July 18, 2023

Professor Plamen Angelov
Editor in Chief
Evolving Systems

Dear Prof. Angelov,

We thank you and Guest Editors very much for assisting us during the review process, as well as for the guidance to comply the original manuscript contents with the reviewers' reports. As a result, the revised version of the manuscript titled *Evolving Fuzzy Predictor with Multivariable Gaussian Participatory Learning and Multi-Innovations Recursive Weighted Least Squares - eFMI* has been resubmitted for review.

We detail below descriptions of the changes made in the original manuscript and responses to the reviewers' comments, suggestions, and concerns. Hopefully, the current version of the manuscript answers the issues raised by the referees properly.

Looking forward to hear from you.

Sincerely yours,

Responses to the reviewers

First of all, we would like to thank the reviewers for the comments and suggestions made in their reports. The constructive comments have certainly contributed to improving the paper content, and they are gratefully acknowledged.

In order to make it easier to identify the changes in the manuscript after revision, we attach the revised paper with highlights in blue showing the changes in the text in response to referee comments.

Reviewer #1: Dear authors, I found your paper interesting, well organized and easy to read. Please find my comments in pdf document attached.

We would like to thank the Reviewer for their valuable comments.

1) You should define the meaning of the short name of the proposed model eFMI. This should be also corrected in the rest of the paper.

Dear reviewer, we defined a short name (eFMI) for the evolving Fuzzy with Multivariable Gaussian Participatory Learning and Multi-Innovations Recursive Weighted Least Squares in the abstract and the last paragraph of the introduction.

2) Are there any problems when you are merging a regular cluster with a micro-cluster?

Dear Reviewer, this is a good point. However, there is no problem merging a regular cluster with a micro-cluster. The merging of a regular cluster and a micro-cluster generates a cluster whose dispersion matrix is calculated using the Mahalanobis distance. Here, we use Mahalanobis distance because the new cluster's number of samples is obtained by the sum of samples of the two merged clusters. To clarify, we included the text below at the end of Section 3.3 and updated Algorithm 2.

The new cluster's number of samples is obtained by the sum of the number of samples of the two merged clusters, i.e., $n^t_{i^* \cup i} = n^t_i + n^t_{i^*}$.

3) I am not sure but I think that zero could lead to some singularities when calculating the inverse covariance matrix. Therefore, maybe a better value is 2*m, where m is the dimensionality of the problem, or length of the vector x.

Dear reviewer, thanks for the comments, and we agree with you. Therefore, we inserted the suggestion in Section 3.6 and reproduced it below.

Suggests $N_{max} \geq 2 * m$, with m the number of input variables, thus avoiding the singularity problem by calculating the inverse covariance matrix using the Mahalanobis distance. Using $N_{max} = 0$, the model will only calculate the distance of Mahalanobis.

4) I suggest to make a short comparison (in Related literature section or here) between

these algorithms. Particulary it is important to see the difference between them, e.g. compare the evolving mechanisms (adding, merging, deleting, etc.) and compare the learning principles of the consequent part (rLS, wRLS, etc.). It could be summarized in a table.

Dear Reviewer, thank you very much for your suggestion. We agree with you, and in the revised version, we include a table in Section 2 (Related Literature) in which the main features of the models are compared. We reproduce the text and the Table added below.

A short comparison between the evolving learning algorithms is shown in Table 1. This table presents the membership functions and the methods to add, merge, and delete rules and/or clusters used in the models. The approach to update the consequent parameters is also presented.

Comparison between the evolving models.

