Journal of Forestry Research (2025) 36:47
https://doi.org/10.1007/s11676-025-01834-x

ORIGINAL PAPER

Pathfinder: a tool for operational planning of forest regeneration on clearcuts

Linnea J. Hansson[1] · Anders Rowell[1] · Mikael Rönnqvist[2] · Patrik Flisberg[1] · Fredrik Johansson[1] · Rasmus Sörensen[1] · Morgan Rossander[1] · Petrus Jönsson[1]

Received: 17 July 2024 / Accepted: 24 November 2024

Abstract Effective forest regeneration is essential for sustainable forestry practices. In Sweden, mechanical site preparation and manual planting is the dominating method, but sourcing labour for the physically demanding work is difficult. An autonomous scarifying and planting system (Autoplant) could meet the requirements of the forest industry and, for this, a tool for regeneration planning and routing is needed. The tool, Pathfinder, plans the regeneration and routes based on the harvested production (hpr) files, soil moisture and parent material maps, no-go areas (for culture or nature conservation), digital elevation models (DEM), and machine data (e.g., working width, critical slope, time taken for different turn angles). The overall planting solution is either a set of capacity constrained routes or a continuous route and could be used for any planting machine as well as for traditional scarifiers as disc trenchers or mounders pulled by forwarders. Pathfinder was tested on eleven regeneration areas throughout Sweden, both with continuous routes and routes based on a carrying capacity of 1500 seedlings. The net operation area, species and seedling density suggestions were deemed relevant by expert judgement in the field. The routes provided by Pathfinder were compared with solutions given by two experienced drivers and a third solution based on the actual soil scarification at the site. Total driving distance did not differ significantly between the suggestions, but Pathfinder included less side-slope driving on steep slopes (≥ 27% or 15°) and medium slopes (15–27%). The chosen threshold value for steep slopes (where side-slope driving should be avoided) affects the routing, and a lower threshold means more turning and longer driving distance. Pathfinder is not only a tool for routing of planting machines, but also helps in planning of traditional regeneration by providing a more correct net area and tree species suggestions based on the growth of the previous stand. It also diminishes the risk of severe soil disturbance by excluding the wettest area in the planning.

Keywords Routing · Site preparation · Planting · Operational planning · Optimization

Project funding: This study was funded by Vinnova, the Swedish Innovation Agency as a part of the Autoplant project (Dnr 2020-04202 and 2023-02747).

The online version is available at https://link.springer.com/.

Corresponding editor: Jingjue Sun.

✉ Linnea J. Hansson
linnea.hansson@skogforsk.se

[1] Skogforsk - the Forestry Research Institute of Sweden, 751 83 Uppsala, Sweden

[2] Département de Génie Mécanique, Université Laval, Québec, QC G1V 0A6, Canada

Introduction

There is a clear shortage of skilled labour in silviculture operations in boreal forestry (Ramantswana et al. 2020; Forsmark and Johannesson 2020). Manual planting is arduous, and traditional mechanical site preparation includes a high dose of whole-body vibrations for the operator when traversing the terrain. Efficient regeneration methods are crucial to sustainable forestry, and in Sweden, where 99% of the planting is manual, interest in mechanized planting is rising. Autoplant is one of several Swedish collaborative projects among forest companies, manufacturers and researchers within silvicultural technology (Hansson et al. 2024). The overall aim of the Autoplant project is to develop an integrated autonomous site preparation and planting system with high precision, low environmental

Published online: 18 March 2025

impact, and a good work environment. A decision support tool is needed for operational planning of the forest regeneration, including route planning for such a machine.

Several decision support tools already exist for path planning in forestry (Flisberg et al. 2007, 2022, 2021; Hosseini et al. 2019, 2023; Hansson et al. 2022; Ovaskainen and Riekki 2022; Holmström et al. 2023). The tools either design the network for the routing or determine the routes given an existing network. The main challenge for a decision support tool for operational planning of planting is to design the network and routing simultaneously.

During forest harvest operations, data is collected about every tree location and all timber bucking in harvested production (hpr) files (Anon 2021a). Even if the exact original tree position is not collected by all harvesters, at least the harvester position is stored. This information allows for a digital reconstruction of the entire forest, extrapolating diameters along the stem to model total tree height. This, together with the age of the stand, can be used to estimate site index, i.e., the dominant tree height at 100 years of a specific tree species which corresponds to site productivity (Hägglund and Lundmark 1977), and supports subsequent regeneration planning.

If a clearcut area is to be regenerated with autonomous methods, an operational plan for the machine is needed, including operating area, which tree species to plant where and at what density, and finally how to achieve time-efficient routes that consider no-go areas and critical slopes. The Pathfinder tool was designed to meet these criteria. The aims of this study are to:

1. Describe the Pathfinder tool for forest regeneration planning and routing, based on harvested production data, soil moisture and parent material maps, digital elevation models, and machine specification. The tool should work not only for planting machines but also for traditional scarifiers, and it should provide both shorter routes where, for example, seedling carrying capacity is limited, and continuous routes where capacity does not restrict route lengths.
2. Test the tool on several clearcut regeneration areas throughout Sweden. The sites should vary in terms of characteristics such as size, slope, tree species, production capacity, and no-go areas.
3. Compare the solutions presented by the Pathfinder tool to manual solutions produced by two experienced operators, as well as the actual solution executed at the site.

Materials and methods

The pathfinder tool

The Pathfinder tool consists of two main parts: the seedling-planning tool, here called *PlantP,* and the routing tool, here called *PathP*. The results from *PlantP* serve as input to *PathP*.

PlantP

The *PlantP* module was developed during previous studies under the Swedish name Plantbeställning for planning species and number of seedlings for a forest regeneration area (Sörensen et al. 2023; Friberg et al. 2019). The *PlantP* module uses the data from the hpr files (i.e., tree height, diameter, root rot, species, and location in the landscape, Fig. 1b) to model the pre-harvest forest. Site indices (SI) are then averaged over units of approximately 1 ha (Fig. 1e). Site index is calculated from the hpr data by an algorithm described in Möller et al. (2011), by connecting the dominant tree height in each unit of calculation with the given or calculated stand age for the same unit (Offenbacher 2024). The dominating tree species for site index is based on the highest basal area at 1.3 m (here P for Scots pine (*Pinus sylvestris* L.) and S for Norway spruce (*Picea abies* (L.) H. Karst.)). Decay occurrence in stems, mainly caused by root or butt rot (*Heterobasidion* spp., *Armillaria* spp., or *Stereum sanguinolentum*), can be interpreted from the hpr files as manually chosen cutting lengths as opposed to lengths matching the currently preferred assortment. Site index is used for suggesting tree species and density except in the following cases (Table 1, Möller et al. 2011). Spruce is suggested instead of pine if there are at least 25% spruce and if SI > P28 (i.e., where pine is the dominating species and will be over 28 m at the age of 100 years). Pine is suggested instead of spruce if decay in the spruce stems exceeds a threshold level (here 70% of the total number of stems); if the decay is 30–70% an equal mix of pine and spruce is suggested. These threshold levels can be modified by the user in a decision scheme, and more species and thresholds can be added according to user preferences. The harvested area from the hpr files is used to decide the net area for regeneration, including a buffer of 6 m on the outer side of each harvester staging area and 12 m on the inner side (Fig. 1b, e). Non-regeneration areas are excluded based on additional cartographical data:

1. Non-productive land and other areas with a low tree coverage pre-harvest, based on data on basal area, height, volume, and date of scan (Swedish Forest Agency 2024), are excluded from the regeneration planning.
2. The Swedish National Road Database is used to distinguish regeneration areas from roads, and roads are then excluded.
3. A soil moisture index map (Swedish Natural Land Cover Database; Olsson and Ledwith 2020) is used to identify and exclude harvested stems in areas that are too

