

Received May 17, 2020, accepted May 27, 2020, date of publication June 11, 2020, date of current version June 23, 2020.

Digital Object Identifier 10.1109/ACCESS.2020.3001690

A Novel Ensemble Beta-Scale Invariant Map Algorithm

HÉCTOR QUINTIÁN¹, (Member, IEEE), AND EMILIO CORCHADO², (Member, IEEE)

¹Department of Industrial Engineering, University of A Coruña, 15405 A Coruña, Spain

²Department of Computer Science and Automatic, University of Salamanca, 37008 Salamanca, Spain

Corresponding author: Héctor Quintián (hector.quintian@udc.es)

ABSTRACT This research presents a novel topology preserving map (TPM) called Weighted Voting Supervision -Beta-Scale Invariant Map (WeVoS-Beta-SIM), based on the application of the Weighted Voting Supervision (WeVoS) meta-algorithm to a novel family of learning rules called Beta-Scale Invariant Map (Beta-SIM). The aim of the novel TPM presented is to improve the original models (SIM and Beta-SIM) in terms of stability and topology preservation and at the same time to preserve their original features, especially in the case of radial datasets, where they all are designed to perform their best. These scale invariant TPM have been proved with very satisfactory results in previous researches. This is done by generating accurate topology maps in an effectively and efficiently way. WeVoS meta-algorithm is based on the training of an ensemble of networks and the combination of them to obtain a single one that includes the best features of each one of the networks in the ensemble. WeVoS-Beta-SIM is thoroughly analyzed and successfully demonstrated in this study over 14 diverse real benchmark datasets with diverse number of samples and features, using three different well-known quality measures. In order to present a complete study of its capabilities, results are compared with other topology preserving models such as Self Organizing Maps, Scale Invariant Map, Maximum Likelihood Hebbian Learning-SIM, Visualization Induced SOM, Growing Neural Gas and Beta- Scale Invariant Map. The results obtained confirm that the novel algorithm improves the quality of the single Beta-SIM algorithm in terms of topology preservation and stability without losing performance (where this algorithm has proved to overcome other well-known algorithms). This improvement is more remarkable when complexity of the datasets increases, in terms of number of features and samples and especially in the case of radial datasets improving the Topographic Error.

INDEX TERMS Ensembles, topology preserving mapping, quality measures, SOM, SIM, ViSOM, MLHL-SIM, GNG, Beta-SIM, WeVoS.

I. INTRODUCTION

The extraction of information from enormous datasets that are generated by modern experimental and observational methods is increasingly necessary in almost all industrial and scientific fields and business operations nowadays. This “information extraction” [1] is defined as the nontrivial data mining [2], [3] of implicit, previously unknown, and potentially useful information. Among several fields where “information extraction” is not an easy task, Big data [4]–[8] is one of the most recent and important topics where the use of intelligent techniques becomes crucial to be able to extract knowledge from the enormous amounts of information. One

of the many techniques used to extract relevant information is data visualization [9]–[13].

A recent advance in this field is the Beta-Scale Invariant Map (Beta-SIM) [14], which is based on a modification of a topology-preserving map that can be used for scale invariant classification [11], [14], [15, p. 2], [16], [17], by deriving new learning rules from Beta distribution and applying it to the SIM [11], [17].

Another widely used clustering and classification algorithm is the Growing Neural Gas (GNG) algorithm, proposed by Fritzke [18], [19]. It is based on the Neural Gas (NG) algorithm previously proposed by Martinetz *et al.* [20] for finding optimal data representations based on feature vectors, which is in turn a modification of the widely known SOM. The main characteristic of the NG algorithm is that instead of

The associate editor coordinating the review of this manuscript and approving it for publication was Aysegül Ucar¹.

expanding through the data input space as a fixed grid of units (as done by the SOM algorithm), the NG algorithm allows the neighbouring relationships of its units to change, expanding more like a gas over the data space.

In the other hand, one of the main problems related to Artificial Neural Networks (ANN) is the fact that they can be rather instable obtaining different results in spite of being trained with the same dataset and similar parameters [21]. In this kind of algorithms, samples of datasets are presented in a random order and algorithms are also randomly initialized to avoid to benefit the same samples in different runs of the algorithm. Therefore, the final results after several runs of training process can be rather different, in spite of an statistical analysis of the results could show that algorithms are stable [11]. Ensembles are widely used in the ANN and have proved to be useful for Topology Preserving Maps to increase the stability of these algorithms [22]. Ensembles [16], [23], [24] are based on the concept that a group of experts will get a better results for a problem than a single expert. In this study, it is combined the WeVoS meta-algorithm [25] with other recent topology preserving map (Beta-SIM) [14], to get the novel WeVoS-Beta-SIM. The aim is to improve stability of the original algorithm (Beta-SIM) getting a better topology preserving representation and, therefore, without losing performance of Beta-SIM algorithm which provides a better visualization of the internal structure of high dimensional datasets.

Weighted Voting Supervision (WeVoS) [25] combines the final network maps of an ensemble of Topology Preserving Maps in a single one that includes the best features of each network in the ensemble, trying to solve the problem previously described of instability of neural networks.

Therefore, this novel research presents and thoroughly analyses the use of WeVoS meta-algorithm when it is applied to a new family of learning rules called Beta-SIM, giving a novel algorithm called WeVoS-Beta-SIM. As result the new algorithm is able to obtain the same performance of Beta-SIM algorithm (in terms of Classification Error and Mean Quantization Error) and at the same time improve the topology and stability of the generated grids (in terms of Topographic Error). It is also compared with other well-known Topology Preserving Maps and WeVoS versions such as: the Self Organizing Maps (SOM), WeVoS-SOM, Scale Invariant Map (SIM), Maximum Likelihood Hebbian Learning-SIM (MLHL-SIM), Visualization Induced SOM (ViSOM), Growing Neural Gas (GNG) and Beta-SIM. The present study reports the application of these algorithms to 14 diverse real benchmark datasets from the UCI web repository [26]. These group of 14 datasets are related to industrial, science and economy cases of study, and contains different combinations of number of features, number of samples and number of classes in order to test the novel WeVoS-Beta-SIM algorithm with very diverse datasets and validate its behavior.

This study is organized as follows: In Section 2, presents the main Topology Preserving Maps and detail the well-know SIM algorithm, in Section 3 the novel ensemble

WeVoS-Beta-SIM is described in detail. Section 4 outlines the quality measures, previously proposed in the literature, used to evaluate different properties of topology preserving mapping algorithms. Section 5 analyzes the capabilities of the WeVoS-Beta-SIM algorithm by applying it to perform a detailed study over 14 real benchmark datasets with diverse characteristics. Finally, Section 6 contains the conclusions and outlines future lines of research.

II. TOPOLOGY PRESERVING ALGORITHMS

Among the great variety of tools for multidimensional data visualization, several of the most widely used are those belonging to the family of the Topology Preserving Maps [27]–[29]. Probably the best known among these algorithms is the Self-Organizing Map (SOM) [27], [30]–[32], based on a type of unsupervised learning called competitive learning. Other types are related to the SIM algorithm [11], [33], [34] which differs from the SOM as it is designed to perform their best with radial datasets, due to the fact that both create a mapping where each neuron captures a “pie slice” of the data according to the angular distribution of the input data.

A. DIFFERENT TYPES OF TP ALGORITHM

Several extensions of SOM can be found in the literature such as the Generative Topographic Mapping (GTM) [35]–[37] which was developed as a probabilistic version of the SOM, in order to overcome some of its limitations, particularly the lack of an objective function. An important potential application of the GTM is allowing a simpler visualization of high-dimensional data.

Other extensions of SOM are the Topographic Product of Experts (ToPoE) [38], and the Harmonic Topographic Map (HaToM) [38], [39], where the topology preserving map is created from a product of experts.

The Visualization Induced SOM (ViSOM) [16], [40], [41], is a SOM extension proposed for the direct preservation of the local distance information on the map, along with the topology. The ViSOM constrains the lateral contraction forces between units and hence regularizes the inter-unit distances, so that distances between units in the data space are in proportion to those in the input space. The ViSOM does not only takes into account the distance between a unit's weights from one iteration to the next, but also the distance between that unit and the Best Matching Unit within the whole map (BMU). This allows the ViSOM to preserve topology by maintaining distance between neighbours of the winner unit.

Two other interesting topology preserving models are the Scale Invariant Map (SIM) [9], [11], [17] and the Maximum Likelihood Scale Invariant Map (MLHL-SIM) [17], [42]. Both are designed to perform their best with radial datasets, as each mapping neuron catch up only a radial portion of the data based on its angular distribution. However, when SOM is trained, it approximates a Voronoi tessellation of the input space [27]. The Scale Invariant Map is an implementation of the negative feedback network [43] to form a topology

preserving mapping. The main difference between this mapping and the SOM [27], [31] is that this mapping is scale invariant.

B. SCALE INVARIANT MAP

The main target of the family of Topology Preserving Maps [27] is to produce low dimensional representations of high dimensional datasets maintaining the topological features of the input space.

The SIM [9], [11], [34] is an algorithm similar to SOM [27], but training uses a method based on negative feedback network [43], [44]. SIM uses a neighbourhood function and competitive learning in the same way as a SOM. SIM model is defined by (1), (2) and (3).

$$\text{Feedforward: } y_i = \sum_{j=1}^N W_{ij}x_j, \quad (1)$$

$$\text{Feedback: } e = x - W_c y_c \quad (y_c = 1), \quad (2)$$

$$\text{Weightsupdate: } \Delta W_i = h_{ci}\eta (x - W_c), \quad \forall i \in N_c, \quad (3)$$

where, x is a N-dimensional input vector, and y an M-dimensional output vector, with W_{ij} being the weight linking input j to output i ; e is the residual or error, η the learning rate, W_c is the weight connected to the output winner and h_{ci} represents the neighborhood function, which is a Gaussian function in this case.

The input data x_j is feed forward through weights W_{ij} to the output neurons y_i , where a linear summation is performed to obtain the activation of the output neurons (1).

Based on the previous obtained neurons activation, a winner neuron is selected using the minimum Euclidean distance (the neuron whose weight vector is closest to the input neuron wins) or using the maximum activation (the neuron with the highest activation wins).

After selection of an output winner, the winner, c , is deemed to be firing ($y_c = 1$) and all other outputs are suppressed ($y_i = 0, \forall i \neq c$).

The winner's activation is then feedback through its weights and this is subtracted from the inputs, and simple Hebbian learning is used to update the weights of all nodes in the neighborhood of the winner.

III. ENSEMBLES

Ensembles are meta-algorithm used to improve the performance of algorithms mainly used in supervised learning [45]–[47]. An ensemble can be view as a group of experts working together to solve a problem. Therefore, a dataset is divided into several parts and one model is generated by using one of these parts. Finally, the ensemble meta-algorithm combines resulting models to predict a result.

The main strength of ensembles is that are able to get a good balance between small variance and small bias. The main reason is that different classifier designs potentially offer complementary information on the patterns to be classified and could be harnessed to improve the performance of the selected classifier.

There are two main approaches [11], [16], [45] to how each of the classifiers or components of the ensemble are going to be trained:

- Independent Training: Each model is trained without knowledge about other models, being Bagging the most widely used technique [47], [55].
- Coordinated trained: The training of one model takes into account how other models were trained. This concept is applied by Boosting techniques [47], [55].

Currently, the use of ensembles for not supervised learning is not strongly developed, being the main reference, the ensembles used in topology-preserving maps. In all cases a fusion process of the generated maps is performed based on some metrics such as Fusion by Euclidean Distance, Similarity of Voronoi Polygons, etc., however they do not take into account the topology neighborhood. In order to avoid such problem, in the recent years several fusion methods were developed for topology-preserving map fusion [11], [48]–[51].

One of the key points is that the weights of each network in the ensemble are initialized to be able to neurons in the same position of two networks are comparable. Therefore, maps should be as similar as possible. Thus, all maps are trained using the same parameters and each ensemble model is initialized with the same values as the previous model finished its training [11], [16].

IV. WeVoS-Beta-SIM

This section presents a novel ensemble algorithm based on Beta-SIM TPM family, devoted to improve the stability of the members of this family, by means of a weighted voting system to fusion of neurons of the different Beta-SIM maps.

Beta-SIM [14] is a novel version of SIM [17], based on the application of a family of learning rules called Beta Hebbian Learning (BHL) [52], when they are applied to SIM.

The main difference between Beta-SIM and SIM is that in Beta-SIM the BHL is used to update the weights of all nodes in the neighbourhood of the winner. Then, Beta-SIM can be defined by means of (4), (5) and (6). Where (6) is obtained by applying BHL method to SIM to update the weights.

$$\text{Feedforward: } y_i = \sum_{j=1}^N W_{ij}x_j \quad (4)$$

$$\text{Feedback: } e = x - W_c y_c \quad (y_c = 1) \quad (5)$$

Weights update:

$$\begin{aligned} \Delta W_i = & \eta \cdot h_{ci} \cdot \text{sign}(x - W_c) \\ & \cdot \left[|x - W_c|^{\alpha-2} (1 - |x - W_c|)^{\beta-2} (1 - \alpha \right. \\ & \left. + |x - W_c| (\alpha + \beta - 2)) \right] \end{aligned} \quad (6)$$

where, x is an N-dimensional input vector, and y an M-dimensional output vector, with W_{ij} being the weight linking input j to output i ; e is the residual or error, η the learning rate, W_c refers to the weights of the winning neuron, h_{ci} represents the neighborhood function, which is a

Gaussian function in this case, and α and β are the parameters that determine the shape of the PDF curve

Therefore, by maximizing the likelihood of the residual with respect to the actual distribution, it is matched the learning rule to the PDF of the residual (e).

The Beta-SIM algorithm is only stable when the absolute value of the residuals ($|e|$) is lower than 1 [14], as when values of the residuals are beyond this limit, the value of the weights update tends towards infinity. To avoid the possibility that the residuals have values higher than 1, the datasets should be normalized in order to satisfy this limitation and preserve the internal topology between dataset dimensions.

The capability of Beta-SIM algorithm to adapt to sparse clusters or to neglect them, based on combinations of parameters α and β [14], provides to the units of Beta-SIM network more freedom than other Topology Preserving Maps to adapt to datasets, however, it also potentially adds instability to the training. So, the use of ensembles and specifically WeVoS fusion algorithm [15], seems to be the most appropriated method to correct this effect.

WeVoS [25] is a weight voting ensemble system to generate fusion single TPMs. Fusion of such TPMs, generate a final TPM that reduce the complexity and increase the accuracy with respect to single maps without using ensemble techniques. It has been previously successfully applied to other well-known Topology Preserving Maps [11], [16], [25]. Beta-SIM has proved to overcome such topology preserving algorithms in the main aspects aimed for Topology Preserving Maps such Mean Quantization Error (MQE), Classification Error (CE) and Topographic Error (TE) [14].

Then, in this research it is presented a novel algorithm based on the application of WeVoS ensemble to the Beta-SIM, presenting a detailed study and comparison with other well-known Topology Preserving Maps and their ensemble versions, such as WeVoS-SOM, to analyse its impact on aspects such as the stability and topology preservation conditions.

WeVoS-Beta-SIM obtains a final map as combination of different Beta-SIM maps by fusion of the neurons in the same position based on a weighted voting. (7) is applied for this voting process:

$$V_{p,m} = \frac{\sum b_{p,m}}{\sum_{i=1}^M b_{p,i}} \cdot \frac{q_{p,m}}{\sum_{i=1}^M q_{p,i}} \quad (7)$$

where, $V_{p,m}$ is the weight of the vote for the unit included in map m of the ensemble, in its position p , M is the total number of Beta-SIM maps, $b_{p,m}$ is the binary vector used for marking the dataset entries recognized by unit in position p of map m , and $q_{p,m}$ is the value of the desired quality measure for unit in position p of map m .

b is a binary vector of the same length as data samples are in the dataset. It is used to store the samples recognized by a single unit.

Fusion of neurons of the different Beta-SIM maps for WeVoS-Beta-SIM, during the training process, is done based on a quality measure [11] calculated for each Beta-SIM map.

This quality measure is considered during the fusion process where the weights of each neuron is proportional to the value of such quality measure, as to modify the position of the neuron in the fused map, the weights of each of the neurons in that position are fed to the final map (nodes N1, N2, N3 in Fig. 1).

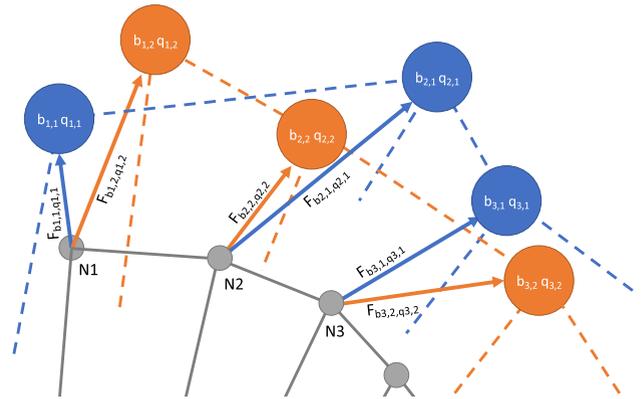


FIGURE 1. Schematic diagram of the weight voting in WeVoS in a 2D map.

In the ensembles, neurons in the same position of different networks are fused, so their weights should be similar in order to be comparable. Therefore, maps should be as similar as possible. Thus, all maps are trained using the same parameters and each ensemble model is initialized with the same values as the previous model finished its training [11], [16].

Briefly, WeVoS-Beta-SIM meta-algorithm works in the following way:

- First of all, an ensemble of Beta-SIM maps is trained.
- Then, the chosen quality/error measure is calculated for each neuron in all Beta-SIM maps.
- The fused map is initialized by calculating the centroids of the neurons in the same position of all the maps, by calculating the superposition of the ensemble.
- For each of the neurons in the fused map, the average neuron quality and the number of total samples recognized in that position for the Beta-SIM maps, are calculated.
- The weight of the vote for each neuron can be calculated with this information by using (7).
- To modify the position of the neuron in the fused map, the weights of each of the neurons in that position are fed to the final map.
- Finally, the learning rate in each case will be the weight of the vote for that neuron.

V. QUALITY AND TOPOLOGY MEASURES

Despite there are several quality measures that are used to measure the capabilities of the Topology Preserving Maps, there is not a global one [11] being some of them complementary by each one and assessing different features of the final maps in different visual representation areas. Among all these quality measures the three ones selected in this research, due to their complementarities, are the following:

Classification Error (CE): Using its inherent pattern matching characteristics, the Topology Preserving Maps in general terms can be used for classification tasks. Intuitively, the samples activating the same neuron of the network are very likely to belong to the same class. When a new sample is presented to the network, the sample can be classified in the same class as the majority of samples activating the same neuron belong to. A consistent behavior when classifying samples points to a correctly trained map. Although this is not the main function of this kind of networks, the measure of how many samples are wrongly classified has been used, to an extent, to assess the quality of the final map in numerous previous studies [11].

Mean Quantization Error (MQE): MQE is related to all forms of vector quantization and clustering algorithms. Thus, this measure completely disregards map topology and alignment. MQE is computed by determining the average distance of the dataset entries to the cluster centroids by which they are represented. In case of SOM, the cluster centroids are the characteristic vectors.

Topographic Error (TE): TE is the simplest of the topology preservation measures. A dataset is also needed to calculate this measure. For all data samples, the respective best and second-BMUs (1st BMU and 2nd BMU) are determined. If these BMUs are not adjacent on the map lattice, it is considered an error. Finally, the total error is normalized to a range from 0 to 1, where 0 means perfect topology preservation.

VI. EXPERIMENTS AND RESULTS

Several experiments have been designed and performed to investigate the capabilities of WeVoS-Beta-SIM and also to compare it with other well-known Topology Preserving Maps such as of SIM, Beta-SIM, SOM, WeVoS-SOM, ViSOM, MLHL-SIM and GNG.

The first type of experiments was designed to present visually some of the main characteristics of the used algorithms such as topology of the grid maps and spread of the grid maps over the data. In the second type, a thoroughly analysis, in terms of the three quality measures used in this research (CE, MQE and TE), was performed to validate the visual results obtained previously.

All the tests were run using a classic ten-fold cross-validation to use the complete dataset for training and testing. The ensembles were trained using one of the simplest meta-algorithms for ensemble training: the bagging meta-algorithm [53].

In the case of WeVoS-Beta-SIM and WeVoS-SOM, the datasets have been reduced to 1/5 of its original size, and a single model and an ensemble of 5 maps are calculated for each one, comparing the performance of the models over datasets with the same inner structure.

A. BENCHMARK DATASETS DESCRIPTION

A total of 14 diverse interesting benchmark datasets related to industry, economy and science cases of study were used to validate the performance of the WeVoS-Beta-SIM algorithm. Datasets were taken from the UCI Machine Repository [26] presenting different characteristics, such as number of samples, features and classes, from low to high dimensional datasets.

In Table 1, a summary of these diverse datasets is presented in terms of sample size, features and number of classes. Different map sizes, and combination of algorithms parameters where tested, presenting the best combination.

