Sensing and Bio-Sensing Research 45 (2024) 100683

Contents lists available at ScienceDirect

Sensing and Bio-Sensing Research

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

Sensors driven system coupled with artificial intelligence for quality monitoring and HACCP in dairy production

Roberto Dragone [a,*], Gerardo Grasso [a,*], Giorgio Licciardi [b], Daniele Di Stefano [a], Chiara Frazzoli [c]

[a] Istituto per lo Studio dei Materiali Nanostrutturati, Sede Roma-Sapienza, Consiglio Nazionale delle Ricerche, P.le Aldo Moro 5, 00185 Rome, Italy
[b] Agenzia Spaziale Italiana (ASI), Via del Politecnico snc, 00133 Rome, Italy
[c] Dipartimento Malattie Cardiovascolari ed Endocrino-Metaboliche, e Invecchiamento, Istituto Superiore di Sanità, Via Giano della Bella, 34, 00162 Rome, Italy

ARTICLE INFO

Keywords:
Food safety
Milk monitoring
Animal welfare
Milk chain
Machine learning

ABSTRACT

The maintenance of good milk quality standards is still a challenge for dairy farmers that requires a rapid control system that is compatible with both the environment and production cost. A patented Hazard Analysis and Critical Control Points-like remote diagnostic (sensor driven) system named BEST was implemented to enable both quality monitoring and traceability in the dairy chain. BEST was daily tested in a dairy farm to identify new reliable indicators of anomalies (safety and quality) in milk production based on a Machine-Learning approach. The database obtained in four months of sensoristic analysis was subjected to a statistical study with AI algorithm to identify outliers. BEST proved ability to spot cows with specific characteristics in the whole herd's database. In particular, AI highlighted the sole cow from a different breed, the only cow that recently gave birth and the only cow in the herd that received treatment with the drug Micospectone® (Lincomycin + Spectinomycin).

1. Introduction

While world milk production is projected to increase by 177 million tons by 2025, dairy farming practices today are required to combine several requirements, including environmental sustainability, animal welfare, food quality and safety, and production efficiency. Food safety and traceability are crucial throughout the food chain, starting with primary production [1]. Ensuring safety involves preventive measures like Hazard Analysis and Critical Control Points (HACCP), which analyzes risks at critical control points (CCPs). Alternatively, a HACCP-like approach focusing on Points of Particular Attention (POPAs) can be used [2]. POPAs are conditions posing threats to animal and public health or on-farm management. Monitoring them, although not mandatory, enables better understanding and management. Currently, HACCP primarily addresses microbiological hazards but should expand to include chemical and physical hazards from farm to fork [2]. The EU advises primary producers, such as dairy farmers, to implement HACCP and HACCP-like programs to prevent milk-borne zoonoses, now including toxicological risks in animal-origin foods [3]. Indeed, milk is a food, an animal health bioindicator, and a process indicator, and hence a One Health matrix.

In the dairy supply chain, milk analyses are not routinely carried out at farm level, but in the laboratory (off-line analyses). In case of unwanted and/or unexpected events, e.g., due to chemical contamination, the disposal of huge quantities of milk would causes financial losses for both dairy farmers and milk processing industry. A 2017 European overview report of the general animal welfare conditions in EU's dairy farms indicates several issues, such as i) the lack of clear indicators useful for farmers to manage herds and improve productivity and animal welfare, ii) the lack of monitoring and assessment systems for existing indicators; iii) the lack of collection and analysis of data at farm level [4].

In recent years, focus shifted to on-line and at-line sensor systems for dairy farms [5,6], known as 'precision livestock farming' (PLF). PLF emphasizes individual animal management through new technologies, enabling a continuous real-time monitoring of animal health, welfare, and environmental impact. If implemented effectively, PLF could enhance efficiency and sustainability in dairy farming, improving animal welfare and facilitating traceability across the food supply chain [7]. To be most effective, PLF tools must be properly adapted to farmers' needs and skills; otherwise, they can even lead to negative impacts [8]. Noticeably, the lack of a proper PLF data collection and data elaboration

* Corresponding author.
E-mail addresses: roberto.dragone@cnr.it (R. Dragone), gerardo.grasso@cnr.it (G. Grasso).

https://doi.org/10.1016/j.sbsr.2024.100683
Received 9 May 2024; Received in revised form 15 August 2024; Accepted 4 September 2024
Available online 5 September 2024

is a hot point. The absence of appropriate adoption of targeted elaboration system for parameters that are continuously monitored would make them useless [9].

In this study, we discuss results from monitoring 11 milk parameters over a 4-month period during standard dairy farm production. Parameters include Temperature (°C), O_2 (ppm), CO_2 (mV), Redox potential (mV), pH, Conductivity (mS), Ca^{2+} (mV), NH_4^+ (mV), NO_3^- (mV), Cl^- (mV), and Milk Yield (L). These parameters have been selected from the literature due to their established interrelations and their potential to indicate changes in milk quality and safety. For instance, they can act as indicators of milk freshness, as the values of measured parameters can influence the stability of vitamins, the bioavailability of bioactive compounds, and consequently, the nutritional profile of the milk. Alterations in these parameters can also affect milk microbiota and indicate potential contamination, spoilage, or the presence of pathogens. Furthermore, they can be linked to (patho)physiological aspects in dairy cows, they are correlated with the technological characteristics of milk and their values can also indicate possible adulterations, such as dilution with water [10–19].

We assessed differences in milk samples identified by BEST as anomalous based on statistical analysis (i.e., outside "daily mean +/- standard deviation"). Our aim was to identify potential new biomarkers in the dairy chain for environmental health, food safety, and animal welfare. Statistical findings were compared with a machine learning approach for faster anomaly detection and explanation. Machine learning tools offer new perspectives in data processing for sensing systems and present novel opportunities in animal farming management [20].

2. Materials and method

The European semi-automated BEST platform (patent EP2304428B1) is a HACCP-like multi-sensor early warning system for continuous monitoring of biological, chemical, and physico-chemical parameters. Developed by Amel Srl (Milan, Italy) during the ALERT project funded by the Italian Ministry of Economic Development, BEST aids in identifying and managing anomalies in milk production, supporting self-monitoring diagnostics and milk chain traceability for dairy farming. It facilitates sampling and field analyses of individual cow's raw milk, detecting and monitoring variations in (bio)markers at critical control points (CCPs) and points of particular attention (POPAs). BEST a valuable tool for dairy farmers due to its user-friendly nature and easy integration into routine practices.

