National Research University Higher School of Economics Faculty of Computer Science

Bachelor’s Programme in Data Science and Business Analytics

TERM PAPER

Chaotic Time Series Prediction

Prepared by the student of Group 191, in Year: 3, Kseniia Lysaniuk

Term Paper Supervisor:
Deputy Head, Professor, Vasilii Gromov

Moscow
2022

Abstract

The problem of predicting the value of time series demonstrating chaotic behavior is currently met with a variety of different methods. However, most of them either lack persistently low error rate, and the error grows with subsequent prediction or demonstrate poor results in terms of errors when it comes to long-term prediction. This research area is highly important as many real-life processes could be mathematically modeled and described with chaotic dynamic systems, hence reflected in chaotic time series. This term paper aims to apply new algorithms to increase the prediction quality. Specifically, it examines self-healing algorithms as a continuation of predictive clustering algorithm with the usage non-predictable points.

For each point which needs to be predicted the algorithm, suggested in the term paper calculates a range of possible prediction values. Then using the clustering algorithms, the unified predicted value is calculated from this set. If it is impossible due to one of the reasons provided in a paper the point is labeled as unpredictable. The quality of the prediction is defined both by the mean squared error on predictable points and number of unpredictable points left after the iteration of the algorithm.

The concept of self-healing algorithm suggests the rise in the quality of prediction by repeatedly going over the points subject to prediction. Each iteration the algorithm updates the values of the points until it reaches the convergence of prediction quality.

The conducted experiment explored the work of the algorithm on classic chaotic time series.

The results prove the convergency of error with growing number of iterations of the self-healing algorithm. The suggested method can also be successfully applied on real-world examples of the chaotic time series.

Key words

Chaotic dynamic systems, prediction of chaotic time series, multi-step time series prediction, non-predictable points, self-healing algorithm, predictive clustering.

Table of Contents

Introduction

The time series is the statistical data on the value of some parameter. Observations are usually sequential, and measurements are made at regular intervals.

The chaotic time series represent nonlinear dynamical systems exhibiting chaotic behavior. Namely it means that the system and hence the corresponding time series is exceptionally sensitive to small changes in initial conditions. Examples of chaotic systems include and not limited by atmospheric dynamics, turbulent fluid, and biological populations. The development of the topic extended its applications even to financial analysis and medicine.

The system, which is chaotic by the mathematical definition, could be fully predictable in the case when the initial values of process are completely known. But in the real-world cases the unknown initial values limit the prediction horizon, and in a short time the system becomes completely unpredictable[1].

Currently the research of chaotic time series is especially relevant in those cases when mathematical description of the studied process is practically impossible, but some characteristic observable quantity had been collected and recorded in a form of time series.

The aim of the current term paper is to rise the quality of prediction algorithm for chaotic time series. Specifically, the application of self-healing algorithm is suggested as an extension of predictive clustering algorithm with the usage non-predictable points[2].

The problem of the multi-step ahead (MSA) prediction is in growing rate of error. Most algorithms used for MSA suggest using the predicted values as regressors

[1] Hirsch M. W., Smale S., Devaney R. Differential Equations, Dynamical Systems, & An Introduction to Chaos : Second ed. – Boston : Academic Press, 2003. – 416 p.
[2] Gromov V. Chaotic Time Series Prediction: Run for the Horizon // Tools and Methods of Program Analysis. – N.Y. : Springer International Publishing, 2021. – pp. 29-43.

and use them for further prediction. Hence even the small errors in the first few predicted values lead to the exponential growth of error in the further calculations[3].

The dynamic systems possess an attractor. It is the set of values the dynamic system will tend to with time. Hence chaotic time series possess repetitiveness in some sense. With the time going the discernible motifs (repeated with a small variations sequencies of values) of the behavior of the systems can be observed. With the training set great enough it is possible to extract these patterns and train the algorithm to recognize them and use for the prediction.

The point which is meant to be predicted can be matched with many motifs, and the problem is to choose the right prediction value. Many possible solutions could be suggested: statistical (the mean, the mode), as well as more effective[4] results of clustering algorithms (then it is the center of the greatest cluster). This work namely uses OPTICS, DBSCAN and Wishart clustering.

If the unified prediction value cannot be formed, either if the point does not match a motif or the problem is with the clustering, the point is said to be unpredictable.

The introduction of the self-healing algorithm which respectively updates the prediction values leads to improvement of the quality of the algorithm.

The term paper is organized in the following way: in the literature overview section, the analysis of scientific works working with the prediction of chaotic time series is presented. Different approaches are analyzed and are compared with the stress made on long-term prediction of real-world series modeled with chaotic dynamic systems. Chapter 1 states the mathematical problem; Chapter 2 presents comprehensive description for the algorithm used in the term paper. Chapter 3 offers the results of numerical experiment conducted with the Lorenz series. The last section is the conclusion.

---

3 Casdagli M. Nonlinear prediction of chaotic time series // Physica D : Nonlinear Phenomena. – May 1989. – Vol. 35. – 335-356 pp.

4 Gromov V. Chaotic Time Series Prediction: Run for the Horizon // Tools and Methods of Program Analysis. – N.Y. : Springer International Publishing, 2021. – pp. 29-43.

Literature overview

Since Lorenz discovered the butterfly effect in a nonlinear system of ordinary differential equations in 1960, several approaches have been proposed for prediction and modeling of chaotic systems[5].

Many studies focus on manifestation of chaotic behavior in purely physical processes, such as river flows[6], but due to the high relevance of the problem, there are many related to this topic articles and sources in the scientific community.

However, many of the studies focus solely on short-term prediction, not considering multi-step ahead prediction and choosing generating only one or few values instead. For this type of prediction any statistical machine learning algorithm would perform well. Interesting instruments suggested in the papers on the topic of the short-term prediction include but not limited to nonlinear models[7] and nonlinear mapping using local approximation[8].

The topic of long-term prediction is covered in comparatively smaller number of papers. For the MSA nearest neighbor predictor[9], artificial networks[10] as well as classic statistical algorithms, such as ARIMA and nonlinear models as multi-layered perceptron and their combinations are used[11]. The most popular idea upon this topic is to repeatedly predict one or a small number of steps ahead and then make a prediction

---

[5] Lorenz E.N. Deterministic non-periodic flow // Journal of Atmospheric Sciences. – 1962. – Vol. 20. – pp. 130-141.
[6] Lisi F., Villi v. Chaotic forecasting of discharge time series : a case study // JAWRA. – April 2001. – Vol. 37. – pp. 271-279.
[7] Casdagli M. Nonlinear prediction of chaotic time series // Physica D : Nonlinear Phenomena. – May 1989. – Vol. 35. – pp. 335-356.
[8] Farmer J.D., Sidorowich J.J. Predicting chaotic time series // Physical Review Letter. – 24 August 1987. – Vol. 59. – pp. 845-848.
[9] Lisi F., Villi V. Chaotic forecasting of discharge time series : a case study // JAWRA. – April 2001. – Vol. 37. – pp. 271-279.
[10] Chandra R., Ong Y.-S., Goh C.-K. Co-evolutionary multi-task learning with predictive recurrence for multi-step chaotic time series prediction // Neurocomputing.
[11] Hirata T., Kuremoto T., Obayashi M., Mabu S., Kobayashi K. Time Series Prediction Using DBN and ARIMA // 2015 International Conference on Computer Application Technologies. – N.Y. : IEEE, 2015. – pp. 24-29.

based on the resulting predictive values. This leads to the increase of error on later predictions.

Problem statement

Consider a chaotic time series $\mathbf{X_t} = \{X_1, X_2\ldots\}$. We assume the series to be normalized, have values from 0 to 1. The goal then is to predict next *h* values given the values from *1* to *t-1*. The series $\mathbf{X_t}$ is consequently split on two not-overlapping set of values which make up the training set $\{X_1, X_2\ldots X_{t-1}\}$ and the testing set $\{X_t, X_{t+1}\ldots X_{t+h-1}\}$. The training set is known to the algorithm and is used for generating patterns. The testing set is not known to the algorithm, as it is the values it is to predict. As the data is not used for training, the good empirical assessment of the quality of the model is received.

Due to the nature of dynamic systems with the time passing they tend to the series attractor, providing repetitiveness of the process. The algorithm generates motives on the training set by choosing the points to include in it. They are not necessarily consequent, but for the sake of computational power it is rational to limit their size. There are two parameters of generated patterns – their size *n* (so the number of points in the pattern) and the maximum distance between the points of the pattern *m*. So, the (*n*, *m*) patterns generated on the training set would be a set of vectors of length *n* such that for each pattern $x_a = \{X_{a1}, X_{a2} \ldots X_{an}\}$ the distance $|a_n – a_i|$ is at most *m*.

For each subsequent point which is needed to be predicted all possible patterns are applied. If the points around the point being predicted match some pattern up to some error $\mathcal{E}$, the matching value from the pattern is considered as a part of the set of possible prediction values $\hat{S}$ for the given point. $\hat{S}_t = \{S_{t1}, S_{t2}, S_{t3} \ldots\}$ would be a set of possible prediction values for the point X_t. The number of possible prediction values for any point is limited from above by the number of all the patterns generated from the training set and by 0 below. The error $\mathcal{E}$ is also a parameter of algorithm and should be chosen carefully, as too high value would lead to too many values, many of them could be irrelevant, and too low value would lead to no matching patterns at all.

