Oct 2013. Vol. 4, No. 4 ISSN2305-8269
International Journal of Engineering and Applied Sciences

www.eaas-journal.org

COMPUTED-AIDED AND MULTIVARIATE REGRESSION MODEL OF DIGITAL THERMOGRAMS FOR DETECTION OF BREAST TUMOURS

José Luis Ruiz-Duarte[1], Carlos Anaya-Eredias[2], Marcelino Barboza-Flores[3], Ricardo Alberto Rodríguez-Carvajal[4]
[1, 2, 3, 4]Universidad de Sonora, Luis Encinas S/N, C. P. 83000, Hermosillo, Sonora, México,
[1, 2, 4]Departamento de Ingeniería Industrial, [3]Departamento de Investigación en Física.
[1, 2, 4]Tel.: +52 (662) 259 21 59, Fax: +52 (662) 259 21 60.
[3]Tel.: +52 (662) 259 21 56, Fax: +52 (662) 212 66 49.
[1]jduarte@industrial.uson.mx, [2]canaya@industrial.uson.mx, [3]mbarboza@cifus.uson.mx,
[4]ricardo@industrial.uson.mx

ABSTRACT

Infrared Thermography has emerged as a complement for X-Ray mammography in the detection of breast cancer. This article reviews the research on various developed systems for thermogram data analysis; and, based on literature scrutiny and basic techniques of inferential and multivariate statistics tools, a multivariate model for analysing thermal images is proposed. The model tests the hypothesis in which a person has higher probability to undergo a cancer related physiological condition if specific threshold temperature does not reach a regression threshold temperature.

Keywords*: Infrared thermography, multivariate analysis, inferential statistics, breast cancer.*

1. INTRODUCTION

Breast cancer has become the second cause of death in women from 34 to 54 years old in Mexico [1], and the second cause of death of women in the world [2]. For its detection, mammography is the gold standard [3], and has contributed for the reduction of breast cancer mortality among 50 year old women, from 25% to 30% [4]. However, mammography has low sensitivity in dense glandular tissue [5], as it has an average of 75% of false positive results (diagnosing illness on a healthy patient), a range from 4% to 34% of false negative results (not diagnosing cancer when it exists) and requires compression and radiation exposure [4]. According to Harris (2009) [6], the risk of developing breast cancer increases 1.5 times in congenital predisposed women and 2.5 times in women that are younger than 20 years old, after being exposed more than five times to radiation.

There are other complementary techniques useful in cancer detection. Ultrasound detects breast lesions and it's useful for finding palpable and not palpable breast anomalies. It has improved detection when used together with mammography [2]. Gamma Ray Scintimammography has been proposed as an ideal complement for X ray mammography and ultrasound because of its ability of evaluate the nature of breast lesions [7]. Microwave imaging is based on recovery of tissue's electric properties, and the dielectric contrast that exists between malign and normal tissue [8]. Electric impedance imaging consists in measuring the ability of tissues for conducting and storing electricity [9]; it has been demonstrated that tumours differs from surrounding tissue [10]. Studies of elastography are based on the premise in which cancerous tissue can be 5 to 20 times harder to bend that glandular adipose tissue [11].

Another complementary option is infrared thermography, which has been utilized in

ISSN2305-8269
International Journal of Engineering and Applied Sciences

www.eaas-journal.org

applications like oncology, vascular disorders, arthritis, rheumatism, and others [12]. This was approved by the US's Food and Drugs Administration in 1982 as a complement for mammography for the detection of breast cancer [4].

Below we discuss these methods and provide a theoretical background, followed by a literature review in practical cases focused on detecting breast cancer. Afterwards, a proposed model for this research is introduced and a pertinent discussion and conclusion is provided.

2. BACKGROUND

Thermography is a non-invasive imaging method that is economic, fast, and harmless for the patient [13]. It can be used repeatedly without harming the patient and can detect subtle physiological changes that indicate a variety of breast diseases and abnormalities [3]. Thermography by its own it's not able for diagnosis generation, but is useful and complementary for diagnosis using other methods [13]. When infrared thermography is used in a multimodal approach [clinic test, mammography and thermography] it can detect up to 95% of cases of early cancer [12].

Thermography in breast cancer detection is based in the principle that if a breast tumour exists, it will use more nutrients that the surrounding tissue to support its growth. As a result, metabolism increases, hence temperature and its transference in all directions [13] increases. A small, but high temperature tumour can indicate a rapid growing of it [4]. Amri et al (2001) [14] research indicates that skin temperature over a malign tumour is frequently over 1 and 3° Celsius warmer than the surrounding skin. Likewise, a tumour can cause changes on the temperature distribution, causing thermal asymmetry between breasts [3]. Also, it is supposed that human body has thermal symmetry, so that thermal variations in both breasts should be similar [15].

