


Article

Solution of Optimal Harvesting Problem by Finite Difference Approximations of Size-Structured Population Model

Johanna Pyy ^{1,*} , Anssi Ahtikoski ², Alexander Lapin ³ and Erkki Laitinen ¹¹ Faculty of Sciences, University of Oulu, FI-90014 Oulu, Finland; erkki.laitinen@oulu.fi² Natural Resources Institute Finland (Luke) Oulu, FI-90014 Oulu, Finland; anssi.ahtikoski@oulu.fi³ Institute of Computational Mathematics & Information Technology, Kazan Federal University, 420008 Kazan, Russia; avlapine@mail.ru

* Correspondence: pyyj@student.oulu.fi

Received: 11 April 2018; Accepted: 24 April 2018 ; Published: 26 April 2018



Abstract: We solve numerically a forest management optimization problem governed by a nonlinear partial differential equation (PDE), which is a size-structured population model. The formulated problem is supplemented with a natural constraint for a solution to be non-negative. PDE is approximated by an explicit or implicit in time finite difference scheme, whereas the cost function is taken from the very beginning in the finite-dimensional form used in practice. We prove the stability of the constructed nonlinear finite difference schemes on the set of non-negative vectors and the solvability of the formulated discrete optimal control problems. The gradient information is derived by constructing discrete adjoint state equations. The projected gradient method is used for finding the extremal points. The results of numerical testing for several real problems show good agreement with the known results and confirm the theoretical statements.

Keywords: size-structured population model; nonlinear partial differential equation; finite difference approximation; optimization; gradient method

1. Introduction

The well-posedness of the continuous size-structured model has been studied in several papers (e.g., [1–4]). In Ref. [1], authors proved the local existence and uniqueness of a solution of the continuous model, where birth and mortality functions depend on total population. In Ref. [2], the authors established the local existence and uniqueness of a solution of the size-structured nonlinear population model, where also growth rate depends on total population. In the papers [3,4], the authors proved global existence and uniqueness of a solution of the continuous nonlinear population model, where all vital rates depend on total population. The total population can be described by e.g., total number of individuals (e.g., [3]), total biomass (e.g., [5]) or basal area.

A continuous nonlinear size-structured population model has been used in a forest management optimization problem (e.g., [6–8]). In a continuous formulation, this nonlinear optimization problem cannot be solved by analytic methods. A natural approach is to solve this problem by approximating a continuous model by a discrete one and further solving a discrete optimization problem by iterative algorithms. In this paper, we focus on development of finite difference schemes to approximate the solution of a continuous nonlinear population model. Efficient schemes are essential for solving optimal control problems or parameter estimation problems as such problems require solving the model numerous times before an optimal solution is obtained.

When continuous population model is approximated by a finite difference scheme, it becomes a matrix population model [9]. In matrix models, trees are divided into classes with respect to their

size—for instance, diameter. The matrix describes how the class division changes at one time step. Matrix population models have also been used for forest management optimizations (e.g., [10,11]).

In optimization, using iterative algorithms is inevitable. Higher-order algorithms are usually sensitive to the regularity of the solution, and, therefore, they usually yield a convergence rate of first order as soon as the compatibility conditions are not satisfied. Moreover, in practice, the vital rates are determined on a statistical basis and the compatibility conditions required for high-order convergence are hardly valid with real-life data. These suggest that, in most cases, a first-order method should be the most adequate. Hence, it is desirable to have a robust scheme that can produce many useful qualitative and quantitative properties of the solutions of the differential problem but requires minimum regularity of the solution [12]. Unfortunately, one could not derive the explicit formula for the optimal strategy since the strategy, the state and the costate are coupled into a complex system. The results at this stage may be regarded as a middle step to real world applications and serve as a starting point for numerical computations [13].

In our knowledge, the comprehensive theoretical investigation of the forest management optimization problem with a continuous nonlinear population model as the state equation is still lacking, and, in that sense, the problem is an open problem. Hence, in this work, the investigation of the problem in its differential form is omitted, whereas, we consider the finite dimensional counterpart of the problem constructed by finite difference approximations of the state problem and taking for cost function a finite dimensional form used in practice. We prove the stability of constructed finite difference schemes on the set of non-negative solutions and solvability of the optimization problem, and deduce the necessary gradient information for iterative solution methods. We solve several applied problems, where different approximations schemes are used, and compare the computed results. The rest of the article is organized as follows. In Section 2, a mathematical model of optimal harvesting problem for the size-structured forest is formulated. In Section 3, we construct and investigate two finite difference approximation schemes for a nonlinear boundary value problem that simulates the growth and the harvesting of a forest. A gradient method for minimizing the cost function is constructed in Section 4. The theoretical details of this method are set out in the Appendix A to the article. Section 5 is devoted to the numerical solution of a real-life problem and comparative analysis of the computing results. Finally, in Section 6, we present discussions.

2. Formulation of the Optimal Control Problem

In order to formulate the mathematical model for the optimal harvesting problem for the size-structured forest, we define the following notations. In space, we denote by $x \in \Omega := (L_0, L]$ the thickness of the tree, where L_0 and L are the lower and upper bounds of the space domain, respectively. Moreover, $t \in (0, T]$ is the time, where T is the upper limit. By Q , we denote the product space $\Omega \times (0, T]$. We denote by $y(x, t)$ and $h(x, t)$ the number of trees per unit area (state) and the number of removed trees per unit area (control), respectively. Now, the optimal harvesting problem where the cost functional $J(y, h)$ characterizes a net present value (NPV) of ongoing rotation, and $d(x, t)$ is the discounted price function, is formulated as follows:

$$\max_{(y,h) \in \mathcal{K}} J(y, h) := \int_Q d(x, t)h(x, t). \tag{1}$$

Above $\mathcal{K} = \mathcal{Y}_{ad} \times \mathcal{H}_{ad}$ is the set of constraints for the state and the control, where

$$\mathcal{Y}_{ad} = \{y \mid \text{for all } (x, t) \in Q : y(x, t) \geq 0; y \text{ is a solution for Equations (4)–(6)}\}, \tag{2}$$

$$\mathcal{H}_{ad} = \{h \mid \text{for all } (x, t) \in Q : 0 \leq h(x, t) \leq h_{max}, \text{ for all } t \in (0, T] : \int_{\Omega} h(x, t) \geq B \text{ or } \int_{\Omega} h(x, t) = 0\} \tag{3}$$

From the point of real-life problems, it is obvious that there exist constants $h_{max} > 0$ and B , which denote the upper limit for harvesting and lower limit for making profitable thinning of trees at

time event t ; otherwise, the thinning is not done. Notice that harvesting h depends on the state y (via constraint sets), which is defined by the population model

$$\frac{\partial y(x, t)}{\partial t} + \frac{\partial(g(x, P(t))y(x, t))}{\partial x} + m(x, P(t))y(x, t) + h(x, t) = 0, \text{ in } Q, \tag{4}$$

$$g(0, P(t))y(0, t) = 0, \text{ in } (0, T], \tag{5}$$

$$y(x, 0) = y^0(x), \text{ in } \Omega, \tag{6}$$

where $g(x, P(t))$ is growth rate, $m(x, P(t))$ mortality rate and $y^0(x) \geq 0$ is initial diameter distribution of the trees. Growth and mortality rates depend on diameter x of a tree and on the basal area, $P(t)$, of the forest stand, where

$$P(t) = \pi \int_0^L \left(\frac{x}{2}\right)^2 y(x, t) dx.$$

In the case $h = 0$, the problems (4)–(6) are a particular case of the problem that have been investigated in [1–4]. In these articles, the existence of a non-negative continuous solution of this problem has been proved under some "natural" assumptions for input data. They are:

1. $g(x, P)$ is continuous and strictly positive for all x and P and continuously differentiable with respect to x ;
2. $m(x, P)$ is non-negative for x and P and integrable in x ;
3. $g(x, P)$ and $m(x, P)$ are Lipschitzian with respect to P ;
4. $\sup_{x, P} m(x, P) < \infty$.

