Brazilian Journal of Experimental Design, Data Analysis and Inferential Statistics

Federal University of Rio de Janeiro
BJEDIS, Rio de Janeiro, Special Edition, v. 01 (2), 2021
DOI: https://doi.org/10.55747/bjedis.v1i2.48370
ISSN: 2763-6925

Special Edition **Received 04/28/2021. Revised 05/11/2021. Accepted 05/30/2021. Published 07/01/2021.**

Surface modeling as a tool for constructing pseudo ternary diagrams

Juliana V. Silva Camilo[1], Anna A. G. Ferreira[1], Josane A. Costa[1], Claudia R. E. Mansur*[1,2] e Fernando Gomes S. Jr.[1,3]

[1] Instituto de Macromoléculas Professora Eloisa Mano, Universidade Federal do Rio de Janeiro (IMA/UFRJ), Centro de Tecnologia, Bl. J, Cidade Universitária, Ilha do Fundão – Rio de Janeiro – Brazil. [2] Programa de Engenharia Metalúrgica e de Materiais (PEMM) /COPPE, Universidade Federal do Rio de Janeiro. Centro de Tecnologia, Bl. F, Cidade Universitária, Ilha do Fundão – Rio de Janeiro – Brazil. [3.] Programa de Engenharia Civil, COPPE, Centro de Tecnologia - Cidade Universitária, Av. Horacio Macedo, 2030, Bloco I. Universidade Federal de Rio de Janeiro, Brasil, Zip code 21941-914, phone number +55 21 3938-7766

Abstract: Microemulsions have been developed and evaluated for application in various fields. For their development, ternary phase diagrams are initially constructed, obtained by investigating mixtures of the aqueous and oil phases and a surfactant, to determine the regions composed of liquid microemulsions. However, the construction of these diagrams requires many steps, so simplification of this method can be useful for the formulation and application of these microemulsified systems. Therefore, this article proposes a method to obtain these diagrams employing mathematical modeling, by testing systems composed of kerosene (as oil phase), a nonionic surfactant based on ethoxylated nonylphenol (NP80), and salt water. The study aimed to ascertain which models (linear, quadratic or special cubic) were able to describe the empirical observations, i.e., the ternary phase diagram containing 99 points. The results obtained showed that a quadratic model with a smaller sample (31 points) was able to represent these systems with 61% variability of the observed data and a confidence level higher than 99.9%. The smaller sample (31 points) was defined by means extrapolation lines, as described in the standard ISO11358, permitting identifying the liquid microemulsion regions of the system composed of kerosene, NP80 and salt water. This result is very promising since it permits determining the conditions of interest with significantly less experimental work, and commensurate reduction of the final cost of the experiments.
Keywords: Microemulsions, phase diagrams, mathematical modeling.

Adherence to BJEDIS Scope: This work is closely related to the scope of BJEDIS, as it presents a new method able to identify of types of emulsion and represent a ternary diagram with a reduced sets of compositions.

* Corresponce address for this author ot the Biopolymers & Sensors Lab. / Macromolecules Institute / Universidade Federal do Rio de Janeiro, Brazil; Tel/Fax: +55-21-3938-7766; E-mail: celias@ima.ufrj.br

1. INTRODUCTION

A microemulsion is a thermodynamically stable, isotropic and optically transparent microdispersed system composed of an oil phase, aqueous phase and an amphiphilic substance (surfactant), forming droplets with diameters between 10 and 100 nanometers [1; 2].

Microemulsions are attracting keen interest for application in various fields due to their specific properties, such as reduction of the interfacial tension between immiscible liquids (typically oil and water), as long as their composition is adequately adjusted, and large interfacial area. These systems are kinetically and thermodynamically stable due to the small size of their dispersed droplets and good solvency capacity, i.e., ability to promote the simultaneous solubilization of hydrophilic and hydrophobic substances in the organic and aqueous phases, respectively, of the microemulsified system [3-5].

Systems formed by microemulsions can be applied in various areas, such as the food processing, pharmaceutical, cosmetic and petroleum industries, among others, by tailoring them to produce efficient results for each application (6-10).

Microemulsions can be formed spontaneously when the aqueous and oil phases and surfactant are placed in contact [11], making it possible to combine enormous quantities of immiscible liquids into a single homogeneous phase [12]. However, the simple mixture of the constituents of microemulsions does not always lead to a homogeneously dispersed system, instead forming multi-phase systems [13]. A classification system was proposed by Winsor in 1948 to define the various equilibria existing between the oil and aqueous phases, consisting of four types of systems:

- Winsor I region, with two phases, an oil phase in equilibrium with the emulsified phase containing oil, water and most of the surfactant mixture;

-Winsor II region, also with two phases, consisting of an aqueous phase in equilibrium with an emulsion;

- Winsor III region, with three phases, containing an aqueous phase and oil phase, separated by an emulsified phase; and

- Winsor IV region, corresponding to a visually single-phase region [14].

To produce microemulsified systems, ternary phase diagrams are first constructed, obtained by mixing the aqueous and oil phase and surfactants in varying proportions and processing conditions. These diagrams are triangular graphs in which each component is represented by one of the vertices of the triangle, so that the sum of these variables is always 100% [15]. This type of graph allows identifying the microemulsion regions with different compositions, where the modification of the composition of the mixture of immiscible solvents can provide specific information about their chemical structures, by the location of the transition limits between regions of macroemulsions, phase separation, oil-in-water (O/W) and water-in-oil (W/O) microemulsions [16; 17].

These diagrams enable defining the different regions composed by the microemulsions and macroemulsions. In the case of microemulsified systems, the central region of interest is the Winsor IV region, where visual observation indicates the existence of a single phase.

However, the method to construct ternary diagrams demands lengthy experiments, since it is necessary to prepare many samples with different compositions or mixtures of various proportions of the oil phase, aqueous phase and surfactant, besides statistical treatment of the data in each case [18; 19].

In this work, a new method was developed to construct ternary diagrams, in which it was possible to identify the different compositions and microemulsion regions by applying mathematical modeling. The modeling of processes is a tool widely used in engineering, since it permits predicting the results of physical, mathematical, biological and chemical phenomena, among others, through mathematical models that are obtained from experimental and analytical tests [20; 21].

The goodness of fit of the model and the experimental data is determined by calculating the coefficient of determination (R^2), which describes the proportion of the total variance in the data that can be explained by the model. It varies between 0 and 1, where higher values indicate better concordance between the variables. That parameter is interpreted as the proportion of the dependent variable Y that is explained by the variation of the independent variable X. For example, if R^2 is equal to 1, it means that 100% of the variation of Y is explained by the variation of X [22; 23].

Another statistical tool that is very useful to ascertain the quality of the fit of the values predicted by the model and

the experimental data is the root-mean-square error (RMSE), i.e., the error represented by the values estimated by the model concerning the observed values [24]. The ideal RMSE value is near zero, since this means the future estimate of the mathematical model will be significant [25; 26].

The new method reported here allowed identifying the regions of liquid microemulsions and the inversion point of the micelles with a smaller number of experiments, thus minimizing the time spent in the laboratory, making the checking and identification of the microemulsion regions in different systems faster. More specifically, it was possible to identify the smallest number of experiments necessary to identify the regions of liquid microemulsions, in a form equivalent to that obtained using ternary diagrams containing 99 points.

2. METHODOLOGY

2.1. Materials

The materials used in this study were: kerosene, obtained from Vetec (Rio de Janeiro, Brazil) as oil phase; a commercial ethoxylated nonionic surfactant (NP80, produced through the reaction of nonylphenol with ethylene oxide, containing 8 ethylene oxide units), purchased from Oxiteno (São Paulo, Brazil); and as aqueous phase, distilled water containing the following salts (total concentration of 20,000 ppm): NaCl (39,300 ppm), KCl (604 ppm), $MgCl_2$ (1,620 ppm) and $CaCl_2$ (1,700 ppm), obtained from Isofar (Duque de Caxias, Brazil). The aqueous phase is a compound of the saline solution with 25% formation water (water naturally formed inside the reservoir) and 75% injected water (desulfated seawater injected in the reservoir during secondary recovery).

2.2. Methods

2.2.1. Preparation of the compositions for constructions for of the ternary diagram (D99)

For the construction of these ternary diagrams, we followed a method previously defined in the literature [9]. Compositions were prepared, called standard compositions, initially with nine oil/surfactant proportions (1:9, 2:8, 3:7, 4:6, 5:5, 6:4, 7:3, 8:2, 9:1), homogenized at rest for 24 hours. Then we added quantities of water in each flask during 11 consecutive days. After each water addiction, the flasks were homogenized at rest for 24 hours. After each titration of the aqueous phase, we evaluated the macroscopic aspects of formulations. According to the macroscopic features determined by visual inspection, the formulations were classified as microemulsions, macroemulsions (liquid, viscous and gel) or phase-separated. Transparent systems were classified as microemulsions (Mi), and milky formulations were classified as macroemulsions (Ma). The difference between liquid, viscous and gel systems is related to the dispersion's ability to flow when placed at an inclination of 45° (9). These quantities of water were added in each flask during 11 consecutive days to obtain the three components of the system, resulting in a sum of 100% of the three components in all the flasks on all days (Figure 1).

