielding robust Pareto sets for dynamic(bio) chemical process. The authors proposed two approaches: deterministic secularization based on multiple objective approaches and robust optimization techniques for dynamic systems based on Lyapunov differential equations.

Zarei and Bahrmand [45] proposed a method for the solution of a MOOCP in linear distributed parameters system governed by a wave equation. The authors defined a solution for the wave equation and then converted the MOOCP into a MOP by the Laplace transform. Finally, they transformed the obtained MOP with the Zimmerman's fuzzy approach into a single objective problem. The resulting problem is solved with a metaheuristic algorithm.

In [18] the authors presented optimality conditions for the sufficiency of the maximum principle for multiple objective optimal control problems with nonsmooth data. It turns out that these conditions are the most general as possible in the sense that problems in which the set of necessary conditions from the maximum principle are also sufficient, necessarily obey them. The authors also presented a variation of such conditions under which the set of optimal solutions of the multiple objective problem can be determined by resolving a related scalar weighting problem. They deal with the following problem:

$$\min g(x(S), x(T)) \tag{9.1}$$

subject to

$$\dot{x}(t) = f(t, x(t), u(t)), \quad \text{a.e. } t \in [S, T] \tag{9.2}$$

$$(x(S), x(T)) \in C, \tag{9.3}$$

$$u(t) \in U(t), \quad \text{a.e. } t \in [S, T]. \tag{9.4}$$

where $g = [g_1, g_2, \ldots, g_k] : R^n \times R^n \to R^k$, $f : [S, T] \times R^n \times R^m \to R^n$, C is a closed subset of $R^n \times R^n$, U is a multifunction mapping $[S, T]$ to nonempty subsets of R^m, S and T are fixed.

10. CONCLUSION

This paper represents an attempt to present the state-of-the-art of the use of multiple objective optimization techniques within the areas of dynamic programming and optimal control. The main concepts of optimization with vector-valued functions have been introduced in classical dynamic programming and optimal control

settings, including multiple objective dynamic problems in discrete and continuous time, stochastic dynamic programming, multiple objective calculus of variations, and others. Several future avenues could be explored with the aim of extending theoretical results including set-valued functions as well as of finding innovative applications in the areas of economic growth, health management, sustainability, and other area in which the analysis of time-dependent multiple criteria decision making problems play a crucial role.

References

[1] F.B. Abdelaziz and H. Masri, A compromise solution for the multiobjective stochastic linear programming under partial uncertainty. *Eur. J. Oper. Res.* **202** (2010) 55–59.

[2] M.A. Abo-Sinna and M.L. Hussein, Decomposition of multi-objective programming problems by hybrid fuzzy dynamic programming. *Fuzzy Sets Syst.* **60** (1993) 25–32.

[3] M.A. Abo-Sinna and M.L. Hussein, An algorithm for generating efficient solutions of multi-objective dynamic programming problems. *Eur. J. Oper. Res.* **80** (1995) 156–165.

[4] S. Anita, V. Arnautu and V. Capasso, An Introduction to Optimal Control Problems in Life Sciences and Economics: From Mathematical Models to Numerical Simulation with MATLAB. Birkhauser (2011).

[5] M. Arana-Jiménez, F. Ortegón Gallego, Duality and weak efficiency in vector variational problems. *J. Optim. Theory App.* **159** (2013) 547–553.

[6] K.J. Ãström, Introduction to Stochastic Control Theory. Academic Press, Inc., London (1970).

[7] R.E. Bellman, Dynamic Programming. Republished 2003: Dover, ISBN 0-486-42809-5. Princeton University Press, Princeton, NJ (1957).

[8] F. Ben Abdelaziz, Solution approaches for the multiobjective stochastic programming. *Eur. J. Oper. Res.* **216** (2012) 1–16.

[9] F. Ben Abdelaziz, P. Lang and R. Nadeau, Dominance and efficiency in multiobjective decision under uncertainty. *Theory Decis.* **47** (1999) 191–211.

[10] L. Berkovitz and N. Medhin, Nonlinear Optimal Control Theory. Chapman & Hall, CRC Press (2012).

[11] M.A. Branch and A. Grace, MATLAB Optimization Toolbox User's Guide, Version 1.5. The Mathworks (1996).

[12] A.E. Bryson, Optimal control – 1950 to 1985. *IEEE Control Syst. Mag.* **16** (1996) 26–33.

[13] A.E. Bryson and Y.C. Ho, Applied Optimal Control. Hemispheres (1975).

[14] V. Chankong, Y.Y. Haimes and D.M. Gemperline, A multiobjective dynamic programming method for capacity expansion. *IEEE Trans. Autom. Control* **26** (1981) 1195–1207.

[15] A. Charnes and W.W. Cooper, Chance constraints and normal deviates. *J. Am. Stat. Assoc.* **57** (1952) 134–148.

[16] A. Chinchuluun, P.M. Pardalos, R. Enkhbat and I. Tseveendori, Optimization and Optimal Control: Theory and Applications. In Vol. **39**. Springer (2010).

[17] S. Dabia, T. Van Woensel and A.G. De Kok, A dynamic programming approach to multi-objective time-dependent capacitated single vehicle routing problems with time windows. BETA Publicatie: working papers 313 (2010).