We also assume that these assumptions are satisfied. We use growth rate g and mortality m in a bilinear form

$$g(x, P(t)) = g_{11} + g_{12}x + (g_{21} + g_{22}x)P(t),$$

$$m(x, P(t)) = m_{11} + \frac{m_{12}}{x} + \frac{m_{13}}{x^2} + \left(m_{21} + \frac{m_{22}}{x} + \frac{m_{23}}{x^2}\right)P(t),$$

where the constants g_{ij} and m_{ij} are such that $g(x, P) > 0$ and $m(x, P) \geq 0$ for all $x \in \Omega$ and $P \geq 0$. Obviously, because of suppositions $g(0, P(t)) > 0$, the boundary condition (5) reads as $y(0, t) = 0$.

The optimal harvesting problem has been investigated in [6–8]. The authors of these publications considered the case where the harvesting function has the form $h(x, t) = c(x, t)y(x, t)$, where $c(x, t)$ is the control. Thus, they investigated a coefficient identification problem while we solve an optimal control problem with distributed (on the right-hand side) control.

3. Finite Difference Approximations

In this chapter, we derive explicit and (semi)implicit finite difference approximations for the state problems (4)–(6) and prove their stability estimates on non-negative solutions. The investigation of existence, uniqueness and convergence of approximations is beyond the scope of our article. For the size-structured population model with recruitment, the existence, uniqueness and convergence of explicit approximations is investigated in [14] and implicit approximation in [5,15].

The following notations are used throughout the paper: $\Delta t = \frac{T}{M}$ and $\Delta x = \frac{L-L_0}{N}$ denote the temporal and spatial mesh size, respectively. The non-overlapping mesh intervals are $(t^{k-1}, t^k]$, $k = 1, \dots, M$, and $(x_{i-1}, x_i]$, $i = 1, \dots, N$, where $t^0 = 0, t^M = T, x_0 = L_0, x_N = L$.

Let us denote by y_i^k and h_i^k the finite difference approximations of $y(x_i, t^k)$ and $h(x_i, t^k)$, respectively. Moreover, we denote $g_i^k := g(x_i, P^k)$ and $m_i^k := m(x_i, P^k)$ the discrete values of the growth rate and mortality rate, respectively, in size class $[x_{i-1}, x_i]$. The discretized value of the basal area at time t^k is $P^k := \pi \sum_{i=1}^N \left(\frac{x_i}{2}\right)^2 y(x_i, t^k)$.

3.1. Explicit Approximation of the State Equation

For all meshpoints $i = 1, \dots, N$; $k = 1, \dots, M$, the explicit finite difference approximation of the size-structured population model (4)–(6) reads

$$\begin{aligned} \frac{y_i^k - y_i^{k-1}}{\Delta t} + \frac{g_i^{k-1}y_i^{k-1} - g_{i-1}^{k-1}y_{i-1}^{k-1}}{\Delta x} + m_i^{k-1}y_i^{k-1} + h_i^k &= 0, \\ y_0^k &= 0, \\ y_i^0 &\geq 0 \text{ constant.} \end{aligned} \tag{7}$$

Note that we use so-called upwind approximation for the first order derivative in space (variable x) using the positivity of coefficient $g(x, P)$ on the set of non-negative mesh functions y . The explicit scheme (7) can be written in the form:

$$y_i^k - \left(1 - \frac{\Delta t}{\Delta x}g_i^{k-1} - \Delta tm_i^{k-1}\right)y_i^{k-1} - \left(\frac{\Delta t}{\Delta x}g_{i-1}^{k-1}\right)y_{i-1}^{k-1} + \Delta th_i^k = 0.$$

Later on, we denote by $a_i^k = 1 - \frac{\Delta t}{\Delta x}g_i^k - \Delta tm_i^k$, and, $b_i^k = \frac{\Delta t}{\Delta x}g_i^k$. Moreover, we denote by $\mathbf{y}^k := (y_1^k, \dots, y_N^k)$, $\mathbf{h}^k := (h_1^k, \dots, h_N^k)$ the vectors of the nodal values and by

$$\mathbf{A}^k = \begin{bmatrix} a_1^k & 0 & 0 & \dots & 0 & 0 \\ b_1^k & a_2^k & 0 & \dots & 0 & 0 \\ 0 & b_2^k & a_3^k & \dots & 0 & 0 \\ \vdots & \vdots & \ddots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & b_{N-1}^k & a_N^k \end{bmatrix}$$

the matrix of coefficients. Now, we can write explicit difference scheme (7) in the following algebraic form:

$$\mathbf{y}^k - \mathbf{A}^{k-1}\mathbf{y}^{k-1} + \Delta t\mathbf{h}^k = 0, \quad k = 1, \dots, M. \tag{8}$$

Note that this scheme is just the forest growth model studied in [11]. Moreover, the numerical calculation of the next temporal state involves only matrix to vector calculations. The drawback of the explicit scheme is that the following stability condition (9) must be satisfied.

Lemma 1. *Let the condition*

$$\Delta x \geq \Delta t \sup_{x,t} g(x, P(t)) \tag{9}$$

be satisfied. Then, on the set of non-negative mesh functions y , the finite difference scheme (7) is stable

$$\max_k \|\mathbf{y}^k\|_1 \leq C(T)(\|\mathbf{y}^0\|_1 + \sum_{k=1}^M \Delta t \|\mathbf{h}^k\|_1), \tag{10}$$

where $\|v\|_1 = \sum_{i=1}^N |v_i|$.

Proof. On the non-negative mesh functions y , the coefficients $g_i(P)$ are positive and $m_i(P) \geq 0$. For the mesh steps satisfying condition (9), the diagonal entries of matrix \mathbf{A}^k satisfy the inequality

$$|a_i^k| \leq 1 - \frac{\Delta t}{\Delta x}g_i(P^k) + \Delta tm_i(P^k).$$

Because of this inequality, we have the following estimate for $\|\cdot\|_1$ -norm of matrices, connected with $\|\cdot\|_1$ -norm of vectors:

$$\|\mathbf{A}^k\|_1 = \sum_{i=1}^N (|a_i^k| + |b_i^k|) \leq 1 + C_m \Delta t$$

with $C_m = \sup_{x,t} m(x, P(t))$. Due to this estimate and condition (9), we obtain from Equation (8) the inequality

$$\|\mathbf{y}^k\|_1 \leq (1 + C_m \Delta t) \|\mathbf{y}^{k-1}\|_1 + \Delta t \|\mathbf{h}^k\|_1 \text{ for all } k = 1, 2, \dots, M,$$

whence stability estimate (10) follows. \square

The condition (9) means that the length of the time step Δt and width of the size class Δx have to be chosen so that a tree cannot grow over one size class during one time step Δt (compare with [16]).

3.2. Implicit Approximation of the State Equation

For all meshpoints $i = 1, \dots, N$; $k = 1, \dots, M$, the implicit finite difference approximation of the models (4)–(6) is the following linearized problem, with nonlinear coefficients calculated on the previous time level:

$$\begin{aligned} \frac{y_i^k - y_i^{k-1}}{\Delta t} + \frac{g_i^{k-1} y_i^k - g_{i-1}^{k-1} y_{i-1}^k}{\Delta x} + m_i^{k-1} y_i^k + h_i^k &= 0, \\ y_0^k &= 0, \\ y_i^0 &\geq 0 \text{ constant.} \end{aligned} \tag{11}$$

Equation (11) can be rewritten as

$$\left(1 + \frac{\Delta t}{\Delta x} g_i^{k-1} + \Delta t m_i^{k-1}\right) y_i^k - \frac{\Delta t}{\Delta x} g_{i-1}^{k-1} y_{i-1}^k - y_i^{k-1} + \Delta t h_i^k = 0,$$

for all $i = 1, \dots, N$, $k = 1, \dots, M$.