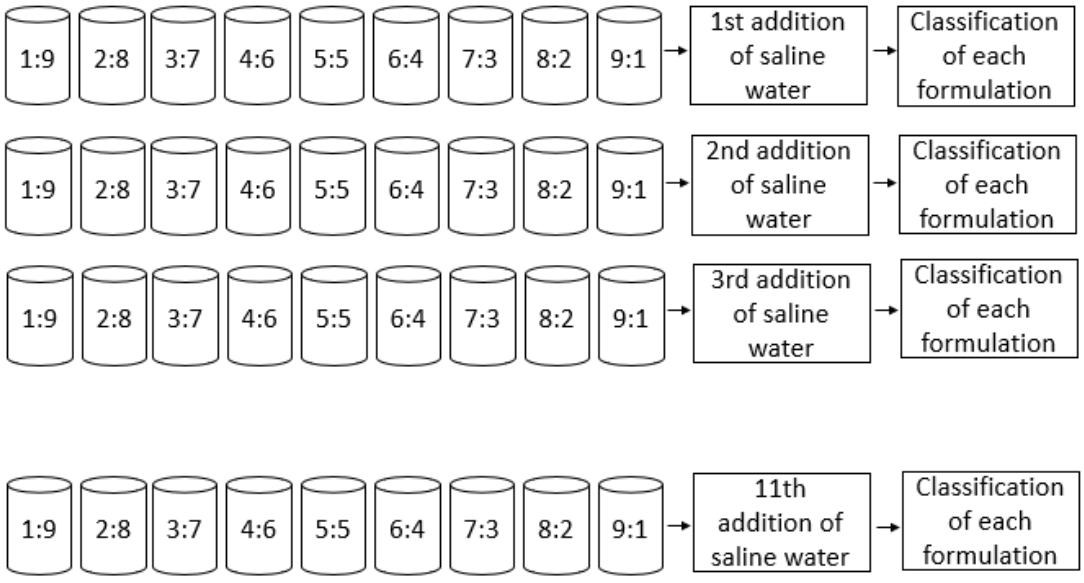

Figure 1 - Illustration of the addition salt water and classification of each formulation

The proportions are prepared and after addition of the water, 99 systems were obtained, enabling mapping the whole area of the phase diagram. The ternary diagrams were constructed with the Qtiplot software, containing all the points obtained and based on visualization of each system, according to the classification of Winsor [9]. Tables 1 and 2 show the mass amount of each component in the formulation**.**

2.3. Mathematical modeling of data

The new method presented here aims to prove the ability of a mathematical model, prepared from a reduced number of compositions, to accurately represent a standard phase diagram containing 99 compositions. Besides this, the novel modeling method presented allows representing and identifying the microemulsion or macroemulsion liquid, viscous or gel or phase separation regions.

2.3.1. Construction of the mathematical models and surface area with the set of 99 formulations

To construct the mathematical models from the empirical observations (phase diagram with 99 formulation), we used the system kerosene/NP80/salt water, according to the following method:

(i) First, a product satisfaction scale was constructed, that is, a value of importance was assigned to each type of formulation making up the phase diagram, as described in Table 3.

Table 1 - Mass amount of oil phase and surfactant in each ratio.

Proportions	Oil Phase (mg)	Surfactant (mg)
1:9	500	4500
2:8	1000	4000
3:7	1500	3500
4:6	2000	3000
5:5	2500	2500
6:4	3000	2000
7:3	3500	1500
8:2	4000	1000
9:1	4500	500

Table 2 - Mass amount of aqueous phase on each day.

Days	Aqueous phase (mg)	Classification
1º	435	1, 2, 3, 4, 5, 6, or 7
2º	952	1, 2, 3, 4, 5, 6, or 7
3º	1579	1, 2, 3, 4, 5, 6, or 7
4º	2353	1, 2, 3, 4, 5, 6 or 7
5º	3333	1, 2, 3, 4, 5, 6, or 7
6º	4615	1, 2, 3, 4, 5, 6, or 7
7º	6364	1, 2, 3, 4, 5, 6, or 7
8º	8889	1, 2, 3, 4, 5, 6, or 7
9º	12857	1, 2, 3, 4, 5, 6, or 7
10º	20000	1, 2, 3, 4, 5, 6, or 7
11º	36667	1, 2, 3, 4, 5, 6, or 7

1 – Liquid Microemulsion; 2 – Viscous Microemulsion; 3 – Gel Microemulsion; 4 – Liquid Macroemulsion; 5 – Viscous Macroemulsion; 6 – Gel Macroemulsion; 7 – Phase Separation.

Table 3 - Value of importance assigned to each formulation of the phase diagram.

Formulations	Value assigned (%)
Liquid Microemulsion	70
Viscous Microemulsion	55
Gel Microemulsion	35
Liquid Macroemulsion	20
Viscous Macroemulsion	10
Gel Macroemulsion	5
Phases Separation	0

(ii) The entry data of each composition were inserted in the R software, in the order: saltwater/kerosene/NP80/classification. The order of the entry data defines the position of each component on the vertex of the equilateral triangle (diagram).

(iii) After inserting the entry data in the R program, three mathematical models were generated: (i) linear – LM99; (ii) quadratic – QM99; and (iii) special cubic – SCM99, to determine which one best fit the input data, i.e., the values of salt water/oil/surfactant/value assigned to the ternary diagram containing 99 formulations.

(iv) The response surfaces and coefficients of determination (R^2) were also generated for each mathematical model.

(v) Then, a standard statistical test was applied to calculate the p-value, in order verify the mathematical models' level of significance.

2.3.2. Choice of points to construct the mathematical models with smallest and crescent number of formulation (Pn)

In this step, points were selected to map the phase diagram, testing nine sets with smallest and crescent numbers of formulations, performed as follows:

(i) First, three external points of each vertex were chosen at the minimum point of kerosene, NP80 and saline water, respectively, namely points 3, 5 and 1. Next, a central point [7] and the means of the vertices (kerosene, NP80 and saline water) were chosen, respectively points 4, 6, 2. The selection of these formulations composed the first set, containing seven points (Figure 2A).

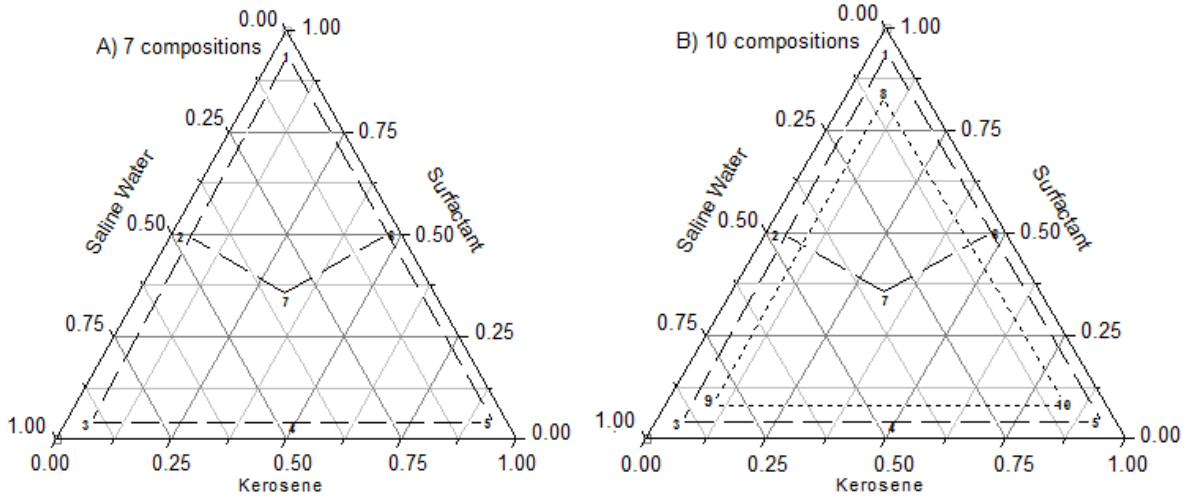

Figure 2 - A) First set of points selected, for a total of 7 compositions. B) Second set of points selected, for a total of 10 compositions.

(ii) Then the other points were chosen by scanning every three points of the vertices in the direction of the center, always with triangular geometry, until attaining 31, as shown in Figures 2 to 6.

(iii) The sets of 16, 28 and 31 points were approximately the mean of sets with 13, 22 and 10 points, respectively.

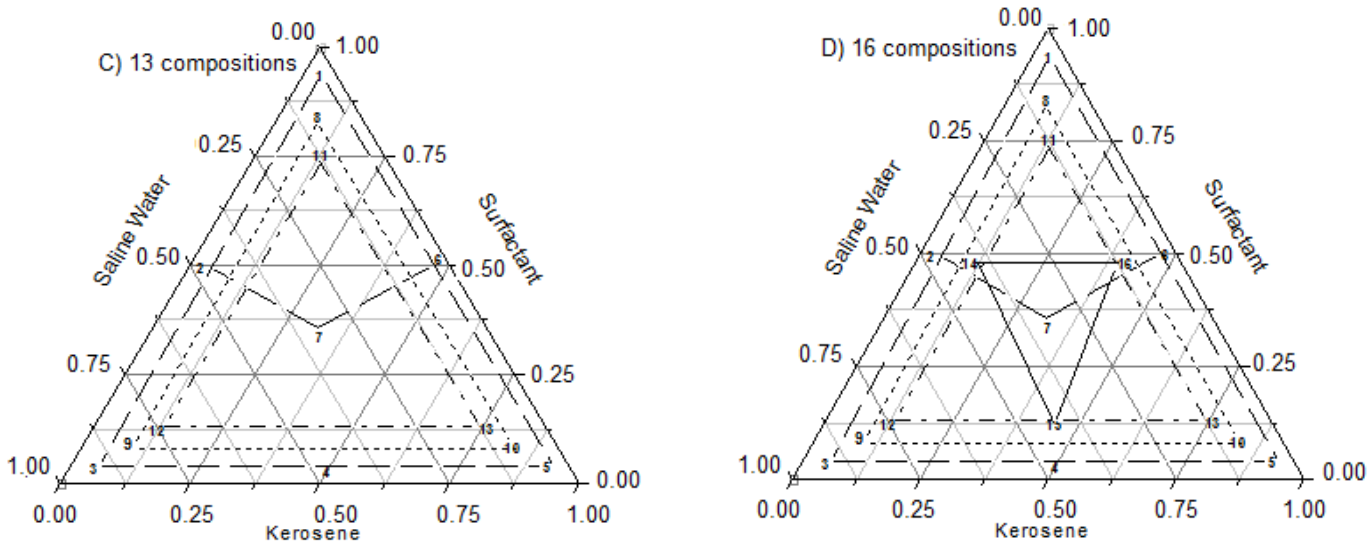

Figure 3 - C) Third set of points selected, for a total of 13 compositions. D) Fourth set of points selected, for a total of 16 compositions.

