Bioresource Technology Reports 30 (2025) 102097

Contents lists available at ScienceDirect

Bioresource Technology Reports

journal homepage: www.sciencedirect.com/journal/bioresource-technology-reports

Optimizing the logistical routing of agricultural side-streams to a biogas plant for a circular bioeconomy implementation

Kasper Eloranta [a], Olli Koskela [a], Elina Tampio [b], Juho Kanniainen [c], Ulla A. Saari [d,*]

[a] Data and Business Analytics Team, HAMK Smart Research Unit, Häme University of Applied Sciences, Visamäentie 35 A, 13100 Hämeenlinna, Finland
[b] Biorefineries and bio-based fertilizers, Production Systems, Natural Resources Institute Finland (Luke), Latokartanonkaari 9, FI-0079 Helsinki, Finland
[c] Computing Sciences, Faculty of Information Technology and Communication Sciences, Tampere University, Korkeakoulunkatu 1, 33720 Tampere, Finland
[d] Industrial Engineering & Management, Faculty of Management and Business, Tampere University, Korkeakoulunkatu 1, 33720 Tampere, Finland

ARTICLE INFO

Keywords:
Circular bioeconomy
Logistic routing
Optimization
Biogas production
Simulation
Genetic algorithm

ABSTRACT

The study addresses a Vehicle Routing Problem (VRP) that incorporates elements of the Biomass Collection Problem (BCP), using geo-spatial biomass data to determine pickup site locations and their content levels. Optimization techniques are proposed for the Biomass Routing and Biogas Production Problem (BRBPP). Biomass potential data was sourced from a national open data service used in Finland. The developed methodology includes models used to simulate and optimize components within a circular supply chain for transporting biomasses to biogas plants for renewable energy production. The study executes spatial information mapping for biomasses near the biogas plant and proposes an optimized routing plan for the optimization problem, aiming to execute a cost-effective logistic routing for continuous biogas production. The main contributions of the study are the tool and methodology developed for geospatial mapping of extensive datasets alongside modeling and optimizing logistic routing for biomass collection and transportation to centralized biogas plants.

1. Introduction

The circular bioeconomy (CBE) offers promising prospects by leveraging waste management for energy production, particularly through the effective utilization of biomasses, i.e., organic side-streams (Bellezoni et al., 2022). Waste-to-energy systems, in general, enhance operational efficiency and cost-effectiveness, enabling industries to convert waste into energy in alignment with the circular economy (CE) principles (Monteiro and Ferreira, 2022). In addition, this type of systems, such as biogas production from nutrient-rich side-streams, also provides benefits through nutrient recirculation. However, while bioenergy production presents opportunities, it also raises concerns about its ecological footprint (Wang et al., 2020). This prompts the need to optimize resource flows in bioenergy production, ensuring maximal utilization of biomass side-streams without compromising environmental impact and emissions.

Optimizing resource flows and logistics is pivotal in adhering to CE principles. Efficient transportation, as observed in the case of animal manure utilized as fertilizer (Kamilaris and Prenafeta-Boldú, 2021), lessens resource pollution, increases manure valorization, and aligns with the practices of CE. However, transportation poses environmental and economic costs that demand consideration in optimization (Kamilaris and Prenafeta-Boldú, 2021). When configuring supply chains, balancing supply and demand while considering actors' roles is crucial (Yılmaz Balaman et al., 2018). Given the complexity of circular supply chains, employing digital and analytical tools becomes essential for modeling and optimizing these networks.

Robust digital capabilities and technologies are required for facilitating circular business models. Digital technologies offer opportunities for enhancing resource flows, value creation, and innovation, aligning with the CE practices (Ranta et al., 2021). They contribute to the CE in two key ways: (i) by enabling knowledge creation and sharing for decision-makers to optimize resource flows and value creation and (ii) by providing analytical tools, such as multi-objective decision-making algorithms for waste-to-energy supply chain optimization (Ranta et al., 2021; Yılmaz Balaman et al., 2018). These digital technologies drive various business model and process innovations, fostering alignment with the practices of CE across industries (Ranta et al., 2021). By embracing CE practices and principles and harnessing advanced computational and digital methods, humanity can work towards achieving carbon neutrality and a greener future.

Bijarchiyan et al. (2020) presented a holistic approach for a supply

* Corresponding author.
E-mail address: ulla.saari@tuni.fi (U.A. Saari).

https://doi.org/10.1016/j.biteb.2025.102097
Received 28 October 2024; Received in revised form 16 February 2025; Accepted 11 March 2025
Available online 14 March 2025

chain model where they considered the interactions for energy processes with the society in a regional setting. This study concentrates on leveraging biogas-based bioenergy production and recycling resource flows in the Kanta-Häme region in Finland. The aim involves connecting a bioenergy facility (i.e., a biogas plant) with local farms (i.e., producers of agricultural side-streams) to promote collaboration and optimize logistical operations, including investment planning for an upcoming centralized biogas plant. The study focuses on modeling and optimizing the logistical routing of biomass side-streams collection, which will be utilized by the centralized biogas plant as feedstocks. These biomasses consist of livestock manures, grasses, and straws. Spatial data mapping of biomass potentials located near the biogas plant is conducted for modeling and optimizing the logistical routing of biomass collection. The study seeks to provide insights into organizing biomass routing for a specific biogas plant investment case, aiding their decision-making processes. Additionally, it aims to compare optimized routing results under two simulation scenarios, one considering consistent biomass quality in terms of energy potential over time, and the other incorporating time-critical factors. Unlike Bijarchiyan et al. (2020), this study omits the societal aspects in the model, but, instead, in a real-world like case, consider the economic effects of the biological processes in the quality of the biomasses and respective implications to storage requirements in the network.

The research is structured around two main research questions. The first question (RQ1) focuses on organizing the routing of biomass collection to align logistics with CBE practices and dynamically optimize these routes: How should the routing of collection of biomasses be organized to align with CE practices and dynamically optimize the logistics based on specific criteria? The second question (RQ2) delves into enhancing logistics optimization by considering factors that were previously overlooked, particularly the impact of decreasing biomass quality over time on the optimal routing solution: How does the consideration of diminishing biomass quality over time affect the optimal routing solution compared to the currently optimized result? In the study, the research questions are addressed through organizing municipal level data sources to simulated biomass sources and estimating the optimal routing and mass collection using genetic algorithm implemented with Python. While the study successfully presents a framework for holistic approach to logistical question, it also reveals the sensitivity of the problem to multitude of cost factors and their accurate estimation.

2. Materials and methods

2.1. Biomass routing and biogas production

This study's optimization problem extends the Vehicle Routing Problem (VRP), a broader version of the Traveling Salesman Problem (TSP) for multiple operators (Bochtis and Sørensen, 2009). It incorporates aspects from the Biomass Collection Problem (BCP) and uses geo-spatial biomass data to determine pickup site locations and their content levels (Gracia et al., 2014). However, the pickup site development is not directly optimized within the genetic algorithm, leading to their exclusion from the optimization problem. Hence, the data processing serves as a solution for BCP, assuming pickup sites and their dynamics as given within the optimization problem.

The VRP encompasses dynamic biomass accumulation at pickup sites and their utilization within the biogas plant. Another aspect involves restricting vehicle visits to specific sites based on the vehicle type. Biomass consumption is modeled through continuous biogas production daily, with discrete biomass consumption from plant storage. This study's optimization problem is termed the Biomass Routing and Biogas Production Problem (BRBPP). BRBPP is conceptually introduced in this section.

The VRP involves a fleet of vehicles tasked with visiting a specific set of collection points once while minimizing the total distance traveled (Gracia et al., 2014). Vehicles must adhere to capacity limitations, starting and ending their routes at a depot (Bochtis and Sørensen, 2010). Routes are constrained by a predetermined duration covering travel and site time (Berger and Barkaoui, 2003). In the BRBPP, unlike VRP, the frequency of visits to a site isn't limited to one; it depends on factors such as time windows, vehicle allowances, and site-specific duration parameters. The total route duration is constrained by the length of a standard work shift, which is assumed to be 9 h in this study, including a 45-min long break.

