International Journal of Food Microbiology 141 (2010) 203–212

Contents lists available at ScienceDirect

International Journal of Food Microbiology

journal homepage: www.elsevier.com/locate/ijfoodmicro

Development of a multi-classification neural network model to determine the microbial growth/no growth interface

Francisco Fernández-Navarro [a,*], Antonio Valero [b], César Hervás-Martínez [a], Pedro A. Gutiérrez [a], Rosa M. García-Gimeno [b], Gonzalo Zurera-Cosano [b]

[a] Department of Computer Science and Numerical Analysis, University of Córdoba, Campus, de Rabanales, Albert Einstein Building, 3rd Floor, 14014 Córdoba, Spain
[b] Department of Food Science and Technology, University of Córdoba, Campus de, Rabanales, Darwin Building, 14014 Córdoba, Spain

ARTICLE INFO

Article history:
Received 8 December 2009
Received in revised form 4 May 2010
Accepted 15 May 2010

Keywords:
Multi-Classification
Sensitivity
Growth/no growth interface
Accuracy
Memetic algorithm
Imbalanced datasets
Oversampling method
SMOTE
S. aureus

ABSTRACT

Boundary models have been recognized as useful tools to predict the ability of microorganisms to grow at limiting conditions. However, at these conditions, microbial behaviour can vary, being difficult to distinguish between growth or no growth. In this paper, the data from the study of Valero et al. [Valero, A., Pérez-Rodríguez, F., Carrasco, E., Fuentes-Alventosa, J.M., García-Gimeno, R.M., Zurera, G., 2009. Modelling the growth boundaries of Staphylococcus aureus: Effect of temperature, pH and water activity. International Journal of Food Microbiology 133 (1–2), 186–194] belonging to growth/no growth conditions of *Staphylococcus aureus* against temperature, pH and a_w were divided into three categorical classes: growth (G), growth transition (GT) and no growth (NG). Subsequently, they were modelled by using a Radial Basis Function Neural Network (RBFNN) in order to create a multi-classification model that was able to predict the probability of belonging at one of the three mentioned classes. The model was developed through an over sampling procedure using a memetic algorithm (MA) in order to balance in part the size of the classes and to improve the accuracy of the classifier. The multi-classification model, named Smote Memetic Radial Basis Function (SMRBF) provided a quite good adjustment to data observed, being able to correctly classify the 86.30% of training data and the 82.26% of generalization data for the three observed classes in the best model. Besides, the high number of replicates per condition tested ($n = 30$) produced a smooth transition between growth and no growth. At the most stringent conditions, the probability of belonging to class GT was higher, thus justifying the inclusion of the class in the new model. The SMRBF model presented in this study can be used to better define microbial growth/no growth interface and the variability associated to these conditions so as to apply this knowledge to a food safety in a decision-making process.

1. Introduction

Boundary models have been arisen in the predictive microbiology field as an approach to determine the ability of growth of microorganisms. In this respect, several works have been published in recent years for both spoilage and pathogenic microorganisms, due to the necessity of gaining knowledge, by mathematical models, about the microbial behaviour in limiting conditions that just prevent growth.

Consequently, these mathematical models may lead to more realistic estimations of food safety risks, and can provide useful quantitative data for the development of processes that allow production of safer food products (Koutsoumanis et al., 2005). Several mathematical approaches have been developed based on deterministic estimates of minimal values of environmental parameters at which growth can occur (Pitt, 1992), polynomial and non-linear equations (Presser et al., 1998; Salter et al., 2001; Skandamis et al., 2007; Valero et al., 2007a) that can be built using a logistic regression procedure proposed by Ratkowsky and Ross (1995) or Artificial Neural Networks (ANNs) which can be applied to define growth/no growth interface of microorganisms (Hajmeer and Basheer, 2003).

The importance of the use of ANNs in predictive microbiology (Garcia-Gimeno et al., 2005; Zurera-Cosano et al., 2005; Hervás-Martínez et al., 2006) as an alternative to regression techniques was stated by Basheer and Hajmeer (2000) due to their flexibility and high degree of accuracy to fit to experimental data. It is important to realize that ANN's often contain a high number of parameters (Hajmeer et al., 1997). Although if we compare ANNs to polynomial models, the number of parameters of ANNs is competitive because the number of parameters increases in a non-linearly way when increasing the number of conditions and/or degree of the polynomial model (Geeraerd et al., 1998). However, ANNs offer a number of advantages, including ability to implicitly detect complex non-linear relationships between dependent and independent variables, and the ability to detect all possible interactions between predictor variables. Disadvantages include its "black box" nature, greater computational burden, and proneness to overfitting. Generally, sigmoid functions have been widely used to built the ANNs structure (Multi-Layer Perceptron type networks) (Leshno et al., 1993) together with other types of ANNs such as Gaussian networks,

* Corresponding author. Tel.: +34 957 21 83 49; fax: +34 957 21 83 60.
E-mail address: i22fenaf@uco.es (F. Fernández-Navarro).

0168-1605/$ – see front matter.
doi:10.1016/j.ijfoodmicro.2010.05.013

General Regression Neural Network models or Radial Basis Functions (Musavi et al., 1992; Mulgrew, 1996).

Hajmeer and Basheer (2002) used a Probabilistic Neural Network (PNN) approach for classification of bacterial growth/no growth data and modeling the probability of growth of a pathogenic *Escherichia coli* R31 in response to temperature and water activity. They found that PNNs were shown to outperform linear and non-linear regression models in both classification accuracy and ease. Later on, ANNs were used to predict the growth/no growth interface (Valero et al., 2007b) and survival/death of *E. coli* O157:H7 in mayonnaise model systems (Yu et al., 2006) or *Listeria monocytogenes* in chorizos (Hajmeer et al., 2006).

Many pattern classification systems were developed for two-class classification problems and theoretical studies of learning have focused almost entirely on learning binary functions (Natarajan, 1991) including the well-known support vector machines (SVM) (Vapnik, 1998) or ANN algorithms such as the perceptron or the error back propagation algorithm (Bishop, 1996). For most of these algorithms, the extension from two-class to the multi-class pattern classification problem is non-trivial, and often leads to an unexpected complexity or weaker performances (Anand et al., 1995; Price et al., 1995; Allwein et al., 2001; Gelenbe and Hussain, 2002).

Logistic regression models provide numerical values of probability of growth (p_G) being the categorization into the classes "growth" or "no growth" which are determined by setting a cut point (generally $P=0.5$). The main output of a logistic regression model can be a value of probability equal, for instance, to 0.23. This means that the chance of growth of a given microorganism is 23% (thus, no growth 77%). However, observed probabilities are subjected to different variability sources, which are especially important at limiting conditions (when probability of growth is around 0.5). This is mainly related to the variability in microbial responses since if the observed probability of growth at one specific condition is 0.23 it is highly probable that this value will not be the same if the microorganism is subjected again to the same condition. Besides this, conditions in which binary responses are observed can be catalogued as no growth, but, when repeating the experiment, these conditions could be classified as growth or vice versa. That is, conditions where the probability of growth is neither 0 nor 1 are highly dependent on microbial responses, which are also associated to high variability. Therefore, by adding new information about the probability of growth associated to these conditions, growth/no growth models can provide alternative and more accurate estimations. In the present study, the logistic regression model from the study of Valero et al. (2009) was modified by providing a categorical classification of *Staphylococcus aureus* growth as a function of temperature, pH and water activity (a_w) in three different classes: "growth" (p_G) that included conditions in which the probability of growth is equal to 1; "no growth" which were conditions where the probability of growth is 0, and a new class denominated "growth transition (GT)" that encompassed all conditions where probability of growth was different from 0 to 1. In our approach, we are modeling the probability of belonging of each pattern to each class instead of modeling the probability of growth or no growth.

At the best of our knowledge very few attempts in the predictive microbiology field that includes a multi-classification structure to model microbial growth/no growth have been performed. One approach was the study of Le Marc et al. (2005) that combined the concept of the Minimum Convex Polyhedron (MCP) previously introduced in Baranyi et al., (1996) with a logistic regression method to model microbial growth/no growth boundaries. They obtained predicted probabilities corresponding to zones of the model domain belonging to growth, no growth and uncertainty regions. However, the uncertainty region was built in zones where no data were available. Besides, the MCP was linked to microbial observations in ComBase (where no available data are found in some cases) and is has also to be used in combination with a logistic regression model.

We understand that there are some difficulties that can be encountered to develop a multi-classification procedure. Firstly, it would be necessary to increase the number of replicates per condition tested, and secondly, it is not so easy to determine, a priori, the size of each class evaluated. This fact can produce a non-balanced structure of these classes, which leads to decrease the performance capacity of the standard classifiers.