If the unified prediction value for the given point is impossible to obtain due to reasons discussed in the following chapter, the point is said to be unpredictable, and no prediction value is assigned to it. The indicator function $\zeta(\hat{S}_t)$ is said to be 1 if the point X_t is predictable and 0 otherwise.

$$\zeta\left(\hat{S}_t\right) = \begin{cases} 1, \text{if the point is predictable} \\ 0, \text{otherwise} \end{cases}$$

Form. 1.1. The indicator function for the point predictability

$\hat{S}_t$ – set of possible prediction values

The self-healing algorithm allows to increase the quality of the prediction of chaotic time series for the given set of patterns. The algorithm repeatedly iterates through all values of the testing set, matching the patterns and updating the values of the points. With each iteration of algorithm not only the value of the point, but its status in terms of predictability can be updated. The points which were considered unpredictable could find a pattern to match because of the changed values of the points near it and vice versa.

The mathematical problem statement consists of both minimizing the mean squared error of prediction on predictable points and minimizing the number of unpredictable ones. It differs from the conventional regression goal used in machine learning, where there is only function to minimize (usually some metric like MSE).

$$I_1 = \sum_{i=t+1}^{t+h} \left(1 - \hat{S}_i\right)$$

$$I_2 = \frac{1}{N} \sum_{i=t+1}^{t+h} \zeta\left(\hat{S}_i\right) (\hat{y}_i - y_i)^2$$

Form. 1.2. The functions subject to optimizing (I_1, I_2)

$\hat{S}_i$ – set of possible prediction values

N – the number of points

$\hat{y}_i$ – the predicted value

y_i – the actual value

Prediction algorithms

When it comes to the way of forming a unified prediction value from the given set, there are many possible ways. From the statistical point of view, one could take an average of all observations, but then the prediction will flatten at the average value of the time series itself (for normalized time series it would be about 0.5). Alternatively, one could take the mode, but it is a poor way of addressing the continuous distribution.

The predictive clustering method suggests choosing the center of the largest cluster as a unified prediction value. The clustering algorithm may wary and can be set as a parameter of prediction algorithm. Naturally the possible clustering algorithms which can be used are limited by the fact that the number of clusters is not known before. The term paper examines DBSCAN, OPTICS, and Wishart clustering.

Depending on the result of clustering the point is assigned to be either predictable or non-predictable. The predictable point is the point for which the satisfying conditions unified prediction value can be find.

The point can be considered non-predictable by the following reasons:

a. No possible prediction values are found at all. That corresponds to the empty set of possible prediction values. It means that no patterns were matched for the given point, which can be caused either by poor training set or too small $\mathcal{E}$ chosen.
b. The new prediction value significantly differs from the previous one. This case requires parametrization by passing the maximal change by iteration as a parameter of the algorithm.
c. In the set of predictive values, there are two equally large (up to an error) clusters.
d. There is only one small cluster.
e. The diameter of the largest cluster is greater than some error.

Self-healing algorithm suggests increasing the prediction quality by improving the prediction error on predictable points and making some of unpredictable points predictable. The algorithm iterates through points by patterns consequently updating their value until the convergence of the metrics.

For the quality metrics to observe the chosen ones are mean squared error (MSE) and number of unpredictable points, which we aim to minimize. As the additional metrics there are mean absolute percentage error (MAPE) and mean absolute error (MAE).

The experiment

For the conducted experiment the Lorenz series was used. Lorenz's equations were developed for studying a simple model of atmospheric convection flows but became a classic example of a chaotic time series when it was found that computer simulations of the equations consistently yielded divergent results, as even the small changes in initial conditions yield different outcomes[12].

$$\begin{cases} \dot{x} = \sigma(y - x) \\ \dot{y} = x(r - z) - y \\ \dot{z} = xy - bz \end{cases}$$

Form. 3.1. The Lorenz system

The Lorenz series is proved to show exhibit chaotic behavior for σ = 10, r = 28, b = 8/3 (and nearby values). Such parameter values were used for simulation.

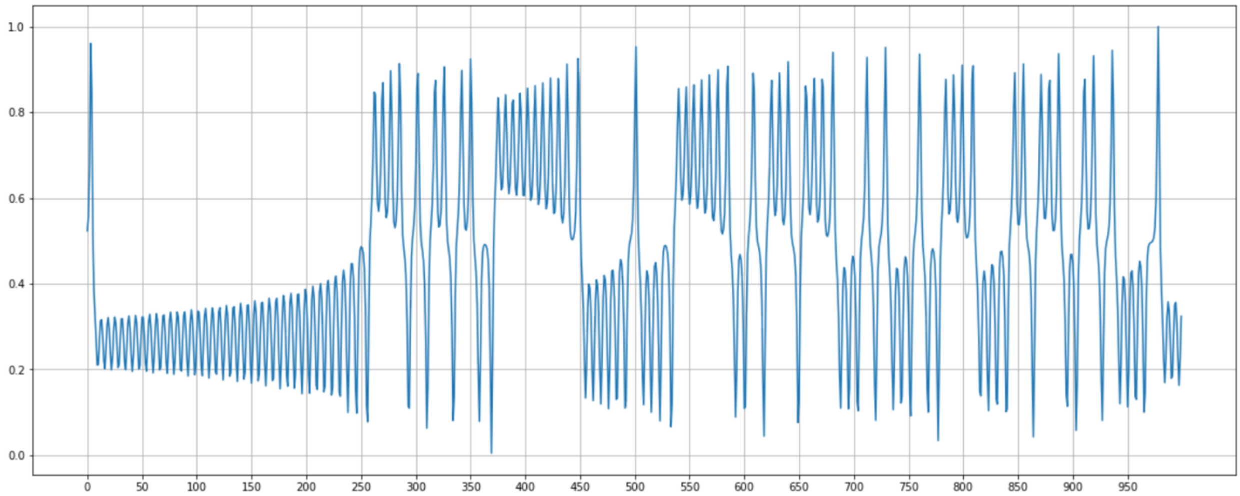


Fig. 3.1. Part of the Lorenz series

Y-axis: the value of the series, X-axis: point number

The series takes the values from 0 to 1 (excluding the endpoints)

[12] Lorenz E.N. The statistical prediction of solutions of dynamic equations // Proceedings of the International Symposium on Numerical Weather Prediction. – Tokyo, 1960. – pp. 628-635.

11000 points were generated, where 10000 points contributed to the training set and 1000 to the testing set (the proportion is 10: 1). Then by all possible combinations the (5, 8) patterns were generated. It required much computing power, so the calculation was performed on the supercomputer.

Firstly, the self-healing algorithm was examined in isolation. We took the test set and excluded randomly 10% of the points, then 20% and so on up to 90%. Then the algorithm tried to recover the missing values by matching the patterns from the training set. The algorithm was repeated until the number of unpredictable points and errors converged to the minimum.

However, such naïve implementation had a serious drawback. It tended to shift the values, which were correct from the very start increasing the error calculated on these points.

Table 3.1. The metrics (MSE, RMSE, MAE, MAPE) on self-healing algorithm depending on the number of points removed (from 10 to 90 with the step of 10)

%	*MSE*	*RMSE*	*MAE*	*MAPE*
10	5,90E-06	0,00242891	0,00057591	0,00088318
20	1,89E-05	0,00434227	0,00138678	0,00241188
30	0,00041145	0,02028422	0,00588307	0,01733945
40	0,00032705	0,0180844	0,00636165	0,01244839
50	0,00770173	0,08775951	0,02327585	0,07571123
60	0,00647274	0,08045335	0,03042171	0,30432897
70	0,00443345	0,06658413	0,03195918	0,06072785
80	0,01875066	0,13693306	0,06105174	0,64467922
90	0,10336399	0,3215027	0,22056787	0,70974734

To solve this problem an idea on point fixation was introduced in two possible implementations:

a. Full fixation. Once a point is considered predictable, it no longer updates its value.
b. Partial fixation. On each iteration the new value assigned to the point is a weighted average of its value on the previous iteration and new unified predicted value.

One may say that a. is a special case of b., when the weight of the new value is 0, and the weight of the previous value is 1.