BEST's analytical strategy relies on the monitoring of chemical-physical parameters namely redox potential, pH, ionic conductivity and temperature, concentrations of main free ions (Ca^{2+}, NH_4^+, NO_3^-, and Cl^-), and concentrations of dissolved gases (O_2 and CO_2) in raw milk. Concentrations of main free ions were measured by using combination ion selective electrodes purchased by Sentek Ltd. (Braintree, Essex, UK), while redox potential, pH, electrical conductivity and temperature sensors were purchased by Amel Srl (Milan, Italy). Dissolved oxygen sensor and dissolved carbon dioxide sensors were purchased by Mettler Toledo (Milan, Italy). Measured values build a dataset and control charts, essential for generating multi-parameter milk footprints for preventive and corrective actions in good production practices [21–23]. BEST identifies anomalous signals by detecting anomalous trends or samples, aiding early anomaly detection in food production chains. Its flexibility allows integration of new probes to enhance anomaly identification in dairy production.

We conducted a statistical analysis on the database collected between March and June by the BEST prototype at "Elio Pascolini" dairy farm in Central Italy (Lazio region, coordinates 41°54′47.94"N, 12°15′48.25″E) during the ALERT project. Up to 15 milk samples per day were automatically collected from one of the 14 milking stations in the milking parlour. Throughout the sampling period, 850 milk samples were collected and 72 different dairy cows passed the sampling station at least once. The dairy farm, known for its high standards in dairy farming, agricultural, and sanitary practices, was previously detailed in a separate study [2].

This study focused on individual cow's raw milk, with sensor data recorded and collected during milking. An in-line automatic sampler, approved by the International Committee for Animal Recording [24], was installed at one of the farm's 14 milking units. The automatic sampler sampled milk from cows passing through the first gate, thus generating random daily milk samples. The milk sampled at the end of each individual cow's milking was collected and convoyed to the BEST by a specially designed automatic system. Statistical analysis of data was performed without knowing the history of the cow. Only after the identification of the 'anomalous' samples, milk samples were associated with the specific milking and therefore with the cow. This was possible because individual cows in the milking parlour weared automatic pedometers on their paws, and each cow was identifiable by a registration number (or ID number). The history of each cow (e.g. ongoing treatments, breed, etc.) was noted in the farm register.

As previously stated, each individual cow's raw milk sample underwent analysis for 11 parameters: Temperature (°C), O_2 (ppm), CO_2 (mV), Redox potential (mV), pH, ionic conductivity (mS/cm), and free inorganic ions (Ca^{2+}, NH_4^+, NO_3^-, Cl^-) here expressed as electrochemical potential (mV), along with Milk Yield (expressed in L, measured by the automated milking system provided by Nutriservice s.r.l., Brescia, Italy). Preliminary laboratory tests advised against daily calibration at the milk farm to save time and minimize chemical usage. These tests revealed that any sensor measurement drift would not impact daily results (sensor drift $\leq$1 %). The procedure for these tests involves verifying the correct response of each sensor's response using three calibration standard solutions with known analyte concentrations. The concentration ranges of these solutions bracket the expected analyte concentration in milk samples and are close to the midpoint of the recommended range for the specific sensor. The signals were plotted against the analyte concentration or, in the case of combination ion-selective electrodes, the logarithm of the ion concentration to establish the correct slope of the calibration curves. A two-point calibration of the dissolved oxygen sensor was conducted at 25 °C and at atmospheric pressure of 760 mmHg by measuring oxygen in open air (O_2 concentration 8.20 mg L^{-1}), and dissolved oxygen in a 10 g L^{-1} sodium sulfite solution (O_2 concentration 0.00 mg L^{-1}) (powder $\geq$98 %, purchased from Merck, Milan, Italy).

Sensor responses were verified using two laboratory-prepared solutions of known composition (three concentrations: 1000, 100, and 10 mg L^{-1}):

Solution a: 1000 mg $L^{-1}$$NH_4CO_3$ + 1000 mg $L^{-1}$$NaNO_3$ (serial dilution for 10 and 1 mg L^{-1}).

Solution b: 1000 mg L^{-1} mg L^{-1} $CaCl_2$ + 1000 mg $L^{-1}$$NaNO_3$ (serial dilution for 10 and 1 mg L^{-1}).

Using these solutions, we verified that the slopes of the sensors' calibration curves remained stable for several months. To streamline field operations, we opted to use daily means as a reference for sensor response. Daily means represent the average values of each parameter measured in milk samples collected on a given day. Each parameter's value in each milk sample is derived from the average of 10 consecutive measurements automatically conducted by BEST, with %RSD < 5 %.

2.1. Statistical approach

The Wilcoxon Signed Rank Test was conducted on all highlighted values of "anomalous samples" to assess their statistical significance ($P < 0.05$). Data analysis utilized Microsoft Excel and the R software package. Daily mean and standard deviation of the 11 parameters were calculated. Parameters showing trends that are more regular and less variability were pH, Ca^{2+}, NH_4^+, NO_3^-, and Cl^-. Statistical analysis focused on these five parameters to identify outlier samples (remaining parameters investigated subsequently). Values of the five parameters in all records were compared with their respective daily mean and standard deviations. Any parameter value outside the "daily mean +/- standard deviation" was highlighted. Records with at least three highlighted values were labeled as "anomalous samples" (AS). The IZSLT (Istituto Zooprofilattico Sperimentale del Lazio e della Toscana) performed additional analysis on cow's milk samples, specifically measuring the percentages of Fat, Proteins, and Lactose.

2.2. Neural network approach

Drifts in sensor calibration slopes may adversely affect statistical analysis. To address this, we propose a neural network (NN) method. NNs, recognized as universal approximators [25], can approximate both linear and nonlinear functions depending on their configuration. Popular in the '80s, NNs have seen renewed interest due to the success of deep learning. Their rapid growth in remote sensing stems from their ability to learn complex patterns [26]. NNs are robust in noisy environments, capable of generalizing even with incomplete or incorrect input data [27]. Additionally, they can flexibly combine different data types without assumptions about data set distributions [28].

