Scientific
Research
Publishing

A Gaussian Multivariate Hidden Markov Model for Breast Tumor Diagnosis

Angelo Raherinirina, Adore Randriamandroso, Aimé Richard Hajalalaina, Rivo Andry Rakotoarivelo, Fontaine Rafamatantantsoa

University of Fianarantsoa, Fianarantsoa, Madagascar
Email: angelo.raherinirina@gmail.com

Abstract

The stage of a tumor is sometimes hard to predict, especially early in its development. The size and complexity of its observations are the major problems that lead to false diagnoses. Even experienced doctors can make a mistake in causing terrible consequences for the patient. We propose a mathematical tool for the diagnosis of breast cancer. The aim is to help specialists in making a decision on the likelihood of a patient's condition knowing the series of observations available. This may increase the patient's chances of recovery. With a multivariate observational hidden Markov model, we describe the evolution of the disease by taking the geometric properties of the tumor as observable variables. The latent variable corresponds to the type of tumor: malignant or benign. The analysis of the covariance matrix makes it possible to delineate the zones of occurrence for each group belonging to a type of tumors. It is therefore possible to summarize the properties that characterize each of the tumor categories using the parameters of the model. These parameters highlight the differences between the types of tumors.

1. Introduction

Breast cancer remains the most common and deadliest cancer worldwide. Statistics show that it affects women much more than men [1]. With an upward trend over the past twenty years, the developed countries seem to be the most affected. Little information is available on the situation in underdeveloped countries [2] [3].

In the case of Madagascar, the non-existence of official statistics does not al-

low an inventory to be made on breast cancer. Data from the Institute Pasteur de Madagascar (IPM) shows that the average age of patients is 48 years. Although the data from the IPM does not represent the whole of the Malagasy reality, they nevertheless show that the number of cases of breast tumor continues to increase; it has become the most frequently observed cancer. The most affected age group in the series is 36 to 55 years old, see Figure 1. The youngest patient is 22 years old and the oldest 90 years old. Of all the cases that could be typed histologically, 2/3 are in the grade 3 infiltrating ductal carcinoma stage of the Scar-Bloom-Richardson histoprognostic classification (SBR) [2].

Usually the diagnosis is very late and most cases are seen clinically only at an advanced stage of the tumor [4]. In addition to socio-economic factors, the structure of the health system makes the early diagnosis of cancerous lesions impossible in all Malagasy territories. Support establishments are almost non-existent outside the main cities (Antananarivo, Fianarantsoa). Indeed, according to IPM data, 80% of the cases observed come from the two large hospitals in the capital Antananarivo. In these situations, the ability to quickly and unambiguously detect the nature of a tumor is of paramount importance. This makes it possible to help specialists in the choice and type of treatment to prescribe from the early stages of its development [5] [6].

However, studies have shown that the diagnosis of the intrinsic nature of cancer is not necessarily reliable at a fairly early stage of its development [7]. Leading to inadequate treatment, this false diagnosis increases the death rate of the disease. The biggest challenge is therefore to provide tools for the rapid and reliable detection of the nature of tumors. In addition to mammographic exams, mathematical and computer modeling tools can help us in this process [8] [9]. Several models have already been proposed to describe the evolution of breast cancer [8] [10] [11].

We are interested in so-called stochastic models to describe the evolution of
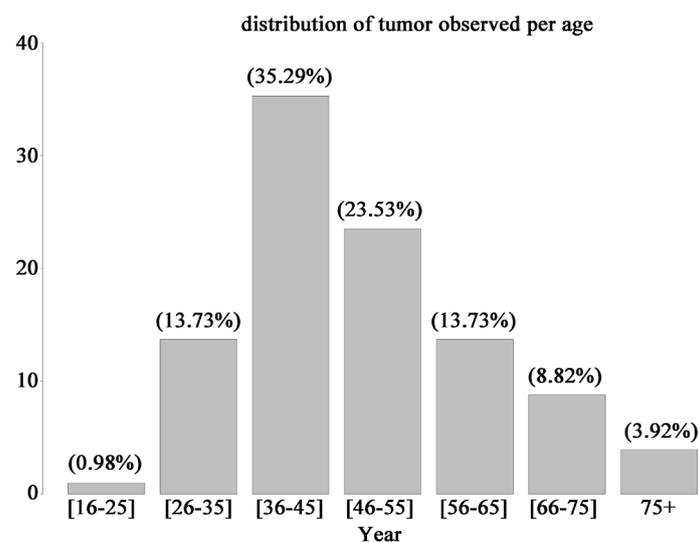

Figure 1. Age of observed breast cancer patients in Madagascar, according to IPM data.

the breast tumor [11] [12] [13]. Given the small amount of information available, these models seem the most appropriate. In [11] the authors use a Markov chain to model the metastatic course of lung cancer. Generally, these models assume that tumor types are unambiguously determined based on observations. This is not the case in reality, appearances are deceptive. Hence the interest to consider the dynamics behind the observed data. The hidden Markov chain is one of the tools used to study the dynamics of a latent phenomenon such as breast cancer. It has already shown its proof in various situations such as the characterization and prediction of the evolution of tumors [14] [15].

For breast cancer, the biomarker provides information on the nature of a tumor: benign or malignant [16]. However, due to the complexity and quantity of data to be exploited, it is necessary to be able to count on an agent who judges the observations objectively and reliably. In this article, our objective is to provide a tool to help diagnose the nature of a breast tumor at a primary stage. By having observations of the biomarker, we consider that the nature of the tumor is a hidden variable. The observable variables consist of the geometric properties of the tumor located on the mammary area. With probabilistic techniques, we will characterize the true nature of the tumor according to the observations. Based on a series of observations, we will determine the probability that the tumor is malignant or benign. This information is important in a country like Madagascar where diagnostic instruments are still very expensive and inaccessible.

2. Gaussian Mixture Hidden Markov Chain Model

A Markov chain is a model which allows the study of sequences of a random quantity. It is based on a very strong assumption that current information is needed to predict the future [17]. Introduced by Rabiner [18], in 1998, the hidden Markov chain is an extension of this model. In most cases, the phenomena that interest us are not observable. The hidden Markov model (HMM) takes into account observed data and hidden events. Consider a markovian process (X_t) in a discrete state space

$$S = (S_1, S_2, \cdots, S_n) \tag{1}$$

Because we are working on a discrete-time Markov chain, it is necessary to introduce the transition matrix A such that

$$A_{ij} = P(X_t = j / X_{t-1} = i). \tag{2}$$

With an initial condition

$$\mu_i = P(X_0 = i) \tag{3}$$

For all *t*, the evolution of (X_t) is unobservable. What we have instead is an observation (Y_t) which can take its value in a finite set

$$O = \{Y_1, Y_2, \cdots, Y_M\} \tag{4}$$

which depends on these hidden variables.