For this reason, in this work, a novel approach will be tackled by considering a non-balanced multi-classification model in three different classes (G, GT and NG). A radial basis function neural network (RBFNN) will be proposed as a classification tool and discriminant functions of G/GT/NG from the best model will be derived.

2. Learning methodology

2.1. Base classifier

We considered radial basis functions neural networks (RBFNNs) (Freeman and Saad, 1995; Orr, 1995; Hwang and Bang, 1997) with softmax outputs and the standard structure as the base classification model. A scheme of these models is given in Fig. 1, where J is the number of classes (in our case, $J=3$, since we had 3 classes: G, GT and NG) and m is the number of hidden nodes or RBFs of the neural net. The inputs of the neural net are represented by the vector $\mathbf{x}$, $f_l(\mathbf{x}, \theta_l)$ is the output of the neural net for each of the l-th class, and, after applying the softmax transformation, these outputs are transformed into probabilities that the pattern $\mathbf{x}$ belong to the corresponding class, $p_l(\mathbf{x}, \theta_l)$. Finally, $\theta=(\theta_1,\ldots, \theta_{J-1})$ is the vector including all the parameters of the neural net.

The activation function of the j-th node in the hidden layer is given by:

$$B_j(\mathbf{x}, \mathbf{w}_j) = \exp\left(-\frac{\|\mathbf{x}-\mathbf{c}_j\|^2}{r_j^2}\right), \quad (1)$$

where $\mathbf{w}_j=(\mathbf{c}_j, r_j)$ is the vector of parameters of the j-th hidden node, $\mathbf{c}_j=(c_{j1},\ldots, c_{jk})$ is the centre of this node, r_j is the corresponding radium and c_{ji} is the weight of the connection between the i-th input

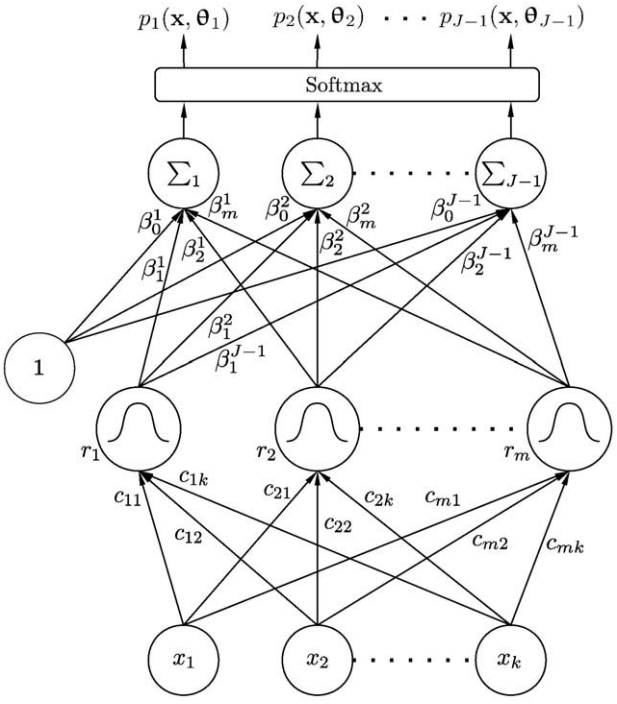

Fig. 1. Structure of Radial Basis Function Neural Networks.

node and the j-th RBF. The activation function of the l-th output node is given by:

$$f_l(\mathbf{x},\theta_l) = \beta_0^l + \sum_{j=1}^{m} \beta_j^l B_j(\mathbf{x},\mathbf{w}_j),\ l = 1,2,3 \tag{2}$$

where $\theta_l = (\beta_0^l, \beta_1^l, \ldots, \beta_m^l, \mathbf{w}_1, \ldots, \mathbf{w}_m)$, β_j^l is the weight of the connection between the j-th RBF hidden node and the l-th output node and β_0^l is the bias of the l-th output node. The transfer function of all output nodes is the identity function.

In order to tackle a multi-classification problem, the outputs of the model have been interpreted from the point of view of probability through the use of the softmax activation function (Richard and David, 1989), which is given by:

$$p_l(\mathbf{x},\theta_l) = \frac{\exp f_l(\mathbf{x},\theta_l)}{\sum_{j=1}^{3} \exp f_j(\mathbf{x},\theta_j)},\ l = 1,2,3 \tag{3}$$

where $f_j(\mathbf{x},\theta_l)$ is the output of the j output neuron for pattern $\mathbf{x}$ and $p_l(\mathbf{x},\theta_l)$ is the probability a pattern $\mathbf{x}$ has of belonging to class j.

Using the softmax activation function presented in Eq. (3), the class predicted by the ANN corresponds to the node in the output layer whose output value is the greatest. In this way, the optimum classification rule $C(\mathbf{x})$ is the following:

$$C(\mathbf{x}) = \hat{l},\ \text{where } \hat{l} = argmax_l p_l(\mathbf{x},\theta_l),\ \text{for } l = 1,2,3. \tag{4}$$

The best RBFNN is determined by means of a memetic algorithm (MA) (detailed in Section 2.4) that optimizes the error function given by the negative log-likelihood for N observations associated with the RBFNN model:

$$L^*(\theta) = \frac{1}{N}\sum_{n=1}^{N}\left[-\sum_{l=1}^{2} y_n^{(l)} f_l(\mathbf{x}_n,\theta_l) + \log\sum_{l=1}^{2}\exp f_l(\mathbf{x}_n,\theta_l)\right]. \tag{5}$$

where $y_n^{(l)}$ is equal to 1 if the pattern $\mathbf{x}_n$ belongs to the l-th class and equal to 0 otherwise. From a statistical point of view, the approach can be seen as a non-linear multinominal logistic regression, where we optimize log-likelihood using a MA.

2.2. Data preprocessing

Sampling strategies, such as over and undersampling, are extremely popular in tackling the problem of class imbalance, i.e., either the minority class is oversampled, the majority classes are undersampled, or some combination of the two is deployed (as is described in (Seiffert et al., 2010)).

In the preprocessing stage, we have applied an oversampling process to the minority class, in our case, the GT class. In this way, the initial dataset has been modified and more synthetic samples of this class have been created. Specifically, the number of minority class patterns (GT patterns) was doubled. The aim was to decrease the problem of imbalanced rate, by selecting the GT class for applying the re-sampling procedure, since this class included originally a half of the number of patterns of the other classes (G and NG).

Synthetic examples were obtained by applying the Synthetic Minority Oversampling Technique (SMOTE) algorithm (Chawla et al., 2002). SMOTE is an oversampling method where the minority class is oversampled by taking each minority class sample and introducing synthetic examples along the line segments joining any/all of the k minority class nearest neighbours. Depending upon the amount of oversampling required, neighbours from the k nearest neighbours are randomly chosen. Our implementation currently uses five nearest neighbours as the maximum value of the k parameter.

2.3. Base evolutionary algorithm

The basic framework of the evolutionary algorithm (EA) is the following: the search begins with an initial population of RBFNNs and, in each iteration, a population-update algorithm which evolves both its structure and weights is applied. The population is subjected to the operations of replication, mutation and recombination.

The general structure of the EA can be supported in the following steps:

1. Generate initial population with randomly generated networks.
2. Evaluate the fitness score for each individual of the population on the objective function.
3. Copy the best individual to the next generation.
4. The best 20% of the population substitutes the worst 20% of individuals.
5. Apply crossover operators to the best 10% of the population.
6. Apply parametric mutation operators to the second best 10% of the population.
7. Apply structural mutation operators to the rest of the population (80%).

We considered $L^*(\theta)$ defined in (Eq. (5)) as the error function of an individual g of the population. The fitness measure needed for evaluating the individuals is a strictly decreasing transformation of the error function $L^*(\theta)$ given by

$$A(g) = \frac{1}{1 + L^*(\theta)};\ 0 < A(g) \leq 1. \tag{6}$$

The crossover operator considered is the binary and multipoint crossover operators. The severity of a mutation to an individual RBFNN model is dictated by the temperature $T(g)$ of the RBFNN model. $T(g)$ is related to $A(g)$ by means of the expression $T(g) = 1 - A(g)$, $0 \leq T(g) < 1$ and, for that reason, $T(g)$ is in descent throughout the evolutionary process, making abrupt changes at the beginning (exploration) and soft changes at the end (exploitation), since it is supposed that the $A(g)$ of the individuals in the population must improve in each iteration of the evolutionary process.

Parametric mutation consists of a simulated annealing algorithm (Martínez-Estudillo et al., 2008). Structural mutation implies a modification in the structure of the RBFNNs and allows the exploration of different regions in the search space, helping to keep the diversity of the population. There are four different structural mutations: hidden node addition, hidden node deletion, connection addition and connection deletion. These four mutations are applied sequentially to each network. More information about genetic operators proposed can be seen in Gutiérrez et al. (2009, 2010).

