



# Improving dynamic treatment unit forest planning with cellular automata heuristics

Wilhelmsson Pär<sup>1</sup> · Lämås Tomas<sup>1</sup> · Wallerman Jörgen<sup>1</sup> · Eggers Jeannette<sup>1</sup> · Öhman Karin<sup>1</sup>

Received: 22 November 2021 / Revised: 12 June 2022 / Accepted: 22 June 2022 / Published online: 3 August 2022  
© The Author(s) 2022

## Abstract

We present a model for conducting dynamic treatment unit (DTU) forest planning using a heuristic cellular automata (CA) approach. The clustering of DTUs is driven by entry costs associated with treatments, thus we directly model the economic incentive to cluster. The model is based on the work presented in the literature but enhanced by adding a third phase to the CA algorithm where DTUs are mapped in high detail. The model allows separate but nearby forest areas to be included in the same DTU and shares the entry cost if they are within a defined distance. The model is applied to a typical long-term forest planning problem for a 1 182 ha landscape in northern Sweden, represented by 4 218 microsegments with an average size of 0.28 ha. The added phase increased the utility by 1.5–32.2%. The model produced consistent solutions—more than half of all microsegments were managed with the same treatment program in 95% of all solutions when multiple solutions were found.

**Keywords** Entry cost · Forest planning · High-resolution data · Spatial optimization

## Introduction

Forest management planning aims at efficient and sustainable use of the forest resource over time, whether it be for economical, social, biological, or other purposes. In this pursuit, the concept of stands has been used in even-aged forest management planning for a very long time (see, e.g., af Ström 1822; Faustmann 1849; Nilsson et al. 2012). A stand is a delimited area where the forest is homogeneous in some regard, making the whole stand suitable for the same forest management. Stands are typically used as description units (DUs), which we define as the smallest unit for collection and storage of data, and modeling of ecosystem

processes. Moreover, stands also act as treatment units as it is the unit used for modeling and planning treatments which on holding level are aligned to fulfil stakeholder goals. An important property of the traditional stand approach is that the stand borders are usually fixed and permanent during the planning horizon in planning processes (Nelson and Brodie 1990; Davis et al. 2001). Furthermore, stands are generally large enough and delineated such that they may be treated individually; spatial clustering of stand management activities is not a prerequisite for practical forestry. This makes linear programming (LP) a suitable and powerful method for solving the resulting optimization problem in forest management planning using the stand approach (see, e.g., Johnson and Scheurman 1977).

The development of remote sensing techniques for data collection has changed and enhanced the conditions for forestry and forest planning (Maltamo et al. 2014). Objective wall-to-wall data with high spatial resolution based on combinations of remote sensing and field surveys are compiled nationwide in, e.g., Finland (Kotivuori et al. 2016) and Sweden (Nilsson et al. 2017). Even if such data may be useful, transforming fine-grained data into stand-wise data will result in loss of information and consequently suboptimal use of the forest resource (Holmgren and Thuresson 1997). An alternative approach to planning forestry based on the stand-wise approach is therefore to utilize dynamic

---

Communicated by Thomas Knoke.

✉ Wilhelmsson Pär  
par.wilhelmsson@slu.se  
Lämås Tomas  
tomas.lamas@slu.se  
Wallerman Jörgen  
jorgen.wallerman@slu.se  
Öhman Karin  
karin.ohman@slu.se

<sup>1</sup> Department of Forest Resource Management, Swedish University of Agricultural Sciences, Umeå, Sweden

treatment units (DTU). DTU planning aims at more efficient use of the forest resource by maintaining high spatial resolution in both the forest data and throughout the planning process. This is attempted by using DUs much smaller than traditional stands and forming DTUs in the planning process by clustering DUs into larger areas in the optimization. Treatments in specific DTUs do not necessarily reoccur. Thus, DTU are dynamic and exist only momentarily in time.

Previous research has studied the concept and performance of DTUs and presented various models for conducting this type of forest planning (e.g., Pascual et al. 2018; Heinonen et al. 2018; Wilhelmsson et al. 2021). The economic incentive to cluster treatments lies in the entry cost for forest operations. The entry cost is defined as the fixed cost associated with a contiguous cluster scheduled for treatment at a specific time point and includes costs for preparatory fieldwork, road maintenance, moving machinery or personnel to the site, and administrative work (Borges et al. 2017). Borges et al. (2017) showed that the entry cost influences the optimal treatment unit size, which has been a concern for past studies concerned with DTU. However, instead of directly modeling the entry cost, most DTU studies have used proxy variables such as conditional common borders in combination with distance from the nearest road (Pascual et al. 2018). Their model tracked the cut-to-cut and cut-to-uncut borders, and metrics were included in the utility function to drive the clustering of treatments.

Forest planning problems with a DTU approach have seldom been solved with exact solution methods (Wilhelmsson et al. 2021), but due to the combinatorial and complex nature of DTU planning, most often solved with heuristics. Pukkala (2009) showed that heuristic methods can successfully handle large spatial problems. This was shown for DTU purposes by Heinonen et al. (2007) and was further highlighted by Pascual et al. (2018). Noteworthy heuristics applied to DTU problems are threshold accepting (Heinonen et al. 2007), reduced cost (Heinonen et al. 2018; Packalén et al. 2011; Pukkala et al. 2009), genetic algorithm (Lu and Eriksson 2000) and simulated annealing (de Miguel and Pukkala 2013; Öhman 2001). One heuristic that has been argued to be particularly well suited for solving forest planning problems in general, and forest planning problems using a DTU approach in particular, is cellular automata (CA). CA was first presented by von Neumann (1966) and introduced in forest planning contexts by Strange et al. (2002). It has been used to solve forest planning problems with a DTU approach in several studies (Heinonen and Pukkala 2007; Mathey et al. 2007; Pascual et al. 2018; Pukkala 2019). In the general form, a CA is constituted of a lattice of cells where each cell may take on a finite set of states. Rules decide how cells may change states, which is dependent on a utility function and spatial relations between the given cell and a subset of the other cells (neighbors). In a DTU forest

planning context, a cell in the lattice represents a DU. A state is equivalent to a treatment program (TP), which is a sequence of treatments over the entire planning horizon and the resulting development of the forest. Thus, a CA algorithm forms DTUs by swapping TPs for segments of forest with regard to treatments planned in spatially nearby DUs, clustering treatments in space and time. Mathey et al. (2007) argued that two properties make CA particularly well suited for solving forest planning problems: (1) the way that different spatial scales can be integrated into a CA model, and (2) the way that landscape scale patterns emerge due to local spatial rules determining the change in DU state. While in-depth comparisons with other heuristics are lacking, previous studies have shown potential in CA compared to simulated annealing in computational speed, the number of iterations needed when using stop criteria, and solution quality (Mathey et al. 2007; Heinonen and Pukkala 2007).

This study aims to improve long-term forest planning with a DTU approach using CA by introducing explicit entry costs. The first and second phase of the algorithm is inspired by the literature. However, a third phase is added to the CA algorithm where the economic incentive to cluster treatments is modeled directly by calculating entry costs in high detail for potential DTUs. The approach is applied in a planning problem for a 1 192 ha landscape in northern Sweden where the forest is described with high spatial resolution data using DUs much smaller than the area of traditional stands.

## Method

### Overview of the cellular automata algorithm

Our approach is based on the CA model presented by Strange et al. (2002), which was improved by Mathey et al. (2007) and Heinonen and Pukkala (2007). The CA evaluated here is a set of DUs representing a forest. The states are (DU specific) TPs. The CA algorithm consists of three phases: local, global and final, each with a number of iterations. The algorithm starts by randomly selecting a TP for all DUs. All phases follow a similar procedure. In each iteration, a random number is drawn from a uniform distribution between 0 and 1 for each DU. Depending on the number, one of three things occur in accordance to probabilities set by the user. The DU is either mutated, innovated, or remain unchanged. A DU being unchanged means that the current TP remains the same. If a mutation occurs, the TP is swapped for a randomly selected one. If innovation occurs, all potential TPs for the DU are evaluated, and the model swaps to the best TP with regard to a utility function stated by the user. DUs are processed in this manner one at a time until all DUs have been processed, at which point the iteration is completed

and the next iteration begins by starting over from the first DU. When the predefined number of iterations is done, the phase is completed and the next phase is initiated by starting over from the first DU, meaning that the output plan from a phase acts as starting point for the next phase. What separates the phases from each other is the utility function. This is elaborated below. All DUs are processed until the last DU of the last iteration of the last phase is processed, at which point all DU are innovated one last time. When this is done, the algorithm is complete and the solution exported. A conceptual visualization of the algorithm is shown in Fig. 1.