VRP assumes static collection points with constant pickup capacity, eliminating the need for revisits and post-pickups (Bochtis and Sørensen, 2009). Conversely, in BRBPP, biomass accumulation necessitates potential revisits post-emptying, varying from linear daily accumulation to occasional, significant "jumps." The model combines VRP with time window constraints (VRPTW) (Bochtis and Sørensen, 2010), but in BRBPP, sites aren't mandated to be visited within designated time windows. Yet, prolonged site visits could incur additional storage costs, augmenting the overall solution expenses.

The BRBPP undergoes notable complexity due to the inclusion of the biogas production process. In traditional routing problems like TSP, VRP, and VRPTW, the depot primarily serves as the start and end point without distinct attributes (Berger and Barkaoui, 2003; Bochtis and Sørensen, 2010, 2009; Gracia et al., 2014). In BRBPP, the biogas plant acts as the depot and its storage levels are crucial for fulfilling biogas production demands. This process daily consumes stored biomasses; if the storage is depleted and consequently there is a decrease or a halt in production, additional costs will occur. Maintaining specific thresholds in stored biomasses, especially Total Solids (TS), is essential to ensure a proper anaerobic digestion process within the specific reactor type for wet-type digestion (Hadin and Eriksson, 2016). Exceeding the TS limit prompts dilution with water, incurring added expenses. The stored biomasses should align with annual input targets, as surpassing storage capacity or total yearly biomass imports triggers additional expenses.

The dynamic accumulation of biomasses in the BRBPP creates a shifting optimization challenge (Liu et al., 2019). Delaying a site visit influences total solution costs, potentially affecting biogas production if storage depletes. Opting for an early visit to a site with less biomass, postponing a visit to another site with greater biomass, may lead to added storage costs if site capacities are exceeded due to the timing of the visits.

Fig. 1 portrays the BRBPP, showcasing different pickup sites and vehicles, representing a typical day within the simulation period. It offers a partial solution example. Each day starts with vehicle assignment to the routes, outlining the locations to be visited. A route might be empty, signifying that there are no transportation needs for a certain vehicle during the day.

The cost function reflects the impact of multiple variables in the BRBPP, creating a multi-dimensional optimization challenge. Unlike the single-objective approach of the classical VRP, the BRBPP, influenced by various cost variables, necessitates multi-objective optimization methods for seeking solutions (Banasik et al., 2017; Baños et al., 2013; Jozefowiez et al., 2008). By incorporating and minimizing multiple cost terms in the optimized routing, the circular supply chain's side-stream procurement optimization yields more comprehensive solutions. These solutions consider diverse cost-influencing factors, thereby elevating supply chain efficiency (Banasik et al., 2017; Baños et al., 2013; Jozefowiez et al., 2008).

The cost function (elaborated in Section 2.4.2) aims to optimize the logistic routing of the biogas production supply chain from farms to the biogas plant. The goal is to prevent production interruptions, minimize dilution water usage, avoid facility overflows, and excessive site visits. This optimization involves route planning to simultaneously minimize driving distances and incorrect visits, reduce driver overtime, and ensure adequate farm-specific biomass storage capacity for production.

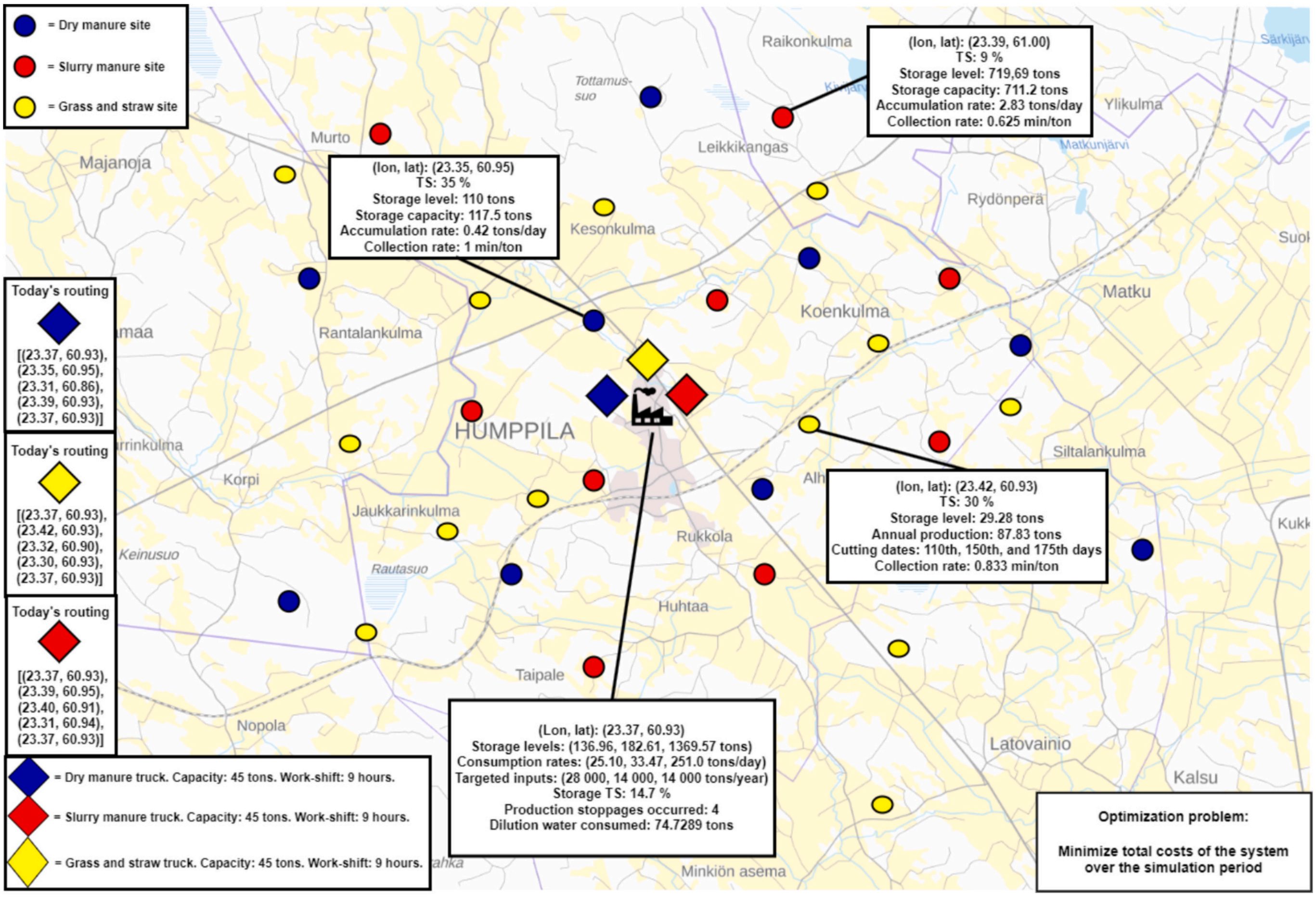


Fig. 1. An example illustration of a partial solution to the biomass routing and biogas production problem (BRBPP).

2.2. Data

Data for this study was sourced from Biomass atlas service developed by Natural Resources Institute Finland (Natural Resources Institute, Finland, 2023), offering geospatial datasets on various biomasses in Finland. Lehtonen et al. (2024) describe the details of the methodology behind generating and computing these datasets. Information on collection year, biomass type, mass in tons, and location coordinates for each biomass point was included in the data extracted from the Biomass atlas.

The biomasses used in this study were slurry manure of cattle and pigs from animal farms (data from years 2015 and 2016), solid manure of horses, cattle, poultry, and pigs (data from 2016), and cereal straw, silage grass, and grass side-streams (green fallows, buffer zones and green manure) generated in 2021. The data selection was based on the expected match with the target company's needs and intended use of biomasses for biogas production.