We then have an observation function

$$f_i(y_k) = P(Y_t = y_k / S_t = i) \tag{5}$$

which gives us the probability of the observation knowing the state.

The Equations (1)-(5) characterize a hidden Markov model. The aim of our modeling work is to determine the nature of a breast tumor in the first phase of its evolution.

We assume that the patient's condition at time t corresponds to a malignant tumor or benign tumor:

$$X \in S = \{\text{Malign}, \text{Benign}\}. \tag{6}$$

The process $(X_t)_{t \geq 0}$ is considered as an unobservable finite state space Markov chain, see **Figure 2**. The variable $Y \in R^n$ represents the observations corresponding to the geometric properties of the tumor. Our goal is to use the series of observations to estimate the different parameters of the process.

Different very powerful algorithms have been developed to do this [19]. The forward-backward algorithm calculates the probability of the current state knowing the sequence of observations $(y_1, y_2, \cdots, y_T)$ [20].

Let us summarize in θ all the parameters of the model: transition matrix, parameter of the probability distribution (mean, variance) for each state. By defining the functions α and β such that

$$\alpha_{t+1}(j) = P\left(O_{1:(t+1)} = o_{1:(t+1)}, S_{t+1} = s_{t+1} / \theta\right) = \sum_{i=1}^{N} \alpha_t(i) a_{ij} f_j(y_{t+1}), \tag{7}$$

$$\beta_t(i) = P\left(O_{t:T} = o_{t:T} / S_t = s_t, \theta\right) = \sum_{j=1}^{N} \beta_{t+1}(i) a_{ij} f_j(y_{t+1}). \tag{8}$$
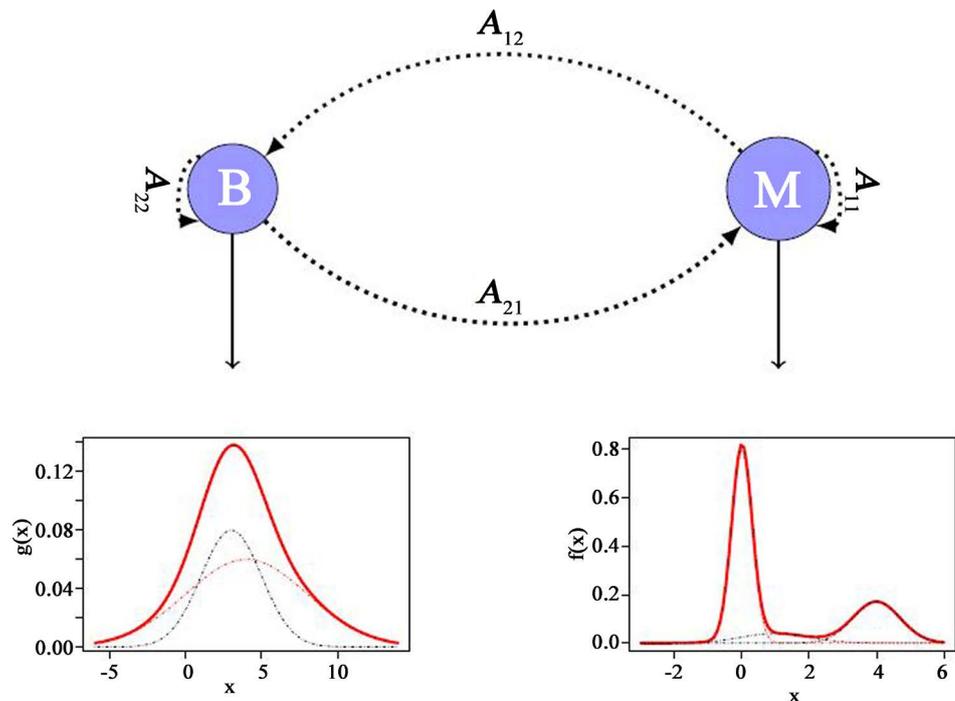
We have

Figure 2. HIdden Markov Model (HMM) for breast tumor. Top: transition graph of the included Markov chain. Bottom: distribution of observations: Gaussian mixture.

$$P\left(Y_{1:T} = y_{1:T}/\theta\right) = \sum_{i=1}^{N} \alpha_t\left(i\right)\beta_t\left(i\right).\tag{9}$$

α and β allow us compute respectively the probability of a given sequences of observations given the parameter θ and a starting time t.

These values will be used in the learning phase (modify the parameter given new observations). The implementation is described by **Table 1**.

In our case, we are interested in the sequence of hidden states knowing the sequence of observations it generated. The Viterbi algorithm, described in **Table 2**, solves this problem [21].

Regarding breast cancer, the observable variable is not a singular value but a d-dimensional vector. This corresponds to the different measurements made on a patient during a diagnosis (temperature, concentration, etc.). Thus, we have

Table 1. Baum-Welch Algorithm for the implementation of the EM algorithm.

Data: $\left(O_1, O_2, \cdots, O_n\right)$ set of observations

Initialization

θ_0

  $\left(a_{ij}\right)$ Matrix of transition

  p_{il} Weight of the l mixture component of state i,

  μ_{il} Mean vector of the l mixture component of state i,

  σ_{il} Covariance matrix of the l mixture component of state i.

Forward $\left(O_1, O_2, \cdots, O_n\right)$,

Backward $\left(O_1, O_2, \cdots, O_n\right)$.

While $\log\left(P\left(O_1, O_2, \cdots, O_n\right)\right) \leq p_0$ (fixed threshold)

Expectation $\xi_t\left(i, j\right) = P\left(s_t = i, s_{t+1} = j/0, \theta\right)$,

$\gamma_t\left(i\right) = P\left(s_t = i/O, \theta\right) = \sum_j \xi_t\left(i, j\right)$,

$\gamma_{il}\left(t\right) = P\left(S_t = i, m = l/O_t, \theta\right) = \gamma_i\left(t\right)\dfrac{p_{il}P\left(O_t/\mu_{il}, \sigma_{il}, \theta\right)}{P\left(O_t/S_t = i, \theta\right)}$.