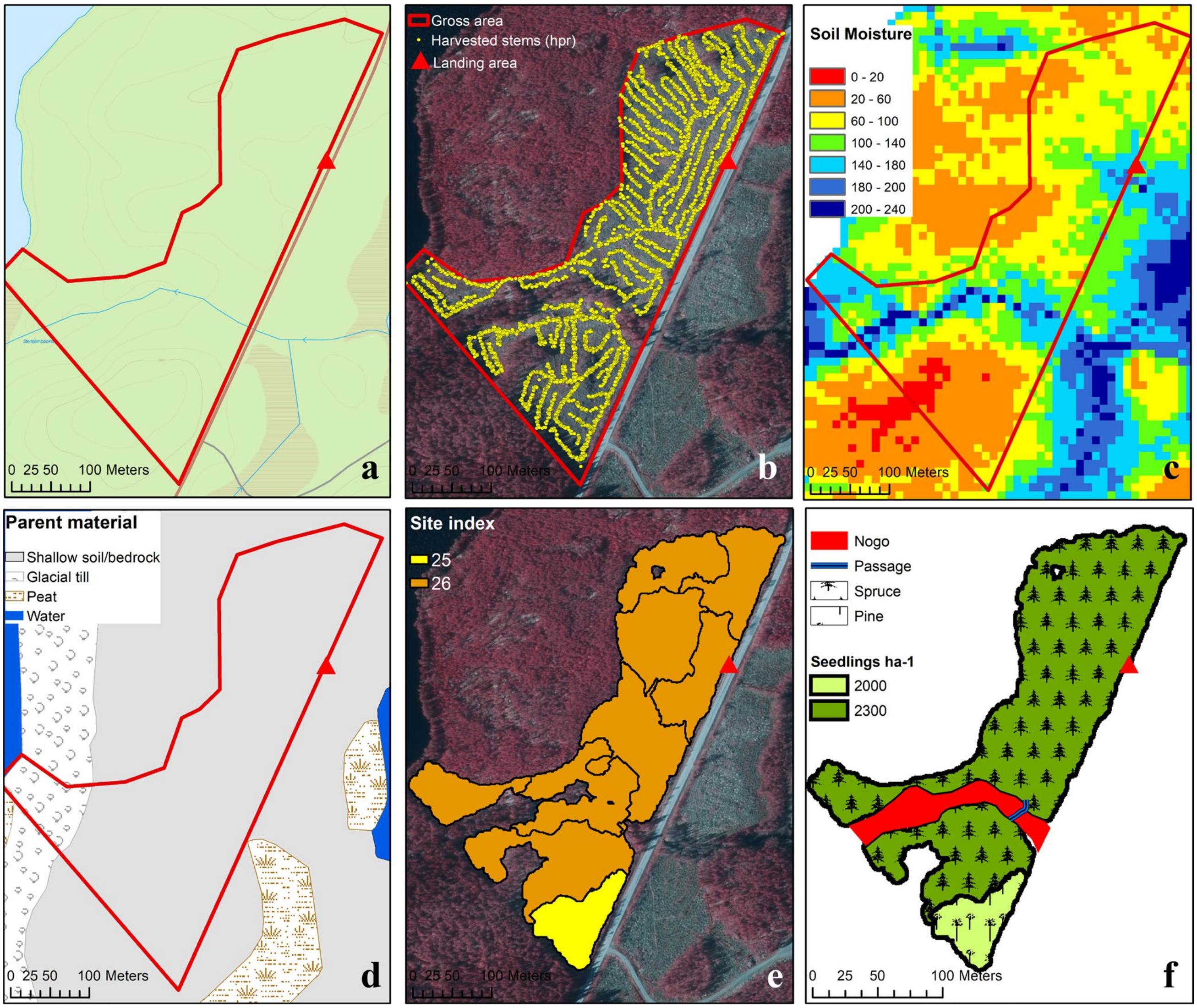

Fig. 1 Input and output from the *PlantP* module at regeneration area 1; **a** planned harvested area (gross area reported to the Swedish Forestry Agency); **b** harvested stems (hpr), i.e., harvester position when felling a tree; **c** soil moisture map; **d** parent material map; **e** site index derived from the stem data in the hpr-file; **f** output from the PlantP module, including net area, species, and seedling density per hectare; the red triangle is the landing area where the seedlings are stored © Lantmäteriet, © The Swedish Environmental Protection Agency, © Swedish Geological Survey

wet, where 240 is saturation and 0 is a theoretical value where the soil is completely dry (Willén et al. 2021, Fig. 1c). In areas drier than 20, Scots pine is suggested and above 200, the area is set as too wet to operate. The chosen threshold values are based on inventory data from ~6000 sample plots throughout Sweden (Sörensen et al. 2023).

4. Parent material maps (Fig. 1d) from the Swedish Geological Survey (Anon. 2021b) are used to exclude peat soils, which was the only parent material class that was proven well enough predicted in the inventory data (Sörensen et al. 2023).
5. Additional no-go areas are drawn manually based on the field planning of the forest company (Fig. 1f).

PathP

There are several aims when establishing efficient vehicle routes. They should be long to avoid time consuming turns and avoid passing the same areas multiple times. At the same time, they should cover the entire area and avoid driving in steep side slopes, thereby forcing the routes in a specific direction. The net regeneration area from *PlantP* (Fig. 1f) is used as input to the *PathP* module together with additional no-go areas (if present), forced passages, a digital elevation

Table 1 How tree species and seedling density are suggested in the *PlantP* module based on site index (SI) and decay occurrence

SI detected	Plant density (ind. ha^{-1})
P25+	2300
P20-P24	2000
< P20	1700
S32+	2300
S24-S31	2000
< S24	1700
SI-species >28 → change to spruce, if at least 25% spruce at harvest	
Decay occurrence when spruce is dominating species	
< 30%, no change of species	
30–70%, pine-spruce mix recommended	
> 70%, change to species from spruce to pine	

P25 means that the highest pine trees are 25 m at 100 years of age; S denotes spruce; P denotes pine

model (DEM, resolution: 1 m), and machine data. The latter include seedling carrying capacity (here 1500), working speed (default 1 m s^{-1}) and working width (here 6 m), critical slope where driving direction is significant (here ≥ 27% or 15°). As sharp turns are difficult both for autonomous terrain vehicles in unknown terrain and for large traditional scarifiers, the tool aims to minimise sharp U-turns. We can set different penalties or additional times for turning in each specific radius. These are input parameters for determining different cost coefficients (examples are given in the case study below).

To find the most efficient routes that minimise overall driving time, the solution approach is divided into four steps. The motivation is to identify important characteristics of each identified region (part of the entire area), and then build up a solution by making overarching or strategic decision first (for example, driving direction) and local operational decisions (for example, turning) later. Each step uses a different model formulation and solution methodology. The steps can be briefly summarised as follows.

Step 1. Area discretisation and critical slope calculation

Step 2. Identify regions with similar slope directions

Step 3. Identify long driving lines in each region that cover the entire area

Step 4. Identify best turnings to connect driving lines

Each of the steps are described in more detail below.

Step 1. Area discretisation and critical slope calculation

We start by dividing the overall area into hexagons (width 10 m, side length 5.785 m, and height 11.6 m, Fig. 2a). We use the DEM to compute the principal directions in each hexagon (Fig. 2b). We also need to identify average directions in areas larger than a single hexagon. Slope within a hexagon is evaluated as a percentage in directions with a step size of 15°. The slope is determined using a circle with a diameter of 10 m, centred at the midpoint of the hexagon. For each direction, the slope is calculated between two points on the circle's circumference passing through the middle in the given direction. For hexagons where the slope exceeds the threshold (here 27%), we enforce a driving direction (Fig. 2c).