Using the notations $a_i^k = 1 + \frac{\Delta t}{\Delta x} g_i^{k-1} + \Delta t m_i^{k-1}$ and $b_i^k = -\frac{\Delta t}{\Delta x} g_{i-1}^{k-1}$, we rewrite Equation (11) in a form of linear algebraic equations

$$\mathbf{B}^{k-1} \mathbf{y}^k - \mathbf{y}^{k-1} + \Delta t \mathbf{h}^k = 0, \quad k = 1, \dots, M, \tag{12}$$

where

$$\mathbf{B}^k := \begin{bmatrix} a_1^k & 0 & 0 & \dots & 0 & 0 \\ b_1^k & a_2^k & 0 & \dots & 0 & 0 \\ 0 & b_2^k & a_3^k & \dots & 0 & 0 \\ \vdots & \vdots & \ddots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & b_{N-1}^k & a_N^k \end{bmatrix}$$

is a matrix of nonlinear coefficients.

Lemma 2. *Finite difference scheme (11) is unconditionally stable on the set of non-negative mesh functions y : for any Δt and Δx the following stability estimate holds:*

$$\max_k \|\mathbf{y}^k\|_1 \leq \|\mathbf{y}^0\|_1 + \sum_{k=1}^M \Delta t \|\mathbf{h}^k\|_1. \tag{13}$$

Proof. By direct calculations, we obtain from Equation (12) the equality

$$\sum_{i=1}^N a_i^k y_i^k + \sum_{i=1}^{N-1} b_i^k y_i^k - \sum_{i=1}^N y_i^{k-1} + \Delta t \sum_{i=1}^N h_i^k = 0.$$

Since $a_i^k + b_i^k \geq 1$ and $a_N^k \geq 1$, then, from this equality, we get

$$\sum_{i=1}^N y_i^k \leq \sum_{i=1}^N y_i^{k-1} + \Delta t \sum_{i=1}^N |h_i^k|.$$

Because of positivity of vectors \mathbf{y}^k and \mathbf{y}^{k-1} , the last inequality can be written in the form

$$\|\mathbf{y}^k\|_1 \leq \|\mathbf{y}^{k-1}\|_1 + \Delta t \|\mathbf{h}^k\|_1 \quad \forall k,$$

whence stability estimate (13) follows. \square

Notice, contrary to the explicit scheme the time step Δt and class width Δx has no mutual dependence, hence the growth of a tree during a time step is not restricted less than one size class. This characteristic of the implicit scheme is useful in the optimal harvesting problem, covered by the models (4)–(6) or parameter identification problem because such problems require solving the model many times before an optimal solution is obtained.

3.3. Approximation of the Optimal Control Problem

We denote $d_i^k := d(x_i, t^k)$ the discounted price for size class $(x_{i-1}, x_i]$ at time t^k , and $\mathbf{d}^k = (d_1^k, \dots, d_N^k)$. Moreover, $(\mathbf{u}, \mathbf{v}) := \sum_{i=1}^N u_i v_i$ is the vector product of vectors $\mathbf{u}, \mathbf{v} \in \mathbb{R}^N$. Approximating the cost function (1) by the right-hand Riemann sum, we get the following approximation for the harvesting problem:

$$\max_{(\mathbf{y}, \mathbf{h}) \in K} \left\{ \bar{J}(\mathbf{y}, \mathbf{h}) := \sum_{k=1}^M (\mathbf{d}^k, \mathbf{h}^k) = \sum_{k=1}^M \sum_{i=1}^N d_i^k h_i^k \right\}. \tag{14}$$

Above, we denote by $K = Y_{ad} \times H_{ad}$, where

$$Y_{ad} = \{(\mathbf{y}, \mathbf{h}) \mid \mathbf{y} \geq \mathbf{0}, \mathbf{y} \text{ is a solution for Equations (7) or (11)}\}, \tag{15}$$

$$H_{ad} = \{\mathbf{h} \mid \mathbf{0} \leq \mathbf{h}^k \leq \mathbf{h}_{\max}, \|\mathbf{h}^k\|_1 \geq B \text{ or } \mathbf{h}^k = \mathbf{0}, k = 1, \dots, M\}. \tag{16}$$

Moreover, $\mathbf{y} = (\mathbf{y}^1, \dots, \mathbf{y}^M)$ and $\mathbf{h} = (\mathbf{h}^1, \dots, \mathbf{h}^M)$.

The following propositions show that the discrete optimal harvesting problem (14) has at least one solution in both cases, i.e., if models (4)–(6) is approximated explicitly or implicitly.

Proposition 1. *Let the mesh steps Δt and Δx satisfy the inequality*

$$1 - \frac{\Delta t}{\Delta x} \sup_{x,t} g(x, P(t)) - \Delta t \sup_{x,t} m(x, P(t)) \geq 0. \tag{17}$$

Then, Problem (14) has at least one solution if \mathbf{y} satisfies Equation (7).

Proof. The set K is non-empty. In fact, due to assumption (17), the solution \mathbf{y} of finite difference scheme (7) with $\mathbf{y}^0 \geq \mathbf{0}$ is non-negative if $\mathbf{h} = \mathbf{0}$. This statement can be easily verified using form (8) of the difference scheme and noting that all entries a_i^k and b_i^k of the matrices \mathbf{A}^k are non-negative.

Obviously, assumption (9) follows from inequality (17), so stability estimate (10) holds. Since vector $\mathbf{h} \in K$ is bounded, then, due to inequality (10), there exists a constant Y such that $\|\mathbf{y}\|_1 \leq Y$, i.e., the set K is bounded. It is closed because of the continuity of functions $g(P)$ and $m(P)$ with respect

to P , while P is obviously continuous with respect to \mathbf{y} . Thus, K is compact. At last, cost function \bar{J} of Problem (14) is continuous, whence the existence of a solution to Problem (14) follows from Weierstrass's theorem. \square

Proposition 2. *Problem (14) has at least one solution if \mathbf{y} satisfies Equation (11).*

Proof. Proof is very similar to the proof of Proposition 1. Namely, the set K is non-empty because, for $\mathbf{h} = \mathbf{0}$, the solution \mathbf{y} of finite difference scheme (11) is non-negative for all Δx and Δt . Since \mathbf{h} is bounded, then, due to stability estimate (13), \mathbf{y} is also bounded, so the set K is bounded. It is closed because of the continuity of functions $g(P)$ and $m(P)$ with respect to P , and continuity of P with respect to \mathbf{y} . Thus, K is compact. At last, cost function \bar{J} of Problem (14) is continuous, whence the existence of a solution to Problem (14) follows. \square

Remark 1. *Since neither the function \bar{J} is strictly concave nor the set K is strictly convex, the optimization problem can have a non-unique solution.*

4. Realization of the Optimal Strategies

In this section, a first order method to approximate the optimal harvesting problem (14) is constructed. In real-life applications, the growth rate g and mortality rate m are determined on a statistical basis and the compatibility conditions required for high-order methods can be hardly validated. Hence, a first order method, which is desirable to have a robust scheme but requires minimum regularity of the solution should be the most adequate. The first order methods require computing of the Fréchet derivatives (Jacobian matrix), which can be computationally expensive. However, when we consider the nonlinear optimization problem, only the gradient of the object function is needed, and the gradient can be computed without the Fréchet derivatives. In this work, the adjoint approach developed in the 1970s in [17] is applied for calculation of the functional gradient. The adjoint method has a great advantage against the direct method because only one linear state problem, so called adjoint state, need to be solved for obtaining the gradient information. Today, it is a well-known method for computing the gradient of a functional with respect to model parameters when this functional depends on those model parameters through state variables, which are solutions of the state problem. However, this method is less well understood in the control of population models, and, as far as we know, no applications to distributed optimal control of harvesting is presented in literature. Duality and adjoint equations are essential tools in studying existence of the optimal pair (\mathbf{y}, \mathbf{h}) , and, for a periodic age-dependent harvesting problem and for age-spatial structured harvesting problem, it is applied for proving the existence of the bang-bang control in [18] and in [19,20], respectively. For continuous size-structured harvesting, problem duality and adjoint equations are applied for proving the existence of the bang-bang control in [6,8].