Maximization $a_{ij} = \dfrac{\sum_{t=1}^{T}\xi_t\left(i, j\right)}{\sum_{t=1}^{T}\gamma_t\left(i\right)}$,

$p_{il} = \dfrac{\sum_{t=1}^{T}\gamma_{il}\left(t\right)O_t}{\sum_{t=1}^{T}\gamma_{il}\left(t\right)}$,

$\sigma_{il} = \dfrac{\sum_{t=1}^{T}\gamma_{il}\left(t\right)\left(O_t - \mu_{il}\right)\left(O_t - \mu_{il}\right)^{\mathrm{T}}}{\sum_{t=1}^{T}\gamma_{il}\left(t\right)}$.

End While

$f_j\left(O_{t+1}\right)$: density of probability of the state i,

$\gamma_{il}\left(t\right)$: probability that the component l of the state i generated the observation at time t,

$\gamma_i\left(t\right)$: probability that the state i generated the observation at time t.

Table 2. Viterbi algorithm (generation of hidden state).

For $i = 1$ to N do

$\quad \delta_1(i) = p_i f_i(O_1)$

End For

For $t = 2$ to T do

$\quad$ For $j = 1$ to N do

$$\delta_t(j) = \max_i \left(\delta_{t-1}(i) a_{ij} \right) f_j(O_t)$$

$$\psi_t(j) = \arg\max_i \left(\delta_{t-1}(i) a_{ij} \right) f_j(O_t)$$

$\quad$ End For

End For

$$P^* = \max_i \left(\delta_T(i) \right)$$

$$Q_T^* = \arg\max_i \left(\delta_T(i) \right)$$

For $t = T - 1$ to 1 do

$$Q_t^* = \psi_{t+1} \left(Q_{t+1}^* \right)$$

End For

---

chosen a multivariate Gaussian density to model the probability of observation of each state [22].

A multivariate Gaussian mixture is a function composed of the sum of several other functions (in most cases Gaussian functions) [23]. Given that our observations are multidimensional (several observed characters), we use a random vector $X = [X_1, X_2, \cdots, X_n]$ distributed according to a mixed Gauss distribution. Each component of the mixture has for average

$$\mu = \left[E(X_1), E(X_2), \cdots, E(X_n) \right], \tag{10}$$

and a covariance matrix

$$\sigma = \begin{pmatrix} V(X_1) & Cov(X_1, X_2) & \cdots & Cov(X_1, X_n) \\ Cov(X_1, X_2) & V(X_2) & \ddots & \vdots \\ \vdots & \vdots & \ddots & \vdots \\ Cov(X_1, X_n) & \cdots & \cdots & V(X_n) \end{pmatrix}, \tag{11}$$

with $\sigma_{ij} = \sigma_{ji}$ and $\sigma_{ii} = Var(X_i) = Cov(X_i, X_i)$.

By associating to each component i a weight p_i with $\sum_i p_i = 1$, the expression of the Gaussian multidimensional distribution is

$$f(x, \mu, \sigma) = \frac{1}{(2\pi)^{n/2} |\sigma|^{1/2}} \exp\left(\frac{1}{2} (x - \mu)^{\mathrm{T}} \sigma^{-1} (x - \mu) \right) \tag{12}$$

Thus, the Gaussian mixture $f_k(x)$ associated with the state k is characterized by:

1) $p_{ik} \in R$ the weight of the component i of the mixture representing the density of the observable function of the states k,

2) $\mu_{ik} \in R^n$ the mean vector of the random vector of the component i of the mixture k,

3) $\sigma_{ij} \in R^{n \times n}$ the covariance matrix of the Gaussian density of the component i of the mixture k.

The parameters are estimated by implementing the Expectation-Maximization algorithm [22] described in Table 3.

3. Result and Discussion

As we do not have any database on breast cancer in Madagascar, we used the published database on breast cancer from Wisconsin to illustrate our model [24]. With a sample made up of 569 individuals, the authors measured 30 parameters which make it possible to determine the state of the tumor (malignant or benign). These 30 observations characterize the two types of cancer that we want to identify. Figure 3 represents the correlation matrix of the characteristic variables.

This figure shows that some of the variables are related. Going from light blue to dark blue, the figure shows the degrees of relationships between the covariates.

It is possible to reduce the number of variables by a principal component analysis. We have plotted the empirical distribution of the geometric characteristics according to the type of tumor, see Figure 4. In particular, we are finding that some of them are easy to distinguish between the tumor types, while others do not.

After switching to principal component analysis, we concluded that a 2-dimensional subspace is sufficient to keep 95% of the variance of the observed data. Figure 5 represents the projection of individuals on the two main.

The number of components of each mixture can be calculated using the Bayesian

Table 3. Expectation Maximization algorithm for a gaussian mixture multivariate model

Observation: $(x_1, x_2, \cdots, x_N)$ with $x_i \in R^n$

Initialization:

μ_0, σ_0, π_0

While $\log(P(x_1, x_2, \cdots, x_N)) \leq p_0 (threshold)$

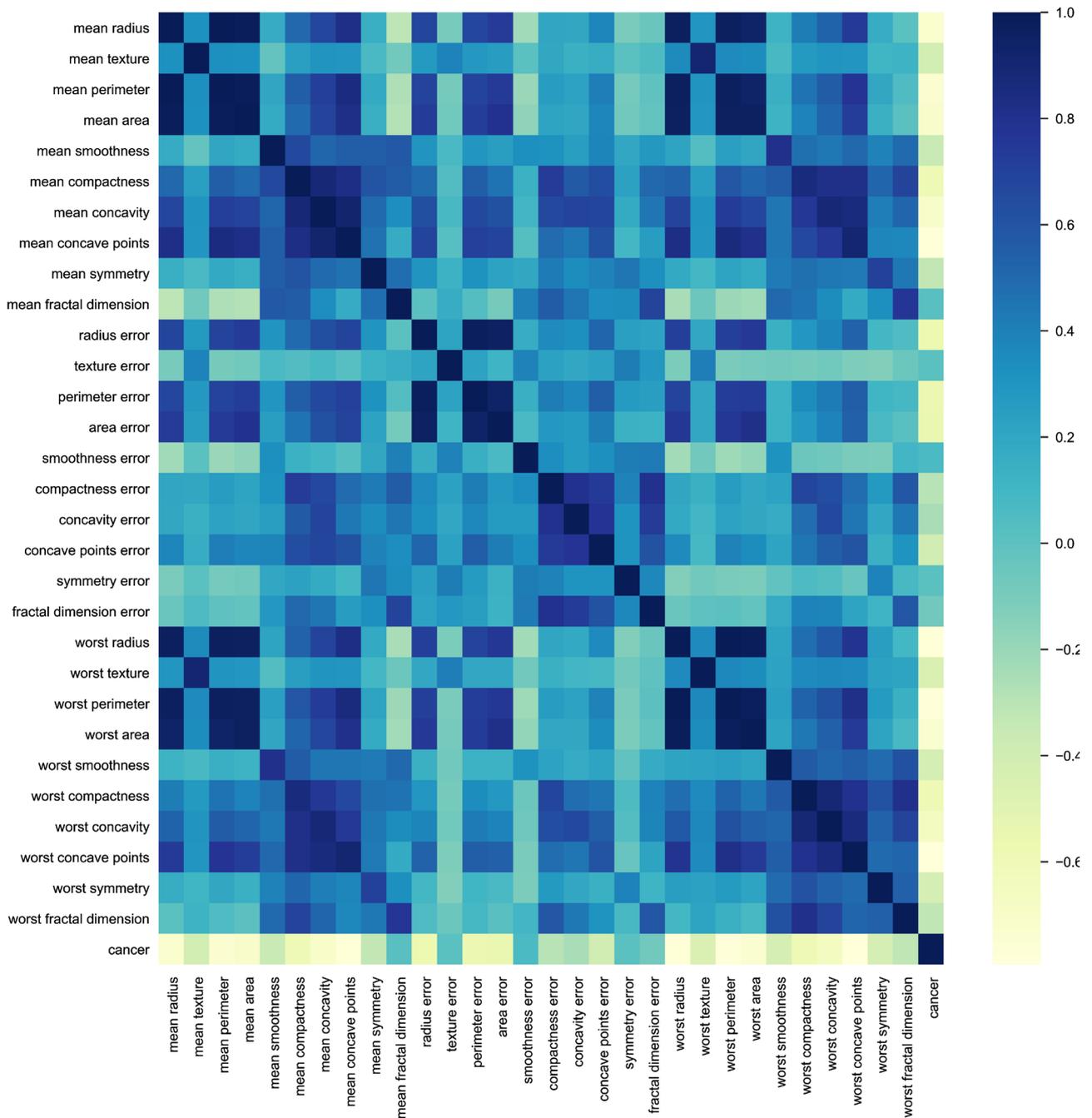|	
Expectation	$r_{nk} = \dfrac{p_k N(x_n / \mu_k, \sigma_k)}{\displaystyle\sum_{j=1}^{K} p_j N(x_n / \mu_j, \sigma_j)}$
Maximization	$\mu_k = \dfrac{\displaystyle\sum_{n=1}^{N} r_{nk} x_n}{\displaystyle\sum_{n=1}^{N} r_{nk}}$
	$\sigma_k = \dfrac{\displaystyle\sum_{n=1}^{N} r_{nk}(x_n - \mu_k)(x_n - \mu_k)^{\mathrm{T}}}{\displaystyle\sum_{n=1}^{N} r_{nk}}$
	$p_k = \dfrac{\displaystyle\sum_{n=1}^{N} r_{nk}}{N}$

End While

Figure 3. Correaltion of characteristics variables.

Information Criterion (BIC) on each of the data that we have for the different types of tumors [25]. For an observation, the point clouds are obtained by making a vector product on each of the two eigenvectors.

We thus have the number of optimal components for each mixture. Our results show that the number of suitable components for the benign type is $k = 2$ while that of the malignant type is $k = 3$. Which is quite logical because of the arrangement of the points in the **Figure 5**. The arrangement on the point clouds of malignant cases is quite dispersive, so it is logical that its mixture has more
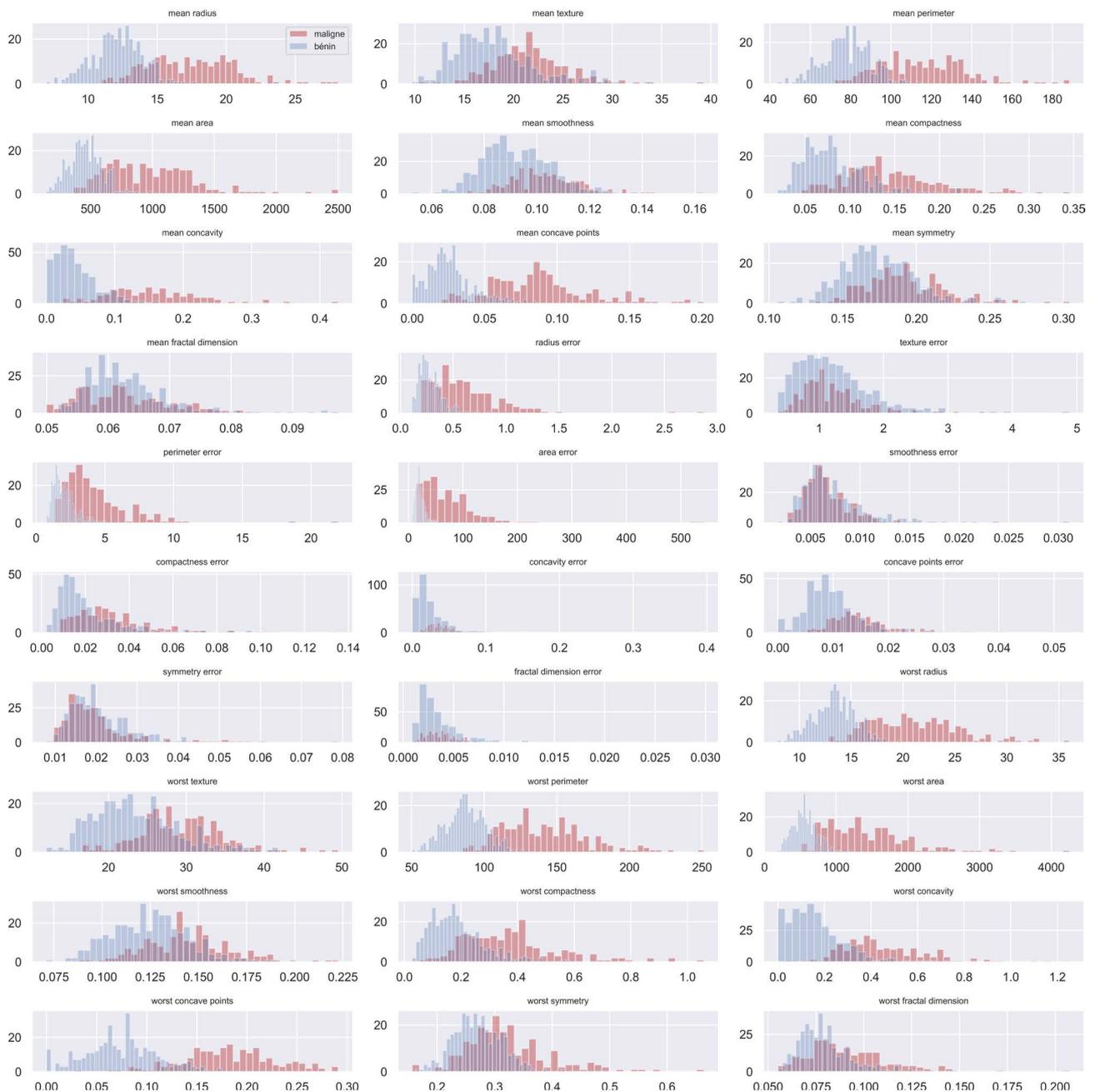
Figure 4. Empirical distribution of geometric characteristics according to the type of tumor.