Models	**Membership functions**	**Add Rules and/or Clusters**	**Merge Rules and/or Clusters**	**Delete Rules and/or Clusters**	**Update Consequent**
ALMMo [46]	Data clouds	Density of the new sample	-	Utility measure	FWRLS
AutoCloud [48]	Data clouds	Eccentricity measure	Number of overlapping samples	-	RLS
eCauchy [20]	Cauchy	Distance between a sample and the cluster centers	Mahalanobis distance between clusted; correlation between fring rates; or information of the local models themselves (regression models)	-	RLS
eFGP [23]	Trapezoidal	Expansion region of a granule	Neighbor granules	Inactivity of granules and rules	RLS
eGAUSS+ [28]	Gaussian	Typicalities of current sample.	Volume of clusters	-	RLS
eIX [45]	Trapezoidal	Grandulo interval	Distance between the centers of the beads	-	Balancing rate
eMG [29]	Gaussian	Compatibility measure of a current sample	Degree of compatibility	-	RWLS
eNFN [34]	Triangular	Mean value of local error and mean global error	-	Age of the rule	Gradient descent with optimal learning rate
eOGS [44]	Gaussian	Grandulo interval	Neighbor granules	Inactivity of granules and rules	RLS
eTS [6]	Gaussian	Potential of a current sample	-	-	RLS
eTS+ [38]	Gaussian	Potential of a current sample	-	Size of the support set of the cluster	FWRLS
FLEXFIS [40]	Gaussian	Distance between a sample and the cluster centers	-	-	WRLS
FLEXFIS+ [42]	Gaussian	Distance between a sample and the cluster centers	Index of overlapping	-	WRLS
Simpl_eTS [36]	Cauchy	Dispersion of a current sample	-	Population of each group	RLS
xTS [37]	Gaussian	Potential of a current sample	-	Age of the rule	RLS

5) As I see from the introduction section and the related works section, evolving systems tend to be autonomous and easy to use. Here we see that in order to obtain the best results of a given algorithm (eFMI, eMG, eNFN,...), we need to perform an exhaustive search for the best initial parameters. What do you say about this problem and are there other suggestions for initialising the parameters?

Dear reviewer, thanks for the comments. We chose the exhaustive search process to find the best parameters configuration for the models for each data set. Another option for initializing the parameters is using a training set base to define the parameters and test the validation set. However, in the literature, there is still a lack of definition in the construction of workflows to guide the development, performance evaluation, testing, validation, and comparison of algorithms in non-stationary environments, as discussed in [14].

14. I. Škrjanc, J. Iglesias, A. Sanchis, D. Leite, E. Lughofer, F. Gomide, Evolving fuzzy and neuro-fuzzy approaches in clustering, regression, identication, and classification: A survey, Information Sciences 490 (2019) 344-368. doi:10.1016/j.ins.2019.03.060.

6) Maybe it is better to plot the error as well, e.g. in a subplot. The same for Fig. 2 and 3.

Dear Reviewer, it was a good contribution to improving the quality of the paper. To fix this issue, we inserted a new error measure (RSS - Residual Sum of Squares) to verify the behavior of the residual modeling error over time. Figures 2, 4, 6, 8, 10, and 12 show the RSS of the models for the evaluated datasets.

Reviewer #2: This paper presents the modification of eMG. In my view, the contribution of this paper is rather limited. Please see my comments below:

Dear Reviewer, the authors would like to thank you for your time and comments.

1) I do not agree that the fuzzy rules are considered as an absolute advantage of evolving fuzzy system compared to evolving neural network. This is not necessarily true because of the following reasons: 1) most fuzzy rules get significantly overlapping after being adjusted; 2) fuzzy rules become incomprehensible if FNN is connected with feature extraction layer, e.g., convolutional layers; 3) most evolving fuzzy systems are based on TS fuzzy models, difficult to interpret.

Dear Reviewer, thank you very much for your comments. We agree with you. The statements about this issue in the second paragraph of Section 1 (Introduction) are strong, and in some scenarios, as mentioned, they are not applicable. Therefore, in the revised text, we removed the paragraph.

2) The following works are relevant SOTA in evolving fuzzy systems which should be reviewed in this paper.

- PALM: An Incremental Construction of Hyperplanes for Data Stream Regression. IEEE Trans. Fuzzy Syst. 27(11): 2115-2129 (2019)
- A novel Spatio-Temporal Fuzzy Inference System (SPATFIS) and its stability analysis. Inf. Sci. 505: 84-99 (2019)

In addition, the concept of evolving systems has been implemented in the deep learning framework as follows:

- Automatic Construction of Multi-layer Perceptron Network from Streaming Examples. CIKM 2019: 1171-1180

Dear Reviewer, we agree with you. Therefore, the references suggested by the reviewers that we believe are more suitable for the paper have been included in this revised version. The list of the references added is reproduced below.

11. M. M. Ferdaus, M. Pratama, S. G. Anavatti and M. A. Garratt, PALM: An Incremental Construction of Hyperplanes for Data Stream Regression, in IEEE Transactions on Fuzzy Systems, vol. 27, no. 11, Nov. 2019,pp. 2115-2129, doi: 10.1109/TFUZZ.2019.2893565.