The utilization of thermography allows the detection of cancer up to 10 years before other imaging methods, evaluate treatment efficiency and predict patient condition. About 44% of patients with abnormal thermograms can develop cancer within the next five years. About 70% of cases, breast cancer can be detected a year earlier than via mammography [16].

A thermogram is a thermal map of a heat radiating body surface [17]. An abnormal thermogram has proven to be a reliable indicator of high risk of breast cancer in early stages [5].

According to Jiang et al. (2005) [12], infrared image processing consists in the application of image processing techniques, for example, resolution improvement, noise reduction.

According to Gonzalez et al. (2004) [18], a digital image is a two-dimensional function $f(x,y)$ where x and y are spatial coordinates and the amplitude of f in each pair of coordinates (x,y) represents the grey level intensity. The values of image intensity depends if it's black and white, grey scaled or colour. For black and white images, $f(x,y)$ is binary, where 0 is black and 1 is white. For grey scaled images, $f(x,y)$ can take values from 0 [black] to 255 [white]. For coloured images, $f(x,y)$ is a vector with three individual components for red, green and blue, with values from 0 to 255 for each component [19]. A pixel is a finite element of an image that has a location and a specific value. Resolution is the number of pixels that form an image, written in an array [Row, Columns] [18].

Image segmentation refers to a pixel classification in background pixels and signal pixels [20]. Once the segmentation has been done, the next stage is the extraction of characteristics, in which measurements of interest are calculated for the image segmented regions: area, volume, mean, temperature. [20].

Inferential statistics allow obtaining conclusions about a population based on a sample [21]. In many problems, according to Montgomery and Runger (2003) [22], it's required to make a decision about whether accepting or not some assertion about a parameter of one or more populations. This is known as Hypothesis Test. A Null Hypothesis (H_0) is the one that will be tested. Rejecting the Null Hypothesis always lead to accepting an Alternative Hypothesis (H_1). A Type I Error consists on rejecting H_0 when it's true. A Type II Error consists on not rejecting H_0 when it's false. The probability of committing a type I error is known as significance level and it's denoted as α.

A special case of hypothesis test is when the underlying distribution of a population is unknown and it's needed to test the null hypothesis

International Journal of Engineering and Applied Sciences

www.eaas-journal.org

that a specific distribution fits appropriately to the population, what is done with the statistic Chi-Squared, and it's called the Chi-Squared test [22].

According to Peña (2002) [23], in order to describe any real situation, it is required to analyse several variables simultaneously. Multivariate data analysis studies multiple variables on each element of a specific population. The multivariate data can be presented on an *n* –sample size- by *p* –number of variables- array. The means vector is a *1* by *p* array that contains each mean of every variable analysed. The covariance matrix contains the variance of each variable in its principal diagonal, and the covariance between two variables in the other cells. The precision matrix –calculated as the inverse of the covariance matrix- contains information about the multivariate regression coefficient between each variable and the others.

A vector **X** is *k*-dimensional normal distributed if its density function is that $f(x) = |\mathbf{V}|^{-½}(2\pi)^{-p/2}\exp\{-(½)(\mathbf{x}-\boldsymbol{\mu})'\mathbf{V}^{-1}(\mathbf{x}-\boldsymbol{\mu})\}$; where **V** represents the covariance matrix, and **μ** its only maximum. The *k*-dimensional normal distribution has, among others, the following properties:

- The distribution is symmetric around **μ**.
- Its marginal distributions are normal distributed.
- If it is 'cut' with hyper planes that are parallel to the one defined by the variables, it can be obtained level curves that are defined by the Mahalanobis distance equation, which is distributed as a Chi-Squared distribution with *p* degrees of freedom. The Mahalanobis distance equation is $D^2 = (\mathbf{x}-\boldsymbol{\mu})'\mathbf{V}^{-1}(\mathbf{x}-\boldsymbol{\mu})$.

Sampling is a set of statistical techniques that study the way of selecting a representative sample of a population whose information allow infer in its parameters. In the multivariate distribution case, Hair Jr. et al. (2009) [24] recommend to utilize a sample size that, in combination with the effect size -considered in their work as the correlation between variables-, and the significance level, α, reach a relatively high power - at least 80%. According to his research, with a sample size of 100 or higher, an acceptable power level can be obtained with a moderate effect level of 0.5 or higher.

A manual analysis of medical studies can be inefficient and human error susceptible [25]. Hence, in real situations it's recommendable to use software in order of automatizing some analysis processes, obtaining a faster performance and reducing human error.