Step 2. Identify regions with similar slope directions

The overall area has regions with different slope characteristics. This step aims to identify these regions, where the slope is consistent throughout each region. The problem can be formulated as a set-partitioning model. This is an integer programming (IP) problem and can be formulated as (Eq. 1a, b, c). The parameters and decision variables are defined in Table 2.

$$\min z = \sum_{r \in R} c_r y_r \tag{1a}$$

$$\sum_{r \in R} s_{hr} y_r = 1, \forall h \in H \tag{1b}$$

$$y_r \in \{0, 1\}, \forall r \in R \tag{1c}$$

The objective function (Eq. 1a) is the total cost of the chosen regions. For each region, the best direction is determined. The cost c_r is designed to increase logarithmically with the number of hexagons in a region. It also increases for each hexagon in the region where the best direction for the region is not the best for the hexagon. Hence, the model finds a solution

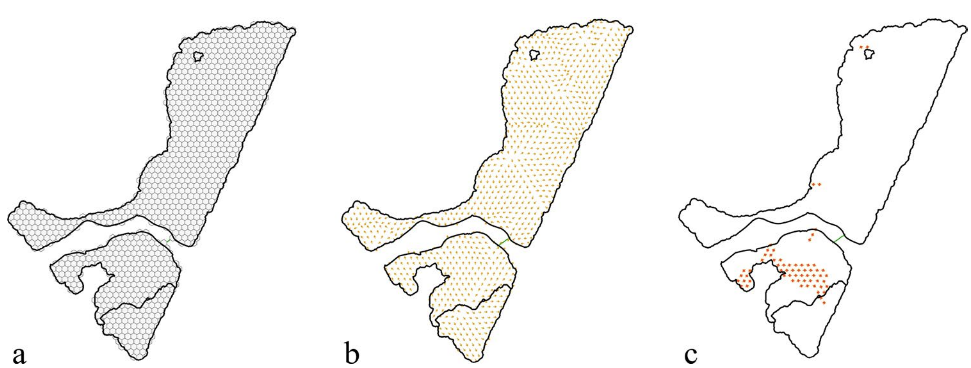

Fig. 2 Results from Step 1 for regeneration area 1. **a** Distribution of hexagons in the area; **b** initial dominating slope directions (arrows); **c** enforced directions due to the slope limit (red arrows)

Table 2 The parameters and decision variables for Eq. 1

Parameter	Variable	Definition
H		Set of all hexagons
R		Set of all possible regions of connected hexagons
s_{hr}		1 if hexagon h is in region r, 0 otherwise
c_r		Cost of region r
	y_r	1 if region r is selected, 0 otherwise

where the overall area is subdivided into regions, each with similar principal directions of the hexagons. The constraints are to ensure that each hexagon is included in exactly one region (Eq. 1b) and that the binary restrictions on the variables are fulfilled (Eq. 1c). The potential number of regions is extremely large due to an exponential growth of combinations, and it is not practical to generate all. Instead, we solve the model heuristically in two phases. The first phase is to solve the set-partitioning model with a set of limited regions made up of different sized hexagons (Fig. 3a). Here, we create regions for each hexagon by creating circles with larger and larger radius around them (in effect, creating larger hexagon regions). All hexagons within the circle radius are included in a region. Only regions where all the hexagons in the region are inside the harvest area are allowed. Also, there must be at least one driving direction in each region that is allowed for all hexagons within the region that have enforced driving directions. The number is limited as each hexagon can only contribute with a linear number of larger hexagons based on the radius of the circles considered. This reduced set-partitioning problem can be solved to optimality using the commercial solver CPLEX (IBM ILOG CPLEX Optimization Studio, IBM, Armonk, N.Y.)

The second phase is to combine the found regions (different sized hexagons) with a repeated matching (RM) algorithm (Flisberg et al. 2007) where the regions are matched together two and two to larger areas until no more matches can be done (Fig. 3b, c). This does not guarantee an optimal solution, but the performance of the RM algorithm is known to produce high quality solutions for this type of problems. The RM algorithm also includes finding the best direction for all matched regions.

Step 3. Identify long driving lines in each region that cover the entire area

For each identified region, we need to determine paths for the vehicle that cover the entire region. Here, we generate many lines (representing straight paths) for each region that start on one side and end at the other side of the region (Fig. 4). From this large set, there is a need to select a subset that covers the region. To formulate constraints for this aspect, we introduce discretised points for each square metre (Fig. 4a, b). Each line will cover a subset of points depending on the assumption of the reach of the planting arm. The optimisation problem is to select a subset of lines so that all discretised points are covered. The problem can be formulated

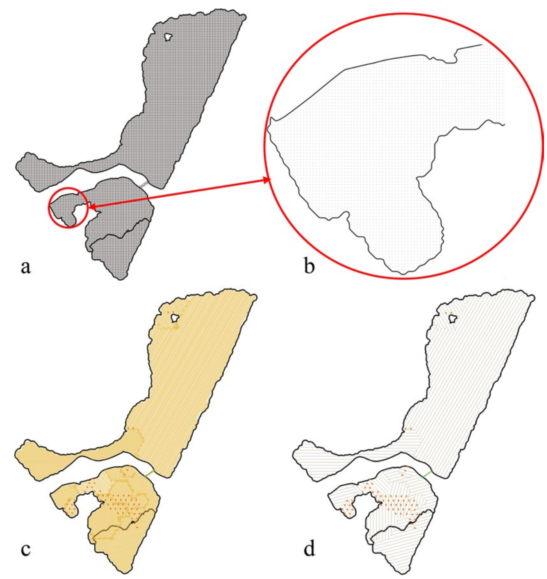

Fig. 4 Results from Step 3 for regeneration area 1. **a** All discretised points to cover; **b** zoomed in part of the area of the discretised points; **c** generated large set of lines; **d** selected lines from the extended set-covering model that cover all points (red arrows indicate forcing direction (slope ≥ 27%))

Fig. 3 Results from Step 2 for regeneration area 1. **a** Solution of mixed sized hexagons with similar slope directions in the first phase; **b** combinations of mixed-sized hexagons to regions with similar slope directions in the second phase; **c** identified regions with similar directions in different colours

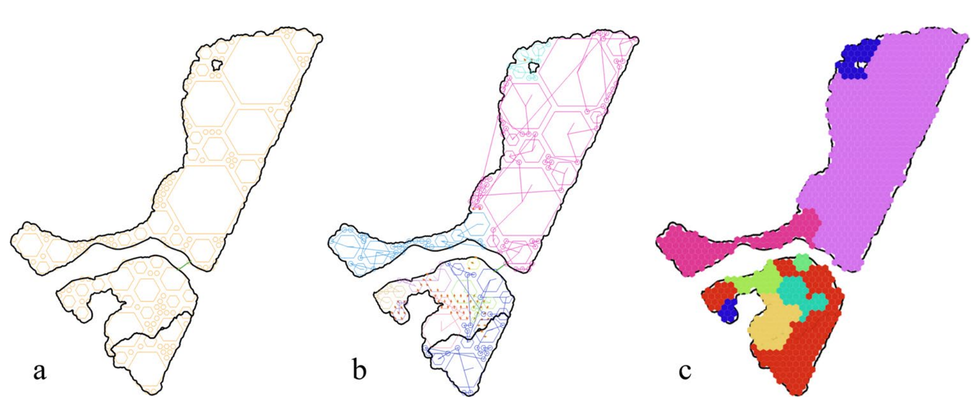

Table 3 The parameters and decision variables for Eq. 2

Parameter	Variable	Definition
P		Set of all discretised points
L		Set of lines
L_l^O		Set of lines that should be avoided if line l is used (since they will overcover some areas)
m_l		Max number of lines from the set L_l^O that can be used without penalty
a_{lp}		1 if line l covers point p, 0 otherwise
c_l		Cost of line l
o		Penalty for overcovering
r_l		Length of line l
r_{max}		Length of the longest line
	z_l	1 if line l is selected, 0 otherwise
	p_l	Penalty if other lines are covering line l

as an extended set-covering model, which is an IP problem, as Eq. 2a, b, c, d, e. The parameters and decision variables are defined in Table 3.

$$\min z = \sum_{l \in L} (c_l z_l + o p_l) \tag{2a}$$

$$\sum_{l \in L} a_{lp} z_l \geq 1, \forall p \in P \tag{2b}$$

$$\sum_{i \in L_l^O} z_i + m_l z_l \leq m_l + p_l, \forall l \in L \tag{2c}$$

$$z_l \in \{0, 1\}, \forall l \in L \tag{2d}$$

$$p_l \geq 0, \forall l \in L \tag{2e}$$

The objective is to minimise the cost of all selected lines and penalty costs for overcovering. The constraints are to ensure that each point in the area is covered at least once (Eq. 2b), limit overcover of lines (Eq. 2c), binary restrictions (Eq. 2d), and non-negativity constraints (Eq. 2e). Due to the detailed discretisation, there will be some overcover in practice. With Eq. 2c, it is possible to reduce the overcover and at the same time manage which lines are used. The cost for a line is defined as Eq. 3.

$$c_l = r_l \left(1 + \log \left(\frac{r_{max}}{r_l} \right) \right) \tag{3}$$

The set-covering problem is relatively easy to solve and can be solved optimally using CPLEX (Fig. 4d).