In this study, we conducted two approaches: supervised and unsupervised. In the supervised approach, we employed the standard multilayer perceptron (MLP) neural network algorithm with sigmoid activation function and two hidden layers [29]. Training of the MLP network utilized a backpropagation approach with a gradient descent algorithm. While NNs are recognized as universal approximators [30], improper use may lead to undesired network overfitting or underfitting, depending on the training set. Overfitting occurs when the network is tailored too much to the learning examples, hindering satisfactory performance on new patterns. Addressing two main issues in supervised NN design is crucial: 1) when to stop the training algorithm, and 2) how many neurons have to be included in the hidden layers of the topology.

For the first issue, we utilized the early stopping algorithm [29]. This algorithm involves both training and test sets, with the test set containing examples not in the training set. Network parameters are adjusted iteratively to minimize error on the training set while simultaneously evaluating network performance on the test set. Training stops when the error on the test set reaches its minimum. Failure to employ this procedure may lead to network overtraining, reducing generalization capability despite smaller training dataset errors. For the second issue, a grid search method aimed to minimize mean-square error (MSE).

In supervised NN training, selecting a representative training dataset is crucial. However, obtaining an extensive dataset to fully train a NN is often impractical, risking ineffective representation of all possible dynamics. An alternative for limited training data is unsupervised classification, commonly known as clustering. Among various clustering techniques, Self-Organizing Maps [31] are considered highly effective. SOMs map multidimensional data onto lower-dimensional subspaces, typically two-dimensional, where geometric relationships between points indicate similarity. They define an ordered mapping onto a regular, two-dimensional grid, with each node associated with a model. Data items are mapped to the node whose model is most similar, i.e., closest in some metric. Models at nearby nodes are more similar than those at distant nodes, forming a similarity graph and structured 'skeleton' of data distribution. The neural network is represented as a grid of neurons, each containing a weight vector and a geometric location. These weight vectors determine the "winning" neuron for each input and are updated based on location during training. Unlike traditional NNs, SOMs are trained using competitive learning rather than error correction learning like gradient descent. SOM training involves randomly initializing weights, iterating over input data, finding the "winning" neuron for each input, and adjusting weights based on its location. The algorithm used to find the "winning" neuron is the Euclidean distance, which minimizes the equation:

$$\sqrt{\sum_{i=0}^{d} (n_i - v_i)^2}$$

where n represents a neuron of dimension d for sample input v. The adjusting of the weights is then performed considering not only the winning neuron, but also its neighbors. Generally, the neighborhood function is designed to have a global maximum at the "winning" neuron and decrease as it gets further away from it. This makes it so that neurons close to the "winning" neuron get scaled towards the sample input the most while neurons far away get scaled the least which creates groupings of similar neurons in the final map (Fig. 1).

3. Results

3.1. Statistical approach

The Wilcoxon Signed Rank Test confirmed statistical significance for all highlighted anomalous samples ($P < 0.05$). Results revealed only three cows producing more than two anomalous milk samples (i.e., records with at least 3 highlighted values). Comparing all anomalous samples in the ALERT database (i.e., those with at least 3 highlighted values) with the farm's paper register provided specific information on the cows. This comparison revealed that these anomalous milk samples came from cows with distinct characteristics from others in the herd. Cow ID number 432, producing four anomalous samples (Table 1), was identified as a Jersey breed, unique in the ALERT database, while other cows were Holstein-Friesian breed. Jersey cow milk consistently showed significantly higher Ca^{2+} and NO_3^- values than the daily mean of all cows sampled that day, with NH_4^+ values consistently lower. Cl^- values tended to decrease over time.

Examining other parameters (Table 2), redox potential values consistently exceed the daily mean, while electrical conductivity values consistently fall below it, although not always significantly. Additionally, compositional analyses by the IZSLT reveal that milk samples from the Jersey cow exhibit significantly higher fat levels compared to those from Holstein-Friesian cows.

Three ASs represent the milk produced by the cow with ID number 577 (Table 3). These ASs were measured on three consecutive days; this cow had given birth 2 weeks earlier (hereinafter referred as "postparturient cow"). According to the milk farm's paper register, this cow was the only postparturient cow for the period considered. It is interesting to note that in milk samples from this cow, the Ca^{2+} and NO_3^- mV values recorded are always significantly higher and the level of Cl^- ion gradually decreases over time, until they become significantly lower

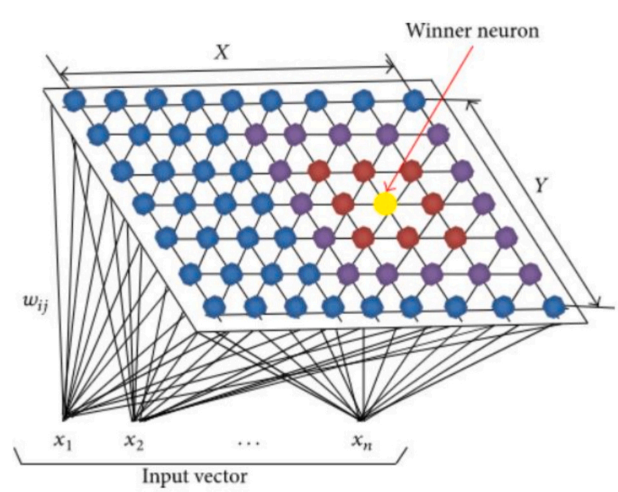

Fig. 1. Structure of the SOM mapping.

Table 1
Parameters recorded in Jersey cow's milk.