3. REVIEW

In an interesting research, Ng et al. (2001a) [3] obtained statistical measures of breast of people with and without the illness; with that they made a Hypothesis Test in order to know if there are significant difference between breast with and without cancer of a sick person, and breasts with cancer and breasts of healthy persons. They have found that there is no significant difference between the temperature mean in any case.

Focusing in the part of image processing, Ng et al. (2001b) [26] developed some algorithms that allow minimizing no real results on thermograms.

In their research, Ng et al. (2002) [15], shows the utilization of artificial neural networks. They obtained a network that had correct diagnosis in 61.54% of the cases. This network has as inputs the mean, median and mode of the breast temperature, also includes age, family history, hormone replacement therapy, age of menarche, presence of palpable lump, previous biopsy or surgery, presence of nipple discharge, breast pain, menopause over age 50, first child over age 30.

In their research, FOK et al. (2002) [25] utilized a combination of thermograms and artificial neural networks, in which image processing techniques could be applied to the images before enter them into the network, and a second neuron in which physiological data is utilized instead of thermograms. They have found that a thermogram based network has good capacity to predict people with cancer, but weren't so good identifying people without it. The opposite occurred in the other network. Besides, Ng and Kee (2008) [13] mention a possible application of artificial neural networks and bio statistical methods, such as utilizing lineal regression for selecting the most important variables when designing a neural network that decides whether a person has cancer or not. This technique obtained 80.95% of global accuracy, with 100% of sensitivity and 70.6% of specificity.

International Journal of Engineering and Applied Sciences

www.eaas-journal.org

Another study is the one of Tan et al. (2009) [27], in which, via the application of neural networks they get the classification of different diagnoses (normal, benign or suspect) with over 80% of accuracy. Meanwhile, Köşüş et al. (2010) [4] compared conventional mammography, digital mammography and digital thermal imaging. They found that screen film mammography has a range from 4% to 34% of false negatives and, in the cases considered as positives, 75% were discarded by biopsy. In their work, Moghbel and Mashohor (2011) [16] make a review of the application of Computer Assisted Detection, based on neural networks and fuzzy logic, finding out that the best result is obtained with a combination of both techniques, obtaining over 90% of global accuracy in the diagnosis.

Amri et al. (2011) [14] utilized the lineal function matrix to determine the behaviour of tumour of various sizes and breast depths, and to evaluate sensitivity levels, via computer simulation. They have found that breast surface temperature significantly increases in depths of approximately 20mm, and noticed that tumour size and location don't affect significantly, according to their model. Kapoor et al. (2012) [28] utilize a computational approach for the detection of tumours based on thermic images, utilizing an automatized method to identify the Region of Interest, developing a methodology for that purpose, in which they use image pre-processing operations in the thermograms, in addition of the development of an automatized system for detecting the Regions of Interest and image segmentation. Afterwards they obtain information of each pixel, and collect the necessary statistical data. From this, they apply a statistical analysis to obtain new information to detect thermal asymmetry in people breasts. Based in this, they make a classification between a 'normal' and 'abnormal' patient. In Table 1, there is a comparison of the analysed researches and the utilized approaches in each technique.

In this research, a different approach is presented. Determining the maximum temperature of a cancerous breast and the temperature in which it is detected, a multivariate regression model is developed and validated in order to determine certain degree of probability a new case has or not a physiological anomaly.

Table 1. Compared researches.

Research	Informatics Tools	Statistics Tools	Image Processing
Ng, et al. (2001a) [3]		X	
Ng, et al. (2001b) [26]			X
Ng, et al. (2002) [15]	X		
FOK, et al. (2002) [25]	X		X
Ng and Kee (2008) [13]	X	X	
Tan, et al. (2009) [27]	X		
Köşüş, et al. (2010) [4]		X	
Amri, et al. (2011) [14]	X	X	
Moghbel and Mashohor (2011) [16]	X		
Kapoor, et al. (2012) [28]	X	X	X

4. METHODS

The applied methods, as shown in Figure 1, use techniques of inferential statistics, multivariate analysis and image processing for improving the detection, via temperature, of a region that has a high probability of having a cancerous lump.

Since 2009, oncologist Enrique Martin del Campo obtained thermal images from women that were diagnosed as suspicious using mammography, ultrasound or magnetic resonance, in the city of Xalapa, Veracruz, Mexico. Based on the thermal data and clinical history, the oncologist diagnosed each case, according to his experience, pinpointing the area of interest. To obtain the thermal images, the patient was sitting, with their breast exposed, arms held high and both hands behind the head, in a room with an environmental temperature between 20° and 22° Celsius.

