APPLICATION OF SUPPORT VECTOR MACHINES AS SUPPORT TO EARLY PREDICTION OF MASTITIS IN SARDA DAIRY EWES

PhD Tutor:
Chia.mo Prof. Antonio Pazzona

PhD Student:
Djangsou Hagassou

Academic Year 2014/2015
# Table of contents

1. Introduction...................................................................................................................... 4

2. Materials and methods .................................................................................................... 10
   2.1. Survey design and flock management...................................................................... 10
   2.2. Milk sampling .......................................................................................................... 10
   2.3. Measurement and laboratory analysis ..................................................................... 11

2.4. Statistical analysis ........................................................................................................ 12
   2.4.1. Statistical instruments and data analysis software ............................................. 12
   2.4.2. The Receiver Operating Characteristic curves (ROC curves)........................... 12
   2.4.3. Support Vector Machines (SVMs):..................................................................... 16
   2.4.4. MSCC threshold for normal ewes milk............................................................... 23

3. Results and discussion ..................................................................................................... 25
   3.1. Milk variables description ......................................................................................... 25
      3.1.1. Milk somatic cell count (MSCC)..................................................................... 26
      3.1.2. Milk electrical conductivity (MEC)................................................................... 29
      3.1.3. Milk Chloride content (MCC) and milk pH (MpH).......................................... 31
      3.1.4. Milk lactose content (MLC)........................................................................... 34
   3.2. Masti-Milk efficiency................................................................................................. 36
   3.3. Relationship between milk first squirts and milk from complete milking.............. 39
   3.4. Relationships between most relevant milk parameter linked with MSCC................. 40
      3.4.1. Relationship between milk electrical conductivity (MEC) and milk somatic cell count (MSCC)..................................................................................................................... 42
      3.4.2. Relationship between Milk Chlorides Content (MCC) and milk somatic cell count (MSCC)......................................................................................................................... 43
      3.4.3. Relationship between milk lactose content (MLC) and milk somatic cell count (MSCC)................................................................................................................................. 44
      3.4.4. Relationship between MEC and milk lactose content (MLC)............................ 45
      3.4.5. Relationship between Milk Chlorides Content (MCC) and Milk Chloride content (MCC)............................................................................................................................. 46
   3.5. ROC curve analysis results ........................................................................................ 48
      3.5.1. ROC curve analysis at MSCC decision threshold of 500*10^3 cells/ml............ 49
      3.5.2. ROC curve analysis at MSCC decision threshold of 750*10^3 cells/ml............ 55
1. Introduction

*Mastitis* is any udder inflammation regardless of its origin, severity or evolution (Bergonier et al., 2003; Olechnowicz et al., 2014). The term *mastitis* comes from the Greek words, *Mastos* = udder/breast and *itis* = inflammation (Viguier et al., 2009; Olechnowicz, 2014). Most often, mastitis is caused by microorganisms, mostly bacterial origin (Albenzio et al., 2002; Bergonier, 2003; Mammadova and Keskin, 2013). Mastitis remains the most frequent and costly disease in the dairy industry, in term of economic losses and animal welfare (Kiossis et al., 2007; Kramer et al., 2009; Miekley et al., 2013). Mastitis is an important issue for three main reasons:

- economic: due to ewe and lamb mortality, treatment costs, reduced milk production, reduced lamb growth, milk payment on cellular quality in certain areas;
- hygienic: causing risk of infection or intoxication of costumers by milk bacteria such as *Escherichia coli*, *Staphylococcus aureus*, *Listeria monocytogenes*, *Salmonella spp.* Etc.;

The Sardinian rural development program (SRDP) has also fixed the threshold of Milk Somatic Cell Count (MSCC) at 1,500*10^3 cells/ml for all farmers willing to participate to the Measure 215 (RDP 2007-2013) and Measure 14 (RDP 2014-2020). The goal was to reduce the MSCC at 1,000*10^3 cells/ml at the end of the program. The MSCC is widely considered as a valuable indicator of udder health status and indirect parameter for sanitary control and management practices (Paape et al., 2001; Vivar-Quintana et al., 2006; Molina et al., 2010).
The ovine livestock in Sardinia consists of 2,400,000 ewes, which is of great cultural and economic importance for the island (SRDP 2007-2013). In fact the production of sheep's milk is one of the main economic activities in Sardinia.

The prevalence of clinical mastitis in ewes is generally below 5%, whereas the prevalence of subclinical mastitis is widely variable and ranges from less than 10 to 50% or more (Bergonier et al., 2003; Contreras et al., 2007; Kiossis et al., 2007; Olechnowicz, 2014).

In Sardinia the ewe milk is mainly destined for high quality pure sheep cheeses making (Pirisi et al., 2007; Molina et al., 2010). Many of pure sheep cheeses produced in Sardinia have the Protected Denomination of Origin (PDO) such as Pecorino Romano, Pecorino Sardo and il Fiore Sardo ect. Those high quality dairy products required a certain quality standards, Which must be guaranteed through the high milk quality. Therefore, quality control systems for milk payment and Good Dairy Farm Practices (GDFP) are promoted in the Sardinian ewe dairy farming system (SRDP 2007-2013, measure 215; SRDP 2014-2020, measure 14). Cuccuru et al. (2011) have reported an annual economic of 69 NZ$ (≈ 42 €) per udder-half over the entire lactation period.

Mastitis induces physiological, physical, chemical changes in milk that influence dairy product quality (Vivar-Quintana et al., 2006). In case of udder infection, the MSCC increases dramatically, resulting from a transfer of white blood cells from the blood into the milk (Pirisi et al., 2000; Vivar-Quintana et al, 2006; Massoud et al., 2009). Mastitis influences also the milk electrical conductivity (MEC) through the increase of Cl− and Na+ concentration in the milk (Kitchen, 1981; Norberg et al., 2004).

Various mastitis detection tools are currently used in dairy industry, especially in dairy cow industry. One of the most widely used and proven indicators for mastitis
assessment and screening test is represented by MSCC (Paape et al., 2001; Bergonier and Berthelot, 2003; Mammadova and Keskin, 2013).

The Recording of successive MSCC improves the efficiency of subclinical mastitis predicting method (Bergonier and Berthelot, 2003). Milk Somatic Cells (MSC) are mainly represented by elements of blood origin (macrophages, polymorphonuclear neutrophil leucocytes (PMN) and lymphocytes) and they play an essential role in the defense mechanisms of the animal, especially udder health (Cuccuru et al., 1997).

Several studies show that healthy udders have regularly a MSCC value lower than 500*10^3 cells/ml (Mavrogenis et al., 1995; Fthenakis, 1996; Gonzalez-Rodriguez et al., 1996; Paape et al., 2001; Bergonier and Berthelot, 2003; Berthelot et al., 2006; Kiossis et al., 2007; Olechnowicz et al., 2014).

Nevertheless, doubts still persist on its usefulness in predicting mastitis. This is mostly related to a large number of factors affecting MSCC such as nutrition, milking technique, season, shelter conditions, age, anatomical and physiological characteristics of udder, stress, type, lactation rank, milk yield and mastitis (Larsgard and Vaabenoe, 1993; Bergonier and Berthelot, 2003; Mammadova and Keskin, 2013).

In the US the legal limit for sheep milk is 750*10^3 cells/ml (Pasteurized Milk Ordonance, 1995). Conversely in EU no official rules have been established for MSCC legal limit in small ruminants. However, the EU has suggested to establish in the future a legal limit for bulk tank milk (European Community/92/46 Directive).

Another important indicator trait for mastitis prediction is represented by the Milk Electrical Conductivity, MEC (Hamann and Zecconi, 1998). It is well-known that mastitis is associated with the increase of conductivity of the udder tissue and changes in milk ionic composition (Mammadova and Keskin, 2013). The MEC depends on the milk concentration of anions and cations and it rises in animal suffering from mastitis.
This is due to the increase of Cl and Na concentration in milk produced by infected udder (Kitchen, 1981). In dairy cows, many studies showed that MEC from cows affected by mastitis is higher than MEC from healthy cows (Hamann and Zecconi, 1998; Norberg et al., 2004). Today automatic milking systems have MEC sensors incorporated for automatic data recording and storage from each quarter during milking. The alarm system is activated when MEC traits deviate from normal aiming at predicting mastitis. Some studies found that a MEC of 5.5 mS/cm could be a good threshold to predict subclinical mastitis in dairy cow (Mammadova and Keskin, 2013). Most often, these alarms have been wrong, therefore using MEC alone to predict mastitis is a suspect method.

Some other studies in scientific literature have shown that MEC can provide a useful indirect estimation of MSCC (Peris et al., 1991; Serra et al., 1995; Foddis F. et al., 2005). Foddis et al. (2005) found that the MSCC in Sarda dairy ewes can be estimated through this regression equation:

\[ y = 2906.5x - 11292 \]

Where: \( y = \text{MSCC} \times 10^3 \) (cells/ml) and \( x = \text{MEC(in mS/cm at 25°C)} \)

In a study conducted by Fthenakis on this topic (somatic cell counts in milk of Welsh-Mountain, Dorset-Horn and Chios ewes throughout lactation) and published on Small Ruminant Research (1996), he has found the following results: from clinically tested healthy udders, 65.20% milk samples had MSCC less than 0.5*10^6 cells/ml, 87.20% had MSCC less than 0.75*10^6 cells/ml and 98.10% had less than 1.0*10^6 cells/ml. From confirmed udders with subclinical mastitis he found that the 0.80% of the samples had MSCC less than 0.5*10^6 cells/ml, 6.30% had MSCC less than 0.75*10^6 cells/ml and the 14.20% of the samples had MSCC less than 1.0*10^6 cells/ml.
Actually the accuracy and precision of mastitis screening methods based on Californian Mastitis (CMT), Milk Somatic Cell Count (MSCC) and Milk Electrical Conductivity (MEC) are almost of the same order of magnitude (McDouglas et al., 2001), but CMT and MEC are particularly useful for the description of the evolution of MSCC in milk throughout the lactation (Scharch et al., 2000).

However, the mastitis predicting system based on MEC is widely spread on many dairy cows farms in EU and US, while similar system is not still reported in literature for dairy ewes.

Satisfactory results in mastitis detection have been reached by using MEC measurements and time-series analysis, using historical data and comparing values from different quarters within milking.

MEC is affected by several factors other than mastitis such as breed, parity, milk fraction and composition (Tangora, 2010).

In dairy ewes, some studies found significant differences between a group of healthy udder halves and infected (Barth et al., 2007; Tangora, 2010).

Mastitis early detection in Sarda dairy ewes is of great importance in term of milk quality improvement, reduction of economic losses and animal welfare.

The goal of the current study is to provide further knowledge on milk quality variables (milk parameters) such as: Milk Somatic Cell Count (MSCC), Milk Electrical Conductivity (MEC), Milk Fat Content (MFC), Milk Total Protein Content (MTPC), Milk Lactose Content (MLC), Milk Chloride Content (MCC), Milk pH (MpH) and Milk Cryoscopic Point (MCP); the correlations between MSCC and other milk quality variables and the usefulness of such milk quality variables to support the mastitis detection techniques. Other important objectives are also addressed in this study, such as: the evaluation of measurement accuracy of the hand-held electrical conductivity,
called Masti-Milk in measuring the MEC on farm; assessing the representativeness of the milk sample from first squirts and milk sample from complete milking; the application of ROC curve analysis and their usefulness in supporting the early mastitis prediction techniques in Sarda dairy ewes; additionally, the applicability of the Support Vector Machines (SVMs) as a monitoring system for the early detection of mastitis based on practically recorded farm data and using previous decision threshold was addressed.
2. Materials and methods

2.1. Survey design and flock management
The survey was conducted on 11 Sarda dairy flocks located in the Northeastern of Sardinia, in the municipalities of Olbia, Golfo Aranci, and Loiri-Porto San Paolo (Italy). The average flock size was 200 to 300 ewes, which was representative of whole Sardinian regional ovine livestock industry. 25 multiparous Sarda dairy ewes in lactation were randomly selected from each flock for the survey purpose. Lambing was mostly concentrated in Autumn and spring, with one lambing per year. The lambs were kept together with the ewes for 30 to 35 days before the start of lactation. During lactation, all ewes were fed on pasture and given hay and concentrate to overcome their nutritional energy need. Ewes were housed on straw litter in prefabricated buildings and milked twice daily in parlors using automatic milking machines, with intervals between milkings of about 12 hours. During milking, ewes were fed with concentrates.

