Computers and Electronics in Agriculture 226 (2024) 109456

Contents lists available at ScienceDirect

Computers and Electronics in Agriculture

journal homepage: www.elsevier.com/locate/compag

Path planning of manure-robot cleaners using grid-based reinforcement learning

Congcong Sun *, Rik van der Tol, Robin Melenhorst, Luis Angel Ponce Pacheco, Peter Groot Koerkamp

Agricultural Biosystems Engineering Group, Wageningen University, 6700 AA Wageningen, The Netherlands

ARTICLE INFO

Keywords:
Path planning
Manure-robot cleaner
Reinforcement learning
Cow behaviors
Cow-robot collisions

ABSTRACT

The use of a robot cleaner for manure removal improves housing conditions for dairy cows in the face of labor shortages. However, current robot cleaners follow programmed fixed routes without considering the dynamic behaviors of cows. This cleaning approach is less efficient and leads to more cow-robot encounters or collisions, thus affecting animal welfare. To address these issues, this paper (1) developed heatmap models for cow locations and defecation behaviors; (2) proposed a dynamic path planning approach for the manure robot cleaner using Grid-based Reinforcement Learning; (3) incorporated cow location information and defecation behavior into the path planning process; (4) compared the performance of the proposed approach with two different cleaning methods: the current fixed programmed cleaning in practice and the ideal path produced by simulated annealing for traveling salesman problem. The simulations mimic the situation in a barn at Dairy Campus of Wageningen Livestock Research located in Leeuwarden (the Netherlands). Obviously, the best performance was achieved when the route was executed without cows present, resulting in no cow-robot collision. However, with cows present, the proposed dynamic path planning strategy achieved a 67.6% reduction in cow-robot encounters while maintaining 85.4% of the cleaning performance compared to the current programmed fixed routes. Compared to the ideal path produced by simulated annealing for traveling salesman problem, the proposed dynamic path planning approach achieved 5% better cleaning performance, at the cost of 25% more cow-robot encounters due to its longer working path. We conclude the proposed grid-based Reinforcement Learning solution for manure robots in barns cleaned most efficient with the least interference with cow traffic.

1. Introduction

The dairy industry plays a pivotal role in global food production, emphasizing the need for efficient and sustainable dairy barn facilities (Britt et al., 2018). One critical aspect influencing the overall functionality of these facilities is the design and management of floor systems. Recently, researchers and practitioners alike have recognized the significant impact that floor systems can have on cow comfort, health, and overall productivity within dairy barns (Van der Tol et al., 2005; Solano et al., 2015).

Many aspects of floor systems in dairy barn facilities are known to affect the well-being of dairy cattle; flooring materials, design considerations, and maintenance practices. In the past few years, consumer demands for ethically produced and high-quality dairy products rise, hence the emphasis on providing optimal living conditions for dairy cows becomes increasingly important (Cardoso et al., 2016; Barker et al., 2010).

Besides that, cleanliness of the dairy barn floor affects the number of claw infections, its slipperiness is another great concern (Van der Tol et al., 2005; Bruijnis, 2006), combinations of both cause the majority of health and animal welfare related problems (Somers et al., 2003). To maintain proper hygiene, housing and management practices play an increasingly important role (Gieseke et al., 2018).

With the selection of flooring systems, an efficient manure removal system is required to maintain optimal hygiene for dairy cattle. The standard slatted concrete floors, characterized by openings that enable manure to fall through. Grooved concrete floors are designed with channels or grooves that facilitate liquid drainage and reduction of ammonia emission. Both allow for effective removal of manure and provide a similar advantage by minimizing direct contact between cows and manure, thereby reducing the risk of bacterial contamination. On the other hand, solid concrete floors, while simpler in design, require a well-planned manure management strategy to prevent accumulation

* Corresponding author.
E-mail address: congcong.sun@wur.nl (C. Sun).

https://doi.org/10.1016/j.compag.2024.109456
Received 31 December 2023; Received in revised form 9 July 2024; Accepted 15 September 2024
Available online 18 September 2024

and maintain a clean environment. The choice between these flooring types depends on various factors, including climate, herd size, and management preferences. With respect to manure removal efficiency, manure scrappers have been the main choice for years, but more recently the robot cleaner gains more popularity. In any system, less contact between manure and cows' hooves is key, which is beneficial to floor friction, claw health and welfare in general (Somers et al., 2003; Van der Tol et al., 2005).

Basically, there are two types of robots for manure removal. Cleaners are used to push manure between the slots of a slatted floor, while collectors can pick up manure and dump the manure at specified locations, such as the Discovery (Lely Industries NV, Maassluis, The Netherlands), which are the subject robots used in this research. Current routines in practice are that at specified times, the manure robots will start a cleaning route. At installation, the frequency per fixed route is set. In addition it allows for several different routes and corresponding frequencies, depending on the requirements at specific locations (Leinweber et al., 2019). Hence, several routes can be planned throughout the barn in a way that every part of the barn is cleaned in a tailor-made time schedule. These fixed and scheduled routes might on one side result in non-optimal cleanings, and on the other side cause cow-robot encounters with possible impaired welfare, as the dynamic environments, i.e. cow behaviors were not considered into the path planning (Corke, 2017). Rushen et al. (2004) found that cows are more likely to have physical injuries when being confronted with obstacles on a slippery floor. Doerfler et al. (2016) stated that experiencing prolonged stress by dairy cows can be associated with a decline in immune competence, health status, milk production and hence impaired welfare.

Since the 1950s, scientists have developed and implemented different path planning algorithms to arrive the target quickly, safely and accurately for various robotic applications (Liu et al., 2023), such as the vacuum cleaner robot (Yakoubi and Laskri, 2016), autonomous vehicles (Geisslinger et al., 2023), as well as the warehouse transport robot (Ishihara et al., 2022). According to their functions, path planning algorithms could be generally divided into three categories: classical algorithms (e.g. Dijkstra algorithm), bionic algorithms (e.g. Genetic Algorithm) and artificial intelligence algorithms (e.g. Neural Network) (Liu et al., 2023). Yakoubi and Laskri (2016) proposed a path planning algorithm for vacuum cleaner robot for coverage region using Genetic Algorithm. The work from (Yakoubi and Laskri, 2016) can make the robot pass through every part of the environment by avoiding obstacles using different sensors, which did not consider priorities of these areas. Geisslinger et al. (2023) proposed an ethical trajectory planning algorithm with a framework that aims at a fair distribution of risk among road users. Long-short-term-memory-based neural network model has been used to train the trajectories of the common road scenarios. Four-step planning method was used for ethical decision-making. However, generalized the proposed approach to different countries, cultures or even individuals has not been achieved yet. Ishihara et al. (2022) applied a path planning algorithm for multiple robots in warehouses using Model Predictive Control. Deadlock can be avoided only for static obstacles when each robot travels on different aisle. Dynamic obstacles have not been considered yet.

Compared to path planning for robots in other domains, development and research on path planning for manure-robot cleaners is limited at the time of writing this article. Currently, most of the robot cleaners drive at predefined and programmed routes, which seems an optimal cleaning strategy, but is not energy efficient and does not consider cow behavior (De Baerdemaeker, 2013). According to the special characteristics of dairy barns with uneven distribution of dirtiness, dynamic obstacles from the moving cows, many varieties of the environment (each barn has different layout), this research is seeking for a dynamic path planning approach which can address these challenges and be applicable to different scenarios.

