Quest Journals
Journal of Research in Agriculture and Animal Science
Volume 10 ~ Issue 5 (2023) pp: 14-28
ISSN(Online) : 2321-9459
www.questjournals.org

Research Paper

Temporal and spatial classification of analytical data of bovine milk, using Self-Organizing Maps

Flávia Queiróz Ferreira[1], Maria Isabel Cristina Batista Mayrink[2], Rita Aparecida Dutra[1], Clausius Duque Gonçalves Reis[1], Paulo Henrique Fidêncio[3], César Reis[1], Efraim Lázaro Reis[1]

[1]*Departamento de Química, Universidade Federal de Viçosa, Viçosa, Brasil*
[2]*Faculdade Dinâmica do Vale do Piranga, Ponte Nova, Brasil*
[3]*Departamento de Química, Universidade Federal dos Vales do Jequitinhonha e Mucuri, Diamantina, Brasil*

ABSTRACT: *The physical-chemical characteristics of milk are of fundamental importance and directly influence the quality of human life. The objective of this work is to map these physical-chemical characteristics of milk, correlating them with geographical locations and collection dates. The tool used was the package called Self-Organizing Maps of Kohonen, which operates within the MatLab environment. This tool allows the creation of a two or three-dimensional arrangement of a certain data set, in order to facilitate the evaluation and visualization of results. The main advantage of this tool is the speed at which the map is generated, even when there is a large number of samples and/or variables.*

KEYWORDS: *Milk, Neural network, Kohonen maps, Self-Organizing Maps*

R*eceived 14 May, 2023;* ***R****evised 24May, 2023;* ***A****ccepted 26 May, 2023 © The author(s) 2023.*
P*ublished with open access at* ***www.questjournals.org***

I. INTRODUCTION

Food quality has been increasingly discussed in the current scenario and plays a fundamental role in human health and quality of life. Milk is a food of great importance in human nutrition because it has high nutritional value. It is rich in carbohydrates, proteins, lipids, various vitamins and minerals and is a source of energy [1-3].

There are several factors that can interfere with milk quality, from animal health problems to product contamination during milking and processing. These factors can cause great economic damage to dairy production, in addition to reducing productivity and the added value of the product, which can interfere with the industrial process of dairy production [4-5].

The physicochemical characteristics of milk are of fundamental importance and have a direct influence on its price. Milk with a higher fat content, for example, has greater added value, as it allows the production of products with better organoleptic characteristics, such as color, aroma and flavor. Fatter milk results in higher yields in cheese and butter, with better texture in the product [6-7].

The present work aims to map the physicochemical characteristics of milk, correlating them to their geographic locations and collection dates. The tool used was Kohonen's Self-Organizing Maps, a computational package that operates within the MatLab environment [8]. Kohonen's Self-Organizing Maps allow the elaboration of an arrangement in one or two dimensions of a given set of data, in order to facilitate the evaluation and visualization of the results [9]. Its advantage lies in the speed with which the map is generated, even when there is a large number of samples and/or variables. Self-Organizing Maps (SOM) is a type of neural network inspired by the neurons of the human cerebral cortex, which is characterized by unsupervised learning [10-14].

Unsupervised learning is based on the principle that the computational algorithm is capable of identifying the classes within the data set by itself, without requiring information from the user. This means that networks are able to learn from training samples, generalize from acquired knowledge and adapt to a new dataset.

*Corresponding Author: Efraim Lázaro Reis (efraimreis@gmail.com)

The entered data is first normalized so that they are all on the same scale, as the algorithm uses the Euclidean distance between the data vectors.

Based on competitive learning, output neurons compete with each other and adjust to a neuron that will be activated after interactions. The activated neuron is called the winning neuron and is the one that indicates the cluster to which the input vector belongs. It has more weight and the neighborhood fits around it (Figure 1).

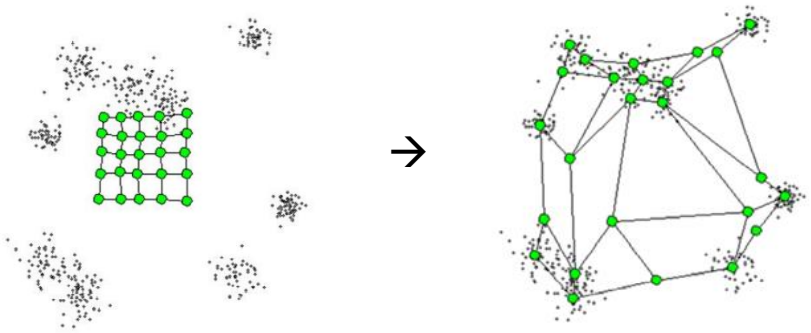

Figure 1: Model fit (green) to experimental data (black).
Source: Bioinf Gif [15], adapted.

The algorithm responsible for the formation of the map consists of four steps, the first being the initialization of the map, followed by three essential processes that are the competitive process, the cooperative process and the synaptic adaptation. A brief description of them can be seen below:

- ✓ Competition: for each input pattern, the grid neurons calculate their respective values of a discriminant function. This discriminant function provides the basis for competition between neurons. The particular neuron with the highest discriminant function value is declared the winner of the competition.
- ✓ Cooperation: the winning neuron determines the spatial location of a topological neighborhood of excited neurons, thus providing the basis for cooperation between neighboring neurons.
- ✓ Synaptic adaptation: allows excited neurons to increase their individual values of the discriminant function in relation to the input pattern through appropriate adjustments applied to their synaptic weights. This process can be decomposed into two phases, namely, a sorting phase followed by a convergence phase. The idea is to carry out a topological ordering of the weight vectors in the first phase and in the second phase to fine-tune the feature map, thus producing an accurate statistical quantization of the input space.

During the iterative training process and due to the neighborhood update, the synaptic weight vectors tend to follow the distribution of the input vectors, leading to a topological ordering of the map. Thus, adjacent neurons will tend to have vectors of similar synaptic weights, thus occurring clusters.

II. MATERIAL AND METHODS

2.1. Obtaining the raw material

Accompanying the vehicle that collects milk in a single tank, samples of raw milk were collected from 17 rural properties, in the city of Bom Jardim de Minas - MG, once a week, for 5 weeks, between 03/27/2020 and 04/30/2020. The samples were stored in sterile 50 mL collectors, placed in a container with reusable ice (non-toxic gel) and duly identified. In the first collection, a Garmini GPS, nüvi 30 model, was used, and the geographic coordinates of each collection point and the altitude were recorded.

The locations of sample collections and the dairy where the samples were analyzed can be seen in table 1, the map in figure 2 shows the route taken during sample collection.

Table 1: Location of sample collection and the dairy where the samples were analyzed.

Producer	Location	Altitude (m)
FRK	21º56'47,0"S 44º08'30.7"W	1178
BAT	21º56'14,0"S 44º07'10.9"W	1198
SER	21°57'02.9"S 44°05'10.4"W	956
FSM	21°56'09.4"S 44°05'30.7"W	956
ZEA	21°57'05.9"S 44°05'07.2"W	855
ZEC	21°57'16.0"S 44°05'02.2"W	912
MAU	21°57'16.0"S 44°04'02.2"W	912
LEO	21°58'31.0"S 44°03'39.1"W	1024
SEB	21°56'09.4"S 44°03'30.7"W	1924
ZEM	21°57'03.4"S 44°02'09.8"W	863
DOM	21°56'58.5"S 44°02'08.6"W	866
BEM	21°56'01.5"S 44°02'11.2"W	952
RAI	21°56'36.0"S 44°01'46.0"W	869
TON	21°56'09.4"S 44°05'30.7"W	872
VIC	21°54'40.8"S 43°59'21.7"W	874
ERP	21°54'40.9"S 43°59'21.8"W	874
RDP	21°56'09.4"S 44°05'30.7"W	956