2.2. Milk sampling
On sampling days, individual milk yield was recorded by means of graduated measuring cylinders attached directly to individual milking units. Sampling was carried out during morning milking. Before sampling, ewes udders were carefully cleaned using cotton wool impregnated with 70% ethanol. From each udder, double milk samples were aseptically recorded: 20 ml for first squirts and 50 ml for udder complete milking. Milk samples were individually collected and brought to laboratory by means of transport tankers at 4° C until laboratory analysis were carried out (standard procedures UNI EN ISO 707). A total of 502 milk samples were collected from the 11 Sarda dairy flocks. Furthermore, 6,641 primiparous milk samples, collected from a 10-years historical database of ARAS laboratory (2004-2014) aiming to compare the results of an
homogenous group with the results of the survey. Both datasets were divided into two
groups for ROC and SVMs data processing. The first data group of the dataset was used
as training set (4,428 milk samples) and the second group was used as test set (2,213
milk samples) for the model validation.

2.3. Measurement and laboratory analysis
Milk samples were brought to the laboratory of Regional Livestock Farmers
Association Of Sardinia (ARAS) where they were analyzed within 24 hours maximum.
Several milk traits such as MSCC, MEC, MFC, MTPC, MLC, MCP, MpH and MCC
were determined.
MSCC was determined following the procedures recommended by International Dairy
Federation (IDF 1995) using the Fossomatic 360, which is an fluor-optical electronic
cell counter. The Fossomatic somatic cell counter (FSCC) is the most Widely used
Somatic Cell Counter in milk-testing laboratory (Gonzalo. et al., 2004). In general, the
Fossomatic is a DNA-specific counter based on the principle of optical fluorescence
(IDF 1995). The ethidium bromide dye penetrates the cell and forms a fluorescent
complex with the nuclear DNA. The sample is then exposed to blue light, which excites
the dyed cells, making them emit red light. These red light pulses are magnified,
counted by a photo multiplier detector, and multiplied by the defined working factor to
compute the number of somatic cells per millimeter.
The measurement of MFC, MTPC, MLC and MpH was carried out using an IR
spectrophotometer.
The MEC of samples was measured twice: once directly on farm using a handheld
electrical conductivity meter, named Masti-milk; and secondly in ARAS laboratory,
using a proven electrical conductivity meter, known as WTW LF 92.
2.4. Statistical analysis

2.4.1. Statistical instruments and data analysis software
All variables (The MSCC, MEC, MFC, MTPC, MLC, MCP, MpH and MCC) were processed through descriptive statistic instruments using Microsoft Office Excel 2010, aiming to describe the distribution and relationships between milk variables. MSCC values were transformed into base 10 logarithmic aiming to normalize the variable before performing statistical analysis.

Furthermore, statistical analysis were carried out on data using ROC curve analysis and SVMs techniques. The main milk variables mostly related to MSCC were processed with the aim of improving the milk abnormality detection (udder inflammation). Rstudio 9.1000 software was used to process data through SVMs techniques, which were implemented using Kernlab, an S4 package for Kernel methods in R (Karatzoglou 2004).

Rstudio is an integrated development environment for R. It consists of a console, a syntax-highlighting editor that supports direct code execution, as well as tools for plotting, history, debugging and workspace management. MEDCALC software was also used for data processing (mainly ROC curves analysis)

2.4.2. The Receiver Operating Characteristic curves (ROC curves)

2.4.2.1. Basic concept
The ROC curves (graph) is a technique of visualizing, organizing and selecting classifiers based on their performance (Fawcett 2006). The ROC methodology was developed in the early 1950s for signal detection theory to depict the tradeoff between hit rates and false alarm rates of classifiers (Egan 1975; Swets et al. 2000). Moreover, the use of ROC methodology has been extended in visualizing and analyzing techniques of behavior and diagnostic systems (Swets 1988). ROC analysis has been increasingly
used for the evaluation of clinical laboratory tests (Metz 1978; Henderson 1993; Schulzer 1994; Smith 1995). There is at least two different main uses of ROC curve analysis. First, it enable the comparison of diagnostic accuracies of several continuous tests through the Areas Under the Curves (AUCs). The AUC can be interpreted as the probability that a case and noncase pair of test values are correctly ranked (Subtil and Rabilloud 2015). Second, it allows the determination of optimal threshold (cut-off value) for dichotomization of a continuous diagnostic test. This uses various methods such as the cut-off point of ROC curve, the closest to the point (0, 1) or the farthest point from the first diagonal of the graph (Perkins and Schisterman 2006).

**2.4.2.2. ROC curve test procedure**

ROC curve is a graphical plot that illustrates the performance of a binary classifier system as its discrimination threshold (cut-off point) is varied. It is used to address a two-class prediction problem in which the outcomes are labeled either as positive (P) or negative (N). In disease detection, for every threshold (cut-off point or criterion value), there are four possible outcomes from a binary classifier. In one hand some cases with disease will be correctly classified as positive (true positive or TP), but some cases with disease will be also classified as negative (false negative or FN). In the other hand, some cases without disease will be classified as negative (true negative or TN), but some case without disease will be classified as positive (false positive or FP). The four outcomes can be formulated in a 2*2 confusion matrix as follow (figure 1):
In Roc curve, the true positive rate (Sensitivity) is plotted in function of false positive rate (100-Specificity) for different cut-off points (criterion value) of the parameter. Each point on ROC curve represents a sensitivity and specificity pair corresponding to a particular decision threshold. The Area Under the ROC Curve (AUC) is a measure of how well a parameter can distinguish between two diagnostic classes (diseased/normal). The figure 2 shows a typical ROC curve.

**Figure 1.** 2*2 confusion matrix for binary classifier. P = real positive, N = real negative, P’ = predicted positive, N’ = predicted negative.
In order to perform the ROC curve analysis, the parameter of interest for the study should be well identified as well as an independent diagnostic classifier which classifies the study subjects into two groups: diseased and normal groups. Before performing the ROC curve analysis, the raw data should be pre-processed through the exclusion of missing, incorrect and noisy values.
2.4.3. Support Vector Machines (SVMs):

2.4.3.1. General concept

The SVMs algorithms are new and innovative artificial-intelligence-based methods of data mining. They constitute a group of supervised learning methods that can be used for classification and regression. They represent an extension to non-linear models of the generalized portrait algorithm developed by Vapnik and Lerner (1963). SVMs are based on statistical learning theory and the Vapnik-Chevonenkis (VC) dimension (Verma et al. 2014). SVMs perform classification by constructing N-dimensional hyperplane that optimally separates the data into two categories (Yu et al. 2010; Verma et al. 2014). SVMs are close cousins to the well-known classical multilayer perceptron Neural Networks. Using a Kernel function, the SVMs are an alternative training method for polynomial radial based function (RBF) and multilayer perceptron classifiers in which the weights of the network are found by solving a quadratic programming problem with linear constraints, rather than by solving a non-convex, unconstrained minimization problem as in standard neural network training. Therefore the goal of the SVMs modeling is to find the hyper-plane that optimally separates clusters of vectors into two classes. In SVMs literature parlance, a predictor variable is called an attribute, a transformed attribute used to define the hyper-plane is called a feature and the task of choosing the most suitable representation is known as feature selection. A set of feature that describes one case is called vector. The nearest vectors used to define the hyper-plane are support vectors. Figure 3 below shows an overview of a standard SVMs process.
The SVMs find the hyper-plane that achieves the maximum separation through the maximization of the distance between the hyper-plane and the support vectors. Given an implicit and embedding $\Phi$ and training data $(x_i, y_i)$ from two classes such that $y_i = \pm 1$, a SVMs finds the hyperplane $w^T \Phi(x) + b = 0$, where $w$ and $b$ represent the parameters of the hyperplane that best separates the two classes (figure 3). The learnt hyperplane is optimal in the sense that it maximizes the margin whereas minimizing some measure of loss in the training data. In most cases, linear separation between two classes is generally a restrictive hypothesis of practical use. Therefore the linearly inseparable data in the input space can be transformed by suitable Kernel functions to a high dimensional feature space, where the data can then separated linearly (Shen et al. 2007; Olsen and Delen 2008) (figure 4).
Figure 4. The best hyperplane which better separate the data into two classes. The red dots have a label $y_i = +1$ and the blue dots have a label $y_i = -1$.

Mapping into high-dimensional can give rise to computational difficulties due to the fact that the dimensionality of the input variable space explodes exponentially (Bennett and Campbell 2000). SVMs avoid this issue thanks to the use of Kernel functions (Miekley et al. 2013). Kernel functions normally includes linear, polynomial, sigmoid and radial basis functions (RBF) that reduce the complexity of dimensionality by avoiding the step of explicitly mapping the data in a high-dimensional space (table 1).

**Table 1.** The most largely used Kernel functions.

<table>
<thead>
<tr>
<th>Kernel function</th>
<th>Expression</th>
</tr>
</thead>
<tbody>
<tr>
<td>The Gaussian RBF Kernel</td>
<td>$k(x, x') = \exp(-\sigma |x - x'|^2)$</td>
</tr>
<tr>
<td>The polynomial Kernel</td>
<td>$k(x, x') = (\text{scale} \cdot &lt;x, x'&gt; + \text{offset})^\text{degree}$</td>
</tr>
<tr>
<td>The linear Kernel</td>
<td>$k(x, x) = &lt;x, x'&gt;$</td>
</tr>
</tbody>
</table>

PhD Student: Djangsou Hagassou, Thesis title: Application of Support Vector Machine as support to early prediction of mastitis in Sarda dairy ewes. Animal Sciences, University of Sassari
The hyperbolic tangent Kernel

\[ k(x, x') = \tanh(\text{scale} \langle x, x' \rangle + \text{offset}) \]

The Laplacian Kernel

\[ k(x, x') = \exp(-\sigma \|x - x'\|) \]

The Bessel Kernel

\[ k(x, x') = (-\text{bessel}_n(\sigma \|x - x'\|^2)) \]

The creation of an optimal SVMs model needs the pre-selection of two key parameters: the penalty parameter (C) and the kernel parameters (gamma).

Those two parameters are known as hyper-parameters. The parameter C controls the over-fitting of the model by specifying the tolerance of misclassification. The kernel parameter (gamma) controls the degree of non-linearity of the model (Olson and Delen 2008). A smaller value of the “soft margin” (C) allows to ignore the points close to the boundary and increases the margin (Verma et al. 2014). In the figure 3, the thick line represents the decision boundary between negative elements (red) and positive (blue). The lighter lines on the margin represent the discriminating value (-1 and +1). The grayscale defines the value of the discriminating function, dark for low value and a light shade for high value (figure 5). Higher is the C value, higher is the penalty assigned to the errors/margin errors. This is shown in figure 5 (left panel), where the two closest points to the hyper-plane affect its orientation, leading to a hyper-plane that comes close to many other data points. When the C value decreases, the hyper-plane’s orientation changes, resulting in a much larger margin for the rest of data (figure 5, right panel).
The Gaussian Kernel parameter (gamma) determines the flexibility of the resulting SVMs in fitting the data. Larger value of gamma leads to over fitting of data points (figure 6). When the gamma value is small, the decision boundary is nearly linear. Conversely the increase of gamma value affect the flexibility of the decision boundary resulting in an over fitting.

Since it is not known C and gamma parameters are the best for one problem, some parameter selection has to be carried out. One of the following two methods: v-fold cross-validation (where v = number of subsets of the training data) and “gri-search” (using cross-validation), is normally applied. Therefore, various pairs of C and gamma are tried and the one with the best cross-validation is chosen. For the unbalances data sets, the receiving operating characteristic (ROC) curve and the related metric area under the curve can be more meaningful performance measures since they allow a difference between errors on positive or negative (Huang and Ling 2005).
2.4.3.2. Procedure of SVMs in milk abnormality detection
The figure 7 below shows the general procedure of milk abnormality detection. First of all, the raw data needs to be randomly divided into training and test datasets. Then, the pre-processing phase is carried out on both datasets, including the exclusion of missing values and scaling. The training dataset is used to develop the model. The first step in model development consists in selecting the Kernel function (e.g. linear, polynomial or RBF), followed by the determination of the Kernel parameters.

Figure 6. The effect of inverse-width parameter of the Gaussian Kernel (γ, known as gamma) for a fixed value of soft-margin constant (C). For a small value of γ (upper left), the decision boundary is nearly linear. As γ increases the flexibility of decision boundary increases. Large values of γ leads to over fitting (Verma et al. 2014).
For the this work the linear kernel was chosen for data processing. The hyperparameters (C and gamma) were determined (best C and gamma). In this work, the cross-validation was applied. the validated SVMs was then used to detect milk normality status in Sarda dairy ewes by using the test dataset.