Reinforcement Learning (RL) is a machine learning method that can learn an optimal decision through interacting with its dynamic environment and can keep improving its decision making policy by rewarding desired behavior or punishing undesired behavior of an agent (Bertsekas, 2019). The motivation of using RL in this research is its ability to learn an optimal path from the data at hand and its adaptability to various use cases and different scenarios due to its trial-and-error planning strategy. There are actually many different types of RL methods applicable for path planning (Singh et al., 2023). Lei et al. (2018) proposed a dynamic path planning of unknown environment for mobile robot using double Q-network deep RL mainly for avoiding static obstacles, without prioritizing different areas. Cui et al. (2021) used actor–critic RL for navigation of autonomous mobile robot in dynamic pedestrian environments based on laser sequence. Areas in the environment are considered evenly. As in the dairy barn, cow behaviors in terms of defecation and location are easily modeled into heatmaps with different priorities for each grid, the path planning problem of manure-robot cleaner can be easily formulated into a single-shot grid-based path finding problem. Therefore, this research is motivated to utilize grid-based Reinforcement Learning (Panov et al., 2017; Notsu et al., 2020; Moon et al., 2022) as its main approach.

The contributions of this research include:

- **Models of Cow Locations and Defecation Behaviors.** Based on observational data collected on cow locations and defecation behaviors, heatmaps were created for three-hour time slots during the day. These models provide detailed insights into cow movements and manure distribution patterns.
- **A Dynamic Path Planning Strategy for a Robot Cleaner.** The path planning strategy is achieved using grid-based RL, which can address variable and dynamic planning environment using only data inputs.
- **Incorporation of Cow Location Information and Defecation Behaviors.** The heatmaps of cow location information and defecation behaviors have been used as inputs for path planning algorithm, to generate a cleaning route for each time slot, which is efficient and minimally disruptive to the cows.
- **Simulation these proposed approaches.** The proposed approaches were tested in a simulated environment with various scenarios and compared with both the current programmed fixed cleaning routes and the ideal routes generated by simulated annealing for traveling salesman problem (TSP) (Zhan et al., 2016).

The simulation were crucial to assess the algorithms' effectiveness before implementing them on physical robots in real life situations. To ensure the reliability of our simulation results, the experimental setup and data collection was ensured to closely mirror real-world conditions from an experimental barn at the Dairy Campus of Wageningen Livestock and Research located in Leeuwarden, the Netherlands. The simulation performance were compared with the current cleaning approach in practice and the simulated annealing method for traveling salesman problem. Based on the simulation results, this research provides valuable insights and recommendations to the livestock robot industry, since current manure cleaning robots lack flexible path planning capabilities that can be tailored to specific needs.

2. Material and methods

Observational data was obtained from the Dairy Campus of Wageningen Livestock and Research located in Leeuwarden, The Netherlands. The Dairy campus functions as a commercial dairy farm as well as a research center for applied projects. The farm has the capacity for about 550 milking cows in 7 barns. Milking was done using a 16-unit rotary parlour, in which cows were milked two times a day between 05:00–08:00 and between 15:30–18:00. During these times, all cows were gathered into a waiting area before getting milked. On average, the cows at the dairy campus produce 9198 kg of milk per

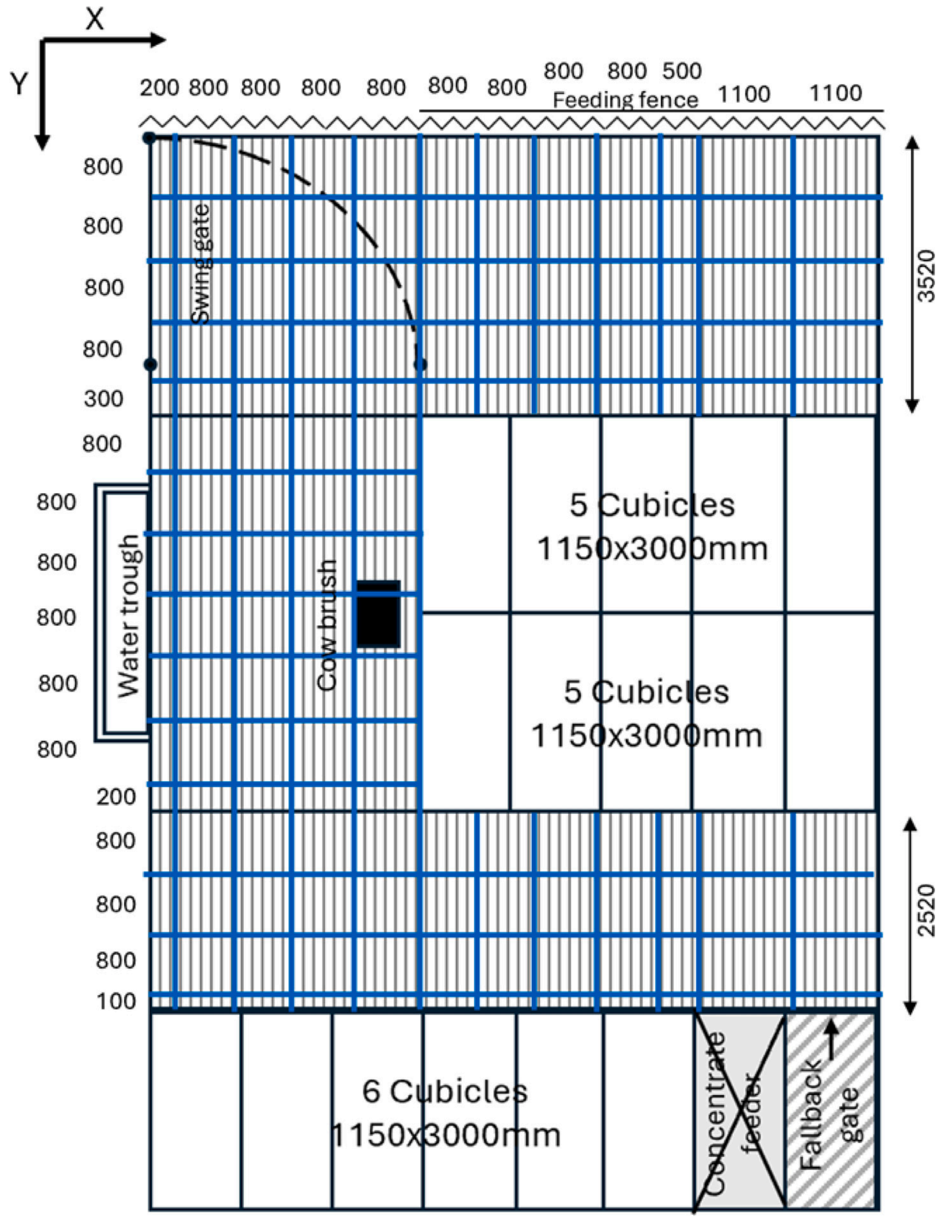


Fig. 1. Overview of space with the locations of cubicles, slatted floor, concentrate feeder, water trough and cow brush. The slatted floor is divided in smaller areas indicated by the blue lines. All sizes are in mm.

year (2022) with 4.54% fat and 3.58% protein. Dairy cows exhibit social behaviors that influence their defecation patterns, which varies across different breeds (Rocha et al., 2020). The defecation behavior may not be random, as certain areas in the barn, known as 'hotspots' are used more frequently (Oudshoorn et al., 2008). Therefore, a detailed description of the cows is deemed necessary and provided below.