About 149 thermal images of these suspicious cases, previously diagnosed, were obtained. 130 of these images were randomly selected for analysis and the other 19 were reserved for validation. The cases are accompanied with a Comma Separated Values [*.csv] file that contains the temperature value for each pixel on an array the same size as the resolution of the digital image of the thermogram.

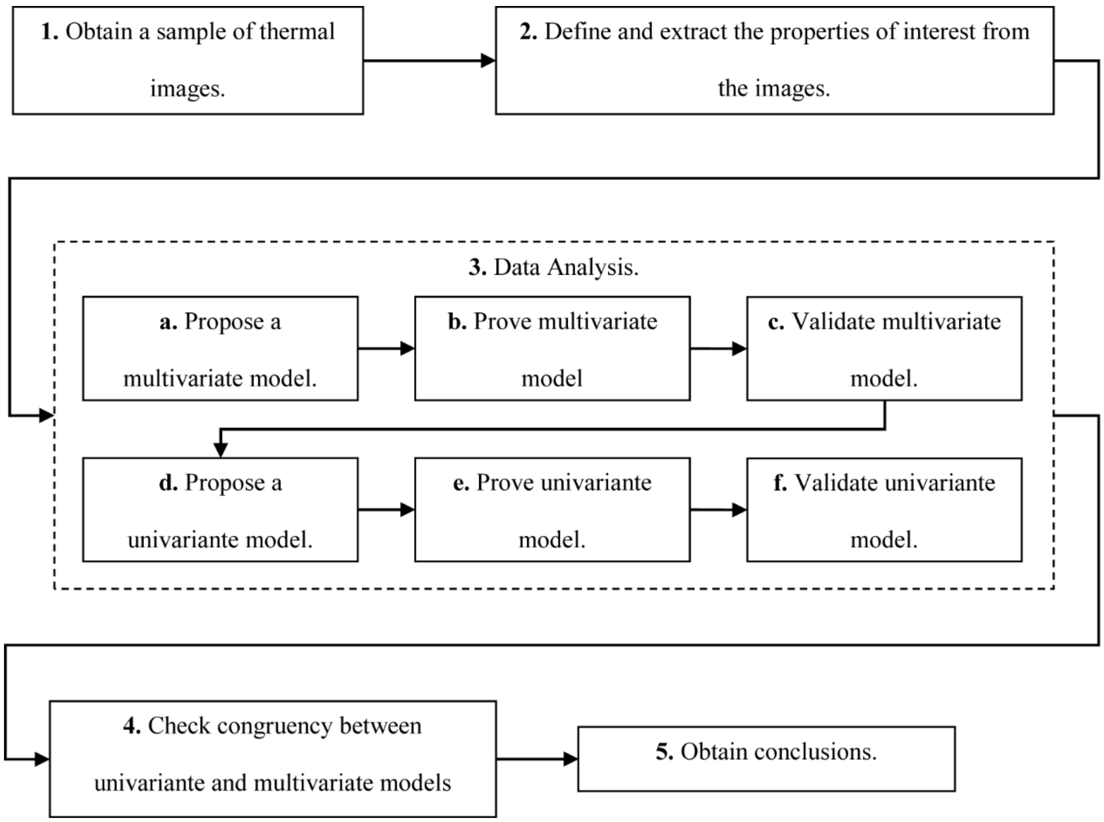

Figure 1. Methods.

There were 2 properties of interest for each image: the maximum temperature of the breast and the threshold temperature which the healthy tissue was not detected but the cancerous tissue could be visible. The maximum temperature was obtained directly from the image, selecting the region of the breast. In order to obtain the second property, an image processing code was used, in which, if the pixel temperature is over a threshold value, the pixel is coloured white, otherwise, it is coloured black, producing a binary image. Then, the diagnosis was checked and, if the area detected by the oncologist was the same that the area coloured white, then the threshold was correctly selected, otherwise, the threshold value had to be adjusted until the white coloured area in the breast was coincident with the diagnosis. An example is shown in the Figure 2.