components than the other type. For each of the two states (malignant and benign), we estimate the parameters of their distribution using the Expectation Maximization algorithm. Learning is supervised because we already know for each observation the state that generated it. The parameters of the Gaussian distribution for each mixture are given in **Table 4** and **Table 5**.

Thus, we have a probability density which makes it possible to associate a probability value for each observation, see **Figure 6**. For each observation $X \in R^2$, we calculate its probability conditioned by all the parameters $P(X/\theta)$. The "+" sign represents the averages. Its size is proportional to the weight of the component
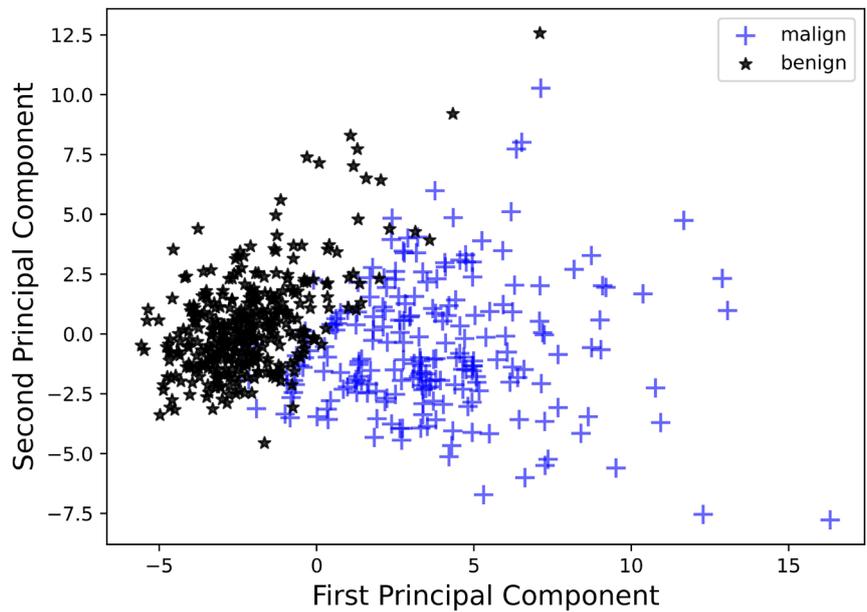
Figure 5. Projection of the sample on the principal axes after the principal component analysis.

Table 4. Parameters of the distribution of the state "malignant" tumor.

Malign Tumor	Eigenvectors	Eigenvalues
Covariance		
$\begin{pmatrix} 5.473 & -0.12 \\ -0.12 & 1.86 \end{pmatrix}$	$\begin{pmatrix} -0.033 & -0.99 \end{pmatrix}$	1.856
Mean	$\begin{pmatrix} -0.99 & 0.033 \end{pmatrix}$	5.476789
$\begin{pmatrix} 2.712 & -1.835 \end{pmatrix}$		
Weight: 0.55		
Covariance		
$\begin{pmatrix} 8.68 & -2.66 \\ -2.66 & 17.67 \end{pmatrix}$	$\begin{pmatrix} -0.96 & -0.26 \end{pmatrix}$	7.9519
Mean	$\begin{pmatrix} -0.26 & 0.96 \end{pmatrix}$	18.39
$\begin{pmatrix} 7.731 & -0.081 \end{pmatrix}$		
Weight: 0.20		
Covariance		
$\begin{pmatrix} 2.38 & 1.01 \\ 1.01 & 2.69 \end{pmatrix}$	$\begin{pmatrix} -0.7588 & 0.6512 \end{pmatrix}$	1.513
Mean	$\begin{pmatrix} 0.6512 & 0.7588 \end{pmatrix}$	3.55
$\begin{pmatrix} 2.87 & 1.735 \end{pmatrix}$		
Weight: 0.25		

in the mixture.

With the estimated parameters, we performed a simulation on the model. The result of the principal component analysis of the simulated data are represented in the **Figure 7**.

By comparing this result with the real data in **Figure 5**, we can deduce that the result is reliable. This concerns the efficiency of the EM algorithm in estimating

Table 5. Parameter of the distribution of the state "benign" tumor.