Different map sizes have been tested (20×20 , 20×25 and 30×30), however the final maps are similar and improvements in CE; MQE and TE are not significant, therefore the best combination of parameters for each algorithm for each experiment are presented in Table 2 (APPENDIX).

TABLE 1. Benchmark UCI datasets description in term of number of samples, features and classes.

	Name of the dataset	Samples	Features	Classes	Source
Dataset 1 (D1)	Liver Disorders	345	6	2	UCI repository
Dataset 2 (D2)	Fertility	100	10	2	UCI repository
Dataset 3 (D3)	Bank Marketing	45211	16	2	UCI repository
Dataset 4 (D4)	Iris	150	4	3	UCI repository
Dataset 5 (D5)	Wine	178	13	3	UCI repository
Dataset 6 (D6)	Contraceptive	1473	9	3	UCI repository
Dataset 7 (D7)	Car	1728	6	4	UCI repository
Dataset 8 (D8)	Dermatology	358	33	6	UCI repository
Dataset 9 (D9)	Image Segmentation	2310	19	7	UCI repository
Dataset 10 (D10)	Landsat Satellite	6435	36	7	UCI repository
Dataset 11 (D11)	Yeast	1484	8	10	UCI repository
Dataset 12 (D12)	Pen-Based Handwritten Digits	10992	16	10	UCI repository
Dataset 13 (D13)	Optical Recog. of Handwritten Digits	5620	64	10	UCI repository
Dataset 14 (D14)	Letter Recognition	20000	16	29	UCI repository

WeVoS-Beta-SIM and WeVoS-SOM use the same parameters as their respective original models, Beta-SIM and SOM.

B. VISUALIZATION RESULTS

In this subsection, the visualization results obtained for 2 of the previously described benchmark datasets are presented (“Iris” dataset and “Landsat Satellite” dataset). The aim is to present visually some of the main characteristics of the used algorithms such as topology of the grid maps and spread of the grid maps over the data. Then, a complete analysis of the results, in terms of three quality measures (see Section 4), is performed over all benchmark datasets.

The 2 datasets used in this subsection were selected due to their different levels of complexity. The first dataset is the well-known “Iris” dataset (with low complexity; only 3 classes and 345 samples, 6 features) and the second one is the “Landsat Satellite” dataset (with high complexity; 7 classes and 6435 samples, 36 features). The graphs presented in Fig. 2, 3, 4 and 5 illustrate the performance of each model for each dataset, presenting the analytical results in the next subsection (Section “5.3 Analytical results”).

Fig. 2 and 4 represent the adaptation of each map to its structure in representation of the dataset under analysis (“Iris” and “Landsat Satellite” datasets). It depicts the lattices composing the maps embedded in a 2D input space. Fig. 2 and Fig. 4 present the datasets projected onto their first three principal components and the final grid maps of the models are also embedded in the space of the three principal components. This approach has been previously satisfactory applied [11], [15], [40], [41], [54] in order to visually support the analytical results presented in Tables 3-8.

Fig. 3 and 5 show the final unit map for each algorithm where only BMUs are displayed. Each BMU of each map is labelled based on the training inputs to which they are reacting. This means that if one neuron (BMU) is activated by 20 inputs samples and 19 of them belongs to class 1, this neuron is labelled as class 1 (red circles in Fig. 4 and 5) [11]. In spite of in boundaries the BMU could easily belong to one or other class, it only happens in the limits of class boundaries and only a few neurons could be misclassified.

GNG is not suitable for this 2D map representation, as some units are disregarded from the final model and therefore the topology preservation is lost.

1) VISUALIZATION RESULTS FOR IRIS DATASET

It is easily observed in Fig. 2 that Beta-SIM and WeVoS-Beta-SIM grid maps are more widely spread throughout the Iris dataset (represented as magenta dots), covering the input space better than the other algorithms. This better coverage over the dataset corresponds to a better MQE result (see Table 7 APPENDIX).

However, SOM, WeVoS-SOM and ViSOM conserve the topology of the map very well as their grid maps contains just few twists and folds (Fig. 2d, 2e and 2g). Therefore, they have the lowest TE values among all algorithms (see Table 8 APPENDIX).

Comparing Beta-SIM and WeVoS-Beta-SIM, Fig. 2b and 2c show how WeVoS-Beta-SIM obtains a better topology of the grid map versus Beta-SIM (which means better TE results), due to the fact that WeVoS-Beta-SIM grid map contains less twists and folds than Beta-SIM. At the same time, the WeVoS-Beta-SIM grid map adapts slightly better their structure to the dataset, covering the Iris dataset more adequately (which means a lower MQE value).

Fig. 3 shows that, in general, the WeVoS-Beta-SIM algorithm (Fig. 3c) provides the map with more compact and clearly separated groups. However, differences with the other algorithms are minor. All algorithms obtain maps where class 1 (red circles Fig. 3) is clearly separated from the other 2 classes (class 2 –blue squares– and class 3 –green triangles–). Differences between maps are only appreciable when they are compared in terms of separation of classes 2 and 3, where WeVoS-Beta-SIM algorithm presents these 2 classes in more compact and clearly defined groups (Fig. 3c).

2) VISUALIZATION RESULTS FOR LANDSAT SATELLITE DATASET

Fig. 4 shows how Beta-SIM (Fig. 4b), GNG (Fig. 4h) and WeVoS-Beta-SIM (Fig. 4c) algorithms outperform the other algorithms by distributing the units of their grid maps over the Landsat Satellite dataset (represented as red dots) in the best possible way. Units of the grid maps are close to the input samples over the whole dataset, which leads to lower MQE values than the other algorithms.

A similar situation occurs with SIM and MLHL-SIM algorithms (Fig. 4a and 4f) as they obtained grid maps which adapt well their structure to the dataset but not as well as the previously mentioned algorithms.

Again, SOM, WeVoS-SOM and ViSOM algorithms are the ones which better preserve the topology of the grid maps (Fig. 4d, 4e and 4g) as their maps have less twist and folds than the other algorithms. Therefore, they obtain the lowest TE values (see Table 8 APPENDIX).

WeVoS-Beta-SIM (Fig. 4c) and Beta-SIM (Fig. 4b) algorithms obtain similar final grid maps, but WeVoS-Beta-SIM preserves the topology of the grid map better than the Beta-SIM algorithm. Comparing both figures (Fig. 4b and 4c), the WeVoS-Beta-SIM grid map presents less twists and folds than the Beta-SIM grid map, having then lower TE values.

It can be seen in Fig. 5 that WeVoS-Beta-SIM (Fig. 5c) provides the best visual representation through a smoother map compared to the rest of the algorithms. The map presents, in general, compact and unmixed groups (there is no mixing of BMUs from different classes). Beta-SIM also obtains compact groups (Fig. 5b), but some of the neurons associated to different classes are mixed. This fact was observed in Fig. 4 where the Beta-SIM grid map (Fig. 4b) contained more twists and folds than WeVoS-Beta-SIM (Fig. 4c).

In the case of SOM and WeVoS-SOM (Fig. 5d and 5e), both algorithms produced maps where groups are clearly defined,

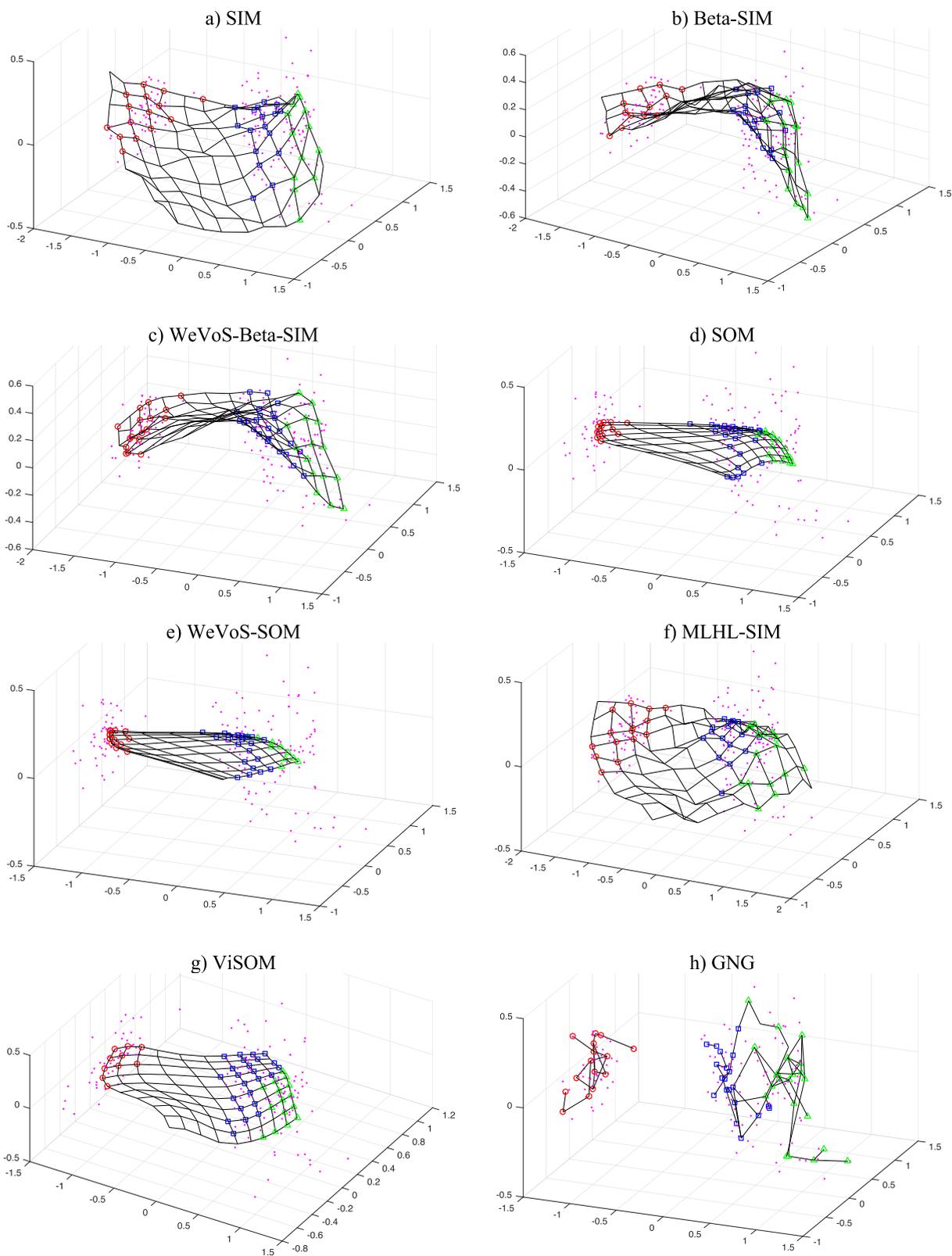


FIGURE 2. Algorithms' final network grids adaptation over the Iris dataset.

however some groups present mixed classes. For instance, in the top of both maps (Fig. 5d and 5e), class 2 (blue squares) is divided into two groups separated by a group of class 1

(red circles). In the case of classes 2 and 6 (blues squares and magenta asterisks respectively), in both maps several neurons appear lost in the middle of the map.

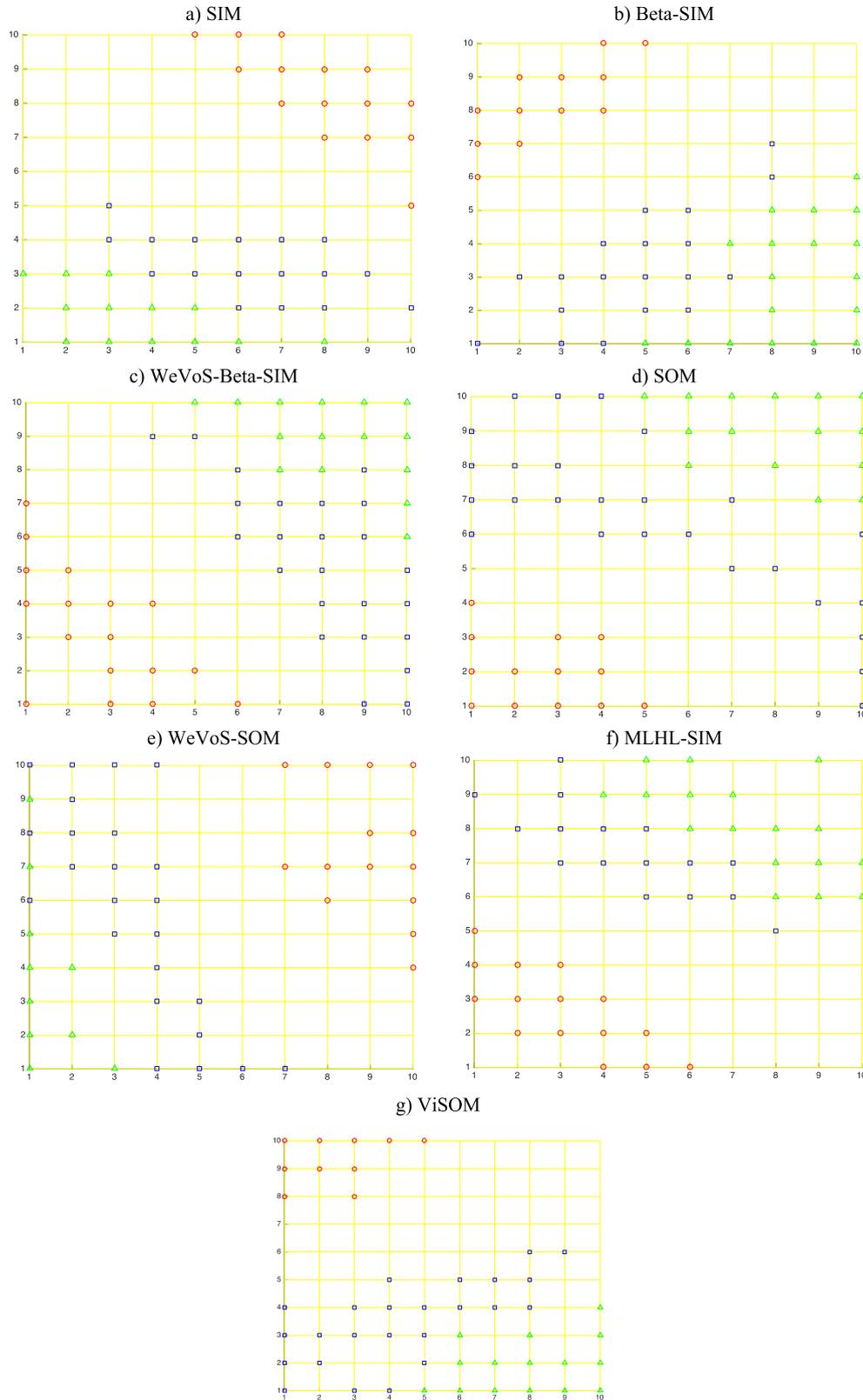


FIGURE 3. Algorithms' final unit maps for Iris dataset.

3) CONCLUSIONS OF THE VISUALIZATION RESULTS

The results suggest that WeVoS-Beta-SIM provides a better visual representation of the datasets than the other algorithms,

as it is able to widely spread its grid map covering the input space better than the other tested models. At the same time, WeVoS-Beta-SIM obtains grid maps with less twists and

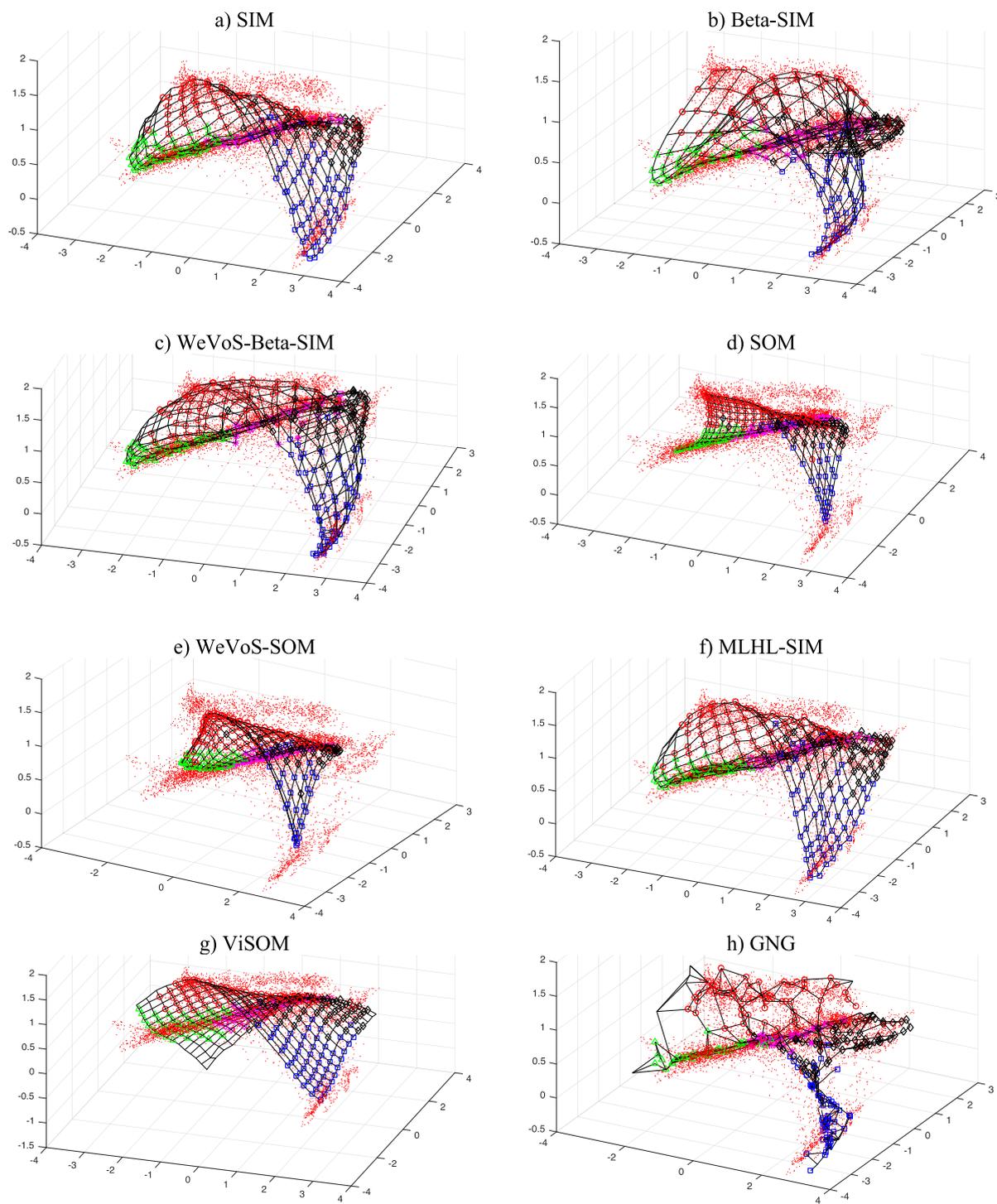


FIGURE 4. Algorithms' final network grid adaptations over the Landsat Satellite dataset.

folds than the Beta-SIM algorithm, which signifies a better topology of the map.

The improvement on visual representation, achieved by WeVoS-Beta-SIM, is notably higher when the complexity of the datasets increases. Using the previous examples, differences between maps were minor for the Iris dataset (low complexity), whereas differences were higher for the Landsat Satellite dataset (high complexity).

C. ANALYTICAL RESULTS

In order to validate the results obtained, statistical tests for the three quality measures were performed consisting of an ANOVA + post-hoc analyses. The statistical results for CE, MQE and TE are presented in Tables 3 to 5 (p-values) and Tables 6 to 8 (average testing values ± standard deviation).

