Optimization of forest management strategies using clustering method and mathematical programming: a case study in Misiones, Argentina

María Emilia Dussel[1], Frank Piedra-Jimenez[1], Juan M. Novas[2,3], and María Analía Rodríguez[1*]
analia.rodriguez@unc.edu.ar
https://ipqa.unc.edu.ar/

[1] Instituto de Investigación y Desarrollo en Ingeniería de Procesos y Química Aplicada (IPQA), Universidad Nacional de Córdoba-CONICET, Córdoba 5000, Argentina
[2] Centro de Investigación y Estudios de Matemática (CIEM), Universidad Nacional de Córdoba-CONICET, Córdoba 5006, Argentina
[3] Centro de Investigación, Desarrollo y Transferencia de Sistemas de Información (CIDS) Universidad Tecnológica Nacional- Facultad Córdoba, Córdoba 5000, Argentina

Abstract. In this study, a novel decision-making approach is proposed for the forest management planning process. Even in small-scale cases of study, the relationship between the dataset size and the complexity of mathematical optimization models (in terms of constraints and variables) is factorial, resulting in exponential increases in computational complexity. Thus, while acknowledging large size and realistic data is crucial to account for reasonable conclusions, it is also a challenge itself. Hence, a procedure is proposed to approach this strategic problem. First, random data is generated to assume an ongoing forest inventory. Second, data is processed applying three successive grouping steps to enhance the utilization of large datasets. Within this stage, clustering techniques are applied using the Scikit-learn library for a large group of stands with several characteristics. Last, a mathematical framework is presented, rooted in Generalized Disjunctive Programming (GDP) and reformulated as a Mixed Integer Linear Programming (MILP) model, to address optimal forest management strategy to maximize the net present value (NPV). The MILP model is implemented in Pyomo library in Python and solved using GAMS-CPlex. The feasibility of the proposed model is assessed using data obtained from the Desarrollo Foresto Industrial web page of the Secretaría de Agricultura, Ganadería y Pesca (Ministerio de Economía de la República Argentina). Computational analysis demonstrates the versatility of the framework as a decision-making tool, highlighting its ability to generate diverse and viable solutions for forest management.

Keywords: Forestry Planning · Generalized Disjunctive Programming · Clustering Method

* Corresponding author.

1 Introduction

Forestry planning is crucial due to the long-term impact of the decisions involved. Effective forest management practices are essential for achieving profitability while satisfying heterogeneous demand targets from multiple wood-based industries. Management plans typically contain basic information such as field description, history, management goals, and forest composition. Forests are often divided into stands (which are the silviculture units) or Forest Management Units (FMUs) for planning purposes, each with specific characteristics, like past activities, forest species, age, timber volume, and more. Management regimes, known as prescriptions, are established outlining activities and their scheduling for each stand or FMUs [1].

Forestry planning offers numerous benefits, particularly in optimizing profitability, reducing cost, enhancing efficiency, and achieving environmental goals. Several studies have focused on developing optimization models for forest management. Nguyen et al., create an interesting open-source tool named PRISM for optimizing and analyzing long-term forest management strategies. Developed in collaboration with the United States Forest Service (USFS), its objective is to integrate advanced technological tools to improve efficiency and effectiveness in sustainable forestry planning decision-making [2]. Troncoso et al., propose and compare decoupled and integrated forestry planning models to illustrate the impacts of a demand-driven and make-to-stock approaches in the forest industry [3]. Piedra-Jimenez et al. [4], propose a model to determine optimal forest management alternatives and timber assortment production.

Due to the huge number of stands in the forest planning problems, grouping techniques provides opportunities to reduce the complexity of the models. Smaltschinski et al. propose a minimum spanning tree method for clustering stands in the Eucalyp-tus plantations of Aracruz, Brazil. According to the authors, this would ultimately improve harvesting efforts when dealing with log transport and machinery deployment problems [5]. Olofsson et al. developed an algorithm to divide forest stands using Airborne Laser Scanner (ALS) data to calculate local height sites, which serve as centroids for Voronoi cells. These cells are subsequently merged to generate segments representing forest stands. This technique enables the delineation of forest patches [6].

In the present study, an original approach is proposed to determine the optimal management strategy for forest planning over a long-term horizon. This work extends a previous work from the research group [4] by adding some new features to capture in more detail the decision-making process: 1) random assignment prescriptions to stands are implemented to simulate a more realistic scenario at the beginning of the time horizon, 2) pre-processing of raw data is done to increase the capability to work with large datasets, 3) an analysis of previous demand for different log assortment is considered to assume a realistic forecast, and 4) the production of stands with ongoing random prescriptions is taken into account.