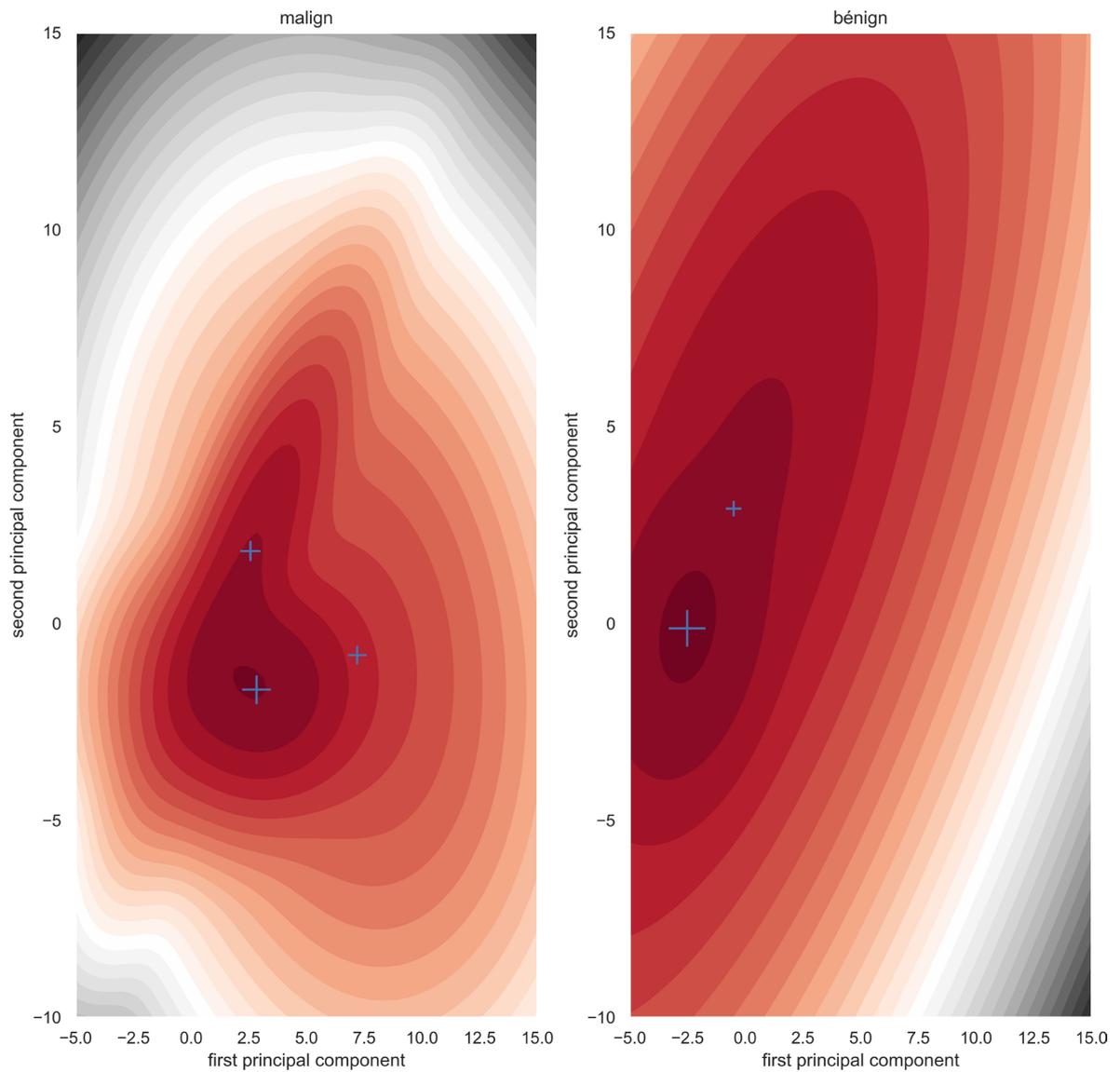| Benign Tumor	Eigenvectors	Eigenvalues
Covariance		
$\begin{pmatrix} 1.56 & 0.47 \\ 0.47 & 2.07 \end{pmatrix}$	$\begin{pmatrix} -0.859 & 0.5114 \end{pmatrix}$	1.2802
Mean	$\begin{pmatrix} 0.5114 & 0.859 \end{pmatrix}$	2.34
$\begin{pmatrix} -2.515 & -0.121 \end{pmatrix}$		
Weight: 0.84		
Covariance		
$\begin{pmatrix} 5.43 & 4.51 \\ 4.51 & 9.05 \end{pmatrix}$	$\begin{pmatrix} -0.829 & 0.5589 \end{pmatrix}$	2.389
Mean	$\begin{pmatrix} 0.5589 & 0.8291 \end{pmatrix}$	12.1202
$\begin{pmatrix} -0.501 & 2.92 \end{pmatrix}$		
Weight: 0.16		

Figure 6. Projection of the probability density of the two principal components after normalization.
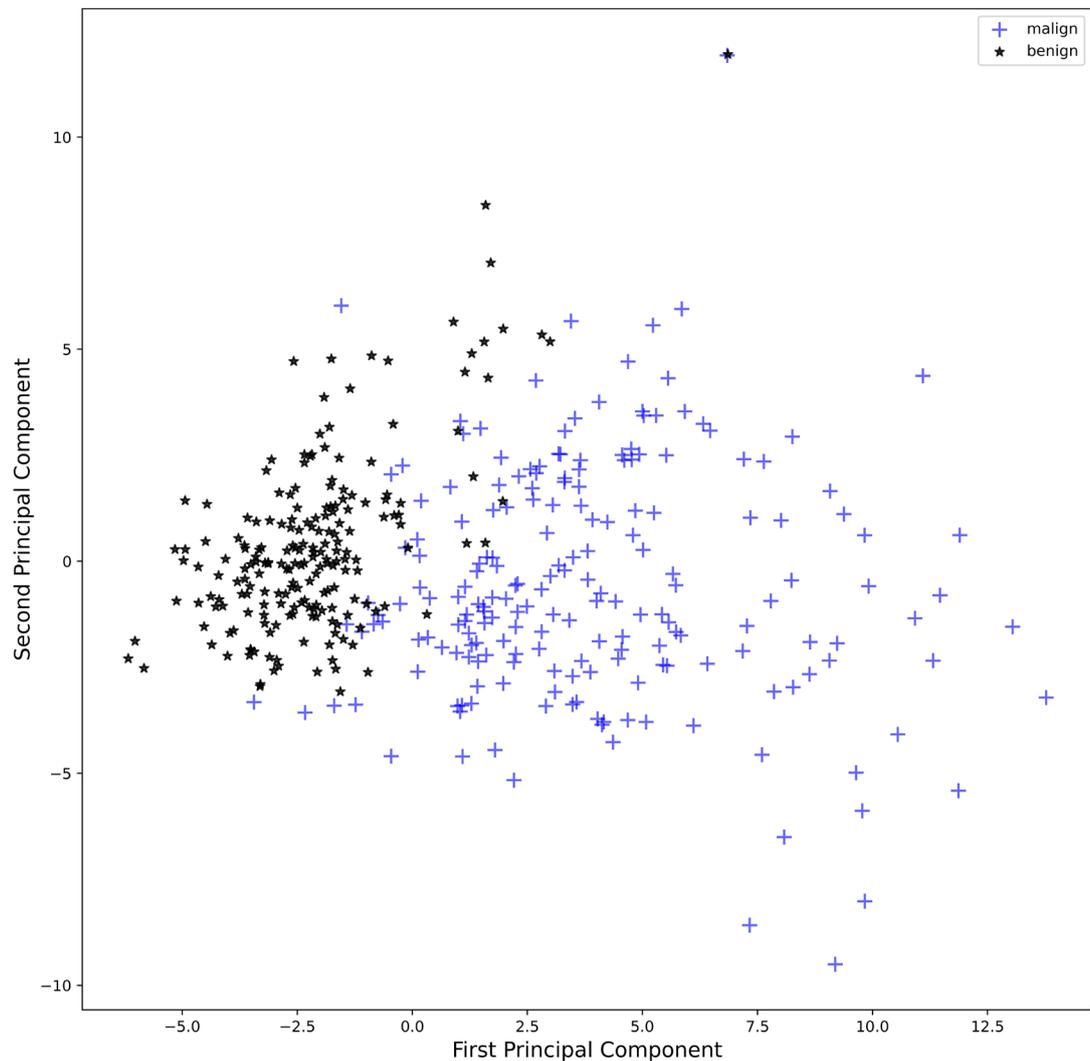
Figure 7. Principal component analysis with simulated data.

the parameters of a Hidden Markov model.