All measures presented are error measures for the testing dataset, so the desired value is always as close to 0

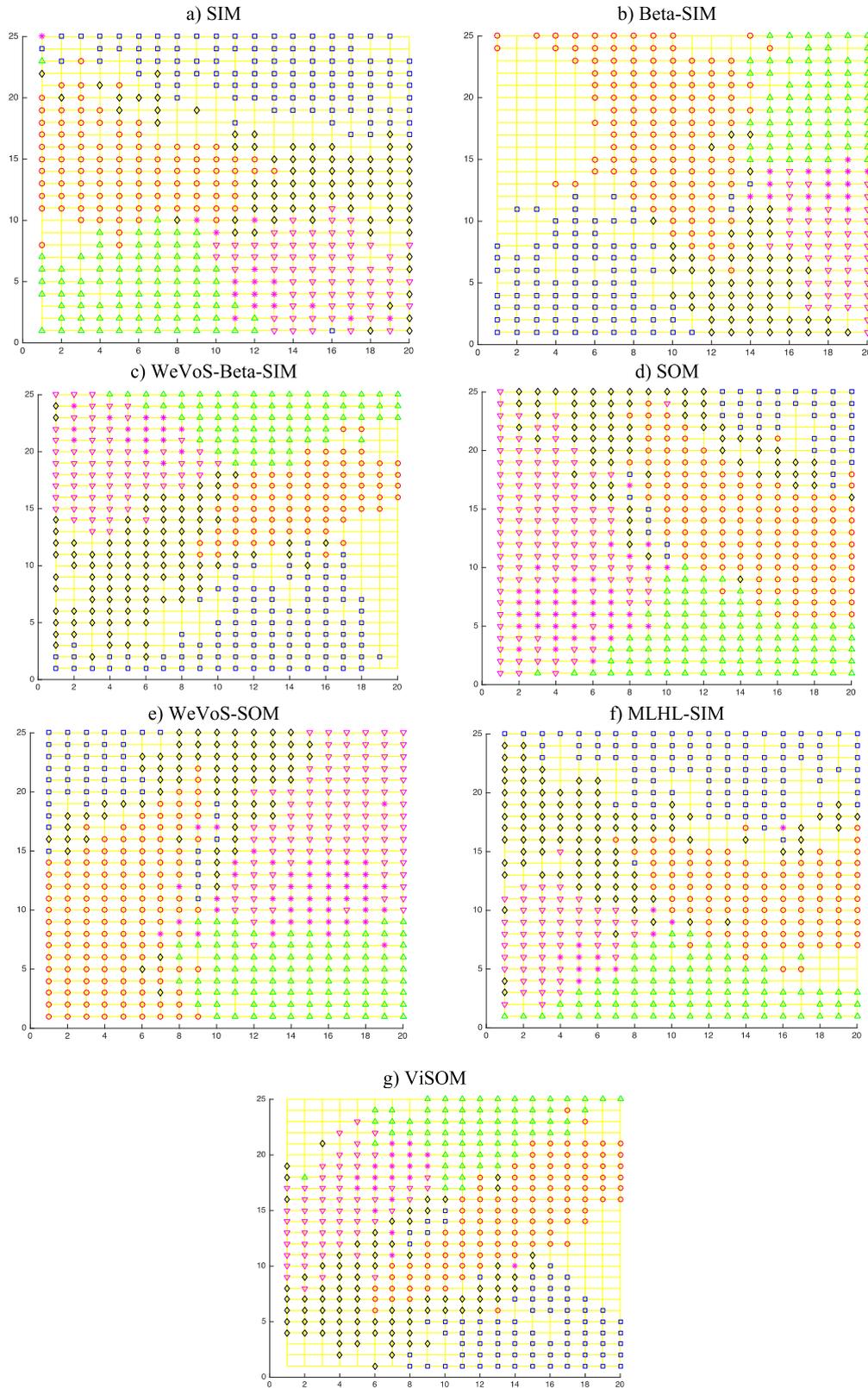


FIGURE 5. Algorithms final network grid adaptations over the Landsat Satellite dataset.

as possible. The CE is presented in percentage form and normalized between 0 and 1, whereas the rest of the measures are expressed as absolute values.

These series of experiments analyze 2 different aspects of the novel WeVoS-Beta-SIM and the other tested algorithms:

- The performance of WeVoS-Beta-SIM in comparison with the other 7 topology preserving models, in terms of CE, MQE and TE quality measures.
- The effect of modifying the number of data samples used during the training process for all algorithms under study. This was done in order to emulate the addition of noise or instability in the datasets [16].

1) ANALYSIS OF RESULTS IN TERMS OF CE

Results presented in Fig. 6 show that GNG, Beta-SIM and WeVoS-Beta-SIM often obtain better results than the other algorithms (SOM, ViSOM, WeVoS-SOM, SIM, MLHL-SIM), in terms of CE values. It should be noted that these differences in CE can only be seen when the complexity of the datasets is high. For low complexity datasets (in this novel research those having less than 5 classes; datasets from D1 to D8), CE results obtained by the different algorithms were not statistically significant (see results of Table 3 APPENDIX).

Results for this experiment confirm the conclusions obtained by the visual representation test (figures from Fig. 2 to 5). For example, in the case of the Iris dataset (a dataset with low complexity), all algorithms presented very similar final maps (Fig. 3), so similar CE values for all algorithms were expected. In the case of the Landsat Satellite dataset (high complexity), the WeVoS-Beta-SIM algorithm obtained a map (Fig. 5c) with more compact groups and less mixed classes, therefore obtaining a better CE value than the other algorithms.

It can also be seen in Fig. 6 that when the complexity of the datasets increases, the WeVoS-Beta-SIM obtains better CE results than the Beta-SIM algorithm. However, when a statistical test of the results is performed (Table 3 APPENDIX) differences between CE values were not statistically significant.

Finally, the change in CE values when the number of samples is increased was analyzed (Fig. 6). The effect of adding such instability in this case is not particularly evident, as the change in CE values did not follow a clear tendency.

2) ANALYSIS OF RESULTS IN TERMS OF MQE

Results for MQE values are presented in Fig. 7, where it is clear that SOM and WeVoS-SOM algorithms obtain the highest MQE values for all datasets. This is expected based on results obtained in the visual representation test (Fig. 2 and 4), where WeVoS-SOM and SOM grid maps do not spread over the datasets as well as the other algorithms.

The rest of the algorithms behave in a similar way, in terms of MQE, where none of which outperform the others when complexity of the dataset is low. However, when complexity of the datasets increases, 2 algorithms stand out over the others in terms of MQE results: WeVoS-Beta-SIM and Beta-SIM algorithms. Again, these results confirm the conclusions obtained by the visual representation tests (Section 5.2), where when the complexity of the datasets is high, the maps of these algorithms cover the input space better than the other algorithms.

Fig. 7 shows that the MQE results obtained by WeVoS-Beta-SIM are often better than the simple model Beta-SIM, even when the complexity of the dataset is low. However, these differences are not always statistically significant (see Table 4 APPENDIX).

Finally, the change in MQE values when the number of samples is increased was analyzed. In this case, adding such instability does not present a particular effect on the algorithms, with the MQE being more dependent on the total size of each dataset.

3) ANALYSIS OF RESULTS IN TERMS OF TE

Finally, the TE results are analyzed and presented in Fig. 8. The TE is related with the topology of the final maps, as was shown in the visual representation section (Section 5.2). Based on those results, it was expected that WeVoS-SOM, SOM and ViSOM would obtain the lowest TE, as these algorithms produced maps with less twists and folds. Results of Fig. 8 confirm it, but differences can only be seen when the complexity of the dataset is high.

When complexity of the datasets increases and differences between algorithms are clear (datasets from D8 to D14), in terms of TE values, GNG and Beta-SIM algorithm obtain the worst TE values. In the case of GNG, this is due to the fact that some units of the final map are disregarded from the final model and therefore the topology preservation is lost. In the case of the Beta-SIM algorithm, it focuses on distributing the units of the grid map over the dataset, so the final map contains several twists and folds (Fig. 2b and 4b).

WeVoS-Beta-SIM (Fig. 8) consistently improves the TE results of the simple model Beta-SIM, especially when the complexity of the datasets increases. This means that the final map provides a better visualization regarding the topology preservation of the map (better TE) whereas keeping similar MQE values. The statistical tests presented in Table 5 (APPENDIX) validate these results when complexity of the datasets is high (as p-values are lower than the significance level of 0.05).

The effect of adding instability by decreasing the number of samples for the training process was analyzed for WeVoS-Beta-SIM and its simple version Beta-SIM. When the complexity of the datasets is high, WeVoS-Beta-SIM presents very similar TE values as the number of samples is increases. However, Beta-SIM algorithm (Fig. 8) obtains unstable TE values (increasing and decreasing without a clear tendency). Therefore, it can be concluded that WeVoS-Beta-SIM, in terms of TE, is less sensitive to noise than the Beta-SIM algorithm. Therefore, WeVoS-Beta-SIM obtains maps with less distortion which in turn provides a better visual representation of the internal dataset structure.

4) CONCLUSION OF THE ANALYTICAL RESULTS

Analytical results of this subsection confirm the conclusions obtained by the visual representation tests (Section 5.2). When complexity of the datasets is high, WeVoS-Beta-SIM algorithm often obtains the best MQE results,

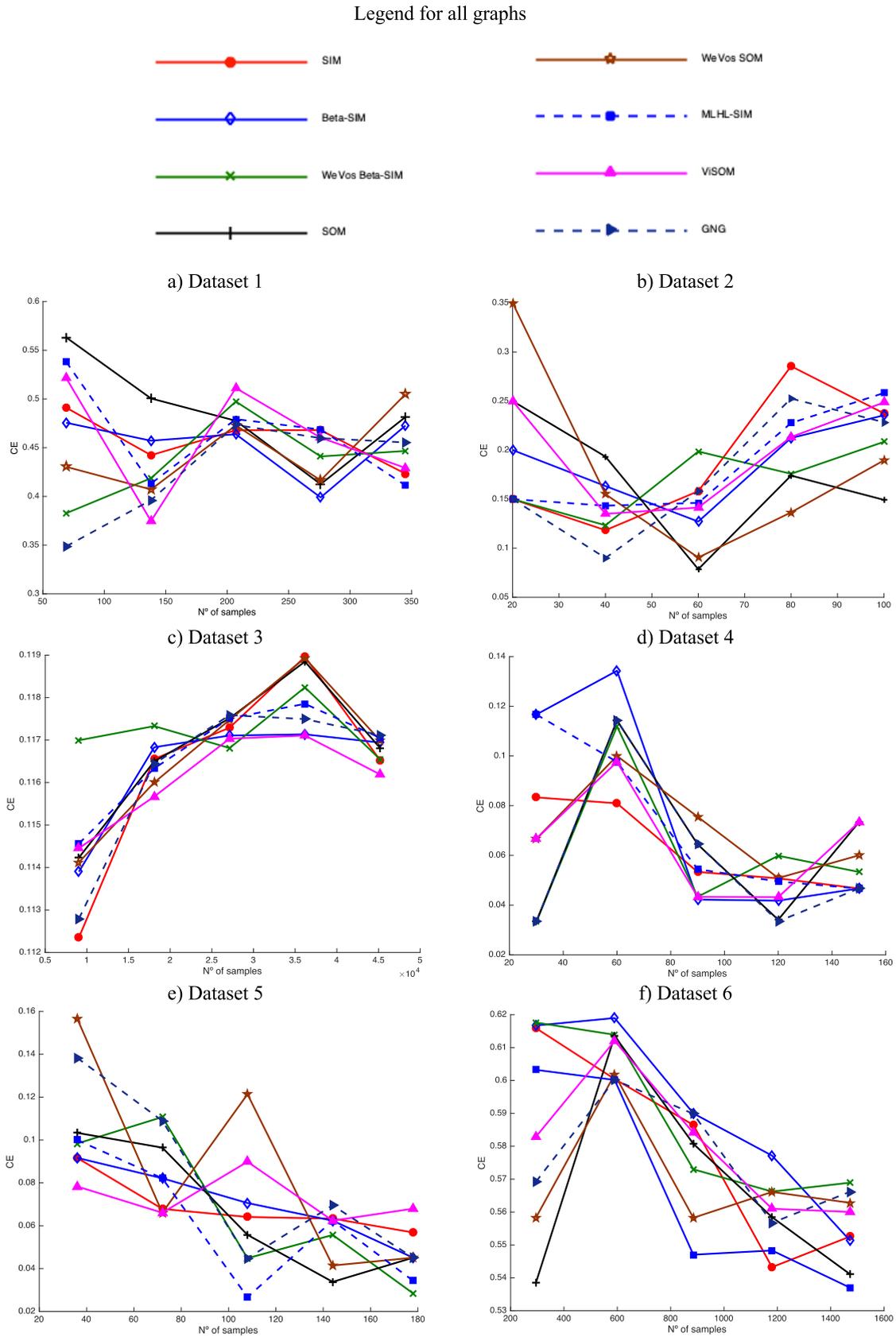


FIGURE 6. CE measures vs number of samples of each algorithm for each of the 14 datasets.

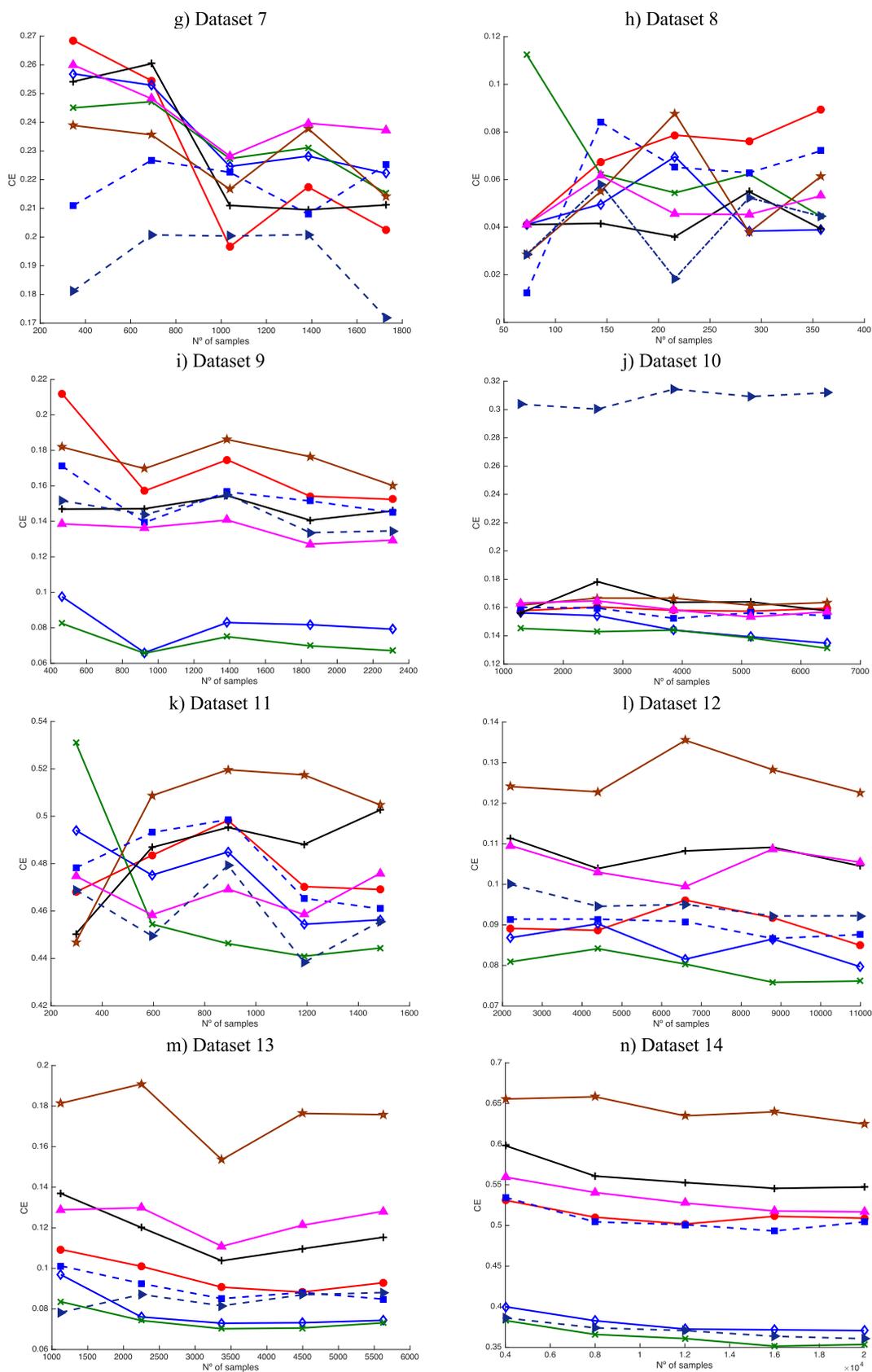


FIGURE 6. (Continued.) CE measures vs number of samples of each algorithm for each of the 14 datasets.

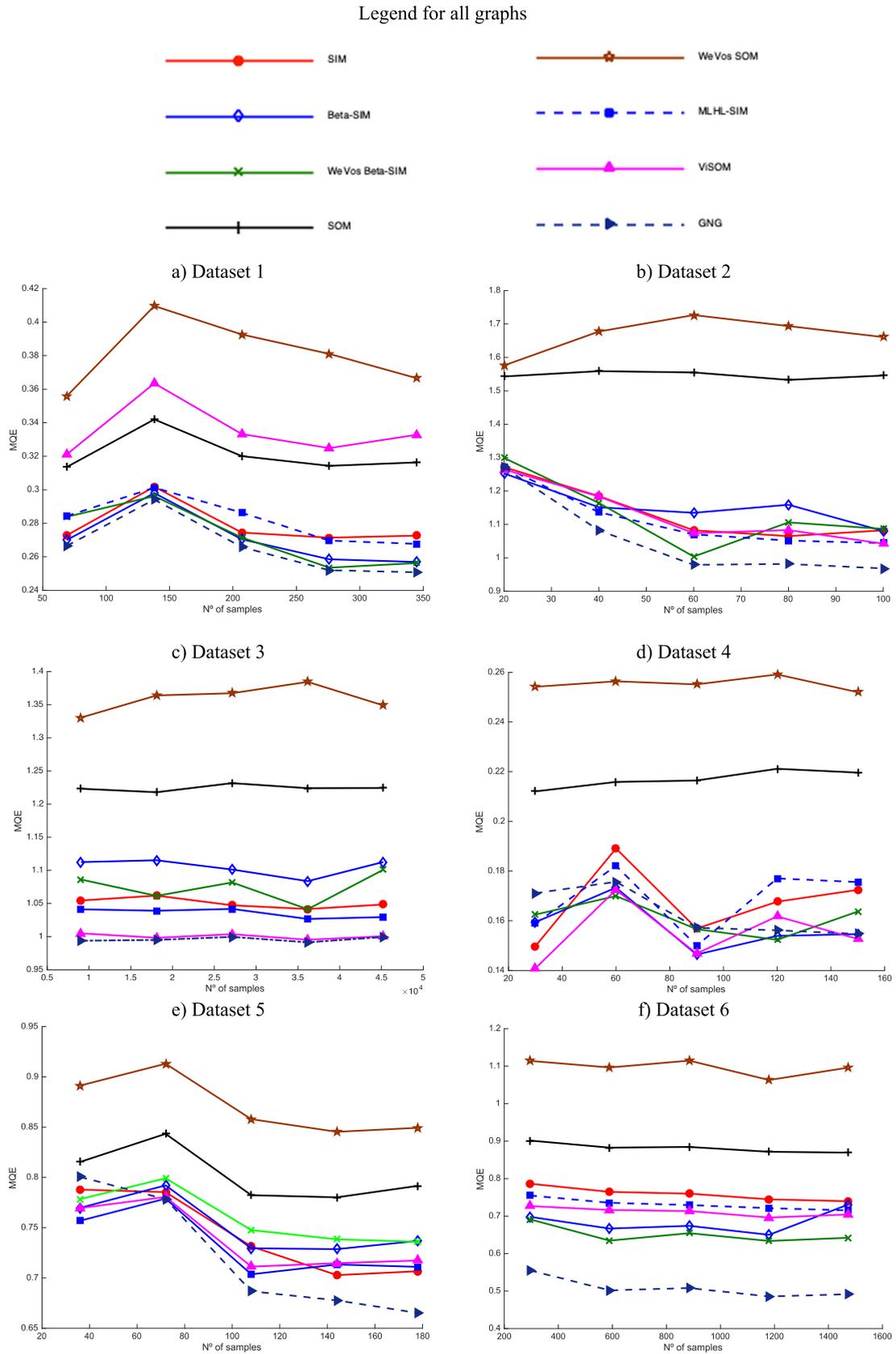


FIGURE 7. MQE measures vs number of samples of each algorithm for each of the 14 datasets.