An essential component of the utility function used here is the net present value (NPV) of a TP, which is calculated by discounting income and costs from different management activities, both within the planning horizon and beyond for an infinite time horizon (Faustmann 1849). In addition to the productivity-based costs for treating a DU, the fixed entry cost is applied to all DTUs within the planning horizon and included in the calculation of NPV. The calculation of entry cost is one of two properties that changes over the phases of the CA algorithm (see Table 1), thus affecting the NPV. A simplified entry cost calculation is conducted in the local and global phases, whereas a higher detail mapping of the DTUs and calculation of entry cost are carried out in the final phase. The other property that changes over the phases is the inclusion of a harvest coefficient. This coefficient is excluded in the local phase but included in the global and final phases. The purpose of the harvest coefficient is to prevent the model from overharvesting, which is a concern when spatiotemporal clustering is beneficial, and the use of NPV may drive the model to harvest large volumes in early periods.

The mapping of DTUs is based on the concept of neighbors. DUs are considered neighbors if the minimum distance between their edges is below a specified distance (neighborhood distance), i.e., neighbors are not necessarily immediately adjacent. In the local phase, the fixed entry cost is calculated using a local scope (see Fig. 2). The entry cost used here is scaled down from a realistic value by multiplying it by 0.02. The local phase is followed by the global phase, which uses the output plan from the local phase as input. Here, the same local scope is used for calculating entry cost, but the utility function is changed to account for harvested volume over time by introducing the harvest coefficient (Fig. 3). coefficient is period-specific and punishes solutions with harvest volumes larger than the harvest target. The resulting plan from the global phase is used as input in the final phase. The final phase, introduced in the current study, maintains consideration to forest level goals using the harvest coefficient but the calculation of NPV changes, as the DTUs are fully mapped in each iteration (see Fig. 4). The complete mapping means that the potential number of

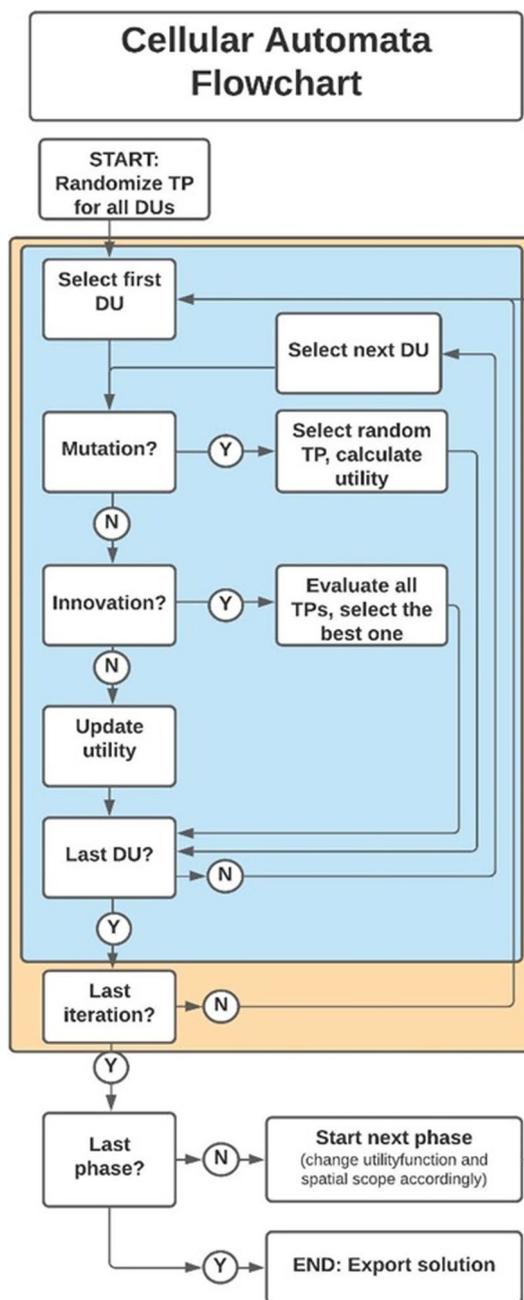
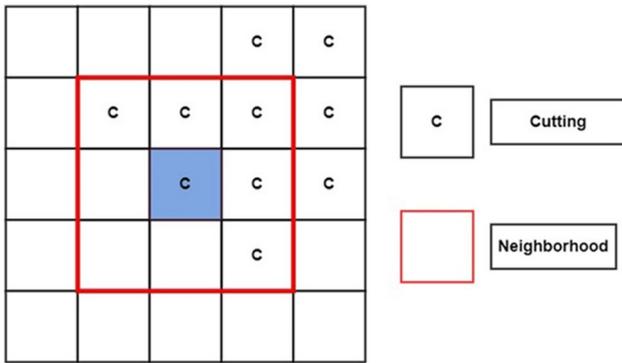


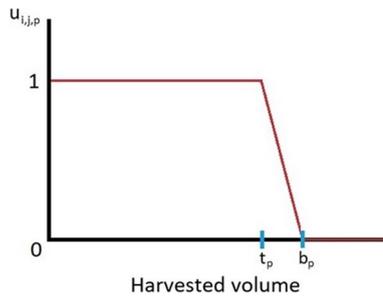
Fig. 1 Conceptual flowchart of the cellular automata approach used in the study

Table 1 How input, output, spatial scope and forest level goals change over the phases over the CA algorithm

Phase	Local	Global	Final
Input	Random start	Plan L	Plan G
Entry cost	Simplified	Simplified	High detail
Harvest coefficient	Not included	Included	Included
Output	Plan L	Plan G	Plan F



**Fig. 2** Consider the lattice a representation of a forest. Treatments in period  $p$  is shown, where  $c$  represents cutting. The red square marks the neighborhood of the blue DU. In the local and global phases,  $d_{i,j,p}$  represents the number of DUs in the neighborhood of DU  $i$  (blue) prescribed for the same treatment in period  $p$  of treatment program  $j$  as the blue DU (including the centering DU itself, here 5). The entry cost  $n_{i,j,p}$  is shared equally among the DUs marked with cutting within the neighborhood (see Eq. 3)

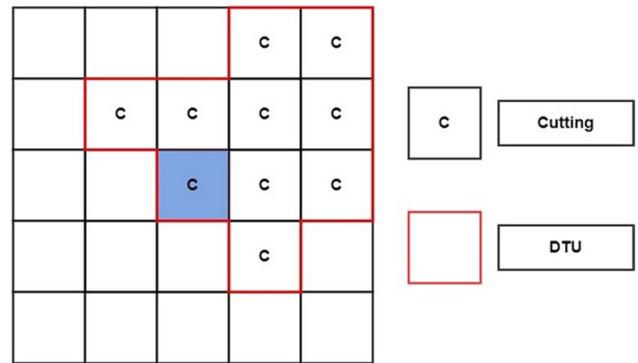


**Fig. 3** Graphical representation of the relation between  $u_{i,j,p}$ ,  $t_p$  and  $b_p$ . The utility coefficient  $u_{i,j,p}$  will have a value of 1 between 0 and  $t_p$   $m^3$  harvested volume, at which point it linearly declines toward 0, which it reaches at  $b_p$   $m^3$  harvested volume. For harvests levels even higher,  $u_{i,j,p}$  will have a value of 0, considering attempted contributions to the utility function from such harvests as worthless

DUs that constitute the DTU and share the entry cost is increased manifold. The entry cost is therefore increased to a realistic value to model the true costs as accurately as possible and let that cost drive the clustering. Each phase ends after the predefined number of iterations is completed. In one final iteration, all DUs are innovated in order to remove isolated and small DTUs, which may be the result of chance via mutation. The resulting plan is exported as the solution.

**Phases and the spatial rule**

Here follows a model formulation for the cellular automata as it progresses over the three phases, as shown in Fig. 1.



**Fig. 4** Consider the lattice a representation of a forest. The lattice shows the treatments in a time period  $p$ , where “C” represents cutting. The red shape shows the DTU. In the final phase, the variable  $D$  represents the total area of DUs included in the DTU (red area) and  $a_i$  represents the area of DU  $i$  (blue). The entry cost  $n_{i,j,p}$  is distributed among the DUs constituting the DTU in proportion to their area (see Eq. 5)

**Local phase: forming simple DTUs**

All phases aim to maximize the value of the utility function. In the local phase, it is defined as follows:

$$\text{Max}Z = \sum_{i=1}^I \sum_{j=1}^{J_i} z_{i,j} * x_{i,j} \tag{1}$$

$$z_{i,j} = \sum_{p=1}^P \frac{g_{i,j,p} - (c_{i,j,p} + n_{i,j,p})}{(1 + r)^{m_p}} + v_{i,j} \tag{2}$$

where  $x_{i,j} = \{0,1\}$ , the share of DU  $i$  managed with TP  $j$ ;  $m_p$  = midperiod year of period  $p$ ;  $g_{i,j,p}$  = gross revenue in period  $p$  of TP  $j$  in DU  $i$ ;  $c_{i,j,p}$  = spatially independent cost in period  $p$  of TP  $j$  in DU  $i$ ;  $n_{i,j,p}$  = entry cost in period  $p$  of TP  $j$  in DU  $i$ ;  $r$  = discount rate;  $v_{i,j}$  = terminal value; the discounted value of the forest management beyond the last planning period;  $I$  = set of DUs;  $J_i$  = set of TPs for DU  $i$ .