2.1. Animals breed and description

Over a period of 7 days, between 1 February 2023 until 8 February 2023, we gathered experimental data in one specific section housing 16 Holstein-Friesian cows. The cows in this section on average had 188 lactation days and were 5 years old. The section provided an area of 11.0 m times 9.4 m, 16 cubicles and slatted floor walking paths of sufficient width (Fig. 1). The slatted floor does not have uneven terrain. The cows did not have access to pasture and were fed with grass and maize, with access to one concentrate feeder during the day. Water was provided ad libitum via a water trough. The cows could also use a cow brush located opposite to the water trough (Fig. 1). According to the Dutch legislation on animal experiments (Wet op de Dierproeven; WoD), this observational study was not considered as an animal experiment. Therefore, no approval from the legal authorities was required.

2.2. Experimental setup and data collection

A video camera (Axis, P1375-E) was mounted above the water trough, from which all walking areas of interest (being the slatted floor) could be seen. The camera recorded continuously and directly stored the material in an online database in sections of one hour. During the darker moments, normal management procedures, lights were automatically on that ensured good recordings without interfering with the normal cow behavior. To determine the number of defecations per grid area, video recordings were made of the barn section from 05:00 to 19:00. The recordings were watched with a playback speed four times faster than the original recording.

Sewio Leonardo Personal tags (SEWIO, Brno, Czech Republic) were used to determine cow traffic and the spatial density of cows in time

Fig. 2. Map of the section of consideration. The green area is the entrance of the barn section. The red area is the concentrate feeder, the yellow areas are the cubicles and the grey area is the slatted floor. Note: the water trough is not visible in this figure.

by means of location measurements. The tags were mounted on the neck collars of the cows and were positioned at the dorsal side of the neck. The SEWIO UWB system consisted of anchors, tags and the software RTLS Studio and used a UWB technology based on the time difference of arrival. Sewio tags allow to measure cow location up to 0.1 meter and continuous measurements are possible (D'Urso et al., 2023). The cow locations were averaged per minute and stored in a database together with the corresponding identification number of the cow, and date–time stamp of the respective measurements.

2.3. Methods

A simple grid was made in Microsoft Excel and put as an overlay on the barn section (Fig. 2) to be used for the heatmaps of cow locations and defecations. The initial grid size was set to be 0.1 m wide and 0.1 m in length, which matches the scale of measurable cow location from Sewio tags.

In this research the Discovery 90S (Lely Industries NV, Maassluis, The Netherlands), was the targeted robot, whose dimensions are 88 cm wide and 127 cm long. To improve planning and computational efficiency, the grid sizes were reorganized. In that way the majority of grid locations had the dimensions of 0.8 m × 0.8 m related to the cleaning width of the manure robot. The grid sizes only deviated at the concentrate feeder and barn entrance, measuring 1.1 m × 0.8 m, and at the edges of the slatted floor, where they had variable lengths and wides (Fig. 1). This barn lay-out is realistic and can be used to simulate a route for the manure robot.

2.3.1. Modeling cow behaviors using heatmaps

In the Microsoft Excel grid the cow locations at group level were added to this grid overlay for every hour, which resulted in 13 heatmaps of cow densities per day. Cow locations, from 05:00 to 18:00 only, were put into heatmaps through assigning the number of cows per area for a time period of 3 h (density per 3-h) to the corresponding grid location. Only the final hour of the day (17:00 to 18:00) was not summed to a 3-h period. This resulted in 4 cow location densities heatmaps of 3 h and 1 heatmap of one hour per day. Subsequently these were averaged over the week of the experimental period to get a good impression of daily cow behavior. In addition to that, it reduced variability and required less real-time data collection and computational power. The weekly average day behavior was determined according to Eq. (1):

$$C^j(t) = 1/d \times \sum_{i=1}^{d} C_i^j(t) \tag{1}$$

where $C^j(t)$ is number of cows at the jth grid at period t, d is number of days ($d = 7$). The remaining time per day (19:00 to 05:00) was not used for the analysis nor in the cow location heatmap due to less activity of the cow group.

The number of defecations per area were manually counted from the video recordings made. The number of defecations were assigned to the corresponding grid. In case the animal was walking, only the start location of defecation was noted. The defecation behavior changed over the day (Robichaud et al., 2011) and was observed in 13 periods of 1-h. The dirtiness per grid location was obtained by Eq. (2):

$$DS^j(t) = 1/d \times \sum_{i=1}^{d} DS_i^j(t) \tag{2}$$

where $DS^j(t)$ is the number of defecations for the jth grid at period t, $t = 13$, while d is the number of days ($d = 7$).

2.3.2. Resize of heatmaps

In practice, the cleaning schedule in a dairy barn was every 3 h on average. Therefore, we resized and summed all droppings per 3-h (except the last one-hour heatmap), maintaining the same temporal-spatial resolution as the cow spatial density map. Due to the uneven amount of heatmaps before addition, the last heatmap formed an exception and no additional hours could be added. In total, 5 heatmaps were produced using Eq. (3):

$$D^j(t) = 1/3 \times \sum_{i=1+3(t-1)}^{t*3} DS^j(i) \tag{3}$$

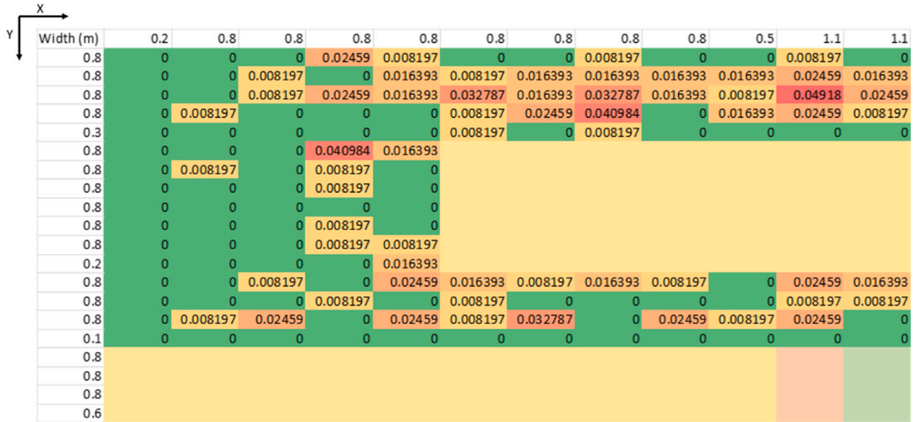
where $D^j(t)$ is the average number of defecations for the jth grid at period t, here $t = 4$, when $t = 5$, we used $D^j(5) = DS^j(13)$ to compute the defecation number after resizing. After resizing and averaging, the values per grid for a heatmap were divided by the total number of defecations for that timeslot. This resulted in a heatmap showing the distribution of manure over the total floor area (Fig. 3). The higher probabilities are indicated with red colors. Data on cow locations were used to assign the number of cows to a specific grid. Here, grid sizes were equal to the resized grids. The values per grid were divided by the total number of cow location measurements during the timeslot, resulting in a heatmap showing the distribution of cows over the total area. Finally, this process yielded a heatmap similar to the heatmap shown earlier in this section.

Afterwards, the heatmaps on manure distribution and cow locations were used to yield the optimal path planning given the distribution of cows and manure through a developed Python script.

2.3.3. Cleaning procedure

The grid with locations was transformed to distance matrix in which the fixed distances between the center of grids was entered. This way a route for cleaning calculated in computer simulations contained real distances and would be as efficient as in the real world.