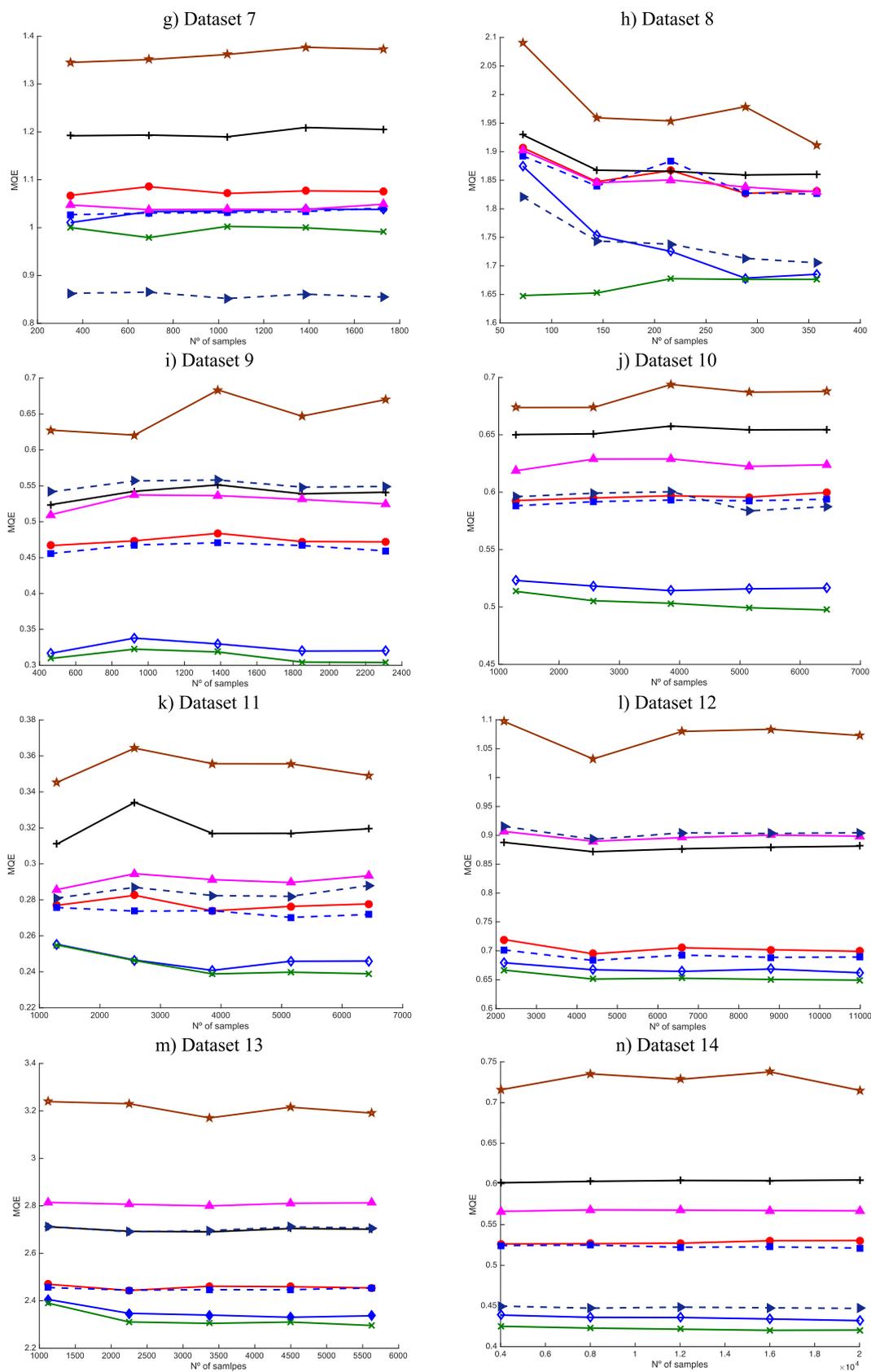


FIGURE 7. (Continued.) MQE measures vs number of samples of each algorithm for each of the 14 datasets.

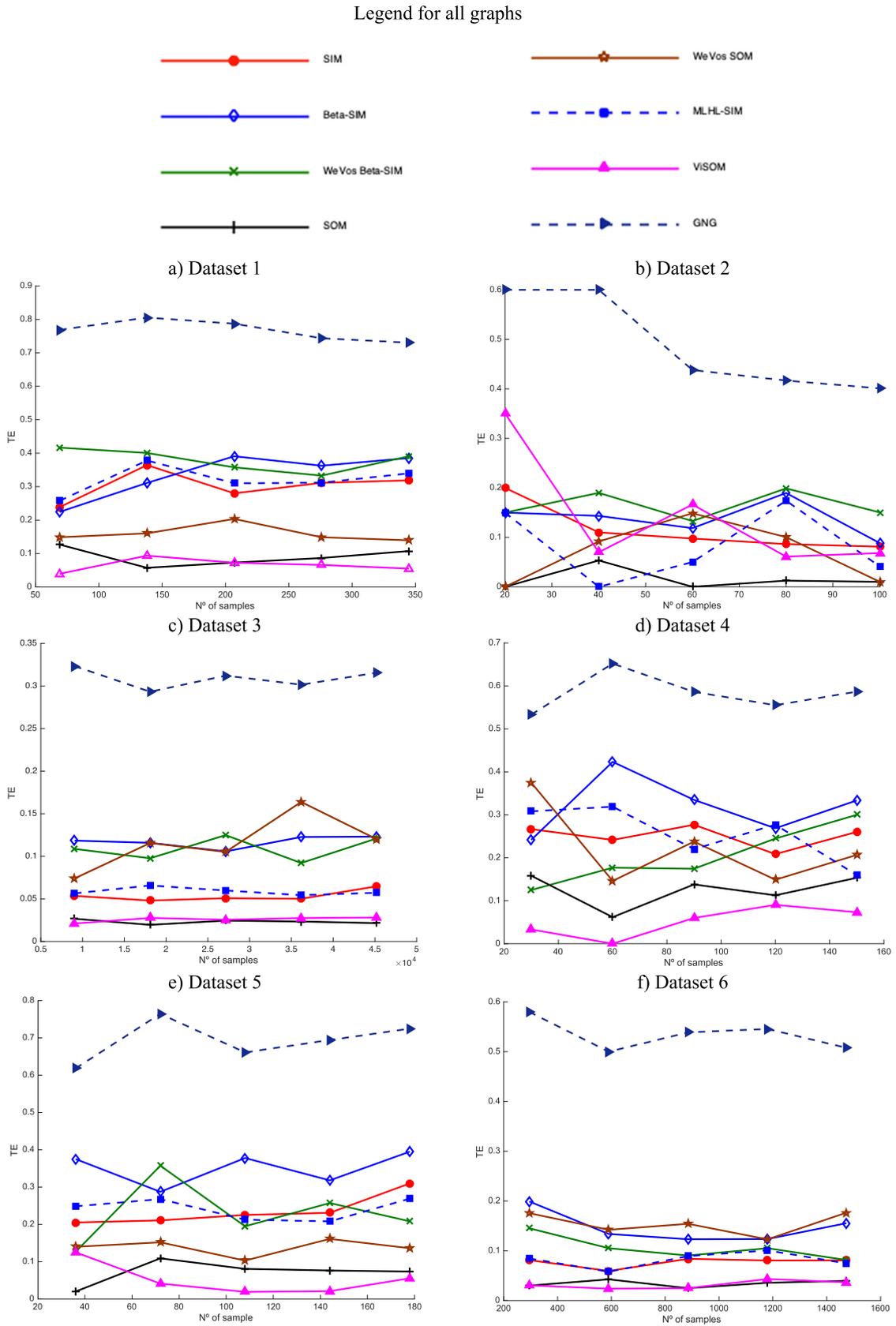


FIGURE 8. TE measures vs number of samples of each algorithm for each of the 14 datasets.

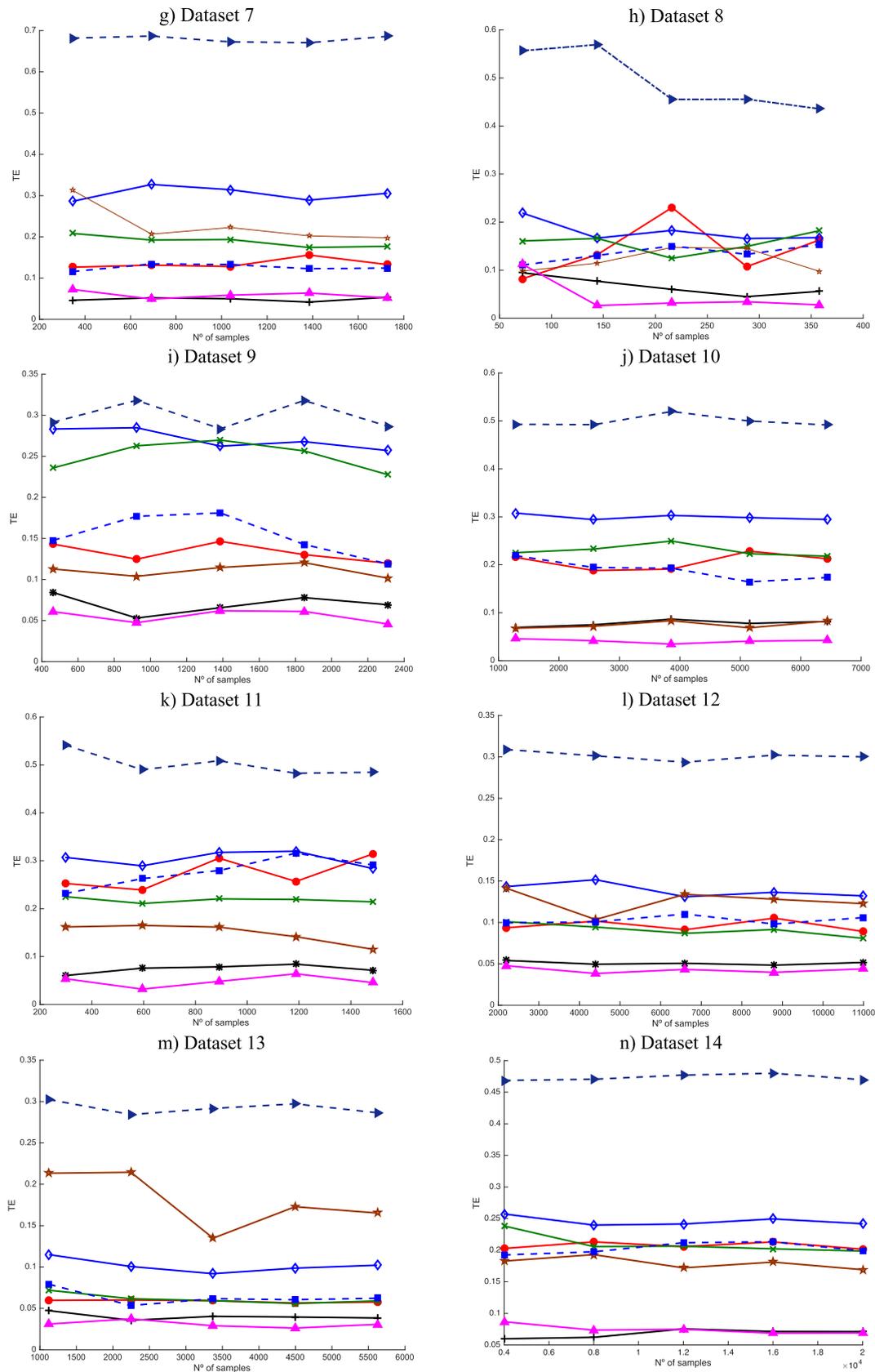


FIGURE 8. (Continued.) TE measures vs number of samples of each algorithm for each of the 14 datasets.

TABLE 2. Algorithms parameters for each Benchmark dataset.

Dataset	Parameters	SOM	ViSOM	SIM	MLHL-SIM	GNG	Beta-SIM
1	Size	[15,10]	[15,10]	[15,10]	[15,10]	-	[15,10]
	Iterations	3,000	2,000	3,000	3,000	3,000	3,000
	Learning rate	0.1	0.05	0.05	0.05	-	0.05
	Neighborhood	12	10	10	10	-	6
	p	-	-	-	0.5	-	-
	$\alpha;\beta$	-	-	-	-	-	$\alpha=3;\beta=4$
	λ_{ViSOM}	-	0,25	-	-	-	-
	$\lambda_{GNG}; \alpha_{GNG}; \beta_{GNG}; a_{max}; Wl; Wn$	-	-	-	-	200; 0.5; 0.9; 5; 0.01; 0.001	-
2	Size	[15,10]	[15,10]	[15,10]	[15,10]	-	[15,10]
	Iterations	5,000	5,000	5,000	5,000	10,000	5,000
	Learning rate	0.1	0.05	0.1	0.1	-	0.1
	Neighborhood	12	10	10	10	-	6
	p	-	-	-	0.9	-	-
	$\alpha;\beta$	-	-	-	-	-	$\alpha=3;\beta=3.5$
	λ_{ViSOM}	-	0,25	-	-	-	-
	$\lambda_{GNG}; \alpha_{GNG}; \beta_{GNG}; a_{max}; Wl; Wn$	-	-	-	-	200; 0.5; 0.9; 5; 0.01; 0.001	-
3	Size	[15,10]	[15,10]	[15,10]	[15,10]	-	[15,10]
	Iterations	30,000	30,000	30,000	30,000	30,000	50,000
	Learning rate	0.05	0.05	0.05	0.05	-	0.05
	Neighborhood	12	10	10	10	-	6
	p	-	-	-	0.5	-	-
	$\alpha;\beta$	-	-	-	-	-	$\alpha=3;\beta=3.5$
	λ_{ViSOM}	-	0,25	-	-	-	-
	$\lambda_{GNG}; \alpha_{GNG}; \beta_{GNG}; a_{max}; Wl; Wn$	-	-	-	-	100; 0.5; 0.9; 5; 0.01; 0.001	-
4	Size	[15,10]	[15,10]	[15,10]	[15,10]	-	[15,10]
	Iterations	5,000	5,000	5,000	5,000	5,000	5,000
	Learning rate	0.1	0.1	0.1	0.1	-	0.1
	Neighborhood	10	-	5	10	-	12
	p	-	-	-	0.9	-	-
	$\alpha;\beta$	-	-	-	-	-	$\alpha=3;\beta=3$
	λ_{ViSOM}	-	0,1	-	-	-	-
	$\lambda_{GNG}; \alpha_{GNG}; \beta_{GNG}; a_{max}; Wl; Wn$	-	-	-	-	100; 0.5; 0.9; 5; 0.1; 0.001	-
5	Size	[15,10]	[15,10]	[15,10]	[15,10]	-	[15,10]
	Iterations	5,000	5,000	5,000	5,000	5,000	5,000
	Learning rate	0.1	0.1	0.1	0.1	-	0.1
	Neighborhood	12	10	12	10	-	12
	p	-	-	-	0.9	-	-
	$\alpha;\beta$	-	-	-	-	-	$\alpha=4;\beta=4$
	λ_{ViSOM}	-	0.1	-	-	-	-
	$\lambda_{GNG}; \alpha_{GNG}; \beta_{GNG}; a_{max}; Wl; Wn$	-	-	-	-	50; 0.5; 0.9; 5; 0.1; 0.0001	-
6	Size	[15,10]	[15,10]	[15,10]	[15,10]	-	[15,15]
	Iterations	10,000	10,000	10,000	10,000	10,000	50,000
	Learning rate	0.1	0.1	0.1	0.1	-	0.1
	Neighborhood	12	10	12	10	-	5
	p	-	-	-	1.2	-	-
	$\alpha;\beta$	-	-	-	-	-	$\alpha=4;\beta=4$
	λ_{ViSOM}	-	0.2	-	-	-	-
	$\lambda_{GNG}; \alpha_{GNG}; \beta_{GNG}; a_{max}; Wl; Wn$	-	-	-	-	50; 0.5; 0.9; 5; 0.1; 0.0001	-
7	Size	[15,10]	[15,10]	[15,10]	[15,10]	-	[15,10]
	Iterations	10,000	5,000	10,000	10,000	10,000	10,000
	Learning rate	0.1	0.1	0.1	0.1	-	0.1
	Neighborhood	10	10	12	10	-	6
	p	-	-	-	0.9	-	-
	$\alpha;\beta$	-	-	-	-	-	$\alpha=5;\beta=3$
	λ_{ViSOM}	-	0.2	-	-	-	-
	$\lambda_{GNG}; \alpha_{GNG}; \beta_{GNG}; a_{max}; Wl; Wn$	-	-	-	-	100; 0.5; 0.9; 5; 0.01; 0.0001	-

TABLE 2. (Continued.) Algorithms parameters for each Benchmark dataset.

Dataset	Parameters	SOM	ViSOM	SIM	MLHL-SIM	GNG	Beta-SIM
8	Size	[15,10]	[15,10]	[15,10]	[15,10]	-	[15,10]
	Iterations	5,000	5,000	5,000	5,000	5,000	5,000
	Learning rate	0.1	0.1	0.05	0.05	-	0.1
	Neighborhood	10	10	10	10	-	6
	p	-	-	-	0.9	-	-
	$\alpha;\beta$	-	-	-	-	-	$\alpha=4;\beta=3$
	λ_{ViSOM}	-	0.3	-	-	-	-
	$\lambda_{GNG}; \alpha_{GNG}; \beta_{GNG}; a_{max}; Wl; Wn$	-	-	-	-	50; 0.5; 0.9; 5;0.01; 0.0001	-
9	Size	[20,15]	[20,15]	[20,15]	[20,15]	-	[20,15]
	Iterations	10,000	5,000	10,000	10,000	10,000	10,000
	Learning rate	0.1	0.1	0.05	0.1	-	0.1
	Neighborhood	10	10	10	10	-	6
	p	-	-	-	0.9	-	-
	$\alpha;\beta$	-	-	-	-	-	$\alpha=4;\beta=3$
	λ_{ViSOM}	-	0.25	-	-	-	-
	$\lambda_{GNG}; \alpha_{GNG}; \beta_{GNG}; a_{max}; Wl; Wn$	-	-	-	-	100; 0.5; 0.9; 5;0.01; 0.0001	-
10	Size	[25,20]	[25,20]	[25,20]	[25,20]	-	[25,20]
	Iterations	20,000	20,000	20,000	20,000	20,000	20,000
	Learning rate	0.1	0.1	0.1	0.1	-	0.1
	Neighborhood	10	10	10	10	-	6
	p	-	-	-	1.1	-	-
	$\alpha;\beta$	-	-	-	-	-	$\alpha=4;\beta=3$
	λ_{ViSOM}	-	0.25	-	-	-	-
	$\lambda_{GNG}; \alpha_{GNG}; \beta_{GNG}; a_{max}; Wl; Wn$	-	-	-	-	100; 0.5; 0.9; 5;0.01; 0.0001	-
11	Size	[15,10]	[15,10]	[15,10]	[15,10]	-	[15,10]
	Iterations	10,000	10,000	10,000	10,000	10,000	10,000
	Learning rate	0.1	0.1	0.1	0.1	-	0.1
	Neighborhood	12	10	10	10	-	6
	p	-	-	-	0.5	-	-
	$\alpha;\beta$	-	-	-	-	-	$\alpha=4;\beta=3$
	λ_{ViSOM}	-	0.1	-	-	-	-
	$\lambda_{GNG}; \alpha_{GNG}; \beta_{GNG}; a_{max}; Wl; Wn$	-	-	-	-	50; 0.5; 0.9; 5;0.01;0.0001	-
12	Size	[20,15]	[20,15]	[20,15]	[20,15]	-	[20,15]
	Iterations	50,000	50,000	50,000	50,000	50,000	50,000
	Learning rate	0.1	0.1	0.1	0.1	-	0.1
	Neighborhood	12	10	8	8	-	6
	p	-	-	-	0.9	-	-
	$\alpha;\beta$	-	-	-	-	-	$\alpha=4;\beta=3$
	λ_{ViSOM}	-	0.1	-	-	-	-
	$\lambda_{GNG}; \alpha_{GNG}; \beta_{GNG}; a_{max}; Wl; Wn$	-	-	-	-	100; 0.5; 0.9; 5;0.001;0.0001	-
13	Size	[20,10]	[20,10]	[20,10]	[20,10]	-	[20,10]
	Iterations	40,000	40,000	40,000	40,000	40,000	40,000
	Learning rate	0.1	0.1	0.1	0.1	-	0.1
	Neighborhood	12	10	12	12	-	6
	p	-	-	-	0.9	-	-
	$\alpha;\beta$	-	-	-	-	-	$\alpha=4;\beta=3$
	λ_{ViSOM}	-	-	-	-	-	-
	$\lambda_{GNG}; \alpha_{GNG}; \beta_{GNG}; a_{max}; Wl; Wn$	-	-	-	-	100; 0.5; 0.9; 5;0.001;0.0001	-
14	Size	[25,20]	[25,20]	[25,20]	[25,20]	-	[25,20]
	Iterations	80,000	80,000	80,000	80,000	80,000	80,000
	Learning rate	0.1	0.1	0.1	0.1	-	0.1
	Neighborhood	12	10	12	12	-	6
	p	-	-	-	0.5	-	-
	$\alpha;\beta$	-	-	-	-	-	$\alpha=4;\beta=3$
	λ_{ViSOM}	-	0.1	-	-	-	-
	$\lambda_{GNG}; \alpha_{GNG}; \beta_{GNG}; a_{max}; Wl; Wn$	-	-	-	-	100; 0.5; 0.9; 5;0.005;0.0001	-

TABLE 3. p-values for CE (WeVoS-Beta-SIM against all algorithms).