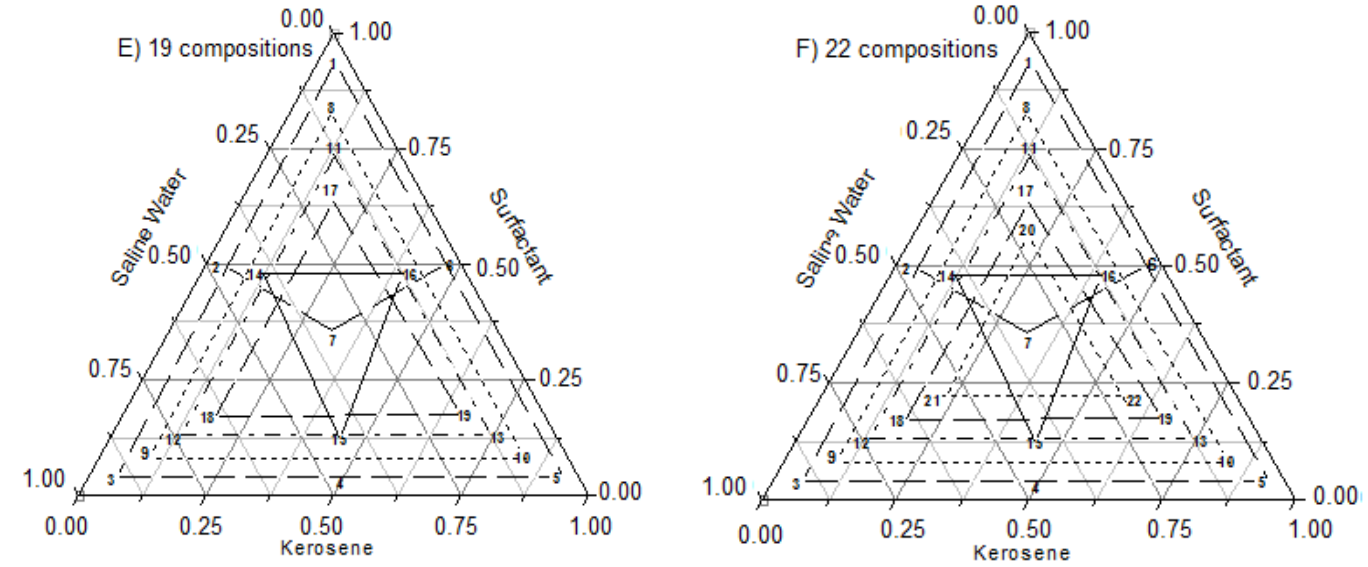

Figure 4 - E) Fifth set of points selected, for a total of 19 compositions. F) Sixth set of points selected, for a total of 22 compositions.

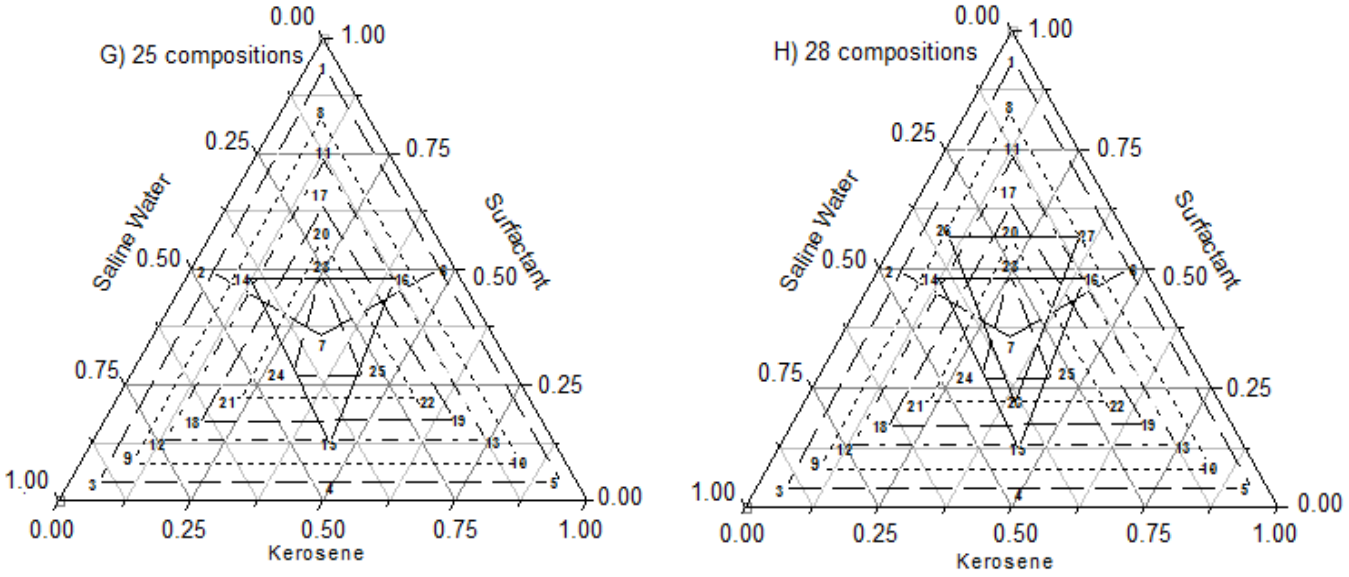

Figure 5 - G) Seventh set of points selected, for a total of 25 compositions. H) Eighth set of points selected, for a total of 28 compositions.

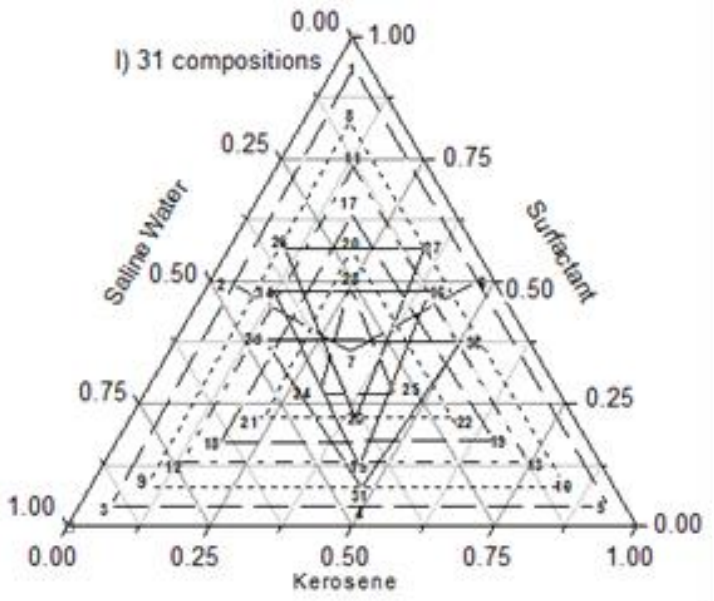

Figure 6 - I) Ninth set of points selected, for a total of 31 compositions.

The sets of points tested to construct the mathematical models with each iteration of sets 7, 10, 13, 16, 19, 22, 25, 28 and 31 points. The first set of points tested was with seven formulations, in order to map the ends and means of the diagram. The other eight sets were built every three points, forming a triangle at the center of the phase diagram. The numbers of of point sets tested in this modeling method were determined according to the mean square error, which decreased with increasing formulations. However, the maximum number of formulations was determined when the mathematical model presented good fit to the data, reduction of the mean square error and minimum amplitude variation after the mapping of the phase diagram.

2.3.3. Preparation of the formulations of the sets tested with each iteration of sets (Fn).

Nine sets with 7, 10, 13, 16, 19, 22, 25, 28, and 31 compositions were tested. For example, the first set contained seven compositions, each of which was reproduced in the laboratory, i.e., the components were blended according to the content of each formulation selected in the previous section. For all the sets of points, each composition was reproduced in the laboratory following the method below:

(i) First, the oil phase and surfactant were placed in a 50 mL flask, which was shaken manually until complete homogenization, after which the flasks were left at rest for 24 hours at room temperature.

(ii) Next, quantities of salt water were added slowly in each flask, to assure the incorporation of the aqueous phase, which remained in equilibrium in the system. Then 2 grams of salt water at room temperature was added to each flask, and the solution was submitted to magnetic stirring for 30 minutes. This step was repeated until the addition of the entire quantity of the composition's aqueous phase (approximately four saltwater additions, each one of 2 grams, every half hour).