2.4. Memetic algorithm: Smote Memetic Radial Basis Function

The Smote Memetic Radial Basis Function (SMRBF) consists of applying the previously described base evolutionary algorithm but including a local search to some specifically selected individuals. This memetic algorithm (MA) includes an optimization clustering process applied in specific stages of the evolutionary process. In this clustering process, each RBFNN model or individual is represented by the set of its accuracies per class. For example, an individual with the Confusion Matrix of the Eq. (7), is represented for the vector [52/60 = 0.86, 23/29 = 0.79, 51/57 = 0.89], since these values are its accuracies per class. The clustering algorithm is able to obtain groups of individuals that have a similar behaviour for the different classes.

$$Confusion\ Matrix = \begin{pmatrix} 52 & 6 & 2 \\ 3 & 23 & 3 \\ 0 & 6 & 51 \end{pmatrix}. \tag{7}$$

After that, we apply *iRprop* + algorithm (Igel and Hüsken, 2003) to the individual closest to the centroid obtained in each cluster. It is

important to note that each cluster has been determined by means of the standard k-means (Fukunaga, 1999) applied over the specific space previously stated. Finally, the optimized individuals are returned to the population with its fitness and values updated since our MA is based on the Lamarckian model (Whitley et al., 1994).

The SMRBF algorithm is detailed in 2. In order to evaluate the effect of the preprocessing step where the GT class is oversampled, we also considered what we call the Memetic Radial Basis Function (MRBF) algorithm. This algorithm follows the same flow diagram that SMRBF methodology but it does not include the preprocessing stage.

3. Computational experiments

3.1. Database description

The original dataset was taken from Valero et al. (2009) describing the growth/no growth boundaries of a five strain cocktail of *S. aureus* as a function of temperature, pH and a_w by an ordinary logistic regression model. Data were collected at 8, 10, 13, 16 and 19 °C at pH levels from 4.5 to 7.5 (0.5 intervals) and at 19 levels of a_w (from 0.856 to 0.999 at regular intervals). The conditions in which *S. aureus* always grows have been labeled as growth (G), those in which never grows as no growth (NG), and finally, those for which sometimes grows and sometimes not (it grows between 1 and 29 times of the 30 replicates tested per condition), as growth transition (GT). For data processing, out of 287 conditions performed, 146 were selected for model training and 141 were chosen for model generalization. From the 146 conditions selected to train the model, 60 conditions were classified as G, 29 as GT and 57 as NG. For the conditions used to validate the performance of the model (141 conditions), 57 were classified as G, 28 as GT, and 56 as NG. More details can be found in Valero et al. (2009).

3.2. Alternative statistical and artificial intelligence methods used for comparison purposes

Different state-of-the-art statistical and artificial intelligence algorithms have been implemented for comparison purposes. In this way, the proposed method was compared to the following algorithms:

- The MRBF method (detailed in Section 2.4). As our approach applies an oversampling procedure in the preprocessing stage, it is necessary to compare its performance to the original MRBF method.
- Multi-logistic regression methods (Witten and Frank, 2005):
  - MultiLogistic (MLogistic): It is an algorithm for building a multinomial logistic regression model with a ridge estimator to guard against overfitting by penalizing large coefficients, based on the work by le Cessie and van Houwelingen (le Cessie and van Houwelingen, 1992). In order to find the coefficient vector, a Quasi-Newton Method is used.
  - SimpleLogistic (SLogistic): It is based on applying LogitBoost algorithm with simple regression functions and determining the optimum number of iterations by a five fold cross-validation: the data is equally split five times into training and test, LogitBoost is run on every training set up to a maximum number of iterations (500) and the classification error on the respective test set is logged. Afterwards, LogitBoost is run again on all data using the number of iterations that gave the smallest error on the test set averaged over the five folds. Further details about the algorithm can be found in Landwehr et al. (2005).

We consider the pH, water activity (a_w) and temperature (T) as the initial co-variates (MLogistic(standard model) and SLogistic(standard model)). The model can be expressed as:

$$Y = b_0 + b_1 \cdot T + b_2 \cdot pH + b_3 \cdot a_w \tag{8}$$

where Y is the dependent variable, b_0 the intercept of model, and b_1, b_2, b_3 the partial regression coefficients. In order to allow a fair comparison between the new developed model and existing classic approaches, we applied the SMOTE algorithm in the patterns of the GT class in combination with the standard logistic regression models.

Furthermore, we compare our approach to logistic regression methods with square and cross product terms in the model (as suggested in Le Marc et al. (2005))): MLogistic(Le Marc et al. (2005) model) and SLogistic(Le Marc et al. (2005) model). This model is expressed as:

$$Y = b_0 + b_1 \cdot T + b_2 \cdot pH + b_3 \cdot a_w + b_4 \cdot T \cdot pH + b_5 \cdot T \cdot a_w + b_6 \cdot pH \cdot a_w + b_7 \cdot T \cdot pH \cdot a_w. \tag{9}$$

As we did with the standard logistic regression models, the SMOTE algorithm was also applied to the pattern of the GT class in combination with the Le Marc et al. (2005) approach.

- A Gaussian RBF Network (RBFN) (Nabney, 2004), deriving the centres and width of hidden units using k-means and combining the outputs obtained from the hidden layer using logistic regression. k-means is applied separately to each class to derive k clusters for each class.
- The C - SVM algorithm (Hastie et al., 2001) with RBF kernels (SVM). From a structural point of view, the SVMs are related to RBFNNs and they have become one of the most popular and developed methods nowadays. In order to face the multi-class case, a "1-against-1" approach has been considered, following the recommendations of Hsu and Lin (Chen et al., 2009).

We also compared our proposal to specific methods for imbalanced data: the OverSampling and SmoteOverSampling methods proposed in Zhou and Liu (2006). These methods have been selected due to their similarities to the model proposed. They use MLP neural networks as the base classifier, and the model is trained by the RProp algorithm. The main differences with our approach are the following: our model is trained by a MA and we used RBFNN as the base classifier.

3.3. Performance measures and algorithms' parameters

Classifiers were evaluated by two measures derived from the confusion matrix: the Correct Classification Rate (*CCR*) and the Minimum Sensitivity (*MS*) over the generalization dataset.

The contingency or confusion matrix $M(g)$ for a classification problem with J classes, N training or generalization patterns and g as classifier is given by the following expression:

The *CCR* measure or accuracy is defined as:

$$M = \left\{ n_{ij};\ \sum_{i,j=1}^{J} n_{ij} = N \right\} \tag{10}$$

where n_{ij} represents the number of times the patterns are predicted by classifier g to be in class j when they really belong to class i. The diagonal corresponds to correctly classified patterns and the off-diagonal to mistakes in the classification task.

$$CCR = (1/N) \sum_{j=1}^{J} n_{jj}, \tag{11}$$

that is, the rate of all the correct predictions.

Let $S_i = n_{ii}/f_i$ be the number of patterns correctly predicted to be in class i with respect to the total number of patterns in class i (sensitivity for class i). The *MS* measure is defined as:

$$MS = \min \{S_i;\ i = 1, \ldots, J\} \tag{12}$$

that is, the accuracy for the class that is worst classified.

The parameter values used in the hybrid techniques proposed were the following: we performed a simple linear rescaling of the input variables in the interval $[-2, 2]$, X_i^* being the transformed variables. The connection between hidden and output layer were initialized in the $[-5, 5]$ interval. The initial value of the radii r_j was obtained in the interval $(0, d_{max}]$, where d_{max} is the maximum distance between two training input examples.

The size of the population was $N=500$. For the structural mutation, the number of nodes that can be added or removed was within the [1,2] interval, and the number of connections to add or delete in the hidden and the output layer during structural mutations was within the [1,7] interval. The number of clusters was $k=6$ for the k-means algorithm. The *iRprop+* local improvement procedure was performed every 50 generations, 8 times during the evolution. In this way, the algorithm stopped when 400 generations were completed. For the *iRprop+* algorithm, we consider a maximum of 75 cycles.

For the selection of the SVM hyperparameters (regularization parameter, C, and width of the Gaussian functions, γ), a grid search algorithm was applied with a ten-fold cross-validation, using the following ranges: $C \in \{2^{-5}, 2^{-3}, \ldots, 2^{15}\}$ and γ $\{2^{-15}, 2^{-13}, \ldots, 2^{3}\}$.

For the MRBF, SMRBF and the specific methods for imbalanced data, we ran the procedures 30 times because they are based on random values and do not return the same result for each execution. For the other methods, the results were obtained running them only one time, because all are deterministic methods.