ISSN2305-8269
International Journal of Engineering and Applied Sciences

www.eaas-journal.org

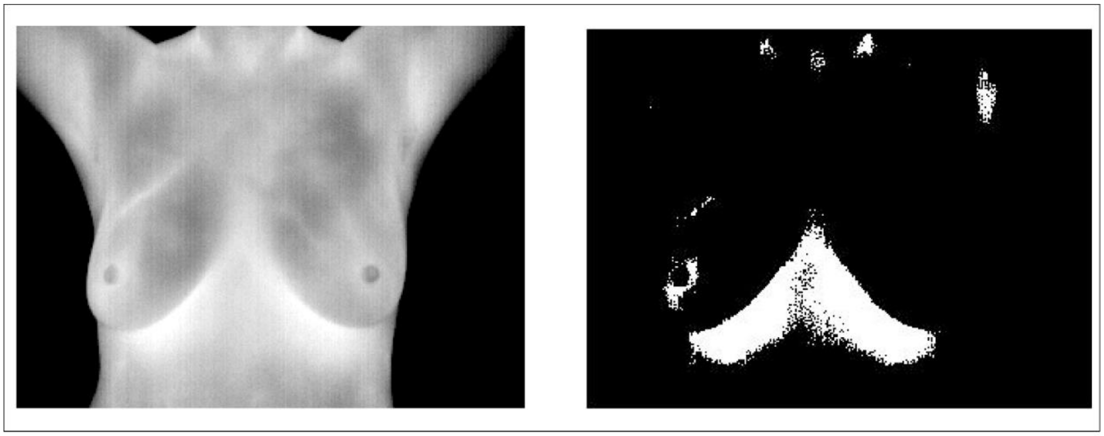

Figure 2. Original image (left) and threshold image (right).

The means vector **μ** obtained from the 130 images is:

$$\mu = \begin{bmatrix} 35.7446 \\ 34.9046 \end{bmatrix}$$

The covariance matrix **S** and the precision matrix $\mathbf{S}^{-1}$ are:

$$S = \begin{bmatrix} 1.2233 & 1.0413 \\ 1.0413 & 1.0541 \end{bmatrix}$$

$$S^{-1} = \begin{bmatrix} 5.1370 & -5.0745 \\ -5.0745 & 5.9615 \end{bmatrix}$$

Once these two properties of the 130 sample data are obtained, the hypothesis that the data is bivariate normally distributed was proved. As shown in Figure 3, where *x* represents the maximum temperature and *y* the threshold temperature for each case, most of the cases are inside the level curve defined as a Chi-Squared distributed level curve defined for 95% of confidence.

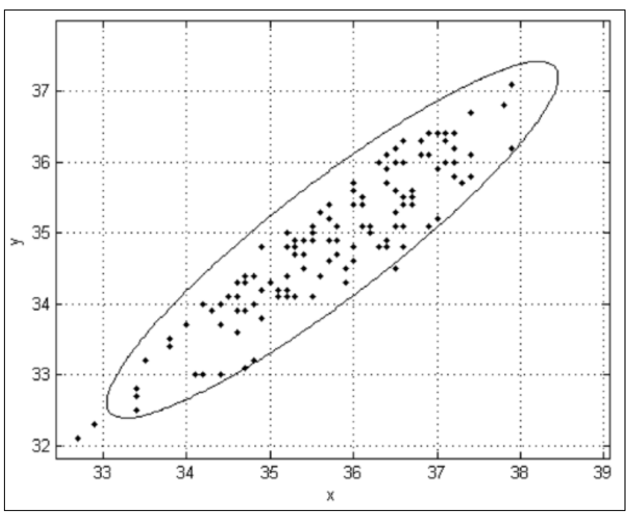

Figure 3. Chi-Squared test for bivariate normal distribution.

With this information we developed a multivariate regression model to predict threshold for new cases, expressed in Equation (1).

$$y = 34.9046 + 0.8512(x - 35.7446) \qquad (1)$$

Oct 2013. Vol. 4, No. 4 ISSN2305-8269
International Journal of Engineering and Applied Sciences

www.eaas-journal.org

Where x represents the maximum breast temperature on a new case, and y represents the threshold temperature obtained with the regression model.

The validation process consists on proving this model on new cases, in order to produce valuable information. As shown in Figure 4, the regression model is a multivariate regression line. It can be supposed that an ideal case of a healthy person, the maximum temperature will be equal to the threshold temperature -these ideal cases are shown as a dashed line in Figure 4-, hence, if the difference of the maximum and threshold temperature is considerably high, it can be supposed that it's a case with high probabilities of having abnormal physiological processes.