In this work, we apply the Lagrange method and give a recipe to systematically define the adjoint state equations and gradient information. We formulate the Lagrangian of the problem (14) with respect to the state constraint (15) only, and use the projection method regarding the control constraint (16). In the projection method, if solution goes outside the constraint set (16), it is projected back to there. Let us generalize and denote by $A(\mathbf{y}, \mathbf{h}) = 0$ the operator Equation (8) (or Equation (12)). Moreover, $A^k(\mathbf{y}, \mathbf{h}) = 0$ is the operator equation at the time level k , $k = 1, \dots, M$.

Suppose the functional \bar{J} and operator A to be differentiable in the sense that there exist the following partial derivatives:

$$\begin{aligned} \bar{J}_y \delta \mathbf{y} &= \lim_{t \rightarrow 0} \frac{\bar{J}(\mathbf{y} + t\delta \mathbf{y}, \mathbf{h}) - \bar{J}(\mathbf{y}, \mathbf{h})}{t}, \quad \bar{J}_h \delta \mathbf{h} = \lim_{t \rightarrow 0} \frac{\bar{J}(\mathbf{y}, \mathbf{h} + t\delta \mathbf{h}) - \bar{J}(\mathbf{y}, \mathbf{h})}{t}, \\ A_y \delta \mathbf{y} &= \lim_{t \rightarrow 0} \frac{A(\mathbf{y} + t\delta \mathbf{y}, \mathbf{h}) - A(\mathbf{y}, \mathbf{h})}{t}, \quad A_h \delta \mathbf{h} = \lim_{t \rightarrow 0} \frac{A(\mathbf{y}, \mathbf{h} + t\delta \mathbf{h}) - A(\mathbf{y}, \mathbf{h})}{t} \end{aligned}$$

for all vectors $\delta \mathbf{y}$ and $\delta \mathbf{h}$ (or at least for the vectors such that $\mathbf{y}, \mathbf{y} + \delta \mathbf{y} \in Y_{ad}$ and $\mathbf{h}, \mathbf{h} + \delta \mathbf{h} \in H_{ad}$). Note that for the fixed \mathbf{h} and \mathbf{y} , $\bar{J}_y \equiv \bar{J}_y(\mathbf{y}, \mathbf{h})$ and $\bar{J}_h \equiv \bar{J}_h(\mathbf{y}, \mathbf{h})$ are vectors while $A_y = A_y(\mathbf{y}, \mathbf{h})$ and $A_h = A_h(\mathbf{y}, \mathbf{h})$ are matrices. By A_y^* and A_h^* , we denote the corresponding transpose matrices.

Let us define Lagrange function, \mathcal{L} , of the problem (14) by

$$\mathcal{L}(\mathbf{y}, \mathbf{h}, \lambda) = \bar{J}(\mathbf{y}(\mathbf{h}), \mathbf{h}) - \sum_{k=1}^M (\lambda^k, A^k(\mathbf{y}, \mathbf{h})),$$

where $\lambda^k \in \mathbb{R}^N$. Now, for all feasible pair (\mathbf{y}, \mathbf{h}) holds $A(\mathbf{y}, \mathbf{h}) = 0$, and, for any λ , we have:

$$\mathcal{L}(\mathbf{y}, \mathbf{h}, \lambda) = \bar{J}(\mathbf{y}(\mathbf{h}), \mathbf{h}),$$

and, since λ does not depend on \mathbf{h} , we have

$$\frac{\partial \bar{J}}{\partial \mathbf{h}} = \frac{\partial \mathcal{L}(\mathbf{y}, \mathbf{h}, \lambda)}{\partial \mathbf{y}} \frac{\partial \mathbf{y}}{\partial \mathbf{h}} + \frac{\partial \mathcal{L}(\mathbf{y}, \mathbf{h}, \lambda)}{\partial \mathbf{h}}. \tag{18}$$

Above, one method to approximate $\frac{\partial \mathbf{y}}{\partial \mathbf{h}}$ is to compute N finite differences over control variable \mathbf{h} . However, each computation requires solving the equation $A(\mathbf{y}, \mathbf{h}) = 0$, and, for large N , this method is computationally expensive. In the *adjoint method*, we can avoid to compute $\frac{\partial \mathbf{y}}{\partial \mathbf{h}}$ by solving the linear adjoint state equation only once.

The theory of constrained optimization, see [21], says that (\mathbf{y}, \mathbf{h}) is the optimal pair for the problem (14) if $(\mathbf{y}, \mathbf{h}, \lambda)$ is a saddle point of \mathcal{L} . The derivatives of \mathcal{L} with respect to \mathbf{y} , \mathbf{h} and λ are:

$$\begin{aligned} \frac{\partial \mathcal{L}(\mathbf{y}, \mathbf{h}, \lambda)}{\partial \mathbf{y}} &= \frac{\partial \bar{J}(\mathbf{y}, \mathbf{h})}{\partial \mathbf{y}} - \left(\frac{\partial A(\mathbf{y}, \mathbf{h})}{\partial \mathbf{y}} \right)^* \lambda, \\ \frac{\partial \mathcal{L}(\mathbf{y}, \mathbf{h}, \lambda)}{\partial \mathbf{h}} &= \frac{\partial \bar{J}(\mathbf{y}, \mathbf{h})}{\partial \mathbf{h}} - \left(\frac{\partial A(\mathbf{y}, \mathbf{h})}{\partial \mathbf{h}} \right)^* \lambda, \\ \frac{\partial \mathcal{L}(\mathbf{y}, \mathbf{h}, \lambda)}{\partial \lambda} &= -A(\mathbf{y}, \mathbf{h}). \end{aligned}$$

Now, $\frac{\partial \mathcal{L}(\mathbf{y}, \mathbf{h}, \lambda)}{\partial \lambda} = 0$ gives the state equation, $\frac{\partial \mathcal{L}(\mathbf{y}, \mathbf{h}, \lambda)}{\partial \mathbf{y}} = 0$ gives the adjoint state equation and $\frac{\partial \bar{J}}{\partial \mathbf{h}} = \frac{\partial \mathcal{L}(\mathbf{y}, \mathbf{h}, \lambda)}{\partial \mathbf{h}}$ gives the gradient.

Now, the calculation of the gradient can be summarized by the following steps when the Lagrangian $\mathcal{L}(\mathbf{y}, \mathbf{h}, \lambda) = \bar{J}(\mathbf{y}, \mathbf{h}) - (\lambda, A(\mathbf{y}, \mathbf{h}))$ is first formulated:

- I Solve the state equation $A(\mathbf{y}, \mathbf{h}) = 0$;
- II Solve the adjoint state equation

$$\frac{\partial \mathcal{L}(\mathbf{y}, \mathbf{h}, \lambda)}{\partial \mathbf{y}} = \frac{\partial \bar{J}(\mathbf{y}, \mathbf{h})}{\partial \mathbf{y}} - \left(\frac{\partial A(\mathbf{y}, \mathbf{h})}{\partial \mathbf{y}} \right)^* \lambda = 0;$$

- III Compute the gradient

$$\frac{\partial \bar{J}}{\partial \mathbf{h}} = \frac{\partial \bar{J}(\mathbf{y}, \mathbf{h})}{\partial \mathbf{h}} - \left(\frac{\partial A(\mathbf{y}, \mathbf{h})}{\partial \mathbf{h}} \right)^* \lambda.$$

Partial derivatives of $\bar{J}(\mathbf{y}, \mathbf{h})$ and $A(\mathbf{y}, \mathbf{h})$ are presented in Appendix A. Gradient $\frac{\partial \bar{J}}{\partial \mathbf{h}}$ we used in projected gradient method [22], which we applied for iteration of a solution of the optimal harvesting problem.