Step 4. Identify best turnings to connect driving lines

The lines provide paths where the machine is to drive in the area, but the driving direction along a line has not been decided yet. Given the lines, we now need to identify the driving direction and turns or change of directions needed to pass along each line. Essentially, we need to identify how to combine the lines in such a way that we construct one continuous route. This is the case if we can refill the machine with seedlings during the journey along the path, or if the machine can carry all seedlings needed for the site. If we need to return to a reloading point, multiple routes are needed. In practice, we solve a travelling salesman problem (TSP, in the case of one continuous route) or a capacitated vehicle routing problem (VRP, in the case of multiple routes). To solve the TSP (or VRP) problem, we must formulate a network problem with nodes and arcs.

In a TSP model formulation, the nodes represent the "cities" the "salesman" must visit, and the TSP route, consisting of arcs or edges, represents the sequence that minimises the length or cost. Also, as the directional driving costs are not symmetric, we need arcs (edges with a direction) in the network formulation for the TSP. In our case, the driving lines represent such cities or nodes, and the possible turns represent arcs. Figure 5 describes a small example when we study the turning options from the red node. There are four different turning options to the edges 1–4 and the time (or cost) for the turning differs. For example, making a 180°turn to edge 4 takes longer than driving straight to edge 2. Turning to edge 3 is relatively easy while making a 90°turn to edge 1 takes longer time. To formulate the network for the TSP, each edge must be represented by three nodes and four arcs. The reason for this is that we must guarantee that the machine follows the entire line and by using the three nodes and four arcs we enforce this requirement. We make sure that all middle nodes for each edge will only have arcs to the other nodes associated to the edge and no turning can start or end at the middle node. This means that the TSP route is forced to drive along each driving line, and we can identify

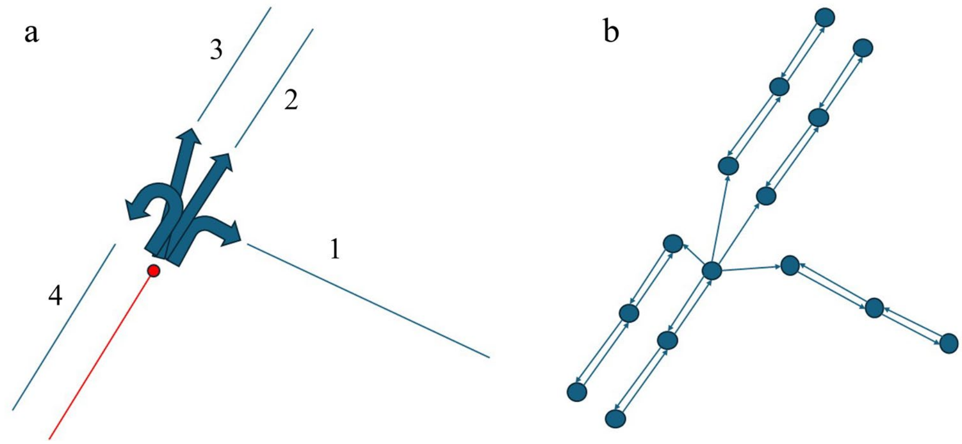

Fig. 5 Turning alternatives and network formulation. **a** Four turning alternatives from the red dot to the four possible driving lines (edges); **b** network representation with nodes and arcs of the edges and turning alternatives

the direction with the arcs. The time for each turning alternative is precomputed for every arc in the network.

The TSP model, which is an IP problem, can be formulated as Eq. 4a, b, c, d, e. The parameters and decision variables for the TSP formulation are defined in Table 4.

$$\min z = \sum_{i \in N} \sum_{j \in N} c_{ij} w_{ij} \tag{4a}$$

$$\sum_{j \in N} w_{ij} = 1, \forall i \in N \tag{4b}$$

$$\sum_{i \in N} w_{ij} = 1, \forall j \in N \tag{4c}$$

$$\sum_{i \in S} \sum_{j \in S} w_{ij} \leq S - 1, \forall S \subset N, S \geq 2 \tag{4d}$$

$$w_{ij} \in \{0, 1\}, \forall i, j \in N \tag{4e}$$

The objective (Eq. 4a) is to minimise the total cost (time). The constraints are to ensure that each node has an arrival and departure arc (Eq. 4b, c), no subtours are allowed (Eq. 4d), and the restriction of the binary restrictions (Eq. 4e). The set S is defined as all subsets of N and are used to eliminate subtours in the formulation. The TSP problem is very hard to solve. This formulation assumes all arcs are included in the network formulation. Arcs that cannot be used have a large cost. We use the well-known Concorde software (https://www.math.uwaterloo.ca/tsp/concorde.html), which is known to be a very efficient solution methodology for TSP problems. The VRP formulation is used when the planting machine must return to a central location to resupply seedlings. Therefore, we need to find a set of routes such that the machine visits all planting locations, with each route limited by the seedling capacity of the machine. The VRP is a generalization of the TSP problem and there are many models and solution methods available (Audy et al. 2023). To solve the VRP problem, we use a heuristic from the Google developer platform (http://developers.google.com/optimization/routing/vrp).

Case study

Regeneration areas

Eleven clearcut regeneration areas spread out over Sweden were selected for this study based on their representativeness of the regeneration areas of that forest owner, and ensuring that different slope steepness, tree species, soil moisture, and parent material were represented (Table 5, Fig. 6). The areas were owned by four forest companies (Holmen, Sveaskog, StoraEnso and SCA) and two forest owners' associations (Södra and Mellanskog). They were visited in the field by a group of forest regeneration experts from the participating organisations, to verify that the suggestions from the *PlantP* module were feasible.

Table 4 The parameters and decision variables for Eq. 4

Parameter	Variable	Definition
N		Set of nodes
S		Subsets of N
c_{ij}		Cost (extra time compared to driving straight) of using arc between node i and j
	w_{ij}	1 if arc between node i and j is in the TSP tour, 0 otherwise

Table 5 A summary of the eleven regeneration areas included in the study: soil material derived from parent material maps (Anon. 2021b); hpr-based site index where P denotes Scots pine, S Norway spruce and the number their height at 100 years; gross area, i.e. the planned harvest area (ha); net area (ha) based on the hpr-files, total number of spruce seedlings, total number of pine seedlings, spruce density per hectare, and pine density per hectare

Object ID	Soil material	Site index$_{hpr}$	Gross area (ha)	Net area$_{hpr}$ (ha)	Net area *PlantP* (ha)	Tot. Spruce	Tot. Pine	Spruce (ind. ha^{-1})	Pine (ind. ha^{-1})
1	Bedrock, glacial till	P25 S26	9.6	8.4	8.1	17,066	1400	2300	2000
2	Clay/silt, bedrock	S27	7.7	6.7	5.8	13,227	0	2300	0
3	Glacial till, clay/silt	P25 S26	9.4	6.6	6.3	11,745	2200	2300	2000
4	Glacial till, clay	P27 S31-33	15.1	12.0	11.9	21,093	6240	2300	2300
5	Glacial till, peat	S32 S33	2.9	2.5	2.1	4851	0	2300	0
6	Sandy till, bedrock, peat	S33	6.1	5.9	5.9	13,467	0	2300	0
7	Bedrock, glacial till, peat	P28 S33	4.7	4.7	4.3	5718	4265	2300	2300
8	Glacial clay, sandy till	S28-30	6.9	6.9	6.9	15,791	0	2300	0
9	Glacial till	S31-32	3.4	2.3	2.3	5400	0	2300	0
10	Bedrock, sandy till, clay, peat	P28 S33	10.1	9.3	8.3	0	19,194	0	2300
11	Till, glaciofluvial sediment	P23 S23	8.4	8.1	8.0	2137	14,144	2300	2000

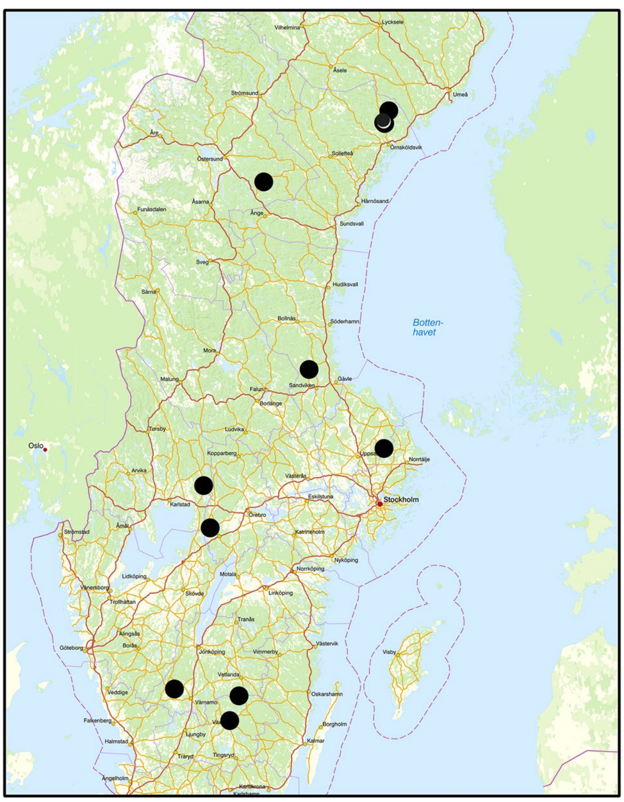

Fig. 6 The location of the eleven study sites (black circles) throughout Sweden © Lantmäteriet