Using a simulation, we illustrate the approximation of the real data by the simulated data from the model. Considering the two tumor states: malignant and benign, Figure 8 represents a realization of the Viterbi algorithm (Table 2).

The first red curve represents the actual condition of the tumor (1 for benign and 0 for malignant) and which is assumed to be hidden. The blue curve below corresponds to the state estimated by the algorithm and is the component most likely to have generated the observations. The estimate was made using the observation sequences projected onto the two eigenvectors and represented by the two lowest blue curves. According to this figure, the model manages to bring closer the true natures of the tumors during its evolution.

This model is therefore a real additional tool which can help specialists in the diagnosis of breast tumors at an early stage. Simulations with breast cancer data show that inference with a Gaussian mixture hidden Markov model can help us characterize the internal nature of the disease.
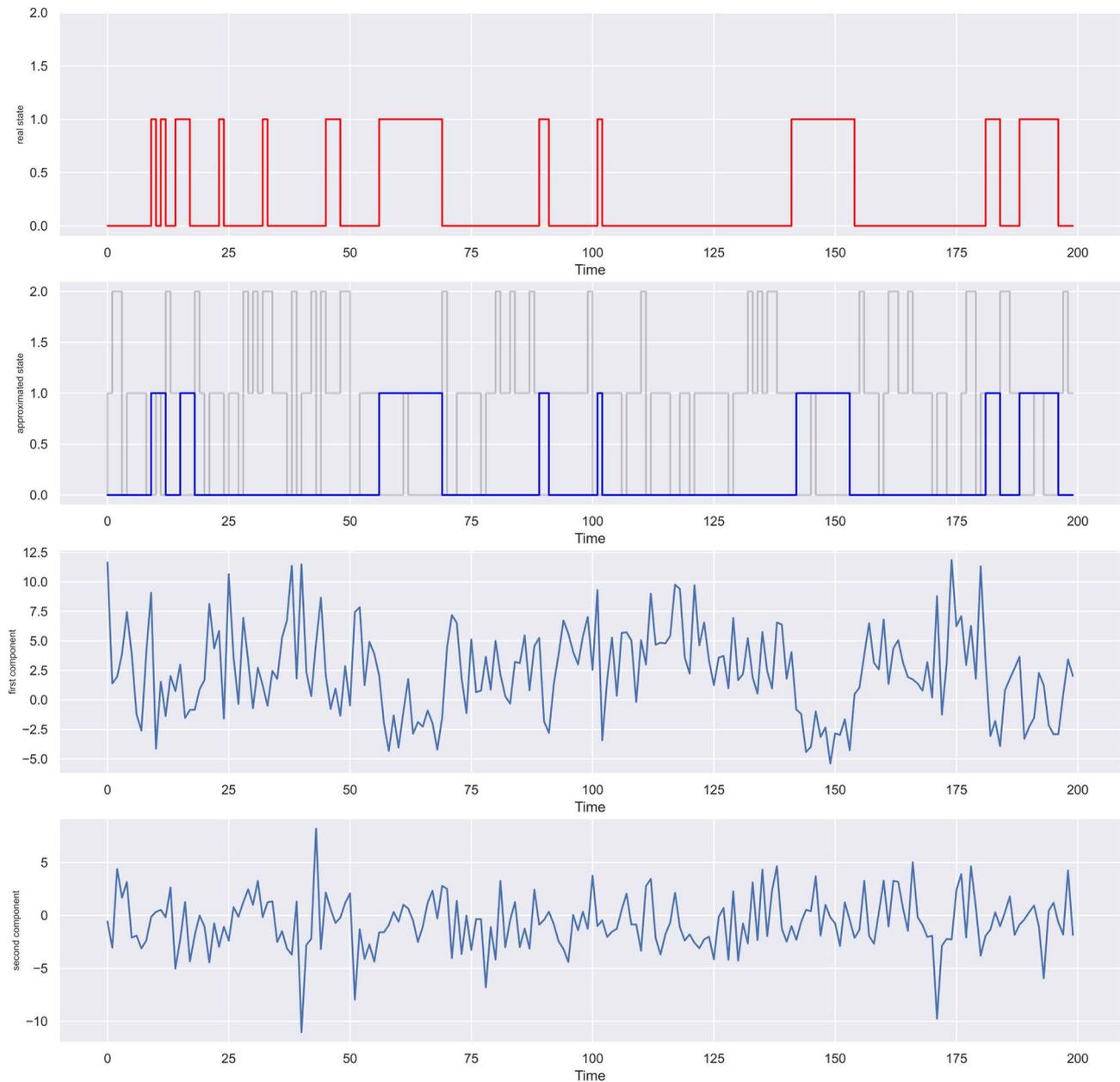
Figure 8. Simulation of the model with the Viterbi algorithm.

4. Conclusions

Faced with the difficulties in diagnosing breast cancer, specialists need other decision-making tools. Especially in an underdeveloped country like Madagascar, patient reception structures remain insufficient. The cost of health examinations is very high and the support is done late. This article constitutes our contribution to the anticipation of decision-making according to the geometric characteristics of the tumor. As we do not have a database on breast cancer in Madagascar, the inference of our hidden Makov model was made with data on breast cancer from Wisconsin, available on the web. As the observations relate to several variables related to the geometric characteristics of the tumor, we chose the mul-

tivariate Gaussian distribution for the distribution of observations corresponding to each state.

By implementing the Viterbi algorithm, we succeeded in identifying the parameters of the distribution of observations corresponding to each state of the tumor: benign or malignant. With the estimated parameters, we reproduced the initial data set under the same conditions. The result is significant.

The interest of our modeling work is both to characterize the nature of tumors as early as possible and also to predict its evolution. This is important especially in the event that our health system is not yet ready for support. Even if the data from the Institut Pasteur de Madagascar (IPM) are not representative, the frequency of the observed cases shows that the situation is critical. To be able to carry out more in-depth studies, specialists recommend the establishment of a database on breast cancer in Madagascar.