2.4.3.3. Test procedure
The performance of the SVMs model was assessed considering 1000*10³ cells/ml as the upper limit in normal ewes’ milk (threshold).

The performance of the SVMs model was evaluated through the assessment of the sensitivity, specificity.

The sensitivity represents the percentage of correctly detected milk samples by the model with MSCC over 1000*10³ as follows:

\[
Specificity = \frac{TP}{TP + FN} \times 100
\]
The specificity indicates the percentage of truly identified milk samples with MSCC below $1000 \times 10^3$ cells/ml by the model as normal as follows:

$$Specificity = \frac{TN}{TN + FP} \times 100$$

### 2.4.4. MSCC threshold for normal ewes milk

A number of scientific studies have been carried out on MSCC determination in ewes milk but an universally accepted threshold has not been yet established and existing views are contradictory. Several authors have considered different MSCC in their studies (table 2). Even the European Union Directive 92/46 (1992) did not set a standard MSCC threshold for ewes milk as it has been defined for cow milk, $400 \times 10^3$ cells/ml for milk quality standards. Nevertheless several studies indicated that the threshold value of MSCC in ewe milk is greater than that in cow milk.

If the value of $1000 \times 10^3$ cells/ml is considered as the threshold, less than 1.9% of milk samples from healthy udders and less than 14.2% from unhealthy ones would be incorrectly classified as unhealthy or healthy udders respectively. Fthenakis (1995) has found that using a MSCC threshold of $0.75 \times 10^6$ cells/ml would increase the sensitivity of the method and simultaneously decrease the its specificity. If $1,500 \times 10^3$ cells/ml is considered, the sensitivity of method would decrease and at the same time it specificity would increase.

Keeping in mind that our goal in this study is to verify the accuracy of SVMs in classifying milk samples from normal udder and udder with inflammation in Sarda dairy ewes, it is extremely important to choose an accurate MSCC threshold as upper limit of MSCC in healthy ewe milk. So we are interested in knowing how many abnormal milk samples have been exactly classified as unhealthy. In a study conducted...
in 1996 on three ewe breeds (Welsh-Mountain, Dorset-Horn and Chios ewes) through entire lactation, Fthenakis has found that on a total amount of 1984 milk samples from confirmed healthy udder, 1294 (65.2%) had MSCC less than $500 \times 10^3$ cells/ml. He found also that, from a total amount of 254 milk samples from udders with confirmed subclinical mastitis, 2 (0.8%) had MSCC less than $500 \times 10^3$ cells/ml and 36 (14.2%) had less than $1000 \times 10^3$ cells/ml. In a study conducted on the “effects of intramammary infections (IMI) on somatic cell score and milk yield in Sarda sheep” (Cuccuru et al. 2011), it is highlighted that milk from infected and uninfected udder half has significantly different milk somatic cell score (MSCS), a different criterion to assess the milk cell content. A MSCS is equivalent to MSCC of 25,000 cells/ml. Each increase of MSCS is associated with the double of MSCC. They have found that the mean and standard deviation (SD) of MSCS was greater in udder half with IMI ($7.71 \pm 0.82$) than in udder half without IMI ($5.53 \pm 1.02$) and $P$-value $< 0.01$. The formula used to perform this calculation was $\log_2(MSCC/100,000)+3$ (Kirk 1984).

<table>
<thead>
<tr>
<th>MSCC threshold (cells/ml)</th>
<th>Authors</th>
</tr>
</thead>
<tbody>
<tr>
<td>$250 \times 10^3$ cells/ml</td>
<td>Heredia and Iturritza (1988); Romeo et al. (1993); Cruz et al. (1994).</td>
</tr>
<tr>
<td>$300 \times 10^3$ cells/ml</td>
<td>Fruganti et al. (1985)</td>
</tr>
<tr>
<td>$500 \times 10^3$ cells/ml</td>
<td>Travnicek et al. (1978); Vitkov and Vitanov (1980); Contreras et al. (1999); Berthelot et al. (2006).</td>
</tr>
<tr>
<td>$750 \times 10^3$ cells/ml</td>
<td>Fthenakis (1996)</td>
</tr>
<tr>
<td>$1,000 \times 10^3$ cells/ml</td>
<td>Green (1984); Mackis and Rodgers (1986); El-Masannat (1987); Fthenakis (1988); Ipsiladis et al. (1988); Fthenakis et al. (1991); Stefanakis et al. (1995).</td>
</tr>
<tr>
<td>$1,500 \times 10^3$ cells/ml</td>
<td>Maisi et al. (1987); Mavrogenis et al. (1995)</td>
</tr>
</tbody>
</table>
3. Results and discussion

3.1. Milk variables description

First of all, the distribution of milk variables was analyzed using histograms and box-plots (Box-and-Whiskers plots). MSCC dataset was transformed into base 10 logarithm to normalize its distribution. After that, the overall means and standard deviations of all Sarda dairy ewes milk variables considered in this survey are calculated on all data recorded during the survey and summarized in the table 3 to better appreciate an overview of data variability. Office Excel 2007 version was used to highlight the distributions of milk variables. All milk variable distributions showed below are based on 502 milk samples collected in this survey. Data were recorded during the entire period of lactation (from October 2013 to June 2014). Over all means and standard deviations of all milk variables considered in the survey are shown in the table 4 below.

<table>
<thead>
<tr>
<th>Variable</th>
<th>Mean ± SD.</th>
</tr>
</thead>
<tbody>
<tr>
<td>MEC.Mastimilk (mS/cm)</td>
<td>4.83 ± 0.46</td>
</tr>
<tr>
<td>MECwtwl92. (mS/cm)</td>
<td>4.75 ± 0.52</td>
</tr>
<tr>
<td>MCP (*10^-3 °C)</td>
<td>576.88 ± 12.50</td>
</tr>
<tr>
<td>MFC (%)</td>
<td>6.43 ± 1.24</td>
</tr>
<tr>
<td>MTPC (%)</td>
<td>5.54 ± 0.68</td>
</tr>
<tr>
<td>MLC (%)</td>
<td>4.74 ± 0.47</td>
</tr>
<tr>
<td>Log.(MSCCfs *10^3)</td>
<td>2.27 ± 0.69</td>
</tr>
<tr>
<td>Log.(MSCCcm*10^3)</td>
<td>2.33 ± 0.66</td>
</tr>
<tr>
<td>MpH</td>
<td>6.59 ± 0.13</td>
</tr>
<tr>
<td>MCC (g/l)</td>
<td>146.18 ± 63.44</td>
</tr>
</tbody>
</table>

Definition of the abbreviations: MEC (Milk Electrical Conductivity); MFC (Milk Fat Content); MTPC (Milk Total Protein Content); MLC (Milk Lactose Content); MSCC/fs (Milk Somatic Cell Count first squirts); MSCC/cm (Milk Somatic Cell Count complete milking); MCP (Milk Cryoscopic Point); MpH (Milk pH) and MCC (Milk Chlorides Content).
3.1.2. Milk somatic cell count (MSCC)
The MSCC was determined following the procedures recommended by International Dairy Federation (IDF 1995) using the Fossomatic 360, which is an fluor-optical electronic cell counts. The Fossomatic SCC method (FSCC) is the most Widely used Somatic Cell Counter (SCC) in milk-testing laboratory. The determination of MSCC on 540 milk samples was carried out in ARAS laboratory. The figures 8 shows the distribution of MSCC. It highlights the extremely asymmetrical distribution of MSCC with the majority of data below $1,500 \times 10^3$ cells/ml. This result leads to consider the normalization of MSCC data using the base 10 logarithm. The figure 9 shows the distribution of the data transformed into base 10 logarithm.
Figure 8. Distribution of MSCCcm*10^3 from dataset of 502 milk samples collected from 11 flocks in Northeastern of Sardinia. The Histogram (panel at left) shows the distribution through 15 classes of 398 cells/class of the frequency. The box-plot (right panel) shows the same distribution.

PhD Student: Djangou Hagassou, Thesis title: Application of Support Vector Machine as support to early prediction of mastitis in Sarda dairy ewes. Animal Sciences, University of Sassari
Figure 9. Distribution of \( \log(\text{MSCCcm} \times 10^3) \) from data of 502 milk samples collected from 11 flocks in Northeastern of Sardinia. The Histogram (panel at left) shows the distribution through 15 classes of 0.2/class of the frequency. The box-plot (right panel) shows the same distribution.
3.1.1. Milk electrical conductivity (MEC)
The MEC was measured in milliSiemens per centimeter (mS/cm) from both udder-halves during milking. 50 ml of milk from complete milking were collected during morning milking. The first measurement of MEC was carried out immediately on farm by using the Masti-milk, a handheld prototype of electrical conductivity meter. When milk samples were brought to ARAS laboratory, MEC was measured again using laboratory electrical conductivity meter (WTW LF 92) for accurate electrical conductivity measurement. Figures 9 shows the MEC distribution of 502 milk sample collected on 11 Sarda dairy flocks in the Northeastern of Sardinia. The Figure 10 show the distribution of MEC of 502 samples collected in 11 flocks.
Figure 10. MEC distribution of 502 milk samples collected from 11 flocks. 15 classes of frequency and class wide of 0.28 mS/cm (panel in the left). In the right panel, the box-plot of the same data.
3.1.2. Milk Chloride content (MCC) and milk pH (MpH)
The assessment of MCC was carried out according to the procedures recommended by International Dairy Federation (IDF 1995), using chemical procedures. The milk pH determination was carried out using the pH-meter. The chemical analysis were performed within 24 hours from every sampling. The 502 milk samples were collected and analyzed during the entire period of lactation on 11 flocks in the Northeastern of Sardinia (from October 2013 to June 2014). The figure 11 (below) shows the distribution of MCC. The MpH was determined by the use of the pH-meter. It was the first milk parameter determined aiming to avoid any post-milking alteration. The distribution of MpH is highlighted in the figure 12 below.
Figure 11. Distribution of MCC from data of 502 milk samples collected from 11 flocks in Northeastern of Sardinia. The Histogram (panel at left) shows the distribution through 15 classes with a class range of 42.866 g/l. The box-plot (right panel) shows the same distribution.

PhD Student: Djangsou Hagassou, Thesis title: Application of Support Vector Machine as support to early prediction of mastitis in Sarda dairy ewes. Animal Sciences, University of Sassari
Figure 12. Distribution of MpH from 502 milk samples collected from 11 flocks in Northeastern of Sardinia. The Histogram (panel at left) shows the distribution through 15 classes with a class range of 0.084. The box-plot (right panel) shows the same distribution.

PhD Student: Djangou Hagassou, Thesis title: Application of Support Vector Machine as support to early prediction of mastitis in Sarda dairy ewes. Animal Sciences, University of Sassari
3.1.4. Milk lactose content (MLC)
Mastitis results in tissue damage therefore the biosynthesis of lactose is decreased (Pyörälä 2003). In addition the tissue damage allows the passage of lactose from alveolar lumen into the blood (osmotic phenomenon) leading to the decrease of the lactose in the milk. Therefore the milk lactose concentration could be used as mastitis indicator since it clearly decreases during the udder inflammation. The association between lactose, the MSCC and mastitis has been studied already since the nineteen-eighties and the results are reviewed by Kitchen (1981). Recently lactose has been proposed to be one of the most useful markers of mastitis for the future use with a proposed threshold value of 4.7% (Hamann and Krömker 1997; Hamann 2002). All the previous studies have been conducted on dairy cow, and any on dairy ewe. Therefore a specific study needs to be implemented on dairy ewe with the aim of verifying the usefulness of lactose as screening indicator.