5. Numerical Example

In this section, we study numerical examples of problem (14). We compared two cases where the state constraint (4) was approximated with explicit approximation (7) and implicit approximation (11). As the discounted price for size class $(x_{i-1}, x_i]$ at time t^k , we used $d_i^k = \frac{c^p v_i^p + c^s v_i^s}{(1+r)^{t^k}}$, where r is the interest rate, c^p and c^s are the prices of the pulpwood and sawlog, respectively, and v_i^p and v_i^s are the volumes of pulpwood and sawlog of a tree in size class $(x_{i-1}, x_i]$, respectively. In the optimizations, we used the following values for parameters: price of pulpwood $c^p = 16.56 \text{ €m}^{-3}$ and sawlog $c^s = 58.44 \text{ €m}^{-3}$, interest rate $r = 3\%$ and lower bound for harvested trees $B = 50 \text{ m}^3 \text{ ha}^{-1}$. The pulpwood and sawlog volumes v_i^p and v_i^s we got from [10]. The optimization results of problem (14) are presented in Tables 1 and 2, and in Figures 1 and 2.

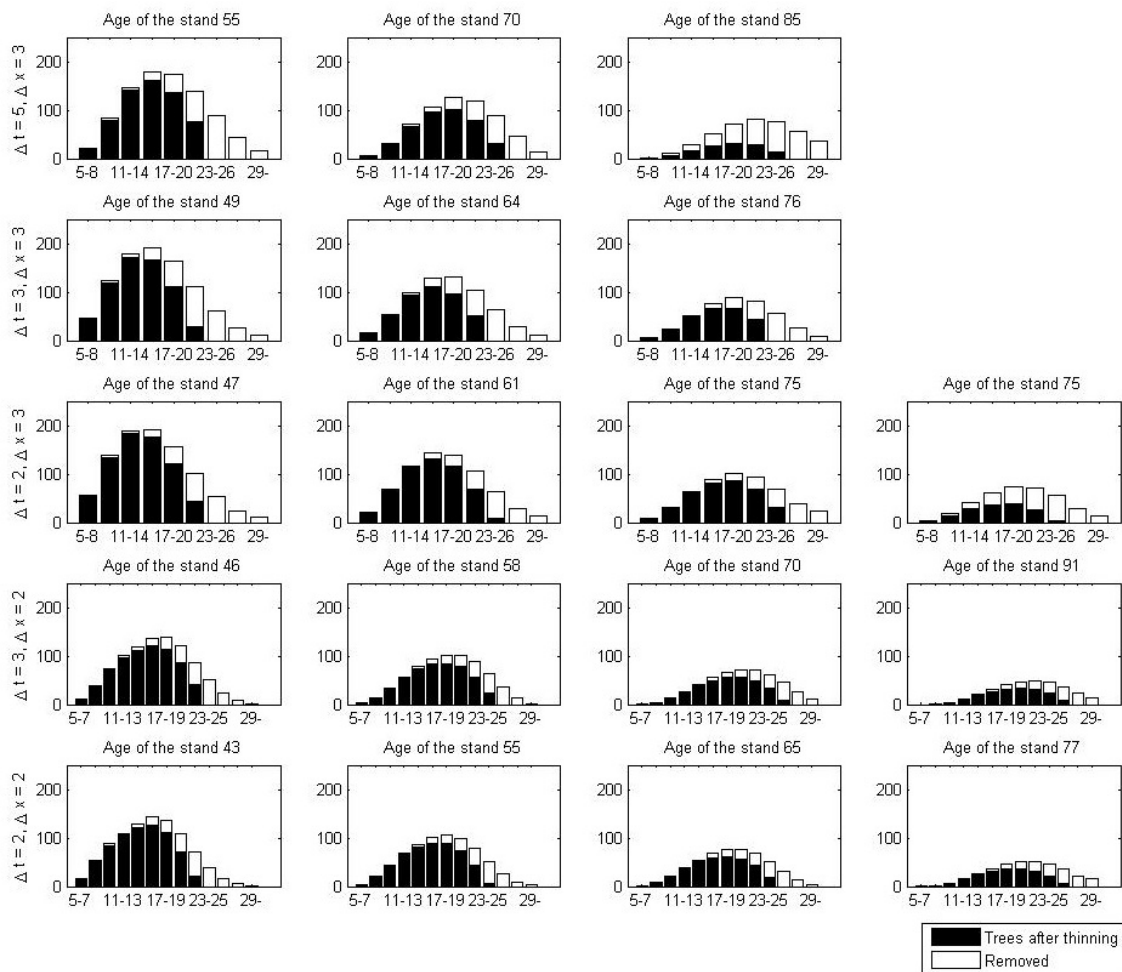


Figure 1. Diameter distributions associated with optimal thinnings of problem (14) with state equation approximated by explicit scheme (7). Numbers (e.g., 5–8, 11–14) represent diameter in centimetres.