(iii) According to the macroscopic features determined by visual inspection, the formulations were classified as microemulsions, macroemulsions (liquid, viscous or gel) or phase-separated, as described in section 2.2.1.

(iv) Finally, a product satisfaction scale was formulated, that is, a value of importance was assigned to each type of formulation that made up each set tested with smallest and crescent number of points, as described in Table 3.

2.3.4. Construction of the surface area and mathematical model for each iteration of the sets tested (Mn).

The mathematical model with each iteration of sets able to describe the set of empirical observations (D99) obtained in the experiment with 99 points was determined as follows. We constructed the surface areas and respective mathematical models of the 9 sets defined with a smallest and crescent number of experiments.

The models were constructed by substituting the independent variables for w, k and s, denoting water, kerosene and surfactant, respectively. The models were: liner f(w, k, s)=A*w+ B*k + C*s); quadratic, f(w, k, s) = A*w + B*k + C*s + D*w*k + E*w*s + F*k*s); and special cubic, f(w, k, s) = A*w + B*k + C*s + D*w*k + E*w*s + F*k*s + G*w*k*s). For each iteration of sets, the model was determined as follows:

(i) First, we inserted the values of the independent variables w, k, s and Va of each point, that is, the quantities (salt water/kerosene/surfactant) and the value assigned to each formulation, in the R software. The entry data were inserted in the order column A (salt water), column B (kerosene), column C (surfactant) and column D (value assigned).

(ii) The mathematical models (LMn, QMn, and SCMn) and the response surfaces for each set of the smaller and crescent number of points were generated individually.

(iii) Then, the predicted values of each model were tested and compared with the predicted values of the mathematical model of the same nature that was obtained from the set of 99 points from the empirical observations (LM99, QM99, and SCM99).

(iv) The comparison between the prediction (LM99, QM99, SCM99) x (LMnp, QMnp, SCMnp), where n is each iteration of sets, permitted calculating the root mean squared error (RMSE).

(v) With the RMSE values, an adjustment of a first-order exponential decay curve with R^2 greater than 0.99 was made.

(vi) By means of extrapolation lines, according to ISO11358, the group with the lowest number of formulations was able to describe the set of empirical observations (phase diagram with 99 formulations) with a significance level of 99.99%.

We repeated the method described above for each iteration of sets of points, i.e., we tested the linear, quadratic and special cubic models for the 9 sets defined (7, 10, 13, 16, 19, 22, 25, 28 and 31 experiments).

3. RESULTS AND DISCUSSION

3.1. Construction of the phase diagrams

The diagrams make it possible to observe various types of dispersions, including translucent systems, such as liquid and viscous microemulsions, and opaque systems, such as liquid and viscous macroemulsions. Microemulsions were classified as systems that were transparent and homogeneous during the phase diagram construction, and can be seen in Figure 7. Both were liquid (transparent and homogeneous fluids, as observed when moving the flask), and gel or viscous (no or little movement when inverting the bottle). Figure 8 shows the appearance of the macroemulsions obtained, which differ from microemulsions by opacity.

In the kerosene/NP80/salt water phase diagram, the formulations were classified as oil-in-water liquid microemulsions after reaching a saltwater content of approximately 55%, generally after the seventh salt water addition. In this system, we did not observe formulations classified as liquid macroemulsions, because when increasing the quantity of salt water in the system, the formulations presented phase separation.

Figure 7 - A) Classification as gel microemulsion in the proportions 1:9 and 2:8 after the second addition of salt water; B) Classification as viscous microemulsion in the proportions 1:9 and 2:8 after the third addition of salt water; and C) Classification as liquid microemulsion in the proportions 1:9 and 2:8 after the eleventh addition of salt water, respectively.

Figure 8 - A) Classification as gel macroemulsion in the proportion 6:4 after the fifth addition of salt water; B) Classification as viscous macroemulsion in the proportion 6:4 after the sixth addition of salt water; and C) Classification as phase separation in the proportion 6:4 after the seventh addition of salt water.

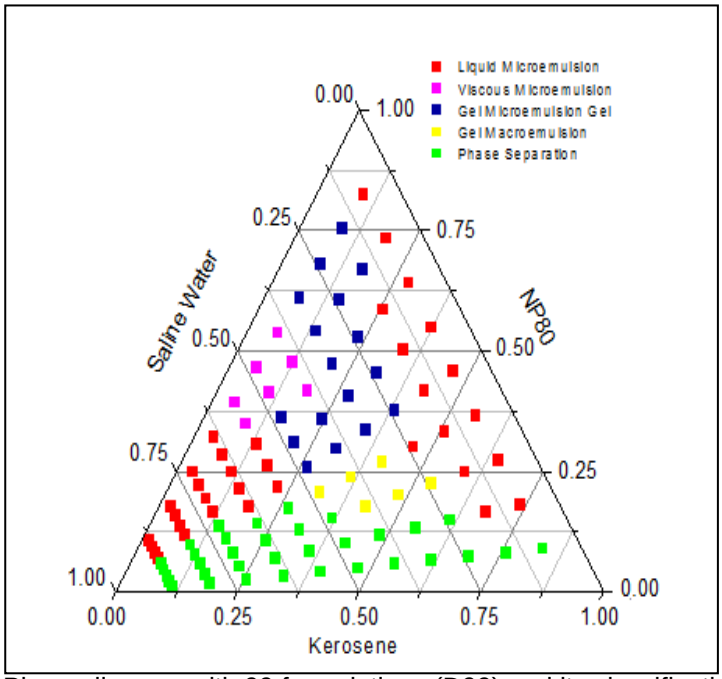

Figure 9 - Phase diagram with 99 formulations (D99) and its classification regions.

This phase diagram (D99) shows the distribution and identification of the regions formed. In general, the diagram constructed experimentally with the 99 formulations of the system formed by kerosene, NP80 and salt water presented characteristics of determined types of emulsions. Water-in-oil (W/O) liquid microemulsions can be observed along the surfactant axis in the ternary diagram, where the water content is below 10%, favoring the dispersion of the water droplets in the continuous phase. In this region, the microemulsions had different concentrations of surfactants, and it could be visually observed that at low surfactant concentrations, approximately 30%, the viscosity of the microemulsion declined. In counterpart, for high surfactant levels, the W/O microemulsion system had higher viscosity (higher than 30%), maintaining its reverse micellar geometry, since the system was still in equilibrium.

Another region, this time characteristic of oil-in-water (O/W) liquid microemulsions, was observed along the saltwater axis of the ternary diagram, where the oil content was below 10%, favoring the dispersion of the oil droplets in the continuous phase. In this case, increasing the surfactant concentration raised the viscosity of the medium, and it was possible to observe the formation of O/W microemulsified systems visually. Therefore, between the W/O and O/W liquid microemulsion regions, there was a viscous microemulsion region where inversion of phases (O/W to W/O) probably occurred, meaning low interfacial tension. In regions where the surfactant concentration was below 20%, phase separation was generally observed.

When the surfactant concentration was between 20 and 30%, it was possible to observe an intermediate region of macroemulsion, in gel form, located between the microemulsion and phase separation regions. This enabled defining the ternary diagram of the kerosene/NP80/salt water system with probable compositions for each region, according to Figure 9. These regions were displaced differently according to the characteristics and affinities of the components of the emulsion (oil phase, aqueous phase, and surfactant).

As can be noted, the experimental method used to identify the regions of O/W or W/O microemulsions of the kerosene/NP80/salt water system is very complex and demands extensive experimental time. Besides this, the identification of the microemulsion regions only occurs after treatment of the data to construct the ternary diagram.

Therefore, the idea behind the new method proposed here is to reduce the number of experiments necessary to identify the microemulsion regions of the kerosene/NP80/salt water system, or any other system with similar characteristics, to find the best microemulsion formulations in less time.

3.2. Mathematical modeling of the data

A new method for the construction of phase diagrams was developed using mathematical models. Thus, it was possible to map the regions of microemulsion or macroemulsion: liquid, viscous, gel or phase separation with smaller and crescent number of formulations. This method reduces the time and quantity of materials spent in the laboratory. The mathematical modeling was tested with a standard kerosene/NP80/salt water system.

3.2.1. Construction of the mathematical models for the set with 99 formulations

The entry data of the phase diagram with 99 formulations of kerosene/NP80/salt water were used to construct three mathematical models. The result was three graphs of the surface area (or response surface), one for each of the three mathematical models: linear (LM99), quadratic (QM99) and special cubic (SCM99), along with the coefficient of determination (R^2) and product satisfaction scale (value assigned – Table 3) through size of contour areas.

Generation of a response surface is a statistical technique to model and analyze problems in which the response variables, i.e., the classifications of the compositions, are influenced by various factors, with the aim of optimizing the response.

In this work, we fitted the different formulations to compose a ternary diagram by multiple regression, in which more than one independent variable (oil/surfactant/water) influences the response variable. Response variable is the type emulsion formed, in which the value was assigned in a scale from 0 to 70%. This scale was a way of separating the types of emulsion formed, that is, to transform the qualitative observations into numerical ones.

Mathematical modeling was done using the least squares technique. Figures 10, 11 and 12 show the response surfaces, mathematical equations and coefficients of determination of the models (linear – LM99, quadratic – QM99 and special cubic – SCM99) for the empirical observations, respectively. To assess which of the mathematical models best described the kerosene/NP80 system with 99 experimental points, we used the parameter that supplies information on the goodness of fit, the coefficient of determination (R^2).