The idea is to first build a list of grid-locations in need of cleaning, then check the cow density to build a "cleaning list" of grid locations that will be used for programming a route. High cow density grids will be avoided, creating a "priority list". A global path is planned in which the robot cleaner goes from one grid to the next on the priority grid



Width (m)	0.2	0.8	0.8	0.8	0.8	0.8	0.8	0.8	0.8	0.5	1.1	1.1
0.8	0	0	0	0.02459	0.008197	0	0	0.008197	0	0	0.008197	0
0.8	0	0	0.008197	0	0.016393	0.008197	0.016393	0.016393	0.016393	0.016393	0.02459	0.016393
0.8	0	0	0.008197	0.02459	0.016393	0.032787	0.016393	0.032787	0.016393	0.008197	0.04918	0.02459
0.8	0	0.008197	0	0	0	0.008197	0.02459	0.040984	0	0.016393	0.02459	0.008197
0.3	0	0	0	0	0	0.008197	0	0.008197	0	0	0	0
0.8	0	0	0	0.040984	0.016393							
0.8	0	0.008197	0	0.008197	0							
0.8	0	0	0	0.008197	0							
0.8	0	0	0	0	0							
0.8	0	0	0	0.008197	0							
0.8	0	0	0	0.008197	0.008197							
0.2	0	0	0	0	0.016393							
0.8	0	0	0.008197	0	0.02459	0.016393	0.008197	0.016393	0.008197	0	0.02459	0.016393
0.8	0	0	0	0.008197	0	0.008197	0	0	0	0	0.008197	0.008197
0.8	0	0.008197	0.02459	0	0.02459	0.008197	0.032787	0	0.02459	0.008197	0.02459	0
0.1	0	0	0	0	0	0	0	0	0	0	0	0
0.8												
0.8												
0.8												
0.6												

Fig. 3. Heatmap showing the proportional distribution of manure for the timeslot of 05:00–08:00 in %. The sum of all cells add up to 100% of all droppings for this timeslot.

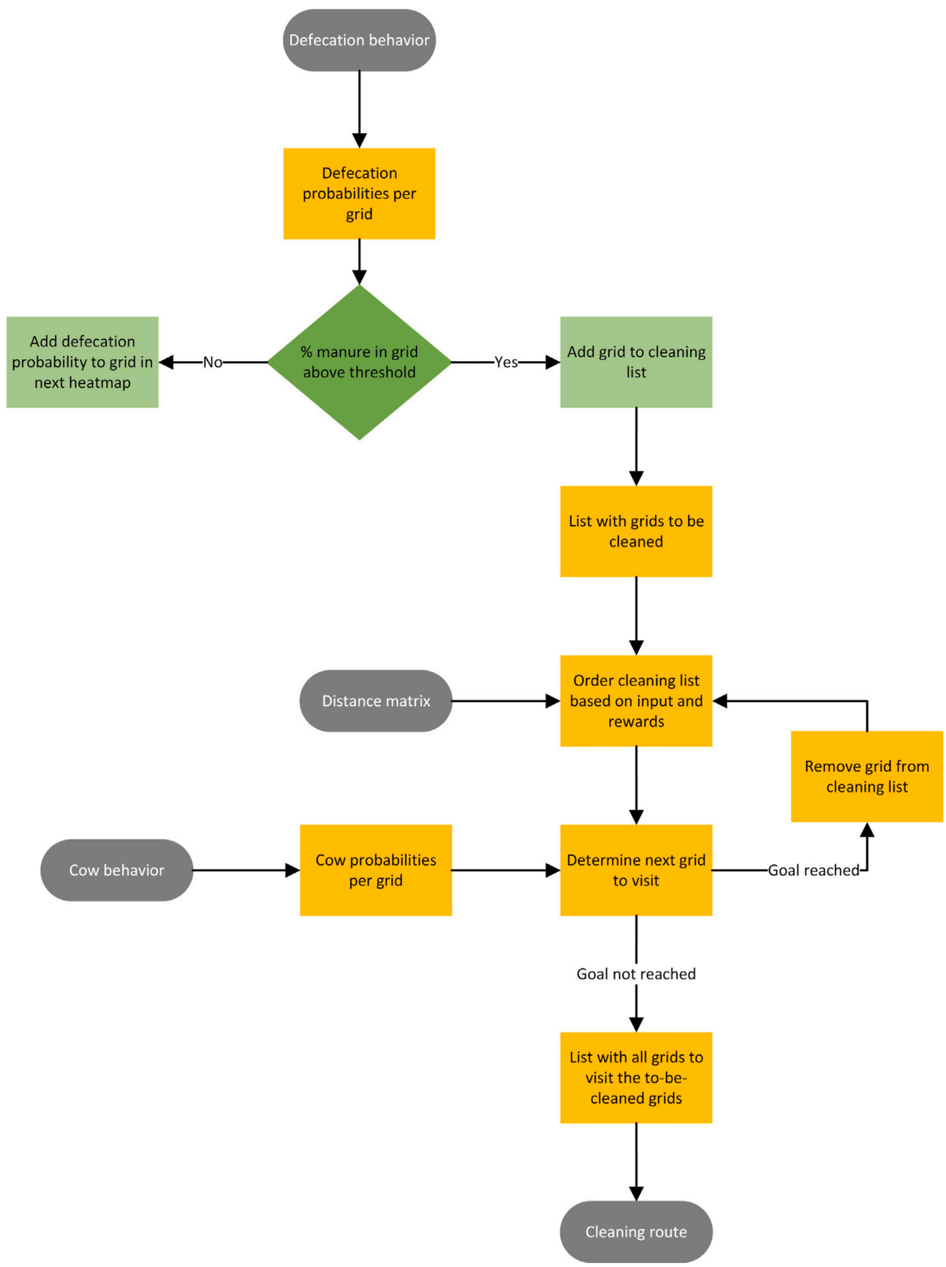


Fig. 4. Flowchart of the cleaning procedure. Rectangles indicate processes, diamonds indicate decisions and round shapes indicate inputs/outputs.

location in the shortest possible route. When priority grids are not next to one another, as a safety measure, the robot cleaner uses local path planning and checks in its adjacent grids what is the best grid possible to follow its route to the next priority grid to clean (≈ lowest cow density). Finally, the remaining grids are passed on the next iteration of route planning. The flowchart in Fig. 4 gives more details in how the system operates using both global and local path planning (Xing et al., 2022).

In detail, when a grid location contains more than 1 percent of the total number of manure droppings, while this area is occupied with less than 1 percent of the total dairy herd standing on the floor, then the area needs cleaning. Hence, the grid location is added to the "cleaning-list" and was used for path planning of the robot cleaner. In case a grid location was not cleaned in the initial route, due to "too high cow density", the number of manure droppings the corresponding grid will be added to the next cleaning in line according to Eq. (4):

$$TD^j(t) = D^j(t) + D^j(t-1) \tag{4}$$

where $TD^j(t)$ is the total number of defecations for the jth grid at timeslot t, $D^j(t)$ is the number of defecations for the jth grid at timeslot t and $D^j(t-1)$ is the number of defecations for the jth grid at the previous timeslot $t-1$. Afterwards, the distance between subsequent priority location grids was calculated using Eq. (4).

For all grids, through Eq. (4), the shortest distance between one grid to another grid was determined, which was used in the path planning to find order of grids to clean defining the most efficient route, hence the shortest distance to travel. The calculation of the shortest distance D and the matrix formulation were achieved by Eq. (5):

$$D_{x\to y} = min(D_{x\to y}, D_{x\to z} + D_{z\to y}) \tag{5}$$

where $D_{x\to y}$ is the distance from point x to point y, $D_{x\to z}$ is the distance from point x to point z and $D_{z\to y}$ is the distance from point z to point y.