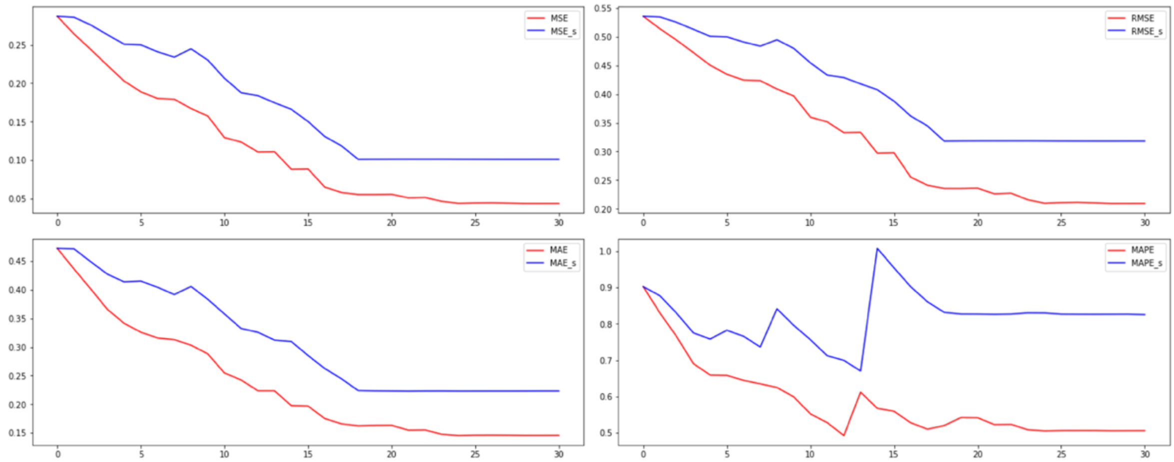


Fig. 3.2. Comparison of performances of algorithms with no fixation (blue) and with full fixation (red); Upper left graph – MSE, upper right graph – RMSE, lower left graph – MAE, lower right graph – MAPE

The result of the experiment showed that the full fixation of the points may lead to the growth of error, as the points which were unpredictable have little influence on each other. Also, in this case algorithm stops in about 20 iterations, as there are no updates of the values.

Table 3.2. The metrics (MSE, RMSE, MAE, MAPE) on self-healing algorithm depending on the number of points removed - partial fixation (from 10 to 90 with the step of 10)

%	*MSE*	*RMSE*	*MAE*	*MAPE*
10	1,69E-04	0,01298938	0,0084582	0,02855475
20	1,46E-04	0,01210071	0,00748086	0,02522349
30	0,00054378	0,02331919	0,0118825	0,03295582
40	0,00079469	0,02819024	0,01679381	0,03530871
50	0,000851	0,02917188	0,01296374	0,0508878
60	0,0064388	0,08024216	0,03296229	0,31479258
70	0,01712526	0,13086352	0,05516578	0,09736329
80	0,03445671	0,18562518	0,09986677	0,78402436
90	0,04385834	0,20942384	0,14563674	0,50585909

Partial fixation of the points demonstrated much better performance. The weights were chosen empirically. The previous value was assigned a weight of 0.8 and the for the new one it is 0.2.

Table 3.2. The metrics on self-healing algorithm depending on the number of points removed - full fixation (from 10 to 90 with the step of 10)

%	*MSE*	*RMSE*	*MAE*	*MAPE*
10	0,00087267	0,02954098	0,02093753	0,11496488
20	0,0006514	0,0255226	0,01820336	0,09339303
30	0,00108576	0,03295082	0,02169536	0,10297759
40	0,00095512	0,03090499	0,02220031	0,09686255
50	0,00121903	0,03491468	0,02068858	0,10211494
60	0,01920387	0,13857804	0,06534026	0,22247578
70	0,05144307	0,22681066	0,1118441	0,28227579
80	0,13218646	0,36357455	0,23486931	0,92747249
90	0,10128167	0,31824781	0,22316004	0,82512775

Introducing partial fixation also led to convergence of the number of unpredictable points to 0.

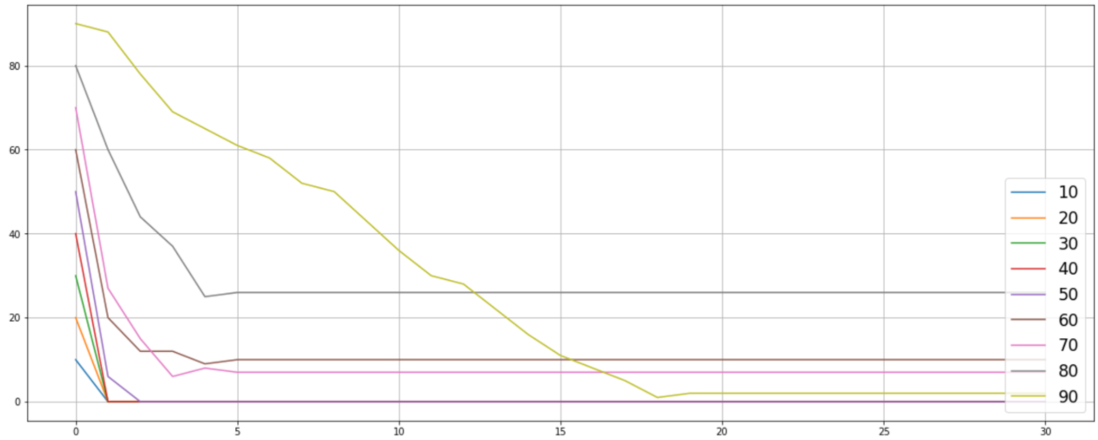


Fig. 3.3. Number of unpredictable points with full fixation. On x-axis: the number of iterations, on y-axis: the number of unpredictable points. Blue: 10%, orange: 20%, Green: 30%, Red: 40%, Purple: 50%, Brown: 60%, Pink: 70%, Grey: 80%, Light green: 90% (proportion of points dropped)

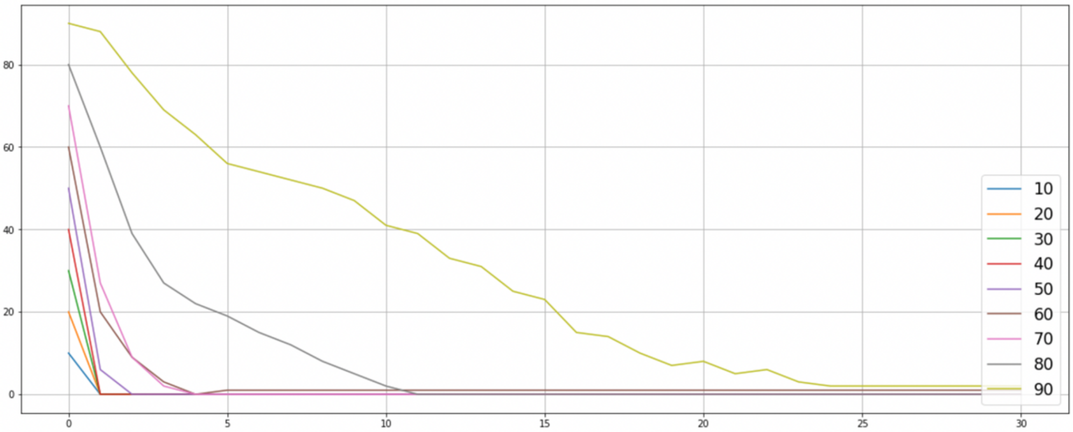


Fig. 3.4. Number of unpredictable points with partial fixation. On x-axis: the number of iterations, on y-axis: the number of unpredictable points. Blue: 10%, orange: 20%, Green: 30%, Red: 40%, Purple: 50%, Brown: 60%, Pink: 70%, Grey: 80%, Light green: 90% (proportion of points dropped)

After the best parameters of self-healing algorithm were deduced, the second part of the experiment consisted of multi-step prediction for 1000 points. The results of the metrics calculated for the three types of point fixation are presented in Table 3.3.

The figure 3.5 represents the results of the multi-step ahead prediction. It can be noted that once again the partial fixation of the points corresponds to better algorithm performance.

Table 3.3. The metrics (MSE, RMSE, MAE, MAPE) on self-healing algorithm for the MSA prediction for three different variations of algorithm (no fixation, full fixation, partial (half-) fixation) after all iterations

Method	MSE	RMSE	MAE	MAPE
No fixation (Total update)	0, 070676283	0, 265850114	0, 174043549	0,646044211
Full fixation	0, 067677963	0, 260149886	0, 157583804	0,781985895
Half-fixation	0, 066345707	0, 257576604	0, 163601835	0,762507231

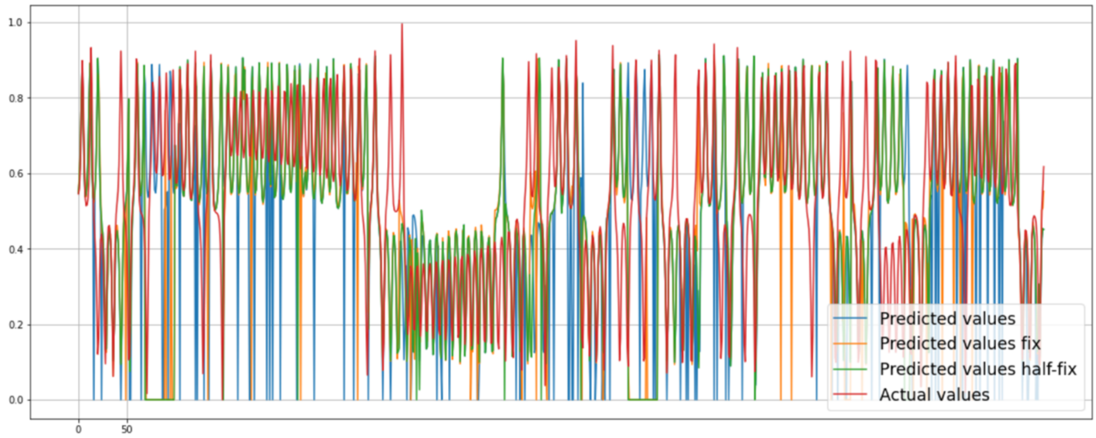


Fig. 3.5. The results of MSA prediction. Y-axis: the value of the series, X-axis: point number. Red: true (actual) values, Blue: prediction values with no fixation, Yellow: prediction values with full fixation, Green: prediction values with partial fixation

Conclusion

The term paper contains research on multi-step prediction of chaotic time series. The study focused on the possibility of the improvement of already existing method of forecasting using clustering algorithms and the concept of non-predictable points by adding self-healing elements. The experiment was conducted on the Lorenz series, and it showed improvements in terms of increasing both mean squared error and number of non-predictable points. It can be concluded that the usage of self-healing algorithm at the worst case can yield equally good predictions and better ones in other cases.