The MRBF algorithm was implemented in JAVA. For the SMRBF method, the MRBF algorithm was slightly modified, applying the oversampling procedure in the preprocessing stage. The base evolutionary algorithm (with several modification to the proposed in this paper) is available in KEEL[1]. KEEL is a software tool to assess evolutionary algorithms for Data Mining problems including regression, classification, clustering, pattern mining and so on (Alcala-Fdez et al., 2008).

We also used "libsvm" (Chang and Lin, 2001) to obtain the results of the SVM method, WEKA (Witten and Frank, 2005) to obtain the results of the RBFN, MLogistic and SLogistic and the CSNN[2] software package to obtain the results of the OverSampling and SmoteOverSampling methods.

4. Results and discussion

4.1. Statistical analysis

A comparison of the SMRBF method with well-known classification techniques given in 3.2 has been carried out. Table 1 shows the results obtained with the different techniques tested. A descriptive analysis of the results leads to the following remarks: the SMRBF method obtained the best result in terms of MS_G and CCR_G over all techniques compared.

To ascertain the statistical significance of the differences between the means (in CCR_G and MS_G for each stochastic methodology: OverSampling, Smote, MRBF and SMRBF), the non-parametric Kolmogorov–Smirnov test (K–S test) was used to evaluate if the CCR_G and MS_G values followed a normal distribution. The K–S test showed that a normal distribution can be assumed because the critical levels, P-values, were higher than 0.05 in all cases. In order to determine the best methodology, an ANOVA statistical method test was carried out. The results of the ANOVA analysis for the CCR_G and MS_G values show that the methodology effect was statistically significant at a level of signification of 5%. Once this test guaranteed that there were significant differences between the results of the different methods, we performed a multiple comparison test on the CCR_G and MS_G values in order to establish an order between the different methods. First, we carried out a Levene test (Miller, 1996) for evaluating the equality of variances. Then, we performed a Tamhane test (Tamhane and Dunlop, 2000), because the variances are not equal (either for CCR_G or MS_G), in order to rank the different methods. Our aim was to find the methodology whose performance (in *CCR* and *MS*) was significantly better than that of the rest of the methodologies.

Table 1
Comparison with other statistical and artificial intelligence methods: Correct Classification Rate and Minimum Sensitivity in the generalization set (CCR_G(%) and MS_G(%), respectively).

Method	CCR_G (%)	MS_G (%)
MLogistic(standard model)	76.60	39.29
SLogistic(standard model)	76.60	32.14
SMOTE + MLogistic(standard model)	71.63	50.00
SMOTE + SLogistic(standard model)	70.92	50.00
MLogistic((Le Marc et al., 2005) model)	80.56	50.00
SLogistic((Le Marc et al., 2005) model)	75.88	32.14
SMOTE + MLogistic((Le Marc et al., 2005) model)	75.17	53.57
SMOTE + SLogistic((Le Marc et al., 2005) model)	74.46	46.42
RBFN	75.18	39.29
SVM	*80.98*	42.86
OverSampling	78.58 ± 2.24	52.14 ± 8.11
SmoteOverSampling	75.60 ± 4.03	60.21 ± 13.29
MRBF	79.71 ± 3.34	53.21 ± 9.59
SMRBF	**81.07 ± 1.60**	**75.32 ± 3.28**

The best result is in bold face and the second best result in italics.

If we analyze the average results for accuracy *CCR* we can observe that the SMRBF methodology obtained better results than those obtained with other methodologies. On the other hand, the results of the average *MS* show that the SMRBF methodology obtained a significant better performance, for a level of signification of 5%, than the other methodologies. Therefore, SMRBF is the classification methodology recommended in this paper for the problem analyzed.

It has been proven in Table 1 that the application of the SMOTE algorithm in combination with logistic regression techniques is not suitable, since the final model obtained less accuracy in correctly classifying the generalization data. The application of the SMOTE algorithm in combination to the logistic regression models improved the Minimum Sensitivity (*MS*) results but decreased the accuracy results (*CCR*). *A priori*, we could think that *MS* and *CCR* objectives could be positively correlated, but while this may be true for small values of *MS* and *CCR*, it is not for values close to 1 on both *MS* and *CCR*, as we can see in the results obtained by the logistic regression models (see Table 1). Despite this, the SMRBF yields improve the *CCR* and the *MS* values obtained by the memetic algorithm (MRBF method).

In general, these results show that the proposed approaches based on RBFNNs are robust to tackle the multi-classification of the growth boundaries of *S. aureus*, obtaining better results than the majority of the existing alternative methods.

4.2. Analysis of the best SMRBF model

4.2.1. Model implementation

We selected the best SMRBF model of the different models obtained in order to perform a further analysis of it. For a correct implementation of the SMRBF multi-classification model it would be necessary: (i) to evaluate its performance capacity, (ii) to compare the observations with predictions, by considering a classification into G, GT and NG classes and finally, (iii) to study the conditions that imply a transition from G to GT and NG and vice versa.

4.2.2. Classification accuracy of the SMRBF model

The outputs of this model are the values of probability that a pattern falls within each class: G (p_G), GT (p_{GT}) and NG (p_{NG}). We considered the softmax activation function in such a way that each pattern is associated to the class with higher probability (Eq. (4)).

It should be highlighted that the robustness of the SMRBF model is given by the number of replicates tested per condition ($n=30$) which provided a more reliable classification in three classes. These results are

[1] http://www.keel.es.
[2] http://lamda.nju.edu.cn/datacode/CSNN.htm.

Table 2
Probability expression of the best SMRBF model. Performance of this model: Correct Classification Rate (*CCR*) on the training set considering the synthetic SMOTE data (CCR_{TS}) and not considering the synthetic SMOTE data (CCR_T), CCR on the generalization set (CCR_G), Minimum Sensitivity (MS) on the training set considering the synthetic SMOTE data (MS_{TS}) and not considering the synthetic SMOTE data (MS_T) and MS on the generalization set (MS_G). Confusion Matrix (CM) for the training set considering the synthetic SMOTE data (CM_{TS}), not considering the synthetic SMOTE data (CM_T) and CM for the generalization set (CM_G).

Best SMRBF *S. aureus* multi-classification model

$$p_{NG}(\mathbf{x},\theta)=\frac{e^{f_1(\mathbf{x},\theta)}}{1+\sum_{i=1}^{2}e^{f_i(\mathbf{x},\theta)}};\ p_{GT}(\mathbf{x},\theta)=\frac{e^{f_2(\mathbf{x},\theta)}}{1+\sum_{i=1}^{2}e^{f_i(\mathbf{x},\theta)}};\ p_G(\mathbf{x},\theta)=\frac{e^{f_3(\mathbf{x},\theta)}}{1+\sum_{i=1}^{2}e^{f_i(\mathbf{x},\theta)}}$$

$$f_1(\mathbf{x},\theta)=-3.05-31.32RBF_1+36.33RBF_2-2.42RBF_3-7.47RBF_4+1.57RBF_5$$

$$f_2(\mathbf{x},\theta)=-3.50-6.59RBF_1+17.84RBF_2-4.30RBF_3+2.14RBF_5$$

$$f_3(\mathbf{x},\theta)=0$$

$$RBF_1=e^{-0.5\cdot\left(\frac{\left((T^*+2.07)^2+(a_w^*+2.27)^2\right)^{0.5}}{2.26}\right)^2}$$

$$RBF_2=e^{-0.5\cdot\left(\frac{\left((T^*-1.04)^2+(pH^*-0.49)^2+(a_w^*-1.77)^2\right)^{0.5}}{1.96}\right)^2}$$

$$RBF_3=e^{-0.5\cdot\left(\frac{\left((T^*+2.03)^2+(pH^*-0.46)^2+(a_w^*+1.04)^2\right)^{0.5}}{1.28}\right)^2}$$

$$RBF_4=e^{-0.5\cdot\left(\frac{\left((T^*-1.34)^2\right)^{0.5}}{0.21}\right)^2}$$

$$RBF_5=e^{-0.5\cdot\left(\frac{\left((pH^*-1.92)^2\right))^{0.5}}{1.53}\right)^2}$$

$T^*, pH^*, a_w^*\ [-2,2]$

$CCR_{TS}=86.28\%$, $CCR_T=86.30\%$, $CCR_G=82.26\%$

$MS_{TS}=82.75\%$, $MS_T=79.31\%$, $MS_G=78.57\%$

$$CM_{TS}=\begin{pmatrix}52&6&2\\6&48&4\\0&6&51\end{pmatrix};\ CM_T=\begin{pmatrix}52&6&2\\3&23&3\\0&6&51\end{pmatrix};\ CM_G=\begin{pmatrix}48&8&1\\4&22&2\\1&9&46\end{pmatrix}$$

described in the confusion matrices associated (Table 2). The classification accuracy of the SMRBF model was high since more than 80% of the cases matched with the observed classes. Regarding the training dataset, 86.30% of the cases were correctly classified (CCR_T), while in the generalization dataset, this percentage was slightly lower (82.26%, CCR_G). Misclassified cases were assigned when the estimated pattern was not associated to the same observed class. As three classes were considered, in this model the errors accounted from the classes G, GT and NG to the adjacent ones; and from the class G to NG and vice versa. This approach allows to a better model accuracy than logistic regression models which consider the two categorical classes; growth and no growth.