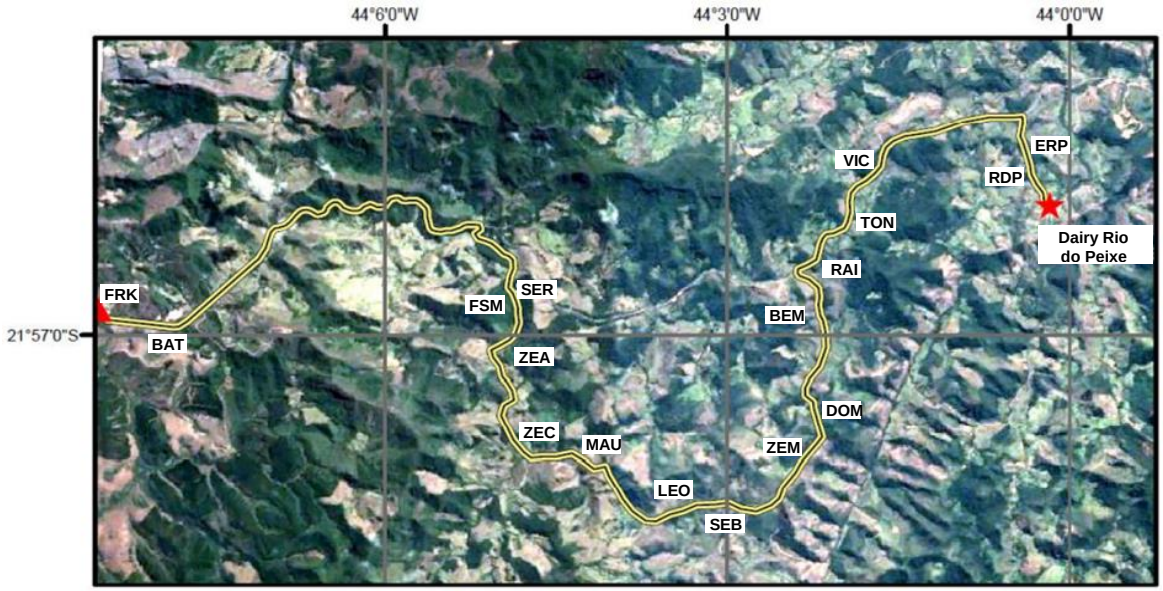

Figure 2: Map of the sample collection region.
Source: Google Earth ASTRIUM 2016. Scale 1:125.000

2.2. Physicochemical analysis

The samples were sent to the physical-chemical analysis laboratory at Dairy Rio do Peixe, located at Sal Derramado Farm, in the rural area of Bom Jardim de Minas – MG, Brazil. They were analyzed with the Milk Analyzer MASTER equipment, SN: 12582 Mode 1. The device has the capacity to perform 60 samples per hour, using a small amount of sample, on average 5 mL [16].

The analyzes carried out were fat content, non-fat solids (NFS), density, proteins, lactose, mineral salts, added water content and freezing point. The values for proteins content, for lactose and for mineral salts were estimated by the device itself from the measured value for non-greasy solids present in the sample [17].

*Corresponding Author: Efraim Lázaro Reis (efraimreis@gmail.com)

The added water content was determined by the cryoscopic index, which is a measure of the freezing point or depression of the freezing point of milk in relation to that of water. It is a test used to detect fraud by adding water to milk. The freezing temperature of milk is lower than that of water due to the effect of substances dissolved in it. The addition of water alters the cryoscopic index, bringing it closer to the freezing point of water. The freezing point of milk is practically constant, although the concentration of soluble constituents can vary substantially [18].

The values expected by the legislation for the aforementioned parameters are shown in Table 2. They may, therefore, present small variations according to the lactation period, the season of the year, the climate, the diet, the breed and diseases of the animals [2, 19-21].

Table 2: Composition of bovine milk.

Fat	Density	Proteins	Lactose	Mineral Sats	Freezing Point
3.90 %	1.032 g mL^{-1}	3.40 %	4.80 %	0.80 %	-0.545 °C

2.3. Data processing

All data obtained were entered in the specified format exemplified in figure 4. (Figure 4), together with the names of the variables analyzed and the producers.

```
8
#n      Fat     NFS     Dens    Prot    Lac     Salts   Water   Freezing
        4.15    8.31    1.029   3.04    4.56    0.68    2.85    -0.530  FRK
        4.00    8.01    1.028   2.93    4.40    0.65    6.64    -0.509  BAT
        5.40    8.41    1.028   3.08    4.62    0.69    0.15    -0.546  FSM
        4.43    8.25    1.028   3.02    4.53    0.67    3.18    -0.529  SER
        4.95    8.16    1.027   2.98    4.48    0.67    3.82    -0.525  RDP
        5.46    8.37    1.028   3.06    4.60    0.68    0.61    -0.543  ZEA
        5.38    8.35    1.028   3.05    4.59    0.68    0.94    -0.554  ZEC
        4.97    8.42    1.028   3.08    4.63    0.69    0.49    -0.544  MAU
        4.91    8.46    1.029   3.09    4.64    0.69    0.12    -0.546  LEO
        4.39    8.15    1.028   2.98    4.47    0.67    4.49    -0.521  SEB
        5.17    8.42    1.028   3.08    4.63    0.69    0.27    -0.545  ZEM
        5.04    8.42    1.028   3.08    4.62    0.69    0.47    -0.544  DOM
        4.60    8.46    1.029   3.09    4.64    0.69    0.47    -0.544  BEM
        4.30    8.01    1.027   2.93    4.40    0.65    6.29    -0.511  RAI
        4.86    8.44    1.029   3.09    4.63    0.69    0.47    -0.544  VIC
        3.86    8.55    1.030   3.13    4.69    0.70    0.21    -0.545  ERP
        4.39    8.34    1.029   3.05    4.58    0.68    2.12    -0.535  TON
```

Figure 3: Example of entering data in text format in the SOM program. The variables were coded as Fat, NFS: Non-Fat Solids, Dens: Density, Prot: Proteins, Lac: Lactose, Salts: Mineral Salts, Add Water: Water, Freezing: Freezing Point.

The SOM Toolbox uses files in ASCII format, specific to the program [11, 22]. Thus, the files were assembled in the Notepad program, as in figure 3.

The first line of the file, just above and to the left, informs the program how many variables will be inserted in the study, in this case there were eight variables. On the second line and preceded by **#n**, the names (or codes) of the eight variables appear. In the last column are the labels, which are the identification of each sample, in this case, the acronym of each producer. It is important to note that the file does not contain commas, only points.

After formatting the files, they were treated in the MATLAB program with the help of the SOM_Toolbox tool, making it possible to create Kohonen's Self-Organizing maps.

With the MATLAB program open, the following commands were entered:

>> som_gui

With this command it is possible to load ASCII files from a specific directory. The following windows are opened in the order shown in figure 4.

*Corresponding Author: Efraim Lázaro Reis (efraimreis@gmail.com)

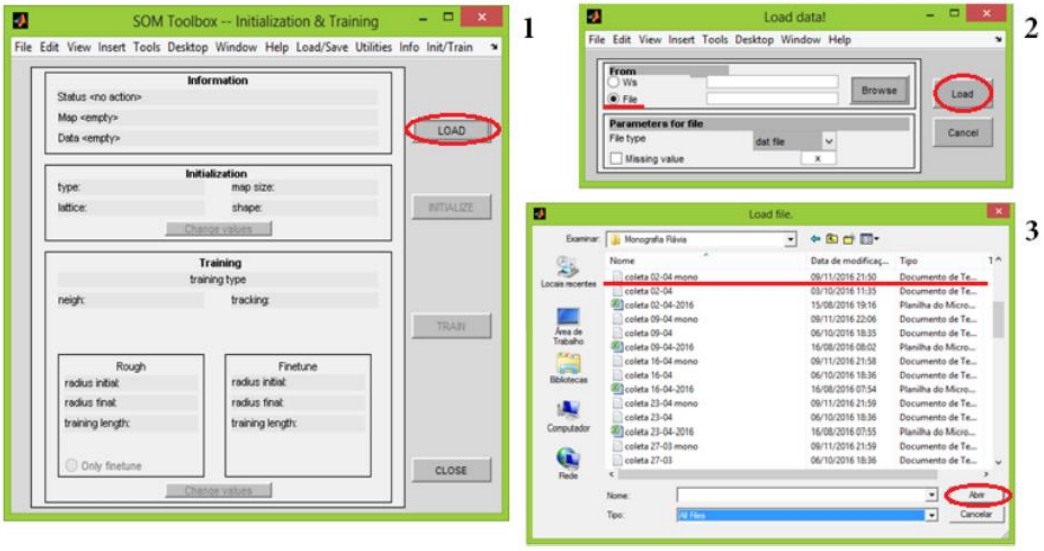

Figure 4: Data insertion scheme in MATLAB.