Given the problem complexity for large-size instances, the proposed approach consists of three parts. First, random data generation is carried out to test realistic scenarios using previous simulations [4]. Then, data processing is conducted to reduce the dimensions of the problems. For this purpose, successive clustering steps are applied and finally, the mathematical model is developed and tested.

The remainder of this study is structured as follows: In the next section, the problem statement is outlined. The proposed approach is then described in the third section. Section four presents the model formulation for the forest management plan. A case study is presented in the fifth section, followed by the demonstration of the results in the sixth section. Conclusions are drawn in section seven. The final part contains the nomenclature used in the model formulation.

2 Problem statement

When planning management operations for large forest areas it is usual to assume a forest inventory is already available. That is, at the beginning of the planning horizon, some stands are in a virgin state (new stands or those whose previous prescriptions have been completed), while others will be undergoing previously assigned prescriptions. As described in the Introduction, a prescription is an ordered sequence of forest operations (thinning and clear-cutting), each of which has a specific stand age of realization.

The objective is to define a forest management plan for a given large area to satisfy the demand of different wood-based products from the industry. For this purpose, the plan defines the assignment of prescriptions to each stand or FUM once the ongoing prescription is finished, or from the beginning if the stand is in a virgin state. In this study, a 40-year horizon planning is selected to ensure that at least one prescription can be assigned to each stand.

In each forest operation, different products are obtained. Typically, these products are defined by log length and diameter. To estimate timber yields in each forest operation within each prescription, growth simulations are used. In this study, five different products are considered.

Finally, it is important to note that each operation incurs in management costs depending on the prescription and tree species, while each log assortment has a corresponding demand that must be satisfied and a selling price.

Considering all these factors, the following approach is proposed.

3 Proposed approach

3.1 Random data generation

The original dataset consists of thousands of stands with diverse information, including area (ha), coordinates, year of plantation, species, etc. Given that the dataset provides the plantation year, a set of prescriptions is randomly assigned to each stand to estimate a forest inventory based on a real-size scenario. After this random assignment, the forest inventory is formed by a set of virgin stands and others that will complete their prescriptions within the horizon planning.

3.2 Data processing

Given the large dataset considered and to guarantee convergence of the model without exceeding the computing capacity, three successive grouping steps are proposed. The main idea is that decisions regarding forest management can be defined for each cluster instead of for each stand, thus, reducing the model size. To guarantee coherence in the decision process, the following steps take into account specific domain characteristics of the forest management aspects:

1. The dataset is grouped according to soil type. This first grouping method is based on the following idea: stands with the same soil type will have similar growing patterns. Thus, it is assumed that each forest management plan can be independently proposed for each type of soil.
2. Each group of soil dataset, is divided according to the year of clear-cutting of previous prescriptions (those that were randomly assigned). This second grouping method is based on the availability of the stands and the capability to assign prescriptions that will start in the same year for all stands within each group.
3. Finally, subgroups obtained from the previous step were further clustered based on the coordinates of each stand applying an unsupervised clustering method. For this purpose, a couple of techniques were evaluated. While there are several metrics to evaluate clustering algorithms, the most common ones are the Silhouette [8], Davies-Bouldin [9], and Calinski-Harabasz [10] indices. The three metrics consider distances between points within clusters and between clusters.
   The Calinski-Harabasz index can range from zero to infinity, where higher values indicate better clustering. While useful for comparing algorithms or cluster quantities, it's challenging to determine if a clustering is inherently "good"' or "bad" based solely on this index. Similarly, the Davies-Bouldin index can take the same range of values, but a high value means poorer clustering (i.e., if the index tends to zero, the clustering is better). However, it's unclear whether a Davies-Bouldin index of 1.2 indicates satisfactory clustering. Conversely, the Silhouette index ranges from -1 to 1. Higher values indicate better clustering, while values close to zero imply poorly defined clusters, with points being interchangeable. Values nearing -1 suggest erroneous clustering.
   The Silhouette index's bounded nature makes it more appropriate to interpret the quality of the proposed clusters.

The objective of this approach is to be able to run a model for each soil type classification and assign prescriptions to clusters instead of stands as mentioned above.

3.3 Forestry planning

The decision-making problem aims at defining the optimal forest management plan for a widespread forest area to maximize the Net Present Value (NPV).

To achieve this, the decision-maker needs to define the optimal scheduling of forest management alternatives (combination of forestry operations) for each cluster and the timber assortment production.