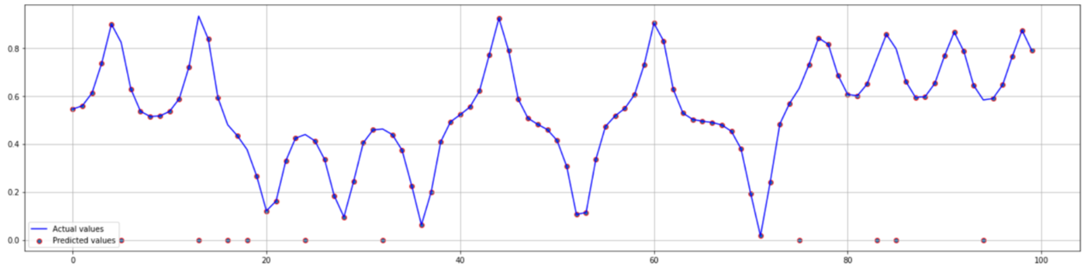
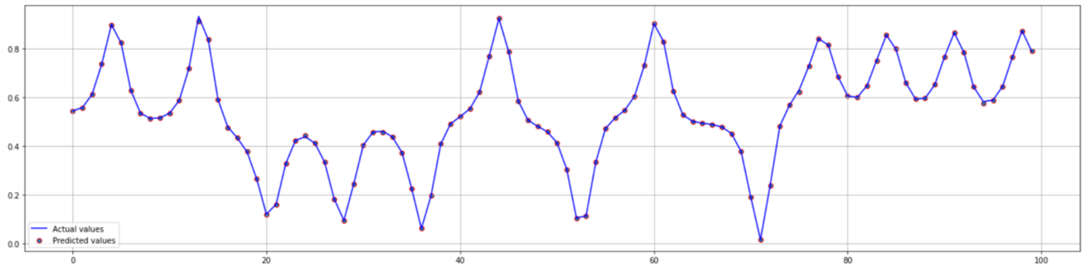
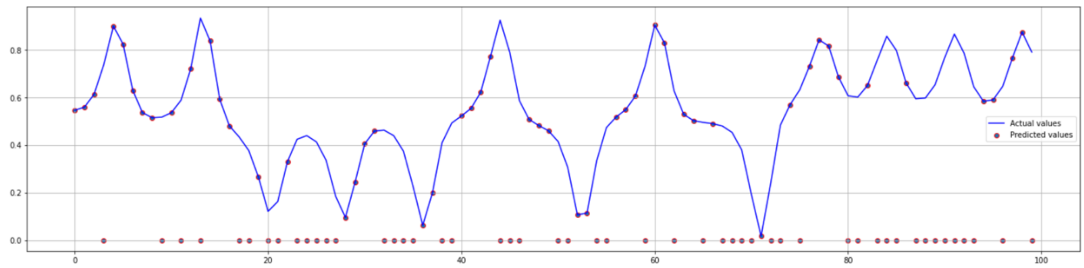
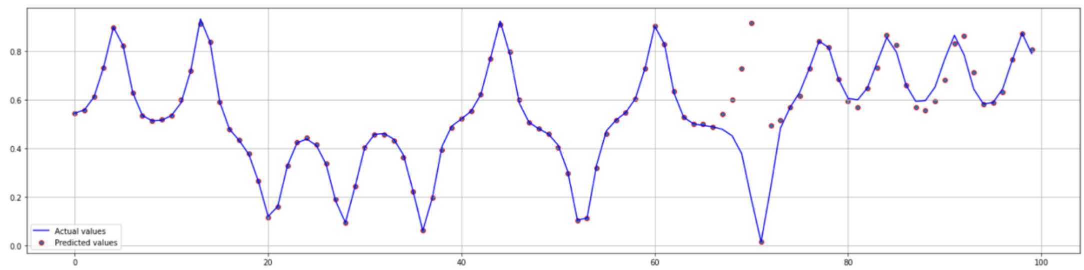
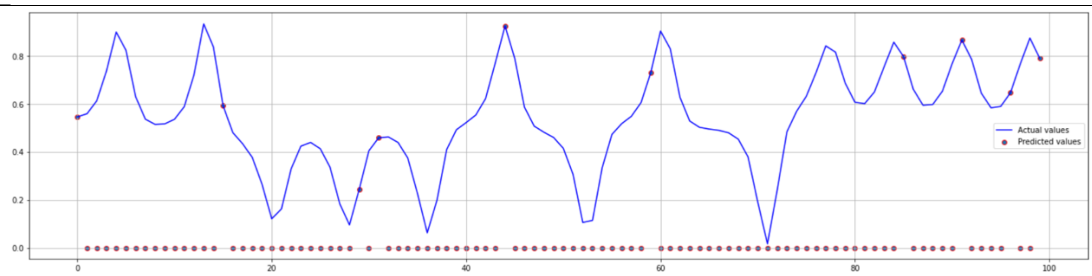
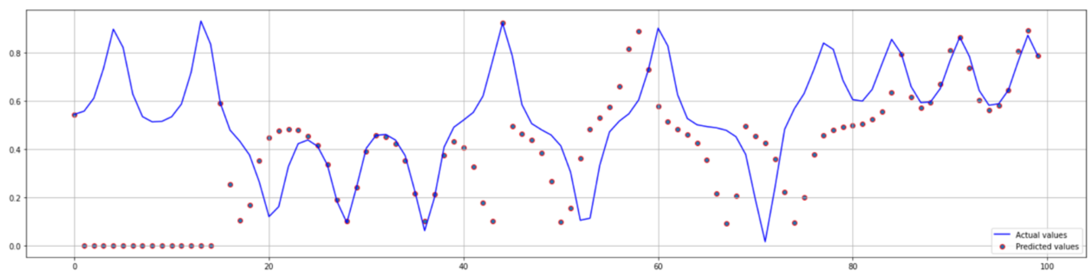
Bibliography