Based on this, the percentage of errors that were misclassified in the adjacent class was 12.32% (18/146) for training data and 15.30% (23/141) for generalization data. It should be remarked that two cases in training data: (i) T 8 °C; pH 5; a_w 0.995; (ii) 10 °C; pH 4.5; a_w 0.989, and one case in generalization data (i) T 8 °C; pH 5; a_w 0.999, were misclassified as NG when G was observed; and one case in generalization data (i) T 10 °C; pH 5; a_w 0.977, was misclassified as G when NG was observed. These two cases accounted from the 1.41% and belonged to conditions that approached growth limits. However the majority of growth/no growth studies agreed that limiting conditions, in which microbial responses could be more variable, are recognized to be difficult to predict accurately (Gysemans et al., 2007a; Gysemans et al., 2007b).

4.2.3. Model predictions as a function of the environmental factors studied

The probability of falling within the class G (probability of growth = 1), GT (probability of growth between 0 and 1) and NG (probability of growth = 0) can be obtained by solving the general equation of the SMRBF model. Note that this is a pure categorical model, against the ordinary regression logistic models previously published in the field of predictive microbiology. As an example, at 13C, pH 5.50 and a_w 0.93; the outputs of the SMRBF model would be: $p_G=0.28$; $p_{GT}=0.64$ and $p_{NG}=0.08$. This means that the model gives a probability (certainty) of 0.28 that it is a

Table 3
Observed conditions from the study of (Valero et al., 2009) in which the replicates tested in training data showed simultaneously growth and no growth and comparison to the predictions given by the SMRBF model.

T(°C)	pH	a_w	p_{obs_log}	p_{pred_log}	p_G	p_{GT}	p_{NG}	p_{max}	Classification
8	7.50	1.00	26/30 = 0.87[a]	1.00	0.42	0.47	0.12	0.47	C[b]
8	7.50	0.99	7/30 = 0.23	0.02	0.20	0.59	0.21	0.59	C
8	6.50	0.99	24/30 = 0.81	0.93	0.59	0.34	0.07	0.59	I
8	5.50	0.99	11/30 = 0.37	0.38	0.29	0.32	0.38	0.38	I
8	7.00	0.98	11/30 = 0.37	0.06	0.27	0.58	0.15	0.58	C
8	6.00	0.98	6/30 = 0.20	0.26	0.35	0.41	0.24	0.41	C
10	5.50	0.96	16/30 = 0.53	0.96	0.33	0.50	0.18	0.50	C
10	7.00	0.96	23/30 = 0.77	0.95	0.17	0.74	0.08	0.74	C
10	6.00	0.96	17/30 = 0.57	0.96	0.30	0.56	0.14	0.56	C
10	7.50	0.95	14/30 = 0.47	0.02	0.01	0.65	0.34	0.65	C
10	6.50	0.95	16/30 = 0.53	0.76	0.10	0.68	0.22	0.68	C
13	4.50	0.96	19/30 = 0.63	0.99	0.27	0.44	0.28	0.44	C
13	7.50	0.94	28/30 = 0.93	1.00	0.51	0.48	0.01	0.51	I
13	5.50	0.94	29/30 = 0.97	1.00	0.84	0.15	0.00	0.84	I
13	5.00	0.93	29/30 = 0.97	0.39	0.06	0.59	0.35	0.59	C
13	7.50	0.92	26/30 = 0.87	0.99	0.02	0.87	0.11	0.87	C
13	5.50	0.92	26/30 = 0.87	0.79	0.08	0.70	0.22	0.70	C
13	7.00	0.92	28/30 = 0.93	0.93	0.01	0.83	0.16	0.83	C
13	6.00	0.92	26/30 = 0.87	0.73	0.03	0.70	0.27	0.70	C
16	7.50	0.92	28/30 = 0.93	1.00	0.77	0.23	0.00	0.77	I
16	5.50	0.92	27/30 = 0.90	1.00	0.94	0.06	0.00	0.94	I
16	5.00	0.92	16/30 = 0.53	0.84	0.21	0.60	0.19	0.60	C
16	7.50	0.91	16/30 = 0.53	1.00	0.09	0.86	0.05	0.86	C
16	6.00	0.90	16/30 = 0.53	0.96	0.12	0.77	0.11	0.77	C
16	6.50	0.89	16/30 = 0.53	0.50	0.02	0.76	0.23	0.76	C
16	7.00	0.88	14/30 = 0.47	0.01	0.00	0.51	0.49	0.51	C
19	7.50	0.88	23/30 = 0.76	0.98	0.13	0.55	0.32	0.55	C
19	6.50	0.88	26/30 = 0.87	0.95	0.38	0.45	0.17	0.45	C
19	7.00	0.87	7/30 = 0.23	0.09	0.09	0.51	0.40	0.51	C

p_{obs_log}: Observed Probabilities.
p_{pred_log}: Predicted Probabilities from the logistic model of (Valero et al., 2009).
p_G: Probability to pertain to the G class.
p_{GT}: Probability to pertain to the GT class.
p_{NG}: Probability to pertain to the NG class.
p_{max}: $maxp_l(\mathbf{x}, l)$, for $l=1, 2, 3$.
[a] Observed probabilities expressed as number of grown replicates/total replicates and the numerical value of probability growth.
[b] Classification criteria: C (Correct); I (Incorrect).

growth condition, 0.08 that it is a no growth condition and 0.64 that it is in a range where both growth and no growth results can be obtained. Therefore, the information originated from this model is the probability of pertaining to one of these three classes, and not a numerical value of probability of growth.

In Tables 3 and 4, the observed probabilities of growth obtained at boundary conditions were compared to the predicted probabilities of the logistic regression model of Valero et al. (2009) and the SMRBF model. These conditions should theoretically approach to the class GT. Regarding training data, 23 out the 29 conditions obtained the highest value of probability for the class GT, while in generalization data, this response was observed in 17 out the 25 conditions obtained. However, by considering three different classes, the errors committed in the misclassified cases are not severe. As it can be seen in Tables 3 and 4, the misclassified cases by the SMRBF model fell in the adjacent class where observed probabilities were obtained. For example, in Table 3 (8 °C, pH 6.50 and a_w 0.990), the observed probability of growth was 0.81 (which is in fact catalogued as "growth" by the logistic regression model). In this case, the SMRBF predicts that 34% of cases will belong to the GT class. However, the largest probability predicted by the SMRBF model was 59% of cases that belong to the G class. The observed probability indicated that growth mostly occurs, thus, the error of the SMRBF model is deviated to the class that approached the most to the real observation.

It should be highlighted that predictions given by the logistic regression model of Valero et al. (2009) approached to 0 and/or 1 in several cases at the boundary conditions represented in Tables 3 and 4, probably due to the convergence of the polynomial model in these zones of the model domain. In these cases, the use of the SMRBF model

Table 4
Observed conditions from the study of (Valero et al., 2009) in which the replicates tested in generalization data showed simultaneously growth and no growth and comparison to the predictions given by the SMRBF model.