25. X. Zhou, P. Angelov, Autonomous visual self-localization in completely unknown environment using evolving fuzzy rule-based classifier, in: 2007 IEEE Symposium on Computational Intelligence in Security and Defense Applications, 2007, pp. 131–138. doi:10.1109/CISDA.2007.368145.

26. S. Subhrajit, M. Pratama and S. Sundaram. A novel Spatio-Temporal Fuzzy Inference System (SPATFIS) and its stability analysis, in Inf. Sci. 505, 2019, pp 84-99, doi:10.1016/J.INS.2019.07.056.

30. P. Angelov, X. Zhou, On line learning fuzzy rule-based system structure from datastreams, in: Proceedings of the IEEE International Conference on Fuzzy Systems(IEEE World Congress on Computational Intelligence), IEEE, 2008, pp. 915–922. doi: 10.1109/FUZZY.2008.4630479.

3) I am not really clear the novelty and contribution of this paper. If fuzzy rule pruning and merging are claimed to be novelty, these approaches have been proposed in the literature for a long time.

We agree and are aware that the pruning and merging methods are not new and are used in several models. This is supported by the literature review presented in Section 2. However, we use these methods in a new model based on multivariate Gaussian participatory learning in our approach. The novelties and the main contributions are listed in Section 1 and reproduced below.

- A new clustering algorithm with participatory learning, multivariable Gaussian membership functions, and two distances in cluster estimation. Using the two distances avoids the singularity problem in calculating the inverse of the cluster's dispersion matrix when the cluster has a small number of samples.
- Mechanism to exclude old and non-representative rules. The procedure is based on the inactivity time of the clusters and the number of samples associated with the clusters. The exclusion of rules relates to the model's ability to forget old knowledge whenever it becomes useless. Besides that, it reduces the computational cost of the algorithm.
- Mechanism to merge similar clusters based on the overlap of the two clusters. Cluster merging aims to improve the model's accuracy and reduces the computational cost.
- A new evolving fuzzy algorithm focused on solving general regression problems with good accuracy;
- The approach was proposed for the regression problems but can be modified for use in classification tasks.

4) problem description section should be added.

Dear Reviewer, thank you very much for your comments. It was an excellent contribution to improving the paper, and as suggested, we added a paragraph with the problem description in Section 1 (Introduction). The new paragraph is reproduced below.

The time series forecasting and nonlinear systems identifications problems are challenging tasks [26] and can be described by $\hat{y}^t = f(y^{t-1}, y^{t-2}, ..., y^{t-n})$, in which the prediction in the current time t is performed using the past lags of the series and a function f suitable for the previous observations [27]. In a online context, the dataflow can be expressed as $S = y^1, y^2, ..., y^t, ..., y^\infty$, where y is sequence of data [11]. Generally, the data streams have the following caracteristics:

- the data stream has unlimited size, and the samples arrive online, and occasionally, in real-time [14];
- the samples are processed only once and then discarded, i.e., need learning with a single pass in the data [11];
- the system should quickly deal with situations of variations in environmental system conditions, such as system drifts/shifts or non-stationary environments [26];
- the large amount of data to be processed [28].

In this context, this paper suggests a new evolving fuzzy model to regression tasks, such as forecasting and system identification. The proposed approach is ...

11. M. M. Ferdaus, M. Pratama, S. G. Anavatti and M. A. Garratt, PALM: An Incremental Construction of Hyperplanes for Data Stream Regression, in IEEE Transactions on Fuzzy Systems, vol. 27, no. 11, Nov. 2019,pp. 2115-2129, doi: 10.1109/TFUZZ. 2019.2893565.

14. I. Škrjanc, J. Iglesias, A. Sanchis, D. Leite, E. Lughofer, F. Gomide, Evolving fuzzy and neuro-fuzzy approaches in clustering, regression, identication, and classification: A survey, Information Sciences 490 (2019) 344-368. doi:10.1016/j.ins.2019.03.060.

26. S. Subhrajit, M. Pratama and S. Sundaram. A novel Spatio-Temporal Fuzzy Inference System (SPATFIS) and its stability analysis, in Inf. Sci. 505, 2019, pp 84-99, doi:10.1016/J.INS.2019.07.056.

27. H. L utkepohl, New Introduction to Multiple Time Series Analysis, Springer ScienceBusiness Media, 2005.

28. I.Škrjanc, Cluster-volume based merging concept for incrementally evolving fuzzy Gaussian clustering - eGAUSS+, IEEE Transactions on Fuzzy Systems 28 (9) (2020) 2222–2231. doi:10.1109/TFUZZ.2019.2931874.