The definition of  $n_{i,j,p}$  changes over the phases. In the local and global phases, it is defined as

$$n_{i,j,p} = \left( \frac{e}{d_{i,j,p}} \right) \text{ if DU } i \text{ has a treatment in period } p, \text{ TP } j, \text{ otherwise } 0 \tag{3}$$

where  $e$  = fixed entry cost;  $d_{i,j,p}$  = number of DUs in DU  $i$ 's neighborhood treated with the same treatment in period  $p$  as DU  $i$  in TP  $j$  in period  $p$ .

The variable  $d_{i,j,p}$  is visualized in Fig. 2. This is what drives clustering in the local and global phase of the CA.

## Global phase: introduction of forest level goals

Optimizing the utility function with a local scope does not always lead to solutions that satisfy forest level goals and constraints. Hence, Mathey et al. (2007) introduced the inclusion of forest level goals in a CA model. The present study deals with the flow of harvested volumes over time. In the global phase, the utility function is therefore defined as follows:

$$z_{i,j} = \sum_{p=1}^P \frac{g_{i,j,p} - (c_{i,j,p} + n_{i,j,p})}{(1+r)^{m_p}} * u_{i,j,p} + v_{i,j} * w_{i,j} \quad (4)$$

where  $z_{i,j}$  = utility of TP  $j$  for DU  $i$ ;  $g_{i,j,p}$  = gross income in period  $p$  of TP  $j$  in DU  $i$ ;  $c_{i,j,p}$  = spatially independent cost in period  $p$  of TP  $j$  in DU  $i$ ;  $n_{i,j,p}$  = entry cost in period  $p$  of TP  $j$  in DU  $i$ ;  $r$  = discount rate;  $v_{i,j}$  = present value of net revenue from beyond the last planning period ( $P$ ) to infinity of TP  $j$  for DU  $i$ ;  $t_p$  = harvest target in period  $p$ ;  $b_p$  = upper harvest bound in period  $p$ ;  $u_{i,j,p}$  = utility coefficient in period  $p$  of TP  $j$  of DU  $i$ . The parameter takes the value

$$\begin{cases} 1 & \text{if } h_p \leq t_p, \\ 1 - (h_p - t_p) / (t_p - b_p) & \text{if } t_p < h_p \leq b_p, \text{ and} \\ 0 & \text{if } b_p < h_p, \end{cases}$$

where  $h_p$  is harvested volume in period  $p$ ;  $w_{i,j} = \min\{u_{i,j,p}\}$  of any period  $p$  for the given DU  $i$  and TP  $j$ ;  $m_p$  = midperiod year of period  $p$ .

The inclusion of the harvest coefficient,  $u_{i,j,p}$ , in the model means that the utility of a TP is lowered if that TP results in overharvesting within the planning period. Thus, while there are no constraints (besides the decision variable  $x_{i,j}$  being binary), the inclusion of the harvest coefficient will effectively constrain harvest levels. The parameter  $w_{i,j}$  is included since the harvest in each period is connected to the net revenue in each period using the harvest coefficient ( $u_{i,j,p}$ ), when meantime, the NPV from beyond the planning horizon is not connected to any harvest level.  $w_{i,j}$  prevents the model from selecting solutions leading to overharvest within the planning horizon in the search for NPV from beyond the planning horizon (terminal value  $v_{i,j}$ ). In such solutions,  $w_{i,j}$  will assume the value of 0, which is multiplied with the NPV from beyond the planning horizon. If all TPs available in a DU have the utility of 0, the model will choose the TP without treatments, and overharvest is avoided.

The entry cost is estimated in the global phase using the same spatial scope as the local phase (see Fig. 2).

## Final phase: high detail mapping of DTUs

The final phase of the algorithm will maintain the consideration of harvested volumes introduced in the global phase as defined by Eq. 4. What changes is the spatial scope of

the calculation of entry cost. Instead of looking only at the window defined by the neighborhood distance around the DU, the model checks if simultaneous treatments are found in a neighboring DU and if so, looks onward into the neighbor's neighbors and so on (into infinity, in theory). Thus, complete mapping of the DTU is conducted, and the entry cost is divided among the DUs constituting the DTU. Stepping stone effects may appear as the model allows separate (non-adjacent) areas to be part of the same DTU as long as all subareas are connected to another subarea according to the definition of neighbor. This change in the definition of entry cost is stated in Eq. 5 and visualized in Fig. 4. Each DU is charged with a share of the entry cost in proportion to its share of the total area of the DTU.

$$n_{i,j,p} = \left( \frac{E}{\frac{a_i}{D}} \right) \quad (5)$$

if DU  $i$  has a treatment in period  $p$  in TP  $j$ , otherwise 0.

where  $E$  = constant representing the entry cost;  $a_i$  = area of DU  $i$ ;  $D$  = area of the DTU, which DU  $i$  is part of in period  $p$  for TP  $j$  (according to principle shown in Fig. 4).

Note that, the value of  $D$  may be high, as the model maps the full area constituting the DTU.

## Analysis area, segmentation of DUs, and simulation of treatment programs

The model was evaluated by developing long term plans for a forest of 1192 ha located northwest of Sundsvall, Sweden. The forest is represented by 4218 microsegments, where each microsegment represents a DU with an average size of 0.28 ha. The microsegments are derived from remote sensing data in a two-step process. First, a region growing expansion model (Grilli et al. 2017) merges adjacent and similar  $12.5 \times 12.5 \text{ m}^2$  raster elements into possible DUs. Similarity is measured as distance in a 5D space with basal area, Lorey's mean height, proportion of pine, proportion of spruce, and proportion of broadleaves as dimensions. Merging is repeated until the smallest difference between neighboring microsegments is higher than a user defined level. This does not limit the size of segments, however, and they may become very large. Therefore, the second step is carried out using mixed integer programming (MIP). The MIP model selects microsegments from a large set of possible ones generated with the region growing method. The goal function of the MIP model is to minimize the sum of standard deviation within segments. The constraints are (1) each raster element must be assigned to a microsegment and (2) the maximum size of microsegments must not exceed a user defined limit. The MIP solution is the final spatial layout for microsegments (DUs). The wall-to-wall ALS data

used for the segmentation were provided by the Swedish Forest Agency and were described by Nilsson et al. (2017).

**Initial state**

The analysis area forest is comprised mainly of Norway spruce (*Picea abies*, 49% of the standing volume), but also of Scots pine (*Pinus sylvestris*, 30%) and birch (*Betula pubescens* and *Betula pendula*, 19%). The mean productivity is 4.9 m<sup>3</sup> ha<sup>-1</sup> year<sup>-1</sup> in the initial state, and the mean age is 58 years. The distribution between age classes is shown in Fig. 5. The case study area was chosen from a larger area, and a reasonably even age distribution was sought after.

**Generation of TPs**

The generation of TPs for each DU over the 50 years planning horizon was conducted with the PlanWise application in the Heureka decision support system (Wikström et al. 2011). Associated with each TP is a projected state of the DU in each period, including, e.g., growing stock and harvested volume. The core of the Heureka system is a collection of empirical models for projecting stand dynamics, e.g., growth, mortality and yield, in 5 year time steps. Heureka PlanWise generates a set of TPs for each DU within a

user-defined forest management framework. The management system was set to even-aged silviculture for all DUs, and the TPs differ with regard to the timing and manner of soil preparation, planting, pre-commercial thinning, thinnings, and final felling. A TP without any treatments was also generated for all DUs. On average, 12.7 potential TPs were generated for each of the 4218 DUs, that is, a total of 53 473 TPs.

**Model settings and hardware**

The cellular automata algorithm for solving the management problem was written in the programming language Python v3.8 using the IDE IntelliJ Pycharm v2019.3.3. The model was run on a Intel Core i7 2.8 GHz computer with 32 GB RAM and 64-bit Windows 10 as operating system. The model was run using three different neighborhood distances; 1 m, 50 m, and 200 m. These are henceforth called “cases.” Because of the stochastic nature of the cellular automata algorithm, where optimal solutions are not guaranteed, the solving procedure was repeated 40 times for the 1 m and 50 m neighborhood cases and 20 times for the 200 m neighborhood case. These sets of solutions are henceforth called “analyses.”