>> sD=ans

The **som_gui** command loads the file and names it as **ans**, so before moving on to the command in the next step, the **ans** file is renamed as **sD**.

>> sD=som_normalize(sD,'var')

The data must be normalized before processing, as you have values on very different scales in the dataset. The normalization of the variables is of special importance, since the SOM algorithm uses Euclidean metrics to measure distances between vectors. If one variable has values in the range of [0,...,1000] and another in the range of [0,...,1] the first will almost completely dominate the organization of the map due to its greater impact on measured distances. The standard way to get all variables to be equally important is to linearly scale all of them so that their variances equal one.

Before proceeding with the elaboration of the map, it is important to define the size of the map. There are two tools for that, the Mean Quantization Error, which represents the average of the distances between each input data vector and the corresponding weight vector of the winning neuron, and the Topographic Error, which quantifies the map's ability to represent the topology of the Input data. For all maps produced, the application of these two tools was taken into account.

>> sM = som_make(sD, 'msize', [6 6], 'lattice', 'hexa')

For the creation of the map and its dimensioning, there is this function that creates, initializes and trains the map. The approximation algorithm, which uses information and intuition about the problem instance and its structure to solve it quickly, is used to determine the number of neurons in the map. The words “lattice” for lattice and “hexa” for hexagonal refer to the visual shape to be adopted for each sample on the map.

>> sM=som_autolabel(sM,sD,'add')

This command adds the labels to the created map.

>> som_show(sM,'umat','all','comp',1:8,'empty','Labels', 'norm', 'd')

Displays the U matrix and component plane. It is important to highlight that in the commands in which the maps are displayed, new windows are opened and must remain so, to guarantee the perfect functioning of all the commands that follow.

>> som_show_add('label',sM,'subplot',10)

Displays the maps with their respective labels.

>> som_show(sM,'umat','all','empty','Samples')

Displays sample map without labels.

>> som_show_add('label',sM,'subplot',2)

Displays map of samples with labels.

The commands presented above were executed for all samples, obtaining the results shown in figures 6 to 10 and 13; for the results shown in figure 12, 4x4 dimensions were used, respectively, which objectively illustrate the results of this work.

III. RESULTS AND DISCUSSION

Table 3 shows the averages of the determined values of the variable's density, freezing point, percentages respectively of fat, non-fat solids, proteins, lactose, mineral salts and added water for all 17 producers in the 6 days of collection.

Table 3: Mean of the variables Fat (%), Non-Fat Solids - NFS (%), Density (g mL^{-1}), Proteins (%), Lactose (%), Mineral Salts (%), Added water (%), Freezing Point (ºC), of each sampling day for all 17 producers.

Days	Variable							
	Fat %	NFS %	Density g/mL	Proteins %	Lactose %	Salts %	Water %	Freezing °C
27/março	4.65±0.84	8.33±9.21	1.029±0.002	3.05±0.08	4.57±0.12	0.68±0.02	2.34±1.71	-0.534±0.011
02/abril	4.69±0.57	8.24±0.25	1.028±0.002	3.02±0.09	4.52±0.14	0.67±0.02	3.32±2.31	-0.528±0.013
09/abril	4.61±0.57	8.27±0.22	1.028±0.001	3.03±0.08	4.54±0.12	0.67±0.02	2.91±2.12	-0.531±0.015
16/abril	4.78±0.37	8.27±0.23	1.028±0.001	3.03±0.08	4.54±0.13	0.67±0.02	2.94±2.45	-0.532±0.016
23/abril	4.27±0.40	8.37±0.22	1.030±0.002	3.06±0.08	4.59±0.12	0.68±0.02	2.67±1.92	-0.532±0.011
30/abril	4.68±0.49	8.33±0.16	1.028±0.001	3.04±0.06	4.57±0.09	0.68±0.01	1.98±2.19	-0.536±0.013
Average	4.61±0.18	8.30±0.04	1.028±0.001	3.04±0.02	4.56±0.02	0.68±0.00	2.69±0.48	-0.532±0.003

As shown in Table 3, the average values do not differ over the six days of collections, only the average values of water percentage on April 2nd and April 30th showed discrepancies in relation to the other collections.

Table 4: Mean of the variables Fat (%), Non-Fat Solids - NFS (%), Density (g mL^{-1}), Proteins (%), Lactose (%), Mineral Salts (%), Added water (%), Freezing Point (ºC), from each producer for all 6 days.

Average	Variable							
	Fat %	NFS %	Density g/mL	Proteins %	Lactose %	Salts %	Water %	Freezing °C
FRK	4,52±0,36	8,32±0,10	1,029±0,001	3,04±0,03	4,56±0,06	0,68±0,01	2,36±1,23	-0,533±0,007
BAT	4,75±0,44	8,33±0,19	1,028±0,001	3,05±0,07	4,57±0,10	0,68±0,02	2,09±2,47	-0,536±0,016
FSM	4,05±1,15	8,43±0,19	1,029±0,002	3,09±0,07	4,63±0,10	0,69±0,02	1,56±1,22	-0,539±0,009
SER	4,53±0,34	8,22±0,37	1,028±0,002	3,01±0,14	4,51±0,20	0,67±0,03	4,28±2,80	-0,527±0,025
RDP	4,66±0,42	8,10±0,16	1,027±0,001	2,96±0,06	4,44±0,09	0,66±0,01	4,90±2,04	-0,519±0,012
ZEA	4,87±0,45	8,26±0,13	1,028±0,000	3,02±0,04	4,54±0,07	0,67±0,01	2,64±1,84	-0,532±0,011
ZEC	4.87±2,34	8,62±0,24	1,032±0,004	3,16±0,09	4,73±0,13	0,71±0,02	1,25±1,37	-0,542±0,010
MAU	4,81±0,29	8,22±0,18	1,028±0,001	3,01±0,07	4,51±0,10	0,67±0,02	3,20±2,19	-0,529±0,012
LEO	5,03±0,44	8,21±0,14	1,028±0,001	3,00±0,05	4,51±0,07	0,67±0,01	3,09±1,87	-0,529±0,011
SEB	4,94±0,35	8,38±0,20	1,028±0,001	3,06±0,07	4,60±0,11	0,68±0,02	1,68±1,73	-0,540±0,015
ZEM	3.80±0,45	8,33±0,27	1,029±0,003	3,05±0,10	4,57±0,15	0,68±0,03	2,94±2,47	-0,530±0,014
DOM	4,84±0,33	8,39±0,07	1,028±0,000	3,07±0,02	4,61±0,04	0,69±0,01	1,06±0,87	-0,541±0,005
BEM	4,78±0,31	8,40±0,09	1,029±0,001	3,07±0,03	4,61±0,05	0,69±0,01	1,07±0,94	-0,541±0,006
RAI	4,61±0,30	8,25±0,12	1,028±0,000	3,02±0,04	4,53±0,06	0,67±0,01	3,05±1,62	-0,529±0,009
VIC	4,15±0,73	8,23±0,17	1,028±0,001	3,01±0,07	4,52±0,10	0,67±0,02	3,60±2,26	-0,526±0,013
ERP	4,08±0,68	8,23±0,29	1,029±0,002	3,01±0,10	4,52±0,16	0,67±0,03	3,88±3,20	-0,525±0,018
TON	4,69±0,73	8,24±0,20	1,028±0,001	3,01±0,07	4,52±0,11	0,67±0,02	3,14±1,71	-0,529±0,010
Average	4,61±0,18	8,30±0,04	1,028±0,001	3,04±0,02	4,56±0,02	0,68±0,00	2,69±0,48	-0,532±0,003

Table 4 shows the averages of the determined values of the variables density, freezing point, percentages respectively of fat, non-fat solids, proteins, lactose, mineral salts, added water content, of each

producer in the six days of collection. Only the average values of fat percentage of the ZEC producer showed greater discrepancies in the deviation.

In this table it is observed that for the variable percentage of added water there was a variation for all producers.

In figures 5 to 12, it is possible to compare the variable maps with the sample maps, bearing in mind that the number of neurons is always equal to 36 (6x6) in both types of maps, the disposition of the variable values follows the disposition of the samples. In the maps of the variables, their real values are placed in a color scale for better visualization.

Only the maps in Figure 11 differ in terms of the number of neurons, which was equal to 16, in a layout on the maps in 4x4 format.