5a) please be direct to your contribution in describing the methodologies.

Dear reviewer, thanks for the comments. The novelties and the main contributions can be found in Section 1 Introduction (after the brief description of the proposed approach). Several works in literature also do this, such as the references used in the paper and listed below:

11. M. M. Ferdaus, M. Pratama, S. G. Anavatti and M. A. Garratt, PALM: An Incremental Construction of Hyperplanes for Data Stream Regression, in IEEE Transactions on Fuzzy Systems, vol. 27, no. 11, Nov. 2019,pp. 2115-2129, doi: 10.1109/TFUZZ. 2019.2893565.

25. X. Zhou, P. Angelov, Autonomous visual self-localization in completely unknown environment using evolving fuzzy rule-based classifier, in: 2007 IEEE Symposium on Computational Intelligence in Security and Defense Applications, 2007, pp. 131–138. doi:10.1109/CISDA.2007.368145.

26. S. Subhrajit, M. Pratama and S. Sundaram. A novel Spatio-Temporal Fuzzy Inference System (SPATFIS) and its stability analysis, in Inf. Sci. 505, 2019, pp 84-99, doi:10.1016/J.INS.2019.07.056

5b) the same parts as [27] and [40] do not need to be re-described again.

Dear reviewer, thanks for the comments. We re-described the same parts as [27] and [40] to make the text easier to read and self-contained.

6) rule merging and rule pruning based on age are not new at al.

Dear Reviewer, thank you for the comments. We agree that using the age or inactivity concept to eliminate or merge clusters is not new; how we can see in Section 2 and more clearly in Table 1, this method is used in several models. Besides that, it proves to be an efficient technique, as it can detect changes and variations in the data. In our approach, the concepts of age and population used together performed well in excluding rules.

7) what happens to section 3.5 when rules are added or removed?

Dear Reviewer, thank you for the comments. When rules are added, we show in Section 3.2 that the consequent parameters $h^t_{c^t}$ *are obtained by the weighted average method of the parameters of the active clusters by equation (15), reproduced below.*

$$h^t_{c^t} = \frac{\sum_{i=1}^{c^t} h^t_i \gamma^t_i}{\sum_{i=1}^{c^t} \gamma^t_i} \qquad (15).$$

Whenever rules are merged (Section 3.3), the consequent parameters of the newly generated cluster are obtained by equation (22), reproduced below.

$$h^t_{i^* \cup i} = \frac{h^t_{i^*}\gamma^t_{i^*} + h^t_i \gamma^t_i}{\gamma^t_{i^*} + \gamma^t_i} \qquad (22).$$

On the other hand, the consequent parameters are also removed when a cluster/rule is *removed (Section 3.4).*

8) I doubt the current hyper-parameter selection can be acceptable in practise. Authors need to set one set of parameters and use it for every problem in this paper.

Dear reviewer, this is a good point. We believe that a continuous challenge is the searching for algorithms with good accuracy, adaptability, autonomy, and without user-defined parameters. However, in this paper, as well as in the works of the models used as a benchmark (eMG [29], eNFN [34], eOGS [44], xTs [37], and eTS [6]), we chose to find the best set of model parameters for each dataset, allowing that all models present their best results and fair comparison between the models.

6. P. Angelov, Evolving rule-based models: a tool for design of flexible adaptive systems,Springer-Verlag Berlin Heidelberg, 2002. doi:10.1007/978-3-7908-1794-2

29. A. Lemos, W. Caminhas, F. Gomide, Multivariable Gaussian Evolving FuzzyModeling System, IEEE Transactions on Fuzzy Systems 19 (1) (2011) 91–104.doi: 10.1109/TFUZZ.2010.2087381

34. A. Silva, W. Caminhas, A. Lemos, F. Gomide, A fast learning algorithmfor evolving neo-fuzzy neuron, Applied Soft Computing 14 (2014) 194–209.doi: 10.1016/j.asoc.2013.03.022

37. P. Angelov, X. Zhou, Evolving Fuzzy Systems from Data Streams in Real-Time, in:Proceedings of the International Symposium on Evolving Fuzzy Systems, 2006, pp.29–35. doi: 10.1109/ISEFS.2006.251157.