Dataset	N° samples	SIM	Beta-SIM	SOM	WeVoS-SOM	MLHL-SIM	ViSOM	GNG
1	69	0.8958	0.9516	0.3850	0.9991	0.5790	0.7094	0.9999
	138	0.9999	0.9983	0.8807	1.0000	1.0000	0.9962	1.0000
	207	0.9986	0.9968	0.9999	0.9996	0.9999	1.0000	0.9996
	276	0.9983	0.9746	0.9976	0.9991	0.9980	0.9998	0.9998
2	345	0.9989	0.9979	0.9873	0.8236	0.9861	0.9998	1.0000
	20	1.0000	0.9999	0.9938	0.7762	1.0000	0.9938	1.0000
	40	1.0000	0.9990	0.9711	0.9998	1.0000	1.0000	0.9997
	60	0.9992	0.9754	0.7170	0.8110	0.9961	0.9933	0.9992
3	80	0.4724	0.9974	1.0000	0.9959	0.9790	0.9970	0.8462
	100	0.9996	0.9998	0.9705	1.0000	0.9889	0.9970	1.0000
	9043	0.0446 [◇]	0.4195	0.5673	0.5171	0.7138	0.6666	0.0960
	18086	0.9885	0.9993	0.9828	0.8168	0.9530	0.5913	0.9752
4	27129	0.9951	0.9998	0.9559	0.9670	0.9559	1.0000	0.9263
	36172	0.9964	0.9557	0.9987	0.9972	0.9999	0.9495	0.9954
	45211	1.0000	0.9907	0.9993	0.9827	0.9624	0.9954	0.9290
	30	0.9952	0.9162	1.0000	0.9996	0.9162	0.9996	1.0000
5	60	0.9974	0.9997	1.0000	1.0000	1.0000	1.0000	1.0000
	90	1.0000	1.0000	0.9991	0.9871	1.0000	1.0000	0.9991
	120	0.9999	0.9934	0.9505	0.9999	0.9998	0.9960	0.9438
	150	1.0000	1.0000	0.9942	1.0000	1.0000	0.9942	1.0000
6	36	1.0000	1.0000	1.0000	0.9905	1.0000	1.0000	0.9991
	72	0.9897	0.9992	1.0000	0.9869	0.9992	0.9869	1.0000
	108	0.9994	0.9962	1.0000	0.4038	0.9996	0.9120	1.0000
	144	1.0000	1.0000	0.9937	0.9996	1.0000	1.0000	0.9997
7	178	0.8318	0.9895	0.9895	0.9895	1.0000	0.4805	0.9895
	295	1.0000	1.0000	0.5788	0.8551	1.0000	0.9915	0.9453
	590	0.9998	1.0000	1.0000	0.9999	0.9997	1.0000	0.9997
	885	0.9995	0.9976	1.0000	0.9990	0.9701	0.9999	0.9977
8	1180	0.9430	0.9993	0.9999	1.0000	0.9855	1.0000	0.9997
	1473	0.9897	0.9838	0.8343	1.0000	0.7163	0.9998	1.0000
	346	0.9946	0.9999	1.0000	1.0000	0.9546	0.9997	0.4479
	692	1.0000	1.0000	0.9988	0.9995	0.9828	1.0000	0.4248
9	1038	0.7513	1.0000	0.9896	0.9993	1.0000	1.0000	0.8534
	1384	0.9932	1.0000	0.9154	0.9999	0.8827	0.9997	0.6597
	1728	0.9879	0.9997	1.0000	1.0000	0.9973	0.7986	0.0737
	72	0.5625	0.5625	0.5625	0.3524	0.1571	0.5625	0.3524
10	144	1.0000	0.9997	0.9928	1.0000	0.9892	1.0000	1.0000
	216	0.9345	0.9954	0.9854	0.7354	0.9995	0.9999	0.6554
	288	0.9923	0.8485	0.9999	0.8346	1.0000	0.9730	0.9989
	358	0.1169	1.0000	1.0000	0.9646	0.6784	0.9994	1.0000
11	462	0.0000*	0.9978	0.1091	0.0012*	0.0056*	0.2371	0.0662
	924	0.0000*	1.0000	0.0000*	0.0000*	0.0000*	0.0000*	0.0000*
	1386	0.0000*	0.9957	0.0000*	0.0000*	0.0000*	0.0000*	0.0000*
	1848	0.0000*	0.9613	0.0000*	0.0000*	0.0000*	0.0000*	0.0000*
12	2310	0.0000*	0.9228	0.0000*	0.0000*	0.0000*	0.0000*	0.0000*
	1287	0.9713	0.9860	0.9864	0.8818	0.9265	0.8279	0.0000*
	2574	0.4136	0.8679	0.0014*	0.0927	0.4720	0.1590	0.0000*
	3861	0.8162	1.0000	0.4380	0.2670	0.9873	0.8016	0.0000*
13	5148	0.0188*	1.0000	0.0003*	0.0015*	0.0348*	0.1331	0.0000*
	6435	0.0000*	0.9967	0.0000*	0.0000*	0.0005*	0.0001*	0.0000*
	297	0.7997	0.9865	0.5333	0.4788	0.9089	0.8765	0.8101
	594	0.9817	0.9976	0.9663	0.6593	0.9159	1.0000	1.0000
14	891	0.4254	0.7712	0.5029	0.0757	0.4165	0.9821	0.8832
	1188	0.9159	0.9991	0.4903	0.0364*	0.9675	0.9953	1.0000
	1484	0.9475	0.9993	0.1509	0.1201	0.9947	0.8352	0.9996
	2199	0.9973	0.9997	0.2056	0.0144*	0.9881	0.2753	0.7632
15	4398	0.9982	0.9883	0.1169	0.0000*	0.9691	0.1542	0.8215
	6597	0.2249	1.0000	0.0010*	0.0000*	0.7232	0.0692	0.3049
	8796	0.2290	0.7177	0.0001*	0.0000*	0.7059	0.0001*	0.2002
	10992	0.7618	0.9983	0.0001*	0.0000*	0.4568	0.0000*	0.0926
16	1124	0.7862	0.9929	0.0434*	0.0000*	0.9645	0.1407	1.0000
	2248	0.0487*	1.0000	0.0010*	0.0000*	0.6557	0.0000*	0.9196
	3372	0.2678	1.0000	0.0052*	0.0000*	0.6724	0.0003*	0.8906
	4496	0.4075	1.0000	0.0001*	0.0000*	0.4241	0.0000*	0.5245
17	5620	0.0379*	1.0000	0.0000*	0.0000*	0.6031	0.0000*	0.2933
	4000	0.0000*	0.8880	0.0000*	0.0000*	0.0000*	0.0000*	1.0000
	8000	0.0000*	0.4548	0.0000*	0.0000*	0.0000*	0.0000*	0.9745
	12000	0.0000*	0.7614	0.0000*	0.0000*	0.0000*	0.0000*	0.8839
18	16000	0.0000*	0.2165	0.0000*	0.0000*	0.0000*	0.0000*	0.8033
	20000	0.0000*	0.4712	0.0000*	0.0000*	0.0000*	0.0000*	0.9900

* Beta-SIM is significantly better (significance level of 0.05)

◇ Beta-SIM is significantly worse (significance level of 0.05)

TABLE 4. p-values for MQE (WeVoS-Beta-SIM against all algorithms).

Dataset	N° samples	SIM	Beta-SIM	SOM	WeVoS-SOM	MLHL-SIM	ViSOM	GNG
1	69	0.9999	0.9995	0.9493	0.1308	1.0000	0.8491	0.9978
	138	1.0000	1.0000	0.7775	0.0061*	1.0000	0.3215	1.0000
	207	1.0000	1.0000	0.0935	0.0000*	0.9871	0.0108*	1.0000
	276	0.9384	1.0000	0.0043*	0.0000*	0.9646	0.0004*	1.0000
	345	0.9972	1.0000	0.2095	0.0004*	0.9997	0.0409*	1.0000
2	20	1.0000	1.0000	0.9206	0.8572	1.0000	1.0000	1.0000
	40	1.0000	1.0000	0.0018*	0.0000*	1.0000	1.0000	0.9872
	60	0.9768	0.7329	0.0000*	0.0000*	0.9911	0.9880	1.0000
	80	0.9992	0.9962	0.0000*	0.0000*	0.9954	1.0000	0.6938
	100	1.0000	1.0000	0.0000*	0.0000*	0.9984	0.9978	0.6521
3	9043	0.4056	0.6412	0.0000*	0.0000*	0.0639	0.0000 [◇]	0.0000 [◇]
	18086	1.0000	0.0325*	0.0000*	0.0000*	0.8685	0.0058 [◇]	0.0030 [◇]
	27129	0.1249	0.7730	0.0000*	0.0000*	0.0421 [◇]	0.0000 [◇]	0.0000 [◇]
	36172	1.0000	0.0363*	0.0000*	0.0000*	0.9498	0.0162 [◇]	0.0065 [◇]
	45211	0.0027 [◇]	0.9836	0.0000*	0.0000*	0.0000 [◇]	0.0000 [◇]	0.0000 [◇]
4	30	0.9994	1.0000	0.4264	0.0054*	1.0000	0.9837	1.0000
	60	0.9821	1.0000	0.3470	0.0019*	0.9989	1.0000	1.0000
	90	1.0000	0.9985	0.0114*	0.0000*	0.9999	0.9988	1.0000
	120	0.9118	1.0000	0.0000*	0.0000*	0.4734	0.9940	1.0000
	150	0.9995	0.9993	0.0234*	0.0000*	0.9962	0.9975	0.9993
5	36	1.0000	1.0000	0.9999	0.9471	1.0000	1.0000	1.0000
	72	1.0000	1.0000	0.9914	0.4144	0.9999	1.0000	0.9999
	108	0.9997	0.9995	0.9713	0.0387*	0.9025	0.9626	0.6379
	144	0.1953	0.9969	0.0727	0.0000*	0.6331	0.6876	0.0012 [◇]
	178	0.9605	1.0000	0.4684	0.0023*	0.9836	0.9973	0.1763
6	295	0.0018*	1.0000	0.0000*	0.0000*	0.1054	0.7699	0.0000 [◇]
	590	0.0000*	0.8916	0.0000*	0.0000*	0.0026*	0.0285*	0.0000 [◇]
	885	0.0010*	0.9922	0.0000*	0.0000*	0.0504	0.2314	0.0000 [◇]
	1180	0.0000*	0.9721	0.0000*	0.0000*	0.0000*	0.0069*	0.0000 [◇]
	1473	0.0197*	0.0475*	0.0000*	0.0000*	0.1789	0.3657	0.0000 [◇]
7	346	0.0066*	0.9989	0.0000*	0.0000*	0.7893	0.1484	0.0000 [◇]
	692	0.0000*	0.0074*	0.0000*	0.0000*	0.0141*	0.0030*	0.0000 [◇]
	1038	0.0000*	0.1972	0.0000*	0.0000*	0.3101	0.1014	0.0000 [◇]
	1384	0.0000*	0.0219*	0.0000*	0.0000*	0.0713	0.0224*	0.0000 [◇]
	1728	0.0000*	0.0005*	0.0000*	0.0000*	0.0002*	0.0000*	0.0000 [◇]
8	72	0.3207	0.4928	0.2220	0.0052*	0.3928	0.3437	0.7936
	144	0.0007*	0.2966	0.0001*	0.0000*	0.0012*	0.0008*	0.4246
	216	0.0000*	0.9768	0.0000*	0.0000*	0.0000*	0.0001*	0.8949
	288	0.0000*	1.0000	0.0000*	0.0000*	0.0000*	0.0000*	0.9927
	358	0.0000*	1.0000	0.0000*	0.0000*	0.0000*	0.0000*	0.9865
9	462	0.0000*	0.9997	0.0000*	0.0000*	0.0000*	0.0000*	0.0000*
	924	0.0000*	0.9759	0.0000*	0.0000*	0.0000*	0.0000*	0.0000*
	1386	0.0000*	0.9898	0.0000*	0.0000*	0.0000*	0.0000*	0.0000*
	1848	0.0000*	0.9420	0.0000*	0.0000*	0.0000*	0.0000*	0.0000*
	2310	0.0000*	0.9067	0.0000*	0.0000*	0.0000*	0.0000*	0.0000*
10	1287	0.0000*	0.9907	0.0000*	0.0000*	0.0000*	0.0000*	0.0000*
	2574	0.0000*	0.9386	0.0000*	0.0000*	0.0000*	0.0000*	0.0000*
	3861	0.0000*	0.8347	0.0000*	0.0000*	0.0000*	0.0000*	0.0000*
	5148	0.0000*	0.0458*	0.0000*	0.0000*	0.0000*	0.0000*	0.0000*
	6435	0.0000*	0.0030*	0.0000*	0.0000*	0.0000*	0.0000*	0.0000*
11	297	0.6356	1.0000	0.0006*	0.0000*	0.6983	0.2171	0.4205
	594	0.0305*	1.0000	0.0000*	0.0000*	0.2113	0.0009*	0.0092*
	891	0.0111*	0.9991	0.0000*	0.0000*	0.0107*	0.0000*	0.0005*
	1188	0.0060*	0.9982	0.0000*	0.0000*	0.0404*	0.0000*	0.0008*
	1484	0.0000*	0.9394	0.0000*	0.0000*	0.0000*	0.0000*	0.0000*
12	2199	0.2761	0.9992	0.0000*	0.0000*	0.7677	0.0000*	0.0000*
	4398	0.2332	0.9844	0.0000*	0.0000*	0.6176	0.0000*	0.0000*
	6597	0.4586	0.9998	0.0000*	0.0000*	0.7756	0.0000*	0.0000*
	8796	0.3557	0.9939	0.0000*	0.0000*	0.7204	0.0000*	0.0000*
	10992	0.3036	0.9989	0.0000*	0.0000*	0.5867	0.0000*	0.0000*
13	1124	0.3653	0.9999	0.0000*	0.0000*	0.6107	0.0000*	0.0000*
	2248	0.0000*	0.4692	0.0000*	0.0000*	0.0000*	0.0000*	0.0000*
	3372	0.0000*	0.4019	0.0000*	0.0000*	0.0000*	0.0000*	0.0000*
	4496	0.0000*	0.9553	0.0000*	0.0000*	0.0000*	0.0000*	0.0000*
	5620	0.0000*	0.5695	0.0000*	0.0000*	0.0000*	0.0000*	0.0000*
14	4000	0.0000*	0.2528	0.0000*	0.0000*	0.0000*	0.0000*	0.0011*
	8000	0.0000*	0.5156	0.0000*	0.0000*	0.0000*	0.0000*	0.0114*
	12000	0.0000*	0.0120*	0.0000*	0.0000*	0.0000*	0.0000*	0.0000*
	16000	0.0000*	0.0443*	0.0000*	0.0000*	0.0000*	0.0000*	0.0000*
	20000	0.0000*	0.4932	0.0000*	0.0000*	0.0000*	0.0000*	0.0005*

* Beta-SIM is significantly better (significance level of 0.05)

◇ Beta-SIM is significantly worse (significance level of 0.05)

TABLE 5. p-values for TE (WeVoS-Beta-SIM against all algorithms).

Dataset	N° samples	SIM	Beta-SIM	SOM	WeVoS-SOM	MLHL-SIM	ViSOM	GNG
1	69	0.1601	0.0993	0.0013 [◊]	0.0040 [◊]	0.2797	0.0000 [◊]	0.0000*
	138	0.9981	0.7576	0.0000 [◊]	0.0016 [◊]	0.9999	0.0000 [◊]	0.0000*
	207	0.6231	0.9940	0.0000 [◊]	0.0143 [◊]	0.9553	0.0000 [◊]	0.0000*
	276	0.9989	0.9926	0.0000 [◊]	0.0001 [◊]	0.9990	0.0000 [◊]	0.0000*
	345	0.3572	1.0000	0.0000 [◊]	0.0000 [◊]	0.7626	0.0000 [◊]	0.0000*
2	20	0.9998	1.0000	0.8518	0.8518	1.0000	0.5728	0.0017*
	40	0.9617	0.9985	0.6080	0.8920	0.1990	0.7492	0.0000*
	60	0.9985	1.0000	0.2883	1.0000	0.8267	0.9989	0.0000*
	80	0.5303	1.0000	0.0392 [◊]	0.6881	0.9999	0.2693	0.0080*
	100	0.6678	0.7758	0.0168 [◊]	0.0168 [◊]	0.1298	0.4541	0.0000*
3	9043	0.0022 [◊]	0.9957	0.0000 [◊]	0.1702	0.0047 [◊]	0.0000 [◊]	0.0000*
	18086	0.0373 [◊]	0.9351	0.0001 [◊]	0.9390	0.4378	0.0005 [◊]	0.0000*
	27129	0.0006 [◊]	0.9370	0.0000 [◊]	0.9267	0.0039 [◊]	0.0000 [◊]	0.0000*
	36172	0.3571	0.7424	0.0121 [◊]	0.0074*	0.4952	0.0229 [◊]	0.0000*
	45211	0.0116 [◊]	1.0000	0.0000 [◊]	1.0000	0.0024 [◊]	0.0000 [◊]	0.0000*
4	30	0.9603	0.9866	1.0000	0.5595	0.8594	0.9969	0.0539
	60	0.9879	0.0284*	0.7728	0.9999	0.5427	0.2593	0.0000*
	90	0.7122	0.1744	0.9989	0.9689	0.9955	0.5864	0.0000*
	120	0.9986	0.9999	0.3425	0.7344	0.9993	0.1675	0.0000*
	150	0.9931	0.9978	0.0879	0.5923	0.1198	0.0007 [◊]	0.0000*
5	36	0.9876	0.1204	0.9329	1.0000	0.8742	1.0000	0.0000*
	72	0.3931	0.9694	0.0111 [◊]	0.0655	0.8914	0.0004 [◊]	0.0000*
	108	0.9978	0.0038*	0.2146	0.4834	0.9999	0.0062 [◊]	0.0000*
	144	0.9992	0.8896	0.0049 [◊]	0.4388	0.9630	0.0001 [◊]	0.0000*
	178	0.3531	0.0028*	0.0796	0.7623	0.8855	0.0282 [◊]	0.0000*
6	295	0.1927	0.3963	0.0005 [◊]	0.9257	0.2522	0.0005 [◊]	0.0000*
	590	0.1917	0.7446	0.0177 [◊]	0.4426	0.1582	0.0006 [◊]	0.0000*
	885	1.0000	0.7656	0.0446 [◊]	0.0563	1.0000	0.0445 [◊]	0.0000*
	1180	0.8835	0.9722	0.0078 [◊]	0.9782	1.0000	0.0267 [◊]	0.0000*
	1473	1.0000	0.0056*	0.3620	0.0001*	1.0000	0.2847	0.0000*
7	346	0.0804	0.1081	0.0000 [◊]	0.0080*	0.0288 [◊]	0.0002 [◊]	0.0000*
	692	0.0771	0.0000*	0.0000 [◊]	0.9967	0.1065	0.0000 [◊]	0.0000*
	1038	0.0236 [◊]	0.0000*	0.0000 [◊]	0.8051	0.0471 [◊]	0.0000 [◊]	0.0000*
	1384	0.9576	0.0000*	0.0000 [◊]	0.6972	0.0562	0.0000 [◊]	0.0000*
	1728	0.0311 [◊]	0.0000*	0.0000 [◊]	0.7735	0.0043 [◊]	0.0000 [◊]	0.0000*
8	72	0.9396	0.9883	0.9761	0.9826	0.9954	0.9963	0.0000*
	144	0.4076	1.0000	0.0248 [◊]	0.2021	0.3858	0.0006 [◊]	0.0000*
	216	0.1169	0.7906	0.6833	0.9988	0.9975	0.2318	0.0000*
	288	0.8899	0.9554	0.0000 [◊]	1.0000	0.9739	0.0000 [◊]	0.0000*
	358	0.9985	0.9998	0.0025 [◊]	0.1206	0.9817	0.0001 [◊]	0.0000*
9	462	0.0579	0.7664	0.0001 [◊]	0.0029 [◊]	0.0842	0.0000 [◊]	0.5963
	924	0.0000 [◊]	0.9797	0.0000 [◊]	0.0000 [◊]	0.0091 [◊]	0.0000 [◊]	0.2574
	1386	0.0000 [◊]	1.0000	0.0000 [◊]	0.0000 [◊]	0.0028 [◊]	0.0000 [◊]	0.9982
	1848	0.0000 [◊]	0.8863	0.0000 [◊]	0.0000 [◊]	0.0000 [◊]	0.0000 [◊]	0.0151*
	2310	0.0000 [◊]	0.6161	0.0000 [◊]	0.0000 [◊]	0.0000 [◊]	0.0000 [◊]	0.0122*
10	1287	1.0000	0.0459*	0.0000 [◊]	0.0000 [◊]	1.0000	0.0000 [◊]	0.0000*
	2574	0.6093	0.2119	0.0000 [◊]	0.0000 [◊]	0.7702	0.0000 [◊]	0.0000*
	3861	0.0426 [◊]	0.0461*	0.0000 [◊]	0.0000 [◊]	0.0555	0.0000 [◊]	0.0000*
	5148	1.0000	0.0198*	0.0000 [◊]	0.0000 [◊]	0.1377	0.0000 [◊]	0.0000*
	6435	1.0000	0.0276*	0.0000 [◊]	0.0000 [◊]	0.5172	0.0000 [◊]	0.0000*
11	297	0.9867	0.1683	0.0000 [◊]	0.4861	1.0000	0.0000 [◊]	0.0000*
	594	0.8594	0.0126*	0.0000 [◊]	0.5720	0.2472	0.0000 [◊]	0.0000*
	891	0.0860	0.0204*	0.0000 [◊]	0.0162 [◊]	0.6542	0.0000 [◊]	0.0000*
	1188	0.6577	0.0003*	0.0000 [◊]	0.0121 [◊]	0.0007*	0.0000 [◊]	0.0000*
	1484	0.0000*	0.0013*	0.0000 [◊]	0.0000 [◊]	0.0002*	0.0000 [◊]	0.0000*
12	2199	0.9997	0.1034	0.0489 [◊]	0.1372	1.0000	0.0146 [◊]	0.0000*
	4398	0.9987	0.0002*	0.0084 [◊]	0.9946	0.9994	0.0004 [◊]	0.0000*
	6597	0.9898	0.0439*	0.7897	0.0272*	0.5190	0.5132	0.0000*
	8796	0.9971	0.2112	0.1674	0.4805	1.0000	0.0506	0.0000*
	10992	0.9967	0.0011*	0.2249	0.0148*	0.4110	0.0494 [◊]	0.0000*
13	1124	0.9922	0.1021	0.7303	0.0000*	0.9997	0.1430	0.0000*
	2248	0.9996	0.0470*	0.1349	0.0000*	0.9522	0.1879	0.0000*
	3372	1.0000	0.0355*	0.5704	0.0000*	1.0000	0.0682	0.0000*
	4496	1.0000	0.0000*	0.5283	0.0000*	0.9963	0.0114 [◊]	0.0000*
	5620	1.0000	0.0018*	0.4892	0.0000*	1.0000	0.1211	0.0000*
14	4000	0.0630	0.7376	0.0000 [◊]	0.0003 [◊]	0.0049 [◊]	0.0000 [◊]	0.0000*
	8000	1.0000	0.3346	0.0000 [◊]	0.6199	0.8500	0.0000 [◊]	0.0000*
	12000	0.9996	0.0008*	0.0000 [◊]	0.0001 [◊]	1.0000	0.0000 [◊]	0.0000*
	16000	0.8154	0.0000*	0.0000 [◊]	0.0976	0.8015	0.0000 [◊]	0.0000*
	20000	1.0000	0.0003*	0.0000 [◊]	0.0404 [◊]	1.0000	0.0000 [◊]	0.0000*

* Beta-SIM is significantly better (significance level of 0.05)

◊ Beta-SIM is significantly worse (significance level of 0.05)

TABLE 6. Average testing CE \pm STD over the 14 benchmark datasets.