Data was presented as kilogram tons geospatially distributed to a uniform one-by-one kilometer grid. For grass and straw, the mass estimations followed this grid, but, for manure, the original estimation was the municipal's annual sum for the target location that had been resampled to the grid. Initial coordinate system EUREF-FIN was converted to WGS84 for further processing. The data files contained some redundancy since manure data was also presented by using the municipality polygon representation, but they were removed and only the grid representation was used.

2.3. Data filtering

The target biogas plant is envisioned to be located in an agricultural region in Finland with proximity to highways to enable both biomass logistics and market for biomethane. Only biomasses within a 50-km radius from the plant were considered. This range was deemed capable of adequately meeting the plant's resource needs and limit the costs of logistics. The target input values for grasses and straws were 28,000 tons per year, achievable by the straw biomass potential alone, and the target for manures was 14,000 tons, also potentially feasible within the area. Distance was calculated using Euclidean distances for each grid square center point.

In the real world, the biomasses would be collected to a pick-up point, most likely by the farmer, from where a transportation company would collect the biomass and transport it to the biogas plant. These pick-up points would later be determined by practical working conditions, but in this study, they were estimated by using KMeans clustering, an unsupervised learning technique (Likas et al., 2003), to identify subareas later to be used as pick-up sites in the model. During clustering, the masses were represented through replicating the points with the number of tons to estimate the KMeans mass weighted cluster centroids. The number of clusters was set to 100, and by using the elbow method (Liu and Deng, 2021), this choice was verified to have a low within cluster sum of squared distances to the cluster centroid. Fig. 2 demonstrates the initial clustering. To facilitate the visualization, the mass of biomass was converted to truckloads (45 tons per truckload) and transformed to a logarithmic scale.

The number of clusters, *k*, was reduced for faster computations. The initial 50 km clustered dataset was divided into five tiers, each 10 km

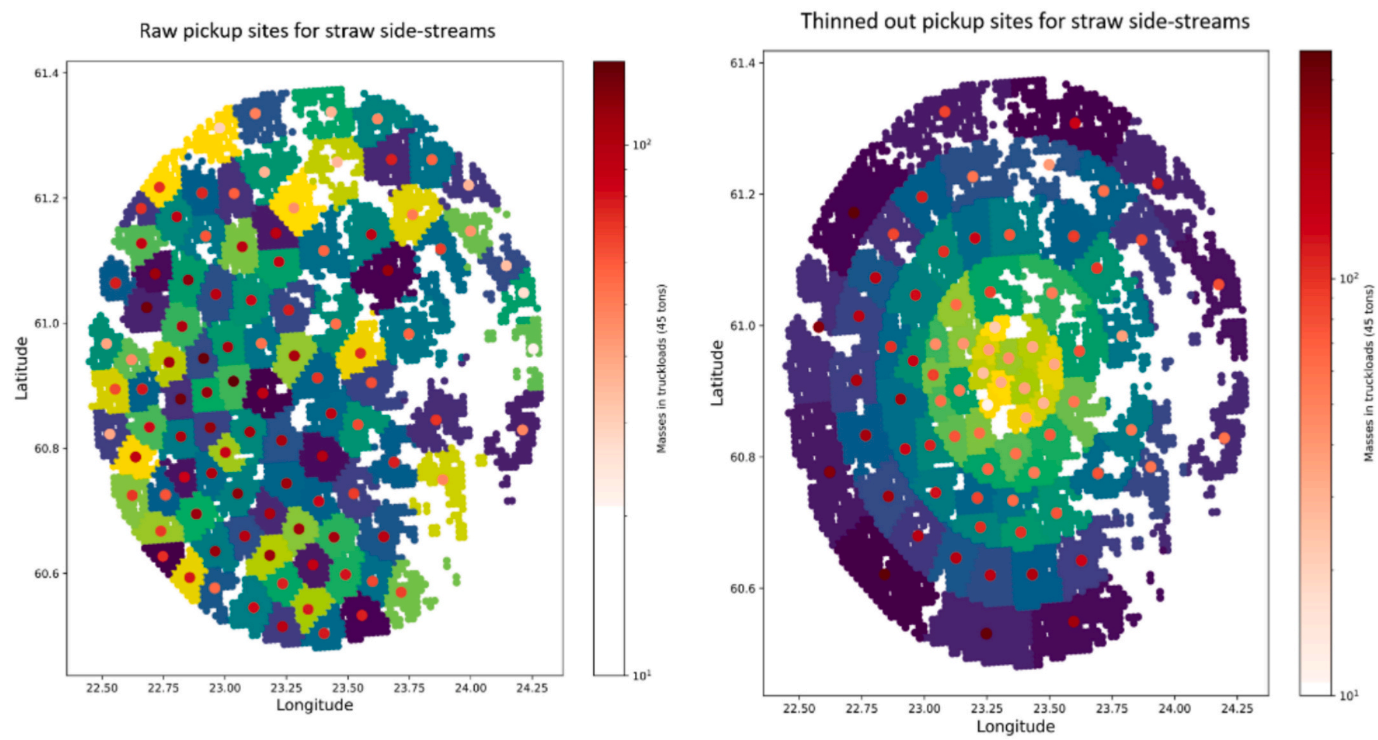


Fig. 2. Left: An example of the initial clustering phase in data processing. Right: An example of the methodology of cluster thinning.

thick. Calculating new k values within each tier involved dividing the original cluster count by a divisor coefficient raised to the tier's sequence number. Reclustering within each tier based on the new k values led to a reduction in the total number of clusters. To further reduce the excess number of clusters, only the closest ones to the facility were used so that their combined biomasses met the annual need for each biomass type (grass, straw, and manure). Fig. 3 shows the resulting estimated pickup sites.

2.4. Computational model

2.4.1. Biomass accumulation model

The linear accumulation model used for manure would not suit grasses and straws due to their harvest requirements, as their harvesting primarily occurs in the summer in Finland. Thus, it was assumed that farms conduct three harvests during the summer: mid-June, early August, and mid-September. After each harvest, the storage level at a pickup site increases from zero to one-third of its annual production. Hence, during a two-week period, the accumulation process was modeled as:

$$\Delta m_{t,i} = \frac{1}{3} M_i \varepsilon_{t,i} \tag{1}$$

where $\Delta m_{t,i}$ is the accumulation of grass or straw at the day t on the pickup site i; M_i is the annual accumulation of grass or straw at site i; and ε is a vector with 9 zeros and one randomly positioned 1, indicating the jump occurrence date, totaling ten business days. The equation governs accumulation only during the expected two-week harvest periods, remaining zero otherwise. $\varepsilon_{t,i}$ is site specific for each two-week period, introducing real-world like variation in the harvest times between sites. To maintain consistency, ε values were pre-determined.

Cluster centroids were modeled as pickup site locations, serving as collection points for the biomasses within the designated area of each cluster. For simulation purposes, initial storage levels of the pickup sites were selected randomly to prevent potential systematic errors that could result from uniform or constant level settings. These initial levels were determined for manure pickup sites by multiplying the site's capacity by a random number between 0 and 0.8. Regarding grass and straw pickup sites, 40 % of the pickup sites were randomly selected to include a storage level that represents the amount of collectable grass and straw generated from a single harvest. This level corresponds to a third of the annual accumulation of grass and straws, considering that three harvests occur within a year.

Regarding manures, pickup site capacity was determined as the cumulative mass within clusters, aligned with regulations mandating farms to accommodate annual manure production. Clustering datasets followed an annual timeframe. Conversely, for grasses and straws, a theoretical capacity was established; however, its optimization impact was disregarded as these biomasses, stored as bales in fields, were assumed to possess infinite storage capacity.