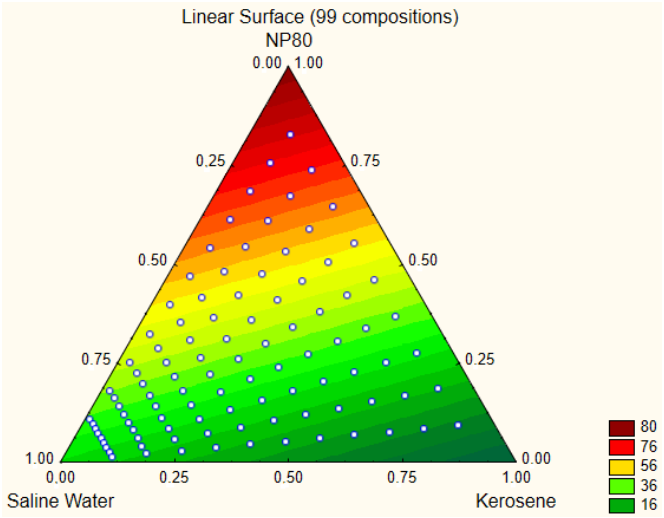

Figure 10 - Response surfaces of linear mathematical model (LM99) of set with 99 formulations.

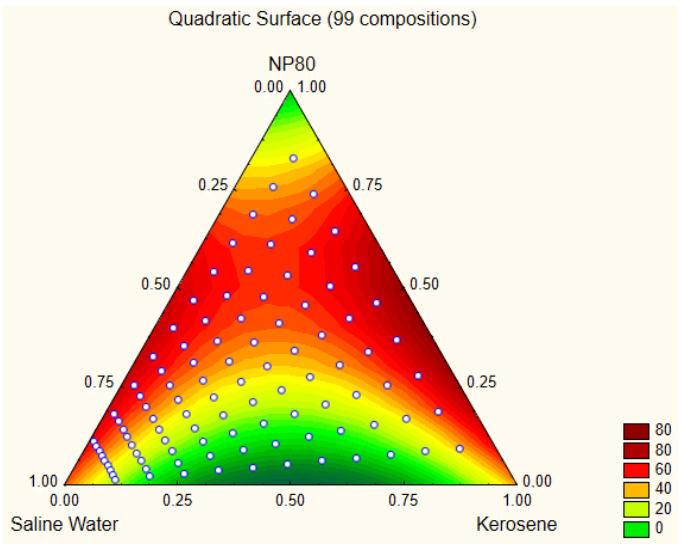

Figure 11 - Response surfaces of quadratic mathematical model (QM99) of set with 99 formulations.

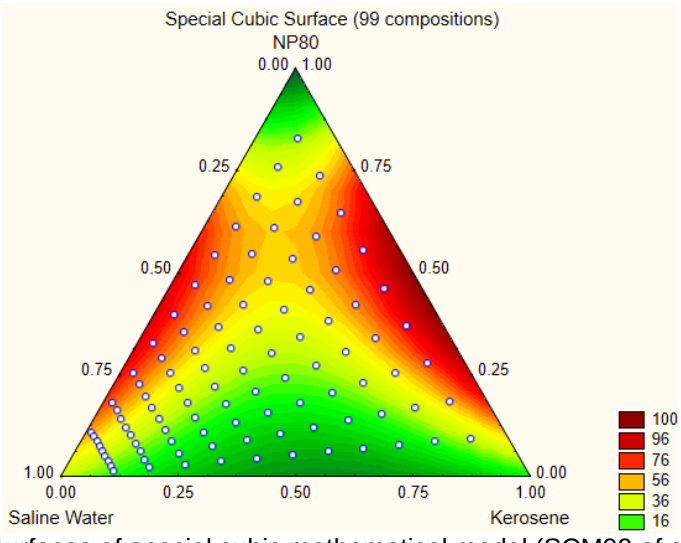

Figure 12 - Response surfaces of special cubic mathematical model (SCM99 of set with 99 formulations.

The mathematical models, R^2 and p-values generated for the QNP80/salt water system with 99 points were as follows:

$$F(w, k, s) = 23.4927*w-4.6549*k+95.7047*s \quad (1)$$

$$F(w, k, s) = 49.3673*w+41.4229*k-10.0471*s-344.8685*w*k+195.5358*w*s+276.5345*k*s \quad (2)$$

$$F(w, k, s) = 30.195*w+6.925*k-44.545*s-165.2454*w*k+375.1588*w*s+536.5786*k*s-1285.2998*w*k*s \quad (3)$$

- Linear model (LM99) - equation 1 with determination coefficient equal to 0.2729 and p-value 0.0061.
- Quadratic model (QM99) - equation 2 with determination coefficient equal to 0.7835 and p-value < 0.0001.
- Special model (SCM99) – equation 3 with determination coefficient equal to 0.8015 and p-value < 0.0001.

The response surfaces in the quadratic (QM99) and special cubic (SCM99) models were similar and the coefficient of determination for the special cubic model was slightly higher (R^2=0.8015) than that in the quadratic model (R^2=0.7835). By conventional statistical test criteria (p-value test of statistical significance), these determination coefficients are considered statistically significant because both models were constructed with confidence degree of the 99.99%, that is, since the quadratic model is less complex then special cubic. The quadratic model was able to represent the phase diagram containing 99 points with 99.99% confidence. There was an error with probability of 0.0001% (p-value), so this quadratic model did not fit the data [27]. Since the special cubic is more complex, we decided that the quadratic model produced the best fit the entry data of the ternary diagram for Q/NP80/salt water with 99 compositions. The surface area of the quadratic model (Figure11) shows the region with the highest value of the product satisfaction scale (classification). This was similar to the region of microemulsions formed in the ternary diagram with the set of 99 experimental points – D99 (Figure 9). That confirms that the response surface of the quadratic model (QM99) more significantly represented the classifications formed in the ternary diagram (D99).

3.2.3. Construction of the surface area and the mathematical model for each iteration of the sets tested

The formulations selected were reproduced in the laboratory according to the method described in section 2.3.3. Then the mathematical models and surface area were constructed. As described in section 3.2.1 for the quadratic model (QM99) that best described the ternary diagram with 99 formulations because of the determination coefficient equal to 0.7835 and significance of 99.99%. In this way, the response surfaces and the quadratic models were constructed for each iteration of sets (MQn), where n is the number of experiments. Those response surfaces were constructed according to the method described in section 4.2.1.

The response surfaces and quadratic models (QMn) of each iteration of sets (QM7, QM10, QM13, QM16, QM19, QM22, QM25, QM28 and QM31) can be seen in Figures 13 to 21.

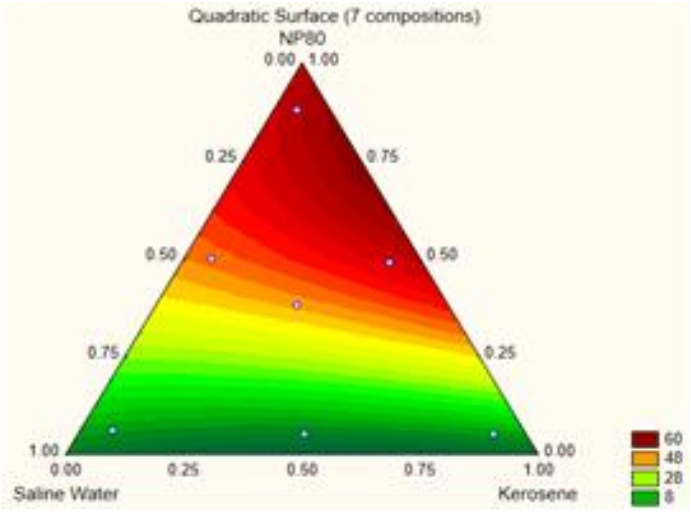

Figure 13 - A) Surface area and quadratic model (QM7) of the first set of points selected, for a total of 7 compositions.

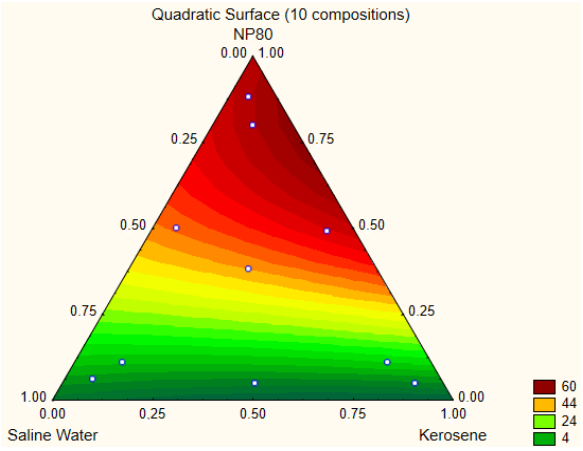

Figure 14 - Surface area and quadratic model (QM10) of the second set of points selected, for a total of 10 compositions.