To model this problem the following assumptions are considered:

1. The soil and climate characteristics of all stands assigned to a given cluster are similar, allowing them to be treated as the minimum unit for assigning a prescription.
2. The set forest management prescriptions available for each cluster are known.
3. Clear-cutting is considered the final forestry operation in all forest management prescriptions.
4. A forest management prescription is considered complete when its final prescribed forestry operation is finished.
5. Given that the forest inventory already includes stands with pre-assigned prescriptions at the beginning of the planning horizon, operations associated with these prescriptions will continue to produce products until they are completed. This production must be accounted for in the final product demand, resulting in a product requirement equal to the maximum between zero and the difference between demand and pre-established production.
6. The requirement of logs must be strictly satisfied by own or third-party production.

4 Model formulation

4.1 Generalized Disjunctive Programming (GDP) model

The decision-making problem is proposed as a GDP formulation model [7] in equations (1) to (9). In Section 8, is described the nomenclature used.

The objective function is to maximize NPV, as follows:

$$NPV = \sum_{t \in T} \frac{1}{(1+\rho)^{t-1}} \left(\sum_{\substack{o,i,p,c,e \\ (i,p) \in IP \\ (c,p) \in CP \\ (p,e) \in PE}} (\epsilon_o - \alpha_i)\, V_{o,i,p,c,e,t} - \sum_{o \in O} (\epsilon_o - \delta_o)\, V_{o,t}^{third} \right). \tag{1}$$

The factor at right-hand side represents, in the first addend of the parenthesis, the price from sales minus the forest operation costs and, in the second addend, the price of products minus the costs of buying from third-parties. These factors are multiplied by the volume of products obtained from forest operations and bought from third-parties, respectively. The present value is obtained by dividing these terms by the updated factor, considering the interest rate ρ, and the corresponding period t.

$$\begin{bmatrix} Z_{i,p,c,e,t} \\ V_{o,i,p,c,e,t} = \beta_{c,e,p,i,o} \cdot \lambda_c \end{bmatrix} \veebar \begin{bmatrix} \neg Z_{i,p,c,e,t} \\ V_{o,i,p,c,e,t} = 0 \end{bmatrix}, \tag{2}$$

$$\forall\ (i,p) \in IP,\ (c,p) \in CP,\ (p,e) \in PE,\ o \in O,\ t \in T.$$

Equation (2) presents a disjunction where a boolean variable $Z_{i,p,c,e,t}$ is introduced to determine whether forest operation i of forest prescription p and tree species e is executed in cluster c during time period t. If the decision is "yes", $Z_{i,p,c,e,t} = true$ and the equation defines the volume of timber assortment o harvested ($V_{o,i,p,c,e,t}$), which equals the yield of timber assortment harvested using forest operation i prescribed in p ($\beta_{i,p,o}$) multiplied by the area of cluster c (λ_c). If the decision is "no", $Z_{i,p,c,e,t} = false$ and the volume of timber assortment o harvested is equal to zero.

$$Y^{start}_{p,c,e,t} \Rightarrow \neg Y^{start}_{p',c,e',t'}, \tag{3}$$

$$\forall\ p \neq p',\ (p,e),(p',e') \in PE,\ (c,p),(c,p') \in CP,\ t,t' \in T, t \leq t' \leq t + \gamma_{c,p} - 1.$$

$$Y^{start}_{p,c,e,t} \Rightarrow \neg Y^{start}_{p,c,e',t'}, \tag{4}$$

$$\forall (p,e),(p,e') \in PE,\ (c,p) \in CP,\ t,t' \in T, t < t' \leq t + \gamma_{c,p} - 1.$$

Constraints (3) and (4) prevent overlapping prescriptions. Another boolean variable is introduced: $Y^{start}_{p,c,e,t}$. When this variable is true, it allows to start a prescription p in the cluster c with tree specie e in time period t. If $Y^{start}_{p,c,e,t} = false$, the prescription cannot start. Constraint (3) prevents overlapping between different prescriptions in the same cluster, and constraint (4) prevents overlapping a given prescription in a cluster while it is performed. Parameter $\gamma_{c,p}$ is introduced to establish the duration of the forest management prescription p for specie e.

$$Y^{start}_{p,c,e,t} \Leftrightarrow Z_{i,p,c,e,t'}, \tag{5}$$

$$\forall\ (i,p) \in IP,\ (c,p) \in CP,\ (p,e) \in PE,\ t,t' \in T, t' \leq |T|, t' = t + \mu_{c,p,i} - 1.$$

Constraint (5) establishes that the initiation of a forest management prescription implies that all available forestry operations within the planning horizon must be scheduled and vice versa, if a forest operation is executed it implies that a forest management prescription started. Parameter $\mu_{c,p,i}$ is introduced to establish the moment the forestry operation i of the prescription p must be performed when the prescription begins in period t.