2.4.2. Cost function

The cost function utilized within the optimization of this study is defined as follows:

$$\mathrm{TC} = \mathrm{C}_{\mathrm{vehicles}} + \mathrm{C}_{\mathrm{bgp}} + c_{ovld} \sum_{i=1}^{m} \mathrm{OVLD}_i, \tag{2}$$

where TC is the total costs caused by the system with the given routing, c_{ovld} is the assumed cost of one day of overload at a pickup site, OVLD_i is the number of days of overloading at pickup site i and m is the number of pickup sites. Moreover, $\mathrm{C}_{\mathrm{vehicles}}$ consist of costs caused by vehicles, and is defined as follows:

$$\mathrm{C}_{\mathrm{vehicles}} = p_f c_f \sum_{i=1}^{n} \mathrm{ODO}_i + c_{ovth} \sum_{i=1}^{n} \mathrm{OVT}_i + c_{wv} \sum_{i=1}^{n} \mathrm{WV}_i, \tag{3}$$

where p_f is the assumed price of fuel (€/l), c_f is the assumed fuel consumption (l/100 km), n is the number of vehicles collecting and transporting the biomasses from farms to the biogas plant, ODO_i is the total mileage driven by vehicle i with the given routing, c_{ovth} is the assumed cost of the overtime work (€/h), OVT_i is the total overtime work of vehicle i with the given routing, c_{wv} represents the unit cost for a vehicle

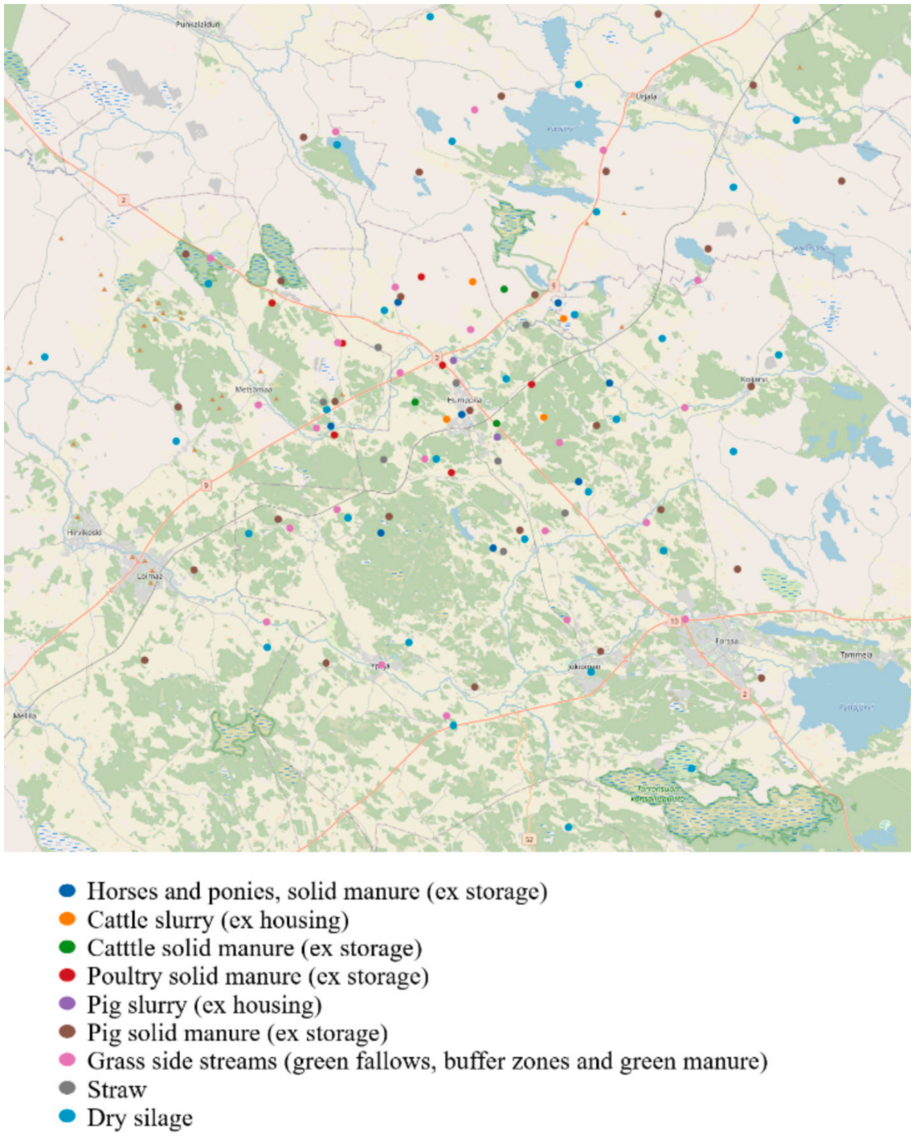


Fig. 3. Visualization of the spatial distribution of pickup sites approximated from biomass datasets.

visiting the wrong type of pickup site, and WV_i signifies the number of wrong visits carried out by vehicle i based on the given routing.

Additionally, C_{bgp} represents costs caused by inefficiencies in biogas production at the biogas plant defined as follows:

$$C_{bgp} = c_{pdst}PDST + c_{ovf}OVF + c_{exim}EXIM + c_{dw}DW \quad (4)$$

where c_{pdst} is the assumed cost of a production stoppage, i.e. biogas plant running out of biomass input, PDST is the number of production stoppages that occurred within the simulation with the given routing, c_{ovf} is the assumed cost of overfilling the biogas plant's storage, which is assumed to have a capacity equal to the targeted annual inputs, OVF is the number of overstocking occurrences that happened with the given routing, c_{exim} is the assumed cost of one unnecessary import of biomasses to the biogas plant after the targeted annual input is reached, EXIM represents the number of unnecessary imports that occurred within the simulation with the given routing, c_{dw} represents the assumed cost for one ton of dilution water consumed, and DW signifies the total amount of dilution water, in tons, consumed at the biogas plant based on the given routing.

Eq. (1) accounts for overloads at pickup sites yet overfilling at grass and straw sites was omitted from the optimization. Given the assumption of infinite storage due to baling in fields, the cost function term $OVLD_i$ remained unchanged if storage surpassed capacity at these specific pickup sites, thereby disregarding overloads at grass or straw sites.

The distances and travel times between the pickup sites and other locations were collected by utilizing a routing API (openrouteservice, n.d.), which returns distance and duration matrices for the given geographical coordinates. Grass and straw bales are assumed to be transported secured with straps as piece goods, while slurry manure involves specialized intake equipment in a vehicle container. Dry manures are collected in a closed boxcar. Each vehicle is assumed to have a transport capacity of 45 tons.

2.4.3. Genetic algorithm

Genetic algorithms (GA) are metaheuristic approaches widely used in operations management for problem-solving and optimization. Inspired by biological evolution, GAs start with an initial population of solutions, referred to as chromosomes, with genes representing parts of these solutions. Genetic operators are applied to create new potential solutions (children) by combining genes from existing solutions (parents). GAs are iterative methods where the population of solutions are used to create better ones (Gracia et al., 2014; Katoch et al., 2021).

Here, the task of the GA was to find a minimizing solution to the cost function. Genes refer to locations a vehicle visits and their order in the chromosome was set to be the order of driving the solution proposal route. Population size of 12,384 was used corresponding to four times the number of required genes. The initial population was randomly generated.

During each iteration, each chromosome acts as a parent pair without limitation, allowing for broader exploration across potential solutions. Selection of parent pairs was randomized. Crossover operator selects two random crossover points from parent p_1's chromosome and transmits the genetic information between these points to the beginning of the child. The remaining genetic material is inherited from parent p_0, in the order of appearance excluding the genes between the previously selected crossover points. This method generates new offspring with genetic sequences comprised of segments from both parental sources. Child replaces one of its parents in the population if its genetic sequence, i.e., driving route, is more efficient than parents'. Child is rejected if their cost exceeds that of either parent. (Niemitalo and Ekkerman, 2022).