2.3.4. Path planning using grid-based RL

Single-shot grid-based path finding is an important problem with the application in robotics. Typically in the Artificial Intelligent community, heuristic search methods are used to solve it. In this research, the grid-based path finding tasks were coped with using the well-known RL statement.

Reinforcement Learning is a subset of machine learning that evaluates actions based on rewards or punishments. Q-learning, a specific RL algorithm, aims to determine an optimal action-selection policy for any finite Markov decision process (MDP). It achieves this by maximizing the total reward over time through repeated interactions with the environment, even when the model of that environment is not known (Maoudj and Hentout, 2020). In Q-learning, $Q(S_t, A_t)$ represents the value of taking action A_t in state S_t at time t. This Q value is updated using information from the environment and a reward r as follows (Sutton and Barto, 2018):

$$Q(S_t, A_t) = Q(S_t, A_t) + \alpha[R_{t+1} + \gamma \max_a Q(S_{t+1}, a) - Q(S_t, A_t)] \tag{6}$$

where α is the learning rate and γ is the discount factor.

In the path planning problem of a manure-robot cleaner in a diary barn, The set of states consists of all the grids on the barn floor. The set of actions (the optimal cleaning order) is determined by the number and location of the grids that need to be cleaned, as defined by Eqs. (2) and (4). The RL agent in this case is the manure-robot cleaner, and the environment is the heatmap for a specific time period, which changes throughout the day. Information about which grids needed cleaning and their mutual distances was included. With knowledge of the grids to be cleaned and the distances between them, path planning could be simulated.

To find the most optimal path, the order of the grids to be cleaned was determined by using a Q-value function. The value function compared each grid (the state variable in this RL problem) during the route, and then chose the grid which was closest to the current grid (state). This finally yielded the most optimal order of grids to clean, given their mutual distances. Therefore, the output after using RL was a list, containing the same grids as the cleaning-list, but now sorted to yield to the most efficient path when executed. The Q-value function used is defined in Eq. (7), based on the basic format of the value function (6) introduced by Sutton and Barto (2018), where the learning rate α and discount factor γ are respectively chosen to be 0.1 and 0.4 through trial-and-error.

$$Q_t(S, A) = \alpha \times (ir + \gamma \times dr - Q_{t-1}(S, A)) \tag{7}$$

where $Q_t(S, A)$ is the value for a state and action at time step t and $Q_{(t-1)}(S, A)$ is the previous value given the state and action. The ir is the immediate reward which can be expressed as $1/(distance\ between\ state\ and\ action)$. The dr is the delayed reward which can be expressed as $1/(highest\ reward\ of\ the\ previous\ state)$. The equation also includes the learning rate α and discount factor γ.

To ensure a continuous and efficient path planning given the locations of cows and the distribution of manure, intermediate points between grids which needed cleaning (intermediate state) were also considered. For each of the four neighboring grids, it was checked whether mutual distance to the grid which needed cleaning decreased. In case the distance decreased, the best option was then selected based on the lowest probability of cow robot encounters and added to the path planning list. Finally, this yields a continuous and efficient path planning given the locations of cows and the distribution of manure.

2.3.5. Simulated annealing for traveling salesman problem

The traveling salesman problem, a member of the class of Non-deterministic Polynomial-Time Complete problems, seeks to find the shortest route that visits each city exactly once and returns to the city it started in Zhan et al. (2016) (Bookstaber, 1997). In the context of path planning of manure-robot cleaners on grids that should to be cleaned, this problem can be effectively formulated as a traveling salesman problem. Simulated annealing, originally introduced as a search algorithm for combinatorial optimization problems (Kirkpatrick et al., 1983), is a widely adopted iterative metaheuristic. Its distinguishing feature lies in its ability to escape local optima by allowing hill-climbing moves to potentially find a global optimum (Zhan et al., 2016). In this study, simulated annealing is employed to evaluate and compare the performance of a proposed dynamic path planning approach using grid-based RL, in addition to the current fixed programmed cleaning method.

To apply simulated annealing to the path planning problem, the state space consists of the set of grids that need to be cleaned, similar to applying the RL approach. Any path that includes grids to be traversed is considered valid. The detailed mathematical prototype can be refereed from (Bookstaber, 1997). The cleaning routes produced by simulated annealing for TSP is considered as the ideal benchmark with shortest route to evaluate performance of the grid-based RL.

2.3.6. Scenarios for cleaning

Four scenarios will be used for path planning; Scenario (1) standard fixed programmed route, Scenario (2) grid-based RL at feeding time, and Scenario (3) grid-based RL at milking time, (4) simulated annealing for traveling salesman problem. The current cleaning approach is performed by cleaning all the areas every 3 h. The new strategies in Scenario 2, 3 and 4 is cleaning only the areas which comply with the thresholds mentioned previously in two ways. Every 3 h these 4 scenarios were used for comparisons. For the Scenario 2, the heatmaps of cow locations were all set to be equal to one specific heatmap in which most cows were next to the feeding fence. Based on this, path planning was performed and performance was evaluated. For the Scenario 3, all values of the heatmaps of cow locations were set to 0, to mimic situations in which cleaning is performed when cows are being milked.

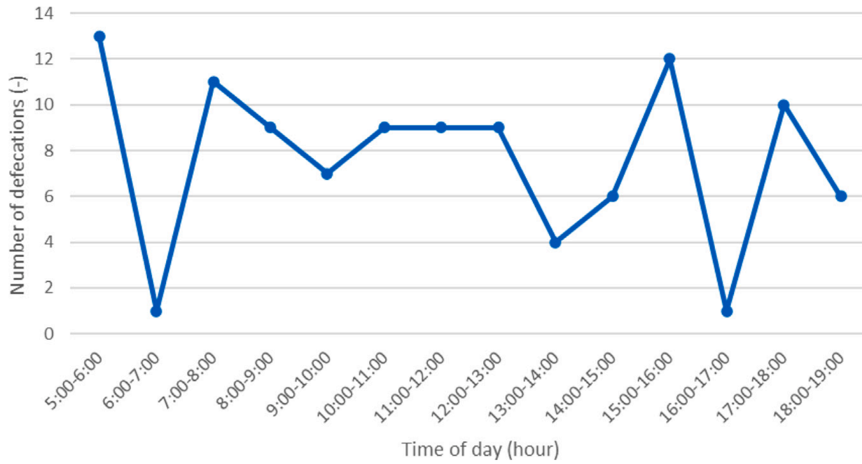


Fig. 5. Total number of defecations on 2 February 2023 during different times of the day. Cows were milked between 05:00–08:00 and between 15:30–18:00, roughage was provided at 8:00 and feed was pushed to feedfence at 15:00.

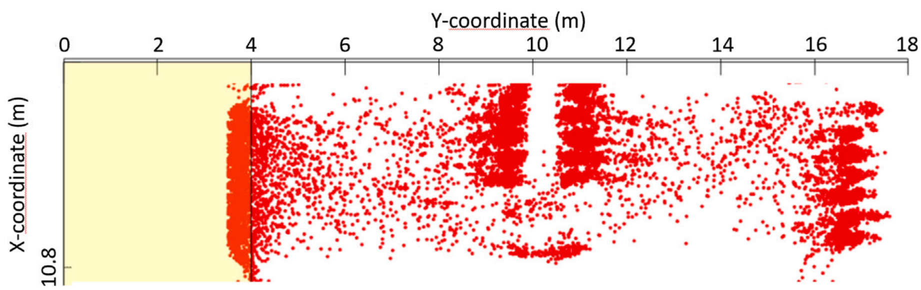


Fig. 6. Overview of cow locations of 2 February 2023.