$$\neg Y^{start}_{p,c,e,t}, \qquad \forall\ c \in C,\ p \in P,\ (p,e) \in PE,\ t \in T,\ t < \tau_c. \tag{6}$$

Constraint (6) prevents prescription p from starting while another prescription, assigned before the planning horizon started, is still ongoing, where parameter τ_c determines the t when previous prescription finishes.

$$\neg Z_{i,p,c,e,t}, \tag{7}$$

$$\forall\ (i,p) \in IP,\ (c,p) \in CP,\ (p,e) \in PE,\ t \in T,\ t \leq \mu_{c,p,i} - 1.$$

Constraint (7) prevents a certain forestry operation i from being scheduled in a previous time period than the one established in its forest management prescription p, where parameter $\mu_{c,p,i}$ establishes the period of time that the forest operation i must be scheduled after the prescription started.

$$Y^{end}_{p,c,e,t} \Leftrightarrow Z_{i,p,c,e,t}, \tag{8}$$

$$\forall\ (i,p) \in IP,\ (c,p) \in CP,\ (p,e) \in PE,\ t \in T,\ i = |I|.$$

Constraint (8) introduces boolean variable $Y^{end}_{p,c,e,t}$ that determines that a forest prescription p ends in a certain time period t if and only if its last operation i associated with the prescription p (i.e., clear-cutting) is scheduled in the same time period t.

$$\sum_{\substack{(i,p)\in IP\\(c,p)\in CP\\(p,e)\in PE}} V_{o,i,p,c,e,t} + V^{third}_{o,t} \leq \max\left\{0, \omega_{o,t} - \nu_{o,t}\right\}, \qquad \forall\ t \in T,\ o \in O. \tag{9}$$

Finally, constraint (9) ensures the fulfillment of log demand by the own production and third-party. Note that given that there is a number of ongoing prescriptions when the planning horizon starts, the corresponding timber assortment obtained from the established management operations ($\nu_{o,t}$) are discounted from the demand parameter ($\omega_{o,t}$).

5 Case study

5.1 Data generation

The dataset for this work was obtained from the Desarrollo Foresto Industrial web page of the Secretaría de Agricultura, Ganadería y Pesca (Ministerio de Economía de la República Argentina). The study area consists of 44,437.38 ha of *Pinus elliottii* and *Pinus taeda* composed of 10,141 stands located in the province of Misiones (Argentina). The dataset includes the geo-location of the stands, their area, and the plantation year (from 2000-2015). In order to assume a realistic forest inventory, 90 prescriptions were randomly assigned to each stand using the plantation year as the initial moment of the assigned prescription. These assignments simulate a realistic forest planning scenario, formed by a set of virgin stands and a set of stands that are performing a previous prescription at the beginning of the horizon planning. To reduce the size of the problem, stands are grouped based on three step grouping method, described below.

5.2 Data processing

The following method is proposed for reducing the dataset size:

1. *Soil type grouping.* A Geographical Information System (GIS) map, obtained from Land and Water Development Division, FAO, Rome, depicting different soil type was utilized in QGIS (Quantum GIS) software. Each stand was labeled according to its soil type (the first two letters) and several composition characteristics, resulting into seven categories:
   1.1. Humic Ferralsols (Fh1-3a) with 3531 stands,
   1.2. Rhodic Ferralsols (Fr4-3b) with 1 stand,
   1.3. Luvic Phaeozems (Hl2-3c) with 4068 stands,
   1.4. Eutric Fluvisols (Je13-a) with 58 stands,
   1.5. Dystric Nitosols (Nd1-3b) with 1783 stands,
   1.6. Eutric Nitosols (Ne1-3a) with 450 stands, and
   1.7. Wet zones (WAT) with 250 stands.
2. *Clear-cutting grouping.* Stands grouped according to soil type were further subdivided based on the year of clear-cutting of randomly assigned prescriptions. Prescriptions completed between 2016 and 2023 were consolidated into the same group assuming they are virgin at the beginning of the planning horizon. This resulted in the subdivision of each soil category into 1 to 18 subgroups according to the number of different years of pre-assigned prescriptions ending, yielding a total of 103 subgroups.
3. *Coordinates grouping.* Finally, the last grouping method applied is based on clustering method using the Scikit-learn library. This method consists of clustering each previous subgroup according to the coordinates of each stand provided by the dataset.
   Two clustering methods from the Scikit-learn library, K-Medoids and K-Means [11], were compared. The comparison involved calculating the Silhouette index for cluster numbers ranging from 9 to 14 for each algorithm in every clustering task. Across all evaluations, the K-Means algorithm consistently achieved higher scores compared to K-Medoids considering the Silhouette index.
   Once the algorithm was chosen, the number of clusters per group was determined, employing the Silhouette index once again. The optimal cluster quantity was chosen from a range between 9 and 14. Ultimately, a total of 296 clusters were generated, with 55 clusters for Fh1-3a soil, 1 cluster for Fr4-3b soil, 45 clusters for Hl2-3c soil, 14 clusters for Je13-a soil, 53 clusters for Nd1-3b soil, 74 clusters for Ne1-3a soil, and 54 clusters for WAT soil.