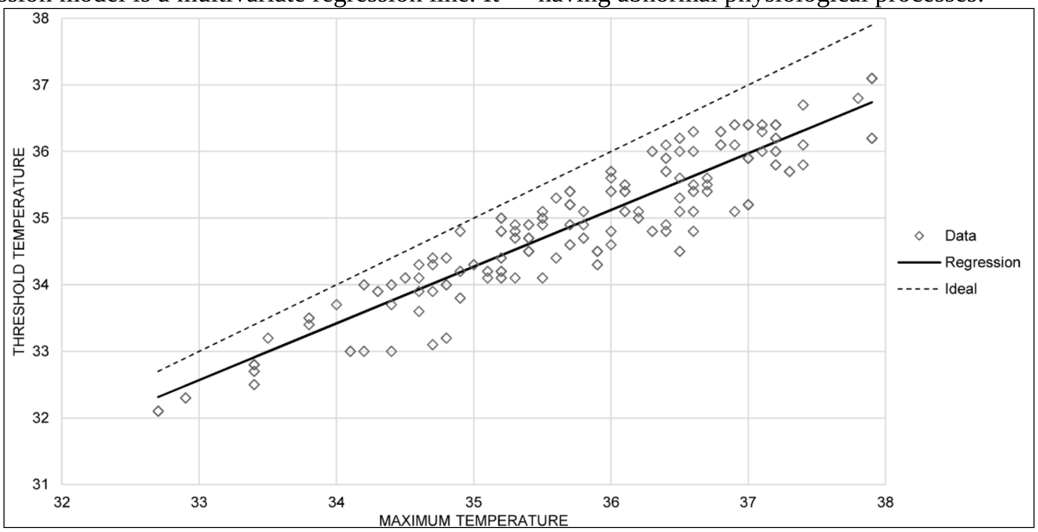

Figure 4. Multivariate regression line.

In order to validate the model, it was calculated the regression threshold temperature (Y), depending on the maximum temperature (X) of the 19 cases reserved for validation. As new cases aren't supposed to be complemented with biopsy, mammography or ultrasound diagnosis, the real threshold temperature (Z) was selected as the minimum of the two maximums temperatures of the breast, supposing that only one breast has an anomaly. This selection of threshold temperature leads to know if there is a peak temperature and thermal asymmetry, if the difference of the two maximum temperatures is considerably high. Thermal anomaly was defined for this research as an area with a temperature higher than the surrounding tissue, which is found in one breast but not in the other one; this represents the presence of a peak temperature and thermal asymmetry.

The hypothesis validated was that, if the adjusted threshold temperature appears over the regression line -closer to the 'ideal' dashed line on figure 4-, it represents a low probability of having an abnormal physiological process. Otherwise, if the adjusted threshold is under the regression line, it represents a high probability of having an abnormal physiological process. As shown in figure 5, some of the validation cases appear over the regression line, and other under it. The sample number appears next to each case point.

ISSN2305-8269
International Journal of Engineering and Applied Sciences

www.eaas-journal.org

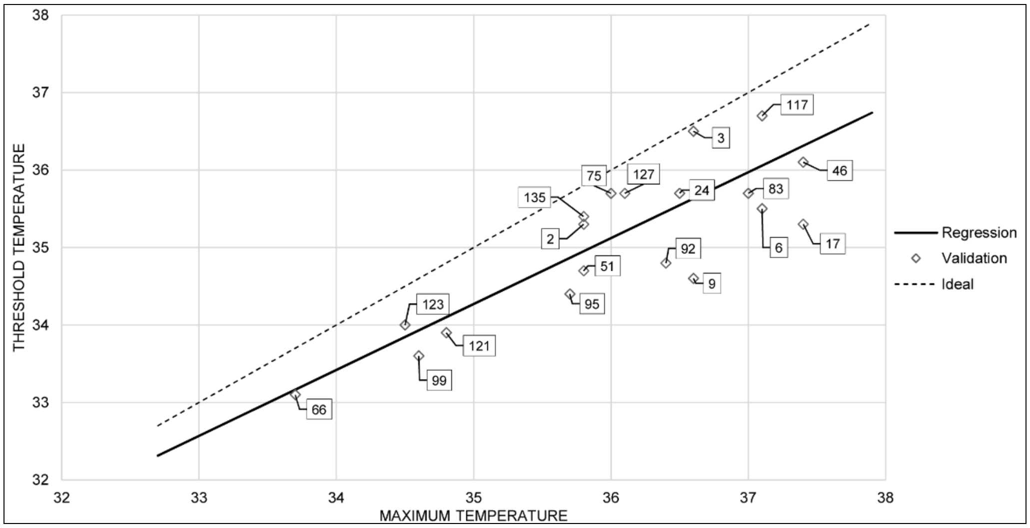

Figure 5. Validation cases in regression line.

Once the difference between Z and Y was calculated, the diagnosis given by the oncologist for each case was written down. The results, as shown in table 2, demonstrate that, while the difference of $Z - Y$ is a positive value, the diagnosis given by the oncologist is in terms of not detected, benign cases or risk cases, and, when the difference results a negative value, the diagnosis were in terms of high risk of malignancy or inflammatory finding.