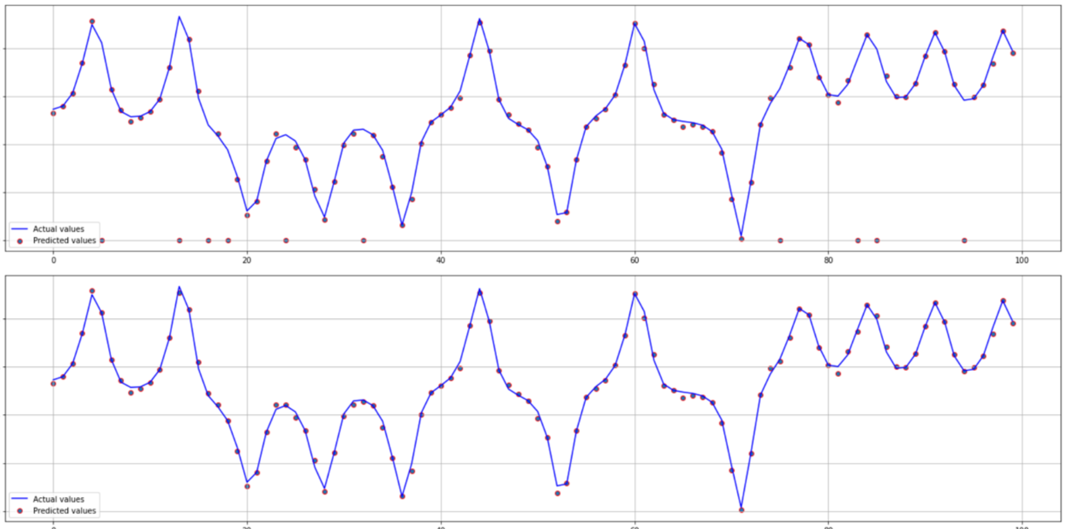
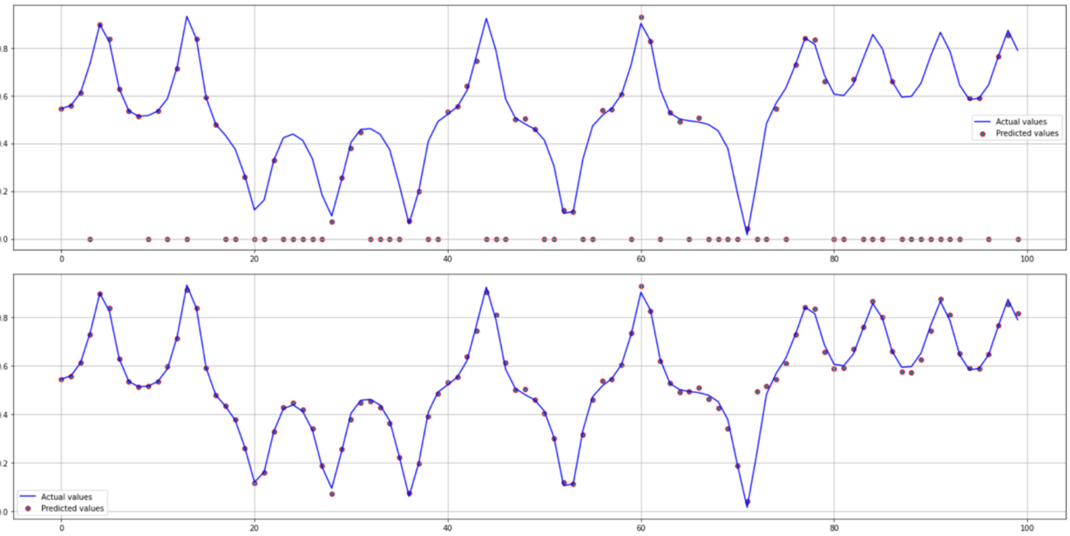
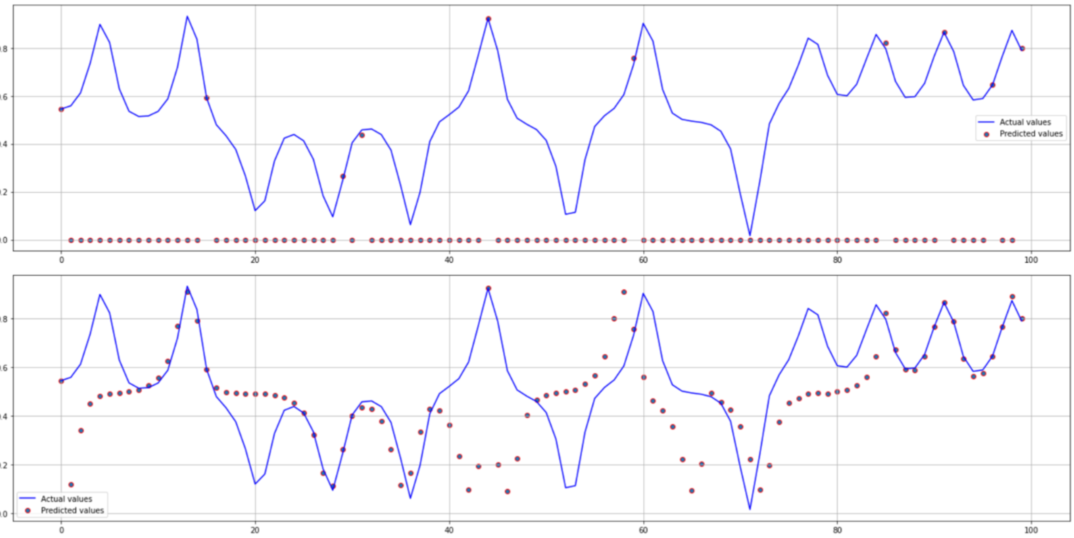
1. Hirsch M. W., Smale S., Devaney R. Differential Equations, Dynamical Systems, & An Introduction to Chaos : Second ed. – Boston : Academic Press, 2003. – 416 p.
2. Gromov V. Chaotic Time Series Prediction: Run for the Horizon // Tools and Methods of Program Analysis. – N.Y. : Springer International Publishing, 2021. – pp. 29-43.
3. Casdagli M. Nonlinear prediction of chaotic time series // Physica D : Nonlinear Phenomena. – May 1989. – Vol. 35. – 335-356 pp.
4. Gromov V. Chaotic Time Series Prediction: Run for the Horizon // Tools and Methods of Program Analysis. – N.Y. : Springer International Publishing, 2021. – pp. 29-43.
5. Lorenz E.N. Deterministic non-periodic flow // Journal of Atmospheric Sciences. – 1962. – Vol. 20. – pp. 130-141.
6. Lisi F., Villi V. Chaotic forecasting of discharge time series : a case study // JAWRA. – April 2001. – Vol. 37. – pp. 271-279.
7. Casdagli M. Nonlinear prediction of chaotic time series // Physica D : Nonlinear Phenomena. – May 1989. – Vol. 35. – pp. 335-356.
8. Farmer J.D., Sidorowich J.J. Predicting chaotic time series // Physical Review Letter. – 24 August 1987. – Vol. 59. – pp. 845-848.
9. Lisi F., Villi V. Chaotic forecasting of discharge time series : a case study // JAWRA. – April 2001. – Vol. 37. – pp. 271-279.
10. Chandra R., Ong Y.-S., Goh C.-K. Co-evolutionary multi-task learning with predictive recurrence for multi-step chaotic time series prediction // Neurocomputing. –
11. Hirata T., Kuremoto T., Obayashi M., Mabu S., Kobayashi K. Time Series Prediction Using DBN and ARIMA // 2015 International Conference on Computer Application Technologies. – N.Y. : IEEE, 2015. – pp. 24-29.
12. Lorenz E.N. The statistical prediction of solutions of dynamic equations // Proceedings of the International Symposium on Numerical Weather Prediction.

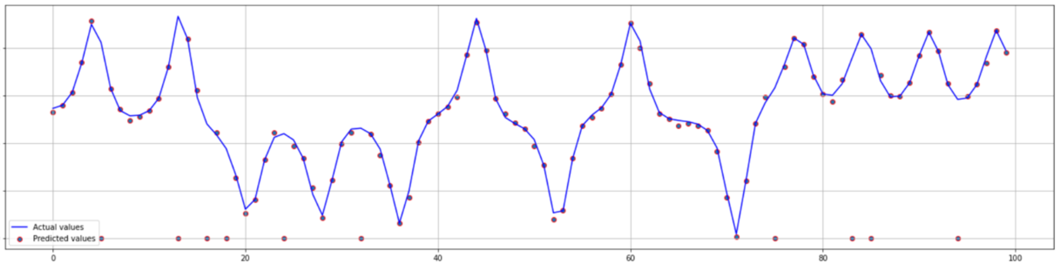
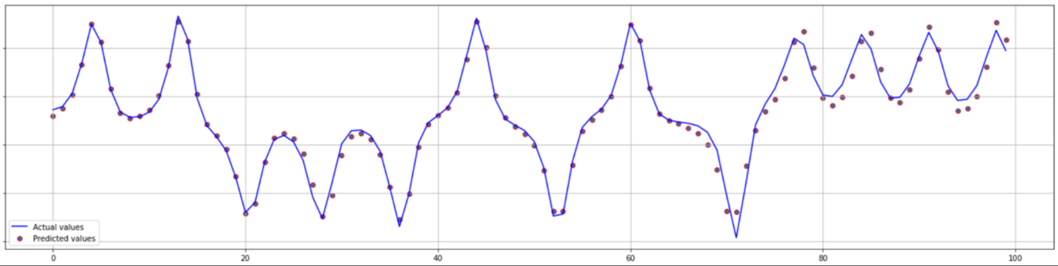
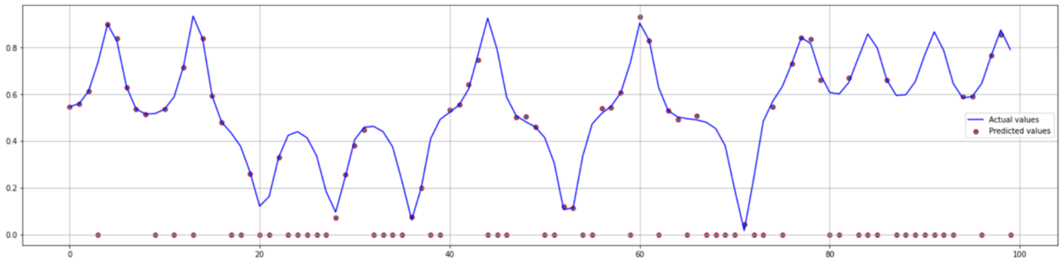
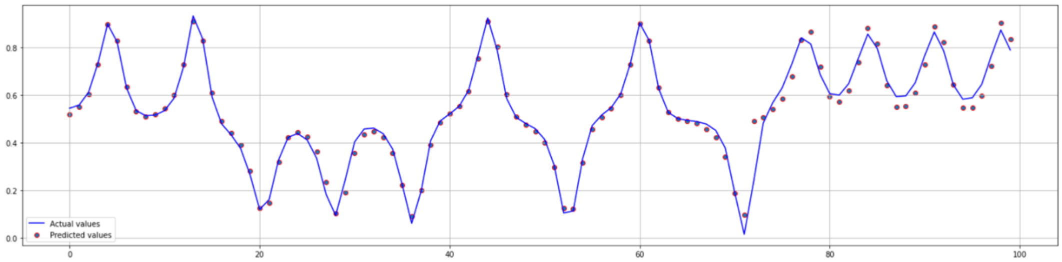
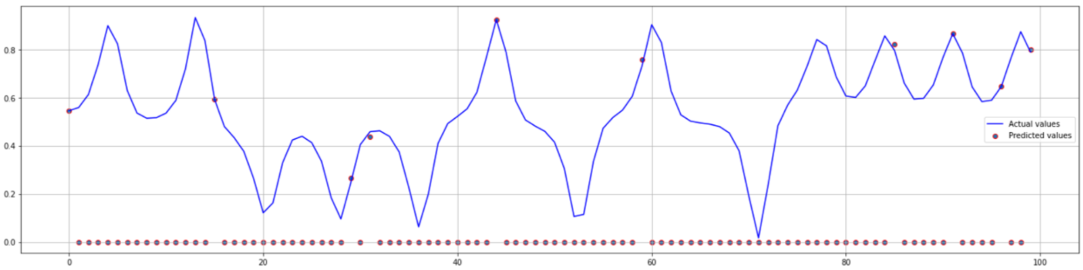
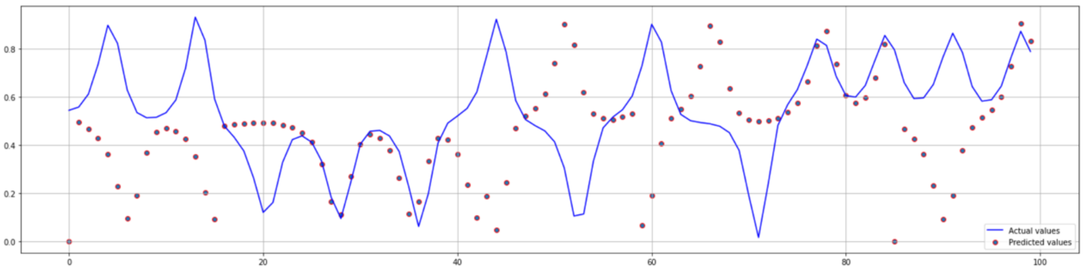
– Tokyo, 1960. – pp. 628-635.