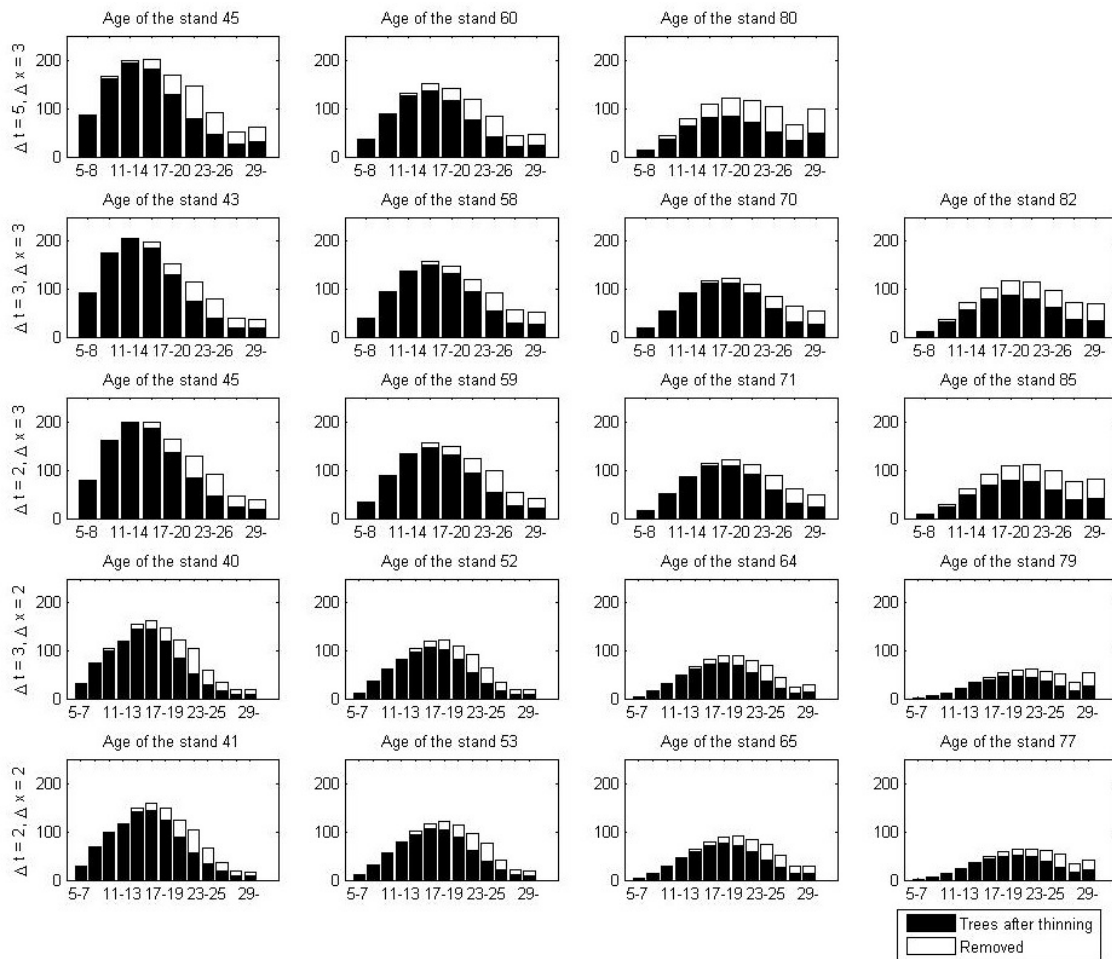


Figure 2. Diameter distributions associated with optimal thinnings of problem (14) with state equation approximated by implicit scheme (11). Numbers (e.g., 5–8, 11–14) represent diameter in centimetres.

Table 1. Maximum net present values, MaxNPVs, (i.e., optimal cost function values of the problem (14)) and mean annual increments (MAI) associated with optimal stand-level managements. Initial density is 1000 stems ha⁻¹.

	MaxNPV (€ha ⁻¹)		MAI (m ³ ha ⁻¹ a ⁻¹)	
	Explicit	Implicit	Explicit	Implicit
Time step $\Delta t = 5$ years and class width $\Delta x = 3$ cm	4095	3996	2.94	2.83
Time step $\Delta t = 3$ years and class width $\Delta x = 3$ cm	4301	4231	2.94	3.11
Time step $\Delta t = 3$ years and class width $\Delta x = 2$ cm	4676	4648	2.96	3.00
Time step $\Delta t = 2$ years and class width $\Delta x = 3$ cm	4384	4311	2.93	3.04
Time step $\Delta t = 2$ years and class width $\Delta x = 2$ cm	4776	4726	3.09	3.09

Table 2. Optimal stand-level managements generated by explicit (7) and implicit (11) approximation of the forest growth model (4). Initial density is 1000 stems ha⁻¹.

	Stand Age (a)		Removal (m ³ ha ⁻¹)		Thinning Intensity (% of Basal Area Removed)		Saw Log Proportion (%)	
	Explicit	Implicit	Explicit	Implicit	Explicit	Implicit	Explicit	Implicit
Time step $\Delta t = 5$ years and class width $\Delta x = 3$ cm								
1st thinning	55	45	91.0	77.2	48	48	87	85
2nd thinning	70	60	68.9	59.3	46	46	88	87
3rd thinning	85	80	98.6	90.6	79	72	87	86
final thinning	105	105	50.0	69.9	100	100	90	86
total			308.5	297.0			88	86
Time step $\Delta t = 3$ years and class width $\Delta x = 3$ cm								
1st thinning	49	43	78.9	52.4	48	35	83	86
2nd thinning	64	58	68.2	54.3	50	35	86	89
3rd thinning	76	70	55.4	50.0	56	35	87	91
4th thinning	-	82	-	72.4	-	59	-	86
final thinning	100	100	91.5	82.2	100	100	88	85
total			293.9	311.3			86	87
Time step $\Delta t = 3$ years and class width $\Delta x = 2$ cm								
1st thinning	46	40	60.6	60.5	36	40	83	81
2nd thinning	58	52	56.5	51.8	38	40	85	83
3rd thinning	70	64	59.7	52.1	45	46	88	86
4th thinning	79	79	50.0	55.0	52	56	89	90
final thinning	94	97	72.6	71.4	100	100	89	86
total			299.5	290.7			87	85
Time step $\Delta t = 2$ years and class width $\Delta x = 3$ cm								
1st thinning	47	45	62.2	59.1	39	38	84	86
2nd thinning	61	59	59.9	54.8	41	37	87	88
3rd thinning	75	71	58.9	50.0	44	37	90	90
4th thinning	83	85	64.4	78.0	70	64	84	87
final thinning	101	103	50.0	71.0	100	100	86	86
total			295.5	313.3			86	87
Time step $\Delta t = 2$ years and class width $\Delta x = 2$ cm								
1st thinning	43	41	55.3	58.5	36	38	80	82
2nd thinning	55	53	55.6	53.3	39	39	84	84
3rd thinning	65	65	50.0	52.0	42	44	86	87
4th thinning	77	77	53.0	50.2	53	51	89	89
final thinning	95	95	79.5	79.2	100	100	89	87
total			293.4	293.1			86	86

The results show that the maximum net present value (NPV) associated with the explicit approximation (7) was higher than the corresponding of the implicit approximation (Table 1). When class width or time step decreased, maximum NPV increased in both cases. The difference of maximum NPVs between the two cases decreased when class width or time step decreased. Only when time step decreased from three to two years, the difference of maximum NPVs increased. The difference was biggest (99 €) when time step $\Delta t = 5$ years and class width $\Delta x = 3$ cm and smallest (28 €) when time step $\Delta t = 3$ years and class width $\Delta x = 2$ cm.

With both approximations, three or four intermediate thinnings were made (Table 2). Number of thinnings increased when time step and class width decreased. When implicit approximation (11) was used, first thinnings were made 1–2 time steps earlier, while the last few thinnings were made 0–2 time steps later than when explicit approximation (7) was used. The thinning intensities were almost identical between the two approximations. If there was some difference, intensity was usually bigger when explicit approximation (7) was used (Table 2). The thinning pattern was in all optimal managements quite similar: in each thinning, more big trees than small ones were removed indicating a thinning from-above method (for different thinning types, see e.g., [23], pp. 727, 733). Thinning from

above has proven to be the best thinning type in stand-level optimizations of even-aged boreal forests (e.g., [24]). When explicit approximation (7) was used, all trees from two or three of the biggest size classes were removed (Figure 1). On the other hand, when implicit approximation (11) was applied, only part of the trees from those size classes were removed (Figure 2).

6. Discussion

This study contributes to existing literature on forest management by providing a theoretically sound framework to solve nonlinear optimization problem of even-aged stands. We compared the results of forest management optimizations, when the explicit and implicit approximations of the forest growth model was used. The optimization results show that the differences of the results between approximations are diminutive. This was expected as solutions of both approximation equations are proved to converge to solutions of continuous equation [5,14].

In numerical examples, we used data from the Scots pine (*Pinus sylvestris* L.) stands that were located in Northern Ostrobothnia, Finland, on nutrient-poor soil type. The data was the same as in [11]. The difference is that, in [11], data was fitted directly to the matrix model, as, in this study, we first fitted data to the continuous model and then approximated it with a matrix model. In [11], the time step was five years and class width 3 cm. The results are in line with each other. Both methods gave four thinnings in optimal management and thinning from above dominated as the thinning type. In [11], the optimal net present value was slightly higher and, in the optimal management, the thinnings were made slightly earlier than in this study.

The optimal harvesting problem with a continuous size-structured population model was studied in [6–8]. In those papers, harvesting was defined as a proportion of removed trees. The maximum principle for the problem was proved in [6,8]. Moreover, in [7], the strong bang-bang principle under some additional (but realistic) conditions was proved. This means that the optimal solution has the structure, where all trees bigger than some certain size are removed. In our results, the solution of the optimization problem, where state constraint was approximated with explicit approximation, was nearer that structure. In addition, the optimization results were a little better then. However, when explicit approximation is used, the time step and class width have to be chosen so that a tree cannot grow over one size class during one time step [16]. We proved that only then is the explicit approximation scheme stable. For the implicit approximation scheme, we proved that it is unconditionally stable. Thus, in implicit approximation, the time step and class width can be chosen freely. In general, explicit approximation of the population model is more commonly used as a forest growth model [9,16].