44. D. Leite, I. Skrjanc, Ensemble of Evolving Optimal Granular Experts, OWA Aggregation, and Time Series Prediction, Information Sciences 504 (2019) 95–112.doi: 10.1016/j.ins.2019.07.053.

9) please produce numerical results based on batch-wise prequential test-then-train protocol. This is necessary to evaluate the performance of stream-mining algorithms.

Dear Reviewer, thanks for the comments. It is important to emphasize that, as mentioned in [13], there is still a lack of standardization in literature in the construction of workflows to guide the development, performance evaluation, testing, validation, and comparison of algorithms in non-stationary environments. Because of that, in this work, we implemented the traditional methodology widely used in the literature. However, we agree with you. The batch-wise prequential test-then-train is an excellent protocol to evaluate the performance of stream-mining algorithms. So, have included in future works the suggestion of carrying out new experiments based on the batch-wise prequential test-then-train, as suggested by the Reviewer. The text is reproduced below.

Future work should address techniques for automatically selecting the eFMI parameters, making them more autonomous. The investigation of mechanisms to reduce the complexity of the algorithm is also relevant for future work. Performing the experiments using other methods, for example, with batch-wise prequential test-then-train [11, 26]. It also highlights the importance of study new algorithms to adjust the consequent parameters.

11. M. M. Ferdaus, M. Pratama, S. G. Anavatti and M. A. Garratt, PALM: An Incremental Construction of Hyperplanes for Data Stream Regression, in IEEE Transactions on Fuzzy Systems, vol. 27, no. 11, Nov. 2019,pp. 2115-2129, doi: 10.1109/TFUZZ.2019.2893565.

26. S. Subhrajit, M. Pratama and S. Sundaram. A novel Spatio-Temporal Fuzzy Inference System (SPATFIS) and its stability analysis, in Inf. Sci. 505, 2019, pp 84-99, doi:10.1016/J.INS.2019.07.056

10) the source codes of this paper has to be made publicly available to allow convenient further studies.

Dear Reviewer, thank you for the comments. We fully agree with you, and as suggested, we make the eFMI source code available. The text included in the revised version is reproduced below.

eFMI was developed in Matlab, and to contribute to further studies, its source code is available at the URL: shorturl.at/cjvD3.

Reviewer #3: The paper "Evolving Fuzzy Predictor with Multivariable Gaussian Participatory Learning and MultiInnovations Recursive Weighted Least Squares - eFMI" proposes a new evolving model based on an unsupervised recursive clustering algorithm. The proposed method uses participatory learning and multivariable Gaussian membership functions as core mechanisms. The paper is well written, and the author applies the proposed methods to different problems with acceptable results.

Dear Reviewer, the authors would like to thank you for your time and comments.

1) The paper can benefit by more relevant references, e.g.

- https://ieeexplore.ieee.org/stamp/stamp.jsp?arnumber=4219092
- https://ieeexplore.ieee.org/stamp/stamp.jsp?arnumber=4630479
- https://onlinelibrary.wiley.com/doi/full/10.1002/int.2189

Dear Reviewer, we agree with you. Therefore, the references suggested by the reviewers that we believe are more suitable for the paper have been included in this revised version. The list of the references added is reproduced below.

11. M. M. Ferdaus, M. Pratama, S. G. Anavatti and M. A. Garratt, PALM: An Incremental Construction of Hyperplanes for Data Stream Regression, in IEEE Transactions on Fuzzy Systems, vol. 27, no. 11, Nov. 2019,pp. 2115-2129, doi: 10.1109/TFUZZ.2019.2893565.

25. X. Zhou, P. Angelov, Autonomous visual self-localization in completely unknown environment using evolving fuzzy rule-based classifier, in: 2007 IEEE Symposium on Computational Intelligence in Security and Defense Applications, 2007, pp. 131–138. doi:10.1109/CISDA.2007.368145.

26. S. Subhrajit, M. Pratama and S. Sundaram. A novel Spatio-Temporal Fuzzy Inference System (SPATFIS) and its stability analysis, in Inf. Sci. 505, 2019, pp 84-99, doi:10.1016/J.INS.2019.07.056.

30. P. Angelov, X. Zhou, On line learning fuzzy rule-based system structure from datastreams, in: Proceedings of the IEEE International Conference on Fuzzy Systems(IEEE World Congress on Computational Intelligence), IEEE, 2008, pp. 915–922. doi: 10.1109/FUZZY.2008.4630479.