Appendices

Appendix 1. Visualization of the series before and after self-healing algorithm (the first part of the research)

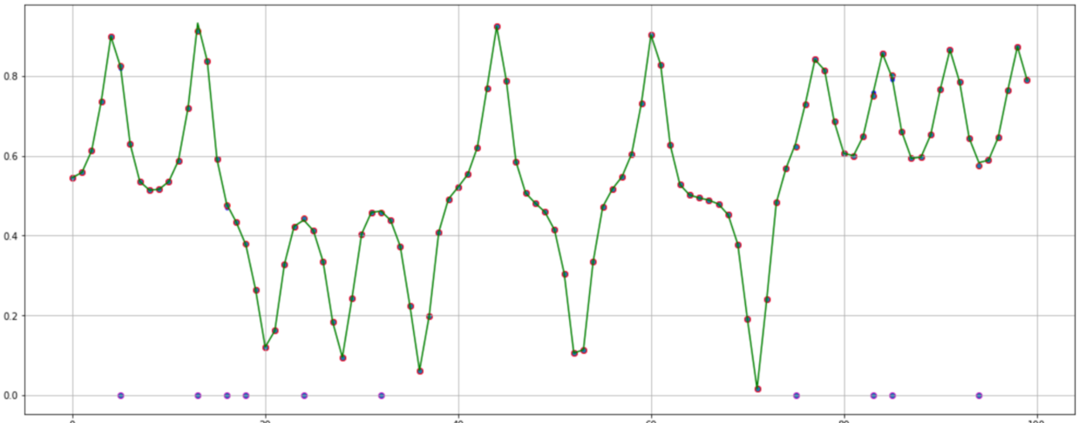
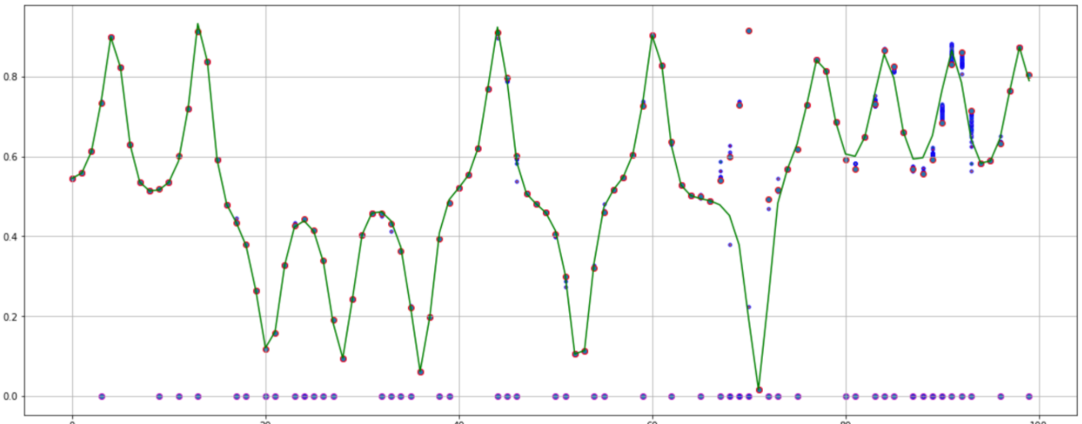
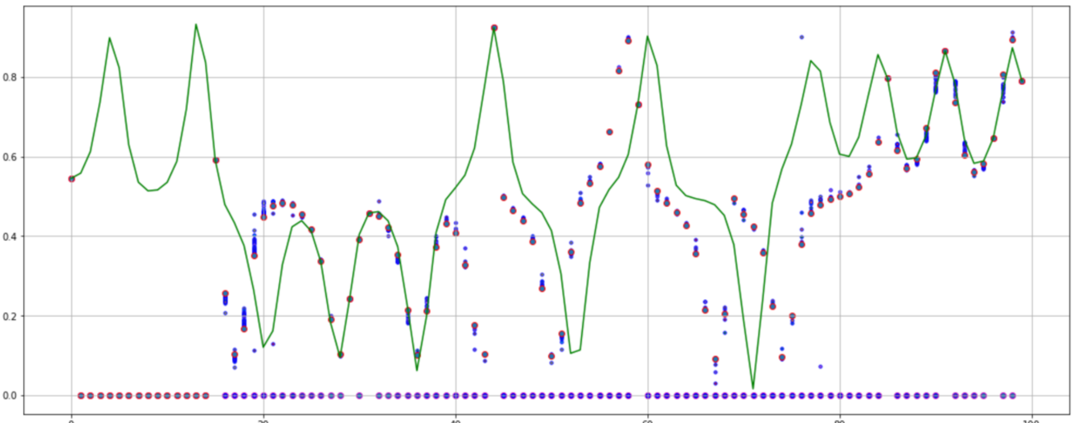
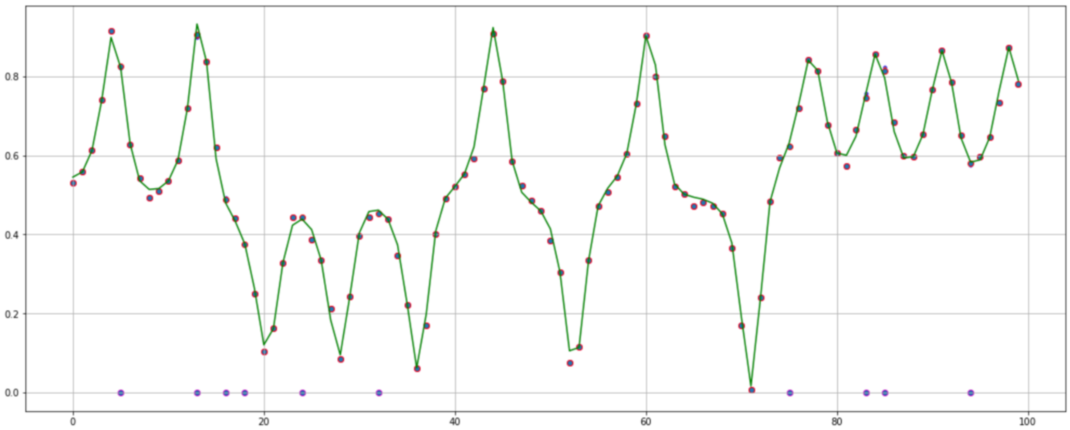
%	Type of point fixation	Graph
10	-	 
50	-	 
90	-	 