The stopping criteria for optimization was set to 1,500,000 iterations. After completing 1,000,000 generations, the algorithm transitioned to a greedy approach for the remaining 500,000 iterations. In the greedy mode, each chromosome with the best-known solution achieved thus far. Ultimately, the best-performing chromosome, which holds the lowest cost (best fitness value), is selected to conclude the optimization process. This strategy avoids reducing the number of potential solutions in the inherited generations, thereby reducing the likelihood of getting trapped in a local optimum.

2.4.4. Truck fleet and biogas plant model

Truck fleet consisted of nine collection trucks, three for each biomass type. Each truck was assumed to have capacity of 45 tons and, per simulation day, was allowed to work a 9-h work shift including a 45-min break. Each driving shift of a truck started and ended at the depot.

The pickup duration at sites was modeled with a fixed element and a component that linearly corresponds to the quantity of biomass for collection as

$$\Delta t_{i,l} = T + v_m l, \tag{5}$$

where $\Delta t_{i,l}$ is the pickup duration for the pickup site i with a load level of l, T is the constant term of the pickup operation, and v_m is the collection rate for the biomass m. T was assumed to be 10 min for all biomasses. The parameter T signifies the setup and dismantling time for pickup operations. The pickup rate v_m was set at 1.6 tons/min (0.625 min/ton) for slurries, 1 ton/min (1 min/ton) for solid manures, and 1.2 ton/min (0.833 min/ton) for grasses and straws based on expert discussions. Despite the seasonal accumulation of grasses and straws, assuming post-harvest baling allows year-round collection. If the vehicle reaches its load limit, it gathers within its capacity, adjusting the operation time according to the loaded volume.

During the simulation, biomass storage levels decrease daily due to consumption rates reflecting resource utilization in biogas production. Upon vehicle arrival for unloading at the biogas plant, the storage levels are updated with the unloaded amount and average rate of total solids (TS) content is recalculated to the storage's weighted rate average.

The depot model logs the cumulative received biomass (triggering when yearly intake targets are met); excessive unloading post-target attainment is penalized within the cost function. Storage overfilling is also a cost factor penalized daily. If storage empties, leading to production halts, the stoppage duration is penalized. To optimize routes with sustained TS rates below 15 % (average of all biomasses to be used as biogas plant feedstock), cumulative dilution water usage in production was tracked. When TS content of the stored biomasses exceed 15 %, dilution water is added to adjust the biogas plant's feedstock mixture, influencing the minimized cost function within the GA, as discussed in the subsequent section.

2.4.5. Performance indicators

In addition to the cost function, the following indicators were chosen to evaluate the solution performance in the scenarios: total mileage (km) driven by the truck fleet, total overtime (h) of the truck shifts, number of days with no routes, total pickup site overload days over all sites, production stoppages at the biogas plant, consumption of dilution water (tons) at the biogas plant, unnecessary imports to the biogas plant, i.e., transporting biomass that was not needed for the desired, set input composition, and overfillings of storage within the biogas plant.

2.5. Experiments

In total, four scenarios were applied and, in all experiments, the simulation period spanned a year, accounting for 251 workdays. Used parameter values were:

- price of fuel $p_f = 2.0$ € and $c_f = 50$ L/100 km,
- penalty for pickup site storage overload $c_{ovld} = 50$ €,
- overtime work cost $c_{ovth} = 50$ €/60 *min*,
- penalty for use of dilution water $c_{dw} = 5$ €/m^3,
- penalty for biogas production stoppage $c_{pdst} = 100000$ €/1 d,
- penalty for storage overfill at biogas plant $c_{ovf} = 100$ €,
- penalty for unnecessary import $c_{exim} = 100$ €,
- penalty for visiting wrong site $c_{ww} = 1000$ €.

Choice of parameter values was based expert discussions and attempt to create representative simulations in given scenarios.

2.5.1. Scenario 1: Efficient routing

Scenario 1 is the GA solution with respect to the cost function such that the biomass quality is stable with respect to time.

2.5.2. Scenario 2: Efficient routing with decreasing biomass quality

Scenario 2 presents an additional aspect to scenario 1: natural decomposition of biomass affecting its quality for further use. This involved modeling the passive drying of all biomass types, adopting van Dyken et al.'s (2010) storage model, where weekly biomass volume reduction of 1 % is attributed to passive drying. This time-sensitive factor manifests as

$$V_{b,w} = V_b\left(1 - r_{l,d}\right)^w, \tag{6}$$

where $V_{b,w}$ is the volume of biomass b at week w, V_b is the initial volume level of biomass at the location of interest, and $r_{l,d}$ is the rate of biomass volume loss caused by passive drying on a weekly basis.

The study approximated a 5 % weekly reduction in moisture content due to passive drying, affecting the TS rate of biomasses. Although van Dyken et al.'s (2010) model detailed moisture changes in timesteps, this study, due to its broader complexity, simplified the moisture level modeling by assuming a consistent 5 % decrease per week. This factor, depicting alterations in moisture levels and subsequently TS rates, is represented as

$$M_{b,w} = M_b\left(1 - r_{m,d}\right)^w, \tag{7}$$

where $M_{b,w}$ is the level of moisture of biomass b at week w, M_b is the initial level of moisture of biomass at the location of interest, and $r_{m,d}$ is the rate of biomass moisture decrease caused by passive drying on a weekly basis. The passive drying dynamics, described in Eqs. (5) and (6), were considered applicable during the storage at pickup sites, transportation, and the storage at the biogas plant, all preceding the utilization of biomass as biogas plant feedstock.

2.5.3. Baseline for scenarios

To benchmark routings in both scenarios, the possible organization of driving resulting from not knowing which sites provide the biomass (semi-random collection orders) or from using a low computational cost organization tool were estimated. Biomass pickup trucks follow these orders to collect and transport biomasses back to the biogas plant. If a truck has multiple orders on a given day, it collects them within its capacity before returning to the plant. Hence, this baseline routing mirrors the practical approach adopted by real-world logistics operators handling biomass collection in an order-oriented manner.

The baseline estimate was calculated by terminating the GA optimization at 100 generations which yields a low optimized solution. The premature termination resulted the vehicles covering unnecessary distances due to erroneous visits in solution proposal (for instance, a vehicle intended for dry manure collection might erroneously visit a site producing straw), but these were filtered from quantifications.

2.6. Used software

The simulation and optimization models utilized in this study stemmed from the groundwork laid by (Niemitalo and Ekkerman, 2022). Modifications were applied to the simulation model for this study and are detailed in Eloranta (2023a). The simulation covers transportation traffic for a user-defined duration using process-based discrete event simulation in Python via Simply (Team SimPy, 2020). For quicker calculations, the optimizer employs SimCpp20, a C++ implementation of the same model (Schütz, 2021). Python code used for data processing is available at (Eloranta, 2023b).

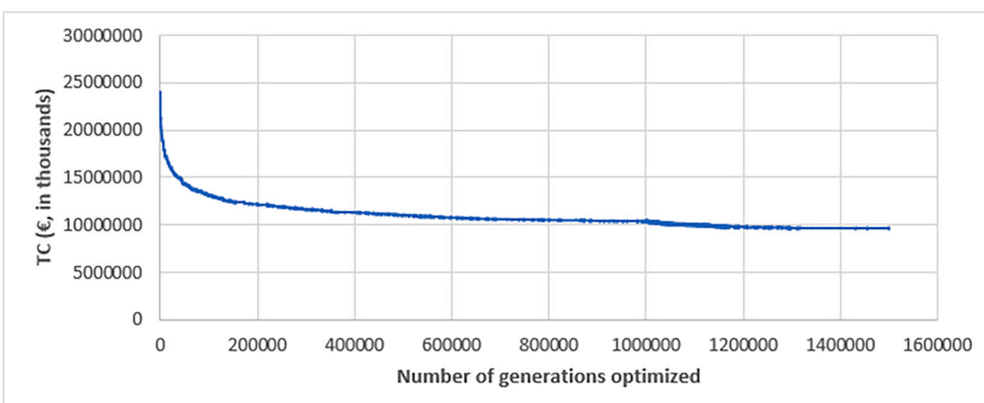
3. Results

Fig. 4 illustrates the optimization trajectories in both scenarios, depicting the most cost-effective route among various proposals. The cost of the best chromosome in the population is stable between the last 30,000 iterations, which suggest that a close to minimum solution has been reached.