Table 2. Validation results.

Sample number	X	Y	Z	Z - Y	Diagnosis
2	35.8	34.95	35.3	0.35	Benign breast lump
3	36.6	35.63	36.5	0.87	Do not detected in these images
6	37.1	36.06	35.5	-0.56	High risk of malignancy with inflammatory component
9	36.6	35.63	34.6	-1.03	High risk of malignancy
17	37.4	36.31	35.3	-1.01	High risk of malignancy with inflammatory component
24	36.5	35.55	35.7	0.15	Suspicious finding
46	37.4	36.31	36.1	-0.21	Inflammatory finding
51	35.8	34.95	34.7	-0.25	High risk of malignancy
66	33.7	33.16	33.1	-0.06	High risk of malignancy
75	36	35.12	35.7	0.58	Suspicious
83	37	35.97	35.7	-0.27	High risk of malignancy
92	36.4	35.46	34.8	-0.66	Confirmatory of malignancy
95	35.7	34.87	34.4	-0.47	High risk of malignancy
99	34.6	33.93	33.6	-0.33	High risk of malignancy
117	37.1	36.06	36.7	0.64	Benign finding
121	34.8	34.10	33.9	-0.20	High risk of malignancy
123	34.5	33.85	34	0.15	Risk of malignancy
127	36.1	35.21	35.7	0.49	Benign finding

International Journal of Engineering and Applied Sciences

www.eaas-journal.org

135	35.8	34.95	35.4	0.45	Risk of malignancy

Besides the multivariate model, the difference of the maximum temperature and the threshold temperature for the 130 cases was calculated in order to know if these differences could be fit into a theoretical probability distribution. The null hypothesis tested the distributions Normal, Exponential, Lognormal and Gamma. It was applied a Chi-Squared test with a significance level α of 5%. All the null hypothesis were rejected.

5. RESULTS

Analysing the data of 130 patients diagnosed by an oncologist -where the maximum temperature of the breast was obtained directly from the image, and the threshold temperature was obtained from the Comma Separated Value file via an algorithm and comparing the result with the diagnosis- a multivariate regression model –shown in Equation (1)- to predict threshold temperature for thermal anomalies from the breast maximum temperature was obtained.

The model was validated with 19 cases. The regression threshold temperature for these cases was calculated from the multivariate regression model, depending on the maximum temperature of the breast. The real threshold temperature was obtained from the minimum of the two maximum temperatures of the breasts. The evidence in these 19 cases shows that, while the difference between the real threshold temperature minus the regression threshold temperature was a positive value, the cases were catalogued by the oncologist as not severe, but, if the difference is a negative value, the patient was classified as severe.

Besides, a univariante model was proposed for the difference of the maximum temperature and the threshold temperature, but the null hypothesis were rejected for the probability distribution tested.

6. CONCLUSIONS AND FUTURE WORK

In breast thermography, a thermal anomaly is defined as an area with a temperature higher than the surrounding tissue, found in one breast, but not in the other one, determining the presence of a peak temperature and thermal asymmetry. A multivariate regression model to provide a threshold temperature from the maximum temperature of the breast has been obtained to detect thermal anomalies. If the actual threshold temperature is under the regression threshold temperature, a high probability of a thermal anomaly case should be considered. Even though, gathered knowledge does not allow us to estimate an accurate probability of having or not such thermal anomaly yet. Moreover, a deeper research involving complementary diagnosis methods is required to confirm that these thermal anomalies are definitely cancer related cases.

7. REFERENCES.

1. Knaul FM, Nigenda G, Lozano R, Arreola-Ornelas H, Langer A, Frenk J. Scielo. Cáncer de mama en México: una prioridad apremiante. Volume 512009. p 335-344.
2. Sree SV, Ng EY-K, Acharya RU, Faust O. Breast imaging: A survey. *World Journal of Clinical Oncology* 2011;2(4):171-178.
3. Ng EY-K, Ung LN, Ng FC, Sim LSJ. Statistical analysis of healthy and malignant breast thermography. *Journal of Medical Engineering & Technology* 2001;25(6):253-263.
4. Köşüş N, Köşüş A, Duran M, Simavlı S, Turhan N. Comparison of standard mammography with digital mammography and digital infrared thermal imaging for breast cancer screening. *Journal of the Turkish-German Gynecological Association* 2010;11:152-157.
5. Jiang L, Zhan W, Loew MH. Modeling static and dynamic thermography of the human breast under elastic deformation. *IOPscience* 2010(56):187-202.
6. Harris J. Meta-analysis: Mammography increased breast cancer risk for high-risk women younger than 30. *Hem/Onc* Today 2009:40-41.
7. Scopinaro F, Pani R, De Vincentis G, Soluri A, Pellegrini R, Porfiri LM. High-resolution scintimammography improves the accuracy of technetium-99m methoxyisobutylisonitrile scintimammography: use of a new dedicated

International Journal of Engineering and Applied Sciences

www.eaas-journal.org

gamma camera. *European Journal of Nuclear Medicine* 1999;26(10):1279-1288.