Conflicts of Interest

The authors declare no conflicts of interest regarding the publication of this paper.

References

[1] Azamjah, N., Soltan-Zadeh, Y. and Zayeri, F. (2019) Global Trend of Breast Cancer Mortality Rate: A 25-Year Study. *Asian Pacic Journal of Cancer Prevention*, **20**, 2015-2020. https://doi.org/10.31557/APJCP.2019.20.7.2015

[2] Sado, A. (2015) Application of Brody Growth Function to Describe Dynamics of Breast Cancer Cells. *American Journal of Applied Mathematics*, **3**, 138-145. https://doi.org/10.11648/j.ajam.20150303.20

[3] Al-Moundhri, M., Al-Bahrani, B., Pervez, I., Ganguly, S., Nirmala, V., Al-Madhani, A., Al-Mawaly, K. and Grant, C. (2004) The Outcome of Treatment of Breast Cancer in a Developing Country—Oman. *The Breast*, **13**, 139-145. https://doi.org/10.1016/j.breast.2003.10.001

[4] Raharisolo, V., Rabarijaona, L., Rajemiarimoelisoa, C., Rasendramino, M. and Migliani, R. (2002) Management of Breast Cancers Diagnosed at the Institut Pasteur de Madagascar from 1995 to 2001. *Archives de l'Institut Pasteur de Madagascar*, **68**, 104-108.

[5] Cianfrocca, M. and Goldstein, L. (2004) Prognostic and Predictive Factors in Early-Stage Breast Cancer. *Oncologist*, **9**, 606-601. https://doi.org/10.1634/theoncologist.9-6-606

[6] Chuwa, E., Yeo, A., Koong, H., Wong, C., Yong, W., Tan, P., Ho, J., Wong, J. and Ho, G. (2009) Early Detection of Breast Cancer through Population-Based Mammographic Screening in Asian Women: A Comparison Study between Screen-Detected and Symptomatic Breast Cancers. *The Breast Journal,* **15**, 133-139. https://doi.org/10.1111/j.1524-4741.2009.00687.x

[7] Bicar, A. (2018) Le cancer du sein chez la jeune femme et sa prise en charge. PhD Thesis, University of Limoge, Limoge, 15.

[8] Botesteanu, D., Lipkowitz, S., Lee, J. and Levy, D. (2016) Mathematical Models of Breast and Ovarian Cancers. *WIREs Systems Biology and Medicine*, **8**, 337-362. https://doi.org/10.1002/wsbm.1343

[9]   Harris, J., Lippman, M., Veronesi, U. and Willett, W. (1992) Breast Cancer. *New England Journal of Medicine*, **327**, 473-480.
https://doi.org/10.1056/NEJM199208133270706

[10]   Enderling, H., Chaplain, M., Anderson, A. and Vaidya, J. (2007) A Mathematical Model of Breast Cancer Development, Local Treatment and Recurrence. *Journal of Theoretical Biology,* **246**, 245-259. https://doi.org/10.1016/j.jtbi.2006.12.010

[11]   Newton, P., Mason, J., Bethel, K., Bazhenova, L., Nieva, J. and Kuhn, P. (2012) A Stochastic Markov Chain Model to Describe Lung Cancer Growth and Metastasis. *PloS ONE*, **7**, e34637. https://doi.org/10.1371/journal.pone.0034637

[12]   Lee, S. and Zelen, M. (2006) A Stochastic Model for Predicting the Mortality of Breast Cancer. *JNCI Monographs*, **2006**, 79-86.
https://doi.org/10.1093/jncimonographs/lgj011

[13]   Speer, J., Petrosky, V., Retsky, M. and Wardwell, R. (1984) A Stochastic Numerical Model of Breast Cancer Growth That Simulates Clinical Data. *Cancer Research*, **44**, 4124-4130.

[14]   Mohammadreza, M., Mohammadreza, S. and Hossein, R. (2020) Using Hidden Markov Model to Predict Recurrence of Breast Cancer Based on Sequential Patterns in Gene Expression Profiles. *Journal of Biomedical Informatics,* **111**, Article ID: 103570.
https://doi.org/10.1016/j.jbi.2020.103570

[15]   Mahmoudzadeh, E., Montazeri, M., Zekri, M. and Sadri, S. (2015) Extended Hidden Markov Model for Optimized Segmentation of Breast Thermography Images. *Infrared Physics and Technology,* **72**, 19-28.
https://doi.org/10.1016/j.infrared.2015.06.012

[16]   Zaman, K., Ambrosetti, A., Perey, L., Jeanneret-Sozzi, W., Delaloye, J. and Ziegler, D. (2007) Cancer du sein chez la jeune femme: Traitements adjuvants et désir de grossesse. *Revue Médicale Suisse*, **7**, 1298-1304.

[17]   Norris, J. (1998) Markov Chains. Cambridge University Press, Cambridge.
https://doi.org/10.1017/CBO9780511810633

[18]   Rabiner, L. and Juang, B. (1986) An Introduction to Hidden Markov Models. *IEEE ASSP Magazine*, **3**, 4-6. https://doi.org/10.1109/MASSP.1986.1165342

[19]   Rabiner, L. (1989) A Tutorial on Hidden Markov Models and Selected Applications in Speech Recognition. *Proceedings of the IEEE*, **77**, 257-286.
https://doi.org/10.1109/5.18626

[20]   Devijver, P. (1985) Baum's Forward-Backward Algorithm Revisited. *Pattern Recognition Letters*, **3**, 369-373. https://doi.org/10.1016/0167-8655(85)90023-6

[21]   Lou, H. (1995) Implementing the Viterbi Algorithm. *IEEE Signal Processing Magazine*, **12**, 42-52. https://doi.org/10.1109/79.410439

[22]   Xuan, G.R., Zhang, W. and Chai, P.Q. (2001) EM Algorithms of Gaussian Mixture Model and Hidden Markov Model. *International Conference on Image Processing*, **1**, 145-148.

[23]   Pinto, R. and Engel, P. (2015) A Fast Incremental Gaussian Mixture Model. *PLoS ONE*, **10**, e0139931. https://doi.org/10.1371/journal.pone.0139931

[24]   Dua, D. and Gra, C. (2019) UCI Machine Learning Repository.University of California, School of Information and Computer Science, Irvine.
http://archive.ics.uci.edu/ml

[25]   Deisenroth, M., Faisal, A. and Ong, C. (2020) Mathematics for Machine Learning. Cambridge University Press, Cambridge. https://mml-book.com/
https://doi.org/10.1017/9781108679930