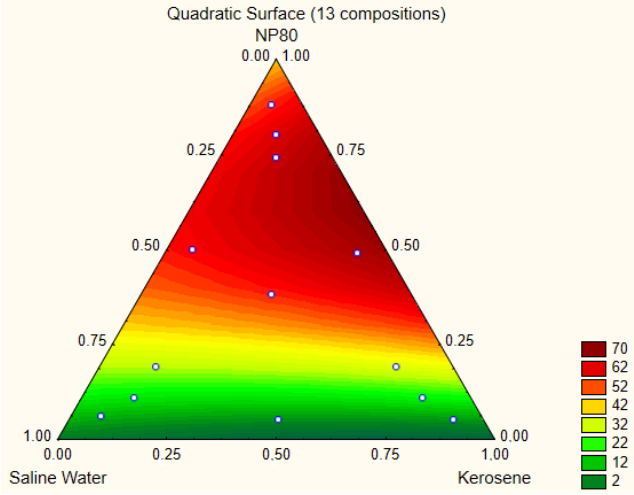

Figure 15 - Surface area and quadratic model (QM13) of the third set of points selected, for a total of 13 compositions.

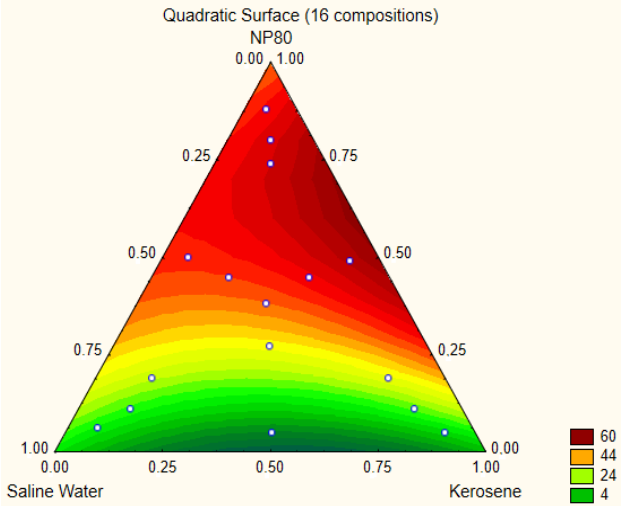

Figure 16 - Surface area and quadratic model (QM16) of the fourth set of points selected, for a total of 16 compositions.

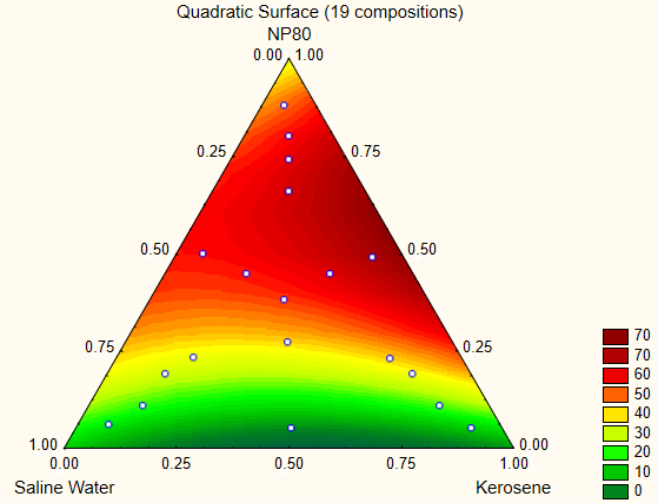

Figure 17 - Surface area and quadratic model (QM19) of the fifth set of points selected, for a total of 19 compositions.

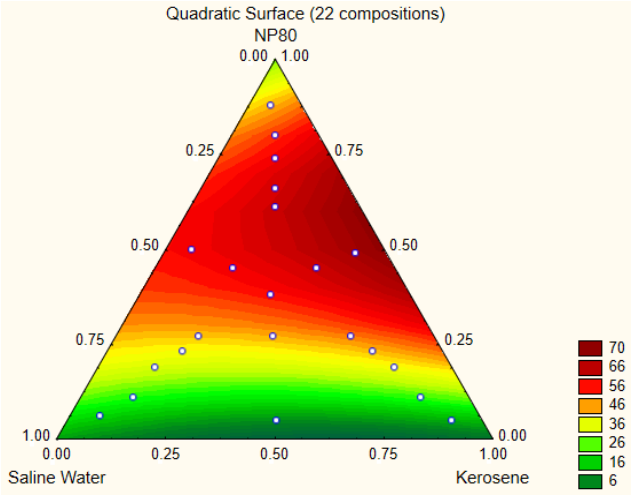

Figure 18 - Surface area and quadratic model (QM22) of the sixth set of points selected, for a total of 22 compositions.

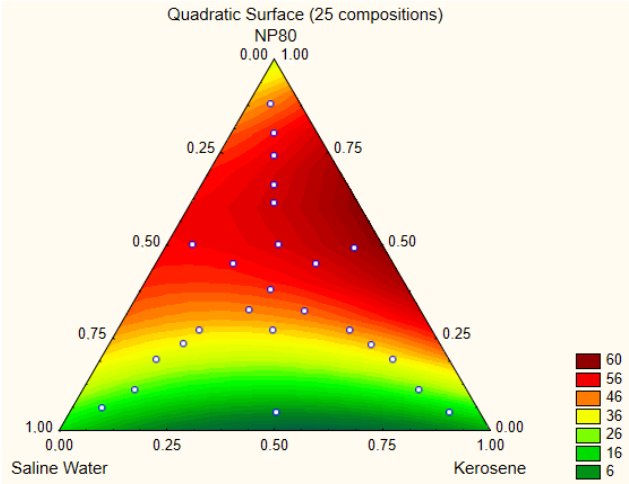

Figure 19 - Surface area and quadratic model (QM25) of the seventh set of points selected, for a total of 25 compositions.

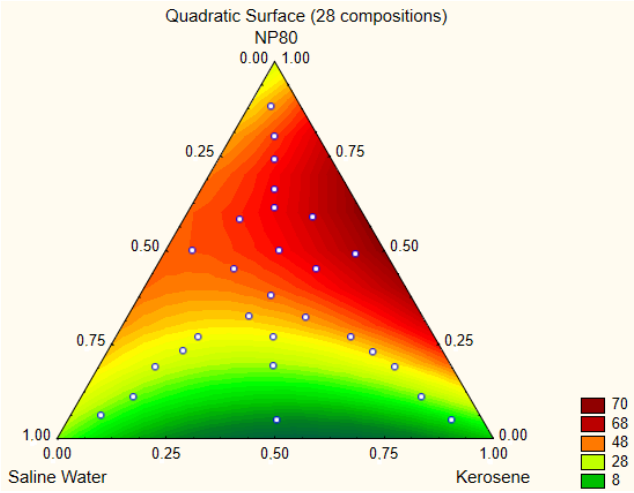

Figure 20 - Surface area and quadratic model (QM28) of the eighth set of points selected, for a total of 28 compositions.

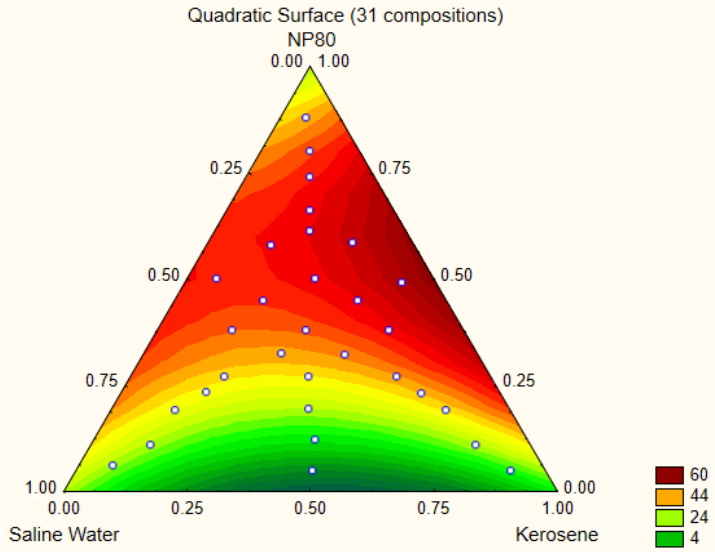

Figure 21 - Surface area and quadratic model (QM31) of the ninth set of points selected, for a total of 31 compositions.

The quadratic mathematical models generated for each set of points of the kerosene/NP80/salt water system are shown below.

7 points	F(w,k,s) = -5.7641*w-8.5569*k+69.7019*s-25.0382*w*k+63.6950*w*s+156.7282*k*s	(4)
10 points	F(w,k,s) = -9.8601*w-15.3241*k+69.6280*s-12.2457*w*k+68.0464*w*s+160.6213*k*s	(5)
13 points	F(w,k,s) = -1.4228*w-7.8237*k+41.8857*s-13.0401*w*k+141.9993*w*s+222.1535*k*s	(6)
16	F(w,k,s) = 9.4250*w+0.6156*k+46.7954*s-	(7)

points	99.1563*w*k+104.9553*w*s+208.1180*k*s	
19 points	F(w,k,s) = 10.4792*w+1.9110*k+33.9709*s-69.5254*w*k+134.5303*w*s+227.9118*k*s	(8)
22 points	F(w,k,s) = 8.6669*w+0.5628*k+25.9313*s-35.5092*w*k+154.2043*w*s+238.8159*k*s	(9)
25 points	F(w,k,s) = 15.0861*w+7.2160*k+30.9648*s-66.3720*w*k+123.5392*w*s+202.0923*k*s	(10)
28 points	F(w,k,s) = 23.4047*w+8.5009*k+30.4454*s-111.6200*w*k+91.6860*w*s+225.8215*k*s	(11)
31 points	F(w,k,s) = 24.1887*w+10.8432*k+23.7248*s-148.5607*w*k+130.1521*w*s+244.7265*k*s	(12)

3.2.4. Comparison between quadratic models of sets tested with each iteration of sets (QMn) and the quadratic model with 99 formulations (QM99).