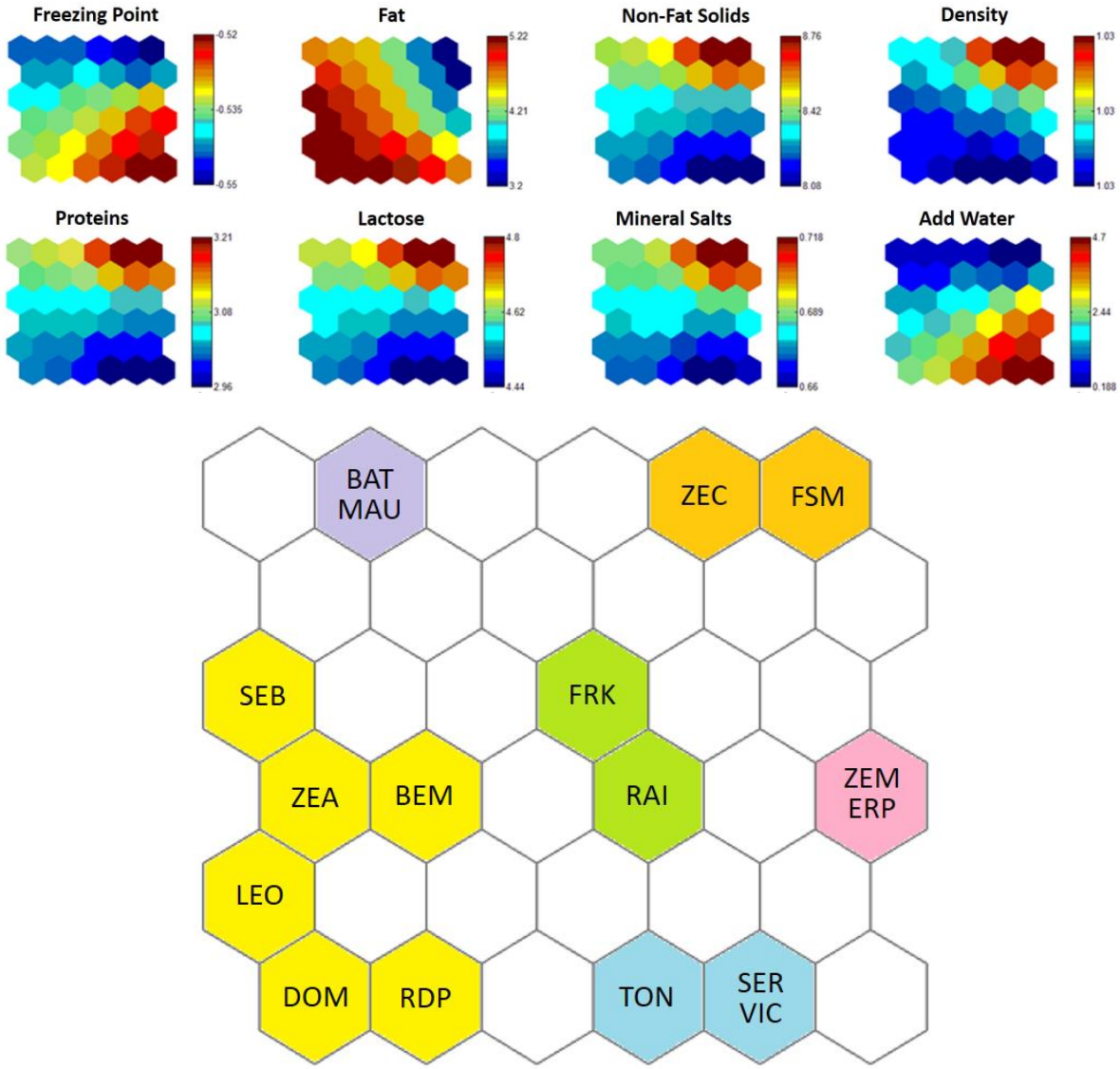

Figure 5: Results of the first collection on March 27th, for the 17 producers.

Thus, for figure 5, for example, the variables non-fat solids, density, proteins, lactose and mineral salts, were the ones that most contributed to the location of the ZEC and FSM samples in the May 27 collection, as they present the highest values in the same region of the map for these samples. On the other hand, the fat variable showed close values for a larger set of samples formed by SEB, ZEA, BEM, LEO, DOM, RDP, TON, SER and VIC. The variables freezing point and added water content were the ones that had the greatest influence on the positioning of the SER, TON and VIC samples.

As can be seen in all the figures, the smallest values in the maps indicate the smallest dissimilarities between the variables and the samples and are represented by cold colors (blue). On the other hand, the highest values indicate the greatest dissimilarities and are represented by warm colors (red). Therefore, clusters are

typically visualized as homogeneous regions with shades close to red. Separations between clusters are marked by more linear regions of colors closer to blue.

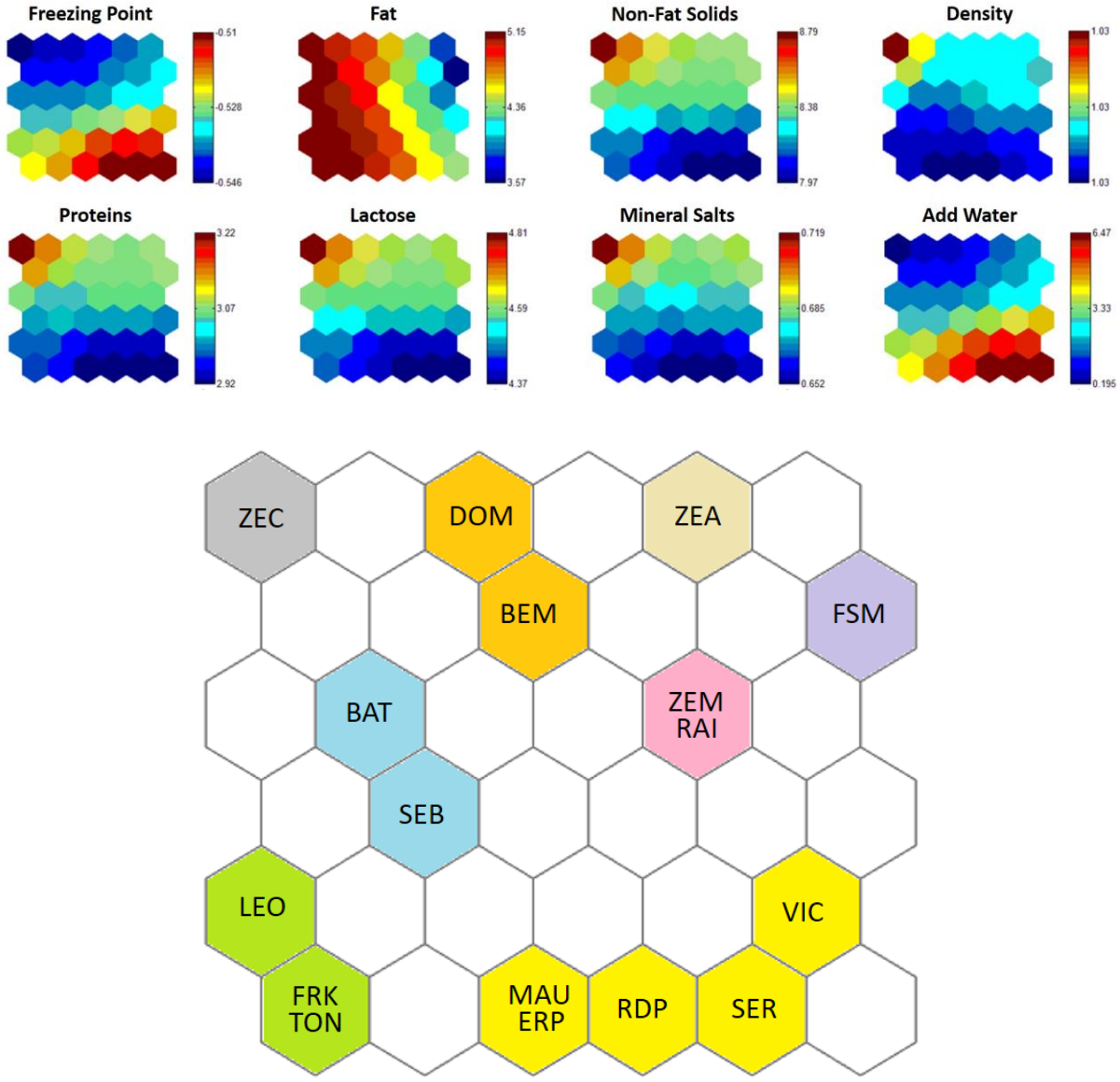

Figure 6: Results of the second collection on April 2, for the 17 producers.