All cases and analyses used the following parameter settings. Each phase included 50 iterations. The harvest target, i.e.,  $t_p$ , was set to 50,000 m<sup>3</sup> and the harvest bound, i.e.,  $b_p$ , was set to 55 000 m<sup>3</sup>. The probability of mutation and innovation was set to 5% and 90%, respectively, leaving a 5% probability that the TP for a given DU remains unchanged in an iteration. A discount rate of 3.0% was used for calculating the NPV, i.e.,  $r$  was set to 0.03.

**Model stability**

There is an element of chance in the solutions produced since mutation means that a TP is randomly drawn from the set of available ones for a given DU. We investigated the impact of said chance on final plans by studying how consistently the model selected TPs for each DU when the model is run repeatedly. We call the consistency in TP selection “stability” and Table 2 shows how this was investigated.

We describe model stability by reporting descriptive statistics from the column named Stability in Table 6.

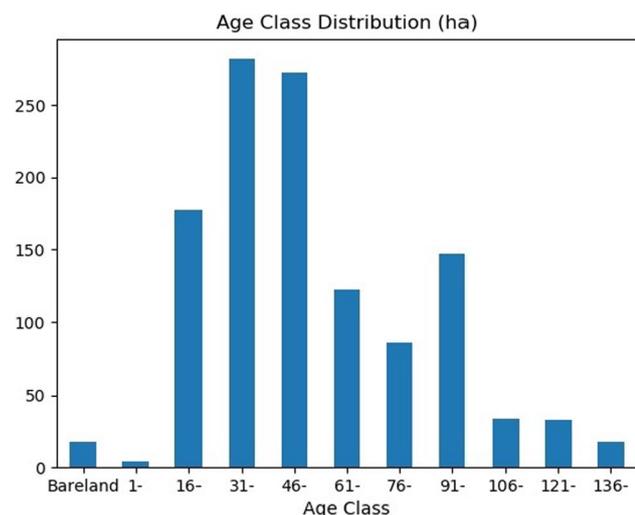


Fig. 5 Initial age class distribution of the analysis area

**Table 2** Example of how stability is measured for three DUs given six repetitions and the selected TPs

DU	Repetition						Most frequent TP	Stability (%)
	1	2	3	4	5	6		
1	TP <sub>12</sub>	TP <sub>12</sub>	TP <sub>13</sub>	TP <sub>12</sub>	TP <sub>13</sub>	TP <sub>12</sub>	TP <sub>12</sub>	67
2	TP <sub>25</sub>	TP <sub>23</sub>	TP <sub>24</sub>	TP <sub>23</sub>	TP <sub>21</sub>	TP <sub>24</sub>	TP <sub>23</sub> , TP <sub>24</sub>	33
3	TP <sub>34</sub>	TP <sub>31</sub>	TP <sub>34</sub>	83				
							Average	61

### Stand-based planning with linear programming

The CA solutions were compared with the outcome of long-term planning problems based on the traditional stand approach and solved by LP (Johnson and Scheurman 1977). The same DTUs were used in the LP planning problem. Maximum NPV was used as objective function, and the LP problem was solved using the optimization tool available in Heureka PlanWise. Four different solutions were produced using different constraints on harvested volume. One solution had no constraints, and the other three were forced to follow the period-specific harvest profile of each of the three corresponding solutions found with the CA. Note that, the LP solutions do not include entry fixed costs in the manner that the CA does, therefore no clustering is conducted here.

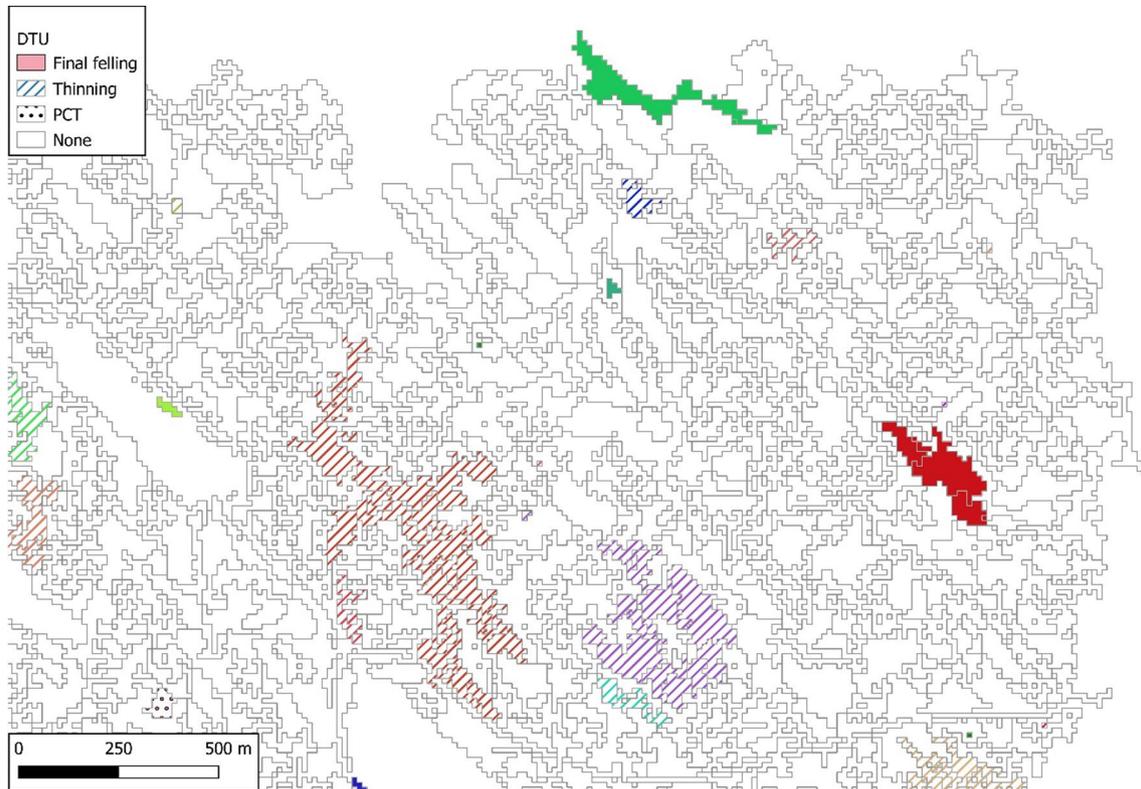
### Results

#### Cases

Table 3 show how utility and NPV increased with higher neighborhood distance. Harvested volume followed the same pattern. The different cases also resulted in different spatial layouts of treatments. The final plans from each case are visualized by showing the treatments and DTUs in the northeast corner of the analysis area in period 2 (Figs. 6, 7, 8). With an increase in the neighborhood distance, the DTUs became larger and fewer but were also more dispersed over the landscape. An example of this can be seen when comparing Figs. 6, 7, 8: The striped

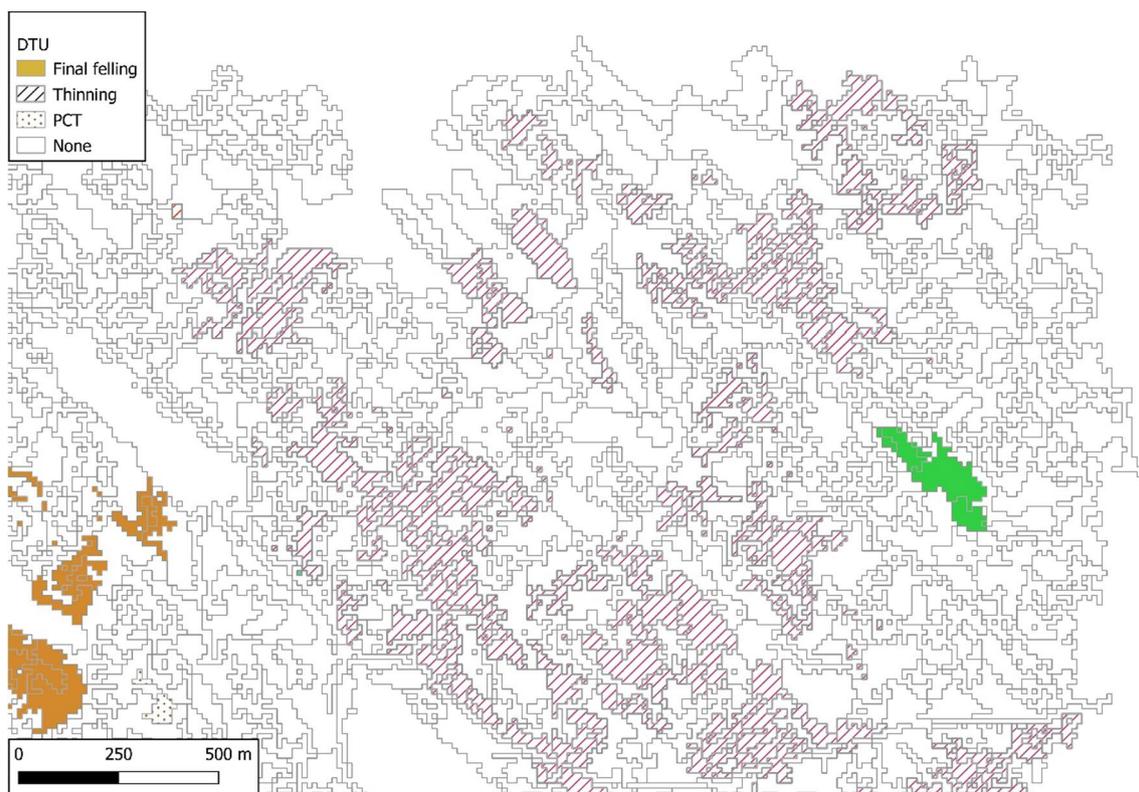
**Table 3** Summary of results for the three cases