Dataset	N° samples	SIM	Beta-SIM	WeVoS-Beta-SIM	SOM	WeVoS-SOM	MLHL-SIM	ViSOM	GNG
1	69	0.491±0.2007	0.476±0.1791	0.383±0.2025	0.563±0.2089	0.430±0.1539	0.538±0.2178	0.521±0.1393	0.349±0.1735
	138	0.442±0.0860	0.457±0.1524	0.419±0.0938	0.501±0.1204	0.407±0.1761	0.413±0.1482	0.375±0.1206	0.396±0.1676
	207	0.468±0.1189	0.464±0.0761	0.497±0.0994	0.478±0.1292	0.473±0.1099	0.479±0.1254	0.511±0.0820	0.473±0.1086
	276	0.468±0.0717	0.399±0.0775	0.441±0.1164	0.413±0.0792	0.417±0.1062	0.469±0.1010	0.460±0.0981	0.460±0.1015
2	345	0.423±0.0720	0.473±0.1093	0.447±0.0832	0.482±0.1008	0.505±0.0835	0.411±0.0815	0.429±0.0852	0.455±0.0906
	20	0.150±0.2415	0.200±0.3496	0.150±0.2415	0.250±0.3536	0.350±0.3375	0.150±0.2415	0.250±0.2635	0.150±0.2415
	40	0.118±0.1306	0.163±0.1927	0.123±0.1350	0.193±0.1607	0.155±0.1442	0.143±0.1294	0.135±0.1667	0.090±0.1663
	60	0.158±0.1660	0.127±0.1568	0.199±0.2467	0.079±0.0833	0.090±0.1405	0.146±0.1168	0.141±0.1733	0.158±0.1660
3	80	0.286±0.1802	0.212±0.1288	0.176±0.1071	0.174±0.1185	0.137±0.0664	0.228±0.1084	0.213±0.1043	0.253±0.1256
	100	0.237±0.1530	0.235±0.1499	0.208±0.0957	0.149±0.0958	0.189±0.1078	0.258±0.1784	0.249±0.1219	0.228±0.1131
	9043	0.112±0.0048	0.114±0.0023	0.117±0.0026	0.114±0.0013	0.114±0.0007	0.115±0.0053	0.114±0.0022	0.113±0.0040
	18086	0.117±0.0008	0.117±0.0019	0.117±0.0029	0.116±0.0011	0.116±0.0011	0.116±0.0010	0.116±0.0026	0.116±0.0030
4	27129	0.117±0.0015	0.117±0.0017	0.117±0.0029	0.118±0.0004	0.117±0.0007	0.118±0.0005	0.117±0.0013	0.118±0.0005
	36172	0.119±0.0005	0.117±0.0037	0.118±0.0026	0.119±0.0003	0.119±0.0003	0.118±0.0023	0.117±0.0026	0.117±0.0027
	45211	0.117±0.0011	0.117±0.0013	0.117±0.0009	0.117±0.0009	0.117±0.0002	0.117±0.0009	0.116±0.0017	0.117±0.0010
	30	0.083±0.1800	0.117±0.1933	0.033±0.1054	0.033±0.1054	0.067±0.1405	0.117±0.1933	0.067±0.1405	0.033±0.1054
5	60	0.081±0.0856	0.134±0.1066	0.112±0.1106	0.114±0.1116	0.100±0.1165	0.098±0.0843	0.098±0.0843	0.114±0.1116
	90	0.053±0.1068	0.042±0.0892	0.043±0.0767	0.064±0.0898	0.076±0.0740	0.054±0.0778	0.043±0.0560	0.064±0.0764
	120	0.051±0.0588	0.042±0.0442	0.060±0.0693	0.034±0.0441	0.051±0.0440	0.049±0.0561	0.043±0.0456	0.033±0.0433
	150	0.047±0.0549	0.047±0.0632	0.053±0.0422	0.073±0.0584	0.060±0.0663	0.047±0.0450	0.073±0.0663	0.047±0.0632
6	36	0.092±0.1493	0.092±0.1493	0.098±0.1622	0.103±0.1372	0.157±0.1421	0.100±0.1610	0.078±0.1301	0.138±0.2061
	72	0.068±0.1147	0.082±0.1487	0.111±0.0832	0.096±0.1117	0.066±0.1237	0.082±0.1142	0.066±0.0915	0.109±0.1066
	108	0.064±0.0941	0.071±0.0777	0.045±0.0475	0.056±0.0781	0.121±0.0971	0.027±0.0584	0.090±0.1074	0.045±0.0616
	144	0.063±0.0710	0.062±0.0677	0.056±0.0560	0.034±0.0475	0.041±0.0357	0.062±0.0775	0.062±0.0612	0.070±0.0744
7	178	0.057±0.0547	0.045±0.0440	0.028±0.0300	0.045±0.0354	0.045±0.0440	0.034±0.0405	0.068±0.0648	0.045±0.0238
	295	0.616±0.1162	0.617±0.1116	0.618±0.0650	0.539±0.0783	0.558±0.1048	0.603±0.0929	0.583±0.0861	0.569±0.0927
	590	0.600±0.0717	0.619±0.0551	0.614±0.0584	0.613±0.0757	0.602±0.0815	0.600±0.0415	0.612±0.0726	0.600±0.0463
	885	0.586±0.0613	0.590±0.0533	0.573±0.0514	0.581±0.0421	0.558±0.0899	0.547±0.0483	0.584±0.0597	0.590±0.0318
8	1180	0.543±0.0385	0.577±0.0338	0.566±0.0464	0.558±0.0302	0.566±0.0413	0.548±0.0526	0.561±0.0479	0.557±0.0591
	1473	0.553±0.0475	0.551±0.0158	0.569±0.0510	0.541±0.0404	0.563±0.0438	0.537±0.0381	0.560±0.0391	0.566±0.0575
	346	0.269±0.0497	0.257±0.0753	0.245±0.0529	0.254±0.0387	0.239±0.0754	0.211±0.0785	0.260±0.0750	0.181±0.0914
	692	0.254±0.0283	0.253±0.0645	0.247±0.0544	0.260±0.0713	0.236±0.0369	0.227±0.0469	0.248±0.0446	0.201±0.0314
9	1038	0.197±0.0296	0.225±0.0545	0.227±0.0545	0.211±0.0443	0.217±0.0392	0.223±0.0337	0.228±0.0333	0.200±0.0474
	1384	0.217±0.0332	0.228±0.0304	0.231±0.0531	0.210±0.0373	0.238±0.0262	0.208±0.0398	0.240±0.0455	0.201±0.0386
	1728	0.203±0.0262	0.222±0.0360	0.215±0.0360	0.211±0.0283	0.214±0.0265	0.225±0.0385	0.237±0.0292	0.172±0.0372
	72	0.041±0.0663	0.041±0.0663	0.113±0.1754	0.041±0.0663	0.029±0.0602	0.013±0.0395	0.041±0.0663	0.029±0.0602
10	144	0.068±0.0614	0.050±0.0482	0.062±0.0404	0.042±0.0361	0.055±0.0535	0.084±0.0811	0.062±0.0767	0.058±0.0476
	216	0.079±0.0453	0.070±0.0557	0.054±0.0410	0.036±0.0347	0.088±0.0537	0.065±0.0561	0.046±0.0472	0.019±0.0240
	288	0.076±0.0553	0.038±0.0306	0.062±0.0141	0.055±0.0400	0.038±0.0407	0.063±0.0430	0.045±0.0369	0.052±0.0299
	358	0.089±0.0389	0.039±0.0322	0.044±0.0338	0.039±0.0269	0.061±0.0362	0.072±0.0458	0.053±0.0304	0.045±0.0295
11	462	0.212±0.0723	0.098±0.0297	0.082±0.0270	0.147±0.0555	0.182±0.0436	0.171±0.0620	0.139±0.0491	0.152±0.0551
	924	0.157±0.0359	0.066±0.0164	0.066±0.0245	0.147±0.0313	0.170±0.0389	0.139±0.0427	0.136±0.0300	0.144±0.0256
	1386	0.175±0.0287	0.083±0.0147	0.075±0.0128	0.154±0.0291	0.186±0.0404	0.157±0.0149	0.141±0.0175	0.155±0.0231
	1848	0.154±0.0277	0.082±0.0205	0.070±0.0119	0.141±0.0248	0.176±0.0267	0.151±0.0309	0.127±0.0245	0.134±0.0276
12	2310	0.152±0.0210	0.079±0.0189	0.067±0.0146	0.146±0.0247	0.160±0.0289	0.145±0.0178	0.129±0.0293	0.135±0.0174
	1287	0.158±0.0280	0.156±0.0325	0.145±0.0358	0.156±0.0194	0.162±0.0249	0.160±0.0329	0.163±0.0218	0.304±0.0195
	2574	0.160±0.0231	0.154±0.0133	0.143±0.0220	0.178±0.0198	0.167±0.0153	0.160±0.0153	0.165±0.0196	0.300±0.0171
	3861	0.158±0.0163	0.144±0.0184	0.144±0.0255	0.164±0.0199	0.167±0.0238	0.152±0.0231	0.158±0.0225	0.314±0.0183
13	5148	0.157±0.0133	0.139±0.0071	0.139±0.0106	0.164±0.0149	0.162±0.0140	0.156±0.0088	0.153±0.0162	0.309±0.0090
	6435	0.159±0.0131	0.135±0.0088	0.131±0.0110	0.158±0.0145	0.163±0.0083	0.154±0.0135	0.157±0.0112	0.312±0.0086
	297	0.468±0.0729	0.494±0.1023	0.531±0.1070	0.450±0.1084	0.447±0.0723	0.478±0.0971	0.475±0.0697	0.469±0.1051
	594	0.483±0.0634	0.475±0.0942	0.454±0.0788	0.487±0.0499	0.509±0.0464	0.493±0.0721	0.458±0.0803	0.450±0.0578
14	891	0.498±0.0552	0.485±0.0302	0.446±0.0542	0.495±0.0662	0.520±0.0502	0.499±0.0646	0.469±0.0485	0.479±0.0654
	1188	0.470±0.0433	0.454±0.0297	0.441±0.0666	0.488±0.0351	0.517±0.0612	0.465±0.0568	0.459±0.0530	0.438±0.0640
	1484	0.469±0.0471	0.456±0.0462	0.444±0.0470	0.503±0.0569	0.505±0.0372	0.461±0.0425	0.476±0.0594	0.456±0.0508
	2199	0.089±0.0154	0.087±0.0312	0.081±0.0291	0.111±0.0276	0.124±0.0333	0.091±0.0293	0.110±0.0179	0.100±0.0278
15	4398	0.089±0.0178	0.090±0.0132	0.084±0.0079	0.104±0.0187	0.123±0.0184	0.091±0.0176	0.103±0.0159	0.095±0.0150
	6597	0.096±0.0079	0.082±0.0085	0.080±0.0160	0.108±0.0110	0.136±0.0266	0.091±0.0135	0.099±0.0104	0.095±0.0110
	8796	0.092±0.0173	0.087±0.0103	0.076±0.0103	0.109±0.0120	0.128±0.0213	0.087±0.0125	0.109±0.0178	0.092±0.0096
	10992	0.085±0.0126	0.080±0.0118	0.076±0.0102	0.105±0.0151	0.123±0.0163	0.088±0.0125	0.105±0.0104	0.092±0.0094
16	1124	0.109±0.0354	0.097±0.0333	0.084±0.0292	0.137±0.0408	0.181±0.0515	0.101±0.0431	0.129±0.0327	0.078±0.0289
	2248	0.101±0.0165	0.076±0.0192	0.074±0.0191	0.120±0.0189	0.191±0.0382	0.093±0.0187	0.130±0.0236	0.087±0.0251
	3372	0.091±0.0097	0.073±0.0095	0.070±0.0193	0.104±0.0170	0.154±0.0338	0.085±0.0147	0.111±0.0185	0.082±0.0206
	4496	0.088±0.0155	0.073±0.0077	0.070±0.0158	0.110±0.0132	0.176±0.0230	0.088±0.0167	0.121±0.0199	0.087±0.0144
17	5620	0.093±0.0109	0.074±0.0056	0.073±0.0131	0.115±0.0120	0.176±0.0245	0.085±0.0146	0.128±0.0167	0.088±0.0094
	4000	0.531±0.0199	0.400±0.0176	0.383±0.0313	0.598±0.0404	0.656±0.0357	0.534±0.0322	0.559±0.0199	0.386±0.0245
	8000	0.510±0.0172	0.383±0.0168	0.366±0.0165	0.561±0.0114	0.658±0.0238	0.504±0.0227	0.541±0.0181	0.374±0.0179
	12000	0.502±0.0183	0.372±0.0164	0.361±0.0126	0.553±0.0235	0.635±0.0140	0.501±0.0134	0.528±0.0178	0.371±0.0133
18	16000	0.511±0.0149	0.372±0.0223	0.351±0.0146	0.546±0.0218	0.640±0.0130	0.493±0.0195	0.518±0.0158	0.364±0.0204
	20000	0.509±0.0198	0.371±0.0205	0.354±0.0182	0.547±0.0193	0.625±0.0233	0.504±0.0171	0.517±0.0202	0.361±0.0090

TABLE 7. Average testing MQE ± STD over the 14 benchmark datasets.