In figure 6, the variables fat, non-fat solids (NFS), density, proteins, lactose and mineral salts were the ones that most contributed to the location of the ZEC sample in the April 2nd collection. Again, the fat variable presents a more uniform value for a larger group of ZEC, BAT, SEB, LEO, FRK, TON, MAU, DOM, BEM and ERP samples. The variables freezing point and added water content were the most influential in the RDP, SER and VIC samples.

*Corresponding Author: Efraim Lázaro Reis (efraimreis@gmail.com)

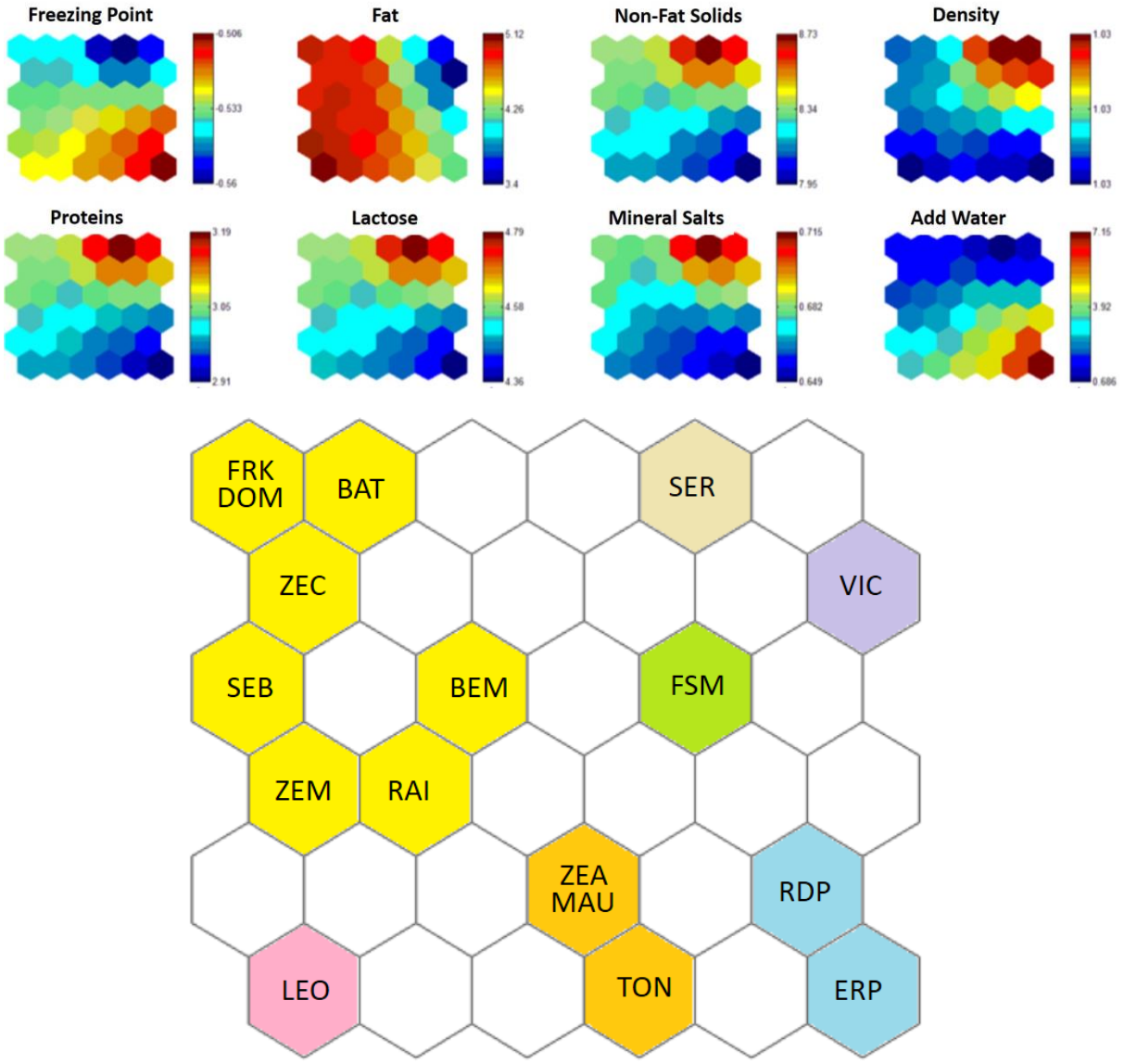

Figure 7: Results of the third collection on April 9, for the 17 producers.

In figure 7, the positioning of the SER sample on the map is more strongly influenced by the variables non-fat solids (NFS), density, proteins, lactose and mineral salts in the April 9 collection. The positioning of the ERP and RDP samples had the strongest influence of the variables added water content and freezing point. The fat variable remains the variable whose values are close for the DOM, FRK, BAT, ZEC, SEB, RAI, ZEM, LEO and TOM samples.

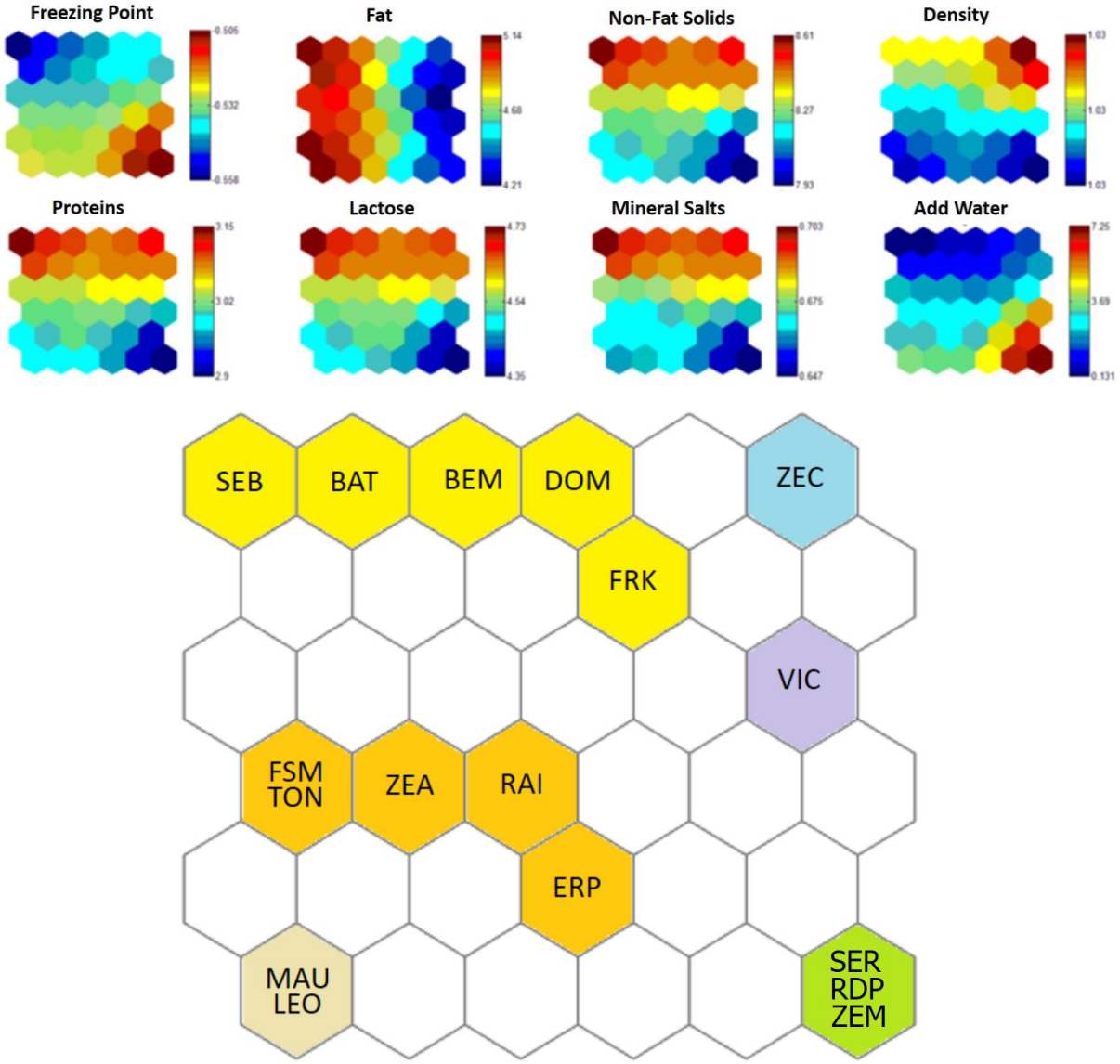

Figure 8: Results of the fourth collection on April 16th, for the 17 producers.