T(°C)	pH	a_w	p_{obs_log}	p_{pred_log}	p_G	p_{GT}	p_{NG}	p_{max}	Classification
8	7.50	0.99	12/30 = 0.40[a]	0.65	0.32	0.53	0.15	0.53	C[b]
8	6.00	0.98	16/30 = 0.53	0.90	0.55	0.32	0.13	0.55	I
8	5.00	0.98	7/30 = 0.23	0.00	0.06	0.19	0.74	0.74	I
8	6.50	0.98	20/30 = 0.66	0.34	0.40	0.46	0.14	0.46	C
10	6.00	0.96	23/30 = 0.77	1.00	0.64	0.33	0.03	0.64	I
10	7.50	0.95	18/30 = 0.60	0.30	0.04	0.77	0.19	0.77	C
10	6.50	0.95	24/30 = 0.80	0.99	0.34	0.59	0.07	0.59	C
10	7.00	0.94	17/30 = 0.57	0.44	0.04	0.75	0.21	0.75	C
10	6.00	0.94	10/30 = 0.33	0.54	0.07	0.56	0.36	0.56	C
13	4.50	0.96	29/30 = 0.97	1.00	0.56	0.31	0.12	0.56	I
13	5.00	0.94	29/30 = 0.97	0.99	0.31	0.56	0.13	0.56	C
13	7.50	0.93	27/30 = 0.90	1.00	0.15	0.82	0.03	0.82	C
13	5.50	0.93	29/30 = 0.97	1.00	0.45	0.50	0.04	0.50	C
13	5.00	0.92	7/30 = 0.23	0.00	0.01	0.36	0.64	0.64	I
13	6.05	0.91	26/30 = 0.87	0.95	0.04	0.81	0.16	0.81	C
16	5.05	0.93	29/30 = 0.96	1.00	0.99	0.01	0.00	0.99	I
16	4.50	0.93	19/30 = 0.63	0.99	0.30	0.46	0.24	0.46	C
16	7.50	0.91	18/30 = 0.60	1.00	0.36	0.63	0.01	0.63	C
16	5.50	0.91	19/30 = 0.63	1.00	0.70	0.29	0.01	0.70	I
16	7.50	0.89	12/30 = 0.40	0.99	0.02	0.82	0.16	0.82	C
16	6.50	0.89	29/30 = 0.96	1.00	0.14	0.81	0.05	0.81	C
16	7.00	0.88	29/30 = 0.96	0.63	0.01	0.77	0.22	0.77	C
19	7.50	0.88	28/30 = 0.93	1.00	0.35	0.52	0.13	0.52	C
19	6.50	0.88	29/30 = 0.96	1.00	0.74	0.23	0.03	0.74	I
19	7.00	0.87	29/30 = 0.96	0.99	0.28	0.54	0.18	0.54	C

p_{obs_log}: Observed Probabilities.
p_{pred_log}: Predicted Probabilities from the logistic model of (Valero et al., 2009).
p_G: Probability to pertain to the G class.
p_{GT}: Probability to pertain to the GT class.
p_{NG}: Probability to pertain to the NG class.
p_{max}: $max p_l(\mathbf{x}, l)$, for $l = 1, 2, 3$.
[a] Observed probabilities expressed as number of grown replicates/total replicates and the numerical value of probability growth.
[b] Classification criteria: C (Correct); I (Incorrect).

can provide more accurate predictions and give also additional information regarding the variability of microbial responses at limiting conditions.

This growth transition is illustrated at some combinations of temperature (10 °C), pH (6.5) and a_w (0.949), in Fig. 3 for the SMRBF model (Fig. 3a, c and e) and for the logistic regression model of Valero et al. (2009) (Fig. 3a, d and f). The largest predicted probabilities for each model are marked, being each drawing line representing the classes G, GT and NG (SMRBF model); and the classes G and NG (logistic regression model). Regarding temperature, the SMRBF model predicted a smoother transition zone between 9.5 and 11 °C (Fig. 2). The cut point predicted by the logistic regression model (approximately 10 °C, Fig. 2) fell inside this zone. Observed probability at 10 °C, pH 6.5 and a_w 0.949 indicated growth transition ($P = 0.53$). This means that, although the two models correctly explained the microbial behaviour, the SMRBF model provides a temperature range where microbial responses are variable, instead of a deterministic value.

In Fig. 3c, the largest probabilities for the SMRBF model corresponded to GT and NG classes. It can be observed that for pH values above 5.75, GT is predicted. This fact is corroborated by observed probabilities obtained at these conditions. However, the logistic regression model predicted growth between pH 6 and 7 (Fig. 2), and no growth outside this range.

Finally, predictions obtained by the SMRBF model (10 °C pH 6.5) at different a_w values, indicated that a transition zone occurred at a_w between 0.941 and 0.959 (Fig. 3e), while according to the logistic regression model, at a_w 0.946, $p = 0.5$ (Fig. 3f). Again, the SMRBF model predicted a smoother transition zone, gaining accuracy in the estimations of the behaviour of *S. aureus* at conditions approaching the growth limits.

5. Application of the SMRBF model in food microbiology in a decision-making process

The main utility of the SMRBF model is its inclusion as a novel tool in a decision-making process within a Hazard Analysis and Critical Control Point (HACCP) system. The information generated allows risk managers to decide the most proper combinations to be followed in order to prevent microbial growth through giving more information regarding microbial variability associated to specific environmental conditions where a binary response (growth and no growth) can be observed. In Table 5, some combinations of pH and a_w levels at 8 and 10 °C are represented, comparing the estimations of the logistic regression model from Valero et al. (2009), and the SMRBF model (fixing $p_G = 0.1$). One of the main applications of the latter model is that one can decide which formulation is safer at a fixed probability value. For instance, at 8 °C there are some formulations that could potentially generate the same value of $p_G = 0.1$ (10% of cases the probability would be equal to 1). It can be seen that the safest formulation was obtained at 8 °C, pH 5.00 and a_w 0.993. In this case, p_{NG} had the highest value (0.69), while in 21% of cases, the probability will be neither 0 nor 1, but an intermediate (unknown) value. For food safety purposes, this value should be as lowest as possible, what implies that variability associated to the predicted microbial response would be reduced. At T 10 °C, pH 5.00 and a_w 0.967, the information given by the SMRBF model states that $p_G = 0.10$, $p_{GT} = 0.40$ and $p_{NG} = 0.50$ ($P = 1$ in 10% of cases; $0 < P < 1$ in

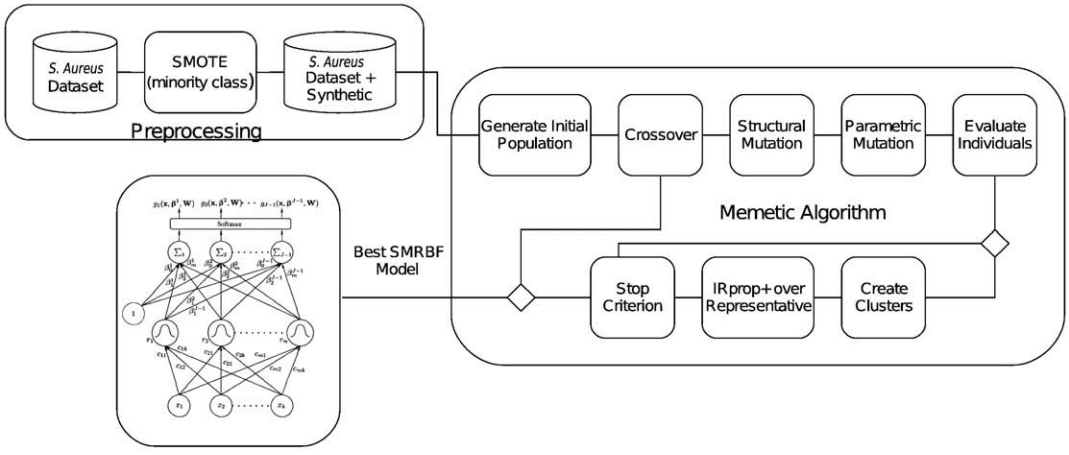

Fig. 2. Flow diagram of the SMRBF method.