5.3 Mathematical model

In this case, a set of prescriptions can be assigned to the set of clusters generated in the previous step. All available prescriptions consider two pre-commercial thinning and a clear cutting which is the last operation. The first one occurs at stand age of 7. The second one occurs on years: 11, 12 and 13. The clear-cutting age ranged from 16 to 25 years. The combination of these alternatives results in 30 prescriptions (1 -1st thinning-,3 -2nd thinning-, 10 -clear cutting-). As it was mentioned, a planning horizon of 40 years was assumed. Additionally, *Pinus elliottii* was the specie taken into account.

Five different log assortments were considered: logs ranging from 8 to 14cm of diameter (for pulp), 14 to 18 cm, 18 to 24 cm, 25 to 29 cm and larger than 30 cm. Timber classification for energy purposes was excluded from the modeling process under the assumption that such products would be obtained from forest post-thinning operations or from clear-cutting residues.

Demand forecast for these products was derived from linear regression analysis, based on consumption patterns between 2000 and 2022. This information was sourced from the Desarrollo Foresto Industrial web page of the Secretaría de Agricultura, Ganadería y Pesca (Ministerio de Economía de la República Argentina) encompassing total consumption activities such as sawmills, compensated, fiber and particle board, and cellulose production. Proportional consumption for each activity was estimated from annual industry survey reports, based on the raw material composition of Pinus logs and cellulose. Additionally, the percentage of demand fulfilled by Misiones was used to adjust product demand, providing a more realistic scenario. Another consideration in this study was the production from randomly assigned prescriptions. This allows to estimate, based on the area of each stand, the moment of the intervention and the simulated yield, the expected production for each product when every forestry operation is performed. As mentioned, the product requirement is calculated as the maximum between zero and the difference the forecasted demand and the expected production from ongoing prescriptions.

Lastly, the group of Je13-a soil was selected to show the model results in the next section.

6 Results

In clustering analysis, both the K-Means and K-Medoids algorithms were applied to the set of clusters, representing the total clustering result across all soil types. The mean Silhouette score for each algorithm was found to be 0.586 for K-Means and 0.460 for K-Medoids. While these scores may appear low, they can be attributed to the relatively small size of certain clusters. In cases where few but far data points were not clustered separately, resulting in a single cluster, the Silhouette score could occasionally be negative or near to zero, contributing to the lower overall scores. The final result for the soil type clustering is shown in Figure 1.

After data processing, a considerable reduction in data size was achieved: from 10,141 stands to 296 clusters, demonstrating the efficiency of the method applied.

Under the proposed approach, the model should be executed for each type of soil. Due to space limitations, the results for the Je13-a soil type are presented. In Fig. 2(a), the clustering result is observed, where the clustering is determined by the year of clear-cutting in the last instance, while in Fig. 2(b), the clustering is determined by K-means algorithm. The optimization model was implemented using the Pyomo library in Python and solved with CPLEX 45.7.0 on an Intel(R) Core(TM) i7-7700K CPU @ 4.20GHz processor with 16.0 GB RAM. The computational performance revealed a CPU time resolution of 59.22 seconds, with a total of 8,047,187 constraints and 327,797 variables, including 81,900 binary variables and 245,897 continuous variables. The relative GAP was calculated to be 0.0, indicating optimal convergence, and the resulting Net Present Value (NPV) was determined to be 18,989,609 USD.