Author Contributions: J.P. made the approximations and conducted the optimizations under supervision of E.L.; A.L. proved the theoretical results; J.P. and A.A. analyzed the numerical results; all authors contributed to the writing of the manuscript.

Acknowledgments: We want to acknowledge the Jenny and Antti Wihuri foundation for financial support.

Conflicts of Interest: The authors declare no conflict of interest.

Appendix A

We used the adjoint method to solve the optimization problem (14). For that, we needed the partial derivatives of the Lagrangian

$$\mathcal{L}(\mathbf{y}, \mathbf{h}, \lambda) = \bar{J}(\mathbf{y}, \mathbf{h}) - \sum_{k=1}^M (\lambda^k, A^k(\mathbf{y}, \mathbf{h})),$$

where $\bar{J}(\mathbf{h}) = \sum_{k=1}^M (\mathbf{d}^k, \mathbf{h}^k)$ is cost function of problem (14) and $A^k(\mathbf{y}, \mathbf{h})$ is constraint (8) or (12).

First, we calculate the partial derivatives of the cost function $\bar{J}(\mathbf{y}, \mathbf{h})$. Since it depends only on \mathbf{h} , obviously $\frac{\partial \bar{J}}{\partial \mathbf{y}} = 0$. The partial derivative of the cost function \bar{J} with respect to \mathbf{h} is

$$\frac{\partial \bar{J}(\mathbf{h})}{\partial \mathbf{h}} = \mathbf{d}.$$

Next, we calculate the partial derivatives of constraint function $A(\mathbf{y}, \mathbf{h})$. In both forms of A (constraints (8) or (12)), the partial derivative with respect to \mathbf{h} is $\frac{\partial A(\mathbf{y}, \mathbf{h})}{\partial \mathbf{h}} = \Delta t$.

Let us calculate the partial derivative of constraint function (8) (explicit approximation of the state Equation (4)) with respect to \mathbf{y}

$$\frac{\partial(\lambda, A(\mathbf{y}, \mathbf{h}))}{\partial \mathbf{y}} = \frac{\partial}{\partial \mathbf{y}} \sum_{k=1}^M (\lambda^k, \mathbf{y}^k - \mathbf{A}^{k-1} \mathbf{y}^{k-1} + \Delta t \mathbf{h}^k). \tag{A1}$$

Let us denote

$$\mathbf{H}^k := \mathbf{A}^k \mathbf{y}^k = \begin{bmatrix} a_{11}^k y_1^k \\ a_{21}^k y_2^k + b_{11}^k y_1^k \\ \vdots \\ a_{N1}^k y_N^k + b_{N-1,1}^k y_{N-1}^k \end{bmatrix}.$$

Then, the partial derivative (A1) can be written in the form

$$\frac{\partial(\lambda, A(\mathbf{y}, \mathbf{h}))}{\partial \mathbf{y}} = \frac{\partial}{\partial \mathbf{y}} \left(\sum_{k=1}^M (\lambda^k, \mathbf{y}^k - \mathbf{H}^{k-1}) \right).$$

By rearranging the terms and defining $\lambda^{M+1} = 0$, we get

$$\frac{\partial(\lambda, A(\mathbf{y}, \mathbf{h}))}{\partial \mathbf{y}} = \frac{\partial}{\partial \mathbf{y}} \left((\lambda^1, \mathbf{H}^0) + \sum_{k=1}^M ((\lambda^k, \mathbf{y}^k) - (\lambda^{k+1}, \mathbf{H}^k)) \right).$$

Now, $\frac{\partial \mathbf{H}^0}{\partial \mathbf{y}} = 0$ by definition of \mathbf{y}^0 and $\frac{\partial \mathbf{H}^k}{\partial \mathbf{y}} = \mathbf{A}^k + \mathbf{H}_1^k, k = 1, \dots, M - 1$, where

$$\mathbf{H}_1^k = \begin{bmatrix} a_{11}^k y_1^k & \dots & a_{1N}^k y_1^k \\ a_{21}^k y_2^k + b_{11}^k y_1^{k+1} & \dots & a_{2N}^k y_2^k + b_{1N}^k y_1^k \\ \vdots & \ddots & \vdots \\ a_{N1}^k y_N^k + b_{N-1,1}^k y_{N-1}^k & \dots & a_{NN}^k y_N^k + b_{N-1,N}^k y_{N-1}^k \end{bmatrix},$$

and

$$a_{ij}^k = \frac{\partial a_i^k}{\partial y_j^k} = -\frac{\Delta t}{\Delta x} (g_{21} + g_{22} x_i) \left(\frac{x_j}{2} \right)^2 \pi - \Delta t \left(m_{21} + \frac{m_{22}}{x_i} + \frac{m_{23}}{x_i^2} \right) \left(\frac{x_j}{2} \right)^2 \pi, \tag{A2}$$

$$b_{ij}^k = \frac{\partial b_i^k}{\partial y_j^k} = \frac{\Delta t}{\Delta x} (g_{21} + g_{22} x_i) \left(\frac{x_j}{2} \right)^2 \pi. \tag{A3}$$

Thus, we can define

$$\frac{\partial(\lambda, A(\mathbf{y}, \mathbf{h}))}{\partial \mathbf{y}} = (\mathbf{w}^1, \mathbf{w}^2, \dots, \mathbf{w}^M),$$

where

$$\begin{aligned} \mathbf{w}^k &= \lambda^k - ((\mathbf{A}^k)^* + (\mathbf{H}_1^k)^*) \lambda^{k+1}, \quad k = 1, \dots, M - 1 \\ \mathbf{w}^M &= \lambda^M. \end{aligned}$$

Then, we calculate the partial derivative of constraint (12) (implicit approximation of the state Equation (4)) with respect to \mathbf{y}

$$\frac{\partial(\lambda, A(\mathbf{y}, \mathbf{h}))}{\partial \mathbf{y}} = \frac{\partial}{\partial \mathbf{y}} \sum_{k=1}^M (\lambda^k, \mathbf{B}^{k-1} \mathbf{y}^k - \mathbf{y}^{k-1} + \Delta t \mathbf{h}^k). \tag{A4}$$

Let us denote

$$\mathbf{G}^k = \mathbf{B}^{k-1} \mathbf{y}^k = \begin{bmatrix} a_1^{k-1} y_1^k \\ a_2^{k-1} y_2^k - b_1^{k-1} y_1^k \\ \vdots \\ a_N^{k-1} y_N^k - b_{N-1}^{k-1} y_{N-1}^k \end{bmatrix}.$$

Then, the partial derivative (A4) can be written in the form

$$\frac{\partial(\lambda, A(\mathbf{y}, \mathbf{h}))}{\partial \mathbf{y}} = \frac{\partial}{\partial \mathbf{y}} \sum_{k=1}^M (\lambda^k, \mathbf{G}^k - \mathbf{y}^{k-1}).$$

Note that function \mathbf{G}^k depends on \mathbf{y}^{k-1} and \mathbf{y}^k so the partial derivative $\frac{\partial \mathbf{G}^k}{\partial \mathbf{y}} = \frac{\partial \mathbf{G}^k}{\partial \mathbf{y}^{k-1}} + \frac{\partial \mathbf{G}^k}{\partial \mathbf{y}^k}$.