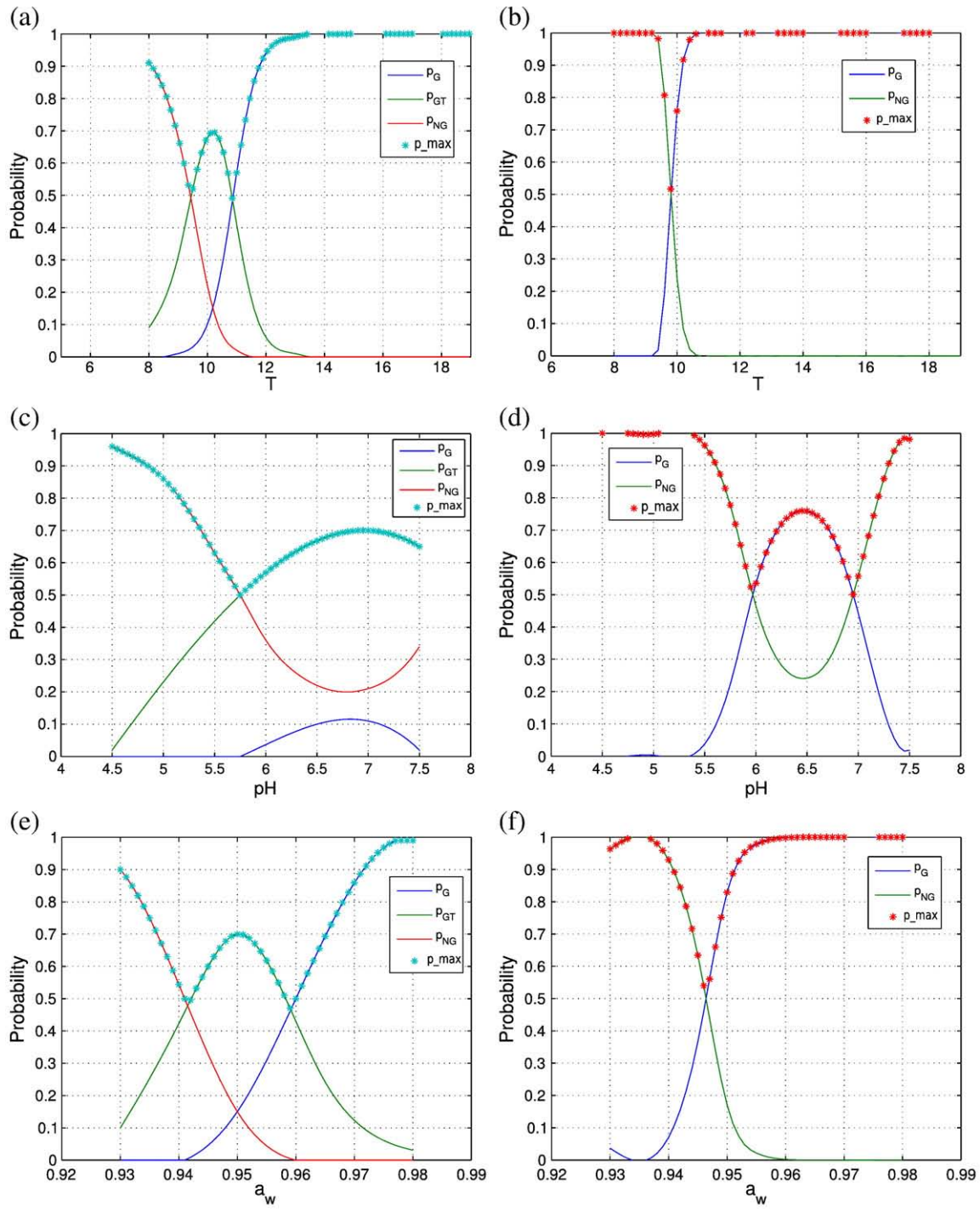

Fig. 3. Evolution of predicted probabilities (p) to pertain at the classes G (Growth), GT (Growth Transition) and NG (No growth), and the maximum of the probabilities (p_{max}) for the SMRBF model (3a, 3c and 3e) and for the logistic regression model of (Valero et al., 2009) (Figs. 3b, d and f), as a function of temperature (Figs. 3a and b: pH 6.5; a_w 0.949), pH (Figs. 3c and d: T 10 °C; a_w 0.949) and a_w (Figs. 3e and f: T 10 °C; pH 6.5).

40% of cases and $P=0$ in 50% of cases). The estimated variability associated to this condition is provided since in 40% of cases, the probability of growth will be between 0 and 1 and in 50% of cases, this probability will be 0. The estimated probability of the logistic regression model is 0.41 ($P=1$ in 41% of cases; $P=0$ in 59% of cases). At T 10 °C and at higher pHs, as seen in Table 5, p_{GT} is getting higher and p_{NG} is lowered. For instance, at pH 7.00 and a_w 0.953, although p_G is 0.1, there is a higher value of probability associated to the class GT (0.77), thus being this condition less safe. This result is in agreement with the logistic regression model, since the estimated value is $P=0.82$. In this latter condition, the logistic regression model predicted growth but the SMRBF model estimated that the probability will be placed within the class GT, being this condition associated to a high variability of the microorganism.

It should be mentioned that this is an alternative (categorical) approach to conventional logistic regression models, especially useful in decision-trees of HACCP systems when formulating new trend and/or minimally processed products.

The SMRBF model proposed in this study has been proven to be a very useful tool to model the microbial behaviour by determining the

Table 5
Comparison of the estimations of the logistic regression model from (Valero et al., 2009), and the SMRBF model (fixing $p_G = 0.1$) in some combinations of pH and a_w levels at 8 and 10 °C.

T(°C)	pH	a_w	P_{pred_log}	p_G	p_{GT}	p_{NG}
8	5.00	0.993	0.05	0.10	0.21	0.69[a]
8	5.60	0.978	0.00	0.10	0.33	0.57
8	6.40	0.972	0.00	0.10	0.50	0.40
8	6.75	0.973	0.00	0.10	0.58	0.32
8	7.00	0.975	0.00	0.10	0.62	0.28
10	5.00	0.967	0.41	0.10	0.40	0.50[a]
10	5.60	0.955	0.57	0.10	0.54	0.36
10	6.40	0.949	0.75	0.10	0.66	0.24
10	6.80	0.951	0.81	0.10	0.74	0.16
10	7.00	0.953	0.82	0.10	0.77	0.13

p_{pred_log}: Predicted Probabilities from (Valero et al., 2009).
p_G: Probability to pertain to the G class.
p_{GT}: Probability to pertain to the GT class.
p_{NG}: Probability to pertain to the NG class.
[a] Safest conditions selected.

probability to pertain at one of the three possible classes that can be obtained from probability models; growth, no growth and growth transition.

6. Conclusions

A novel SMRBF model was proposed which consisted of a radial basis function neural network optimized by using a memetic algorithm combined with an oversampling procedure for dealing with the imbalanced nature of the dataset. This new methodology derived in an interpretable equation that provided accurate predictions in means of the generalization data regarding values of *CCR* (81.07%) and MS (75.32%) as previously shown. The existence of a new class, named GT, has been included in the model mainly due to the high number of replicates per condition (30) used in our study, which produced a smoother transition between growth and no growth zones. This class is clearly justified since in certain zones of the model domain, microbial responses are more variable and therefore, a classification into G or NG cannot be established. In these conditions which approach microbial growth limits, the probability to pertain at the class GT is higher. This approach can help predictive modellers to better define the growth boundaries of microorganisms and to model the microbial variability associated to these conditions.

Acknowledgement

This work has been partially subsidized by the TIN 2008-06681-C06-03 project of the Spanish Inter-Ministerial Commission of Science and Technology (MICYT), FEDER funds, by the Excellence Project AGR-01879 (Junta de Andalucía), by the Research Group AGR-170 HIBRO of the "Plan Andaluz de Investigación, Desarrollo e Innovación" (PAIDI), and the P08-TIC-3745 project of the "Junta de Andalucía" (Spain). The research of Francisco Fernández-Navarro has been funded by the "Junta de Andalucia" Predoctoral Program, grant reference 390015-P08-TIC-3745.