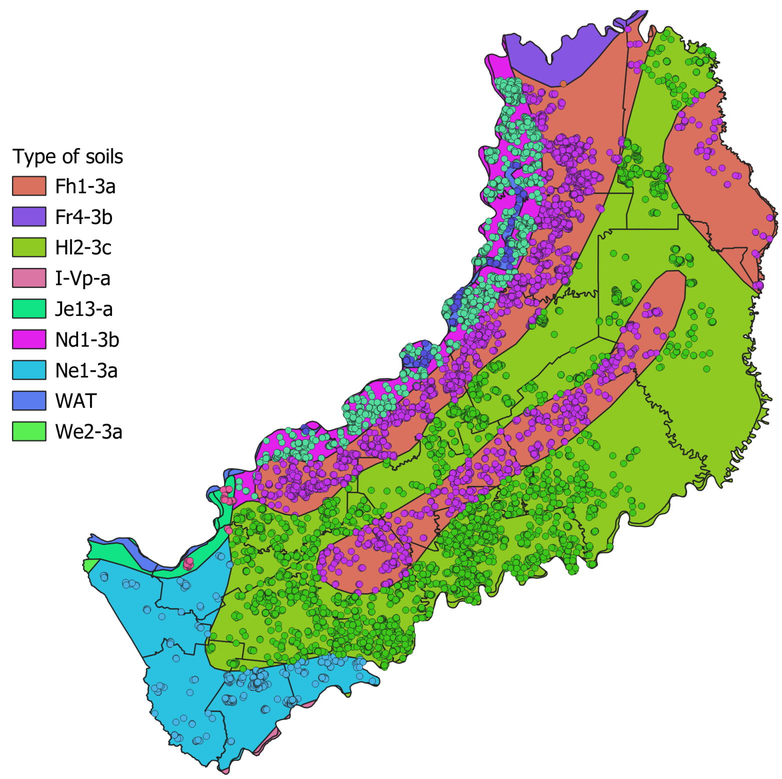
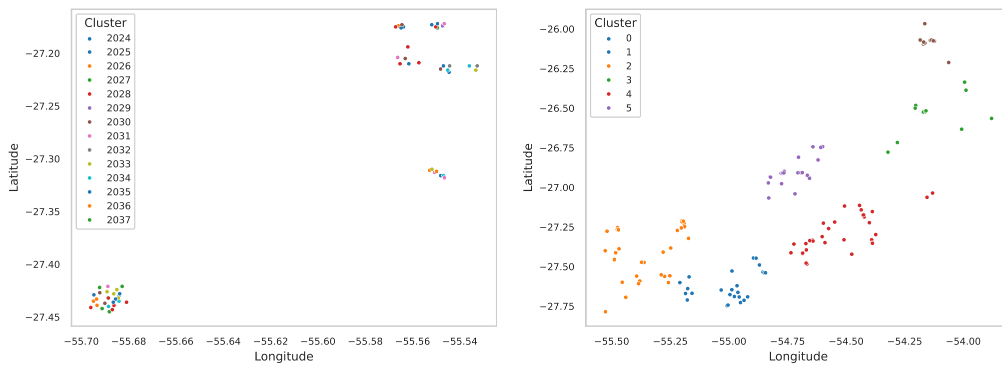


Fig. 1: Map of Misiones showing different types of soils and the stands labeled according to that characterization.

The optimal forest planning strategy is depicted in Figure 3, where labels "T1", "T2", and "CC" represent "First Thinning", "Second Thinning", and "Clear-Cutting", respectively. It can be observed that seven (from thirty) prescriptions were assigned to each cluster, with a preponderance of the shortest prescription.

Figure 4 illustrates the total volume produced by the assigned prescriptions, along with the volume supplied from third parties. It is evident that from $t = 15$ (2038), the production stabilizes. This stabilization can be attributed to the linearly estimated demand and the absence of production from randomly assigned prescriptions, which concluded at $t = 14$ (2037).



(a) Je13-a soil type.

(b) Hl2-3c soil type with clear-cutting in 2036.

Fig. 2: Clustering for Hl2-3c soil type (for clear-cutting in 2036) and for Je13-a soil type.

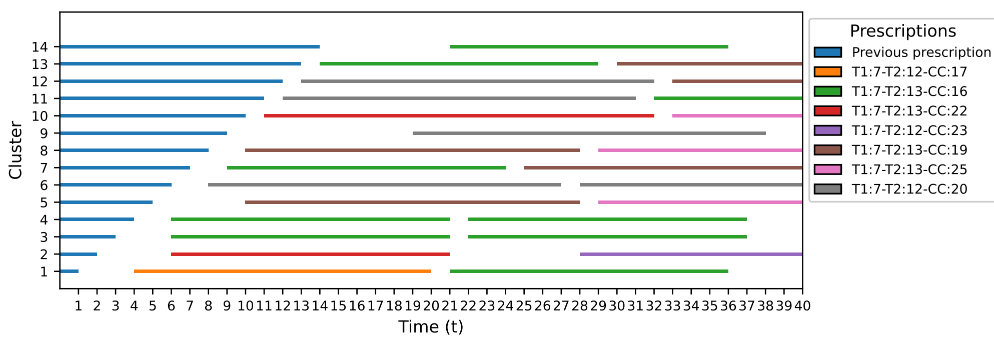


Fig. 3: Graph showing prescriptions assigned to each cluster.