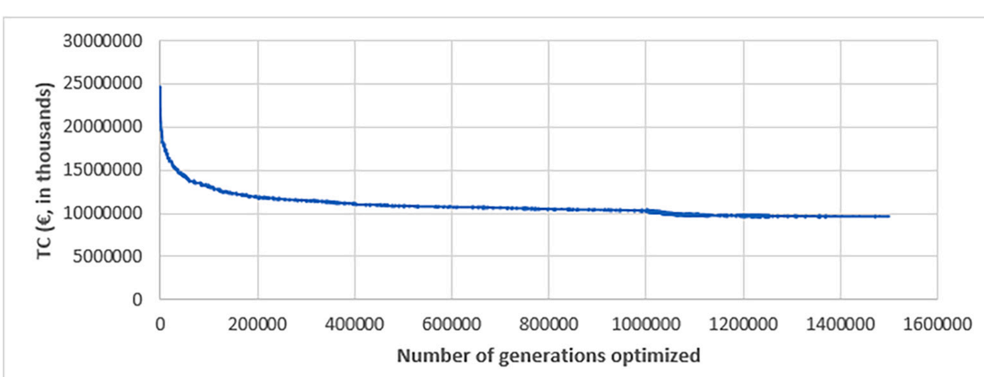
Comparison of performance indicator values between the scenarios

Table 1
Characteristics of scenarios with the values of the cost function terms resulting from optimization and baseline estimation.

	Optimized routing		Baseline routing	
	Scenario 1	Scenario 2	Scenario 1	Scenario 2
Total mileage (km)	3352	3292	2708	2411
Total overtime (h)	0	0	0	0
Number of days with no routes	177	174	182	188
Total pickup site overload days	1879	1763	3391	3283
Production stoppages	0	0	110	126
Consumption of dilution water (tons)	1,910,840	1,907,960	2,188,950	1,921,640
Unnecessary imports to the biogas plant	0	0	0	0
Overfillings within the biogas plant	0	0	0	0
Total costs (€)	9,648,153	9,627,953	22,114,303	22,372,352



Scenario 1 (Graph a)



Scenario 2 (Graph b)

Fig. 4. Routing optimization progress in scenarios 1 (Graph a) and 2 (Graph b). Switch to greedy mode is seen at 1,000,000 iterations and the cost function value of best solution in the population is stable in the last iterations, suggesting that convergence has been reached.

is collected in Table 1. The difference between the scenarios 1 and 2 is practically negligible and their indicator values are aligned. It would have been expected that the scenario 2 has higher cost in biomass collection routing as the biomass sites may be visited slightly less efficiently than in scenario 1 to gather the biomass at higher TS content This was, however, not the case and the difference between indicator values is most likely consequence of randomness in the GA. Comparing optimized routing with the baseline highlights the need for planning stable biogas plant feedstock stream since baseline, relying on one-by-one biomass offers by the sites, suffers from production stoppages and inefficient process parameters (excess use of dilution water). In all cases, overtime, unnecessary imports, and storage overfilling at biogas plant are zero, confirming the convergence of the solution. High dilution water amount is unexpected but difficult to estimate its implication given the used, very rough biogas plant model.

Table 2 describes the behavior of the truck fleet for different biomass types. As is evident from the zero mileage with some trucks, the fleet capacity was not constraining the solution. Scenario 2, where biomasses were assumed to dry as a function of time leading to increased TS content, shows higher mileage for grass and straw and less for solid manure, which is a natural implication of the drying. Since material amount of each site is lower, collection cost is higher. For manure, the material availability at a site is abundant, so the collection of biomass is more efficient.

4. Discussion

4.1. Computed routings

The comparison of the optimized results between scenarios 1 and 2 (one considering consistent biomass quality in terms of energy potential over time, and the other incorporating time-critical factors) underscores the challenge of determining the impact of biomass time-criticality on optimized performance. While routing performance seems slightly enhanced when considering time-criticality, reflected in lower total costs in Table 1, these differences in performance indicators are minimal and could be coincidental. Thus, it is uncertain if these disparities are solely due to accounting for biomass time-criticality, warranting further research on the subject.

Given the integration of biomass time-criticality, incorporating a drying process within various components such as sites, vehicles, and the biogas plant's storage - outlined in Eqs. (5) and (6), influenced by Van Dyken et al.'s model (2010) - results in volume reduction and heightened total solids content due to biomass' spontaneous drying. However, it remains uncertain whether this drying process directly influences the noted enhancements in performance indicators. In scenario 2, there were less driven kilometers, and the amount of dilution water used was lower, which resulted in lower total costs. Future research endeavors should delve into comprehending the dynamics of biomass time-criticality and its correlation with the specified performance metrics. In addition, it should be noted that the model by Van Dyken et al. (2010) does necessarily be valid with agricultural biomasses such as manures, grass and straw due to their lower TS content compared to wood biomasses. However, the model was applied to showcase the effect of changing TS content in the biomasses on the routing optimization and costs.

The results of the study highlight the importance of factoring passive biomass processes into logistic routing for biogas production systems. With optimized routing, both scenarios achieved zero production stoppages annually, however, scenario 1 was not time critical, while scenario 2 was time critical. Additionally, lower overload days at pickup sites suggest time-critical influences on biomass availability for biogas production. This is linked to the assumption of biomass time-criticality, where daily storage reductions at pickup sites result in fewer capacity-exceeding days.

Given the frequent overload days at pickup sites, questioning the sufficiency of nine vehicles for biomass collection arises. However, intentionally setting a high number of pickup sites ensures ample biomass for the biogas plant's annual resource needs. Consequently, using overload days as a performance indicator and cost factor might not be imperative. Therefore, the high values of numbers of overload days observed with optimized routings could be partially disregarded.

The absence of production stoppages in both scenarios emphasizes continuous biogas production and uninterrupted biomass processing within the plant. This suggests potential cost reductions via improved logistic planning, irrespective of whether time-criticality was accounted for. Evidence indicating passive time-critical processes impacting biomass quality for energy production suggests an opportunity to enhance the study's simulation model for a more comprehensive understanding of time-critical dynamics (Hadin and Eriksson, 2016; Mönch-Tegeder et al., 2013; Van Dyken et al., 2010).

Optimized solutions in both scenarios ensure uninterrupted biogas production without overtime work, reducing travel distance and overall costs by targeting specific sites. Yet, considering the absence of evidence supporting their global optimality, it's vital to acknowledge potential suboptimality in the attained results. Identifying the cost factors enhancing these solutions remains ambiguous. Thus, additional research is imperative to devise methodologies for achieving the global optimum solution in this context.

The benchmark results show the expected advantages of employing optimization techniques in planning biomass logistics. Comparison against the optimized routings reveal that a reactive approach to collecting biomasses based on random farm orders jeopardizes continuous biogas production, leading to frequent production stoppages. This reactive method also induces unnecessary kilometers driven, contrasting with planned logistics that consider resource requirements within the biogas plant. Conversely, optimized routes are proactive, considering resource utilization, vehicle mileage, and other pertinent factors, presenting a more reasoned and sophisticated approach to biomass collection than reactive, order-driven methods. This enhanced decision-making process holds potential for significantly reducing total system costs as well the amount of emissions from the logistics chain, ultimately

Table 2
Total mileages of each vehicle of the fleet in scenarios during the simulated one-year period.