By definition of \mathbf{y}^0 , derivative $\frac{\partial(\lambda^1, \mathbf{y}^0)}{\partial \mathbf{y}} = 0$ and derivative $\frac{\partial(\lambda^1, \mathbf{G}^1)}{\partial \mathbf{y}} = \frac{\partial(\lambda^1, \mathbf{G}^1)}{\partial \mathbf{y}^1}$. By rearranging the terms and defining $\lambda^{M+1} = 0$, we get

$$\frac{\partial(\lambda, A(\mathbf{y}, \mathbf{h}))}{\partial \mathbf{y}} = \sum_{k=1}^M \left(\frac{\partial(\lambda^k, \mathbf{G}^k)}{\partial \mathbf{y}^k} + \frac{\partial(\lambda^{k+1}, \mathbf{G}^{k+1} - \mathbf{y}^k)}{\partial \mathbf{y}^k} \right).$$

Derivative $\frac{\partial \mathbf{G}^k}{\partial \mathbf{y}^k} = \mathbf{B}^{k-1}$ and derivative

$$\frac{\partial \mathbf{G}^{k+1}}{\partial \mathbf{y}^k} = \begin{bmatrix} a_{11}^k y_1^{k+1} & \dots & a_{1N}^k y_1^{k+1} \\ a_{21}^k y_2^{k+1} - b_{11}^k y_1^{k+1} & \dots & a_{2N}^k y_2^{k+1} - b_{1N}^k y_1^{k+1} \\ \vdots & \ddots & \vdots \\ a_{N1}^k y_N^{k+1} - b_{N-1,1}^k y_{N-1}^{k+1} & \dots & a_{NN}^k y_N^{k+1} - b_{N-1,N}^k y_{N-1}^{k+1} \end{bmatrix},$$

where a_{ij}^k and b_{ij}^k are derivatives of coefficients a_i^k and b_i^k with respect to y_j^k defined in Equations (A2) and (A3), respectively. Thus, we can define

$$\frac{\partial(\lambda, A(\mathbf{y}, \mathbf{h}))}{\partial \mathbf{y}} = (\mathbf{q}^1, \mathbf{q}^2, \dots, \mathbf{q}^M),$$

where

$$\mathbf{q}^k = (\mathbf{B}^{k-1})^* \lambda^k + \left(\left(\frac{\partial \mathbf{G}^{k+1}}{\partial \mathbf{y}^k} \right)^* - \mathbf{1}_N \right) \lambda^{k+1}, \quad k = 1, \dots, M-1,$$

$$\mathbf{q}^M = (\mathbf{B}^{M-1})^* \lambda^M,$$

and $\mathbf{1}_N$ is $N \times N$ identity matrix.

References

1. Kato, N.; Torikata, H. Local existence for a general model of size-dependent population dynamics. *Abstr. Appl. Anal.* **1997**, *2*, 207–226. [\[CrossRef\]](#)

2. Kato, N. A general model of size-dependent population dynamics with nonlinear growth rate. *J. Math. Anal. Appl.* **2004**, *297*, 234–256. [[CrossRef](#)]
3. Calsina, A.; Saldana, J. A model of physiologically structured population dynamics with a nonlinear individual growth rate. *J. Math. Biol.* **1995**, *33*, 335–364. [[CrossRef](#)]
4. Calsina, A.; Saldana, J. Basic theory for a class of models of hierarchically structured population dynamics with distributed states in the recruitment. *Math. Model. Meth. Appl. Sci.* **2006**, *16*, 16951722. [[CrossRef](#)]
5. Ackleh, A.S.; Deng, K.; Hu, S.A. Quasilinear Hierarchical Size-Structured Model: Well-Posedness and Approximation. *Appl. Math. Optim.* **2005**, *51*, 3559. [[CrossRef](#)]
6. Hritonenko, N.; Yatsenko, Y.; Goetz, R.-U.; Xabadia, A. Maximum principle for size-structured model of forest and carbon sequestration management. *Appl. Math. Lett.* **2008**, *21*, 1090–1094 [[CrossRef](#)]
7. Hritonenko, N.; Yatsenko, Y.; Goetz, R.-U.; Xabadia, A. A bang-bang regime in optimal harvesting of size-structured populations. *Nonlinear Anal. Theory Methods Appl.* **2009**, *71*, e2331–e2336. [[CrossRef](#)]
8. Hritonenko, N.; Yatsenko, Y.; Goetz, R.-U.; Xabadia, A. Optimal harvesting in forestry: Steady-state analysis and climate change impact. *J. Biol. Dyn.* **2013**, *7*, 41–58. [[CrossRef](#)] [[PubMed](#)]
9. Liang, J.; Picard, N. Matrix model of Forest Dynamics: An Overview and Outlook. *For. Sci.* **2013**, *59*, 359–378. [[CrossRef](#)]
10. Rämö, J.; Tahvonen, O. Economics of harvesting uneven-aged forest stands in Fennoscandia. *Scand. J. For. Res.* **2014**, *29*, 777–792. [[CrossRef](#)]
11. Pyy, J.; Ahtikoski, A.; Laitinen, E.; Siipilehto, J. Introducing a non-stationary matrix model for stand-level optimization, an even aged Pine (*Pinus sylvestris* L.) stand in Finland. *Forests* **2017**, *8*, 163. [[CrossRef](#)]
12. Anita, S.; Ianneli, M.; Kim, M.-Y.; Park, E.-J. Optimal harvesting for periodic age-dependent population dynamics. *SIAM J. Appl. Math.* **1998**, *58*, 1648–1666.
13. Xie, Q.; He, Z.-R.; Wang, X. Optimal harvesting in diffusive population models with size random growth and distributed recruitment. *Electron. J. Differ. Eq.* **2016**, *214*, 1–13.
14. Ackleh, A.S.; Farkas, J.Z.; Li, X.; Ma, B. Finite difference approximations for a size-structured population model with distributed states in the recruitment. *J. Biol. Dyn.* **2015**, *9*, 2–31. [[CrossRef](#)] [[PubMed](#)]
15. Ackleh, A.S.; Ito, K. An implicit finite difference scheme for the nonlinear size-structured population model. *Numer. Func. Anal. Opt.* **1997**, *18*, 865–884. [[CrossRef](#)]
16. Picard, N.; Liang, J. Matrix models for size structured populations: Unrealistic fast growth or simply diffusion? *PLoS ONE* **2014**, *9*, e98254. [[CrossRef](#)] [[PubMed](#)]
17. Lions, J. *Nonhomogeneous Boundary Value Problems and Applications*; Springer Verlag: Berlin, German, 1972.
18. Anita, S.; Arnautu, V.; Stefanescu, R. Numerical optimal harvesting for periodic age-structured population dynamics with logistic term. *Numer. Func. Anal. Opt.* **2009**, *30*, 183–198. [[CrossRef](#)]
19. Kang, Y.H.; Lee, M.J.; Jung, I.H. Optimal Harvesting for an Age-Spatial-Structured Population Dynamic Model with External Mortality. *Abstr. Appl. Anal.* **2012**. [[CrossRef](#)]
20. Kim, Y.K.; Lee, M.J.; Jung, I.H. Duality in an Optimal Harvesting Problem by a Nonlinear Age-Spatial Structured Population Dynamic System. *KYUNGPOOK Math. J.* **2011**, *51*, 353–364. [[CrossRef](#)]
21. Ciarlet, P.G. *Introduction to Numerical Linear Algebra and Optimization*; Cambridge University Press: New York, United States, 1989.
22. Anita, S.; Arnautu, V.; Capasso, V. *An Introduction to Optimal Control Problems in Life Sciences and Economics: From Mathematical Models to Numerical Simulation with MATLAB*; Springer: Dordrecht, The Netherlands, 2011.
23. Kuuluvainen, T.; Tahvonen, O.; Aakala, T. Even-aged and uneven-aged forest management in boreal Fennoscandia: A review. *AMBIO* **2012**, *41*, 720–737. [[CrossRef](#)] [[PubMed](#)]
24. Tahvonen, O.; Pihlainen, S.; Niinimäki, S. On the economics of optimal timber production in boreal Scots pine stand. *Can. J. For. Res.* **2013**, *43*, 719–730. [[CrossRef](#)]