Finally, Figure 5 illustrates the behavior of demand, production from assigned prescriptions, the log assortment bought by third-party, and the total production (the sum between assigned prescriptions production and third-party production). It is easy to observe that from t=15 forward, the demand is met with total production while before t=15, the demand is met with total production and production from randomly assigned prescriptions. The graph only displays total production, which appears to indicate overproduction from t=1 to t=15. This phenomenon is attributed to certain log assortments being overproduced due to randomly assigned prescriptions. This shows the importance of forestry planning.

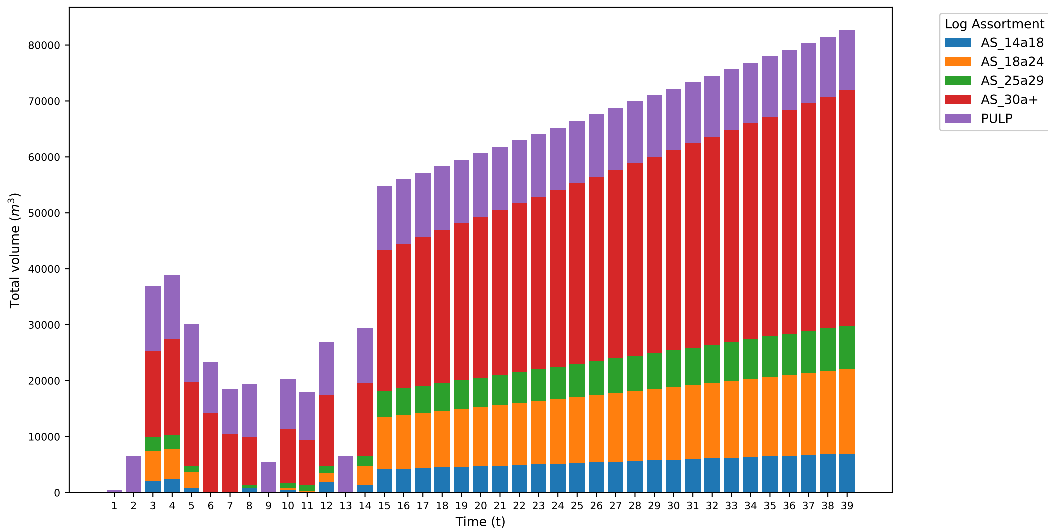


Fig. 4: Graph showing total volume of each log assortment produced during the time horizon.

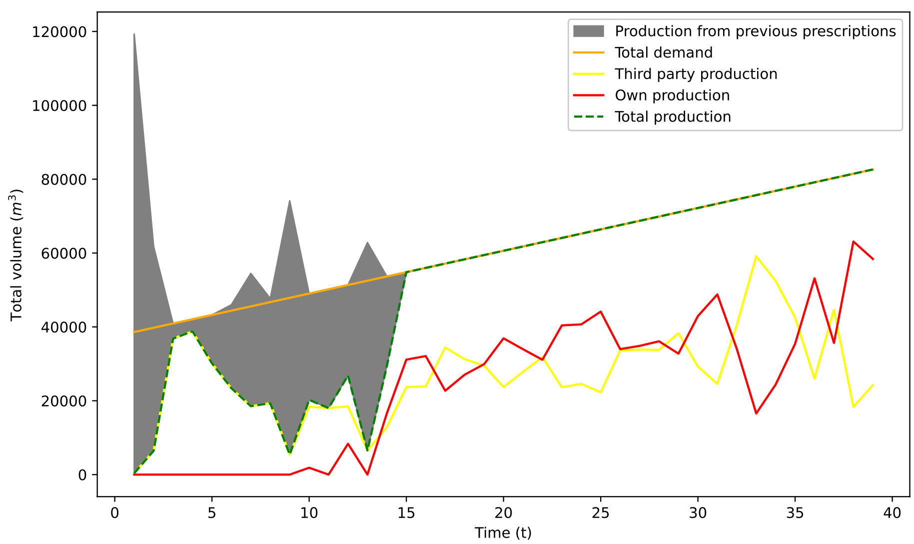


Fig. 5: Graph showing the volume of demand and production during the time horizon.

7 Conclusion

The application of the clustering method resulted in a substantial reduction of the dataset size, decreasing from 10,141 individual stands to just 296 clusters. This reduction to under 3% of the original size underscores the method's exceptional efficiency in managing large datasets and offers a solution to the challenges associated with working with large datasets.

The GDP model demonstrated optimal performance in long-term forest planning. This versatile mathematical programming model can be applied to various datasets and contexts. Future work will involve comparing disaggregated runs with integrated ones.