In figure 8, the group of samples SEB, BAT, BEM, DOM and FRK were positioned due to the variables non-fat solids (NFS), proteins, lactose and mineral salts, while the samples SER, RDP and ZEM had the most pronounced influences of the variables freezing point and added water content for the collection from April 16th. The fat variable appears again with similar values for the SEB, BAT, BEM, FSM, TON, ZEA, MAU, LEO and TON samples.

*Corresponding Author: Efraim Lázaro Reis (efraimreis@gmail.com)

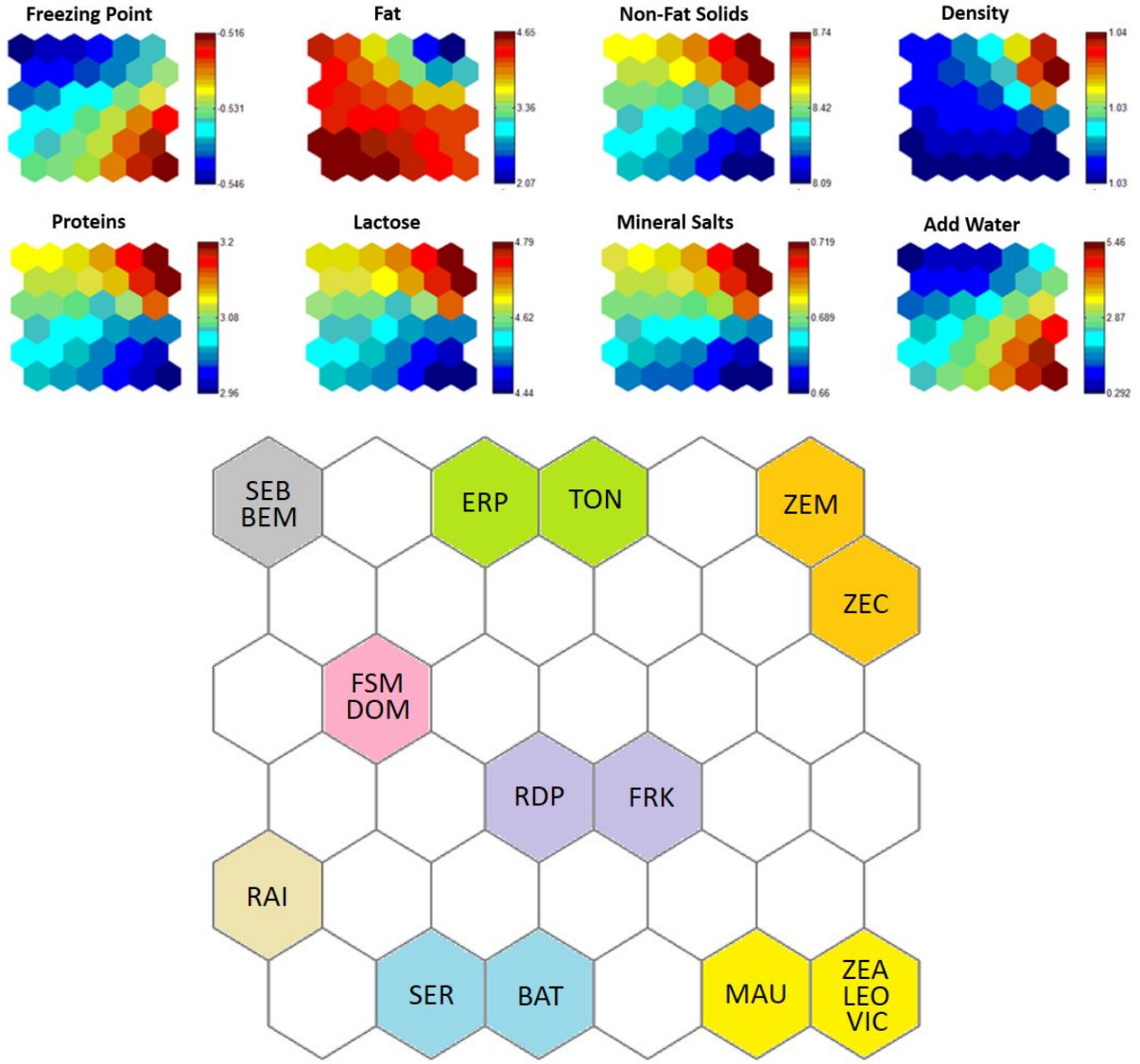

Figure 9: Results of the fifth collection on April 23h, for the 17 producers.

In figure 9, the ZEM and ZEC samples were more influenced in their positions by the variables non-fat solids (NFS), density, proteins, lactose and mineral salts, while the group of samples formed by MAU, ZEA, LEO and VIC had the most pronounced influences of the variables freezing point and added water content for April 23 collection. The fat variable in this case generally influenced the positioning of a larger set of SEB, BEM, FSM, DOM, RDP, FRK, RAI, SER, BAT, MAU, ZEA, LEO and VIC samples.

*Corresponding Author: Efraim Lázaro Reis (efraimreis@gmail.com)