Dataset	N° samples	SIM	Beta-SIM	WeVoS-Beta-SIM	SOM	WeVoS-SOM	MLHL-SIM	ViSOM	GNG
1	69	0.273±0.0581	0.270±0.0588	0.284±0.0547	0.314±0.0601	0.356±0.0836	0.284±0.0529	0.321±0.0582	0.267±0.0349
	138	0.302±0.0593	0.298±0.0544	0.296±0.0575	0.342±0.0771	0.410±0.0949	0.301±0.0557	0.363±0.0693	0.294±0.0482
	207	0.274±0.0342	0.270±0.0392	0.271±0.0355	0.320±0.0427	0.393±0.0469	0.286±0.0393	0.333±0.0325	0.266±0.0285
	276	0.271±0.0234	0.258±0.0368	0.254±0.0319	0.314±0.0424	0.381±0.0540	0.270±0.0301	0.325±0.0188	0.252±0.0240
2	345	0.273±0.0452	0.257±0.0502	0.256±0.0512	0.316±0.0632	0.367±0.0707	0.268±0.0453	0.333±0.0586	0.251±0.0355
	20	1.271±0.4257	1.253±0.4362	1.299±0.4564	1.543±0.3992	1.576±0.4144	1.272±0.4763	1.264±0.4367	1.271±0.4981
	40	1.185±0.2158	1.151±0.3075	1.164±0.1639	1.559±0.1349	1.678±0.1624	1.137±0.2281	1.184±0.2166	1.082±0.2010
	60	1.082±0.1299	1.135±0.1129	1.004±0.3907	1.555±0.0644	1.726±0.1009	1.070±0.1453	1.074±0.1206	0.979±0.1642
	80	1.065±0.1432	1.159±0.1951	1.106±0.2257	1.533±0.1191	1.694±0.1569	1.052±0.1598	1.083±0.1516	0.982±0.1406
3	100	1.083±0.1255	1.080±0.1832	1.087±0.1921	1.546±0.1504	1.661±0.1187	1.044±0.1568	1.042±0.1540	0.968±0.1053
	9043	1.055±0.0212	1.112±0.0442	1.086±0.0461	1.223±0.0185	1.330±0.0460	1.041±0.0311	1.005±0.0259	0.994±0.0142
	18086	1.062±0.0102	1.115±0.0473	1.061±0.0321	1.218±0.0136	1.364±0.0798	1.039±0.0211	0.998±0.0109	0.995±0.0156
	27129	1.047±0.0226	1.101±0.0257	1.082±0.0361	1.232±0.0111	1.367±0.0560	1.042±0.0153	1.004±0.0116	1.000±0.0133
	36172	1.042±0.0216	1.084±0.0194	1.041±0.0218	1.224±0.0181	1.384±0.0683	1.027±0.0104	0.995±0.0126	0.991±0.0155
4	45211	1.048±0.0194	1.113±0.0304	1.101±0.0406	1.224±0.0112	1.349±0.0547	1.030±0.0159	1.001±0.0079	0.999±0.0138
	30	0.150±0.0471	0.159±0.0429	0.162±0.0530	0.212±0.0698	0.254±0.0781	0.159±0.0398	0.141±0.0309	0.171±0.0446
	60	0.189±0.0424	0.173±0.0337	0.170±0.0367	0.216±0.0569	0.256±0.0637	0.182±0.0415	0.172±0.0481	0.176±0.0368
	90	0.157±0.0297	0.146±0.0265	0.157±0.0260	0.216±0.0513	0.255±0.0654	0.150±0.0228	0.147±0.0257	0.157±0.0214
	120	0.168±0.0182	0.154±0.0221	0.152±0.0273	0.221±0.0430	0.259±0.0332	0.177±0.0259	0.162±0.0212	0.156±0.0180
5	150	0.172±0.0232	0.155±0.0255	0.164±0.0251	0.220±0.0536	0.252±0.0694	0.176±0.0224	0.153±0.0208	0.155±0.0199
	36	0.788±0.2192	0.770±0.2263	0.779±0.2192	0.815±0.2219	0.891±0.2175	0.757±0.2419	0.770±0.2309	0.800±0.2016
	72	0.785±0.1029	0.792±0.1327	0.799±0.1349	0.843±0.1338	0.913±0.1594	0.779±0.0977	0.781±0.1124	0.778±0.0630
	108	0.731±0.0834	0.729±0.0765	0.748±0.0813	0.782±0.0741	0.858±0.0710	0.704±0.0747	0.711±0.0885	0.687±0.0604
	144	0.703±0.0165	0.729±0.0248	0.739±0.0268	0.780±0.0322	0.845±0.0562	0.713±0.0321	0.715±0.0233	0.678±0.0225
6	178	0.707±0.0532	0.737±0.0635	0.736±0.0757	0.792±0.0591	0.849±0.0762	0.711±0.0538	0.717±0.0567	0.665±0.0436
	295	0.786±0.0393	0.698±0.0493	0.691±0.0567	0.901±0.0606	1.114±0.0607	0.755±0.0496	0.727±0.0491	0.555±0.0309
	590	0.765±0.0540	0.667±0.0320	0.635±0.0518	0.883±0.0580	1.096±0.0929	0.735±0.0453	0.717±0.0436	0.502±0.0413
	885	0.760±0.0463	0.674±0.0404	0.655±0.0391	0.885±0.0516	1.115±0.1007	0.729±0.0433	0.713±0.0498	0.508±0.0247
	1180	0.745±0.0314	0.650±0.0194	0.634±0.0243	0.871±0.0220	1.063±0.0822	0.721±0.0288	0.696±0.0215	0.486±0.0191
7	1473	0.740±0.0241	0.731±0.1435	0.642±0.0323	0.869±0.0376	1.096±0.0805	0.715±0.0220	0.704±0.0284	0.492±0.0189
	346	1.067±0.0433	1.011±0.0413	1.000±0.0616	1.192±0.0176	1.345±0.0421	1.027±0.0317	1.047±0.0273	0.863±0.0349
	692	1.086±0.0440	1.033±0.0300	0.979±0.0222	1.193±0.0353	1.351±0.0326	1.030±0.0342	1.037±0.0272	0.865±0.0270
	1038	1.071±0.0166	1.034±0.0362	1.003±0.0280	1.190±0.0302	1.362±0.0367	1.031±0.0275	1.038±0.0259	0.852±0.0153
	1384	1.077±0.0333	1.039±0.0408	1.000±0.0182	1.209±0.0232	1.377±0.0212	1.034±0.0268	1.038±0.0145	0.861±0.0119
8	1728	1.075±0.0247	1.038±0.0235	0.991±0.0206	1.205±0.0234	1.373±0.0307	1.041±0.0244	1.049±0.0239	0.855±0.0069
	72	1.907±0.1526	1.875±0.1530	1.648±0.5929	1.930±0.1503	2.090±0.1507	1.893±0.1591	1.902±0.1489	1.821±0.1609
	144	1.848±0.1067	1.753±0.0817	1.653±0.1058	1.868±0.0886	1.959±0.1094	1.840±0.1073	1.846±0.0767	1.744±0.0932
	216	1.867±0.1031	1.725±0.0808	1.677±0.0827	1.866±0.0853	1.954±0.0946	1.883±0.0760	1.851±0.0896	1.738±0.0777
	288	1.827±0.0721	1.679±0.0702	1.676±0.0748	1.859±0.0642	1.979±0.0992	1.828±0.0719	1.838±0.0837	1.713±0.0627
9	358	1.831±0.0650	1.685±0.0633	1.675±0.0679	1.861±0.0616	1.911±0.0635	1.826±0.0809	1.830±0.0584	1.705±0.0587
	462	0.467±0.0206	0.317±0.0187	0.310±0.0217	0.524±0.0290	0.628±0.0732	0.456±0.0190	0.510±0.0270	0.542±0.0229
	924	0.474±0.0337	0.338±0.0322	0.323±0.0266	0.542±0.0355	0.621±0.0468	0.468±0.0323	0.537±0.0341	0.557±0.0341
	1386	0.484±0.0172	0.330±0.0124	0.319±0.0123	0.551±0.0142	0.683±0.0727	0.471±0.0161	0.536±0.0190	0.558±0.0187
	1848	0.472±0.0175	0.320±0.0263	0.305±0.0197	0.539±0.0203	0.647±0.0631	0.467±0.0187	0.531±0.0203	0.548±0.0210
10	2310	0.472±0.0164	0.320±0.0172	0.304±0.0122	0.541±0.0151	0.670±0.0686	0.460±0.0191	0.525±0.0143	0.549±0.0153
	1287	0.593±0.0272	0.523±0.0220	0.514±0.0195	0.650±0.0216	0.674±0.0221	0.588±0.0299	0.619±0.0265	0.596±0.0335
	2574	0.595±0.0286	0.518±0.0191	0.505±0.0221	0.651±0.0238	0.674±0.0212	0.592±0.0218	0.629±0.0324	0.599±0.0244
	3861	0.597±0.0163	0.514±0.0135	0.503±0.0098	0.658±0.0180	0.694±0.0179	0.593±0.0222	0.629±0.0151	0.600±0.0199
	5148	0.596±0.0168	0.516±0.0091	0.499±0.0100	0.654±0.0113	0.687±0.0168	0.593±0.0167	0.622±0.0142	0.584±0.0166
11	6435	0.600±0.0127	0.516±0.0081	0.497±0.0092	0.655±0.0119	0.688±0.0103	0.594±0.0103	0.624±0.0113	0.587±0.0094
	297	0.277±0.0268	0.255±0.0201	0.255±0.0277	0.311±0.0311	0.345±0.0273	0.276±0.0325	0.286±0.0269	0.281±0.0269
	594	0.283±0.0225	0.247±0.0157	0.246±0.0149	0.334±0.0318	0.364±0.0287	0.274±0.0238	0.295±0.0266	0.287±0.0251
	891	0.274±0.0159	0.241±0.0122	0.238±0.0093	0.317±0.0243	0.356±0.0346	0.274±0.0150	0.291±0.0211	0.282±0.0183
	1188	0.276±0.0181	0.246±0.0158	0.240±0.0144	0.317±0.0248	0.356±0.0287	0.270±0.0191	0.290±0.0249	0.282±0.0204
12	1484	0.278±0.0124	0.246±0.0108	0.239±0.0118	0.320±0.0131	0.349±0.0222	0.272±0.0109	0.293±0.0117	0.288±0.0120
	2199	0.719±0.0198	0.679±0.0216	0.667±0.0175	0.888±0.0172	1.097±0.1305	0.702±0.0218	0.907±0.0238	0.916±0.0195
	4398	0.695±0.0129	0.667±0.0147	0.651±0.0094	0.872±0.0140	1.032±0.1071	0.683±0.0140	0.890±0.0121	0.893±0.0150
	6597	0.705±0.0103	0.665±0.0149	0.653±0.0146	0.877±0.0061	1.080±0.1605	0.693±0.0129	0.896±0.0096	0.904±0.0096
	8796	0.702±0.0099	0.669±0.0125	0.651±0.0040	0.880±0.0073	1.084±0.1445	0.689±0.0130	0.901±0.0069	0.903±0.0077
13	10992	0.699±0.0109	0.662±0.0081	0.649±0.0124	0.881±0.0068	1.073±0.1346	0.689±0.0056	0.899±0.0065	0.904±0.0096
	1124	2.470±0.0891	2.406±0.0658	2.390±0.0886	2.712±0.0712	3.239±0.1038	2.456±0.0834	2.814±0.0665	2.714±0.0742
	2248	2.444±0.0340	2.347±0.0402	2.311±0.0376	2.693±0.0374	3.230±0.0561	2.445±0.0339	2.807±0.0380	2.691±0.0361
	3372	2.462±0.0278	2.340±0.0336	2.305±0.0185	2.691±0.0211	3.171±0.0785	2.447±0.0271	2.800±0.0185	2.696±0.0274
	4496	2.460±0.0324	2.331±0.0319	2.311±0.0270	2.705±0.0311	3.216±0.0826	2.447±0.0253	2.811±0.0329	2.711±0.0324
14	5620	2.455±0.0238	2.337±0.0205	2.296±0.0174	2.702±0.0158	3.191±0.1283	2.454±0.0158	2.813±0.0197	2.707±0.0202
	4000	0.526±0.0070	0.439±0.0076	0.425±0.0086	0.601±0.0081	0.716±0.0292	0.524±0.0090	0.566±0.0103	0.450±0.0046
	8000	0.527±0.0114	0.436±0.0054	0.423±0.0060	0.603±0.0048	0.735±0.0387	0.525±0.0068	0.568±0.0047	0.447±0.0050
	12000	0.527±0.0040	0.436±0.0037	0.421±0.0035	0.604±0.0042	0.729±0.0218	0.522±0.0071	0.568±0.0040	0.449±0.0053
	16000	0.530±0.0029	0.434±0.0047	0.420±0.0022	0.604±0.0050	0.738±0.0269	0.523±0.0067	0.567±0.0045	0.448±0.0020
20000	0.530±0.0033	0.432±0.0034	0.420±0.0054	0.605±0.0035	0.715±0.0352	0.521±0.0051	0.567±0.0035	0.447±0.0037	

TABLE 8. Average testing TE \pm STD over the 14 benchmark datasets.

Dataset	N° samples	SIM	Beta-SIM	WeVoS-Beta-SIM	SOM	WeVoS-SOM	MLHL-SIM	ViSOM	GNG
1	69	0.239±0.1829	0.224±0.1579	0.417±0.2234	0.127±0.1420	0.149±0.1203	0.258±0.1316	0.039±0.0634	0.768±0.1341
	138	0.364±0.1955	0.311±0.1050	0.401±0.1315	0.057±0.0656	0.160±0.1298	0.378±0.1434	0.093±0.0827	0.805±0.1124
	207	0.280±0.1090	0.391±0.1368	0.357±0.0948	0.073±0.0714	0.204±0.1149	0.310±0.0887	0.072±0.0337	0.787±0.0873
	276	0.311±0.0615	0.363±0.1388	0.333±0.0792	0.087±0.0535	0.149±0.0434	0.312±0.0741	0.066±0.0597	0.744±0.1065
	345	0.319±0.0861	0.386±0.0780	0.391±0.1045	0.108±0.0524	0.140±0.0720	0.340±0.0634	0.055±0.0495	0.730±0.0589
2	20	0.200±0.2582	0.150±0.2415	0.150±0.2415	0.000±0.0000	0.000±0.0000	0.150±0.2415	0.350±0.3375	0.600±0.3162
	40	0.110±0.1912	0.143±0.1601	0.190±0.1404	0.053±0.1167	0.092±0.1493	0.000±0.0000	0.070±0.1135	0.600±0.3074
	60	0.098±0.1395	0.119±0.1431	0.133±0.0994	0.000±0.0000	0.149±0.1466	0.050±0.0805	0.166±0.1422	0.438±0.1767
	80	0.087±0.0799	0.190±0.1909	0.199±0.1774	0.013±0.0395	0.100±0.1294	0.174±0.1323	0.061±0.0645	0.417±0.1434
	100	0.081±0.0793	0.088±0.1082	0.150±0.0972	0.010±0.0316	0.010±0.0316	0.041±0.0711	0.068±0.0793	0.401±0.1517
3	9043	0.054±0.0133	0.118±0.0277	0.109±0.0483	0.027±0.0155	0.074±0.0350	0.057±0.0192	0.021±0.0081	0.323±0.0437
	18086	0.048±0.0135	0.116±0.0419	0.098±0.0174	0.020±0.0065	0.116±0.0740	0.066±0.0169	0.028±0.0060	0.293±0.0370
	27129	0.051±0.0219	0.106±0.0362	0.125±0.0305	0.025±0.0095	0.105±0.0782	0.060±0.0234	0.026±0.0121	0.312±0.0324
	36172	0.050±0.0187	0.123±0.0347	0.092±0.0239	0.023±0.0096	0.164±0.1018	0.054±0.0132	0.027±0.0109	0.302±0.0392
	45211	0.065±0.0317	0.123±0.0456	0.121±0.0382	0.022±0.0050	0.120±0.0632	0.057±0.0125	0.028±0.0100	0.316±0.0276
4	30	0.267±0.3443	0.242±0.3250	0.125±0.2265	0.158±0.2306	0.375±0.3711	0.308±0.3380	0.033±0.1054	0.533±0.3244
	60	0.241±0.2004	0.423±0.2518	0.177±0.1302	0.062±0.1048	0.146±0.1286	0.319±0.1823	0.000±0.0000	0.652±0.1935
	90	0.277±0.1686	0.335±0.1647	0.174±0.1129	0.138±0.1062	0.238±0.1438	0.220±0.1254	0.060±0.0868	0.586±0.1702
	120	0.209±0.0900	0.269±0.1457	0.245±0.1549	0.113±0.1002	0.149±0.0736	0.277±0.1622	0.091±0.0817	0.555±0.1944
	150	0.260±0.1456	0.333±0.1089	0.300±0.1305	0.153±0.1219	0.207±0.1153	0.160±0.1004	0.073±0.0734	0.587±0.0932
5	36	0.205±0.2577	0.375±0.2443	0.127±0.1705	0.020±0.0632	0.140±0.1818	0.248±0.2217	0.125±0.2265	0.618±0.1729
	72	0.211±0.1196	0.288±0.2056	0.357±0.2234	0.109±0.0827	0.152±0.1546	0.268±0.1742	0.041±0.0663	0.764±0.1157
	108	0.225±0.0851	0.378±0.0943	0.195±0.1107	0.081±0.0784	0.103±0.1185	0.213±0.0982	0.019±0.0403	0.661±0.1558
	144	0.231±0.0845	0.318±0.1290	0.257±0.1208	0.076±0.0832	0.161±0.0908	0.208±0.0892	0.021±0.0338	0.694±0.1519
	178	0.310±0.0748	0.395±0.1535	0.208±0.0897	0.074±0.0538	0.136±0.0868	0.269±0.0991	0.056±0.0741	0.725±0.1448
6	295	0.081±0.0322	0.199±0.0870	0.145±0.0620	0.030±0.0399	0.176±0.0726	0.085±0.0586	0.030±0.0247	0.580±0.0437
	590	0.059±0.0293	0.134±0.0564	0.105±0.0429	0.043±0.0259	0.142±0.0580	0.058±0.0379	0.024±0.0198	0.500±0.0329
	885	0.084±0.0249	0.123±0.0483	0.091±0.0521	0.025±0.0167	0.154±0.0879	0.089±0.0180	0.025±0.0175	0.539±0.0540
	1180	0.081±0.0302	0.124±0.0254	0.105±0.0454	0.036±0.0164	0.123±0.0785	0.101±0.0436	0.043±0.0177	0.546±0.0371
	1473	0.081±0.0437	0.155±0.0669	0.081±0.0353	0.039±0.0199	0.176±0.0693	0.075±0.0179	0.037±0.0167	0.508±0.0344
7	346	0.127±0.0570	0.286±0.0815	0.208±0.0407	0.046±0.0281	0.312±0.0952	0.116±0.0522	0.072±0.0543	0.681±0.0613
	692	0.132±0.0392	0.327±0.0594	0.192±0.0633	0.052±0.0268	0.207±0.0633	0.134±0.0244	0.049±0.0316	0.687±0.0386
	1038	0.128±0.0322	0.314±0.0612	0.194±0.0684	0.050±0.0256	0.223±0.0388	0.133±0.0194	0.059±0.0202	0.672±0.0498
	1384	0.156±0.0443	0.289±0.0420	0.174±0.0338	0.042±0.0122	0.202±0.0496	0.123±0.0294	0.064±0.0267	0.670±0.0458
	1728	0.133±0.0309	0.306±0.0425	0.177±0.0293	0.054±0.0244	0.198±0.0346	0.124±0.0255	0.052±0.0160	0.686±0.0286
8	72	0.082±0.0979	0.219±0.1702	0.161±0.1943	0.095±0.1111	0.098±0.1170	0.111±0.1119	0.113±0.1123	0.557±0.2346
	144	0.132±0.1237	0.167±0.0991	0.166±0.0738	0.077±0.0653	0.115±0.0961	0.131±0.0826	0.026±0.0453	0.569±0.1583
	216	0.230±0.1105	0.183±0.0808	0.125±0.0848	0.060±0.0507	0.148±0.0939	0.150±0.0914	0.032±0.0306	0.455±0.1053
	288	0.107±0.0496	0.166±0.0832	0.239±0.0712	0.045±0.0232	0.145±0.0598	0.133±0.1113	0.034±0.0363	0.456±0.1382
	358	0.163±0.0798	0.168±0.0649	0.182±0.0774	0.056±0.0460	0.097±0.0673	0.154±0.0729	0.028±0.0229	0.436±0.0924
9	462	0.143±0.0751	0.283±0.0583	0.236±0.0671	0.084±0.0276	0.113±0.0513	0.148±0.0980	0.061±0.0433	0.292±0.0906
	924	0.125±0.0674	0.285±0.0359	0.263±0.0586	0.053±0.0202	0.104±0.0398	0.177±0.0794	0.047±0.0215	0.318±0.0581
	1386	0.146±0.0474	0.263±0.0554	0.269±0.0525	0.066±0.0330	0.115±0.0381	0.181±0.0774	0.062±0.0240	0.284±0.0408
	1848	0.130±0.0421	0.268±0.0420	0.256±0.0498	0.078±0.0229	0.121±0.0632	0.143±0.0743	0.061±0.0213	0.318±0.0481
	2310	0.120±0.0460	0.257±0.0400	0.228±0.0446	0.069±0.0122	0.102±0.0314	0.119±0.0447	0.046±0.0165	0.287±0.0384
10	1287	0.215±0.0873	0.308±0.0815	0.225±0.0727	0.069±0.0241	0.068±0.0239	0.219±0.0745	0.046±0.0208	0.493±0.0584
	2574	0.188±0.0954	0.294±0.0655	0.233±0.0622	0.075±0.0186	0.071±0.0188	0.194±0.0725	0.042±0.0126	0.492±0.0299
	3861	0.191±0.0446	0.303±0.0444	0.249±0.0593	0.087±0.0256	0.083±0.0176	0.193±0.0597	0.034±0.0089	0.520±0.0376
	5148	0.228±0.1048	0.298±0.0506	0.223±0.0411	0.078±0.0153	0.068±0.0181	0.164±0.0451	0.041±0.0089	0.500±0.0320
	6435	0.213±0.0810	0.295±0.0639	0.218±0.0477	0.082±0.0191	0.083±0.0257	0.173±0.0759	0.042±0.0084	0.492±0.0345
11	297	0.253±0.0686	0.307±0.0544	0.225±0.0768	0.060±0.0302	0.162±0.0770	0.232±0.0959	0.054±0.0324	0.542±0.0945
	594	0.239±0.0507	0.289±0.0627	0.216±0.0325	0.076±0.0252	0.165±0.0642	0.263±0.0550	0.032±0.0243	0.490±0.0684
	891	0.305±0.0572	0.317±0.0442	0.220±0.0434	0.079±0.0287	0.162±0.0593	0.280±0.0680	0.048±0.0145	0.509±0.0608
	1188	0.257±0.0616	0.320±0.0418	0.219±0.0620	0.084±0.0327	0.141±0.0525	0.316±0.0561	0.064±0.0271	0.482±0.0369
	1484	0.314±0.0253	0.284±0.0333	0.214±0.0387	0.071±0.0258	0.115±0.0383	0.291±0.0373	0.046±0.0172	0.485±0.0573
12	2199	0.094±0.0203	0.143±0.0244	0.101±0.0266	0.054±0.0140	0.141±0.0696	0.100±0.0289	0.048±0.0116	0.309±0.0349
	4398	0.102±0.0260	0.152±0.0238	0.094±0.0157	0.050±0.0139	0.103±0.0479	0.101±0.0112	0.038±0.0083	0.301±0.0397
	6597	0.091±0.0189	0.131±0.0138	0.086±0.0131	0.051±0.0079	0.134±0.0982	0.110±0.0228	0.043±0.0116	0.294±0.0239
	8796	0.105±0.0319	0.137±0.0144	0.091±0.0189	0.048±0.0079	0.128±0.0955	0.098±0.0239	0.040±0.0076	0.302±0.0212
	10992	0.089±0.0171	0.132±0.0142	0.081±0.0145	0.052±0.0086	0.123±0.0631	0.106±0.0141	0.044±0.0098	0.300±0.0223
13	1124	0.060±0.0225	0.115±0.0309	0.072±0.0293	0.047±0.0147	0.213±0.0600	0.079±0.0393	0.031±0.0178	0.302±0.0348
	2248	0.060±0.0105	0.101±0.0245	0.061±0.0279	0.036±0.0134	0.215±0.0447	0.053±0.0111	0.037±0.0161	0.284±0.0294
	3372	0.059±0.0158	0.092±0.0127	0.059±0.0188	0.040±0.0123	0.135±0.0448	0.062±0.0157	0.029±0.0105	0.292±0.0277
	4496	0.056±0.0107	0.099±0.0122	0.056±0.0145	0.040±0.0102	0.173±0.0300	0.060±0.0123	0.026±0.0086	0.297±0.0277
	5620	0.058±0.0152	0.102±0.0154	0.059±0.0138	0.038±0.0059	0.165±0.0549	0.062±0.0134	0.031±0.0076	0.286±0.0166
14	4000	0.203±0.0233	0.257±0.0254	0.221±0.0187	0.060±0.0141	0.183±0.0418	0.193±0.0266	0.087±0.0170	0.468±0.0317
	8000	0.213±0.0181	0.240±0.0143	0.206±0.0204	0.062±0.0105	0.193±0.0553	0.197±0.0166	0.074±0.0144	0.471±0.0270
	12000	0.206±0.0135	0.241±0.0206	0.206±0.0207	0.075±0.0084	0.172±0.0191	0.212±0.0192	0.075±0.0096	0.477±0.0124
	16000	0.213±0.0113	0.249±0.0107	0.202±0.0135	0.071±0.0085	0.181±0.0316	0.213±0.0186	0.069±0.0085	0.480±0.0157
	20000	0.202±0.0152	0.242±0.0171	0.199±0.0132	0.071±0.0101	0.169±0.0459	0.199±0.0152	0.069±0.0101	0.470±0.0123

which corresponds to a more open and widely spread grid map over the dataset.

WeVoS-Beta-SIM also obtains the best CE results providing final maps with clearly defined groups, again when complexity of the datasets is high.

At the same time, the TE results of WeVoS-Beta-SIM are better than those obtained by the simple model Beta-SIM, providing to the final map a better topology preservation. Therefore, it can be concluded that the WeVoS-Beta-SIM is able to provide the best visual representation of the internal structure of datasets when their complexity is high (i.e. in this research means more than 5 classes).

Finally, in terms of stability, WeVoS-Beta-SIM is less sensitive to noise in terms of TE than the simple model Beta-SIM, which lead to obtain maps with less distortion effect which in turn provides a better visual representation of the internal dataset structure. However, CE and MQE results do not show a clear tendency when instability is added.