Neighborhood distance (m)	Utility (million)	NPV (million SEK)	Average DTU size (ha)	No. of DTUs per period	Harvested volume (m <sup>3</sup> ha <sup>-1</sup> year <sup>-1</sup> )
1	50.3	50.9	2.04	49.6	7.4
50	51.2	53.4	5.90	19.4	8.1
200	53.9	54.4	12.7	9.8	8.3



**Fig. 6** Spatial layout of DTUs in the second five-year period in the northeast corner of the analysis area, using 1 m neighborhood distance. Filled DTUs are scheduled for final felling, striped DTUs are

scheduled for thinning and dotted DTUs are scheduled for pre-commercial thinning (PCT). Individual DTUs are colored the same and non-marked polygons are micro segments (DTUs) or non-forest



**Fig. 7** Spatial layout of DTUs in the second five-year period in the northeast corner of the analysis area, using 50 m neighborhood distance. Filled DTUs are scheduled for final felling, striped DTUs are

scheduled for thinning and dotted DTUs are scheduled for pre-commercial thinning (PCT). Individual DTUs are colored the same and non-marked polygons are micro-segments (DTUs) or non-forest

thinnings in the middle of Fig. 6 are considered separate DTUs when using 1 m neighborhood distance (Fig. 6), whereas the cases with 50 and 200 m, respectively, mapped this area as a single, but dispersed DTU (Figs. 7 and 8). Small, isolated DTUs did occur in all cases (see Fig. 6).

### Analysis of the cellular automata algorithm

The final phase contributed to both utility and NPV (see Table 4). Utility increased with 1.5–32.2% and NPV increased with 3.6–33.8% depending on neighborhood distance. The coefficient of variation (standard deviation divided by the mean) varied with neighborhood distance but was 0.0161 or lower for the utility function and 0.0013 or lower for the NPV.

Utility and NPV increased in only the few first iterations of each phase, see Figs. 9, 10, 11, 12. The overall trends were the same for all analyses. Improvements in later iterations were very small. The last iteration of the final phase (innovation of all DTUs) made noticeable contributions to both the NPV and the utility function.

Local phase solutions showed clear overharvest in the first period (see Fig. 12, line P1) and a shortage of harvests in the following couple of periods. The introduction of

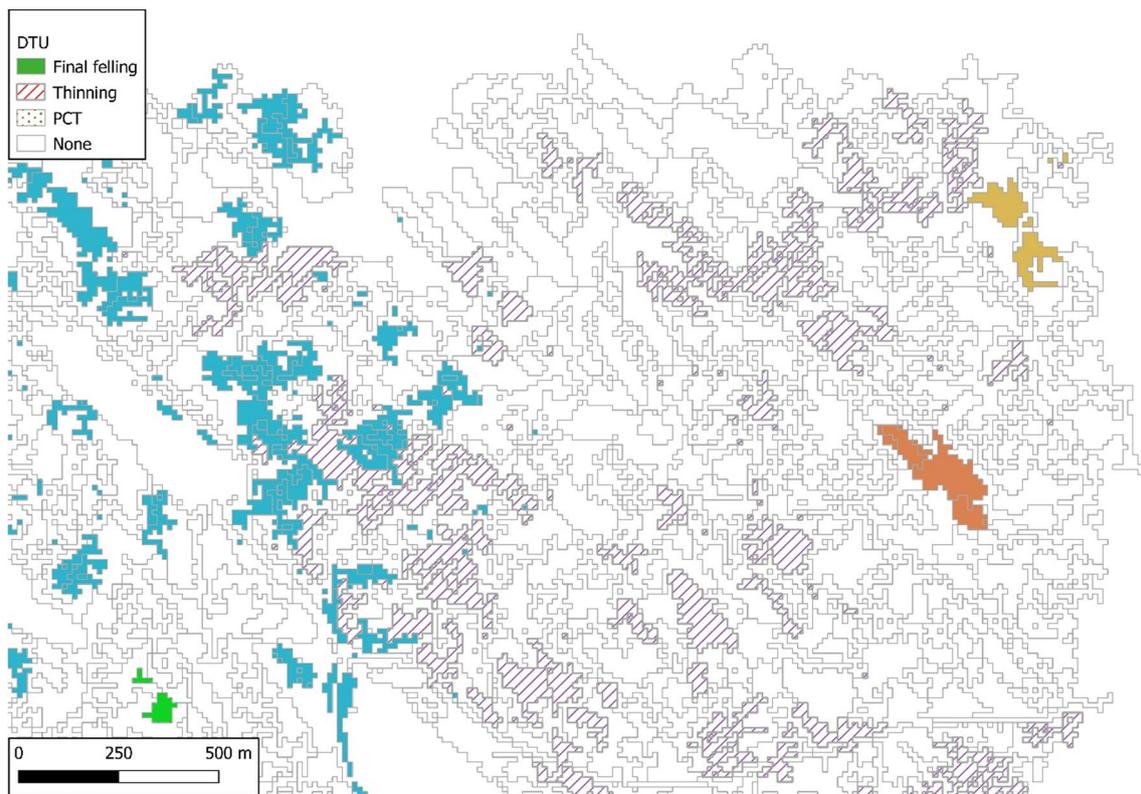
the harvest coefficient in the global phase mitigated this, but the final phase further decreased harvests, i.e., after introducing full entry costs, in some periods (see Fig. 13, lines P2 and P4).

### Model runtime

Runtimes increased with longer neighborhood distances, which is logical since the computational burden increases for our model when mapping DTUs that are large in terms of the number of included DTUs. The final phase was computationally time demanding, especially so when using 50 and 200 m neighborhoods. The model needed 800–6500 s per plan (repetition) depending on neighborhood distance (Table 5).

### Model stability

We investigated the stability of each analysis. On average, the model selected the same TP in ~87% of solutions for any neighborhood distance (see Table 6).



**Fig. 8** Spatial layout of DTUs in the second five-year period in the northeast corner of the analysis area, using 200 m neighborhood distance. Filled DTUs are scheduled for final felling, striped DTUs are

scheduled for thinning and dotted DTUs are scheduled for pre-commercial thinning (PCT). Individual DTUs are colored the same and non-marked polygons are micro-segments (DTUs) or non-forest

**Table 4** Utility and NPV in the analyses (20–40 solutions). Delta represents changes in solutions achieved by the final phase

Neighborhood distance (m)	Utility or NPV (SEK)	Global phase		Final phase		
		Average value	Average value	Delta	SD	Coeff of variation
1 <sup>1</sup>	Utility	37.3 M	49.9 M	32.2%	804 936	0.0161
	NPV	38.5 M	51.0 M	33.8%	65 129	0.0013
50 <sup>1</sup>	Utility	47.5 M	53.1 M	11.7%	156 174	0.0029
	NPV	48.9 M	53.5 M	9.2%	60 844	0.0011
200 <sup>2</sup>	Utility	51.0 M	51.7 M	1.5%	611,017	0.0118
	NPV	5.25 M	54.4 M	3.6%	45 424	0.0008