Data preparation

The following data was collected from all regeneration areas: gross area planned for harvest, hpr-files, GIS data including DEM, soil moisture and parent material maps, additional no-go areas, forced passages, roads, and GNSS (Global Navigation Satellite System) tracking from scarifiers or orthophotos including scarification tracks. We used extra time penalties for 180° turns of different radii: a radius of 0 m gave a time penalty of 35 s, 6 m gave 25 s, and ≥ 12 m gave 15 s. If the radius was between any of those values, the penalty was linearly interpolated. A 90° turn gave a penalty of 5 s which linearly decreased to 0 s when the turn angle was 0°.

Two experienced operators (named A and B) who are used to giving advice to less experienced operators were asked to divide the 11 regeneration areas into smaller sub-areas (if necessary) and provide a driving direction for each sub-area (Fig. 7). They were given digital elevation maps of the areas, with wet areas marked, and a topographic map with contour lines at 5 m intervals (Fig. 7). The tracking data from the actual operations or the directions from the orthophotos were managed in the same way to create a third proposal for driving directions (named C). The directions were transformed into twelve classes, each corresponding to a direction in 15° steps from 0 up to 165°, where the first was 0, the second 15, and the twelfth 165. An example from area 1 is provided in Fig. 7.

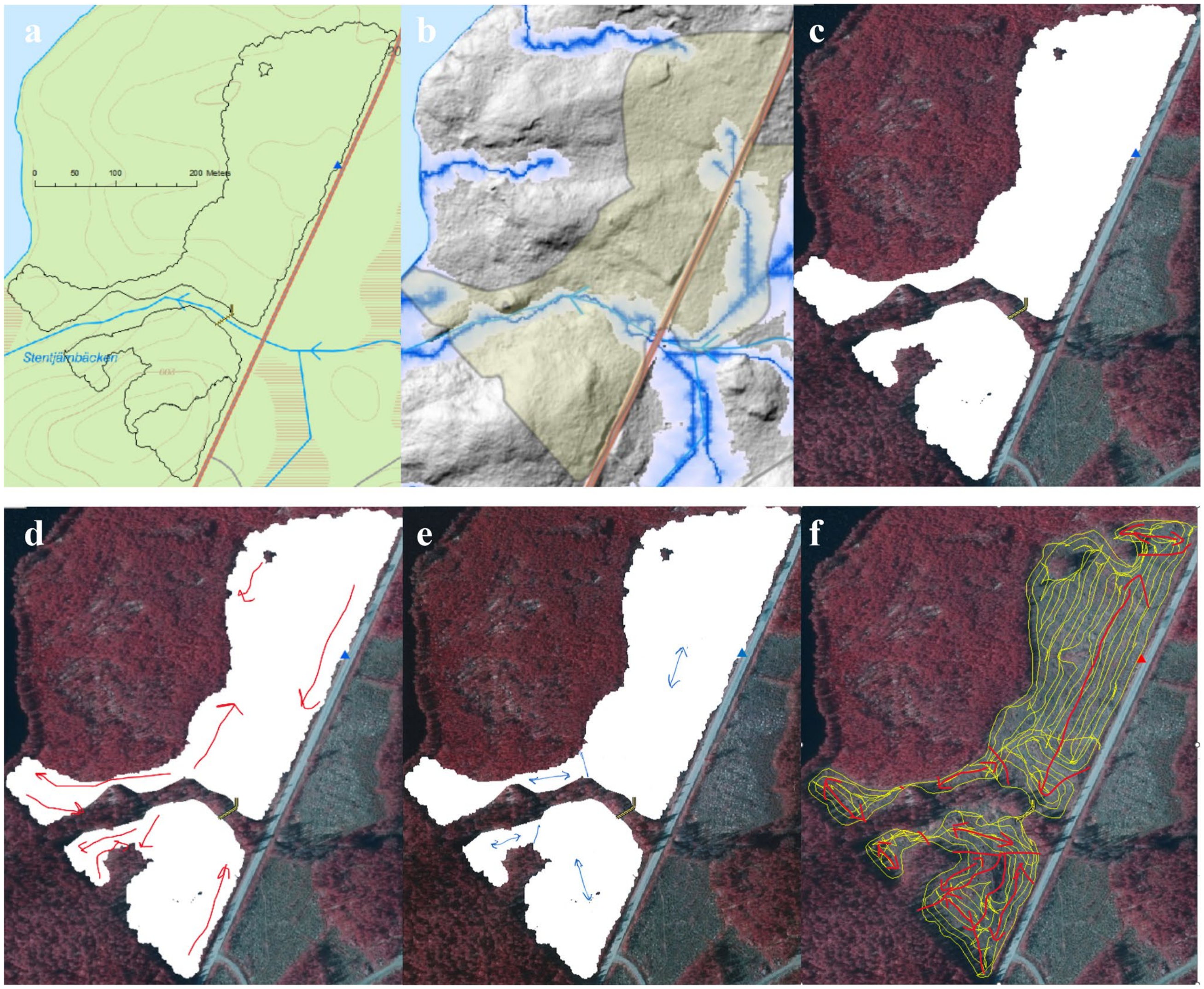

Fig. 7 Upper row: **a** topographic map (contours at 5-m intervals); **b** a digital elevation model with wet areas marked, and **c** the map on which the experienced operators (A and B) drew their sub-areas and driving directions. Lower row: Suggestions by operators A (**d**), B (**e**), and tracking data (**f**) from the actual operation (C) with sub-areas and driving directions marked with red arrows. © Lantmäteriet

To test how the threshold value for critical side-slope driving affected the routing, *PathP* was also run with a lower threshold value of 15%. Unless stated otherwise in the results, the 27% results are presented.

The side-slope driving distance (ℓ_{ss}) when planting was calculated as Eq. 5.

$$\ell_{ss} = \alpha_{max}(1 - |\cos(|\theta - d|)|) \tag{5}$$

where, α_{max} is the steepest slope ratio and $|\theta - d|$ is the difference between the driving direction, θ and the direction of the steepest slope, d.

We define a set of classes for different levels of side slope (similar to the driving direction classes). The total distance is then the summation of the driving distance in each class as Eq. 6.

$$\sum_{\alpha_{max} \in \alpha_k} \alpha_{max}(1 - |\cos(|\theta - d|)|),\ \forall k \tag{6}$$

where, α_k is the slope class.

Statistical analysis

One-way ANOVA analyses were performed to compare the results from the three operators (A, B, and C) against the results from Pathfinder across the 11 regeneration areas using a linear model approach with the Ordinary Least Squares (OLS) method, implemented via the Statsmodels library in Python. The analysis treated each of the 11 cases as distinct categories, with separate assessments conducted for each case. Dependent variables such as 'distance' and 'number of turns' were analysed to determine the statistical differences between each operator (A, B, or C) and Pathfinder across these cases. The treatment groups (i.e., the different operators) were specified with the Pathfinder result as the reference, and each case was treated as a fixed effect in the model to understand its unique impact. Residuals from the model were evaluated through Q-Q plots, histograms with 20 bins, and scatter plots to check the assumptions of normality and homogeneity of variances. A 95% confidence interval and a significance threshold of $p < 0.05$ were used to interpret the findings.

Results

PlantP

The results from *PlantP* are presented in Table 5. On average, the net harvested area derived from the production files (hpr) was 0.3 ha smaller than the gross area planned for final felling (as reported to the Swedish Forest Agency). The net area from *PlantP* was, on average, 1.0 ha smaller than the harvested production area, comprising additional area exclusions from the depth to water maps, soil parent material maps (i.e., peat areas) and additional no-go areas (Table 5). In areas with varying soil conditions resulting from differences in elevation, the suggestions on tree species were assessed as relevant (expert judgement in field, Johansson pers. comm), with Scots pine on the hills and Norway spruce in lower areas.