The predicted values of each quadratic model for each iteration of sets of points (QMnp) were compared with those calculated by the quadratic model with 99 points (QM99p). This comparison predicted values between (QMnp) and (QM99p) and allowed calculating the root mean squared error (RMSE). Thus, we evaluated which quadratic model with each iteration of sets of formulations presented the lowest error. Table 4 below, for example, shows the values predicted by the quadratic model with 7 points (QM7pred) - the sixth column from left to right.

The equation of the quadratic model with 7 points is represented by the fourth equation described above, that is, F(w,k,s) = -5.7641*w-8.5569*k+69.7019*s-25.0382*w*k+63.695*w*s+156.7282*k*s), where the independent variables x, y and z of the quadratic model were replaced by the values of salt water (w), kerosene (k) and surfactant - NP80 (s) and dependent variable F(w,k,s) is the prediction of the values assigned to the formulations.

Table 4 - Calculated values predicted by the quadratic model of 7 points and quadratic model of 99 points with and without correction.

7pts	Aqueous Phase (xw	Oil Phase (k)	Surfactant (s)	Assigned value	QM7p	QM99p	RMSE (QM7p x QM99p)	QM99pc	RMSE (QM7pc x QM99pc)
1	0.07	0.05	0.88	70	71.24	19.69	28.44	19.69	22.22
2	0.44	0.06	0.5	55	49.86	61.39		61.39	
3	0.87	0.07	0.06	0	1.03	35.61		35.61	
4	0.47	0.48	0.05	0	-3.72	-23.99		0	
5	0.07	0.88	0.05	0	1.13	31.01		31.01	
6	0.07	0.44	0.49	70	65.19	72.46		70	
7	0.32	0.3	0.38	35	45.28	46.60		46.60	

The predicted values of each quadratic model, with each iteration of sets - MQn were calculated by substituting the variables x, y and z for the values of salt water - w, kerosene - k and surfactant (NP80 - s, respectively. The same was done for the quadratic model with 99 points (MQ99), which is represented by the last equation of Table 5: F (a, q, t) = 49.3673*w + 41.4229*k - 10.0471*s-344.8685*w*k + 195.5358*w*s + 276.5345*k*s.

The root mean square error (RMSE) was obtained according to the method calculated by the Qtiplot program. The values predicted by the quadratic model with the smallest and largest number of points, (MQnp) and (MQ99p), were inserted into the program and the linear adjustment was made.

As can be seen in Table 6, the RMSE was calculated in two ways: quadratic model with and without correction. In the case of MQn and MQ99 without correction, the predicted values below and above the assigned values of importance were considered, that is, between 0 - 70%, depending on the formulation. The RMSE was 28.44 (Mg7p x Mg99p). In the case of MQnpc and MQ99pc, the RMSE ranged from 28.44 to 22.22 (MQ7pc x MQ99pc). When the predicted value of the quadratic model was corrected, the RMSE declined, indicating that the estimated error around the observed values was minimized.

The quadratic models were aligned (corrected) with minimum and maximum values, so that the values predicted by quadratic models below zero or greater than 70 were set with values of 0 and 70%, respectively. The alignment of the predicted values was set with the values of minimum (zero) and maximum (seventy) for all quadratic models of this work.

Thus, all RMSE values were calculated based on the values predicted by quadratic models with each iteration of sets (MQnpc) versus the values predicted by the corrected quadratic model of 99 points (MQ99pc). Table 5 shows the RMSE values of each set of points with 99.99% reliability and Table 6 shows the RMSE for the set of 99 points that was calculated between the experimentally observed data versus the values predicted by MQ99pc.

Table 5 - Results of the root mean squared error (RMSE) for the sets of points.

N = no. of points	RMSE (QMnpc) X (QM99pc)
7	22.23
10	18.55
13	14.13
16	12.75
19	10.50
22	9.27
25	8.71
28	8.13
31	7.51

Table 6 - Results of root mean square error (RMSE) for sets of 99 points.

Points	RMSE (Assigned Values) x $QM99_{pc}$)
99	3.57

.

To identify the number of experiments capable of representing significantly the phase diagram with 99 formulations, a first-order exponential decay curve was generated in the Qtiplot program, plotting the RMSE value of each set points versus the number of experiments. By means of extrapolation lines as described in standard ISO11358 [28], an estimate was made of the number of experiments required to represent the phase diagram of empirical observations significantly. Figure 22 shows the first-order exponential decay curve. The estimated number of formulations was 31 after extrapolation of the lines, which was confirmed by the amplitude results generated by the program after the plot of this curve. The results showed that the quadratic model with 31 points was able to adequately describe the phase diagram of the kerosene/NP80/ saline system containing the 99 empirical observations. The new method for the construction of the phase diagram developed this work showed that with the increase of the number of experiments, the RMSE declined, that is, the quadratic model containing 31 points (MQ31) presented good adjustment of the empirical observations. In addition, the result of the amplitude was equal

to 3.11x10[1], indicating that additional gains in fit adjustment would require 31 more experiments, that is, the RMSE would decrease significantly after 62 experiments, which would be experimentally impractical.

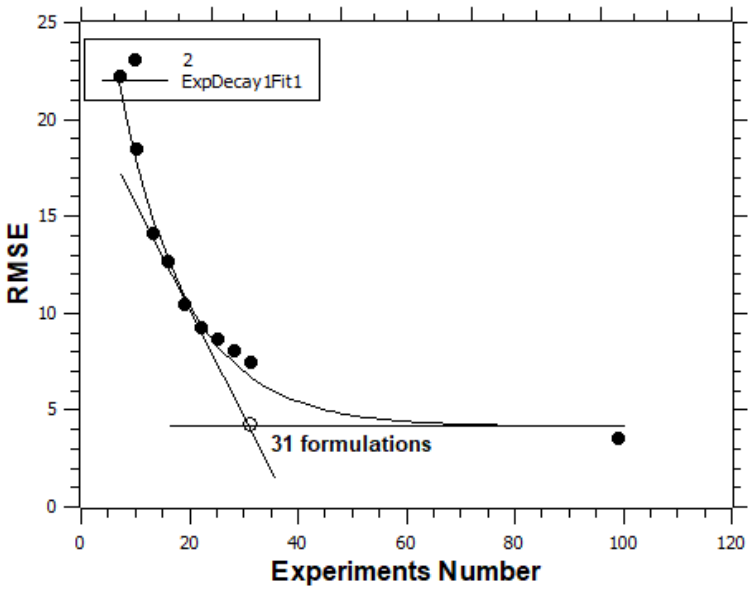

Figure 22 - First-order exponential decay curve.

For show the significance of the reduced method was compared the RMSE of the reduced set and the enlarged set. Table 8 shows the RMSE for the set of 31 and 99 points that was calculated between the experimentally observed data versus the values Predicted by MQ31pc [29-31]. Both sets presented the approximately the same RMSE, that is, the method of reduced formulations is excellent and reproduces very well the enlarged set. Therefore, the latter is unnecessary.

Table 7 - Results of root mean square error (RMSE) for sets of 31 and 99 points.

Sets/model	Assigned value of the sets 31 compositions x MQpc31	Assigned value of the sets 99 compositions x MQpc31
RMSE	13.62	13.50

4. MODEL VALIDATION

To validate the quadratic model with 31 points, five compositions were chosen in the region of the liquid microemulsion and reproduced in the laboratory, to test a specific region of the model to confirm the expected result. The formulations were three oil-in-water types and two water-in-oil types. At the end of the experiments, all the formulations were microemulsions (Figure 23).

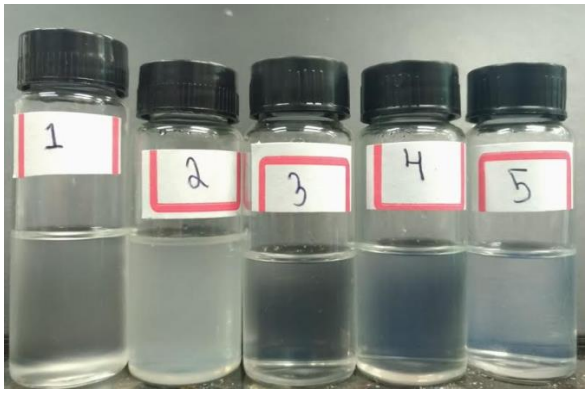

Figure 23 - Formulations in regions of liquid oil-in-water (points 1,2,3) and water-in-oil (points 4, 5) microemulsions.

Table 7 shows the values predicted by the quadratic model with 5 points chosen in the region of the liquid microemulsions, along with the percentage of the aqueous phase, oil phase, surfactant and the p-value predicted by the quadratic model with 31 points. The results show that the prediction of the quadratic model was in the range of 30 to 70%. The quadratic model was able to predict the formulations of the liquid microemulsions and showed a lower RMSE.

Table 7 - Calculated values predicted by the quadratic model with 5 formulations to validate the quadratic model of 31 points.