The figure 13 shows the distribution of the MLC from 502 milk samples collected on 11 flocks in the Northeastern of Sardinia.
Figure 13. Distribution of MLC from data of 502 milk samples collected from 11 flocks in Northeastern of Sardinia. The Histogram (panel at left) shows the distribution through 15 classes with a class range of 0.30 (%). The box-plot (right panel) shows the same distribution.
3.2. Masti-Milk efficiency
Several handheld electrical conductivity meters are currently used in dairy cow farms for MEC measurement. No electrical conductivity meter has been design and realized specifically for MEC measurement in ovine milk. The one what was adopted and used for MEC measurement in farm in this survey is called Masti-milk (figure 15). The Misti-milk is already used in dairy cow for MEC measurement in farm. It is equipped with five lights which have different colors and light according to the MEC value. The color of light indicates the mastitis risk status. Actually the device is also provided with a reading scale of mastitis risk status of the animal as shown in table 5. The aim in choosing this device was to test its accuracy in measuring MEC in ewes milk before adopting it as support to mastitis first screening in Sarda dairy ewes. For this purpose, the MEC measured on farm through the Masti-milk was compared with the once measured in ARAS laboratory (WTW LF 92). The table 6 shows a highly significant correlation between MEC recorded using Masti-Milk and MEC recorded using the WTW LF 92 ($r = 0.954$). Furthermore, the linear regression between MEC measured in farm by using Masti-Milk and the MEC measured in laboratory by the means of the proven electrical conductivity meter known as WTW LF 92 which was used in ARAS laboratory (figure 14), highlighted the significant accuracy of the device in measuring MEC ($R^2 = 0.9124$):

$$y = 1.0413x - 0.2827$$

Where: $y =$ MEC (mS/ml at 25°C) and $x =$ MEC (mS/ml at 25°C measured on farm through Masti-Milk).
As introduced above, the handheld EC-meter used for MEC measurement in the farm was the one normally used for MEC measurement in dairy cow industry. It is provided with the following scale (table 4) for data interpretation. It uses MEC as input data and provides in output the mastitis risk status in term of percentage and the chloride milk content status. The basic concept is that, high is the MEC value, largely altered is the milk chloride content and higher is the risk of udder inflammation (mastitis risk). The figure 15 shows the handheld EC-meter used during this survey, conducted on 11 flocks in the North-eastern of Sardinia. The high $R^2$ value could be interpreted as the higher accuracy of the device in measuring MEC.

Table 4. Masti-Milk interpretation scale, provided by the manufacturer for data interpretation

<table>
<thead>
<tr>
<th>MEC (mS/cm)</th>
<th>Milk chloride status</th>
<th>Mastitis risk (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>3.00</td>
<td>3.00</td>
<td></td>
</tr>
<tr>
<td>3.50</td>
<td>3.50</td>
<td></td>
</tr>
<tr>
<td>4.00</td>
<td>4.00</td>
<td></td>
</tr>
<tr>
<td>4.50</td>
<td>4.50</td>
<td></td>
</tr>
<tr>
<td>5.00</td>
<td>5.00</td>
<td></td>
</tr>
<tr>
<td>5.50</td>
<td>5.50</td>
<td></td>
</tr>
<tr>
<td>6.00</td>
<td>6.00</td>
<td></td>
</tr>
<tr>
<td>6.50</td>
<td>6.50</td>
<td></td>
</tr>
<tr>
<td>7.00</td>
<td>7.00</td>
<td></td>
</tr>
</tbody>
</table>

Figure 14. Linear regression between MEC measured by Masti-Milk and MEC by WTW FL 92. Data from 502 milk samples collected from 11 flocks in Northeastern of Sardinia
<table>
<thead>
<tr>
<th>Chlorides range</th>
<th>Condition</th>
<th>MEC value</th>
</tr>
</thead>
<tbody>
<tr>
<td>&lt;4.35</td>
<td>Normal milk</td>
<td>0</td>
</tr>
<tr>
<td>4.35-4.65</td>
<td>Chlorides slightly altered</td>
<td>20-40</td>
</tr>
<tr>
<td>4.65-4.95</td>
<td>Chloride moderately altered</td>
<td>40-60</td>
</tr>
<tr>
<td>4.95-5.25</td>
<td>Chlorides highly altered</td>
<td>60-80</td>
</tr>
<tr>
<td>&gt;5.25</td>
<td>inflammatory status</td>
<td>80-100</td>
</tr>
</tbody>
</table>

**Figure 15.** Handheld electrical conductivity meter, known as Masti-milk, used in MEC measurement in farm during the survey in 11 flocks located in the North-eastern of Sardinia
3.3. Relationship between milk first squirts and milk from complete milking

After laboratory analysis of milk first squirts samples and milk from complete milking samples it emerges that the milk first squirts samples are highly representatives of milk from complete milking in terms of MSCC (table 2) with \( r = 0.955 \) and \( r = 0.915 \) respectively. This result leads to consider that the milk first squirts can be used to determine the MSCC with no significant differences in using samples from complete milking. However using the milk first squirt samples has the advantage to avoid the huge waste of milk currently collected for laboratory analysis (20 ml for milk first squirts against 50 ml for milk complete milking). The figure 16 below, show the linear regressions between milk first squirts and milk from complete milking. The linear regression highlighted un important coefficient of determination \( (R^2 = 0.9887) \):

\[
y = 0.9734x + 26,676
\]

Where: \( y = (\text{MSCC.cm}) \times 10^{-3} \) and \( x = (\text{MSCC.fs}) \times 10^{-3} \)

![Linear regression between MSCC.fs (from milk first squirts) and MSCC.cm (from complete milking). Data from 502 milk samples collected from 11 flocks in North-eastern of Sardinia](image)

**Figure 16.** Linear regression between MSCC.fs (from milk first squirts) and MSCC.cm (from complete milking). Data from 502 milk samples collected from 11 flocks in North-eastern of Sardinia

PhD Student: Djangsou Hagassou, Thesis title: Application of Support Vector Machine as support to early prediction of mastitis in Sarda dairy ewes. Animal Sciences, University of Sassari
3.4. Relationships between most relevant milk parameter linked with MSCC

The overall correlation coefficients between the most relevant milk variables were assessed and summarized in the table 5. They show how significantly the milk variables are linked to each other. The processing data software such as: Microsoft office excel 2010, Medcal and Rstudio 9.1000 were used to assess the relationships between the milk variables and their significances.

The milk electrical conductivity measured on farm using the masti-milk and the one measured in laboratory are significantly correlated with a high Pearson product-moment correlation coefficient or Pearson's r (r = 0.954). This result leads to consider that the MEC measured in farm using the handheld electrical conductivity meter is a good estimation of MEC measured in laboratory (table 5). Also the degree of linear dependence between MSCC from milk first squirts sample (MSCCfs) and MSCC form complete milking (MSCCcm) is highly significant and leads to consider that the MSCCfs is representative of MSCC from complete milking with r = 0.915 (table 5).

There is an interesting correlation between MSCC and MEC, highlighted by Pearson’s r in the table 6 showing r = 0.428-0.472 according to electrical conductivity meter used for MEC measurement and milk samples from first squirts or complete milking.

Milk lactose content is negatively highly correlated with MEC (r = - 0.687 or - 0.717 according to measurement method), MSCC (r = - 0.532 or - 0.579 according to milk sample) and chlorides content (r = - 0.942). It is also correlated positively with the milk cryoscopic point (r = 0.667) and the milk pH (r = 0.656). Furthermore, the milk chlorides content is also highly correlated with the MSCC (r = 0.531 or 0.572 according to the milk sample used in the analysis).
Table 5. Overall correlation coefficients between milk variables, calculated from a dataset of 02 samples collected from 11 flocks during conducted in the Northeastern of Sardinia.

<table>
<thead>
<tr>
<th></th>
<th>Masti-Milk</th>
<th>WTWLF92</th>
<th>MFC</th>
<th>MPC</th>
<th>MLC</th>
<th>Log.(MSCCfs)</th>
<th>Log.(MSCCcm)</th>
<th>MCP</th>
<th>MpH</th>
<th>MCC</th>
</tr>
</thead>
<tbody>
<tr>
<td>Masti-Milk</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>WTWLF92</td>
<td>0.954</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MFC</td>
<td>-0.190</td>
<td>-0.220</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MPC</td>
<td>-0.069</td>
<td>-0.120</td>
<td>0.394</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MLC</td>
<td>-0.688</td>
<td>0.719</td>
<td>-0.228</td>
<td>-0.321</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Log.(MSCCfs)</td>
<td>0.462</td>
<td>0.479</td>
<td>0.144</td>
<td>0.304</td>
<td>-0.585</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Log.(MSCCcm)</td>
<td>0.471</td>
<td>0.483</td>
<td>0.134</td>
<td>0.285</td>
<td>-0.585</td>
<td>0.973</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MCP</td>
<td>-0.411</td>
<td>-0.490</td>
<td>0.138</td>
<td>0.018</td>
<td>0.672</td>
<td>-0.411</td>
<td>-0.422</td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>MpH</td>
<td>-0.277</td>
<td>-0.323</td>
<td>-0.144</td>
<td>-0.213</td>
<td>0.649</td>
<td>-0.210</td>
<td>-0.215</td>
<td>0.530</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>MCC</td>
<td>0.824</td>
<td>0.868</td>
<td>-0.007</td>
<td>0.100</td>
<td>-0.943</td>
<td>0.575</td>
<td>0.580</td>
<td>-0.645</td>
<td>-0.530</td>
<td>1</td>
</tr>
</tbody>
</table>

Definition of the abbreviations: MEC (Milk Electrical Conductivity); MFC (Milk Fat Content); MPC (Milk Protein Content); MLC (Milk Lactose Content); MSCCfs (Milk Somatic Cell Count first squirts); MSCCcm (Milk Somatic Cell Count complete milking); MCP (Milk Cryoscopic Point); MpH (Milk pH) and MCC (Milk Chlorides Content). P value < 0.0001.
3.4.1. Relationship between milk electrical conductivity (MEC) and milk somatic cell count (MSCC)
The MSCC and MEC are considered as markers of udder inflammation specially in dairy cow production system.

Automatic measurement of MEC in milking systems have been studied in cattle providing good sensitivity and specificity results: 100 and 95% (De Mol et al. 1999), 88 and 97% (Mele et al. 2001), 92 and 93% (Cavero et al. 2006). These two milk parameters are useful for mastitis detection, particularly in clinical form.

Despite the great potential of MSCC and MEC highlighted in dairy cows field, there is little scientific studies published on MSCC and MEC in dairy ewes production sector in general and in Sarda dairy ewes in particular. The table 5 of this work shows an interesting correlations between MSCC and MEC (r = 0.428-0.472). The linear regression between MSCC and MEC assessed using a dataset of 502 milk samples collected in 11 flocks in the Northeastern of Sardinia is represented by the following equation and the figure 17 ($R^2 = 0.1665$):

$$y = 2913.3x - 12713$$

where: $y = \text{MSCC} \times 10^3$ and $x = \text{MEC (mS/cm at 25° C)}$.

This result is in accordance with the previous studies conducted by Peris et al. (1991), Serra et al. (1995) and Foddis et al. (2005):

$$y = 2906.5x - 11292$$

where: $y = \text{MSCC} \times 10^3 \text{ cells/ml}$ and $x = \text{MEC (mS/cm at 25° C)}$. 

PhD Student: Djangsou Hagassou, Thesis title: Application of Support Vector Machine as support to early prediction of mastitis in Sarda dairy ewes. Animal Sciences, University of Sassari
3.4.2. Relationship between Milk Chlorides Content (MCC) and milk somatic cell count (MSCC)

The table 5 above shows the degree of dependence between MSCC and MCC highlighted by r value (r = 0.531 or 0.572 depending on type of sample). The figure 18 below shows the regression equation between the two variables (R² = 0.3246).

The highly significant relationship between MCC and MSCC is due to the passage of chlorides from blood into the alveolar lumen, caused mostly by the deterioration of the membranes of epithelial cells in the mammary gland (Hamann and Zecconi, 1992; Diaz et al., 2011). Furthermore, Milk Somatic Cells are mainly represented by elements of blood origin (macrophages, polymorphonuclear neutrophil leucocytes (PMN) and lymphocytes) and they play an essential role in the defence mechanisms of the animal, especially udder health (Cuccuru et al., 1997). Therefore in case of udder inflammation the concentration of defense elements coming from the blood increase rapidly in milk.
leading to the increase of the MSCC. Therefore, knowing the MCC value, it is possible to estimate the MSCC through the following regression equation:

\[ y = 42,15x - 4924,2 \]

Where: \( y = \text{MSCC} \times 10^{-3} \) and \( x = \text{MCC (g/l)} \).

Figure 18. Regression between MSCC and MCC calculated on 502 milk samples collected during the survey conducted on 11 flocks in the Northeastern of Sardinia.

### 3.4.3. Relationship between milk lactose content (MLC) and milk somatic cell count (MSCC)

The milk trait showed a significant correlation with the MSCC (table 5). The Pearson correlation coefficient calculated on 502 milk samples from 11 Sarda dairy flocks is \( r = -0.579 \).

The highly negative relationship between MLC and MSCC is probably due to the passage of lactose from the alveolar lumen into the blood, caused by the deterioration of the membranes of epithelial cells in the mammary gland (osmotic phenomenon). In
addition, the biosynthesis of lactose could be affected in udder severely affected by udder inflammation (such as mastitis) contributing to the decrease of MLC. The figure 19 shows the regression equation between the two milk variable ($R^2 = 0.3155$):

$$y = -5506.3x + 27331$$

Where $y = \text{MSCC} \times 10^{-3}$ and $x = \text{lactose content (\%)}$. 