2.4. Evaluation performance

A comparison was made to evaluate performance among the current way of cleaning by manure-cleaning robots, the newly proposed dynamic path planning strategy, as well as the simulated annealing for traveling salesman problem. The metrics used are: percentage of manure droppings cleaned, the estimated probability of cow-robot encounters, the total length of the programmed route, as well as the cleaning efficiency, which is computed by dividing the percentage of manure droppings cleaned by the total length of the programmed route. The performance measures were selected since they provided a thorough and quantifiable assessment of path planning performance in terms of cow health, based on cleaning performance and collisions, as well as cleaning efficiency, in terms of route length. The percentage of manure droppings was quantified by the proportion of droppings that were removed. The estimated probability of cow-robot encounters was found by summing the cow-robot probabilities per grid for the entire proposed path. The length of the program was found by summing up the number of visited grids. Distances between grids were assumed to be equal.

3. Results and discussion

As one recording of 1 February between 05:00 and 06:00 is missing. Therefore, for the accuracy of averaging, the decision is made to include only 6 out of the 7 days from this time period. In total, 90 h of video recordings were stored and used during later processing.

3.1. Video observations

Fig. 5 presents an overview of defecation numbers on February 2, 2023 during different times of the day. The figure shows that the number of defecation increases in the time slot before milking, for example around 15:00. This increase can be attributed to the higher level of activity and stress within the herd before milking, which is align with results presented by Aland et al. (2002). Conversely, the number of defecation activities decreased after milking. Additionally, it was found that most defecation activities occurred near the main walking route of cows, often just after standing up and leaving their cubicle. In contrast, almost no defecation were found in the first two lines (1.6 m) from the feeding fence. The cows spent a significant amount of time lying or standing idle in the cubicles.

Data on cow locations were also examined, and in some cases, likely due to interference of the triangulation of tags caused by ironwork in the barn, the cow locations were adjusted to fit the grid of the walking area. For example, when the location of a cow was found beyond the feeding fence ($y \approx 3.5\ m$), 0.5 m was added to the y-locations to correct for the cow's position at the feeding fence. As the cows did not wander around much, their activities were generally purposeful, such as going to the feed fence, water trough, or concentrate dispenser. The density plot of cow locations (Fig. 6) intuitively shows where most activities took place. In this figure, a clear pattern can be seen of cows either standing or lying inside cubicles or standing near the feeding fence or concentrate feeder.

3.2. Cleaning strategies results

The video recordings were analyzed to track defecation time and locations over time, and this data was processed into heatmaps based on the layout of the barn section. Path planning was simulated five times for specific time periods, and the results were averaged to determine the percentage of manure cleaned, the probability of cow-robot collisions, and the total route length, using three different path planning approaches throughout the day.

Table 1
Comparison of the current path planning strategy with the proposed new path planning strategy.

Timeslot	Cleaning percentage (%)			Collision probability (%)			Route length (m)			Cleaning Efficiency (%)		
	S.1	S.2	S.4	S.1	S.2	S.4	S.1	S.2	S.4	S.1	S.2	S.4
05:00–08:00	100	86.1	83.2	100	19.2	16.4	142	92.4	70.2	70	93	118
08:00–11:00	100	88.2	85.3	100	34.5	20.8	142	123.8	85.4	70	72	99
11:00–14:00	100	87.4	82.3	100	41.3	30.1	142	150.8	95.4	70	60	86
14:00–17:00	100	82.5	77.5	100	34.0	32.1	142	122.4	82.2	70	67	94
17:00–18:00	100	82.9	79.7	100	33.0	22.1	142	143.6	93.4	70	58	85
Average	100	85.4	81.6	100	32.4	24.3	142	126.4	85.3	70	69.6	96

From 5:00 to 18:00, the average percentage of manure cleaned for the Scenario 1: current, Scenario 2: new dynamic at feeding time, Scenario 3: new dynamic at milking time, and Scenario 4: simulated annealing for traveling salesman problem was 100%, 85.4%, 85.5% and 81.6% respectively. The probability of cow-robot collisions was 100%, 32.4%, 0%, and 24.3% for the four scenarios respectively. In scenario 3, when no cows were present in the barn, the probability of cow-robot collisions was set to 0 for each time period.

Table 1 provides an overview of the results for all simulations and time periods, comparing Scenario 1 (abbreviated to $S.1$), Scenario 2 (abbreviated to $S.2$) and Scenario 4 (abbreviated to $S.4$). As Scenario 3 is a special case of Scenario 2, which will not be included in this overview comparison. It is evident that the number of collisions decreased to approximately one-third with the fixed programming route, while the cleaning proportions were similar or even improved during higher cow density periods.

In Scenario 1, the total probability of cow-robot encounters was higher when more cows were present at the feeding fence. Conversely, when no cows were present (e.g. cows were in the waiting area or being milked), more effective and undisturbed cleaning was possible. Compared to the fixed routes in Scenario 1, the grid-based RL incorporating cow behavior in Scenario 2 resulted in a potential 67.6% decrease in the total probability of cow-robot encounters, leading to fewer collisions. While Scenario 1 achieved a 100% manure cleaning rate by covering the entire section, it resulted in longer paths, potential energy inefficiency, and more cow-robot encounters. Dynamic path planning in Scenario 2 achieved a manure cleaning percentage ranging from 82.5% to 88.2%. Utilizing features such as cleaning locations in subsequent iterations could further increase the cleaning percentage and establish a minimal threshold in the reward function of the RL path planning. There was little difference in the total path length needed to clean the barn between the current and dynamic scenarios on average. The cleaning percentage of Scenario 2 using Grid-based RL is slightly better than Scenario 4, the simulated annealing for traveling salesman problem with 5% improvement in average. Scenario 4 has less collision probability due to the shortest routes it takes according to its planning strategy. However, if we compare the collision probabilities for these three scenarios at each unit of length, which are 70%, 25% and 27%, Scenario 2 is the solution with the lowest collision probability.

Figs. 7, 8, and 9 illustrate the final path planning results in a graphical manner for Scenarios 1, 2 and 4 during the time slot of 05:00–08:00 as an example, providing further interpretations of the analyzed results. Among these figures, the *Cleaning list* sub-figures provide the grids that need to be cleaned at different scenarios. The grids with blue nodes at the *Final route* sub-figures are the grids that are to be cleaned, while grids with orange nodes are the grids that are passed by the robots. In this work, it is assumed that in all scenarios, the robot cleaner starts from the same location near its charging station and returns to the charging point after completing the cleaning. As shown in Figs. 8 and 9, there are more grids passed by scenario 3 than scenario 4, which is the ideal route with minimal distance for this TSP problem.

3.3. Path planning and animal behavior

Path planning involves finding an optimal route between start and final destinations, considering specific routing criteria (Karur et al., 2021). The current method, pre-programmed routes, usually the robot has cow-avoidance algorithms on board to take a detour to prevent major collisions. In this research, global path planning, utilizing previous map knowledge, was combined with local path planning (Xing et al., 2022). Local path planning checked adjacent grids' cow density to determine the path, aiming to lower the probability of robot-cow encounters. Future approaches could benefit from involving intermediate steps in global path planning based on cow densities. Our proposed combination of global and local path planning, using averaged cow densities and relatively coarse grid dimensions, demonstrated effectiveness.