Data	Registration number	T (°C)	pH	Ca^{2+} (mV)	NH_4^+ (mV)	NO_3^- (mV)	Cl^- (mV)	
May 4, 2017	432	26.97	6.32	419.46	249.69	303.88	116.43	
			6.62	413.92	252.44	280.99	109.21	Daily mean
			0.12	2.25	2.52	10.3	4.95	SD
May 10, 2017	432	27.81	6.52	413.95	250.22	289.48	114.66	
			6.58	411.68	253.37	275.81	111.14	Daily mean
			0.06	2.13	1.73	6.28	3.36	SD
May 24, 2017	432	24.55	6.5	412.11	249.45	254.73	99.48	
			6.52	409.67	253.32	244.73	98.18	Daily mean
			0.05	1.92	3.01	5.55	3.88	SD
May 30, 2017	432	28.12	6.55	413.59	251.39	249.32	93.47	
			6,51	409,55	251.56	242.82	97.97	Daily mean
			0.04	2.07	0.96	4.27	3.87	SD

Table 2
Other interesting parameters recorded in the Jersey cow's milk.

Data	ID number	T (°C)	ORP (mV)	Cond (mS)	Fat Bov (%)	
May 4, 2017	432	26.97	73.05	5.86	/	
			24.75	6.18	/	Daily mean
			16.71	0.31	/	SD
May 10, 2017	432	27.81	34.27	5.9	4.84	
			20.95	5.95	4.13	Daily mean
			6.69	0.29	0.44	SD
May 24, 2017	432	24.55	27.95	5.96	4.86	
			18.95	6.36	3.70	Daily mean
			5.29	0.29	0.80	SD
May 30, 2017	432	28.12	54.41	6.17		
			30.65	6.48		Daily mean
			11.64	0.26		SD

than the daily mean; reverse trend occurs for the level of NH_4^+ ion. In addition, the pH values always significantly lower than the respective daily mean, obtained from the milk samples of all the cows measured that day.

The laboratory analyses conducted by the IZSLT showed that this cow's milk sample was significantly richer in %protein and significantly less rich in %lactose respect to daily means (Table 4).

A second cow, with serial number 500, produced three AS. The milk farm's paper register did not mention any particular condition for this cow.

Cow serial number 680 (Table 5) produced only one AS. Upon review of the farm's paper register, we discovered that this cow had previously been treated with the drug Micospectone® (Lincomycin + Spectinomycin). After completing the drug treatment and the subsequent withdrawal period, the cow was reintegrated into the herd. The anomalous sample was collected on 24 March, 4 days after the cow's reintroduction. This sample exhibited significantly higher levels of Ca^{2+}, NH_4^+, and NO_3^- compared to the respective daily mean. Two other milk samples from this cow were found in the database, collected 7 and 8 days after reintegration. The sample taken after 7 days had outliers for two parameters (Ca^{2+} and NH_4^+), while the sample taken after 8 days had none.

According to the milk farm's paper register, cow ID number 680 was the only one in our database that had received drug treatment.

3.1.1. Supervised neural network approach

A supervised classification approach based on MLP was tested. In particular, it was used a NN architecture with two hidden layers and sigmoidal activation function in each node:

$$\sigma(x) = \frac{1}{1 + e^{-x}}$$

where x is the weighted sum of the output from the nodes of the previous layer:

$$x = \sum_i w_i g_i$$

The NN is trained by the Adam algorithm in order to minimize the sum-of-squares error of the form:

Table 4
Other interesting parameters of the postparturient cow.

Data	ID number	T (°C)	% Protein	%Lactose	
May 16, 2017	577	27.63	–	–	
			–	–	Daily mean
			–	–	SD
May 17, 2017	577	30.75	3.94	4.46	
			3.47	4.81	Daily mean
			0.34	0.19	SD
May 18, 2017	577	24.63	4.00	4.31	
			3.32	4.85	Daily mean
			0.25	0.21	SD

Table 3
Postparturient dairy cow.

Data	ID number	T (°C)	pH	Ca^{2+} (mV)	NH_4^+ (mV)	NO_3^- (mV)	Cl^- (mV)	
May 16, 2017	577	27.63	6.24	412.44	250.50	279.38	104.06	
			6.55	407.67	251.36	260.06	99.13	Daily mean
			0.11	2.91	1.47	7.48	7.62	SD
May 17, 2017	577	30.75	6,29	415.07	259.97	265.45	103.44	
			6.56	405.80	254.7	258.40	104.32	Daily mean
			0.12	3.83	3.00	6.75	5.49	SD
May 18, 2017	577	24.63	6.35	412.19	258.19	265.02	92.13	
			6.50	405.48	253.97	255.43	102.64	Daily mean
			0.07	3.08	3.33	4.59	5.36	SD

Table 5
Cow treated with Micospectone®.

Data	ID number	°C	pH	Ca^{2+} (mV)	NH_4^+ (mV)	NO_3^- (mV)	Cl^- (mV)	
March 24, 2017	680	24.30	6.67	458.18	303.31	354.30	122.95	
			6.68	452.51	294.39	349.98	122.34	Daily mean
			0.08	2.92	4.71	3.07	4.18	SD
March 27, 2017	680	26.68	6.67	436.73	279.18	358.09	130.67	
			6.64	434.93	274.28	356.96	127.78	Daily mean
			0.06	1.74	4.06	2.29	9.23	SD
March 28, 2017	680	25.85	6.64	432.87	270.97	357.31	135.23	
			6.65	431.73	267.92	357.47	135.32	Daily mean
			0.04	2.49	3.77	2.60	5.08	SD

No other cows present in the milk farm's paper register could be associated to any other different characteristics respect to the herd.

$$E = \frac{1}{2}\sum_{n=1}^{N}\sum_{k=1}^{d}\left[y_k(x^n) - x_k^n\right]^2$$

where y_k (k = 1, 2, d) is the output vector. The Adam algorithm is a first-order gradient-based optimization of stochastic objective functions, based on adaptive estimates of lower-order moments [32].

The input dataset was composed of 9 parameters out of the initial 11, by excluding “Temperature” and “Milk yield”. In fact, the measurement procedures did not allow the measurement of milk temperature immediately after milking, and measured temperature was more affected by environmental temperature than by the cows' body.

Milk yield was not considered in the analysis because not always measured.

Four main classes were associated to the anomalies identified during the analysis carried out in the statistical approach, namely “Post-parturient”, “Post antibiotic-treatment”, “Jersey Breed”, and “Normal”.

Since the original data was affected by daily drift, day-by-day data standardization was performed, characterized by zero-mean and unitary standard deviation.