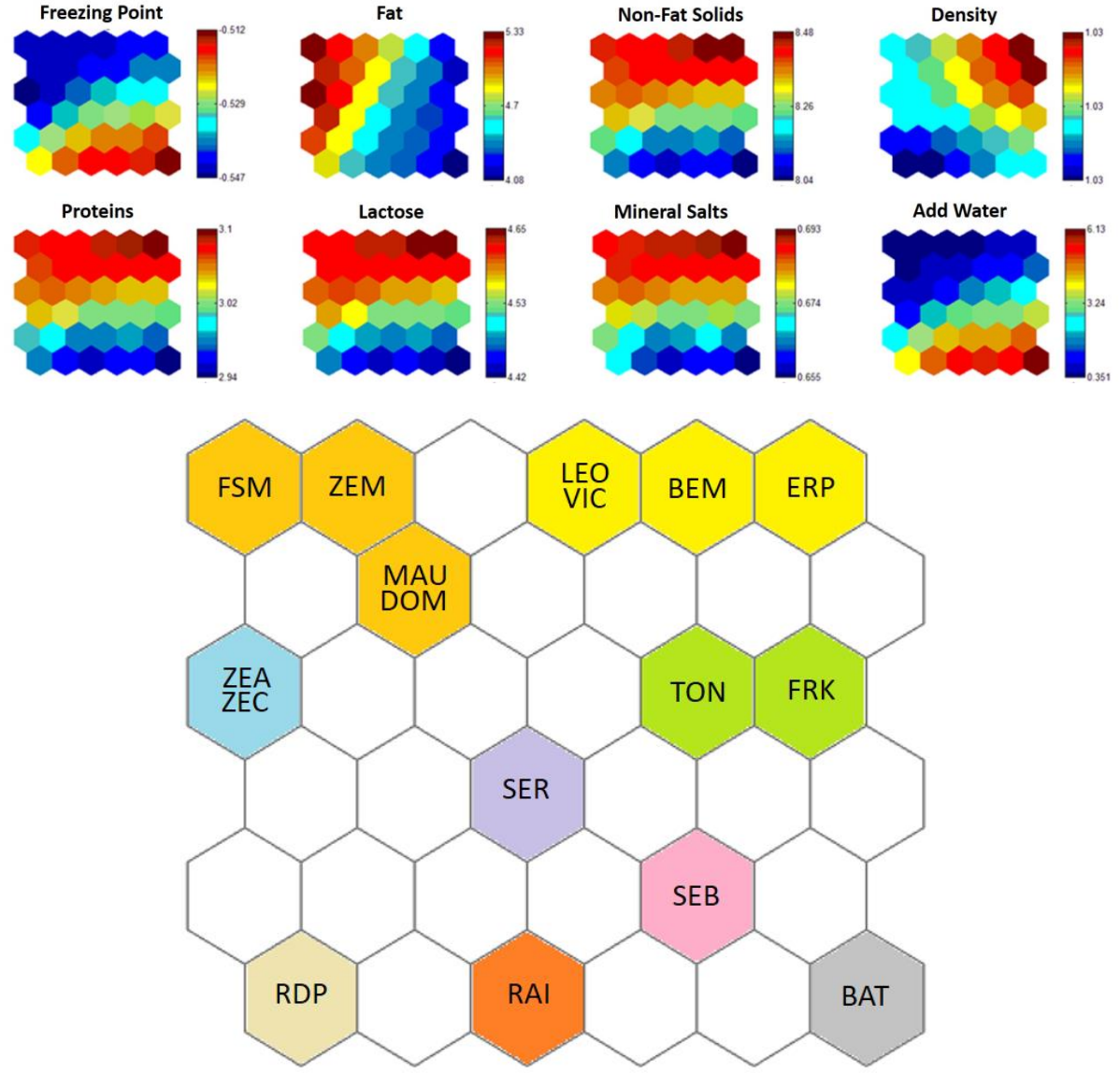

Figure 10: Results of the sixth collection on April 30th, for the 17 producers.

In figure 10, a large group of samples formed by FSM, ZEM, MAU, DOM, LEO, VIC, BEM and ERP were positioned more effectively due to the variables non-fat solids (NFS), proteins, lactose and mineral salts, while samples RAI, SEB and BAT had the same more pronounced influences of the added water content variable for the April 30 collection. The fat variable had a more pronounced influence on the FSM, ZEM, MAU, DOM, ZEA and ZEC samples on this date.

Figure 11 shows the self-organizing map for the temporal averages of all 17 producers for the 6 collections: (1) March 27th, (2) April 2nd, (3) April 9th, (4) April 16th, (5) April 23rd and (6) April 30th, according to data in Table 3. In this case, it is observed that the averages of the variables for the dates of April 9th and 16th, presented very close values, which is shown by the location in the same hexagon, and the variables non-fat solids, density, proteins, lactose and mineral salts were the ones that most presented values that contributed, more strongly the positioning of the samples for these two dates. Again, the fat variable stands out with the highest value for the collection on April 23, but also influencing the samples of March 27, April 2 and April 30, while the variables freezing point and added water content presented different values for the April 30 collection.

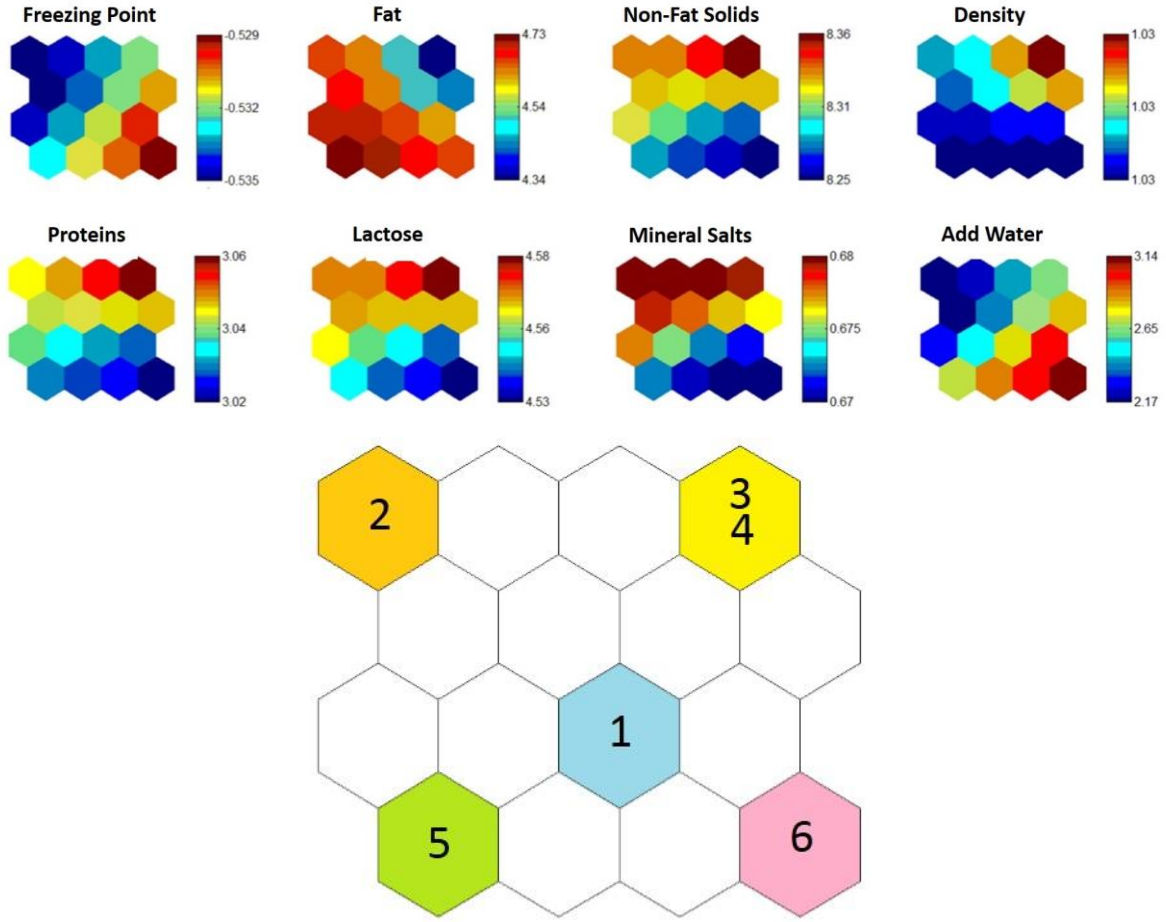

Figure 11: Temporal average of all 17 producers for the 6 collections: (1) March 27, (2) April 2, (3) April 9, (4) April 16, (5) April 23 and (6) April 30, according to data in table 3.

According to the results found, there was a significant variation in the physical-chemical parameters of each sample. Due to these variations, the maps created did not show well-defined patterns regarding the location and time of collection of each sample.