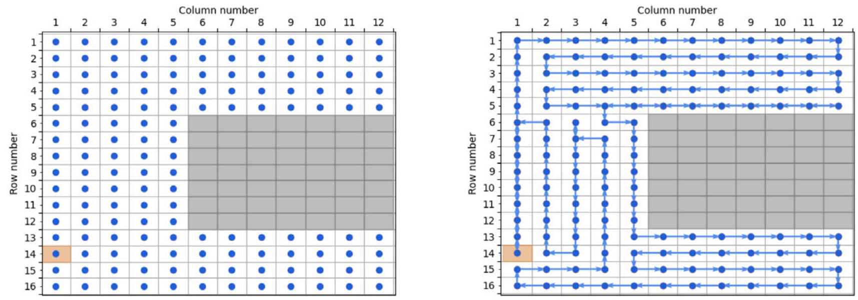
Despite Leinweber et al.'s findings 2019, which showed avoidance behaviors in cow-robot encounters, our cows, already familiar with a manure-cleaning robot, did not exhibit avoidant behavior. The Sewio tags and location data, validated by D'Urso et al. (2023), provided accurate information. Defecation observations, scored in the initial grid, were considered valid for path planning, even with grid resizing.

3.4. Limitations and recommendations

As an initial exploration of data-based dynamic path planning methods, the current approach indeed has limitations that can be addressed in future research:

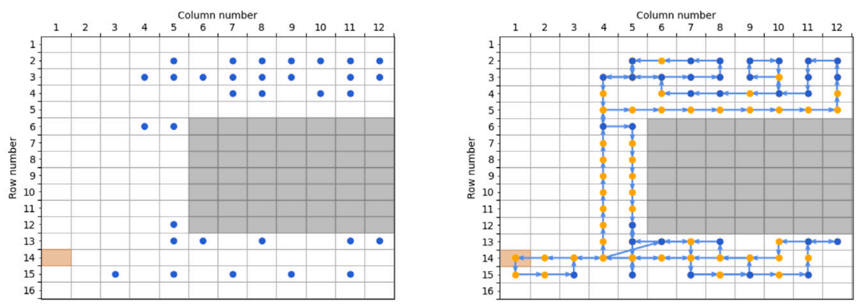
- The current experiment was conducted in a single barn during a week in one season. Given the significant influence of barn layout, milking and feeding systems on cow behavior, more experiment are needed to achieve more generalized findings.
- The modeling of defecation behavior could benefit from including more dairy herds to enhance the generalizability of the approach.
- The current implementation remains primarily in a simulation environment and requires further validation in practical settings.
- Only single operating robot is considered, which might not be the case in larger scale commercial barns.

In future research, real-time communication with sensors and automatic detection of defecation could significantly enhance path planning. This includes utilizing real-time behavior measurements for dynamic path adjustments and investigating how environmental factors influence behavior and path planning. Addressing the unpredictability of the manure-robot cleaner's path could reduce stress among dairy herds. It is recommended to use consistent protocols for the robot cleaner how to avoid animal encounters, which helps cows better understand as it makes predictable movements. Further refinement of cow behavior models and the consideration of urination patterns, solid floors, and grazing areas could lead to more effective cleaning and path planning strategies. Additionally, in large-scale dairy operations with numerous cows (e.g., averaging around 120 animals in the Netherlands), increased monitoring data is essential for accurately modeling cow behaviors. Looking ahead, implementing multiple cleaning robots that work cooperatively could offer a more robust and efficient solution, promoting energy efficiency and improving the overall welfare of cows in barn environments.



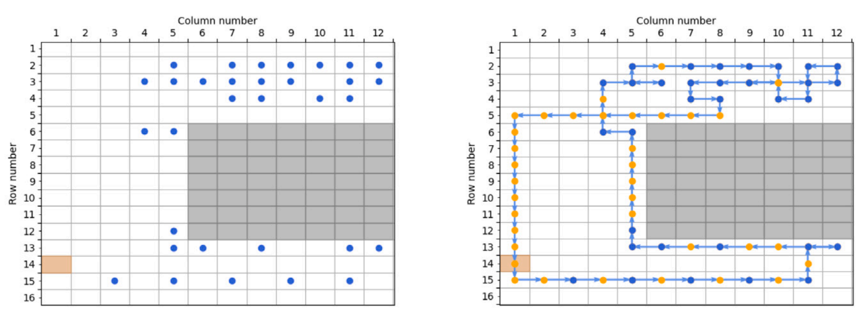
(a) Cleaning list (b) Final route

Fig. 7. Scenario 1: Fixed routes, (a) grids to clean (blue dots), and (b) the programmed route (blue arrows connecting the dots) and both graphs show the start grid (orange fill).



(a) Cleaning list (b) Final route

Fig. 8. Scenario 3: cleaning approach using grid-based RL, (a) grids to clean (blue dots), and (b) the programmed route (blue arrows connecting the blue dots). In addition the grids visited, but not on cleaning list (orange dots) and both graphs show the start grid (orange fill).



(a) Cleaning list (b) Final route

Fig. 9. Scenario 4: ideal routes produced by simulated annealing for TSP, (a) grids to clean (blue dots), and (b) the programmed route (blue arrows connecting the blue dots). In addition the grids visited, but not on cleaning list (orange dots) and both graphs show the start grid (orange fill).

4. Conclusion

In this research, dynamic path planning was executed while considering cow behaviors based on heatmaps of cow and defecation locations using grid-based RL. After validation in a simulated environment with multiple scenarios defined based on a barn at Dairy Campus of Wageningen Livestock Research, the proposed path planning approach could achieve a 67.6% decrease in cow-robot encounters while maintaining 85.4% of the cleaning performance compared to the current programmed fixed routes. Compared with the ideal routes generated by simulated annealing for traveling salesman problem, the proposed grid-based RL solution could still achieve 5% better cleaning performance. The collision probability at each unit length of grid-based RL is the lowest, which is 25%, comparing with 70% and 27% for the Scenario 1 and 4 respectively. The detailed results at the above sections demonstrated the effectiveness and efficiency of the proposed cleaning method. Moreover, due to the characteristics of RL, the proposed path planning approach can be implemented solely on data without the need for a process model. This enables more generalization of the proposed approach to different layouts and scenarios.

CRediT authorship contribution statement

Congcong Sun: Writing – review & editing, Writing – original draft, Validation, Resources, Project administration, Methodology, Investigation, Conceptualization. **Rik van der Tol:** Writing – review & editing, Writing – original draft, Methodology, Investigation, Resources. **Robin**

Melenhorst: Writing – review & editing, Writing – original draft, Validation, Software, Investigation, Methodology, Resources. **Luis Angel Ponce Pacheco:** Methodology, Software, Validation. **Peter Groot Koerkamp:** Writing – review & editing, Investigation, Conceptualization, Project administration, Resources.

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.

Data availability

Data will be made available on request.

Acknowledgments

The authors want to thank dr.ir. Wijbrand Ouweltjes from Wageningen Livestock Research group for his support, and Dairy Campus for providing observational data. The authors also want to thank Dr. Tim Hoogstad, Salma Rian, Dr. Rachel van Ooteghem for their valuable inputs on preparing an informative graphical abstract.

References

Aland, A., Lidfors, L., Ekesbo, I., 2002. Diurnal distribution of dairy cow defecation and urination. Appl. Animal Behav. Sci. 78 (1), 43–54.