We used a grid search method to determine the optimal number of hidden layers and neurons, minimizing the mean-square error (MSE). Then, an early stopping approach was carried out, to avoid data overfitting. This has been obtained by dividing the whole measurements into three datasets, namely “training”, “test” and validation datasets. The training, test and validation datasets were composed respectively of about 10 %, 5 % and 85 % of the whole measurements. The optimal topology was composed by 9 input nodes, two hidden layers with 20 nodes each, and an output layer with 4 nodes.

3.1.2. Unsupervised neural network approach

As for the unsupervised classification of the data (clustering) based on SOMs, a NN architecture was used to map the original data into a square grid. In addition, in this case, “Temperature” and “Milk yield” were not included. A day-by-day data standardization was performed to mitigate the daily drift in the input data. Several training runs, 200 epochs each, were performed in order to find the optimal network topology by varying the size of the square grid.

The optimization of the SOM algorithms involved several key steps to ensure accurate and robust results. The optimal grid size was determined through iterative training runs, with the final configuration being a 12 × 12 grid. The SOM training process utilized the Euclidean distance metric for updating the neurons' weight vectors. The neighborhood function used in these experiments is a Gaussian function, which is commonly used due to its smooth and continuous nature. This function was selected to ensure that neurons closer to the winning neuron (best matching unit) are updated more significantly than those further away, creating a smoother mapping of the data. As unsupervised approach, the labeling of the results was carried out by analyzing the neighbor of each neuron corresponding to “Normal” cows and those affected by the three types of anomalies identified during the analysis carried out in the statistical approach, namely namely “Postparturient”, “Post antibiotic-treatment”, “Jersey Breed”, and “Normal”.

From a first visual analysis of the grid map in Fig. 2, it is possible to identify two extended yellow areas on the top left and lower right parts of the grid. Because of their extensions, these two areas can be associated to the Normal class. Similarly, the neurons that have darker connections (larger distances) represent data less correlated with the neighbors, which can be associated with the anomalies Classes. This distance is based on the Euclidean metric, not the correlation coefficient.

Moreover, it is possible to analyze the neurons associated with the anomalies in order to understand the accurateness of this proposed approach.

Figure 3 depicts the nodes associated with the measurements related to the classes “Jersey breed” (ID number 432, circled in blue), “Post-parturient” (ID number 577, circled in cyan), and “Post antibiotic-treatment” i.e., cow treated with Micospectone® ID number 680, circled in green).

As it can be noted, all the measurements present weak connections with the neighbors. Moreover, most of these measurements, with the exception of the measurement of May 24, 2017, present fewer connections with the neighbors being on the border, confirming the anomaly behavior. Regarding the neurons associated with the milk measurements from the postparturient cow (ID number 577), all the measurements lie on the border and present very weak connection with the neighbors' neurons. It is interesting to note that since the three measurements are

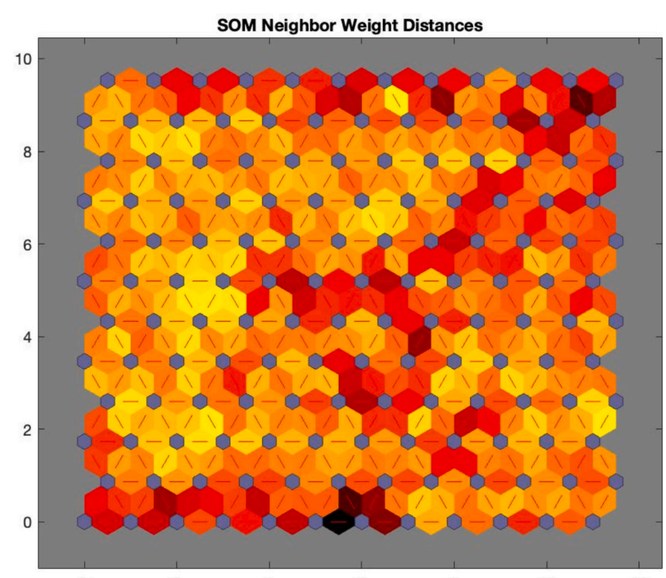

Fig. 2. SOM neighbor weight distance map. The red lines represent the relations between a neuron and each neighbor. The neighbor patches are colored from black (far) to yellow (close) to show the Euclidean distance between each neuron's weight vector and its neighbors, indicating similarity. Darker colors represent larger distances (less similarity), while lighter colors indicate smaller distances (more similarity). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

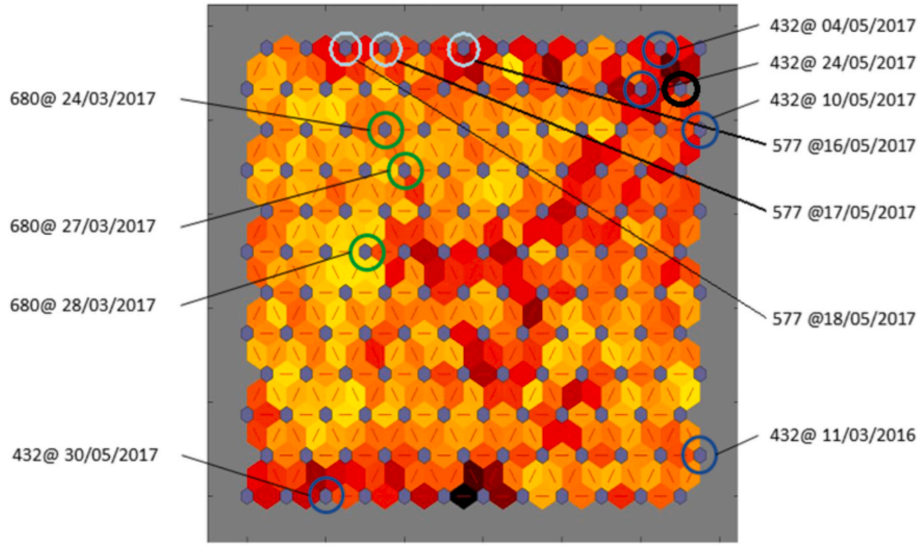

Fig. 3. Positions on the grid of the neurons associated with the classes "Jersey breed" (Blue), "Postparturient" (Cyan), and "Post antibiotic-treatment" (Green). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

very close (positions 135, 136 and 138) it is possible to suppose that the neuron at position 137 (Fig. 3, circled in black) could be associated to a cow presenting similar conditions to the cow ID number 577. Regarding the neurons associated with the cow treated with Micospectone® (ID number 680), it is interesting to note that the connections of these three neurons tends to become stronger as the time pass from the last dose of Micospectone®. This means that while the March 24, 2017 measurement presents slightly weak connections with the neighbors, the other two present a higher level of correlation with the surrounding neurons, thus allowing us to suppose that on March 28, 2017 the effect of the Micospectone® is almost completely disappeared.