8 Nomenclature

Sets	**Description**
C	Forest clusters
P	Forest management prescription
I	Forest management operation
E	Tree specie
T	Time period
O	Log assortment
CP	Forest management prescription p assigned to cluster c
PE	Forest management prescription p that can be scheduled to tree species e
IP	Forest operation i prescribed in forest prescriptions p

Table 1: Description of sets of GDP and MILP models.

Parameters	**Description**	**Unit**
ρ	Annual discount rate	[%]
ϵ_o	Selling price of the log assortment o	$\left[USD\ m^{-3}\right]$
α_i	Unitary cost of forest operation i	$\left[USD\ m^{-3}\right]$
δ_o	Unitary cost of buying the the log assortment o	$\left[USD\ m^{-3}\right]$
$\beta_{c,e,p,i,o}$	Yield of log assortment o and forest operation i prescribed p for tree specie e and cluster c	$\left[ton\ ha^{-1}\right]$
λ_c	Area of cluster c	$[ha]$
$\nu_{o,t}$	Total volume production of log assortment o in time period t	$\left[m^3\ year^{-1}\right]$
$\gamma_{c,p}$	Length of prescription p for cluster c	$[year]$
$\mu_{c,p,i}$	Age established for scheduling the prescribed intervention i prescribed p for cluster c	$[year]$
$\omega_{o,t}$	Demand of log assortment o in time period t	$\left[m^3\right]$
τ_c	Time where cluster c is available to assign a prescription p	$[year]$

Table 2: Description of parameters of GDP and MILP models.

Variable	Description	Type of variable
$Y^{start}_{p,c,e,t}$	Equal to 1 if in time period t where prescription p starts in cluster c with specie e, 0 otherwise	Boolean
$Y^{end}_{p,c,e,t}$	Equal to 1 if in time period t where prescription p ends in cluster c with specie e, 0 otherwise	Boolean
$Z_{i,p,c,e,t}$	Equal to 1 if in time period t where operation i prescribed in p is carried out in cluster c with specie e, 0 otherwise	Boolean
$V_{o,i,p,c,e,t}$	m^3 of log assortment o produced in cluster c with specie e with forestry operation i prescribed in p	Nonnegative Reals
$V^{third}_{o,t}$	m^3 of log assortment o sent to third parties in time t $[ton]$	Nonnegative Reals

Table 3: Description of variables of GDP and MILP models.

References

1. Khanal, P. and Straka, T.: Fundamentals of Forest Resource Management Planning. *Clemson (SC): Clemson Cooperative Extension, Land-Grant Press by Clemson Extension* 2 (2020).
2. Nguyen, D., Henderson, E., Wei, Y.: PRISM: A decision support system for forest planning. *Environmental Modelling & Software,* **157**, 105515 (2022)
3. Troncoso, J., D'Amours, S., Flisberg, P., Rönnqvist, M., Weintraub, A.: A mixed integer programming model to evaluate integrating strategies in the forest value chain — a case study in the Chilean forest industry. *Canadian Journal of Forest Research*, **45**(7), 937–949 (2015)
4. Piedra-Jimenez, F., Broz, D., Novas, J. M., Rodríguez, M. A.: A MILP Decision Support Approach for the Optimal Forest Planning Management. In Proceedings of the 11th International Conference on Production Research – Americas, pp. 484–490, Springer Nature, Switzerland (2023)
5. Smaltschinski, T., Seeling, U., Becker, G.: Clustering forest harvest stands on spatial networks for optimised harvest scheduling. *Annals of Forest Science* **69**, 651–657 (2012)
6. Olofsson, K., Holmgren, J.: Forest stand delineation from lidar point-clouds using local maxima of the crown height model and region merging of the corresponding Voronoi cells. *Remote Sensing Letters* **5** (3), 268-276 (2014)
7. Raman, R., Grossmann, I. E.: Modelling and computational techniques for logic based integer programming. *Computers and Chemical Engineering* **18**(7), 563–578 (1994)
8. Rousseeuw, P. J.: Silhouettes: a graphical aid to the interpretation and validation of cluster analysis. *Journal of computational and applied mathematics* **20**, 53–65 (1987)
9. Davies, D. L., Bouldin, D. W.: A cluster separation measure. *IEEE transactions on pattern analysis and machine intelligence* **2**, 224–227 (1979)
10. Caliński, T., Harabasz J.: A dendrite method for cluster analysis. *Communications in Statistics-theory and Methods* **3**(1), 1–27 (1974)
11. MacQueen, J.: Some methods for classification and analysis of multivariate observations. In: *Proceedings of the fifth Berkeley symposium on mathematical statistics and probability*, pp. 281–297. Oakland, CA, USA (1967)