![Figure 19](image_url)

**Figure 19.** Regression between MSCC and milk lactose content calculated on 502 milk samples collected during the survey conducted on 11flocks in the Northeastern of Sardinia

### 3.4.4. Relationship between MEC and milk lactose content (MLC)

The negative and highly significant relationship between MEC and MLC with $r = -0.719$ (table 5) is in line with the studies found in scientific literature on this topic (Das and Singh, 2000; Ying et al., 2004; Bansal et al., 2005; Diaz et al., 2011). Several studies in dairy goat industry highlighted that the increase of MEC and MSCC in milk produced by the animals affected by udder inflammation, is related to the decrease of MLC (Barros and Leitão, 1992; Martí et al., 1999; Leitner et al., 2004; Diaz et al., 2011). This significant relationship is due to the passage of chlorides form blood into
the alveolar lumen, caused mostly by the deterioration of the membranes of epithelial cells in the mammary gland and, it is balanced by a decrease of in lactose concentration (Hamann and Zeconni, 1992; Diaz et al., 2011). The figure 20 shows the linear regression between MEC and MLC ($R^2 = 0.472$):

$$y = -0.5004x + 7.1366$$

Where $y = MLC$ (%) and $x = MEC$ (mS/ml at 25° C).

![Figure 20. Regression between MSCC and milk lactose content calculated on 502 milk samples collected during the survey conducted on 11 flocks in the Northeastern of Sardinia](image)

3.4.6. Relationship between milk electrical conductivity (MEC) and Milk Chloride content (MCC)

The MCC is one of the most accurate indirect methods in dairy cow mastitis detection (Fernando R.S. et al., 1985). Combined with the other indirect methods for mastitis screening test, MCC can have considerable potential for mastitis detection also in dairy ewe mastitis detection. Here below, two indirect methods for assessing the MCC are...
determined. MEC and milk lactose content can be used as independent variables for MCC estimation.

The table 5 above shows a great correlation between MCC and MEC (r = 0.868) and between MCC and milk lactose content (r = -0.942). Those two methods can avoid the time-consuming chemical analysis for MCC determination. The figure 21 shows the linear regression between MEC and MCC of 502 milk samples collected in 11 Sarda dairy flocks in the Northeastern of Sardinia (R² = 0.797):

\[ y = 86.161x - 265.4 \]

Where y = MCC (g/l) and x = MEC (mS/cm at 25° C).

**Figure 21.** Linear regression between MCC and MEC of 502 milk samples collected in 11 flocks in the Northeastern of Sardinia
3.5. ROC curve analysis results
The ROC curves analysis was carried out using 6,641 primiparous milk samples, collected from a 10-years historical database of ARAS laboratory (2004-2014). The purpose of choosing only primiparous dairy ewes was to form an homogeneous group aiming at minimizing the effects of factors affecting milk parameters. First, data was divided into two groups: one for the training (4,428 milk samples) and the second group of data for the test (2,213 milk samples). the two datasets were used for training (training set) and the validation of the ROC curves analysis (test set). The same two datasets were also used for the implementation of the SVMs model.

Three different MSCC thresholds (cut-off points or criterion values) were identified and used according to the studies done by previous authors on this topic (tables 3 and 7). In a study conducted by Fthenakis (1996) on three dairy ewes breeds (Welsh-Mountain, Dorset-Horn and Chios), he has found different results considering different thresholds (see table 6 below). In the current study, the same MSCC decision thresholds were used for the ROC curves analysis.

As aforementioned, a ROC curve is a plot of sensitivity against the false-positive rate, defined as 1-specificity. Therefore each point on graph is generated by the different decision thresholds.

In general, the closer the curve follow the left-hand border and the top border of the ROC space, the more accurate is the test. The closer the curve comes to the diagonal of ROC space, the less accurate is the test. Consequently the area under the ROC curve (AUC) represents the measure of the test accuracy. Therefore the AUC summarizes the accuracy of a test by a single number and it do not change according to the prevalence of the udder inflammation. The ROC curves analysis outputs of the three MSCC...
decision thresholds and the three milk variables (MEC, MLC and MpH) are highlighted in the paragraphs below.

Table 6. Results of a study conducted by Fthenakis (1996) on the relationship between MSCC and mastitis. He has found different percentages of healthy and udders with subclinical mastitis, considering three different MSCC (500*10^3, 750*10^3 and 1000*10^3 cells/ml). His study was conducted on three ewe breeds (Welsh-Mountain, Dorset-Horn and Chios).

<table>
<thead>
<tr>
<th>MSCC (cells/ml)*10^-3</th>
<th>Healthy</th>
<th>Diseased</th>
</tr>
</thead>
<tbody>
<tr>
<td>&lt; 500</td>
<td>&lt; 750</td>
<td>&lt; 1000</td>
</tr>
<tr>
<td>Percentage (%)</td>
<td>65.20</td>
<td>87.20</td>
</tr>
<tr>
<td>&lt; 500</td>
<td>&lt; 750</td>
<td>&lt; 1000</td>
</tr>
<tr>
<td>0.80</td>
<td>6.30</td>
<td>14.20</td>
</tr>
</tbody>
</table>

3.5.1. ROC curve analysis at MSCC decision threshold of 500*10^3 cells/ml
At the MSCC decision threshold of 500*10^3 cells/ml, the ROC curves analysis was carried out processing 6,641 primiparous milk samples divided into two datasets: 4,428 milk samples for training set and 2,213 milk samples for test set.

The three ROC curves for the three udder inflammation detection tests (MEC, MLC and MpH) are represented in figures 22, 23 and 24. The different cut-off points (criterions) are denoted by the orange marker on the curve.

Considering the MEC as predictive variable and the MSCC as classification variable with the MSCC decision threshold of 500*10^3 cell/ml, the test highlighted that, from the 4,428 milk samples 906 (20.46%) were found positives and 3,522 (79.54%) were negatives. The sensitivity and specificity of test were 58.50% and 66.72% respectively, with the associated criterion of 4.82 mS/cm (figure 21). The tables 7 and 8 show the different outputs.
Figure 22. Sensitivity and specificity of ROC curve, using MEC as predictive variable and MSCC as classification variable. Of 4428 milk samples and using a threshold $500 \times 10^3$ cells/ml, 906 (20.46%) were found positives and 3,522 (79.54%) were negatives. The test has 4.82 mS/cm as criterion value with a sensitivity of 58.50% and specificity of 66.72%.

Table 7. ROC statistic test related to MEC as predictive variable and MSCC as classification variable with decision threshold of $500 \times 10^3$ cell/ml

<table>
<thead>
<tr>
<th>Area under the ROC curve (AUC)</th>
<th>0.678</th>
</tr>
</thead>
<tbody>
<tr>
<td>Standard Error $^a$</td>
<td>0.0100</td>
</tr>
<tr>
<td>95% Confidence interval $^b$</td>
<td>0.664 – 0.692</td>
</tr>
<tr>
<td>Significance level P (Area=0.5)</td>
<td>$&lt; 0.0001$</td>
</tr>
</tbody>
</table>

$^a$ DeLong et al., 1988  
$^b$ Binomial exact
Table 8. MEC representative parameters calculated through the Youden index with MSCC at 500*10^3 cells/ml

<table>
<thead>
<tr>
<th>Associated criterion</th>
<th>4.82</th>
</tr>
</thead>
<tbody>
<tr>
<td>95% Confidence interval a</td>
<td>4.68 – 4.94</td>
</tr>
<tr>
<td>Sensitivity</td>
<td>58.50</td>
</tr>
<tr>
<td>Specificity</td>
<td>66.72</td>
</tr>
</tbody>
</table>

Considering the MLC as predictive variable and the MSCC as classification variable with the MSCC decision threshold of 500*10^3 cell/ml, the test highlighted that, from the 4,428 milk samples 906 (20.46%) were found positives and 3,522 (79.54%) were negatives. The sensitivity and specificity of test were 67.66% and 70.50% respectively, with the associated criterion of 4.84% (figure 23). The tables 9 and 10 show the different outputs.

Figure 23. Sensitivity and specificity of ROC curve, using MLC as predictive variable and MSCC as classification variable. Of 4428 milk samples and using a threshold 500*10^3 cells/ml, the model found 4.84% as criterion value with a sensitivity of 67.66% and specificity of 70.50%
Choosing the Mph as predictive variable and the MSCC as classification variable with the MSCC decision threshold of 500*10^3 cell/ml, the test highlighted that, from the 4,428 milk samples 906 (20.46%) were found positives and 3,522 (79.54%) were negatives. The sensitivity and specificity of test were 61.48% and 47.42% respectively, with the associated criterion of 6.64 (figure 24). The tables 11 and 12 show the different outputs.

The extremely low specificity is shown by the close to the diagonal (figure 24 below).
Figure 24. Sensitivity and specificity of ROC curve, using MpH as predictive variable and MSCC as classification variable. Of 4428 milk samples and using a threshold $500 \times 10^3 \text{ cells/ml}$, the model found 6.64 as criterion value with a sensitivity of 61.48% and specificity of 47.42%.

Table 11. ROC statistic test related to MpH as predictive variable and MSCC as classification variable with decision threshold of $500 \times 10^3 \text{ cell/ml}$

<table>
<thead>
<tr>
<th></th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Area under the ROC curve (AUC)</td>
<td>0.559</td>
</tr>
<tr>
<td>Standard Error $^a$</td>
<td>0.0104</td>
</tr>
<tr>
<td>95% Confidence interval $^b$</td>
<td>0.544 – 0.574</td>
</tr>
<tr>
<td>Significance level P (Area=0.5)</td>
<td>&lt; 0.0001</td>
</tr>
</tbody>
</table>

$^a$ DeLong et al., 1988

$^b$ Binomial exact
The predictive test using the MEC, MLC and MpH for udder inflammation status detection was validated using a dataset of 2,213 milk samples previously selected. The results of the test validation is shown below in the table 13. As aforementioned, the accuracy of the ROC curve test is assessed using the AUC. The table 14 shows the AUC of the three predictive variables (MEC, MLC and MpH) at MSCC decision threshold of 500*10^3 cells/ml. As it can be denoted through the ROC curve outputs (figure 25), the MLC performs better than MEC (0.757 vs 0.678 AUC) and MpH (0.757 vs 0.559 AUC), and MEC performs better than MpH (0.678 vs 0.559 AUC).

Table 12. MpH representative parameters calculated through the Youden index with MSCC at 500*10^3 cells/ml

<table>
<thead>
<tr>
<th>Associated criterion</th>
<th>≤6,64</th>
</tr>
</thead>
<tbody>
<tr>
<td>95% Confidence interval a</td>
<td>≤6,548160311 to ≤6,7</td>
</tr>
<tr>
<td>Sensitivity</td>
<td>61,48</td>
</tr>
<tr>
<td>Specificity</td>
<td>47,42</td>
</tr>
</tbody>
</table>

Table 13. Test validation for ROC curve analysis at MSCC decision threshold of 500*10^3 cells/ml. The MEC was the predictive variable considered.

<table>
<thead>
<tr>
<th>Type of dataset</th>
<th>Data</th>
<th>Criterion</th>
<th>sensitivity</th>
<th>95% CI</th>
<th>Specificity</th>
<th>95% CI</th>
</tr>
</thead>
<tbody>
<tr>
<td>Training set</td>
<td>4,428</td>
<td>4.82 mS/cm</td>
<td>61.88</td>
<td>58.1 – 65.5</td>
<td>65.83</td>
<td>64.3 – 67.3</td>
</tr>
<tr>
<td>Test set</td>
<td>2,213</td>
<td>4.82 mS/cm</td>
<td>62.09</td>
<td>56.7 – 67.3</td>
<td>63.63</td>
<td>61.4 – 65.8</td>
</tr>
</tbody>
</table>

Table 14. AUCs of the three different tests (MEC, MLC and MpH) at MSCC threshold of 500*10^3 cells/ml

<table>
<thead>
<tr>
<th>Predictive variable</th>
<th>AUC</th>
</tr>
</thead>
</table>

PhD Student: Djangsou Hagassou, Thesis title: Application of Support Vector Machine as support to early prediction of mastitis in Sarda dairy ewes. Animal Sciences, University of Sassari
**3.5.2. ROC curve analysis at MSCC decision threshold of 750*10^3 cells/ml**

At the MSCC decision threshold of 750*10^3 cells/ml, the ROC curves analysis was carried out processing 6,641 primiparous milk samples divided into two datasets: 4,428 milk samples for training set and 2,213 milk samples for test set.