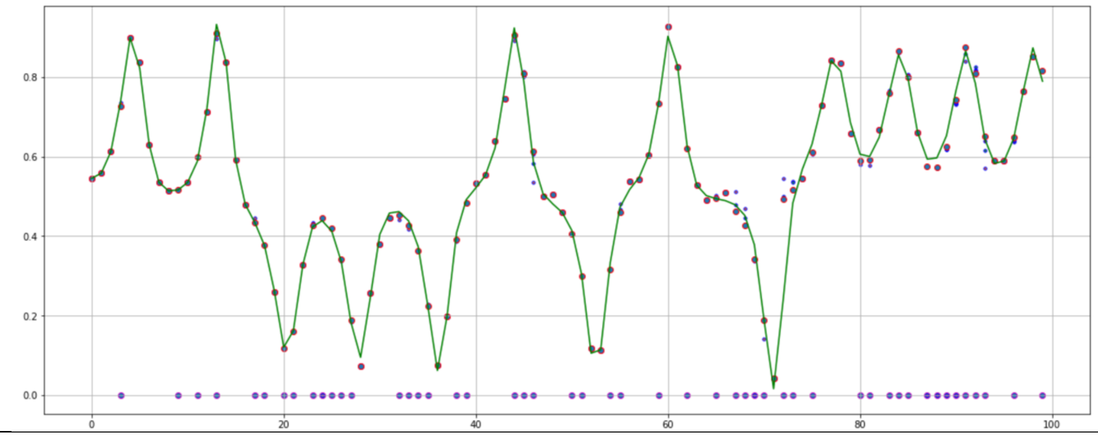
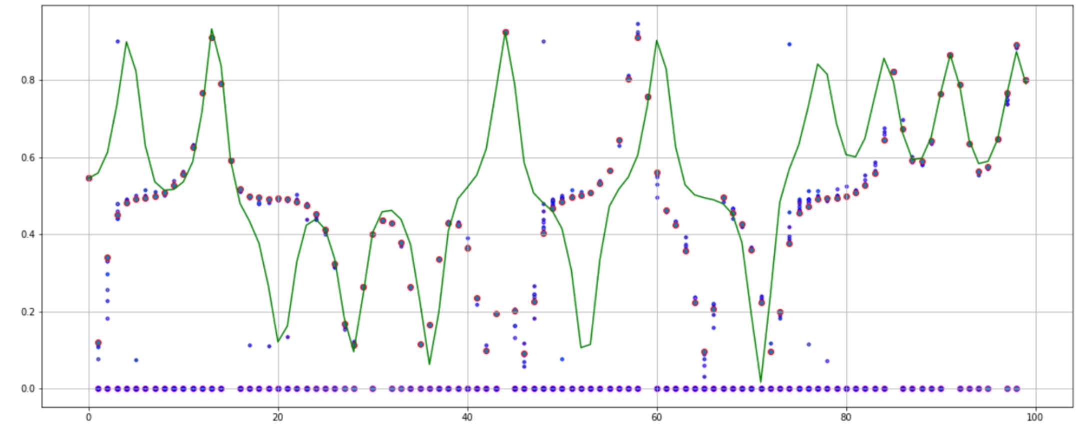
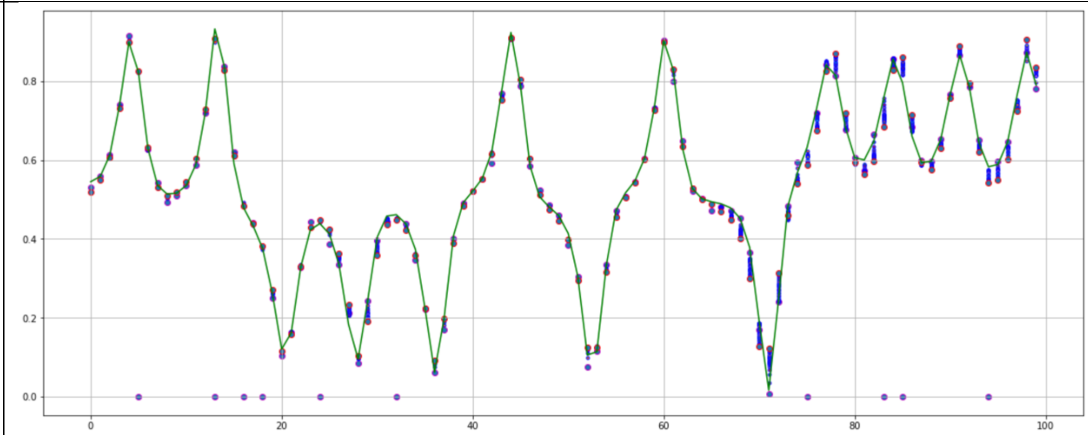
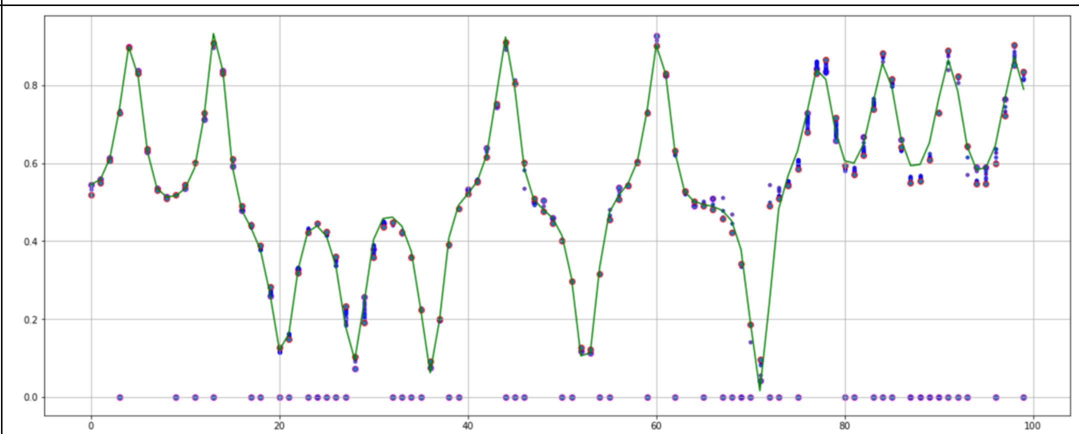
10	f	Actual values Predicted values
50	f	Actual values Predicted values
90	f	Actual values Predicted values