<sup>1</sup>40 solutions; <sup>2</sup>20 solutions

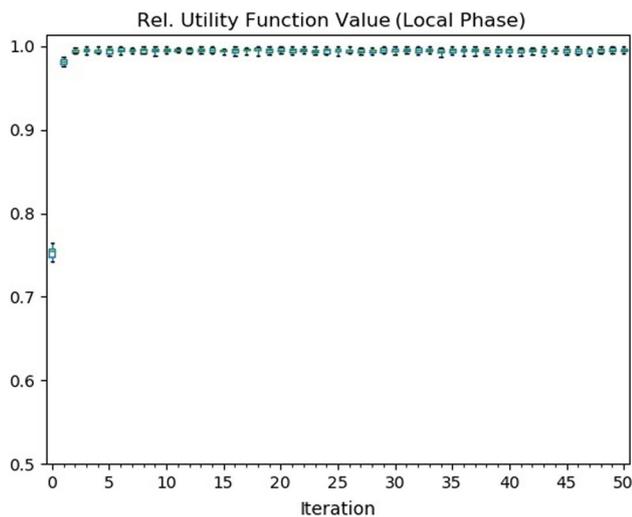
### Comparison with linear programming

Table 7 shows the NPV and harvested volume when using the heuristic and a traditional LP approach to solve similar planning problems. The relative NPV in solutions found by the CA algorithm (including entry costs) were within 91.4–97.5% of the theoretical maximum found using LP (including neither entry costs nor harvest constraints). Three LP solutions were also found including a constraint stating that the harvest profile had to match the harvest profile from

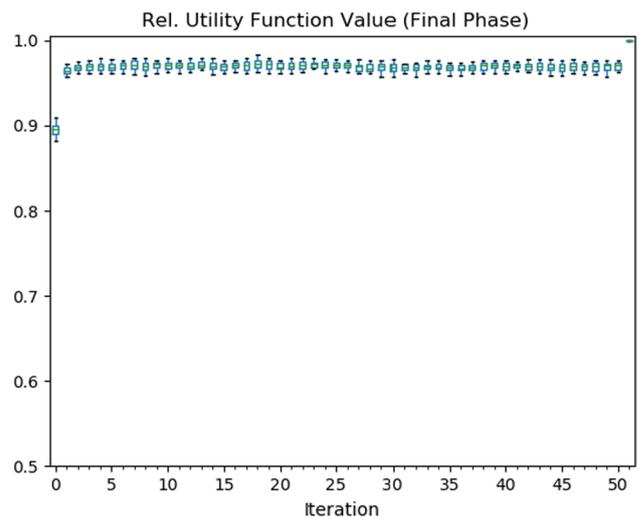
each CA solution. When comparing solutions produced by CA to these LP solutions, the relative NPV from CA then rose to 94.3–98.1%.

### Discussion

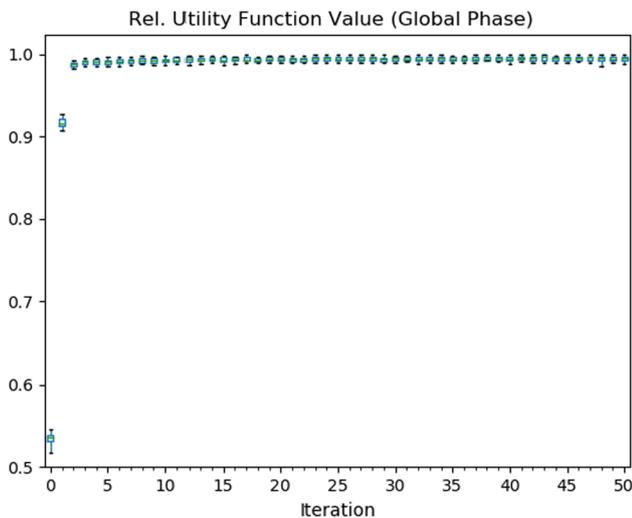
This study is an attempt to improve forest planning with a DTU approach using a cellular automata algorithm. Clustering treatments is necessary for forest management



**Fig. 9** Utility function values in the local phase, relative to the maximum value in the displayed dataset, for the 50 m neighborhood distance case. Boxplots representing results from the 40 repetitions of the CA algorithm. The utility function changes over phases, hence values are not comparable between Figs. 9, 10, and 11. Iteration 0 showing data from the randomly selected initial plan



**Fig. 11** Utility function values in the final phase, relative to the maximum value in the displayed dataset, for the 50 m neighborhood distance case. Boxplots representing results from the 40 repetitions of the CA algorithm. Utility function changes over the phases, hence values are not comparable between Figs. 9, 10, and 11. Iteration 0 showing the solutions produced by the global phase as evaluated by the final utility function



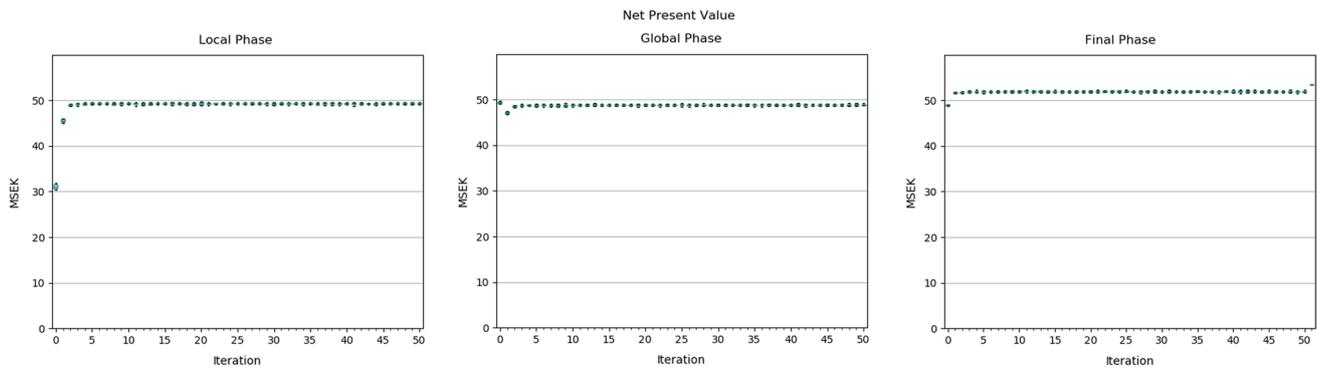
**Fig. 10** Utility function values in the global phase, relative to the maximum value in the displayed dataset, for the 50 m neighborhood distance case. Boxplots representing results from the 40 repetitions of the CA algorithm. Utility function changes over phases, hence values are not comparable between Figs. 9, 10, and 11. Iteration 0 showing the solutions produced by the local phase as evaluated by the global utility function

planning when using DUs of high spatial resolution (Heinonen et al. 2018). From an economic standpoint, clustering is needed due to the fixed costs associated with treatments, e.g., for moving harvest machinery. In contrast to comparable studies in the literature, this study explicitly modeled the economic incentive to cluster harvest

activities, in a high-detail manner in the final phase of the algorithm.

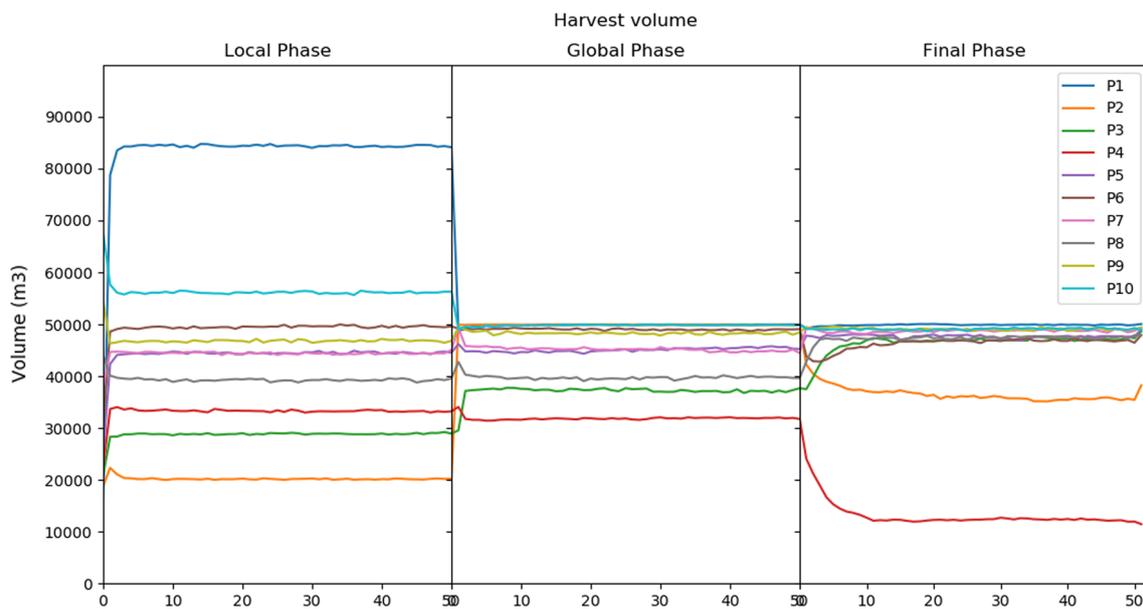
The model successfully solved the planning problem using different parameter settings and neighborhood distances. The model succeeded in the sense that (i) the plans had clustered treatments (ii) overharvesting did not occur, and (iii) all CA-produced plans (including entry cost) were within 5.7% of the comparable plans found with LP (excluding entry costs). The final phase improved the solutions in terms of both utility function value (1.5–32.2%) and NPV (3.8–32.4%), see Table 4. The model was successful in preventing overharvests compared to the harvest level set by the harvest coefficient. Thus, the effect of including the harvest coefficient was similar to inclusion of a constraint limiting harvest. However, the model did not result in even harvest flow in all periods (Fig. 13), which may be a concern for some forest owners. This may be achieved by decreasing the target volume (variable  $t_p$ ), possibly in combination with increasing the upper bound ( $b_p$ ).