The three ROC curves for the three udder inflammation detection tests (MEC, MLC and MpH) are represented in figures 26, 27 and 28. The different cut-off points (criterions) are denoted by the orange marker on the curve.

Focusing on the MEC as predictive variable and the MSCC as classification variable with the decision threshold of 750*10^3 cell/ml, the test highlighted that, from the 4,428 milk samples 682 (15.40%) were found positives and 3,746 (84.60%) were negatives.

---

**Table:**

<table>
<thead>
<tr>
<th>Test</th>
<th>AUC</th>
</tr>
</thead>
<tbody>
<tr>
<td>MEC test</td>
<td>0.678</td>
</tr>
<tr>
<td>MLC test</td>
<td>0.757</td>
</tr>
<tr>
<td>MpH test</td>
<td>0.559</td>
</tr>
</tbody>
</table>

---

*PhD Student: Djangou Hagassou, Thesis title: Application of Support Vector Machine as support to early prediction of mastitis in Sarda dairy ewes. Animal Sciences, University of Sassari*
The sensitivity and specificity of test were 61.88% and 65.83% respectively, with the associated criterion of 4.82 mS/cm (figure 26). The tables 15 and 16 show the different outputs.

**Figure 26.** Sensitivity and specificity of ROC curve, using MEC as predictive variable and MSCC as classification variable. Of 4428 milk samples and using a threshold $750 \times 10^3$ cells/ml, 682 (15.40%) were found positives and 3,746 (84.60%) were negatives. The test has 4.82 mS/cm as criterion value with a sensitivity of 61.88% and specificity of 65.83%.

**Table 15.** ROC statistic test related to MEC as predictive variable and MSCC as classification variable with decision threshold of $750 \times 10^3$ cell/ml

<table>
<thead>
<tr>
<th>Area under the ROC curve (AUC)</th>
<th>0.700</th>
</tr>
</thead>
<tbody>
<tr>
<td>Standard Error $^a$</td>
<td>0.0107</td>
</tr>
<tr>
<td>95% Confidence interval $^b$</td>
<td>0.686 – 0.713</td>
</tr>
<tr>
<td>Significance level P (Area=0.5)</td>
<td>$&lt;0.0001$</td>
</tr>
</tbody>
</table>

$^a$ DeLong et al., 1988  
$^b$ Binomial exact
Table 16. MEC representative parameters calculated through the Youden index with MSCC at $750 \times 10^3$ cells/ml

<table>
<thead>
<tr>
<th>Associated criterion</th>
<th>4.82</th>
</tr>
</thead>
<tbody>
<tr>
<td>95% Confidence interval a</td>
<td>4.71 – 4.96</td>
</tr>
<tr>
<td>Sensitivity</td>
<td>61.88</td>
</tr>
<tr>
<td>Specificity</td>
<td>65.83</td>
</tr>
</tbody>
</table>

The ROC curve test based on the MLC as predictive variable and the MSCC as classification variable with the decision threshold of $750 \times 10^3$ cell/ml, has highlighted that, from the 4,428 milk samples 682 (15.40%) were found positives and 3,746 (84.60%) were negatives. The sensitivity and specificity of test were 69.94% and 68.63% respectively, with the associated criterion of 4.84% (figure 27 and table 17). The tables 17 and 18 show the different outputs.
Figure 27. Sensitivity and specificity of ROC curve, using MLC as predictive variable and MSCC as classification variable. Of 4428 milk samples and using a threshold $750 \times 10^3$ cells/ml, the model found 4.84% as criterion value with a sensitivity of 69.94% and specificity of 68.63%

Table 17. ROC statistic test related to MLC as predictive variable and MSCC as classification variable with decision threshold of $750 \times 10^3$ cell/ml

<table>
<thead>
<tr>
<th>Area under the ROC curve (AUC)</th>
<th>0.766</th>
</tr>
</thead>
<tbody>
<tr>
<td>Standard Error (^a)</td>
<td>0.00967</td>
</tr>
<tr>
<td>95% Confidence interval (^b)</td>
<td>0.754 – 0.779</td>
</tr>
<tr>
<td>Significance level P (Area=0.5)</td>
<td>&lt; 0.0001</td>
</tr>
</tbody>
</table>

\(^a\) DeLong et al., 1988  
\(^b\) Binomial exact

Table 18. MLC representative parameters calculated through the Youden index with MSCC at $750 \times 10^3$ cells/ml

<table>
<thead>
<tr>
<th>Associated criterion</th>
<th>4.84</th>
</tr>
</thead>
<tbody>
<tr>
<td>95% Confidence interval (^a)</td>
<td>4.76 – 4.93</td>
</tr>
</tbody>
</table>

PhD Student: Djangsou Hagassou, Thesis title: Application of Support Vector Machine as support to early prediction of mastitis in Sarda dairy ewes. Animal Sciences, University of Sassari
The ROC curve test based on the Mph as predictive variable and the MSCC as classification variable with the decision threshold of $750 \times 10^3$ cell/ml, highlighted that, from the 4,428 milk samples 682 (15.40%) were found positives and 3,746 (84.60%) were negatives. The sensitivity and specificity of test were 76.54% and 32.97% respectively, with the associated criterion of 6.68 (figure 28 and table 20). The tables 19 and 22 show the different outputs.

The extremely low specificity is shown by the curve close to the diagonal (figure 28 below).

![Figure 28. Sensitivity and specificity of ROC curve, using Mph as predictive variable and MSCC as classification variable. Of 4428 milk samples and using a threshold 750*10^3 cells/ml, the model found 6.68 as criterion value with a sensitivity of 76.54% and specificity of 32.97%.](image-url)
Table 19. ROC statistic test related to MpH as predictive variable and MSCC as classification variable with decision threshold of 750*10³ cell/ml

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Area under the ROC curve (AUC)</td>
<td>0.557</td>
</tr>
<tr>
<td>Standard Error</td>
<td>0.0114</td>
</tr>
<tr>
<td>95% Confidence interval</td>
<td>0.543 – 0.572</td>
</tr>
<tr>
<td>Significance level P (Area=0.5)</td>
<td>&lt; 0.0001</td>
</tr>
</tbody>
</table>

a DeLong et al., 1988
b Binomial exact

Table 20. MpH representative parameters calculated through the Youden index with MSCC at 750*10³ cells/ml

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Associated criterion</td>
<td>6.68</td>
</tr>
<tr>
<td>95% Confidence interval</td>
<td>6.61 – 6.70</td>
</tr>
<tr>
<td>Sensitivity</td>
<td>76.54</td>
</tr>
<tr>
<td>Specificity</td>
<td>32.97</td>
</tr>
</tbody>
</table>

The ROC curve test based on the MEC, MLC and MpH for the prediction of udder inflammation status was validated using a dataset of 2,213 milk samples previously selected. The results of the test validation is shown below in the table 22. As aforementioned, the accuracy of the ROC curve test is assessed using the AUC. The table 22 shows the AUC of the three predictive variables (MEC, MLC and MpH) at MSCC decision threshold of 750*10³ cells/ml. As it can be denoted through the ROC curve outputs (figure 29 and table 21), the MLC performs better than MEC (0.766 vs 0.700 AUC) and MpH (0.766 vs 0.557 AUC), and MEC performs better than MpH (0.700 vs 0.557 AUC).
Table 21. Test validation for ROC curve analysis at MSCC decision threshold of 750*10^3 cells/ml. The MEC was the predictive variable considered.

<table>
<thead>
<tr>
<th>Type of dataset</th>
<th>Data</th>
<th>Criterion</th>
<th>sensitivity</th>
<th>95% CI</th>
<th>Specificity</th>
<th>95% CI</th>
</tr>
</thead>
<tbody>
<tr>
<td>Training set</td>
<td>4,428</td>
<td>4.82 mS/cm</td>
<td>61.88</td>
<td>58.1 – 65.5</td>
<td>65.83</td>
<td>64.3 – 67.3</td>
</tr>
<tr>
<td>Test set</td>
<td>2,213</td>
<td>4.82 mS/cm</td>
<td>62.09</td>
<td>56.7 – 67.3</td>
<td>63.63</td>
<td>61.4 – 65.8</td>
</tr>
</tbody>
</table>

Table 22. AUCs of the three different tests (MEC, MLC and MpH) at MSCC threshold of 750*10^3 cells/ml

<table>
<thead>
<tr>
<th>Predictive variable</th>
<th>AUC</th>
</tr>
</thead>
<tbody>
<tr>
<td>MEC test</td>
<td>0.700</td>
</tr>
<tr>
<td>MLC test</td>
<td>0.766</td>
</tr>
<tr>
<td>MpH test</td>
<td>0.557</td>
</tr>
</tbody>
</table>

Figure 29. The ROC curves of the three different predictive milk variables (MEC, MLC and MpH) at MSCC decision threshold of 750*10^3 cells/ml
3.5.3. ROC curve analysis at MSCC decision threshold of $1000 \times 10^3$ cells/ml
At the MSCC decision threshold of $1000 \times 10^3$ cells/ml, the ROC curves analysis was carried out processing 6,641 primiparous milk samples divided into two datasets: 4,428 milk samples for training set and 2,213 milk samples for test set.

The three ROC curves for the three udder inflammation detection tests (MEC, MLC and MpH) are represented in figures 30, 31 and 32. The different cut-off points (criterions) are denoted by the orange marker on the curve.

Focusing on the MEC as predictive variable and the MSCC as classification variable with the decision threshold of $1000 \times 10^3$ cell/ml, the test highlighted that, from the 4,428 milk samples, 537 (12.13%) were found positives and 3,891 (87.87%) were negatives. The sensitivity and specificity of test were 63.50% and 65.02% respectively, with the associated criterion of 4.82 mS/cm (figure 30). The tables 23 and 24 show the different outputs.
Figure 30. Sensitivity and specificity of ROC curve, using MEC as predictive variable and MSCC as classification variable. Of 4,428 milk samples and using a threshold $10^3$ cells/ml, 537 (12.13%) were positives and 3,891 (87.87%) were negatives. The test has 4.82 mS/cm as criterion value with associated sensitivity and specificity of 63.50% and 65.02% respectively.

Table 23. ROC statistic test related to MEC as predictive variable and MSCC as classification variable with decision threshold of $10^3$ cell/ml

<p>| | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Area under the ROC curve (AUC)</td>
<td>0.705</td>
</tr>
<tr>
<td>Standard Error $^a$</td>
<td>0.0119</td>
</tr>
<tr>
<td>95% Confidence interval $^b$</td>
<td>0.692 – 0.719</td>
</tr>
<tr>
<td>Significance level P (Area=0.5)</td>
<td>&lt; 0.0001</td>
</tr>
</tbody>
</table>

$^a$ DeLong et al., 1988  
$^b$ Binomial exact
Table 24. MEC representative parameters calculated through the Youden index with MSCC at 1000*10^3 cells/ml

<table>
<thead>
<tr>
<th>Associated criterion</th>
<th>4.82</th>
</tr>
</thead>
<tbody>
<tr>
<td>95% Confidence interval a</td>
<td>4.65 – 4.95</td>
</tr>
<tr>
<td>Sensitivity</td>
<td>63.50</td>
</tr>
<tr>
<td>Specificity</td>
<td>65.02</td>
</tr>
</tbody>
</table>

The ROC curve test based on the MLC as predictive variable and the MSCC as classification variable with the decision threshold of 1000*10^3 cell/ml, has highlighted that from the 4,428 milk samples 537 (12.13%) were positives and 3,891 (87.87%) were negatives. The sensitivity and specificity of test were 83.80% and 54.51% respectively, with the associated criterion of 4.3% (figure 31 and table 25). The tables 25 and 26 show the different outputs.