8. Golnabi AH, Meaney PM, Geimer SD, Paulsen KD. Comparison of no-prior and soft-prior regularization in biomedical microwave imaging. *Journal of Medical Physics* 2011;36(3):159-70.
9. Oran OF, Ider YZ. Magnetic resonance electrical impedance tomography (MREIT) based on the solution of the convection equation using FEM with stabilization. *Physics in Medicine and Biology* 2012;57(16).
10. Laufer S, Solomon SB, Rubinsky B. Tissue characterization using electrical impedance spectroscopy data: a linear algebra approach. *Physiological Measurement* 2012;33(6):997-1013.
11. McKnight AL, Kugel JL, Manduca A, Hartmann LC, Ehman RL. MR elastography of breast cancer: preliminary results. *American Journal of Roentgenology* 2002;178(6):1411-1417.
12. Jiang LJ, Ng EY-K, Yeo ACB, Wu S, F. P, Yau WY, Chen JH, Yang Y. A perspective on medical infrared imaging. *Journal of Medical Engineering & Technology* 2005;29(6):257-267.
13. Ng EY-K, Kee EC. Advanced integrated technique in breast cancer thermography. *Journal of Medical Engineering & Technology* 2008;32(2):103-114.
14. Amri A, Saidane A, Pulko S. Thermal analysis of a three-dimensional breast model with embedded tumour using the transmission line matrix (TLM) method. *Computers in Biology and Medicine* 2011;41(2):76-86.
15. Ng EY-K, Fok SC, Peh YC, Ng FC, Sim LSJ. Computerized detection of breast cancer with artificial intelligence and thermograms. *Journal of Medical Engineering & Technology* 2002;26(4):152-157.
16. Moghbel M, Mashohor S. A review of computer assisted detection/diagnosis (CAD) in breast thermography for breast cancer detection. *Artificial Intelligence Review* 2011.
17. Ibarra-Castanedo C, Bendada A, Maldague X. Thermographic Image Processing for NDT. In: Estructurales AAdENDy, editor. IV Conferencia Panamericana de Ensayos No Destructivos. Buenos Aires, Argentina2007.
18. Gonzalez RC, Woods RE, Eddins SL. *Digital Image processing using MATLAB.* Upper Saddle River, N.J.: Pearson Prentice Hall; 2004.
19. Megahed FM, Woodall WH, Camelio JA. A Review and Perspective on Control Charting with Image Data. *Journal of Quality Technology* 2011;43(2):83-98.
20. Laza Dueñas CM, Rego Días A, Carlos Pías N. Aplicación de algoritmos para el procesamiento de imágenes a la tecnología de microarreglos para el diagnóstico. *Revista Investigación Operacional* 2010;31(3):268-274.
21. Montgomery DC. *Introduction to Statistical Quality Control*. Arizona State University; 2005.
22. Montgomery DC, Runger GC. *Applied Statistics and Probability for Engineers*. John Wiley & Sons, Inc.; 2003.
23. Peña D. *Análisis de datos multivariantes*. *1st ed*. McGraw-Hill/Interamericana. Madrid; 2002.
24. Hair Jr. JF, Black WC, Babin BJ, Anderson RE. *Multivariate Data Analysis*. *7th ed*. Prentice Hall. United States; 2009
25. FOK SC, Ng EY-K, TAI K. Early detection and visualization of breast tumor with Thermogram and Neural Network. *Journal of Mechanics in Medicine and Biology* 2002;02(02):185-195.
26. Ng EY-K, Chen Y, Ung LN. Computerized bre ast the rmography: study of image segmentation and tempe rature cyclic variations. *Journal of Medical Engineering &Technology* 2001;25(1):12-16.
27. Tan JMY, Ng EY-K, Archarya U. R, Keith LG, Holmes J. Comparative Study on the Use of Analytical Software to Identify the Different Stages of Breast Cancer Using Discrete Temperature Data. *Journal of Medical Systems* 2009;33(2):141-153.
28. Kapoor P, Prasad SVAV, Patni S. Image Segmentation and Asymmetry Analysis of Breast Thermograms for Tumor Detection. *International Journal of Computer Applications* 2012;50(9).