4. Discussion

As expected, the presence of three cows with different characteristics compared to the herd reflect the good condition of the dairy farm where the BEST was applied.

In particular, this research spotted the only Jersey cow (ID number 432), the only postparturient cow (ID number 577) and the only cow (ID number 680) previously treated with Micospectone® in the whole herd's database. As far as we know, no other cows present in the milk farm's paper register could be associated to any other different condition respect to the other cows of the herd.

The physicochemical properties of raw milk depend on numerous factors, including genetic factors such as breed [33]. Previous studies showed that genetic factors determine differences in the chemical composition and physicochemical properties of milk among cow breeds [34,35]. In milk, ionized Ca in the soluble phase accounts for about 10 % of the total Ca [36–40]. Our findings of higher levels of Ca^{2+} and milk fat in Jersey cow's milk compared to Friesian cow's milk align with literature data [34,41–43]. Compared to other cow's milk, Jersey cow milk is higher in milk fat and calcium levels [43,44]. In particular, the concentrations of all forms of Ca is higher in Jersey cows milk than in Holstein-Friesian cows [42,43].

In our research the postparturient cow (ID number 577), 2 weeks after calving, was in the postcolostrum period (day 6 to 30 of lactation) [45]. Milk samples from that postparturient cow always had an amount of Ca^{2+} significantly higher than the daily mean. A higher ionic calcium value has been previously reported in milk on the first day after birth, which steadily decreases over subsequent days [45].

It was also measured a lower pH in the milk of the postparturient cow, increases over time. The pH of bovine milk (nonmastitic and noncolostrum) commonly ranges 6.4–6.8 [47–49], but it can decrease in the first 5 days after calving and then steadily increased [45].

Our results also indicate a gradual decrease in Cl^- levels over time in the milk of the postparturient cow. Levels of Cl^- in milk may vary with lactation stage, decreasing from colostrum to mature milk but increases sharply towards the end of lactation [19,46,50].

Our research has found that postparturient cow's milk is richer in protein (Table 4). The high protein content observed during early postpartum was not unusual and has been reported before [45,51,52]. Lower percentage of lactose in milk from postparturient cow (Table 4) is also confirmed by literature data [45].

Results from cow serial number 500 (i.e., another cow that produced three AS) suggest an alteration in physiological state of that cow. But no additional information are available on the farm's paper register or reported by the breeder for that cow in the period considered. The cow treated with Micospectone® was identified because its milk had Ca^{2+}, NH_4^+ and NO_3^- significantly higher compared to the daily averages, until the seventh day post-reintegration into the herd. Outlier parameters gradually normalized, with all returning to normal values by the eighth day. Our research suggests that pregnancy, drugs and breed or physiological factors could also influence the amount of NO_3^- and NH_4^+ in raw milk; further studies will investigate this aspect.

The results from this study motivate us to integrate the BEST platform with the MOLOKO multi-parameter optical sensor [53] to expand the range of monitored parameters. Specifically, the MOLOKO sensor will provide quick, semi-quantitative, on-site, and automated analysis of contaminants and indicators of milk quality and safety (e.g., antibiotics and lactoferrin). The use of the integrated Moloko-'Best' system in the farm scenario can promote the application of new tools for on-farm milk analyses.

5. Conclusions

The Neural Network approach confirms the results of the statistical approach, with the advantage of being a more rapid and predictive. The BEST platform has proven to be a valuable tool for identifying cows with different characteristics in the herd, thus opening the way towards early HACCP and traceability in the dairy supply chain. Future research and technological improvement requires increasing participative and tight collaboration between milk producers and scientific researchers.

CRediT authorship contribution statement

Roberto Dragone: Writing – review & editing, Supervision, Funding acquisition, Conceptualization. **Gerardo Grasso:** Writing – review & editing, Writing – original draft, Methodology, Investigation. **Giorgio Licciardi:** Writing – review & editing, Writing – original draft, Visualization, Formal analysis. **Daniele Di Stefano:** Writing – original draft, Visualization, Formal analysis. **Chiara Frazzoli:** Writing – review & editing, Supervision, Resources, Project administration, Funding acquisition, 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.

Data availability

Data will be made available on request.

Acknowledgements

This work was developed within the activities of the Consortia :i) ALERT project (grant MI 00195 - Italian Ministry for Economic Development);ii) MOLOKO project (grant agreement no. 780839 - European Union's Horizon 2020)