Points	Saline Water	Kerosene	NP80	Prediction attributed value
1	0.13	0.30	0.57	65.62
2	0.13	0.30	0.37	59.22
3	0.60	0.13	0.27	40.42
4	0.70	0.09	0.21	37.29
5	0.77	0.07	0.16	33.95

5. CONCLUSIONS

The results obtained were encouraging, since the quadratic regression model, estimated at 99.99% significance level, explained approximately 61% of the variability of the observed data. Besides this, the quadratic model containing 31 formulations was able to represent / map a phase diagram with a reduced number of formulations and with a degree of significance of 99.99%. In other words, by calculating the root mean squared error (RMSE) and the extrapolation procedure (ISO11358), the regions of microemulsions, liquid or viscous macroemulsions, gels or phase separation can be identified in a shorter time and with fewer experiments.

ACKNOWLEDGMENTS

This study was financed in part by the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES) - Finance Code 001. This work was also supported in part by Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPQ) and by Fundação Carlos Chagas Filho de Amparo à Pesquisa do Estado do Rio de Janeiro (FAPERJ).

Sample CRediT author statement

XXXXXXX

REFERENCES

1. SOLANS, C.; KUNIEDA, H. Industrial Applicantions of Microemulsions. New York: Marcel Dekker, INC, 1997. 401p.

2. SCHULZ, P.A. Nanomateriais e a interface entre nanotecnologia e ambiente. Visa em debate, São Paulo, p. 53-58, nov. 2013.

3. BERA, A.; OJHA, K.; KUMAR, T.; MANDAL, A. Water solubilization capacity, interfacial compositions and thermodynamic parameters of anionic and cationic microemulsions. **Colloids and Surfaces A: Physicochemical Engineering Aspects**, v. 404, p.70 – 77, 2012.

4. BERA, A.; KUMAR, T.; OJHA, K.; MANDAL, A. Screening of microemulsion properties for application in enhanced oil Recovery. **Fuel**, v.121, p.198–207, 2014.

5. PAUL, B.K.; MOULIK, S.P., Microemulsions: an overview. **Journal Dispersion Science Technology**. V.18:4, p.301-367, 1997.

6. SILVA, J. D.; SILVA, Y. P.; PIATINICKI, C. M. S.; BOCKEL, W. J.; MENDONÇA, C. R. B. Microemulsions: Componentes, características, potencialidades em química de alimentos e outras aplicações. **Quimica Nova***, v.* 38, n. 9, 1196-1206, 2015.

7. DEL GAUDIO, et al. Process for the recovery of heavy oil from an underground reservoir. US n. 2012/0261120 A1, 3 jul. 2012, 18 out. 2012.

8. MANDAL, S.; GHOSH, S.; BANERJEE, C.; KUCHLYAN, J.; BANIK,D.; SARKAR,N. A Novel Ionic Liquid-in-Oil Microemulsion Composed of Biologically, Acceptable Components: An Excitation Wavelength Dependent Fluorescence Resonance Energy Transfer Study, **Journal of Physical Chemistry B**, v. 117, p. 3221−3231, 2013 DOI: 10.1021/jp4009515

9. BARRADAS, T. N.; CAMPOS, V. E. B.; SENNA, J. P.; COUTINHO, C. S. C.; TEBALDI, B. S.; SILVA, K. G. H.; MANSUR, C. R. E. Development and characterization of promising o/w nanoemulsions containing sweet fennel essential oil and non-ionic sufactants. **Colloids and Surfaces A: Physicochemical Engineering Aspects**, v. 480, p. 214–221, 2015.

10. Izabel C. V. M. Santos; Raquel R. Martelloti; Priscila F. Oliveira; Claudia R. E. Mansur. Development of microemulsions to reduce the viscocity of crude oil emulsions. **Fuel**, Volume 210, 684-694 (2017). DOI: https://doi.org/10.1016/j.fuel.2017.08.088

11. KILPATRICK P. K.; SULLIVAN A. P.; The effects of inorganic solid particles on water and crude oil emulsion stability. **Industrial & Engineering Chemistry Research**, v. 41, p. 3389-3404, 2002. DOI 10.1021/ie010927n

12. PORRAS M.; SOLANS C.; GONZÁLEZ C.; GUTIÉRREZ J. M.; Properties of water-in-oil (w/o) nano-emulsions prepared by a low-energy emulsification method. **Colloids and Surfaces A: Physicochemical and Engineering Aspects**, Barcelona, v. 324, p 181 – 188, 2008. DOI: 10.1016/j.colsurfa.2008.04.012

13. ALBUQUERQUE, H. S.; VALE, T. Y. F.; DANTAS, T. N. de C.; NETO, A. A. D.; SANTANNA, V. C.; COELHO, T. A. M.; Estudo da eficiência de microemulsified systems na recuperação avançada de petrol. **4º PDPETRO**, 2007

14. WINSOR, P.A. Hydrotropy, solubilization,and related emulsification processes. VIII. Effect of constitution on amphiphilic properties. **Transactions of the Faraday Society**, v. 44, p. 463-471, 1948. DOI: 10.1039/TF9484400451

15. OLIVEIRA, A. G.; SCARPA, M. V.; CORREA, M. A.; CERA, F. R.; FORMARIZ, T. P. Microemulsions: estrutura e aplicações como sistema de liberação de fármacos. **Quimica Nova***,* v.27, n. 1, 131-138, 2004.

16 FORMARIZ, T. P.; URBAN, M. C.C.; JÚNIOR, A. A. S.; GERMIÃO, M. P. D.; OLIVEIRA, A. G. Microemulsions e fases líquidas cristalinas como sistemas de liberação de fármacos. **Revista Brasileira de Ciências Farmacêuticas**, v. 41, n. 3, jul./set., 2005.

17. OLIVEIRA, R. B.; LIMA, E. M. Polímeros na obtenção de sistemas de liberação de fármacos. **Revista Eletronica de Farmácia**, 3, 1, 29-35, 2006.

18. SCHULMAN, J. H., STOEKENIUS, W., PRINCE, L. M. Mechanism of formation and structure of microemulsions by electron microscopy. **Journal of Physical Chemistry**., v. 63, p. 1677–1680, 1959.

[19. MALIK, M. A.; WANI, M- Y.; HASHIM, M. A. Microemulsion method: A novel route to synthesize organic and inorganic nano materials. **Arabian Journal of Chemistry**, v. 5, p. 397–417, 2012.

20. SILVA, B.J.; MENEZES, R.R., SANTANA, L.N.L.; MELO, L.R.L.; NEVES, G.A.; FERREIRA, H.C. Uso de técnicas estatísticas para modelar a resistência à flexão de corpos cerâmicos contendo resíduo de granito. **Revista Matéria**, v. 17, n. 1, pp. 919 – 930, 2012.

21. LUNA, A. S.; LIMA, I. C. A.; HENRIQUES, C. A.; ARAÚJO, L. R. R.; ROCHA, W. F.; SILVA, J. V. Prediction of fatty methyl esters and physical properties of soybean oil/biodiesel blends from near and mid-infrared spectra using the data fusion strategy. **Analytical Methods**, 2017. DOI: 10.1039/C7AY01638G,

22. LEGATES, D. R. Evaluating the use of "goodness-of-fit" measures in hydrologic and hydroclimatic model validation. **Water resources research**, vol. 35, no. 1, pages 233-241, january 1999.

23. GUJARATI, D. N.; PORTER, D.C. Econometria Básica. 4° edição, AMGH, 2011.

24. FERRREIRA, L. P.; MOREIRA, A. N.; DELAZARE, T.; OLIVEIRA, G. E.; SOUZA, F. G. Petroleum Absorbers Based on CNSL, Furfural and Lignin – The Effect of the Chemical Similarity on the Interactions among Petroleum and Bioresins. **Macromolecular Simposia,** v. 319, p. 210–221, 2012.

25. WEHRENS, R. Chemometrics with R Multivariate Data Analysis in the Natural Sciences and Life Sciences. 2011

26. LIRA, M. A. T.; SILVA, E.M.; ALVES, J. M.B. M. Estimativa dos recursos eólicos no litoral cearense usando a teoria da regressão linear. **Revista Brasileira de Meteorologia**, v.26, n.3, 349 - 366, 2011;

27. BROWNLEE, K. A.: Statistical Theory and Methodology in Science and Engineering. John Wiley & Sons, New York 1965, 590 S., 70 Abb., Tafelanhang. Disponível em: <https://onlinelibrary.wiley.com/doi/abs/10.1002/bimj.19680100225>Acesso em: 15 de Dezembro de 2018.

28. INTERNATIONAL ORGANIZATION FOR STANDARDIZATION ISO11358-1:2014. Plastics – Thermogravimetry (TG) of polymers – Part 1: General principles. Disponível em: <https://www.iso.org/standard/59710.html>. Acesso em: 16 de Dezembro de 2018.

29. JIANG, J.; CHEN, Y.; CAO, J.; MEI, C. Improved Hydrophobicity and Dimensional Stability of Wood Treated with Para_n/Acrylate Compound Emulsion through Response Surface Methodology Optimization. **Polymers**, ,v. 12, n. 86, 2020. doi:10.3390/polym12010086.

30. HUSAIN, A.; ADEWUNMI, A. A.; MAHMOUD, M., KAMAL, M. S., AL-HARTHI, M. A. Stability of Diesel/Water Emulsions: Experimental and Modeling Investigations. **Journal of Energy Resources Technology**, v.143, n.11, November 2021. doi.org/10.1115/1.4049606.

31. RASHID, T., TAQVI, S. A. A.; SHER, F.; RUBAB, S.; THANABALAN, M.; BILAL, M.; ISALM, B. Enhanced lignin extraction and optimization form oil palm biomass using neural network modelling. **Fuel**, v. 293, June 2021. https://doi.org/10.1016/j.fuel.2021.120485.