[18] V.A. De Oliveira and G.N. Silva, On sufficient optimality conditions for multiobjective control problems. *J. Global Optim.* **64** (2016) 721–744.

[19] L.D. Elsgolc, Calculus of Variations. Courier Corporation (2012).

[20] A. Gambier and M. Jipp, Multi-objective optimal control: an introduction. In: *2011 8th Asian Control Conference (ASCC)*. IEEE (2011) 1084–1089.

[21] A. Gambier, A. Wellenreuther and E. Badreddin, A new approach to design multi-loop control systems with multiple controllers. In: *Proceedings of the 45th IEEE Conference on Decision and Control, San Diego* (2006).

[22] S. Gass and T. Saaty, The computational algorithm for the parametric objective function. *Naval Res. Logistics Q.* **2** (1955) 39–45.

[23] H.P. Geering, Optimal Control with Engineering Applications. Springer (2007).

[24] I.M. Gelfand and S.V. Fomin, Calculus of Variations. Dover Publications Inc., New York (2000).

[25] I. Ginchev, D. La Torre and M. Rocca, Optimality criteria for multiobjective dynamic optimization programs: the vector-valued ramsey model in Banach spaces, In: Nonlinear Analysis and Convex Analysis. Korea edited by S. Akashi, D.S. Kim, T.H. Kim, G.M. Lee, W. Takahashi and T. Tanaka. (2012) 53–73.

[26] Y.Y. Haimes, L.S. Lasdon and D.A. Wismer, On a bicriterion formulation of the problems of integrated system identification and system optimization. *IEEE Trans. Syst. Man Cybern.* **1** (1971) 296–297.

[27] N.B.C Hutauruk and M. Brown, Directed multi-objective optimization for controller design. In: *Proc. International Conference on Instrumentation, Communication and Information Technology (ICICI), Bandung, Indonesia. August 3–5* (2005) 751–756.

[28] S. Jacquin, L. Jourdan and E.G. Talbi, A multi-objective dynamic programming-based metaheuristic to solve a bi-objective unit commitment problem using a multi-objective decoder. *Int. J. Metaheuristics* **5** (2016) 3–30.

[29] P. Kall and S.W. Wallace, Stochastic Programming. Wiley (1994).

[30] S. Krichen and F. Ben Abdelaziz, An optimal stopping problem with two decision makers. *Sequential Anal.* **26** (2007) 467–480.

[31] R. Klotzler, Multiobjective dynamic programming. *Math. Oper. Statist. Ser. Optim.* **9** (1978) 423–426.

[32] Y.J. Lai and C.L. Hwang, Fuzzy Mathematical Programming, Methods and Applications. Springer-Verlag (1992).

[33] S. Lenhart and J.T. Workman, Optimal Control Applied to Biological Models, Mathematical and Computational Biology. Chapman & Hall, Boca Raton, FL, USA; CRC Press, London, UK (2007).

[34] D. Liberzon, Switching in Systems and Control. Birkhauser, Boston (2003).

[35] D. Liberzon, Calculus of Variations and Optimal Control Theory. A Concise Introduction. Princeton University Press, Princeton, NJ, USA (2012).

[36] G.P. Liu, J.B. Yang and J.F. Whidborne, Multiobjective Optimisation and Control. Research Studies Press Ltd., Exeter (2003).

[37] F. Logist, B. Houska, M. Diehl and J.F. Van Impe, Robust multi-objective optimal control of uncertain (bio) chemical processe. *Chem. Eng. Sci.* **66** (2011) 4670–4682.

[38] L.S. Pontryagin, V.G. Boltyanskii, R.V. Gamkrelidze and E.F. Mishchenko (1964), The Mathematical Theory Of Optimal Processes (translated by D. E. Brown. A Pergamon Press Book). The Macmillan Co., New York (2011).

[39] A. Seierstad and K. Sydsaeter, Optimal Control Theory with Economic Applications. North-Holland, Amsterdam (1987).

[40] G.W. Taxue, R.R. Inman and D.M. Mades, Multiobjective dynamic programming with application to a reservoir. *Water Resour. Res.* **15** (1979) 1403–1408.

[41] J. Teghem, D. Dufrane, M. Thauvoye and P. Kunsch, Strange: an interactive method for multi-objective linear programming under uncertainty. *Eur. J. Oper. Res.* **26** (1986) 65–82.

[42] R. Vinter, Optimal Control. Birkhauser (2010).

[43] D.W. Walkup and R.J.B. Wets, Stochastic programs with recourse. *SIAM J. Appl. Math.* **15** (1967) 1299–1314.

[44] Q. Wei, H. Zhang and J. Dai, Model-free multiobjective approximate dynamic programming for discrete-time nonlinear systems with general performance index functions. *Neurocomputing* **72** (2009) 1839–1848.

[45] H. Zarei and M.R. Bahrmand, Multiobjective optimal control of the linear wave equation. *Ain Shams Eng. J.* **5** (2014) 1299–1305.

[46] A.M. Zohrevand, H. Rafiei and A.H. Zohrevand, Multiobjective dynamic cell formation problem: a stochastic programming approach. *Comput. Ind. Eng.* **98** (2016) 323–332.