Barker, Z.E., Leach, K.A., Whay, H.R., Bell, N.J., Main, D.C.J., 2010. Assessment of lameness prevalence and associated risk factors in dairy herds in England and Wales. J. Dairy Sci. 93 (3), 932–941.

Bertsekas, D.P., 2019. Reinforcement Learning and Optimal Control. Athena Scientific, Nashua, USA.

Bookstaber, D., 1997. Simulated annealing for the traveling salesman problem. pp. 1–9. http://dx.doi.org/10.13140/RG.2.2.26488.39682, Preprint.

Britt, J.H., Cushman, R.A., Dechow, C.D., Dobson, H., Humblot, P., Hutjens, M.F., Jones, G.A., Ruegg, P.S., Sheldon, I.M., Stevenson, J.S., 2018. Invited review: Learning from the future—A vision for dairy farms and cows in 2067. J. Dairy Sci. 101 (5), 3722–3741.

Bruijnis, M.R.N., 2006. Weidegang en opstallen van melkvee, in English: Pasture Grazing and Housing for Dairy Cattle. Technical Report, Wageningen University & Research.

Cardoso, C.S., Hötzel, M.J., Weary, D.M., Robbins, J.A., von Keyserlingk, M.A.G., 2016. Imagining the ideal dairy farm. J. Dairy Sci. 99 (2), 1663–1671.

Corke, P., 2017. Robotics, Vision and Control: Fundamental Algorithms In MATLAB, second ed. Springer, Berlin, Germany.

Cui, Y., Zhang, H., Wang, Y., Xiong, R., 2021. Learning world transition model for socially aware robot navigation. In: 2021 IEEE International Conference on Robotics and Automation. ICRA 2021, Xián, China, pp. 9262–9268.

De Baerdemaeker, J., 2013. Precision agriculture technology and robotics for good agricultural practices. IFAC Proc. Vol. 46 (4), 1–4.

Doerfler, R.L., Lehermeier, C., Kliem, H., Möstl, E., Bernhardt, H., 2016. Physiological and behavioral responses of dairy cattle to the introduction of robot scrapers. Front. Vet. Sci. 3 (106), 1–11.

D'Urso, P.R., Arcidiacono, C., Pastell, M., Cascone, G., 2023. Assessment of a UWB real time location system for dairy cows' monitoring. Sensors 23 (10), 4873.

Geisslinger, M., Poszler, F., Lienkamp, M., 2023. An ethical trajectory planning algorithm for autonomous vehicles. Nature Mach. Learn. 5, 137–144.

Gieseke, D., Lambertz, C., Gauly, M., 2018. Relationship between herd size and measures of animal welfare on dairy cattle farms with freestall housing in Germany. J. Dairy Sci. 101 (8), 7397–7411.

Ishihara, S., Kanai, M., Narikawa, R., Ohtsuka, T., 2022. A proposal of path planning for robots in warehouses by model predictive control without using global paths. IFAC-PapersOnLine 55 (37), 573–578.

Karur, K., Sharma, N., Dharmatti, J., 2021. A survey of path planning algorithms for mobile robots. Vehicles 3 (3), 448–468.

Kirkpatrick, S., Gelatt, C.D., Vecchi, M.P., 1983. Optimization by simulated annealing. Science 220 (4598), 671–680.

Lei, X., Zhang, Z., Dong, P., 2018. Dynamic path planning of unknown environment based on deep reinforcement learning. J. Robotics 2018, 5781591.

Leinweber, T., Zähner, M., Schrade, S., 2019. Evaluation of a dung-removal robot for use in dairy housing from an ethological and process-engineering point of view. Landtechnik 74 (3), 55–68.

Liu, L., Wang, X., Yang, X., Liu, H., Li, J., Wang, P., 2023. Path planning techniques for mobile robots: Review and prospect. Expert Syst. Appl. 227 (2023), 120254.

Maoudj, A., Hentout, A., 2020. Optimal path planning approach based on Q-learning algorithm for mobile robots. Appl. Soft Comput. 97 (2020), 106796.

Moon, W., Park, B., Nengroo, S.H., Kim, T., Har, D., 2022. Path planning of cleaning robot with reinforcement learning. In: IEEE International Symposium on Robotic and Sensors Environments. ROSE, Abu Dhabi, United Arab Emirates, pp. 1–7.

Notsu, A., Yasuda, K., Ubukata, S., Honda, K., 2020. Online state space generation by a growing self-organizing map and differential learning for reinforcement learning. Appl. Soft Comput. 97(B), 106723.

Oudshoorn, F.W., Kristensen, T., Nadimi, E.S., 2008. Dairy cow defecation and urination frequency and spatial distribution in relation to time-limited grazing. Livestock Sci. 113 (1), 62–73.

Panov, A.I., Yakovlev, K.S., Suvorov, R., 2017. Grid path planning with deep reinforcement learning: Preliminary results. Procedia Comput. Sci. 123 (2018), 347–353.

Robichaud, M.V., de Passillé, A.M., Pellerin, D., Rushen, J., 2011. When and where do dairy cows defecate and urinate? J. Dairy Sci. 94 (10), 4889–4896.

Rocha, L.E., Terenius, O., Veissier, I., Meunier, B., Nielsen, P.P., 2020. Persistence of sociality in group dynamics of dairy cattle. Appl. Animal Behav. Sci. 223 (2020), 104921.

Rushen, J., de Passillé, A.M., Borderas, F., Tucker, C., Weary, D., 2004. Designing better environments for cows to walk and stand. Adv. Dairy Technol. 16, 55–64.

Singh, R., Ren, J., Lin, X., 2023. A review of deep reinforcement learning algorithms for mobile robot path planning. Vehicles 2023 (5), 1423–1451.

Solano, L., Barkema, H.W., Pajor, E.A., Mason, S., LeBlanc, S.J., Heyerhoff, J.C.Z., Nash, C.G.R., Haley, D.B., Vasseur, E., Pellerin, D., Rushen, J., de Passillé, A.M., Orsel, K., 2015. Prevalence of lameness and associated risk factors in Canadian Holstein-Friesian cows housed in freestall barns. J. Dairy Sci. 98 (10), 6978–6991.

Somers, J.G.C.J., Frankena, K., Noordhuizen-Stassen, E.N., Metz, J.H.M., 2003. Prevalence of claw disorders in dutch dairy cows exposed to several floor systems. J. Dairy Sci. 86 (6), 2082–2093.

Sutton, R.S., Barto, A.G., 2018. Reinforcement Learning An Introduction, second ed. The MIT Press, Cambridge, Massachusetts, USA.

Van der Tol, P.P.J., Metz, J.H.M., Noordhuizen-Stassen, E.N., Back, W., Braam, C.R., Weijs, W.A., 2005. Frictional forces required for unrestrained locomotion in dairy cattle. J. Dairy Sci. 88 (2), 615–624.

Xing, X., Ding, H., Liang, Z., Li, B., Yang, Z., 2022. Robot path planner based on deep reinforcement learning and the seeker optimization algorithm. Mechatronics 88, 102918.

Yakoubi, M.A., Laskri, M.T., 2016. The path planning of cleaner robot for coverage region using genetic algorithms. J. Innov. Dig. Ecosyst. 3 (1), 37–43.

Zhan, S.-h., Lin, J., Zhang, Z.-j., Zhong, Y.-w., 2016. List-based simulated annealing algorithm for traveling salesman problem. Comput. Intell. Neurosci. 2016, 1712630.