PathP

The solution time for all four steps and for each object was short, limited to about four minutes for the largest regeneration areas (12 ha). The results from the *PathP* module were similar to the results based on the two skilled operators (A and B) and the actual operation (C) (Fig. 8). Hence, the route planning in *PathP* can be used for operational planning. Note that the operators only provided the driving directions in areas of their own choice. The routes are then optimised in the same way in all four cases. This means that the difference between operators and *PathP* is essentially the solution from Step 1–2 from the methodology, as Step 3–4 is used for the operators' solutions.

There were no significant differences in total distance among the solutions of continuous routes (TSP, Fig. 9) or routes based on a capacity restriction of 1500 seedlings (VRP, Fig. 10). The total number of routes in all sites were 122, 123, 120 and 94 for *PathP*, operators A, B, and C, respectively. Note that data is missing for site 10 and 11 for operator C, which explains the lower number. The driving distance on side slope greater than 27% was shorter for *PathP* than for operators A, B, and C (Fig. 11). Also, the side-slope driving distances between 15 and 27% for *PathP* were shorter than for the operators (Fig. 11).

The more aggregated solutions provided by operators A and B resulted in a lower number of gentle turns on average than the solutions by *PathP* (time penalties < 30 s, Table 6). There were no statistical differences between the solutions for turns with higher time penalties than 30 s. The average time penalty for turning per hectare was higher for the *PathP* than for the two skilled drivers A and B (Table 6).

When the threshold value for steep slopes was changed from 15°(27%) to 15% in the *PathP* solutions, the routing became more complicated (Fig. 12). The additional turning time (penalties) increased by 27% (TSP) and 18% (VRP) (data not shown). The total driving distance increased slightly (4% for TSP and 2% for VRP, data not shown).

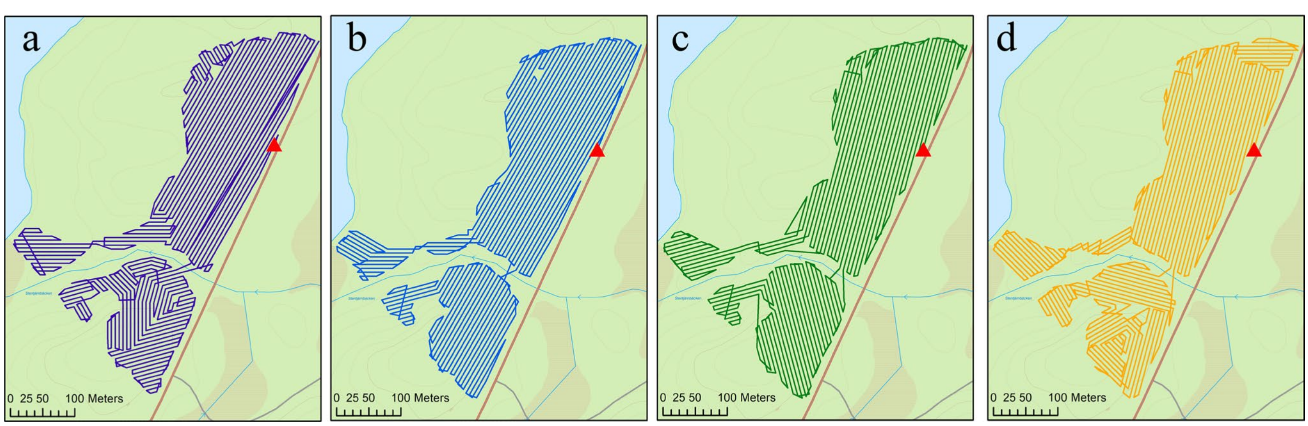

Fig. 8 Example of the different TSP solutions at site 1 from PathP (**a**), Operators A (**b**), B (**c**), and C (**d**)

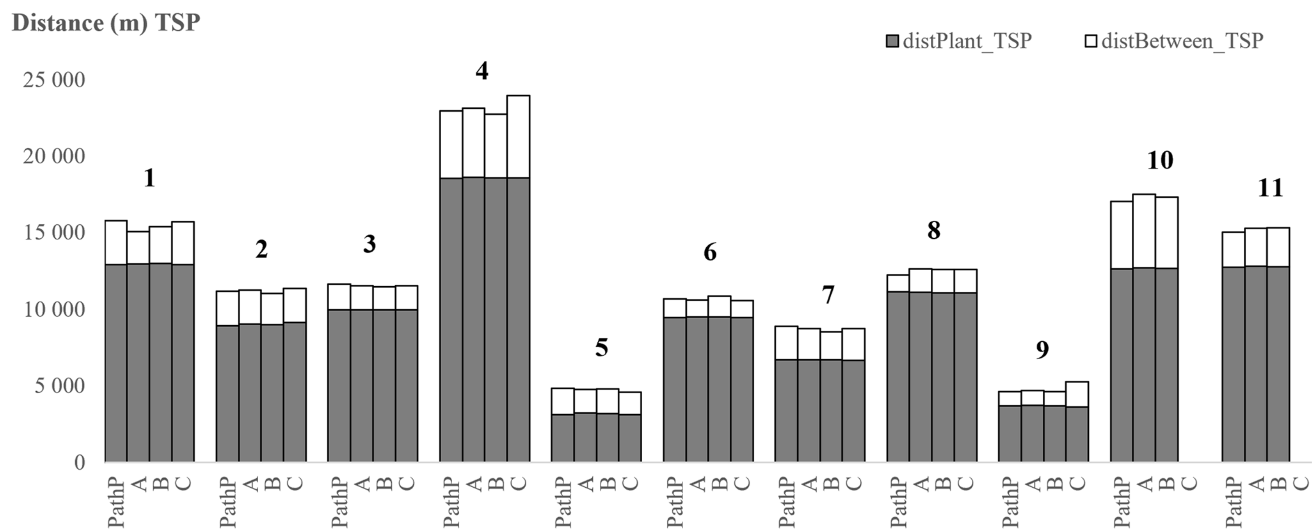

Fig. 9 Total driving distance, divided into distance planting (distPlant) and distance not planting (distBetween) across the 11 different sites, comparing the TSP (Travelling salesman problem, continuous route) solution from Pathfinder (*PathP*) with solutions based on driving directions from operators A, B, and C

Discussion

A decision support tool is needed for operational planning of an integrated autonomous site preparation and planting system. The objective of developing Pathfinder was to produce an operational plan for the machine to follow, which tree species to plant where and at what density, and to achieve efficient routes avoiding no-go areas and critical side slope. Pathfinder is on the way to fulfilling these objectives, but improvements are still needed before it can serve as the global plan for an autonomous machine. However, there is already potential for using the route direction and spacing as a decision support tool for inexperienced scarifier operators as the automatically suggested driving directions were similar to the directions chosen by skilled operators (A and B) and the actual operation (C) (Figs. 8, 10). Pathfinder also helps the operator avoid driving on critical side slope (Fig. 11). However, when the tool is used as a decision support tool for scarifier operators, complicated routes can be hard to follow due to the real-time avoidance of obstacles, such as boulders or wind thrown trees, which shift the routes laterally. The general directions and spacing between the