References

[1] M.J. Vilar, J.L. Rodriguez-Otero, M.L. Sanjuan, F.J. Diéguez, M. Varela, E. Yus, Implementation of HACCP to control the influence of milking equipment and cooling tank on the milk quality, Trends Food Sci. Technol. 23 (2012) 2–12, https://doi.org/10.1016/j.tifs.2011.08.002.
[2] A. Lombardo, C. Boselli, S. Amatiste, S. Ninci, C. Frazzoli, R. Dragone, A. De Rossi, G. Grasso, A. Mantovani, G. Brajon, From invention to innovation: risk analysis to integrate one health Technology in the Dairy Farm, Front. Public Health 5 (2017) 302, https://doi.org/10.3389/fpubh.2017.00302.
[3] C. Frazzoli, A. Mantovani, Toxicants exposures as novel zoonoses: reflections on sustainable development, food safety and veterinary public health, Zoonoses Public Health 57 (2010) 136–142, https://doi.org/10.1111/j.1863-2378.2009.01309.x.
[4] European Commission, Directorate-General for Health and Food Safety, Welfare of Cattle on Dairy Farms – Overview Report, Publications Office, 2017, https://doi.org/10.2875/815860 (accessed 05 December 2023).
[5] C. Frazzoli, B. Bocca, A. Mantovani, The one health perspective in trace elements biomonitoring, J. Toxicol. Environ. Health Part B 18 (2015) 344–370, https://doi.org/10.1080/10937404.2015.1085473.
[6] C. Frazzoli, A. Mantovani, R. Dragone, Local role of food producers' communities for a global one-health framework: the experience of translational research in an Italian dairy chain, JACEN 3 (2014) (2014) 14–19, https://doi.org/10.4236/JACEN.2014.32B00.
[7] G. Grasso, D. Zane, R., Dragone field and remote sensors for environmental health and food safety diagnostics: an open challenge, Biosensors 12 (5) (2022) 285, https://doi.org/10.3390/bios12050285.
[8] N. Hostiou, J. Fagon, S. Chauvat, A. Turlot, F. Kling-Eveillard, X. Boivin, C. Allain, Impact of precision livestock farming on work and human-animal interactions on dairy farms. A review, Biotechnol. Agron. Soc. Environ. 21 (2017) 268–275, https://doi.org/10.25518/1780-4507.13706.
[9] T. Van Hertem, L. Rooijakkers, D. Berckmans, A.P. Fernández, T. Norton, E. Vranken, Appropriate data visualisation is key to precision livestock farming acceptance, Comput. Electron. Agric. 138 (2017) 1–10, https://doi.org/10.1016/j.compag.2017.04.003.
[10] L.W.T. Fweja, M.J. Lewis, A.S. Grandison, The potential of conductivity, redox potential and dissolved oxygen in raw milk quality prediction, Huria, J. Open Univ. Tanzan. 15 (1) (2013) 52–70, https://doi.org/10.4314/HURIA.V15I1.
[11] S. Abraham, R. Cachon, S. Jeanson, B. Ebel, D. Michelon, C. Aubert, C. Rojas, G. Feron, E. Beuvier, P. Gervais, A procedure for reproducible measurement of redox potential (eh) in dairy processes, Dairy Sci. Technol. 93 (2013) 675–690, https://doi.org/10.1007/s13594-013-0134-5.
[12] S. Haratifar, L. Bazinet, N. Manoury, M. Britten, P. Angers, Impact of redox potential electrochemical modification and storage conditions on the oxidation reaction prevention in dairy emulsion, dairy, Sci. Technol. 91 (2011) 541–554, https://doi.org/10.1007/s13594-011-0025-6.
[13] P.F. Fox, T. Uniacke-Lowe, P.L.H. McSweeney, J.A. O'Mahony, Physical properties of milk, in: Dairy Chemistry and Biochemistry, Springer, Cham, 2015, pp. 321–343, https://doi.org/10.1007/978-3-319-14892-2_8.
[14] T. Aydogdu, J.A. O'Mahony, N.A. McCarthy, pH, the fundamentals for milk and dairy processing: a review, Dairy 4 (2023) 395–409, https://doi.org/10.3390/dairy4030026.
[15] M.J. Lewis, The measurement and significance of ionic calcium in milk–a review, Int. J. Dairy Technol. 64 (2011) 1–13, https://doi.org/10.1111/j.1471-0307.2010.00639.x.
[16] F. Gaucheron, Y. Le Graet, Determination of ammonium in milk and dairy products by ion chromatography, J. Chromatogr. A 893 (2000) 133–142, https://doi.org/10.1016/S0021-9673(00)00695-6.
[17] L.W. Gapper, B.Y. Fong, D.E. Otter, H.E. Indyk, D.C. Woollard, Determination of nitrite and nitrate in dairy products by ion exchange LC with spectrophotometric detection, Int. Dairy J. 14 (2004) 881–887, https://doi.org/10.1016/j.idairyj.2004.02.015.
[18] A. Jóźwik, N. Strzałkowska, E. Bagnicka, W. Grzybek, J. Krzyżewski, E. Poławska, A. Kołataj, J.O. Horbańczuk, Relationship between milk yield, stage of lactation, and some blood serum metabolic parameters of dairy cows, Czeh J. Anim. Sci. 57 (8) (2012) 353–360, https://doi.org/10.17221/6270-CJAS.
[19] P.F. Fox, T. Uniacke-Lowe, P.L.H. McSweeney, J.A. O'Mahony, Salts of Milk, in: Dairy Chemistry and Biochemistry, Springer, Cham, 2015, pp. 241–270, https://doi.org/10.1007/978-3-319-14892-2_5.
[20] S. Neethirajan, The role of sensors, big data and machine learning in modern animal farming, Sens. Bio-Sens. Res. 29 (2020) 100367, https://doi.org/10.1016/j.sbsr.2020.100367.
[21] A. De Vries, J.K. Reneau, Application of statistical process control charts to monitor changes in animal production systems, J. Anim. Sci. 88 (13 Suppl) (2010) E11–E24, https://doi.org/10.2527/jas.2009-2622.
[22] K. Mertens, E. Decuypere, J. De Baerdemaeker, B. De Ketelaere, Statistical control charts as a support tool for the management of livestock production, J. Agric. Sci. 149 (2011) 369–384, https://doi.org/10.1017/S0021859610001164.
[23] F. Martelli, C. Giacomozzi, A. Fadda, C. Frazzoli, Understanding seasonal changes to improve good practices in livestock management, Front. Public Health 6 (2018), https://doi.org/10.3389/fpubh.2018.00175.
[24] ICAR (International Committee for Animal Recording) ICAR Guidelines, Section 11: Guidelines for Testing, Approval and Checking of Milk Recording Devices. https://www.icar.org, 2012 (accessed 05 December 2023).
[25] G. Cybenko, Approximation by superpositions of a sigmoidal function math, Control Sign. Syst. 2 (1989) 303–314, https://doi.org/10.1007/BF02551274.
[26] T.P. Dawson, P.J. Urran, S.E. Plummer, The biochemical decomposition of slash pine needles from reflectance spectra using neural networks, Int. J. Remote Sens. 19 (1998) 1433–1438, https://doi.org/10.1080/014311698215540.
[27] F. Del Frate, G. Schiavon, A combined natural orthogonal functions—neural network technique for the radiometric estimation of atmospheric profiles, Radio Sci. 332 (1998) 405–410, https://doi.org/10.1029/97RS02219.
[28] A. Benediktsson, J.R. Sveinsson, Feature extraction for multi-source data classification with artificial neural networks, Int. J. Remote Sens. 18 (1997) 727–740, https://doi.org/10.1080/014311697218728.
[29] C.M. Bishop, Neural networks and their applications, Rev. Sci. Instrum. 65 (1994) 1803–1832, https://doi.org/10.1063/1.1144830.
[30] K. Hornik, M. Stinchcombe, H. White, Multilayer feedforward networks are universal approximators, Neural Netw. 2 (1989) 359–366, https://doi.org/10.1016/0893-6080(89)90020-8.
[31] T. Kohonen, Self-organizing formation of topologically correct feature maps, Biol. Cybern. 43 (1982) 59–69, https://doi.org/10.1007/BF00337288.
[32] D.P. Kingma, J. Ba, Adam: a method for stochastic optimization, arXiv (2015) preprint arXiv:1412.6980, https://arxiv.org/pdf/1412.6980.
[33] A. Singh, A. Pratap, Comparison of physicochemical properties of raw milk from indigenous and exotic cows at Allahabad, IJSR 3 (2014) 2319–7064, https://doi.org/10.56093/ijans.v92i7.133145.
[34] M. Boland, Influences on raw milk quality, in: G. Smith (Ed.), Dairy Processing, Improving Quality C.H.I.P.S, Weimar, USA, 2003, pp. 42–67.
[35] S.C. Nickerson, Milk production: Factors affecting milk composition, in: F. Hardling (Ed.), Milk Quality, An Aspen Publ, Gaithersburg, Maryland, 1999, pp. 3–24.
[36] M.C. Neville, C.D. Watters, Secretion of calcium into milk, J. Dairy Sci. 66 (1983) 371–380, https://doi.org/10.3168/jds.S0022-0302(83)81802-5.
[37] J.N. VanHouten, J.J. Wysolmerski, Transcellular calcium transport in mammary epithelial cells, J. Mammary Gland Biol. Neoplasia 12 (2007) 223–235, https://doi.org/10.1007/s10911-007-9057-1.
[38] C. Holt, An equilibrium thermodynamic model of the sequestration of calcium phosphate by casein micelles and its application to the calculation of the partition of salts in milk, Eur. Biophys. J. 33 (2004) 421–434, https://doi.org/10.1007/s00249-003-0377-9.
[39] R. Jenness, The composition of milk, in: B.L. Larsson, V.R. Smith (Eds.), Lactation: A Comprehensive Treatise, Academic Press, New York, 1974, pp. 3–108.
[40] B. Lonnerdal, C. Glazier, Calcium binding by a-lactalbumin in human milk and bovine milk, J. Nurr. 115 (1985) 1209–1226, https://doi.org/10.1093/jn/115.9.1209.
[41] S.C. Nickerson, Milk production: Factors affecting milk composition, in: F. Hardling (Ed.), Milk Quality, An Aspen Publ. Gaithersburg, Maryland, 1999, pp. 3–24.
[42] M. Czerniewicz, K. Kieczewska, A. Kruk, Comparison of some physicochemical properties of milk from Holstein Friesian and Jersey cows, Pol. J. Food Nutr. Sci. 15 (2006) 61–64.
[43] H. Frckowiak, Cattle from the Jersey Island – history and significance, Medycyna Wet. 60 (2004) 666–667 (in Polish).
[44] I. Antkowiak, J. Pytlewski, D. Stanisawski, Effect of selected factors of Jersey cows breeding on their performance and milk composition, Zesz. Nauk. Prz. Hod. 72 (2004) 101–111 (in Polish).