A high number of iterations in each phase do not seem to be necessary since the culmination of NPV and utility function values were observed after only a few iterations. Past studies using cellular automata models have had utility culminate later during the iterations (e.g., Pascual et al. 2019), possibly because those models shifted utility weighting linearly over the iterations, whereas our model changes the utility function abruptly when a new phase starts. The final phase was the most time-consuming. The presented model solved the planning problem for 1 192 ha (4218 DUs) in



**Fig. 12** NPV over all three phases for the 50 m neighborhood distance analysis. Boxplots representing results from 40 repetitions of the CA algorithm. Iteration 0 in the local phase graph marks the NPV

of the randomly selected initial plans. Iteration 0 in the global and final phase marks the NPV of the input solution in each phase



**Fig. 13** Harvest profiles for the 50 m neighborhood distance case. Average values for 40 solutions. Global phase introduces the harvest coefficient and the final phase charges each DTU by the full entry

cost. Iteration number within each phase on the x axis. P1 – P10 are the ten five year periods over the planning horizon

**Table 5** Average values of runtime for different neighborhood distances

Neighborhood distance (m)	Runtime per iteration			Iterations per phase	Repetitions	Avg runtime per repetition (s)
	Local (s)	Global (s)	Final (s)			
1	4.3	4.4	7.9	50	40	830
50	4.4	4.6	29.6	50	40	1855
200	4.5	4.5	121.2	50	20	6510

830–6510 s (Table 5), depending on neighborhood distance. Unfortunately, computation time is not always reported in the literature. Comparisons are possible with the reduced cost model of Heinonen et al. (2018). Their model solved DTU planning problems for a 13 805 ha area (42,606 DUs)

with solution times of ~2000s (Table 5). Note that, the complexity and computational cost is a result not only of analysis area size and spatial resolution, but also temporal solution, planning horizon, number of treatment units, neighborhood distance, etc. If short solution times is of great

**Table 6** For each DU, we investigated which TP occurred the most often in the analyses. The frequency (in percent) of this TP results in a value between 0 and 100 for each DU. This frequency describes the stability in TP assignment, since a value of 100 for a DU shows that the model selected the same TP in 100% of all repetitions for that DU. The table shows descriptive statistics for this frequency for all DUs of the three analyses

Percentile	Stability (%)		
	1 m <sup>1</sup>	50 m <sup>1</sup>	200 m <sup>2</sup>
Min	25	23	25
Median	95	95	95
Max	100	100	100
Average	87	87	87

<sup>1</sup>Average values over 40 repetitions; <sup>2</sup>Average values over 20 repetitions

importance, our model may be improved by adding stopping criteria based on specific utility function values, possibly allowing for finding good solutions in fewer iterations and hence shorter solution times. Regarding runtimes, it has to be stressed that code was not optimized for solution time.

Heuristic models may not guarantee optimal solutions and solutions produced are to some degree the product of chance. Bettinger et al. (2009) suggest a six-level framework for validation of forest planning heuristics, where level six is the highest. Corresponding to level two in this piece of literature, we solve problems repeatedly and report spread in utility function value and in NPV. We also report the model's tendency to select the same TP for an individual DU over a set of solutions. We repeated the solving procedure 40 times for the 1 m and 50 m neighborhood distance and 20 times for the 200 m ditto. The coefficient of variation for the utility function value was 0.0161 or lower for all analyses and for NPV the same was true for a value of 0.0013. The stability in TP selection was very similar in the different analyses

regardless of neighborhood distance and was relatively high in all analyses—all analyses had an average stability of 87% and a median of 95%. This indicates that the model produces similar solutions over and over when using the same parameters for the same analysis area. This is a mark of quality for the model when nuanced by the comparisons with solutions from LP. Compared to solutions found using LP, the CA produced solutions with relative NPV of 94.3% or higher (Table 7). Bettinger et al. (2009) considers comparisons with optimal solutions generated for similar problems as a high level (five out of six) of validation but our comparison is vague since entry costs play a significant role in the CA but may not be included in the LP. Therefore, we do not consider our analysis as a level five validation. Furthermore, the harvest constraints also differ between the two. The relevance of the comparison lies in the fact that solutions found by LP are always the theoretical optima. Consequently, if the solution found by the CA is close to the solution found by the LP, it indicates that the CA produces high-quality forest plans. Therefore, the comparison was included.

Allowing for longer distances between separate areas within the same DTU increased NPV and utility function value. This is logical since increasing the distance has similarities with the relaxation of a constraint when using an exact solution method and matches the literature (Borges et al. 2017; Heinonen et al. 2018). A wider neighborhood distance allows the model to form more scattered DTU and results in fewer entry costs for a given distribution of treatments in the geography. What the appropriate neighborhood distance is, is however not obvious. The distance should reflect the point at which harvest machinery can no longer move within a harvest area without having transportation assistance by a trailer or similar. This distance is in practice dependent on several factors, many of which were neglected in this study (e.g., water courses, ground conditions, topography, roads). If a very short distance is used, the model will consider many treatments as isolated cuttings,

**Table 7** Comparisons of NPV and harvested volume in different CA, and LP solutions. Note that, the LP solutions do not include entry costs in the calculation of NPV. Therefore, even if the selection of TPs was exactly the same in a plan found using LP and a plan found using CA, the CA solution would have lower NPV due to higher costs

Solution	NPV (SEK ha <sup>-1</sup> )	NPV (rel.)	NPV (rel. to corresponding LP solution)	Avg harvest (m <sup>3</sup> ha <sup>-1</sup> year <sup>-1</sup> )
<i>Linear programming</i>				
No constraints	46 849	100.0	n/a	7.6
1 <sup>1</sup>	45 411	96.9	n/a	6.8
50 <sup>1</sup>	46 298	98.8	n/a	7.4
200 <sup>1</sup>	46 567	99.4	n/a	7.5
<i>Cellular automata</i>				
1 <sup>2</sup>	42 804	91.4	94.3	6.8
50 <sup>2</sup>	44 851	95.7	96.9	7.4
200 <sup>2</sup>	45 699	97.5	98.1	7.5

<sup>1</sup>Using a constraint stating that the harvested volume in each period must equal the harvested volume from the cellular automata solution with the corresponding neighborhood distance

<sup>2</sup>Using a utility function promoting even flow of harvested volumes (Fig. 3)

and thereby as non-profitable. If a very long distance is used, DTUs are to a more significant extent constituted by small and scattered areas, which the underlying harvest productivity models used in Heureka (Eriksson and Lindroos 2014) are not validated against.

The entry cost is a main component in the model and it was scaled down in the local and global phases. The aim here was that (while a lower detail mapping is conducted in the local and global phases) the effective entry cost charged from single DU would approximately correspond to the effective entry cost charged in the final phase. Furthermore, consideration was taken to the fact that a full entry cost would far exceed the income from thinning very small DU (a single  $12.5 \times 12.5 \text{ m}^2$  grid cell). Using a full entry cost, the model might consider all small DTU as unprofitable, not allowing small DTU to establish in the local and global phase, thus potentially getting stuck in local optima. After reasoning and brief testing, the entry cost was therefore scaled down in the local and global phases by multiplying it by 0.02.

Small DTU occurred in all cases, even though an important aim of the model is to avoid these. The fact that TPs (a sequence of treatments) are evaluated here, rather than individual treatments, is a possible explanation for this. A TP may be considered viable to the model, even though an individual treatment at a specific timepoint in the TP is not. Hence, the model may not guarantee that all treatment in all DUs in a solution are economically profitable, even though NPV was a component in the utility function.

Finally, delineation of forests into microsegments, i.e., the DUs in this study, may be of varying quality and is a source of error for forest planners using such data. Furthermore, the microsegments used here consists of sets of squares, forming straight and perpendicular patterns which in principle is a poor representation of the gradual variations that occur in real forests. As an alternative, remote sensing techniques allow for single tree identification and segmentation algorithms can be employed on such tree level data to form microsegments with more fine tuned borders (e.g., Olofsson and Holmgren 2014). Yet another alternative would have been to use the underlying cells as DUs, at the cost of an even more complex planning problem with many more decision variables. After all, the model for carrying out the segmentation and the model for conducting forest planning are two separate things. The former is not the focus of the study, and therefore, no investigation was conducted on the quality of the microsegmentation.