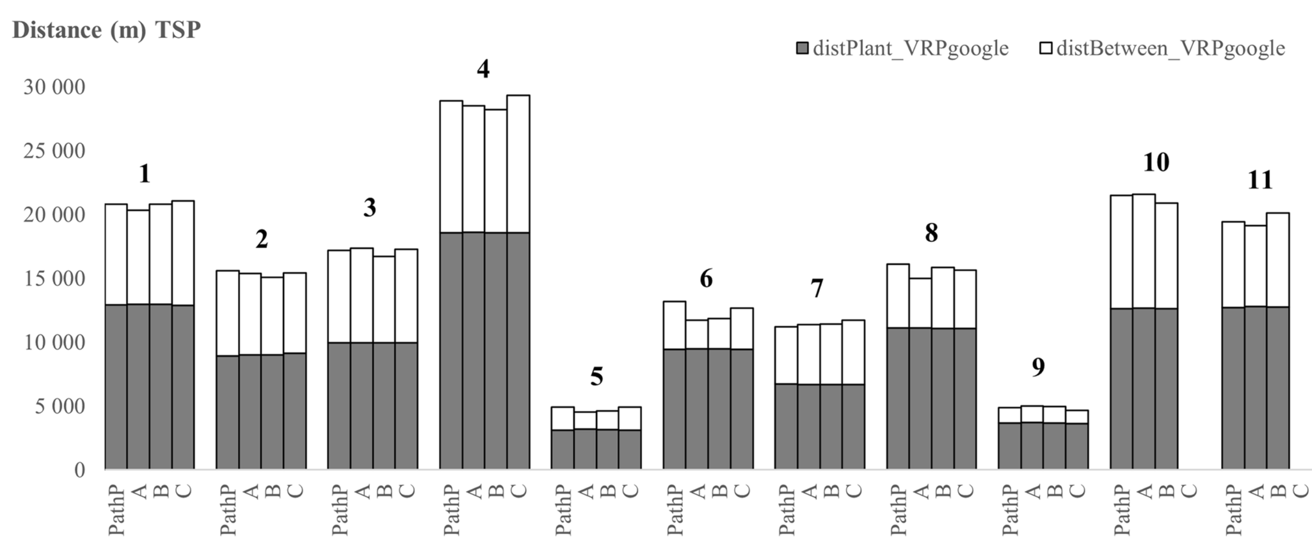

Fig. 10 Total driving distance, divided into distance planting (distPlant) and distance not planting (distBetween) across the 11 different sites, comparing the VRP (Vehicle Routing Problem, routes with a maximum capacity of 1 500 seedlings) solution from Pathfinder (*PathP*) with solutions based on driving directions from operators A, B, and C

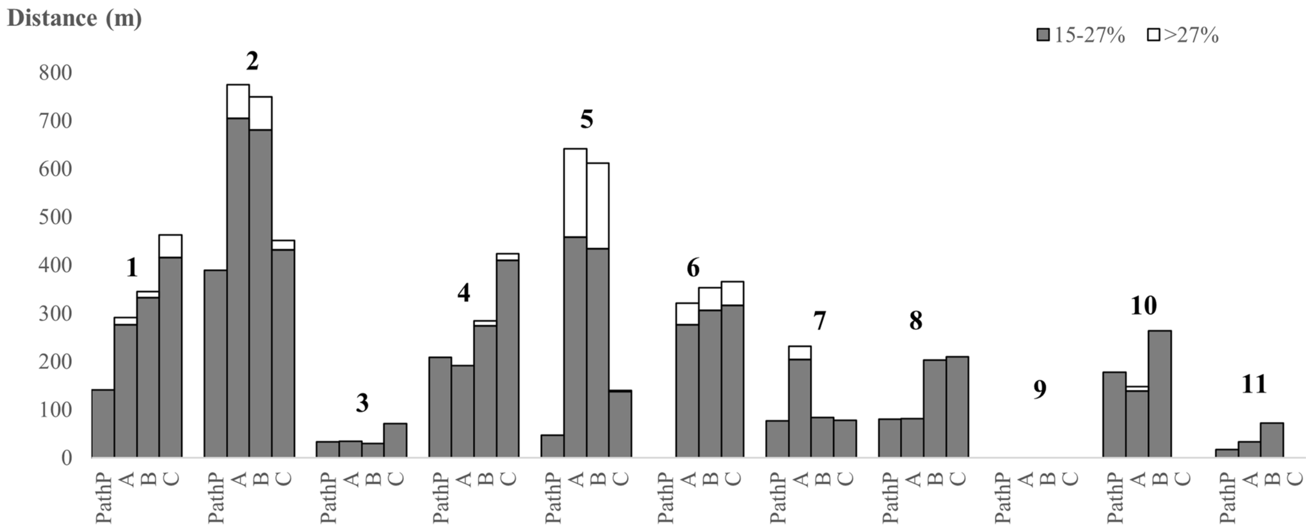

Fig. 11 Driving distance on side slopes (15–26.9% or ≥ 27%) across the 11 different sites, comparing the TSP (Travelling salesman problem, continuous route) solution from Pathfinder (*PathP*) with solutions based on driving directions from operator A, B, and C

Table 6 Average number of turns (n) in different time penalty classes (seconds), average total number of turns, and average time penalty per ha produced by the four solutions (*PathP*, operator A, B, and C); significant differences compared to *PathP* are indicated by asterisks, with no asterisk indicating no significant difference; note that there is no data from operator C for site 10 and 11

Solution (TSP)	0–4.9 s	5.0–9.9 s	10.0–29.9 s	30.0–54.9 s	55.0–79.9 s	≥ 80.0 s	Avg. Tot. turns	Avg. time penalty per ha
PathP	265	7	85	2	8	0	367	400
A	226**	3*	71**	2	9	0	311**	343*
B	213***	1**	73*	2	9	0	298***	341*
C	215	5	74	2	9	0	305	399

*, $p < 0.05$; **, $p < 0.01$; ***, $p < 0.001$

routes may suffice as decision support when an operator is steering the machine.

The *PlantP* module improves area estimation of the regeneration area and facilitates ordering a more correct number of seedlings compared to using the harvested gross area. In forestry, wrong area estimation leads to problems when planning the scarification work, especially for operators that only get paid for the area prepared. The more correct area estimation by the *PlantP* module helps give a more correct basis for the contractors.

The *PlantP* module could be improved by excluding wet areas entirely, not only the stems in that area, as part of or the whole area is included again by the 12 m inward buffer. In this study, the resolution of the soil moisture map was too coarse (10 m × 10 m), and it would be better to use, for example, the Depth to Water maps with a resolution of 1–2 m. In this study, peat soil was subtracted *after* the area was calculated from the hpr-file, but not all that area is normally excluded from regeneration. In this case, it would probably be better to only exclude the cut stems in those areas while retaining the buffers (12 m inwards and 6 m outwards). All threshold values of, e.g., soil moisture, root rot frequence, soil type, plant species and density at different site index values can be adjusted by the user to better fit the local conditions as well as user preferences.

The *PathP* module can be configured according to the planner's preferences and objectives through different parameter selections. The *PathP* module could be improved by allowing segments of driving lines to be included in a VRP route and/or by adding a capacity constraint earlier in the solution approach to avoid driving lines that exceed, for example, half the capacity. This would enable better utilisation of the full capacity of the machine. Another improvement is to introduce curved driving lines and smoother turns to allow a machine to follow the given route more precisely.

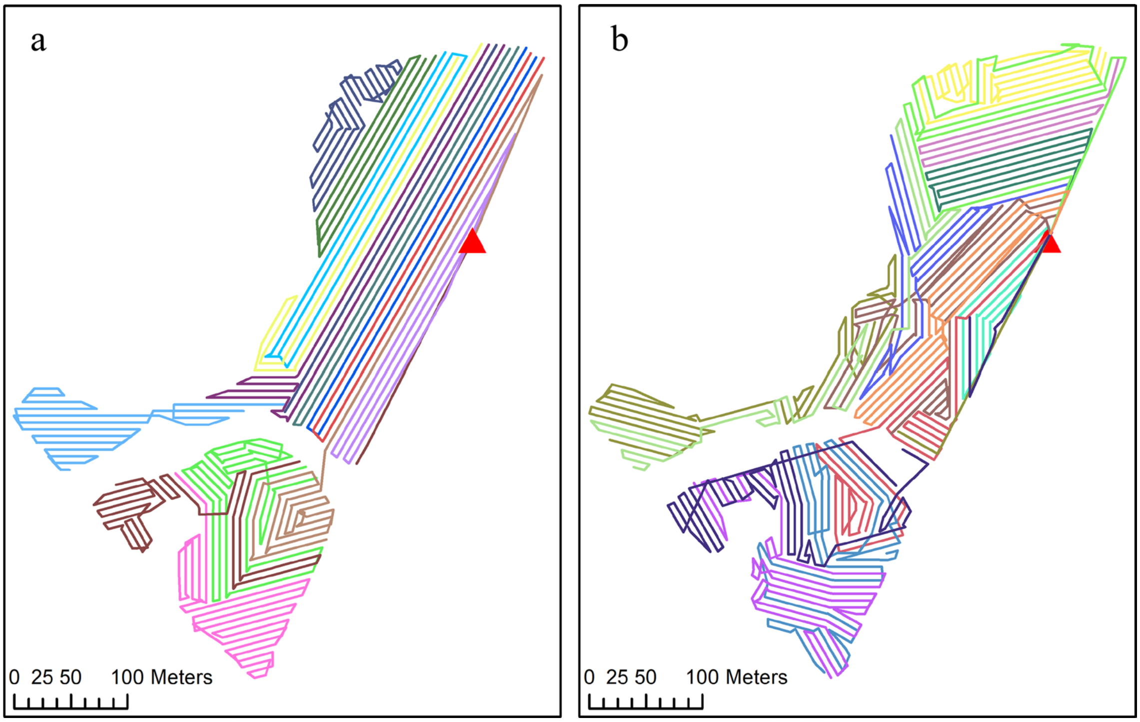

Fig. 12 Vehicle routing problem (VRP) for the capacity restrained routes (maximum 1500 seedlings when the threshold for side-slope driving is 27% (**a**) and 15% (**b**))