		Optimized routing		Baseline routing	
		Scenario 1	Scenario 2	Scenario 1	Scenario 2
	Type of the vehicle	Mileage (km)	Mileage (km)	Mileage (km)	Mileage (km)
Vehicle 1	Slurry	36	135	28	8
Vehicle 2	Grass/straw	0	0	544	581
Vehicle 3	Solid manure	805	532	376	334
Vehicle 4	Slurry	135	36	0	11
Vehicle 5	Solid manure	659	1091	364	345
Vehicle 6	Grass/straw	0	0	409	239
Vehicle 7	Slurry	30	30	55	42
Vehicle 8	Grass/straw	0	0	297	474
Vehicle 9	Solid manure	1686	1467	636	378
Total	Slurry	201	201	83	61
Total	Grass/straw	659	1091	1205	1400
Total	Solid manure	2491	1999	14,201	951
Total	All biomass types	3352	3292	2708	2412

bolstering the efficiency of circular supply chains and refining CBE practices.

4.2. Significance and limitations of the study

The study aimed to use spatial mapping to plan logistics for biomasses near the biogas plant, optimizing routes for continuous production and streamlined flow. It sought to minimize costs while collecting and transporting biomasses efficiently. Additionally, it aimed to explore the impact of passive time-critical processes within biomasses on logistic optimization.

The study achieved its objectives partly. The methodology developed comprehensively models, simulates, and optimizes components within the circular supply chain responsible for transporting biomass to the biogas plant for production. Spatial information mapping for biomasses near the biogas plant was effectively executed. This process involved data processing stages that successfully mapped biomass potentials within a 50 km radius of the plant. It identified specific pickup locations to fulfill the annual biomass demand for biogas production.

The study proposed an optimized routing plan in a particular setting, aiming to execute cost-effective logistic routing for continuous biogas production at the plant. Validating the performance, it compared selected indicators between optimized and practical biomass routing approaches. The optimized routes minimized cost terms in the simulation model. However, these solutions might be suboptimal for the BRBPP, suggesting potential for further optimization with increased computational resources.

The study aimed to assess how passive drying processes in collected biomass affect optimized routing solutions for biogas production. Despite the optimized results displaying similar total cost performance, understanding the influence of time-critical variables on solution quality remains uncertain. Enhancing the simulation model is essential for a comprehensive grasp of biomass time-critical dynamics.

The study's theoretical significance encompasses several aspects: it introduces the definition of BRBPP and applies sophisticated optimization techniques to address it. While acknowledging potential suboptimal outcomes, the study underscores the need for further research, recognizing the complexity of BRBPP that may exceed current optimization capabilities. The comparison of optimized outcomes, considering or neglecting passive drying processes in biomass, leaves the impact of time-critical passive drying on logistic routing optimization uncertain. This underscores the necessity of reassessing BRBPP through a time-critical perspective to achieve a more comprehensive solution.

The study's main practical contribution involves tool and methodology development for geospatial mapping of extensive datasets (Eloranta, 2023b), alongside modeling and optimizing logistic routing for biomass collection and transportation to centralized biogas plants (Eloranta, 2023a; Niemitalo and Ekkerman, 2022). These tools extend beyond the specific biogas plant case, enabling mapping of diverse geospatial biomass datasets and logistics optimization for collecting agricultural biomasses for centralized biogas plants. The study yields an optimized daily routing plan for collecting biomasses destined for the biogas plant. Additionally, the geospatial mapping provides insights into nearby biomass availability, aiding efficient planning of biomass collection locations and distances. Such approach is applicable to many CBE use cases where different types of materials are collected from multiple locations for further processing.

4.3. Future research

The study reveals some future research opportunities as well as limitations concerning the initial biomass data (Natural Resources Institute, Finland, 2023.) used to determine cluster central points as pickup sites. A significant constraint lies in the applicability of the study's findings. The pickup site locations were obtained as weighted central points from clustering biomass data, lacking tailored methods for defining biomass collection points, like the Borvemar model (Gracia et al., 2014; Velázquez-Marti and Annevelink, 2009). Consequently, these modeled pickup sites, representing farms in simulations, are theoretical and may not correspond to actual nearby farms of the case company's biogas plant. Given the biogas plant's planning stage and the absence of an established supplier network, this limitation emphasizes the necessity of revisiting the study's simulation and optimization approach when actual farm suppliers are identified, ensuring more practical routing solutions. Same scholarly methodology applied to choice of modeling parameters (see chapter 2.5) and fine tuning their values closer to reality would be required.

One potential limitation of the study is the absence of constraints on the actual blend of biomasses utilized in biogas production. The model permits daily variations in biomass composition for assumed biogas production. However, this approach may lack practical feasibility, as optimization techniques exist to determine the ideal biomass composition for anaerobic digestion, maximizing methane production (Álvarez et al., 2010). These methods highlight the process's stability and minimal variations in composition.

Additional limitations stem from the logistic routing optimization using GA, known for its significant computational demands. As previously highlighted, the inherent risk of GA optimization involves potential convergence to a local rather than a global optimum (Katoch et al., 2021). The study's possibly suboptimal outcomes may result from the GA being confined to a local optimum. Due to limited computational resources and time constraints, the optimization process was capped at 1,500,000 generations. While further optimization could potentially yield better solutions, the complexity of the problem necessitated this decision within the research timeframe available.

Future research should prioritize advancing optimization techniques for solving the BRBPP. One promising avenue involves implementing reinforcement learning (RL) methodologies, which have already proven effective in optimizing logistic systems (Chen et al., 2022; Sun et al., 2019). RL, a sophisticated machine learning approach, holds promise in efficiently addressing the BRBPP by improving existing solutions and comprehending the system's functionality underlying the optimization process. Integrating RL into logistic routing optimization for BRBPP could provide insights into how passive time-critical processes within biomasses impact optimized solution performance.

Future research could enhance the simulation model's realism by introducing increased noise levels, introducing randomness to deterministic processes like manure accumulation. Similar noise terms could also be incorporated into vehicle collection operations' durations and the daily biomass consumption by the biogas plant. However, augmenting noise levels could substantially complicate the optimization problem, necessitating careful consideration of their inclusion and implications before implementation in the simulation model.

5. Conclusion

This study developed and demonstrated an optimization methodology for biomass collection and routing with the addition of considering the decrease of biomass quality during storage. Here, geospatial biomass data and Vehicle Routing Problem (VRP) solution algorithm were integrated to create scenarios based on a potential real-world location. The proposed optimization framework successfully demonstrates the potential enhances the efficiency and cost-effectiveness of biomass logistics, ensuring a continuous and reliable supply for renewable energy production.

From a broader perspective, this study contributes to the circular bioeconomy by providing insight in reducing transportation costs and emissions or by offsetting them elsewhere based on better resource use in the bioenergy sector. The developed approach and tools offer practical applications for policymakers, bioenergy producers, and logistics operators seeking to enhance the efficiency and sustainability of biomass-based energy production.

The framework allows future explorations in integrating real-time data on biomass availability, weather conditions, and dynamic transport constraints to further refine routing efficiency.

CRediT authorship contribution statement

Kasper Eloranta: Writing – original draft, Visualization, Software, Methodology, Investigation, Formal analysis, Conceptualization. **Olli Koskela:** Writing – review & editing, Software, Resources, Project administration, Methodology, Funding acquisition, Formal analysis, Conceptualization. **Elina Tampio:** Writing – review & editing, Resources, Methodology, Funding acquisition, Conceptualization. **Juho Kanniainen:** Writing – review & editing, Validation, Supervision, Conceptualization. **Ulla A. Saari:** Writing – review & editing, Validation, Supervision, Conceptualization.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

This work was conducted in the "Biokaasua hiiliviisaasti ja ravinteet kiertoon Kanta-Hämeessä (BioKanta)" -project (A80732, A80733) funded by the Regional Council of Häme from the European Regional Development Fund.

Data availability

Data will be made available on request.

References

Álvarez, J.A., Otero, L., Lema, J.M., 2010. A methodology for optimising feed composition for anaerobic co-digestion of agro-industrial wastes. Bioresour. Technol. 101, 1153–1158. https://doi.org/10.1016/j.biortech.2009.09.061.