References

Alcala-Fdez, J., Sánchez, L., García, S., del Jesús, M.J., Ventura, S., Garrell, J.M., Otero, J., Romero, C., Bacardit, J., Rivas, V.M., Fernández, J.C., Herrera, F., 2008. KEEL: a software tool to assess evolutionary algorithms for data mining problems. Soft Computing 13 (3), 307–318.
Allwein, E.L., Schapire, R.E., Singer, Y., 2001. Reducing multiclass to binary: a unifying approach for margin classifiers. Journal of Machine Learning Research 1 (2), 113–141.
Anand, R., Mehrotra, K., Mohan, C.K., Ranka, S., 1995. Efficient classification for multiclass problems using modular neural networks. IEEE Transactions on Neural Networks 6 (1), 117–124.
Baranyi, J., Ross, T., McMeekin, T.A., Roberts, T.A., 1996. Effects of parameterization on the performance of empirical models used in 'predictive microbiology'. Food Microbiology 13 (1), 83–91.
Basheer, I., Hajmeer, M., 2000. Artificial neural networks: fundamentals, computing, design and application. Journal of Microbiological Methods 43, 3–31.
Bishop, C.M., 1996. Neural Networks for Pattern Recognition. Oxford University Press, Oxford, UK.
Chang, C., Lin, C., 2001. Libsvm: A Library for Support Vector Machines.
Chawla, N.V., Bowyer, K.W., Hall, L.O., Kegelmeyer, W.P., 2002. Smote: synthetic minority over-sampling technique. Journal of Artificial Intelligence Research 16, 321–357.
Chen, H., Tino, P., Yao, X., 2009. Probabilistic classification vector machines. IEEE Transactions on Neural Networks 20 (6), 901–914.
Freeman, J.A.S., Saad, D., 1995. Learning and generalization in radial basis function networks. Neural Computation 7 (5), 1000–1020.
Fukunaga, K., 1999. Introduction to Statistical Pattern Recognition, 2nd Edition. Academic Press.
Garcia-Gimeno, R.M., Hervas-Martinez, C., Rodriguez-Perez, R., Zurera-Cosano, G., 2005. Modelling the growth of *Leuconostoc mesenteroides* by artificial neural networks. International Journal of Food Microbiology 105 (3), 317–332.
Geeraerd, A.H., Herremans, C.H., Cenens, C., Van Impe, J.F., 1998. Application of artificial neural networks as a non-linear modular modeling technique to describe bacterial growth in chilled food products. International Journal of Food Microbiology 44 (1–2), 49–68.
Gelenbe, E., Hussain, K.F., 2002. Learning in the multiple class random neural network. IEEE Transactions on Neural Networks 13 (6), 1257–1267.
Gutiérrez, P.A., Hervás-Martínez, C., Carbonero, M., Fernández, J.C., 2009. Combined projection and kernel basis functions for classification in evolutionary neural networks. Neurocomputing 72 (13–15), 2731–2742.
Gutiérrez, P.A., Hervás-Martínez, C., Lozano, M., 2010. Designing multilayer perceptrons using a guided saw-tooth evolutionary programming algorithm. Soft Computing 14 (6), 599–613.
Gysemans, K.P.M., Bernaerts, K., Geeraerd, A.H., Vermeulen, A., Debevere, J., Devlieghere, F., Van Impe, J.F., 2007a. Evaluation of growth/no growth modelling approaches on a non-abrupt growth/no growth interface of *Listeria monocytogenes*. Food Australia 59 (9), 427–432.
Gysemans, K.P.M., Bernaerts, K., Vermeulen, A., Geeraerd, A.H., Debevere, J., Devlieghere, F., Van Impe, J.F., 2007b. Exploring the performance of logistic regression model types on growth/no growth data of *Listeria monocytogenes*. International Journal of Food Microbiology 114 (3), 316–331.
Hajmeer, M., Basheer, I., 2002. A probabilistic neural network approach for modelling and classification of bacterial growth/no-growth data. Journal of Microbiological Methods 51, 217–226.
Hajmeer, M., Basheer, I., 2003. Comparison of logistic regression and neural network-based classifiers for bacterial growth. Food Microbiology 20 (1), 43–55.
Hajmeer, M., Basheer, I., Cliver, D.O., 2006. Survival curves of *Listeria monocytogenes* in chorizos modeled with artificial neural networks. Food Microbiology 23 (6), 561–570.
Hajmeer, M.N., Basheer, I.A., Najjar, Y.M., 1997. Computational neural networks for predictive microbiology ii. application to microbial growth. International Journal of Food Microbiology 34 (1), 51–66.
Hastie, T., Tibshirani, R., Friedman, J.H., 2001. The Elements of Statistical Learning. Springer. August.
Hervás-Martínez, C., Garcia-Gimeno, R.M., Martinez-Estudillo, A.C., Martinez-Estudillo, F.J., Zurera-Cosano, G., 2006. Improving microbial growth prediction by product unit neural networks. Journal of Food Science 71 (2), 31–38.
Hwang, Y., Bang, S., 1997. An efficient method to construct radial basis function neural network classifier. Neural Networks 10 (8), 1495–1503.
Igel, C., Hüsken, M., 2003. Empirical evaluation of the improved rprop learning algorithms. Neurocomputing 50 (6), 105–123.
Koutsoumanis, K., Taoukis, P., Nychas, G., 2005. Development of a safety monitoring and assurance system for chilled food products. International Journal of Food Microbiology 100 (1–3), 253–260.
Landwehr, N., Hall, M., Frank, E., 2005. Logistic model trees. Machine Learning 59 (1–2), 161–205.
le Cessie, S., van Houwelingen, J., 1992. Ridge estimators in logistic regression. Applied Statistics 41 (1), 191–201.
Le Marc, Y., Pin, C., Baranyi, J., 2005. Methods to determine the growth domain in a multidimensional environmental space. International Journal of Food Microbiology 100 (1–3), 3–12.
Leshno, M., Lin, V., Pinkus, A., Shocken, S., 1993. Multilayer feed-forward networks with a nonpolynomical activation function can approximate any function. Neural Networks 6, 861–867.
Martínez-Estudillo, F.J., Hervás-Martínez, C., Gutiérrez, P.A., Martínez-Estudillo, A.C., 2008. Evolutionary product-unit neural networks classifiers. Neurocomputing 72 (1–2), 548–561.
Miller, R., 1996. Beyond ANOVA, Basics of App. Statistics. Chapman & Hall, London.
Mulgrew, B., 1996. Applying radial basis functions. IEEE Signal Processing Magazine 13 (2), 50–65.
Musavi, M.T., Ahmed, W., Chan, K.H., Farms, K.B., Hummels, D.M., 1992. On the training of radial basis function classifiers. Neural Networks 5, 595–603.
Nabney, I.T., 2004. Efficient training of RBF networks for classification. International Journal of Neural Systems 14 (3), 201–208.
Natarajan, B., 1991. Machine Learning: A Theoretical Approach. Morgan Kaufmann, Los Alamitos, CA.
Orr, M.J.L., 1995. Regularisation in the selection of radial basis function centres. Neural Computation 7 (3), 606–623.
Pitt, R., 1992. A descriptive model of mold growth and aflatoxin formation as affected by environmental conditions. Journal of Food Protection 56, 139–146.
Presser, K., Ross, T., Ratkowsky, D., 1998. Modelling the growth limits (growth/no growth interface) of *Escherichia coli* as a function of temperature, pH, lactic acid

concentration and water activity. Applied and Environmental Microbiology 1773–1779.
Price, D., Knerr, S., Personnaz, L., Dreyfus, G., 1995. Pairwise neural network classifiers with probabilistic outputs. In: Tesauro, G., Touretzky, D., Leen, T. (Eds.), Advances in Neural Information Processing Systems 7 (NIPS-94). MIT Press, pp. 1109–1116.
Ratkowsky, D., Ross, T., 1995. Modelling the bacterial growth/no growth interface. Letters in Applied Microbiology 20, 29–33.
Richard, D., David, E.R., 1989. Product units: a computationally powerful and biologically plausible extension to backpropagation networks. Neural Computation 1 (1), 133–142.
Salter, M., Ratkowsky, D., Ross, T., McMeekin, T., 2001. Modelling the combined temperature and salt (NaCl) limits for growth of a pathogenic *Escherichia coli* strain using nonlinear logistic regression. International Journal of Food Microbiology 61, 159–167.
Seiffert, C., Khoshgoftaar, T.M., Van Hulse, J., Napolitano, A., 2010. Rusboost: a hybrid approach to alleviating class imbalance. IEEE Transactions on Systems, Man, and Cybernetics Part A:Systems and Humans 40 (1), 185–197.
Skandamis, P., Stopforth, J., Kendall, P., Belk, K., Scanga, J., Smith, G., Sofos, J., 2007. Modeling the effect of inoculum size and acid adaptation on no growth/no growth interface of *Escherichia coli* O157:H7. International Journal of Food Microbiology 120, 237–249.
Tamhane, A.C., Dunlop, D.D., 2000. Statistics and Data Analysis. Prentice Hall.
Valero, A., Hervás, C., García-Gimeno, R.M., Zurera, G., 2007a. Product unit neural network models for predicting the growth limits of listeria monocytogenes. Food Microbiology 24 (5), 452–464.
Valero, A., Hervás, C., García-Gimeno, R.M., Zurera, G., 2007b. Searching for new mathematical growth model approaches for listeria monocytogenes. Journal of Food Science 72 (1), M16–M25.
Valero, A., Pérez-Rodríguez, F., Carrasco, E., Fuentes-Alventosa, J.M., García-Gimeno, R.M., Zurera, G., 2009. Modelling the growth boundaries of *Staphylococcus aureus*: Effect of temperature, pH and water activity. International Journal of Food Microbiology 133 (1–2), 186–194.
Vapnik, V.N., 1998. Statistical Learning Theory. Wiley.
Whitley, D.L., Gordon, V.S., Mathias, K.E., 1994. Lamarckian evolution, the baldwin effect and function optimization. In: Davidor, Y., Schwefel, H.P., Männer, R. (Eds.), Parallel Problem Solving from Nature — PPSN III. Springer, Berlin, pp. 6–15.
Witten, I.H., Frank, E., 2005. Data Mining: Practical Machine Learning Tools and Techniques. Data Management Systems, 2nd Edition. Morgan Kaufmann (Elsevier).
Yu, C., Davidson, V.J., Yang, S.X., 2006. A neural network approach to predict survival/death and growth/no-growth interfaces for *Escherichia coli* O157:H7. Food Microbiology 23 (6), 552–560.
Zhou, Z.-H., Liu, X.-Y., 2006. Training cost-sensitive neural networks with methods addressing the class imbalance problem. IEEE Transactions on Knowledge and Data Engineering 18 (1), 63–77.
Zurera-Cosano, G., García-Gimeno, R.M., Rodríguez-Pérez, M.R., Hervás-Martínez, C., 2005. Validating an artificial neural network model of leuconostoc mesenteroides in vacuum packaged sliced cooked meat products for shelf-life estimation. European Food Research and Technology 221 (5), 717–724.