[45] A. Tsioulpas, A.S. Grandison, M.J. Lewis, Changes in physical properties of bovine milk from the colostrum period to early lactation, J. Dairy Sci. 90 (2007) 5012–5017, https://doi.org/10.3168/jds.2007-0192.
[46] A. Flynn, Minerals and trace elements in milk, Adv. Food Nutr. Res. 36 (1992) 209–252, https://doi.org/10.1016/s1043-4526(08)60106-0.
[47] M.A. Elbagermi, A.I. Alajtal, H.G.M. Edwards, A comparative study on the physicochemical parameters and trace elements in raw Milk samples collected from Misurata- Libya, SOP Transact. Anal. Chem. 1 (2014), https://doi.org/10.15764/Ache.2014.02002.
[48] R. Kanwal, T. Ahmed, B. Mirza, Comparative analysis of quality of milk collected from buffalo, cow, goat and sheep of Rawalpindi/Islamabad region in Pakistan, Asian J. Plant Sci. 3 (2004) 300–305, https://doi.org/10.3923/ajps.2004.300.305.
[49] A. Enb, M. Abou Donia, N. Abd-Rabou, A. Abou-Arab, M. El-Senaity, Chemical composition of raw milk and heavy metals behavior during processing of milk products, Glob. Vet. 3 (2009) 268–275.
[50] M. Kirchgessner, H. Friesecke, G. Koch, Nutrition and the Composition of Milk. Crosby Lockwood, London, 1967.
[51] R.C. Foley, D.D. Bath, E.M. Diokinson, H.A. Tueker, Dairy Cattle: Principles Practices, Problems, Profits, Lea and Febiger, Philadelphia, USA, 1972 pp. xvi+-574.
[52] S. Sodhi, S.P. Ahuja, S. Singh, Changes in the Levels of Ionic Calcium and Proteins During Assembly of Casein Micelles and Transition of Colostrum to Milk in the Buffalo and Cow, 1996, pp. 101–108.
[53] M. Prosa, M. Bolognesi, L. Fornasari, G. Grasso, L. Lopez-Sanchez, F. Marabelli, S. Toffanin, Nanostructured organic/hybrid materials and components in miniaturized optical and chemical sensors, Nanomater 10 (2020) 480, https://doi.org/10.3390/nano10030480.