Figure 31. Sensitivity and specificity of ROC curve, using MLC as predictive variable and MSCC as classification variable. Of 4,428 milk samples and using a threshold 1000*10^3 cells/ml, 537 (12.13%) were positives and 3,891 (87.87%) were negatives. The test has 4.93% as criterion value with associated sensitivity and specificity of 83.80% and 54.51% respectively.
Table 25. ROC statistic test related to MLC as predictive variable and MSCC as classification variable with decision threshold of $1000 \times 10^3$ cell/ml

<p>| | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Area under the ROC curve (AUC)</td>
<td>0.766</td>
</tr>
<tr>
<td>Standard Error $^a$</td>
<td>0.0106</td>
</tr>
<tr>
<td>95% Confidence interval $^b$</td>
<td>0.754 – 0.779</td>
</tr>
<tr>
<td>Significance level P (Area=0.5)</td>
<td>&lt; 0.0001</td>
</tr>
</tbody>
</table>

$^a$ DeLong et al., 1988  
$^b$ Binomial exact

Table 26. MLC representative parameters calculated through the Youden index with MSCC at $1000 \times 10^3$ cells/ml

<p>| | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Associated criterion</td>
<td>4.93</td>
</tr>
<tr>
<td>95% Confidence interval $^a$</td>
<td>4.87 – 4.96</td>
</tr>
<tr>
<td>Sensitivity</td>
<td>83.80</td>
</tr>
<tr>
<td>Specificity</td>
<td>54.51</td>
</tr>
</tbody>
</table>

The ROC curve test based on the MpH as predictive variable and the MSCC as classification variable with the decision threshold of $1000 \times 10^3$ cell/ml, has highlighted that from the 4,428 milk samples 537 (12.13%) were positives and 3,891 (87.87%) were negatives. The sensitivity and specificity of test were 75.98% and 32.54% respectively, with the associated criterion of 6.68 (figure 32 and table 28). The tables 27 and 28 show the different outputs.

The extremely poor specificity is shown by the curve close to the diagonal (figure 32 below).
Figure 32. Sensitivity and specificity of ROC curve, using MpH as predictive variable and MSCC as classification variable. Of 4,428 milk samples and using a threshold $1000 \times 10^3$ cells/ml, 537 (12.13%) were positives and 3,891 (87.87%) were negatives. The test has 4.93% as criterion value with associated sensitivity and specificity of 75.98% and 32.54% respectively.

Table 27. ROC statistic test related to MpH as predictive variable and MSCC as classification variable with decision threshold of $1000 \times 10^3$ cell/ml

<p>| | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Area under the ROC curve (AUC)</strong></td>
<td>0.546</td>
</tr>
<tr>
<td><strong>Standard Error</strong>&lt;sup&gt;a&lt;/sup&gt;</td>
<td>0.0127</td>
</tr>
<tr>
<td><strong>95% Confidence interval</strong>&lt;sup&gt;b&lt;/sup&gt;</td>
<td>0.531 – 0.561</td>
</tr>
<tr>
<td><strong>Significance level P (Area=0.5)</strong></td>
<td>0.0003</td>
</tr>
</tbody>
</table>

<sup>a</sup> DeLong et al., 1988

<sup>b</sup> Binomial exact

PhD Student: Djangou Nagassou, Thesis title: Application of Support Vector Machine as support to early prediction of mastitis in Sarda dairy ewes. Animal Sciences, University of Sassari
Table 281. MpH representative parameters calculated through the Youden index with MSCC at 1000*10^3 cells/ml

<table>
<thead>
<tr>
<th>Associated criterion</th>
<th>6.68</th>
</tr>
</thead>
<tbody>
<tr>
<td>95% Confidence interval</td>
<td>6.62 – 6.75</td>
</tr>
<tr>
<td>Sensitivity</td>
<td>75.98</td>
</tr>
<tr>
<td>Specificity</td>
<td>32.54</td>
</tr>
</tbody>
</table>

The ROC curve test based on the MEC, MLC and MpH for the prediction of udder inflammation status was validated using a dataset of 2,213 milk samples previously selected. The results of the test validation is shown below in the table 30. As aforementioned, the accuracy of the ROC curve test is assessed using the AUC. The table 31 shows the AUC of the three predictive variables (MEC, MLC and MpH) at MSCC decision threshold of 1000*10^3 cells/ml. As it can be denoted through the ROC curve outputs (figure 33 and table 30), the MLC performs better than MEC (0.766 vs 0.705 AUC) and MpH (0.766 vs 0.546 AUC), and MEC performs better than MpH (0.705 vs 0.546 AUC).

Table 29. Test validation for ROC curve analysis at MSCC decision threshold of 1000*10^3 cells/ml. The MEC was the predictive variable considered.

<table>
<thead>
<tr>
<th>Type of dataset</th>
<th>Data</th>
<th>Criterion</th>
<th>sensitivity</th>
<th>95% CI</th>
<th>Specificity</th>
<th>95% CI</th>
</tr>
</thead>
<tbody>
<tr>
<td>Training set</td>
<td>4,428</td>
<td>4.82 mS/cm</td>
<td>63.50</td>
<td>59.30 – 67.60</td>
<td>65.02</td>
<td>63.50 – 66.50</td>
</tr>
<tr>
<td>Test set</td>
<td>2,213</td>
<td>4.82 mS/cm</td>
<td>63.67</td>
<td>57.50 – 69.60</td>
<td>62.80</td>
<td>60.60 – 64.90</td>
</tr>
</tbody>
</table>

Table 30. AUCs of the three different tests (MEC, MLC and MpH) at MSCC threshold of 1000*10^3 cells/ml

<table>
<thead>
<tr>
<th>Predictive variable</th>
<th>AUC</th>
</tr>
</thead>
</table>

PhD Student: Djangsou Hagassou, Thesis title: Application of Support Vector Machine as support to early prediction of mastitis in Sarda dairy ewes. Animal Sciences, University of Sassari
### 3.5.4. Specific ROC curves test accuracy comparison at different MSCC decision thresholds

As described above, the AUC represents a summary statistic of the overall predictive performance of the test in detecting the udder inflammation status. Therefore, AUCs could provide interesting measures for the comparison of the overall predictive test performances. The figure 34 below, shows the results of MEC test at three different MSCC decision thresholds (500*10^3, 750*10^3 and 1000*10^3 cells/ml). The sensitivity of the MEC test seems to improve slightly when the MSCC decision threshold increases.

---

**Figure 33.** The ROC curves of the three different predictive milk variables (MEC, MLC and MpH) at MSCC decision threshold of 1000*10^3 cells/ml
(58.50%, 61.88% and 63.50%, at 500*10^3, 750*10^3 and 1000*10^3 cells/ml respectively). Inversely the specificity of MEC test decreases slightly with the increase of MSCC decision threshold (66.72%, 65.83% and 65.02% at 500*10^3, 750*10^3 and 1000*10^3 cells/ml). These test results are in accordance with the principles of ROC curve analysis. The MEC test accuracy at three MSCC decision thresholds (500*10^3, 75*10^3 and 1000*10^3 cells/ml) is highlighted in the table 32 below through the AUC value. It seems that the MEC test performs better with the increase of MSCC decision threshold with AUC value of 0.664, 0.700 and 0.705 at 500*10^3, 750*10^3 and 1000*10^3 cells/ml, respectively.

Similarly to the MEC test (above), also MLC test shows the same behavior: the sensitivity of the test increases when the MSCC decision threshold increases (67.66, 69.94 and 83.80% at 500*10^3, 750*10^3 and 1000*10^3 cells/ml respectively). Inversely, the test specificity decreases with the increase of MSCC decision threshold (70.50, 68.63 and 54.51% at 500*10^3, 750*10^3 and 1000*10^3 cells/ml respectively). Also the MLC criterion changed slightly (from 4.84% to 4.93%, figure 34).

The MLC test accuracy at three MSCC decision thresholds (500*10^3, 750*10^3 and 1000*10^3 cells/ml) is highlighted in the table 33 below, through the AUC value. It seems that the MLC test accuracy improves a little from MSCC decision threshold of 500*10^3 cells/ml to 750*10^3 cells (AUC value = 0.757 and 0.766) and remains the same at 750*10^3 and 1000*10^3 cells/ml (AUC value = 0.766).

Considering the MpH test, its sensitivity follow an irregular trend (61.48, 76.54, 75.98% at 500*10^3, 750*10^3, 1000*10^3 cells/ml, respectively), whereas its specificity has a regular trend as in the previous tests, decreasing when the MSCC decision threshold increases (47.42, 32.97, 32.54% at 5000*10^3, 75*10^3 and 1000*10^3 cells/ml respectively, figure 36). The test criterion value remains almost the same (6.64, 6.68 6.76).
and 6.68) at the MSCC decision thresholds (5000*10^3, 750*10^3 and 1000*10^3 cells/ml).

The MpH test accuracy at different MSCC decision thresholds is really poor and highlighted through the AUC values in the table 33.
**Table 31.** MEC test - AUCs at MSCC decision thresholds of 500*10^3, 750*10^3 and 1000*10^3 cells/ml

<table>
<thead>
<tr>
<th></th>
<th>500*10^3 cells/ml</th>
<th>750*10^3 cells/ml</th>
<th>1000*10^3 cells/ml</th>
</tr>
</thead>
<tbody>
<tr>
<td>AUC</td>
<td>0.664</td>
<td>0.700</td>
<td>0.705</td>
</tr>
</tbody>
</table>

**Figure 34.** Roc curves based on MEC test with the related sensitivity, specificity and criterion at MSCC decision thresholds of 500*10^3 cell/ml (left panel), 750*10^3 cells/ml (panel in the middle) and 1000*10^3 cells/ml (right panel)

PhD Student: Djangou Hagassou, Thesis title: Application of Support Vector Machine as support to early prediction of mastitis in Sarda dairy ewes. Animal Sciences, University of Sassari
Table 2. MLC test - AUCs at MSCC decision thresholds of 500*10^3, 750*10^3 and 1000*10^3 cells/ml

<table>
<thead>
<tr>
<th></th>
<th>500*10^3 cells/ml</th>
<th>750*10^3 cells/ml</th>
<th>1000*10^3 cells/ml</th>
</tr>
</thead>
<tbody>
<tr>
<td>AUC</td>
<td>0.757</td>
<td>0.766</td>
<td>0.766</td>
</tr>
</tbody>
</table>

**Figure 35.** Roc curves based on MLC test with the related sensitivity, specificity and criterion at MSCC decision thresholds of 500*10^3 cell/ml (left panel), 750*10^3 cells/ml (panel in the middle) and 1000*10^3 cells/ml (right panel)

PhD Student: Djangsou Hagassou, Thesis title: Application of Support Vector Machine as support to early prediction of mastitis in Sarda dairy ewes. Animal Sciences, University of Sassari
Table 33. MpH test - AUCs at MSCC decision thresholds of $500 \times 10^3$, $750 \times 10^3$ and $1000 \times 10^3$ cells/ml

<table>
<thead>
<tr>
<th></th>
<th>500$\times$10$^3$ cells/ml</th>
<th>750$\times$10$^3$ cells/ml</th>
<th>1000$\times$10$^3$ cells/ml</th>
</tr>
</thead>
<tbody>
<tr>
<td>AUC</td>
<td>0.559</td>
<td>0.557</td>
<td>0.546</td>
</tr>
</tbody>
</table>

Figure 36. Roc curves based on MpH test with the related sensitivity, specificity and criterion at MSCC decision thresholds of $500 \times 10^3$ cell/ml (left panel), $750 \times 10^3$ cells/ml (panel in the middle) and $1000 \times 10^3$ cells/ml (right panel)
3.6. Application of SVMs as support to udder inflammation detection: case study on Sarda dairy ewes

To develop and verify the model of the SVMs, 6,641 milk samples were collected from the historical laboratory database of the Regional Association of Sardinian Farmers (ARAS). The dataset was composed only by primiparous milk samples and was divided into two datasets: 4,428 milk samples for training set and 2,213 milk samples for test set.

The same dataset was used in the previous section addressing the prediction of udder inflammation status using the ROC curve analysis. To develop the SVMs model, several Kernel functions were applied aiming to find the one which fits better the model. In addition, a number of different pairs of the C parameter and gamma kernel parameter were also applied to the SVMs model with the purpose of finding the best pair which provides the best performance of the model.

After several attempts and using the training dataset, a simplified polynomial kernel was found to be more suitable than the normal polynomial kernel:

\[
k(x, x') = (\text{scale} \cdot < x, x' > + \text{offset})^{\text{degree}}
\]

The hyperparameters of the polynomial equation were assumed as follow:
Scale = 0 (default), offset = 1 and degree = 1.

The output was a linear kernel:

\[
k(x, x') = (1 \cdot < x, x' > + 0)^1 = < x, x' >
\]

The linear kernel function was applied to the model and the coast (C) parameter was assumed as 1. During the phase of model training, 4,428 milk samples from only primiparous Sarda dairy ewes were used. From the whole dataset, the SVMs model has used 446 as support vectors. The figure 37 shows the plot of the training set. Then the SVMs model was tested by using 2,213 primiparous milk samples, previously selected.
Of 2,213 milk samples, 159 were identified as true positives (TP), 1,459 true negatives (TN), 498 false positives (FP) and 97 false negatives (FN). Therefore, the sensitivity and specificity of the test were 62.11% and 74.55% respectively. The figures 37 and 38 represent the plots of training and test sets.
Figure 37. SVMs classification plot of training set. 4428 milk samples from Sarda primiparous ewes (primiparous). The model used 446 vectors as support vectors.