## Conclusions

We conclude that the presented approach is an addition to the set of heuristics applicable to forest planning with DTU. The main contribution of the study is the final phase

of the algorithm where the economic incentive to cluster treatments is modeled by calculating entry costs in high detail. The analyses showed that the final phase improves goal function value in solutions. While the magnitude of improvement appears dependent on parameterization in earlier phases of the algorithm and data, modeling a realistic entry cost, instead of using proxy variables, decreases the need for parameterization and expert-usage of decision support systems. While the final phase improved solutions, the phase was also computationally costly. In addition, early culmination of goal function value was observed, which leads us to reason that inclusion of stop criteria may be advisable to reduce solution times. The results also showed that the model produces consistent solutions when a given planning problem is solved repeatedly.

**Authors' contributions** P.W. developed the method, wrote the code, conducted the analyses and wrote the manuscript. K.Ö., T.L., and J.W. all contributed to the development of the method and provided feedback on the manuscript. J.E. processed and prepared forest data and provided feedback on the manuscript.

**Funding** Open access funding provided by Swedish University of Agricultural Sciences. The research resulting in this study was funded by the MISTRA Digital Forests research program, the ÅForsk Foundation, as well as the Swedish University of Agricultural Sciences.

**Availability of data** The data can be provided on request to the corresponding author PW.

**Code availability** The code can be provided on request to the corresponding author PW.

## Declarations

**Conflicts of interest** The authors declare that there are no economical nor academical conflicts of interest that could have affected the work in this study.

**Consent for publication** All authors have given consent for publication.

**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

## References

- af Ström AI (1822) Förslag till en förbättrad Skogshushållning i Sverige. Kungliga Lantbruksakademien, Stockholm
- Bettinger P, Sessions J, Boston K (2009) A review of the status and use of validation procedures for heuristics used in forest planning. *Math Comput for Nat-Resour Sci* 1(1):26–37
- Borges P, Kangas A, Bergseng E (2017) Optimal harvest cluster size with increasing opening costs for harvest sites. *For Policy Econ* 75:49–57. <https://doi.org/10.1016/j.forpol.2016.11.012>
- Davis LS, Johnson KN, Bettinger PS, Howard TE (2001) *Forest management*, 4th edn. McGraw-Hill, New York
- Eriksson M, Lindroos O (2014) Productivity of harvesters and forwarders in CTL operations in northern Sweden based on large follow-up datasets. *Int J for Eng* 25(3):179–200. <https://doi.org/10.1080/14942119.2014.974309>
- Faustmann M (1849) Berechnung des Wertes welchen Waldboden sowie noch nicht haubare Holzbestände für die Waldwirtschaft besitzen. *Allgemeine Forst- Und Jagdzeitung* 15:441–455
- Grilli E, Menna F, Remondino F (2017) A review of point clouds segmentation and classification algorithms, *The International Archives of Photogrammetry, Remote Sensing and Spatial. Inf Sci* 42:339–344. <https://doi.org/10.5194/isprs-archives-XLII-2-W3-339-2017>
- Heinonen T, Pukkala T (2007) The use of cellular automaton approach in forest planning. *Can J for Res* 37(11):2188–2200. <https://doi.org/10.1139/X07-073>
- Heinonen T, Kurttila M, Pukkala T (2007) Possibilities to aggregate raster cells through spatial optimization in forest planning. *Silva Fennica* 41(1):89–103
- Heinonen T, Makinen A, Rasinmaki J, Pukkala T (2018) Aggregating microsegments into harvest blocks by using spatial optimization and proximity objectives. *Can J for Res* 48(10):1184–1193. <https://doi.org/10.1139/cjfr-2018-0053>
- Holmgren P, Thureson T (1997) Applying objectively estimated and spatially continuous forest parameters in tactical planning to obtain dynamic treatment units. *For Sci* 43(3):317–326. <https://doi.org/10.1093/forestscience/43.3.317>
- Johnson KN, Scheurman HL (1977) Techniques for prescribing optimal timber harvest and investment under different objectives—discussion and synthesis. *Forest Science* 23(1):1
- Kotivuori E, Korhonen L, Packalen P (2016) Nationwide airborne laser scanning based models for volume, biomass and dominant height in Finland. *Silva Fennica* 50(4):1567
- Lu F, Eriksson O (2000) Formation of harvest units with genetic algorithms. *For Ecol Manag* 130(1–3):57–67. [https://doi.org/10.1016/S0378-1127\(99\)00185-1](https://doi.org/10.1016/S0378-1127(99)00185-1)
- Maltamo M, Packalen P (2014) Species-specific management inventory in Finland. In: Maltamo M et al. (2014) *Forestry applications of airborne laser scanning: concepts and case studies, managing forest ecosystems*, vol 27. Springer, Dordrecht, p 241
- Mathey A-H, Krcmar E, Tait D, Vertinsky I, Innes J (2007) Forest planning using co-evolutionary cellular automata. *For Ecol Manag* 239(1–3):45–56. <https://doi.org/10.1016/j.foreco.2006.11.007>
- Miguel Magaña S, Pukkala T, Tato JP (2013) Dynamic treatment units: flexible and adaptive forest management planning by combining spatial optimization methods and LiDAR. *Cuadernos De La Sociedad Española De Ciencias Forestales* 37:43–48
- Nelson J, Brodie D (1990) Comparison of a random search algorithm and mixed integer programming for solving area-based forest plans. *Can J for Res* 20(7):934–942. <https://doi.org/10.1139/X90-126>
- Nilsson M, Wasterlund DS, Wahlberg O, Eriksson LO (2012) *Forest Planning in a Swedish Company—a Knowledge Management Analysis of Forest Information*. *Silva Fennica* 46(5):717–731
- Nilsson M, Nordkvist K, Jonzén J, Lindgren N, Axensten P, Wallerman J, Egberth M, Larsson S, Nilsson L, Eriksson J, Olsson H (2017) A nationwide forest attribute map of Sweden predicted using airborne laser scanning data and field data from the National Forest Inventory. *Remote Sens Environ* 194:447–454. <https://doi.org/10.1016/j.rse.2016.10.022>
- Öhman K (2001) *Forest planning with consideration to spatial relationships*. Dissertation, Swedish University of Agricultural Sciences. Acta Universitatis Agriculturae Sueciae. Silvestria 198
- Olofsson K, Holmgren J (2014) Forest stand delineation from lidar point-clouds using local maxima of the crown height model and region merging of the corresponding Voronoi cells. *Remote Sens Lett* 5(3):268–276. <https://doi.org/10.1080/2150704X.2014.900203>
- Packalén P, Heinonen T, Pukkala T, Vauhkonen J, Maltamo M (2011) Dynamic treatment units in eucalyptus plantation. *For Sci* 57(5):416–426. <https://doi.org/10.1093/forestscience/57.5.416>
- Pascual A, Pukkala T, Miguel S, Pesonen A, Packalen P (2018) Influence of timber harvesting costs on the layout of cuttings and economic return in forest planning based on dynamic treatment units. *Forest Systems* 27(1):e001. <https://doi.org/10.5424/fs/2018271-11897>
- Pascual A, Pukkala T, de Miguel S, Pesonen A, Packalen P (2019) Influence of size and shape of forest inventory units on the layout of harvest blocks in numerical forest planning. *Eur J For Res* 138(1):111–123. <https://doi.org/10.1007/s10342-018-1157-5>
- Pukkala T (2019) Using ALS raster data in forest planning. *J for Res* 30:1581–1593. <https://doi.org/10.1007/s11676-019-00937-6>
- Pukkala T, Heinonen T, Kurttila M (2009) An application of a reduced cost approach to spatial forest planning. *For Sci* 55:13–22. <https://doi.org/10.1093/forestscience/55.1.13>
- Strange N, Meilby H, Jellesmark Thorsen B (2002) Optimization of land use in afforestation areas using evolutionary self-organization. *For Sci* 48(3):543–555. <https://doi.org/10.1093/forestscience/48.3.543>
- von Neumann J (1966) *The theory of self-reproducing automata*. In: Burks A (ed) *Theory of self-producing automata*. University of Illinois Press. Urbana and London
- Wikström P, Edenius L, Elfving B, Eriksson O, Lämås T, Sonesson J, Ohman K, Wallerman J, Waller C, Klinteback F (2011) The Heureka forestry decision support system—an overview. *Math Comput for Nat-Resource Sci* 3(2):87–94
- Wilhelmsson P, Sjödin E, Wastlund A, Wallerman J, Lamas T, Ohman K (2021) Dynamic treatment units in forest planning using cell proximity. *Can J For Res* 51(7):1065–1071. <https://doi.org/10.1139/cjfr-2020-0210>

**Publisher's Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.