VII. CONCLUSIONS AND FUTURE WORK

In this research, a novel topology-preserving model known as WeVoS-Beta-SIM has been presented, analyzed and compared with other well-known topology preserving models. This novel algorithm aims to obtain the best topology preserving summary as possible in order to improve the visual representation of high dimensional datasets and to increase the stability of the original model (Beta-SIM).

Therefore, the use of ensemble WeVoS when applying to the Beta-SIM algorithm improves the visual representation of the internal structure of high complex datasets (in this research means more than 5 classes), generating grid maps widely spread and that covers the input space better than the other models (better MQE values). At the same time, WeVoS-Beta-SIM obtains maps with less twists and folds than the simple model Beta-SIM algorithm (better TE values), which signifies a better topology of the map.

As can be seen in the results, the improvement on visual representation, created by WeVoS-Beta-SIM, is notably higher when complexity of the datasets increases. With very simple datasets, it only makes slight improvements or can even obtain worse results. That said, its usefulness has been proven in the case of more complex datasets (more than 5 classes), where the extra complexity of the calculation of the ensemble leads to the considerable increase of performance, obtaining a better organization and visualization of the presented information.

Results also show that WeVoS-Beta-SIM is less sensitive to noise in terms of TE than the simple model Beta-SIM. It leads to maps with less distortion which in turn provides a better visual representation of the internal dataset structure.

All the previous improvements made WeVoS-Beta-SIM a powerful new tool for the data mining community and should take its place among existing Topology Preserving Maps.

Future work will be focused on using WeVoS-Beta-SIM to analyze challenging real datasets to solve problems in the fields of big data, electric vehicles, energy efficiency,

cybersecurity, etc. Also, the use and comparison with other ensembles will be explored for all algorithms tested in this research.

APPENDIX

See Tables 2–8.

REFERENCES

- [1] S. Marcos-Pablos and F. J. García-Peñalvo, "Information retrieval methodology for aiding scientific database search," *Soft Comput.*, vol. 24, no. 8, pp. 5551–5560, Apr. 2020, doi: [10.1007/s00500-018-3568-0](https://doi.org/10.1007/s00500-018-3568-0).
- [2] G. G. Sundarkumar and V. Ravi, "A novel hybrid undersampling method for mining unbalanced datasets in banking and insurance," *Eng. Appl. Artif. Intell.*, vol. 37, pp. 368–377, Jan. 2015, doi: [10.1016/j.engappai.2014.09.019](https://doi.org/10.1016/j.engappai.2014.09.019).
- [3] D. T. Larose, *Discovering Knowledge in Data: An Introduction to Data Mining*. Hoboken, NJ, USA: Wiley, 2014.
- [4] Y. Cheng, K. Chen, H. Sun, Y. Zhang, and F. Tao, "Data and knowledge mining with big data towards smart production," *J. Ind. Inf. Integr.*, vol. 9, pp. 1–13, Mar. 2018, doi: [10.1016/j.jii.2017.08.001](https://doi.org/10.1016/j.jii.2017.08.001).
- [5] X. Wu, X. Zhu, G.-Q. Wu, and W. Ding, "Data mining with big data," *IEEE Trans. Knowl. Data Eng.*, vol. 26, no. 1, pp. 97–107, Jan. 2014, doi: [10.1109/TKDE.2013.109](https://doi.org/10.1109/TKDE.2013.109).
- [6] A. Oussous, F.-Z. Benjelloun, A. Ait Lahcen, and S. Belfkih, "Big data technologies: A survey," *J. King Saud Univ. Comput. Inf. Sci.*, vol. 30, no. 4, pp. 431–448, Oct. 2018, doi: [10.1016/j.jksuci.2017.06.001](https://doi.org/10.1016/j.jksuci.2017.06.001).
- [7] B. B. Gupta, D. P. Agrawal, S. Yamaguchi, and M. Sheng, "Advances in applying soft computing techniques for big data and cloud computing," *Soft Comput.*, vol. 22, no. 23, pp. 7679–7683, Dec. 2018, doi: [10.1007/s00500-018-3575-1](https://doi.org/10.1007/s00500-018-3575-1).
- [8] J. A. Saez and E. Corchado, "KSUFS: A novel unsupervised feature selection method based on statistical tests for standard and big data problems," *IEEE Access*, vol. 7, pp. 99754–99770, 2019, doi: [10.1109/ACCESS.2019.2930355](https://doi.org/10.1109/ACCESS.2019.2930355).
- [9] R. V. Vega, H. Quintián, J. L. Calvo-Rolle, Á. Herrero, and E. Corchado, "Gaining deep knowledge of Android malware families through dimensionality reduction techniques," *Log. J. IGPL*, vol. 27, no. 2, pp. 160–176, Mar. 2019, doi: [10.1093/jigpal/jzy030](https://doi.org/10.1093/jigpal/jzy030).
- [10] Á. Herrero, L. Sáiz-Bárcena, M. A. Manzanedo, and E. Corchado, "A hybrid proposal for cross-sectoral analysis of knowledge management," *Soft Comput.*, vol. 20, no. 11, pp. 4271–4285, Nov. 2016, doi: [10.1007/s00500-016-2293-9](https://doi.org/10.1007/s00500-016-2293-9).
- [11] B. Baruque and E. Corchado, "WeVoS scale invariant map," *Inf. Sci.*, vol. 280, pp. 307–321, Oct. 2014, doi: [10.1016/j.ins.2014.05.005](https://doi.org/10.1016/j.ins.2014.05.005).
- [12] R. V. Vega, H. Quintián, C. Cambra, N. Basurto, Á. Herrero, and J. L. Calvo-Rolle, "Delving into Android malware families with a novel neural projection method," *Complexity*, vol. 2019, pp. 1–10, Jun. 2019, doi: [10.1155/2019/6101697](https://doi.org/10.1155/2019/6101697).
- [13] C. S. Wickramasinghe, K. Amarasinghe, D. Marino, and M. Manic, "Deep self-organizing maps for visual data mining," in *Proc. 11th Int. Conf. Human Syst. Interact. (HSI)*, Jul. 2018, pp. 304–310, doi: [10.1109/HSI.2018.8430845](https://doi.org/10.1109/HSI.2018.8430845).
- [14] H. Quintián and E. Corchado, "Beta scale invariant map," *Eng. Appl. Artif. Intell.*, vol. 59, pp. 218–235, Mar. 2017, doi: [10.1016/j.engappai.2017.01.002](https://doi.org/10.1016/j.engappai.2017.01.002).
- [15] B. Baruque, E. Corchado, and H. Yin, "THE S2-ensemble fusion algorithm," *Int. J. Neural Syst.*, vol. 21, no. 06, pp. 505–525, Dec. 2011, doi: [10.1142/S0129065711003012](https://doi.org/10.1142/S0129065711003012).
- [16] E. Corchado and B. Baruque, "WeVoS-ViSOM: An ensemble summarization algorithm for enhanced data visualization," *Neurocomputing*, vol. 75, no. 1, pp. 171–184, Jan. 2012, doi: [10.1016/j.neucom.2011.01.027](https://doi.org/10.1016/j.neucom.2011.01.027).
- [17] E. Corchado and F. Colin, "The scale invariant map and maximum likelihood Hebbian learning," presented at the Int. Conf. Knowl.-Based Intell. Inf. Eng. Syst., 2002, vol. 82.
- [18] B. Fritzke, "A growing neural gas network learns topologies," in *Proc. Adv. Neural Inf. Process. Syst.*, 1995, pp. 625–632.
- [19] M. Zapater, D. Fraga, P. Malagon, Z. Bankovic, and J. M. Moya, "Self-organizing maps versus growing neural gas in detecting anomalies in data centres," *Log. J. IGPL*, vol. 23, no. 3, pp. 495–505, Jun. 2015, doi: [10.1093/jigpal/jzv008](https://doi.org/10.1093/jigpal/jzv008).

- [20] T. M. Martinez, S. G. Berkovich, and K. J. Schulten, "Neural-gas' network for vector quantization and its application to time-series prediction," *IEEE Trans. Neural Netw.*, vol. 4, no. 4, pp. 558–569, Jul. 1993.
- [21] B. Pes, N. Dessì, and M. Angioni, "Exploiting the ensemble paradigm for stable feature selection: A case study on high-dimensional genomic data," *Inf. Fusion*, vol. 35, pp. 132–147, May 2017, doi: [10.1016/j.inffus.2016.10.001](https://doi.org/10.1016/j.inffus.2016.10.001).
- [22] G. Haixiang, L. Yijing, L. Yanan, L. Xiao, and L. Jinling, "BPSO-AdaBoost-KNN ensemble learning algorithm for multi-class imbalanced data classification," *Eng. Appl. Artif. Intell.*, vol. 49, pp. 176–193, Mar. 2016, doi: [10.1016/j.engappai.2015.09.011](https://doi.org/10.1016/j.engappai.2015.09.011).
- [23] Y. Ren, L. Zhang, and P. N. Suganthan, "Ensemble classification and regression-recent developments, applications and future directions [review article]," *IEEE Comput. Intell. Mag.*, vol. 11, no. 1, pp. 41–53, Feb. 2016, doi: [10.1109/MCI.2015.2471235](https://doi.org/10.1109/MCI.2015.2471235).
- [24] L. Pietruczuk, L. Rutkowski, M. Jaworski, and P. Duda, "How to adjust an ensemble size in stream data mining?" *Inf. Sci.*, vol. 381, pp. 46–54, Mar. 2017, doi: [10.1016/j.ins.2016.10.028](https://doi.org/10.1016/j.ins.2016.10.028).
- [25] B. Baroque and E. Corchado, "A weighted voting summarization of SOM ensembles," *Data Mining Knowl. Discovery*, vol. 21, no. 3, pp. 398–426, Nov. 2010, doi: [10.1007/s10618-009-0160-3](https://doi.org/10.1007/s10618-009-0160-3).
- [26] M. Lichman, "UCI machine learning repository," Ph.D. dissertation, School Inf. Comput. Sci., Univ. California, Irvine, CA, USA, 2013.
- [27] T. Kohonen, "The self-organizing map," *Neurocomputing*, vol. 21, nos. 1–3, pp. 1–6, 1998, doi: [10.1016/S0925-2312\(98\)00030-7](https://doi.org/10.1016/S0925-2312(98)00030-7).
- [28] E. Mohebi and A. Bagirov, "Constrained self organizing maps for data clusters visualization," *Neural Process. Lett.*, vol. 43, no. 3, pp. 849–869, Jun. 2016, doi: [10.1007/s11063-015-9454-1](https://doi.org/10.1007/s11063-015-9454-1).
- [29] M. Vasighi and H. Amini, "A directed batch growing approach to enhance the topology preservation of self-organizing map," *Appl. Soft Comput.*, vol. 55, pp. 424–435, Jun. 2017, doi: [10.1016/j.asoc.2017.02.015](https://doi.org/10.1016/j.asoc.2017.02.015).
- [30] E. K. Haimoudi, H. Fakhouri, L. Cherrat, and M. Ezziyyani, "Towards a new approach to improve the classification accuracy of the Kohonen's self-organizing map during learning process," *Int. J. Adv. Comput. Sci. Appl.*, vol. 7, no. 3, pp. 230–236, 2016.
- [31] T. Kohonen, "Essentials of the self-organizing map," *Neural Netw.*, vol. 37, pp. 52–65, Jan. 2013, doi: [10.1016/j.neunet.2012.09.018](https://doi.org/10.1016/j.neunet.2012.09.018).
- [32] X.-S. Xu, X.-L. Liang, G.-Q. Yang, X.-L. Wang, S. Guo, and Y. Shi, "SOMH: A self-organizing map based topology preserving hashing method," *Neurocomputing*, vol. 236, pp. 56–64, May 2017, doi: [10.1016/j.neucom.2016.08.102](https://doi.org/10.1016/j.neucom.2016.08.102).
- [33] B. Baroque and E. Corchado, "A novel ensemble of scale-invariant feature maps," in *Computer Recognition Systems 3*, M. Kurzynski and M. Wozniak, Eds. Berlin, Germany: Springer, 2009, pp. 265–273.
- [34] L. Sánchez, H. Quintián, H. Pérez, and E. Corchado, "Optimization of MLHL-SIM and SIM algorithm using OpenMP," in *Proc. 17th Conf. Spanish Assoc. Artif. Intell. Adv. Artif. Intell. (CAEPIA)*, Salamanca, Spain, Sep. 2016, pp. 227–236, doi: [10.1007/978-3-319-44636-3_21](https://doi.org/10.1007/978-3-319-44636-3_21).
- [35] C. M. Bishop, M. Svensén, and C. K. I. Williams, "GTM: The generative topographic mapping," *Neural Comput.*, vol. 10, no. 1, pp. 215–234, Jan. 1998, doi: [10.1162/089976698300017953](https://doi.org/10.1162/089976698300017953).
- [36] M. Ghassany and Y. Bennani, "Collaborative fuzzy clustering of variational Bayesian generative topographic mapping," *Int. J. Comput. Intell. Appl.*, vol. 14, no. 1, Mar. 2015, Art. no. 1550001, doi: [10.1142/S1469026815500017](https://doi.org/10.1142/S1469026815500017).
- [37] D. Horvath, I. Baskin, G. Marcou, and A. Varnek, "Generative topographic mapping of conformational space," *Mol. Informat.*, vol. 36, no. 10, Oct. 2017, Art. no. 1700036, doi: [10.1002/minf.201700036](https://doi.org/10.1002/minf.201700036).
- [38] C. Fyfe, "The topographic product of experts," in *Proc. Artif. Neural Netw. Biol. Inspirations (ICANN)*, vol. 3696, Berlin, Germany, 2005, pp. 397–402.
- [39] Y. Jeong, K. Lee, B. Yoon, and R. Phaal, "Development of a patent roadmap through the generative topographic mapping and bass diffusion model," *J. Eng. Technol. Manage.*, vol. 38, pp. 53–70, Oct. 2015, doi: [10.1016/j.jengtecman.2015.08.006](https://doi.org/10.1016/j.jengtecman.2015.08.006).
- [40] W. Huang and H. Yin, "ViSOM for dimensionality reduction in face recognition," in *Proc. 7th Int. Workshop Adv. Self-Organizing Maps (WSOM)*, St. Augustine, FL, USA, Jun. 2009, pp. 107–115, doi: [10.1007/978-3-642-02397-2_13](https://doi.org/10.1007/978-3-642-02397-2_13).
- [41] W. Huang and H. Yin, "On nonlinear dimensionality reduction for face recognition," *Image Vis. Comput.*, vol. 30, nos. 4–5, pp. 355–366, May 2012, doi: [10.1016/j.imavis.2012.03.004](https://doi.org/10.1016/j.imavis.2012.03.004).
- [42] B. Baroque and E. Corchado, *Fusion Methods for Unsupervised Learning Ensembles*, vol. 322. Berlin, Germany: Springer, 2011.
- [43] C. Fyfe, *Hebbian Learning and Negative Feedback Networks*. Berlin, Germany: Springer, 2005.
- [44] C. Fyfe, "A neural network for PCA and beyond," *Neural Process. Lett.*, vol. 6, nos. 1–2, pp. 33–41, 1997, doi: [10.1023/A:1009606706736](https://doi.org/10.1023/A:1009606706736).
- [45] D. Tripathi, D. R. Edla, and R. Cheruku, "Hybrid credit scoring model using neighborhood rough set and multi-layer ensemble classification," *J. Intell. Fuzzy Syst.*, vol. 34, no. 3, pp. 1543–1549, Mar. 2018, doi: [10.3233/JIFS-169449](https://doi.org/10.3233/JIFS-169449).
- [46] D. R. Edla, D. Tripathi, R. Cheruku, and V. Kuppli, "An efficient multi-layer ensemble framework with BPSOGSA-based feature selection for credit scoring data analysis," *Arabian J. Sci. Eng.*, vol. 43, no. 12, pp. 6909–6928, Dec. 2018, doi: [10.1007/s13369-017-2905-4](https://doi.org/10.1007/s13369-017-2905-4).
- [47] D. Tripathi, D. R. Edla, R. Cheruku, and V. Kuppli, "A novel hybrid credit scoring model based on ensemble feature selection and multilayer ensemble classification," *Comput. Intell.*, vol. 35, no. 2, pp. 371–394, May 2019, doi: [10.1111/coin.12200](https://doi.org/10.1111/coin.12200).
- [48] L. A. Pasa, J. A. F. Costa, and M. G. D. Medeiros, "An ensemble algorithm for Kohonen self-organizing map with different sizes," *Log. J. IGPL*, vol. 25, no. 6, pp. 1020–1033, Dec. 2017, doi: [10.1093/jigpal/jzx046](https://doi.org/10.1093/jigpal/jzx046).
- [49] M. Rigamonti, P. Baraldi, A. Alessi, E. Zio, D. Astigarraga, and A. Galarza, "An ensemble of component-based and population-based self-organizing maps for the identification of the degradation state of insulated-gate bipolar transistors," *IEEE Trans. Rel.*, vol. 67, no. 3, pp. 1304–1313, Sep. 2018, doi: [10.1109/TR.2018.2834828](https://doi.org/10.1109/TR.2018.2834828).
- [50] F. Saitoh, "An ensemble model of self-organizing maps for imputation of missing values," in *Proc. IEEE 9th Int. Workshop Comput. Intell. Appl. (IWCIA)*, Nov. 2016, pp. 9–14, doi: [10.1109/IWCIA.2016.7805741](https://doi.org/10.1109/IWCIA.2016.7805741).
- [51] R. Salas, C. Saavedra, H. Allende, and C. Moraga, "Machine fusion to enhance the topology preservation of vector quantization artificial neural networks," *Pattern Recognit. Lett.*, vol. 32, no. 7, pp. 962–972, May 2011, doi: [10.1016/j.patrec.2011.01.020](https://doi.org/10.1016/j.patrec.2011.01.020).
- [52] H. Quintián and E. Corchado, "Beta Hebbian learning as a new method for exploratory projection pursuit," *Int. J. Neural Syst.*, vol. 27, no. 6, Sep. 2017, Art. no. 1750024.
- [53] H. Parvin, M. MirnabiBaboli, and H. Alinejad-Rokny, "Proposing a classifier ensemble framework based on classifier selection and decision tree," *Eng. Appl. Artif. Intell.*, vol. 37, pp. 34–42, Jan. 2015, doi: [10.1016/j.engappai.2014.08.005](https://doi.org/10.1016/j.engappai.2014.08.005).
- [54] M. Chattopadhyay, P. K. Dan, and S. Mazumdar, "Principal component analysis and self-organizing map for visual clustering of machine-part cell formation in cellular manufacturing system," *Syst. Res. Forum*, vol. 5, no. 1, pp. 25–51, Jun. 2011, doi: [10.1142/S179396661100028X](https://doi.org/10.1142/S179396661100028X).
- [55] B. Wang and J. Pineau, "Online bagging and boosting for imbalanced data streams," *IEEE Trans. Knowl. Data Eng.*, vol. 28, no. 12, pp. 3353–3366, Dec. 2016.



HÉCTOR QUINTIÁN (Member, IEEE) received the M.S. degree in computer science from the University of A Coruña, Spain, in 2010, and the Ph.D. degree from the University of Salamanca in 2017.

He is an Assistant Professor with the University of A Coruña. He has published over 30 peer-reviewed articles, 21 conference papers, and 18 conference editorials. He has patented several software models. He has been the co-organizer, the program committee chair, and the session chair of several international conferences. His research interests include neural networks, with a particular focus on exploratory projection pursuit, maximum likelihood Hebbian learning, self-organizing maps, multiple classifier and its applications to control engineering, optimization, and education.

Dr. Quintián is the Vice-Chair of the IEEE-SMC Spanish Chapter. He has actively contributed to several current projects in the EU, including IT4Innovation, ICT Action COST IC1303, and IntelliCIS, and several national and regional projects.



EMILIO CORCHADO (Member, IEEE) received the Ph.D. degree in computer science from the University of Salamanca, Spain.

He is an RTD Expert hired by international organizations, such as the European Commission, the Grant Agency of Czech Republic, and the Spanish National Agency for Assessment and Forecasting. He has collaborated with SMEs and new companies in the innovation field in about 40 projects. He has patented software models, and

he owns the IP of more than ten ICT tools and models. He is currently a Full Professor in computer and automatic science with the University of Salamanca. He has published over 100 peer-reviewed articles in a range of topics from knowledge management and risk analysis, intrusion detection systems, food industry, artificial vision, and the modeling of industrial

processes. His research interests include neural networks, with a particular focus on exploratory projection pursuit, maximum likelihood Hebbian learning, self-organizing maps, multiple classifier systems, and hybrid systems.

Dr. Corchado has been the Organizing Chair, the Program Committee Chair, the Session Chair, and the General Chair of a number of conferences, such as the International Conference on Hybrid Artificial Intelligence Systems (HAIS), the International Conference on Intelligent Data Engineering and Automated Learning (IDEAL), and the International Conference on Knowledge-Based Intelligent Information and Engineering Systems (KES). He was the Chair of the IEEE Spanish Section, from 2014 to 2015, and he has actively contributed to several current projects in the EU, including SOFTCOMP, IT4Innovation, ICT Action COST IC1303, and IntelliCIS NISIS.

• • •