Banasik, A., Kanellopoulos, A., Claassen, G.D.H., Bloemhof-Ruwaard, J.M., Van Der Vorst, J.G.A.J., 2017. Closing loops in agricultural supply chains using multi-objective optimization: a case study of an industrial mushroom supply chain. Int. J. Prod. Econ. 183, 409–420. https://doi.org/10.1016/j.ijpe.2016.08.012.

Baños, R., Ortega, J., Gil, C., Márquez, A.L., De Toro, F., 2013. A hybrid meta-heuristic for multi-objective vehicle routing problems with time windows. Comput. Ind. Eng. 65, 286–296. https://doi.org/10.1016/j.cie.2013.01.007.

Bellezoni, R.A., Adeogun, A.P., Paes, M.X., De Oliveira, J.A.P., 2022. Tackling climate change through circular economy in cities. J. Clean. Prod. 381, 135126. https://doi.org/10.1016/j.jclepro.2022.135126.

Berger, J., Barkaoui, M., 2003. A Hybrid Genetic Algorithm for the Capacitated Vehicle Routing Problem. In: Cantú-Paz, E., Foster, J.A., Deb, K., Davis, L.D., Roy, R., O'Reilly, U.-M., Beyer, H.-G., Standish, R., Kendall, G., Wilson, S., Harman, M., Wegener, J., Dasgupta, D., Potter, M.A., Schultz, A.C., Dowsland, K.A., Jonoska, N., Miller, J. (Eds.), Genetic and Evolutionary Computation — GECCO 2003, Lecture Notes in Computer Science. Springer Berlin Heidelberg, Berlin, Heidelberg, pp. 646–656. https://doi.org/10.1007/3-540-45105-6_80.

Bijarchiyan, M., Sahebi, H., Mirzamohammadi, S., 2020. A sustainable biomass network design model for bioenergy production by anaerobic digestion technology: using agricultural residues and livestock manure. Energy Sustain. Soc. 10, 19. https://doi.org/10.1186/s13705-020-00252-7.

Bochtis, D.D., Sørensen, C.G., 2009. The vehicle routing problem in field logistics part I. Biosyst. Eng. 104, 447–457. https://doi.org/10.1016/j.biosystemseng.2009.09.003.

Bochtis, D.D., Sørensen, C.G., 2010. The vehicle routing problem in field logistics: Part II. Biosyst. Eng. 105, 180–188. https://doi.org/10.1016/j.biosystemseng.2009.10.006.

Chen, L., Hu, B., Guan, Z.-H., Zhao, L., Shen, X., 2022. Multiagent meta-reinforcement learning for adaptive multipath routing optimization. IEEE Trans. Neural Netw. Learning Syst. 33, 5374–5386. https://doi.org/10.1109/TNNLS.2021.3070584.

Eloranta, K., 2023a. "logistics_sim", source code of case-specific, further developed version of the logistics simulation and optimization.

Eloranta, K., 2023b. "data_procession.ipynb", source code of the data procession tool.

Gracia, C., Velázquez-Martí, B., Estornell, J., 2014. An application of the vehicle routing problem to biomass transportation. Biosyst. Eng. 124, 40–52. https://doi.org/10.1016/j.biosystemseng.2014.06.009.

Hadin, Å., Eriksson, O., 2016. Horse manure as feedstock for anaerobic digestion. Waste Manag. 56, 506–518. https://doi.org/10.1016/j.wasman.2016.06.023.

Jozefowiez, N., Semet, F., Talbi, E.-G., 2008. Multi-objective vehicle routing problems. Eur. J. Oper. Res. 189, 293–309. https://doi.org/10.1016/j.ejor.2007.05.055.

Kamilaris, A., Prenafeta-Boldú, F.X., 2021. Examining the perspectives of using manure from livestock farms as fertilizer to crop fields based on a realistic simulation. Comput. Electron. Agr. 191, 106486. https://doi.org/10.1016/j.compag.2021.106486.

Katoch, S., Chauhan, S.S., Kumar, V., 2021. A review on genetic algorithm: past, present, and future. Multimed. Tools Appl. 80, 8091–8126. https://doi.org/10.1007/s11042-020-10139-6.

Lehtonen, E., Anttila, P., Hakala, K., Luostarinen, S., Lehtoranta, S., Merilehto, K., Lehtinen, H., Mäntylä, V., Maharjan, A., Niemeläinen, O., Kolehmainen, M., 2024. An open web-based GIS service for biomass data in Finland. Environ. Model. Softw. 176, 105972. https://doi.org/10.1016/j.envsoft.2024.105972.

Likas, A., Vlassis, N., Verbeek, J.J., 2003. The global k-means clustering algorithm. Pattern Recogn. 36, 451–461. https://doi.org/10.1016/S0031-3203(02)00060-2.

Liu, F., Deng, Y., 2021. Determine the number of unknown targets in open world based on elbow method. IEEE Trans. Fuzzy Syst. 29, 986–995. https://doi.org/10.1109/TFUZZ.2020.2966182.

Liu, S., Zhang, Y., Liu, Y., Wang, L., Wang, X.V., 2019. An 'internet of things' enabled dynamic optimization method for smart vehicles and logistics tasks. J. Clean. Prod. 215, 806–820. https://doi.org/10.1016/j.jclepro.2018.12.254.

Mönch-Tegeder, M., Lemmer, A., Oechsner, H., Jungbluth, T., 2013. Investigation of the methane potential of horse manure. Agric. Eng. Int. 15, 161–172.

Monteiro, E., Ferreira, S., 2022. Biomass waste for energy production. Energies 15, 5943. https://doi.org/10.3390/en15165943.

Natural Resources Institute, Finland, 2023. Biomassa-Atlas [WWW Document]. URL, Natural Resource Data. https://biomassa-atlas.luke.fi/.

Niemitalo, O., Ekkerman, G., 2022. Logistics Simulation and Optimization. Source Code and the Documentation of the Original Simulation and Optimization Models.

openrouteservice, n.d. Routing API [WWW Document]. URL https://openrouteservice.org/ (accessed 7.20.23).

Ranta, V., Aarikka-Stenroos, L., Väisänen, J.-M., 2021. Digital technologies catalyzing business model innovation for circular economy—Multiple case study. Resour. Conserv. Recy. 164, 105155. https://doi.org/10.1016/j.resconrec.2020.105155.

Schütz, F., 2021. SimCpp20 – A Discrete-Event Simulation Framework for C++.

Sun, P., Hu, Y., Lan, J., Tian, L., Chen, M., 2019. TIDE: Time-relevant deep reinforcement learning for routing optimization. Future Gener. Comput. Syst. 99, 401–409. https://doi.org/10.1016/j.future.2019.04.014.

Team SimPy, 2020. Discrete event simulation for Python [WWW Document]. URL. https://simpy.readthedocs.io/en/latest/ (accessed 7.20.23).

Van Dyken, S., Bakken, B.H., Skjelbred, H.I., 2010. Linear mixed-integer models for biomass supply chains with transport, storage and processing. Energy 35, 1338–1350. https://doi.org/10.1016/j.energy.2009.11.017.

Velázquez-Marti, B., Annevelink, E., 2009. GIS application to define biomass collection points as sources for linear programming of delivery networks. Trans. ASABE 52, 1069–1078. https://doi.org/10.13031/2013.27776.

Wang, Z., Bui, Q., Zhang, B., Pham, T.L.H., 2020. Biomass energy production and its impacts on the ecological footprint: an investigation of the G7 countries. Sci. Total Environ. 743, 140741. https://doi.org/10.1016/j.scitotenv.2020.140741.

Yılmaz Balaman, Ş., Wright, D.G., Scott, J., Matopoulos, A., 2018. Network design and technology management for waste to energy production: an integrated optimization framework under the principles of circular economy. Energy 143, 911–933. https://doi.org/10.1016/j.energy.2017.11.058.