Figure 38. SVMs classification plot, test set. 2213 milk samples from Sarda dairy ewes (primiparous) were used for the test. The sensitivity and specificity were 62.11% and 74.55% respectively.

PhD Student: Djangsou Hagassou, Thesis title: Application of Support Vector Machine as support to early prediction of mastitis in Sarda dairy ewes. Animal Sciences, University of Sassari
3.7. Performance of ROC curve analysis method and SVMs model in predicting udder inflammation status in Sarda dairy ewes.

In general, every point on ROC curve has as coordinates the sensitivity (ordinate) and 1-specificity (axis). Two milk variables are considered, one as predictive variable and the other one as classification variable. In this study, the MSCC was considered as classification variable and three mostly correlated with MSCC variables, such as MEC, MLC and MpH were chosen as predictive variables. The ROC curve analysis highlighted that the MLC performs better than MEC and MpH, and MEC performs better than MpH at all three MSCC decision thresholds (500*10³, 750*10³ and 1000*10³ cells/ml). The accuracy of the ROC curve analysis is determined through the evaluation of the AUC. The figures 24, 28 and 32 show the results of ROC analysis of the three milk variables at three MSCC decision thresholds and highlighted the MpH has the poorest accuracy in predicting udder inflammation in Sarda dairy ewes. The overlapping of the MLC curve on MEC curve can lead to consider only the MLC for the udder inflammation prediction with consequent loose of information which can be useful (figure 39 below).
Figure 39. comparison between MEC and MLC curves at 1000*10^3 cells/ml. A sample size of 2,213 milk samples, of which 256 (11.57%) were positives and 1,957 (88.43%) were negatives. The AUCs of MEC and MLC were 0.688 and 0.777 respectively with difference between the two areas corresponding to 0.0889 and P-value < 0.0001.

Therefore it is important to find a system which can combine the two milk variables highly associated with MSCC to avoid the loose of information. Therefore, the SVMs model was applied in this trial aiming to overcome the issue of information loosing and improving the prediction of udder inflammation status in Sarda dairy ewes. Conversely to the ROC curve analysis method, where every point on the curve represents the combination of sensitivity and 1-specificity related to one predictive variable and one classification variable, the SVMs model allows to use simultaneously several predictive variables associated with a classification variable. In the current study, the MEC and MLC were used simultaneously as predictive variables in combination with MSCC as classification variable. In addition, every strip

PhD Student: Djangou Hagassou, Thesis title: Application of Support Vector Machine as support to early prediction of mastitis in Sarda dairy ewes. Animal Sciences, University of Sassari
(parallelogram) in SVMs plotting system (figure 30) represents a density of probability with the same value.

![SVM classification plot](image)

**Figure 40.** SVMs classification plot of 2213 milk samples from Sarda dairy ewes. The discriminability index was converted in percentage of sensibility (blue color) and specificity (light-brown color). The red strip represents the density of probability at which the sensitivity and specificity are equivalent to 62% and 75% respectively.

The figure 41 (below) highlighted some cases in which the combination of MEC, MLC and MSCC can provide more useful information and improve the prediction of udder inflammation:

**Case 1** - represents the condition in which milk has high MLC (5%) and low MEC (4.4 mS/cm). In this situation the animal could be considered in physiologically normal condition considering the high biosynthesis of lactose at udder level. The lowest MEC

PhD Student: Djagsou Hagassou, Thesis title: Application of Support Vector Machine as support to early prediction of mastitis in Sarda dairy ewes. Animal Sciences, University of Sassari
value could be considered as the expression of inexistence tissue damage or other factors affecting MEC; or no relevant udder disturbance. This is a desirable condition in which the animal could be considered in healthy condition.

**Case 2** – high MLC (5%) and high MEC (5.8 mS/ml). The higher value of MLC could be considered as the effect of the normal biosynthesis activity of mammary gland (as explained above). Therefore the higher value of MEC could be interpreted as the beginning of udder inflammation or existence of some MEC influencing factors (such as stage of lactation, parity, age, milking system, feeding system, etc.). In this case further monitoring on udder inflammation is needed.

**Case 3** – low MLC and low MEC. The biosynthesis of lactose is affected but the MEC value seems to be norm. This discrepancy could be explained through error committed during data recording or the animal has chronic disturbance at mammary gland. More monitoring of udder health status is needed and the milk sampling should be repeated.

**Case 4** – low MLC and high MEC. Undesirable situation, the biosynthesis of lactose is affected and the high value of MEC could be explained through the probably damage of udder tissue and consequent udder inflammation occurrence. In this case a prompt and detailed diagnostic analysis (bacteriological analysis) is needed. This situation can be interpreted as the result of the deterioration of the membranes of epithelial cells in the mammary gland which consequently caused the passage of lactose from the alveolar lumen into the blood and conversely it allows the passage of blood elements such as chlorides, from the blood into the alveolar lumen. This abnormal osmotic phenomenon can lead to the decrease of the MLC and the increase of MEC.
Figure 41. SVMs classification plot interpretation. Case 1 - high MLC and low MEC; case 2 – high MLC and high MEC; case 3 – low MLC and low MEC.
4. Conclusion

The current study highlighted that the milk first squirt sample is highly representative of the milk sample from complete milking in term of milk somatic cell count (MSCC). This is demonstrated by the high Pearson correlation coefficient ($r = 0.973$) between the MSCC obtained from milk first squirt samples (MSCC.fs) and the MSCC from milk samples obtained from complete milking (table 5). Using the milk first squirt samples has the advantage to avoid the huge waste of milk normally collected for laboratory analysis (20 ml for milk first squirts against 50 ml for milk complete milking).

The highly significant Pearson correlation between the MEC measured on the field (farms) using a hand-held electrical conductivity meter, named *Masti-Milk* and the MEC measured in laboratory by the means of a proven electrical conductivity meter, WTW LF 92 ($r = 0.954$), leads to considered that the Masti-Milk is highly reliable and can be used in dairy ewe farm for the MEC measurement. The table 6 shows the Pearson’s r between MEC measured by the means of the Masti-Milk and the WTWLF92.

Furthermore, important relationships between several milk variables considered in this survey emerged in the table 6 through the Pearson correlation coefficient. Most of them are related to the MSCC as following: MEC and MSCC ($r = 0.442$), MLC and MSCC ($r = -0.579$), MCC and MSCC ($r = 0.572$) and; MLC and MCC ($r = -0.942$), MLC and MEC ($r = -0.719$), MCC and MEC ($r = 0.868$).

The results of the ROC analysis using the MLC, MEC and MpH as predictive variables for udder inflammation and MSCC as classification variable at three different MSCC decision threshold ($500*10^{3}$, $750*10^{3}$ and $1000*10^{3}$ cells/ml), have highlighted that the MLC performs better than MEC and MpH, and the MEC performs better than MpH in term of sensitivity, specificity and accuracy test represented by the Area Under the ROC
Curves (AUC). The figures 22 to 36 and tables 7 to 33 show the results of the ROC analysis.

The best MEC test through the ROC curve analysis was found at the MSCC decision threshold of $1000 \times 10^3$ cells/ml with the sensitivity of 63.50%, the specificity of 65.02%, associated criterion of 4.82 mS/cm and the AUC of 0.705. That means from all dairy ewes with MSCC truly above $1000 \times 10^3$ cell/ml, 63.50% was truly classified using MEC as predictive variable and the remaining 36.50% was classified as False negative by the ROC curve test.

In addition 65.02% of the dairy ewes with the MSCC below $1000 \times 10^3$ cells/ml was truly classified using MEC as the predictive variable and the remaining 34.98% was classified as false positive.

The best MLC test through the ROC curve analysis was found at MSCC decision threshold of $750 \times 10^3$ cells/ml with the sensitivity of 69.94%, specificity of 68.63%, associated criterion of 4.84% and AUC of 0.766. In this case, 69.94% of the dairy ewes with the MSCC above $750 \times 10^3$ cells/ml was truly classified and the remaining 30.06% was classified as false negative. Moreover the 68.63% of the dairy ewes with the MSCC below $750 \times 10^3$ cells/cm were truly classified and the remaining 31.37% was classified as false positive.

The MpH as predictive variable was the poorest test in separating the dairy ewes at the three different MSCC (figure 36 and table 33).

The SVMs application at the MSCC decision threshold $1000 \times 10^3$ has shown an interesting potential through the combination of the MLC and MEC as predictive variables with the MSCC as classification variable. It has found the good compromise between sensitivity and specificity at 62% and 75% respectively. Conversely to the ROC curve analysis where each point on the curve represents the combination between

PhD Student: Djangou Hagassou, Thesis title: Application of Support Vector Machine as support to early prediction of mastitis in Sarda dairy ewes. Animal Sciences, University of Sassari
sensitivity and specificity, the SVMs plots the sensitivity and specificity as density of probability (figure 41). Every strip represents a density of probability with the same sensitivity and specificity. Therefore it could improve the prediction system for udder inflammation in dairy ewes.

The ROC curve analysis has shown that using the MEC alone as a predictive variable at the MSCC decision threshold of $500*10^3$, $750*10^3$ and $1000*10^3$ cells/ml (figure 33) can lead to a risky decision for udder inflammation detection because of high number of false negatives (41.50, 38.22 and 36.50% at the MSCC threshold of $500*10^3$, $750*10^3$ and $1000*10^3$ cells/ml respectively). This is in accordance with a number of studies done on MEC as indicator for udder inflammation detection (Larsgard A.G. and Vaabenoe A., 1993; Bergonier D. and Berthelot X., 2003; Mammadova N. and Keskin I., 2013). Therefore the daily measurement of the MEC through the installation of MEC sensors in milking units could be of great importance to record day by day the MEC and constitute an historical time-series data which can highlighted the anomalous deviation of MEC from the normal trend.

The ROC curve analysis has pointed out that at the MSCC decision threshold of $500*10^3$, $750*10^3$ and $1000^3$ cells/ml, the MLC curve hidden the MEC and MpH curves (figures 25, 29 and 33). These results could lead to choose the MLC as the most accurate predictive parameter in predicting udder inflammation by loosing information from the other two parameters (MEC and MpH).

To overcome this issue, the SVMs model has been used by combining the two most accurate milk variables for udder inflammation prediction (MLC and MEC) in association with the MSCC as classifier (figure 40). The linear SVMs model has shown better udder inflammation prediction by combining information from both MLC and MEC.
Considering the results from both ROC curve analysis and SVMs, it could be suggested to the farmers to use MEC and MLC as the integrated system for first screening system for udder inflammation detection. This screening method can reduce the number of dairy ewes with suspicious udder inflammation on which the appropriate but expansive and time-consuming laboratory analysis such as bacteriological analysis is needed. However, further and practical studies should be carried out on the individual udder-half aiming to investigate on the normal and abnormal individual udder-half behavior and comparing the milk parameter between normal and abnormal half-udder as it has been done on dairy cow, similarly the udder inflammation prediction method of inter quarter ratio (Norberg et al., 2004; Diaz et al., 2011).
5. Acknowledgements
I am thankful to my tutor Professor Antonio Pazzona for his helpfulness and readiness in giving technical advice to conduct the research and survey. I would like to thank the director of ARAS with his staff for providing laboratory facilities to conduct the in-farm milk sampling, laboratory milk sample analysis and data storage. I thank engineer Giovanni Chessa for his contribution in data processing and analysis. The helpfulness and patience of all farmers during milk sampling are gratefully acknowledged and appreciated.

Thank you Firida, I finally understood that you are the greatest person I never ever met, and I feel really lucky for that. I want to say thanks to my godmother Zinzula Maria Rosaria, I appreciated everything you have done for me and I will never say enough thanks.

I want to say thanks to my lovely girlfriend, Jasmine Luche, her parents and her lovely grandparents. Baby you are the most wonderful flower that made my life more colorful.
6. References


PhD Student: Djangsou Hagassou, Thesis title: Application of Support Vector Machine as support to early prediction of mastitis in Sarda dairy ewes. Animal Sciences, University of Sassari


Subtil F. and M. Rabilloud. 2015. An enhancement of ROC curves made them clinically relevant for diagnostic-test comparison and optimal-threshold determination. 68:752-759.