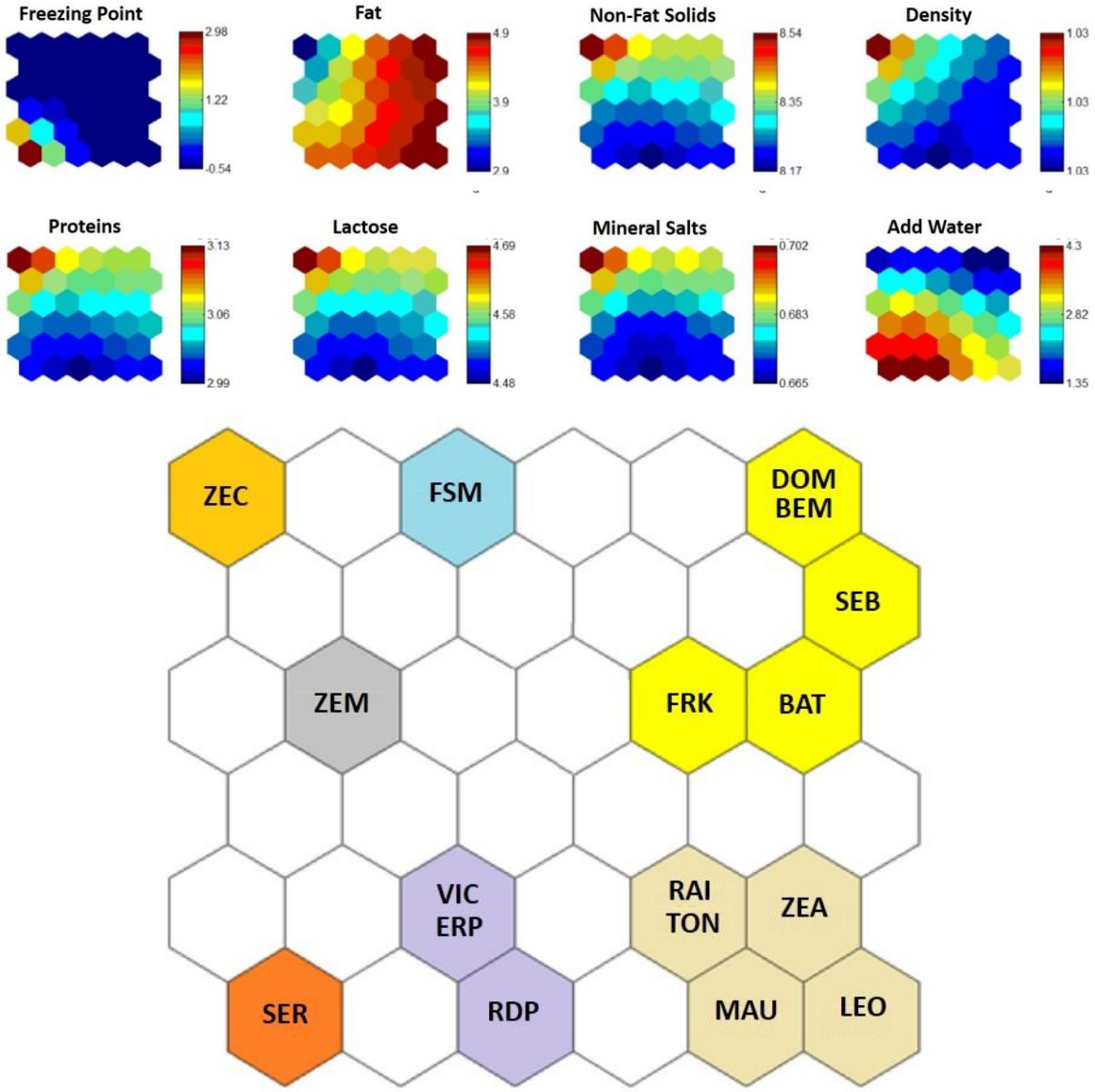

Figure 12: Spatial average of the 6 days of collection for the 17 producers, according to data in table 4.

Samples from the same location, as can be seen in figure 12, did not necessarily have the same physicochemical characteristics over time.

In view of this behavior, it was not possible to show patterns of location and time of collection for the elaboration of patterns for the spatial and temporal mapping of the origin of the milk. But it is important to emphasize that the quality of the final dairy products remained practically fixed, considering that processed milk was always the mixture of all producers.

IV. CONCLUSION

Self-Organizing Maps (SOM) are powerful and promising resources when combined with work on the determination of physical-chemical parameters.

Specifically, it allows viewing the similarities and differences between the samples, which the algorithm groups considering the similarities according to the evaluated chemical profile and separates or distances samples considered distinct, also according to the evaluated profile.

Another important aspect is the ability to form groups of samples in the network, which are also considered because they present the same characteristics evaluated by the location in different, but neighboring,

neurons. This is based on Kohonen's network processing method, which considers that the more distant a sample is placed in relation to the other, the greater differences can be pointed out between them.

For this work on the determination of physical and chemical parameters of milk collected at different locations and at different times through exploratory data analysis, separate behavior patterns were not evident. This occurred due to the uniformity of analyzed parameters. This result may have occurred because the characteristics of the samples from the same place differ in time, making it impossible to draw up a spatial profile of the samples. Dairy cattle feeding, breed and variable temperature may help explain the absence of this temporal and consequently spatial profile of the milk samples.

ACKNOWLEDGMENTS

The authors thank Fundação de Apoio à Pesquisa do Estado de Minas Gerais - FAPEMIG, for the financial support.

REFERENCES

[1]. Souza, M. R., Rodrigues, R., Fonseca, L. M., Cerqueira, M. M. O. P. (1995). Pasteurização do leite. Caderno Técnico da Escola de Veterinária, UFMG, 13, 85-93.
[2]. Brasil. (2011). Ministério da Agricultura, Pecuária e Abastecimento: Gabinete do Ministro. Instrução Normativa Nº 62.
[3]. Lay-Ang, G. (2022). A importância do leite para a saúde. Brasil Escola. Available at: https://brasilescola.uol.com.br/saude/a-importancia-leite-para-saude.htm. Accessed October 10, 2022.
[4]. Brasil. (2002). Ministério da Agricultura, Pecuária e Abastecimento. Secretaria de Defesa Agropecuária. Departamento Nacional de Inspeção de Produtos de Origem Animal. Regulamento Técnico de Identidade e Qualidade de Leite. Instrução Normativa n° 51.
[5]. Planinić, A.; Čelan, S.; S. Orjana; Pilić, Z. (2022). Modern Methods of milk Annalysis. International Journal of Research - Granthaalayah, 10(8), 11–19.
[6]. Dias, J. A.; Antes, F. G. (2014). Qualidade físico-química, higiênico sanitária e composicional do leite cru: indicadores e aplicações práticas da Instrução Normativa 62. Embrapa Rondônia.
[7]. Júnior L. C. R.; Salles M. S. V.; Fernando P. (2016). Leite com mais sólidos rende mais. Available at: http://cienciadoleite.com.br/noticia/3341/leite-com-mais-solidos-rende-mais. Accessed September 12, 2016.
[8]. Matlab (2012b). The MathWorks, Inc., Natick, Massachusetts, United States.
[9]. Faria, E. L.; Albuquerque, M. P.; Afonso, J. L. G.; Albuquerque, M. P.; Cavalcante, J. T. P. (2010). Introdução ao Toolbox de Redes Neurais de Kohonen. CBPFNT-001/10.
[10]. KOHONEN, T. (2001). SELF-ORGANIZING MAP. 3ª. ED., SPRINGER-VERLAG, BERLIM. GERMANY.
[11]. Vesanto, J.; Himberg, J.; Alhoniemi, E.; Parhankangas. J. (2000). SOM Toolbox for Matlab 5. Helsinki University of Technology, Technical Report.
[12]. Pereira Leite, E.; Souza Filho, C. R. (2010). Mapas auto-organizáveis aplicados ao mapeamento do potencial mineral na região de serra leste, província mineral de Carajás, Pará. Revista Brasileira de Geofísica. 28(3), 397- 409.
[13]. Drumond, A. R.; Reis, E. L.; Veado, M. A. V. R.; Araujo, M. M. R.; Moreira, J. M. Q. (2014). Utilização de mapas auto-organizáveis para verificação da distribuição de metais na atmosfera de Ipatinga, Minas Gerais, através de biomonitoramento com Sphagnum Capillifolium. Eclética Química, 39, 175-191.
[14]. Dutra, R. A. Reis, E. L.; Reis, C.; Fidêncio, P. H.; Reis, C.; D. G.; Damasceno, O. I. C. (2021). Análise espacial e temporal de dados físicos e químicos de águas superficiais, por mapas auto organizáveis (SOM). Brazilian Journal of Development, 7(6), 57578-57594.
[15]. Bioinf Gif. Available at: https://gifer.com/pt/83cQ. Accessed May 10, 2023.
[16]. Venturoso, P. C.; Almeida, K. E.; Rodrigues, A. M.; Damin, M. R; Oliveira, M. N. (2007). Determinação da composição físico-química de produtos lácteos: estudo exploratório de comparação dos resultados obtidos por metodologia oficial e por ultra-som. Revista Brasileira de Ciências Farmacêuticas, 43(4), 607-613.
[17]. Bobbio, F. O., bobbio, P. A. (2003). Introdução à química de alimentos. 3ª. ed. São Paulo.
[18]. Tronco, V. M. (2010). Manual para Inspeção da Qualidade do Leite. 4. ed. Santa Maria: Ed. UFSM.
[19]. Venturini, K. S; Sarcinelli, M. F; Silva L. C. (2007). Características do Leite. Universidade Federal do Espírito Santo – UFES. Programa Institucional de Extensão, Boletim Técnico.
[20]. Instituto Adolfo Lutz. (2008). Métodos físico-químicos para análises de alimentos. 4ª. ed. São Paulo.
[21]. Lanagro/RS. (2011). Depressão do Ponto de Congelamento do Leite Fluido. MAPA/SDA/CGAL Laboratório Nacional Agropecuário - LANAGRO/RS. Laboratório de Produtos de Origem Animal Método de Ensaio - MET Código: MET POA/10/01/01.
[22]. SOM Toolbox 2.0. (2005). Available at: http://www.cis.hut.fi/projects/somtoolbox. Accessed January 6, 2023.