The complexity of routing is affected by the chosen threshold at which driving on side slopes should be avoided (Fig. 12). The threshold at which driving is possible on side slopes is affected by the frequency of surface obstacles such as boulders (Malmberg et al. 1980). If there are few surface obstacles, a threshold of 27% could be reasonable for a machine the size of a forwarder but, as the number of large obstacles increases the threshold should be lowered (Malmberg et al. 1980). In our case, the 27% threshold provided a similar solution for route directions to the directions and subareas chosen by operators A, B and C, who were all experienced at working with large scarification or planting machines (Fig. 8). For a smaller autonomous vehicle, such as the terrain vehicle platform used in the Autoplant project (Hansson et al. 2024), a threshold of 15% or lower is probably more realistic.

It is difficult to make quantitative comparisons with manually planned and executed operations. In our case, we have attempted this by asking operators to indicate the preferred driving directions. Then, we use the proposed *PathP* software for the routing part (Step 3–4) to compute all KPI used for the comparisons. This enables us to compare the impact of selecting the driving directions. However, as we solve the routing to optimality, it is not possible to see the impact of better routing. In future analysis, we would collect more GNSS-measured tracks for more detailed comparisons. It is also difficult to compare with other decision support systems under development. Here, it would be interesting to create a set of standardised cases with all information available in a standard format and defined times for different turning options.

Conclusion

The proposed Pathfinder supports ordering the correct numbers and right species of seedlings. It also presents a better area estimation for regeneration. The continuous routes provided are similar to the routes presented by skilled operators, but with less side-slope driving on critical slopes. Pathfinder is an important building block in the system of an autonomous site preparation and planting machine.

Acknowledgements We would like to thank Holmen, SCA, Södra, Sveaskog, StoraEnso, and Mellanskog for sharing their harvest and regeneration data, the language editor Leslie Walke, and Helena Gålnander and Victoria Forsmark for assistance in the work.

Data availability Raw data are available upon request. The Pathfinder code is not open source, but it is possible to set up a similar model by following the steps and equations in this paper.

Declarations

Conflict of interest The authors declare there are no competing interests.

References

Anon (2021a) StanForD 2010: modern communication with forest machines. Technical report, Skogforsk, Uppsala

Anon (2021b) Soil Types [parent material maps] 1:25 000-1:100 000: Geological Survey of Sweden

Audy JF, Rönnqvist M, D'Amours S, Yahiaoui AE (2023) Planning methods and decision support systems in vehicle routing problems for timber transportation: a review. Int J For Eng 34(2):143–167. https://doi.org/10.1080/14942119.2022.2142367

Flisberg P, Forsberg M, Rönnqvist M (2007) Optimization based planning tools for routing of forwarders at harvest areas. Can J For Res 37(11):2153–2163. https://doi.org/10.1139/x07-065

Flisberg P, Rönnqvist M, Willén E, Frisk M, Friberg G (2021) Spatial optimization of ground-based primary extraction routes using the BestWay decision support system. Can J For Res 51(5):675–691. https://doi.org/10.1139/cjfr-2020-0238

Flisberg P, Rönnqvist M, Willén E, Forsmark MV, Davidsson A (2022) Optimized locations of landings in forest operations. Can J For Res 52(1):59–69. https://doi.org/10.1139/cjfr-2021-0032

Forsmark V, Johannesson T (2020) Skogsvårdsföretagens rekrytering: förusättningar, nuläge och konsekvenser [Silvicultural companies recruitment]. Arbetsrapport 1039–2020, Skogforsk, Uppsala, pp 32. (in Swedish)

Friberg G, Jacobson S, Möller JJ, Bhuiyan N, Willén E (2019) Föryngringsplanering med hjälp av skördarinformation och geodata [Regeneration planning using harvester information and geodata]. Arbetsrapport 1002–2019, Skogforsk, Uppsala, pp 16. (in Swedish)

Hägglund B, Lundmark JE (1977) Site index estimation by means of site properties Scots pine and Norway spruce in Sweden. Studia Forestalia Suecica 138:38

Hansson LJ, Forsmark V, Flisberg P, Rönnqvist M, Mörk A, Jönsson P (2022) A decision support tool for forwarding operations with sequence-dependent loading. Can J For Res 52(12):1513–1526. https://doi.org/10.1139/cjfr-2022-0011

Hansson LJ, Sten G, Rossander M, Lideskog H, Manner J, van Westendorp R, Li SY, Eriksson A, Wallner A, Rönnqvist M, Flisberg P, Edlund B, Möller B, Karlberg M (2024) Autoplant—autonomous site preparation and tree planting for a sustainable bioeconomy. Forests 15(2):263. https://doi.org/10.3390/f15020263

Holmström E, Nikander J, Backman J, Väätäinen K, Uusitalo J, Jylhä P (2023) A multi-objective optimization strategy for timber forwarding in cut-to-length harvesting operations. Int J For Eng 34(2):267–283. https://doi.org/10.1080/14942119.2022.2149003

Hosseini A, Lindroos O, Wadbro E (2019) A holistic optimization framework for forest machine trail network design accounting for multiple objectives and machines. Can J For Res 49(2):111–120. https://doi.org/10.1139/cjfr-2018-0258

Hosseini A, Wadbro E, Ngoc Do D, Lindroos O (2023) A scenario-based metaheuristic and optimization framework for cost-effective machine-trail network design in forestry. Comput Electron Agric 212:108059. https://doi.org/10.1016/j.compag.2023.108059

Malmberg CE, Hansen R, Svensson A (1980) Körning i brant terräng [Driving in steep terrain]. Forskningsstiftelsen Skogsarbeten, Oskarshamn. (in Swedish)

Möller JJ, Arlinger J, Barth A, Bhuiyan N, Hannrup B (2011) Ett system för beräkning och återföring av skördarbaserad information till skogliga register-och planeringssystem [A system for calculation and feedback of harvester-based information to forestry planning systems]. Arbetsrapport 2011-756, Skogforsk, Uppsala, pp 55. (in Swedish)

Offenbacher C (2024) Segmentation Algorithm. Technical Report, Skogforsk, Uppsala, pp 14

Olsson B, Ledwith M (2020) National Land Cover Database (NMD) - product description. Swedish Environmental Protection Agency, Stockholm

Ovaskainen H, Riekki K (2022) Computation of strip road networks based on harvester location data. Forests 13(5):782. https://doi.org/10.3390/f13050782

Ramantswana M, Guerra SPS, Ersson BT (2020) Advances in the mechanization of regenerating plantation forests: a review. Curr For Rep 6(2):143–158. https://doi.org/10.1007/s40725-020-00114-7

Sörensen R, Johansson F, Gålnander H (2023) Ökad skogsproduktion och förbättrad miljöhänsyn: genom anpassning till lokala förutsättningar [Increased forest production and improved conservation measures: through adaptation to local conditions]. Arbetsrapport 1175-2023, Skogforsk, Uppsala, pp 67. (in Swedish)

Swedish Forest Agency (2024) Nedladdning av geodata [Download Geo Data]. https://www.skogsstyrelsen.se/sjalvservice/karttjanster/geodatatjanster/nerladdning-av-geodata/. Accessed on 2024

Willén E, Johansson F, Jacobson S, Keskitalo C, Friberg G (2021) Kartering av skog på felaktig ståndort—Studie med nationellt tillgängliga geodata [Mapping of forests with improper site indexes - study using nationally available geodata]. Arbetsrapport 1091-2021, Skogforsk, Uppsala, pp 37. (in Swedish)

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