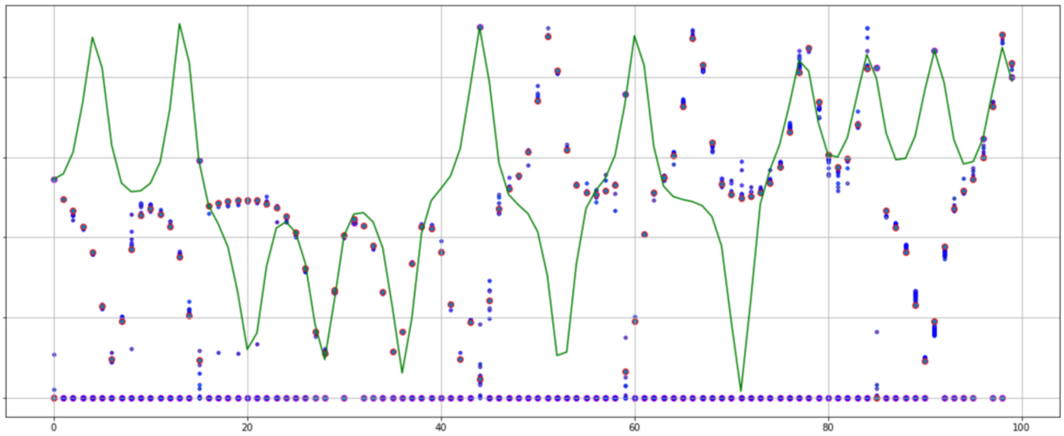
10	hf	 
50	hf	 
90	hf	 

Appendix 2. Change of prediction value by iteration

%	Type of the point fixation	Graph

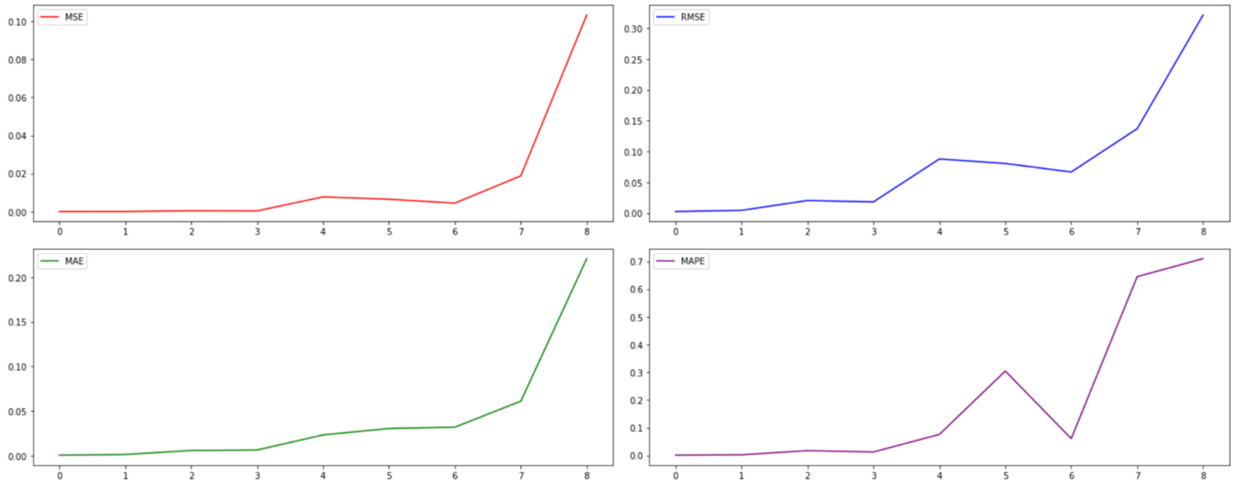
10	-	
50	-	
90	-	
10	f	

50	f	
90	f	
10	hf	
50	hf	

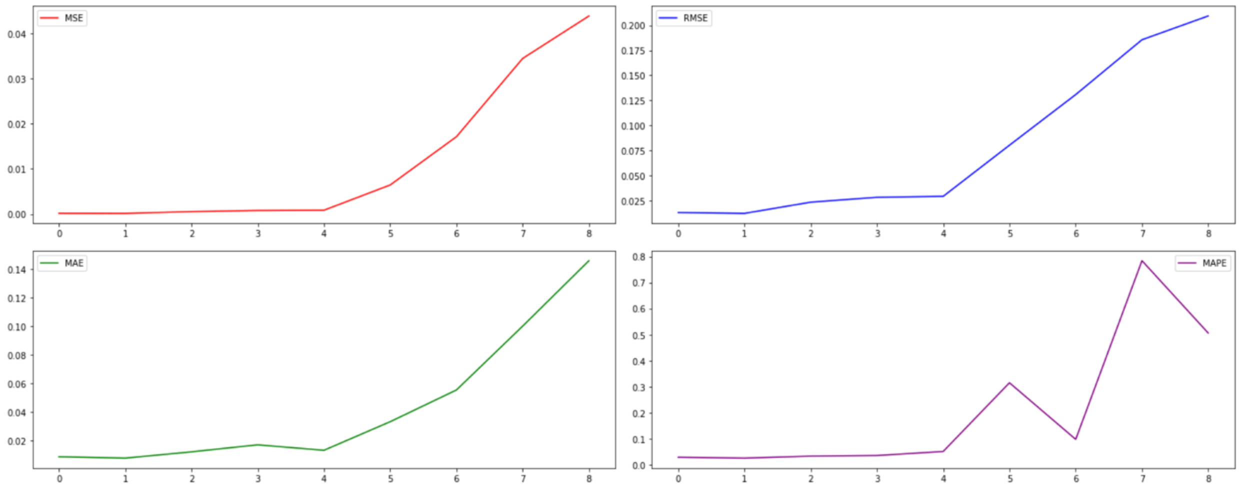
90	hf	

Appendix 3. The comparative graphs of error rate depending on the number of removed points.

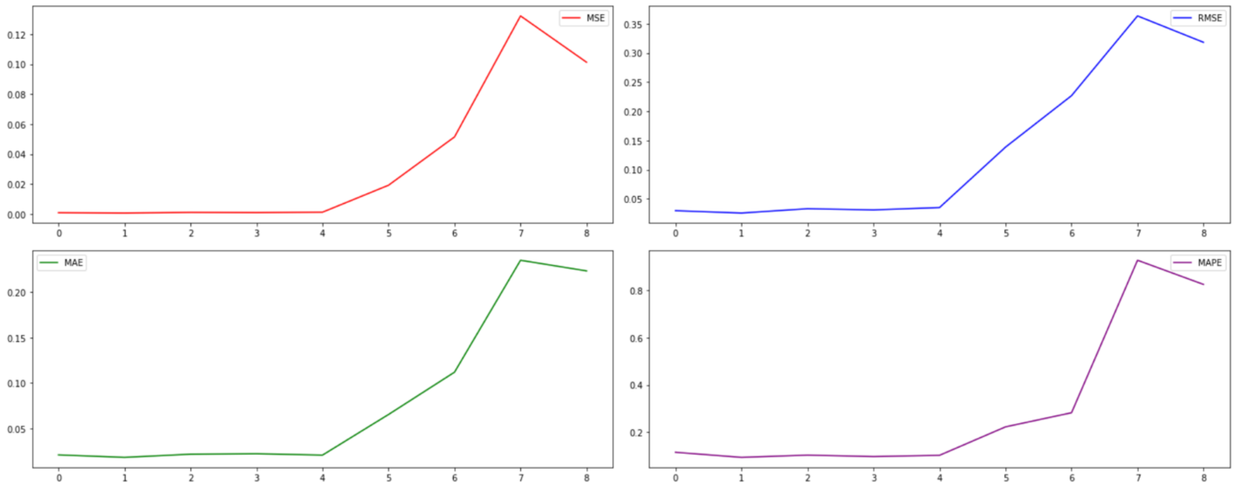
No fixation:



Complete fixation:

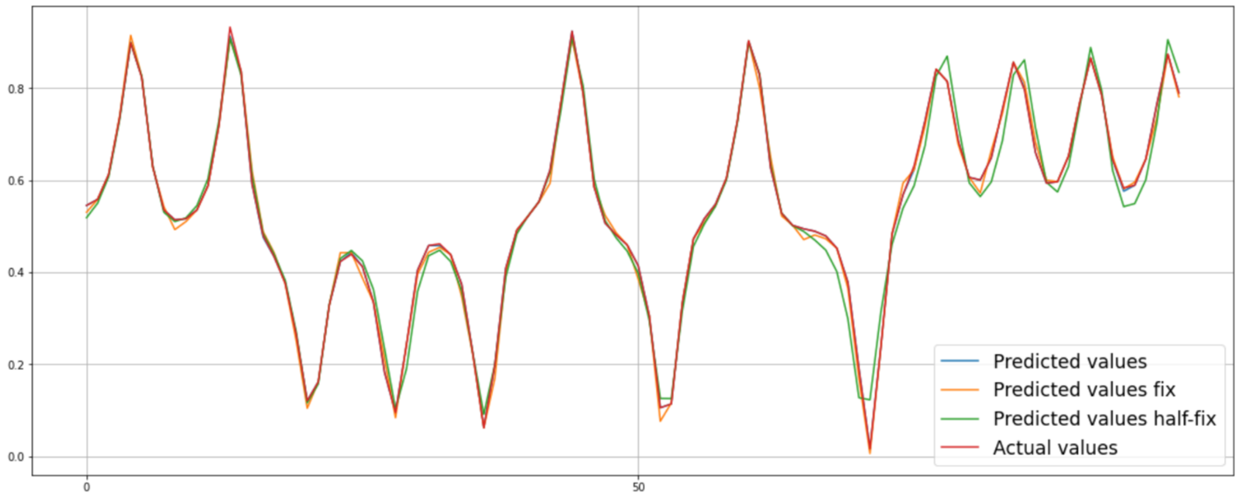


Partial fixation:

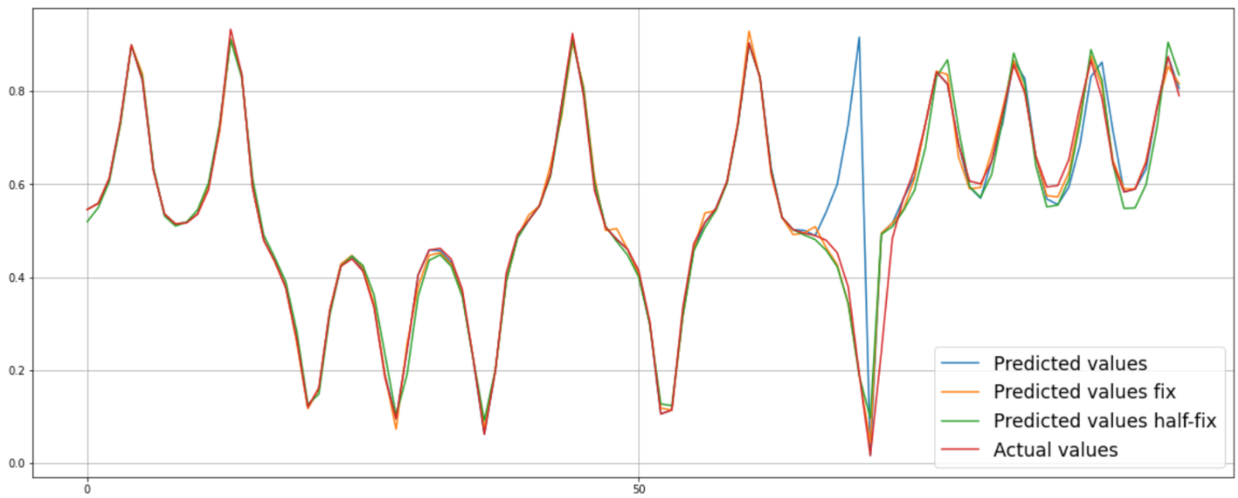


Appendix 4. Comparison graphs on the prediction values depending on number of removed points.

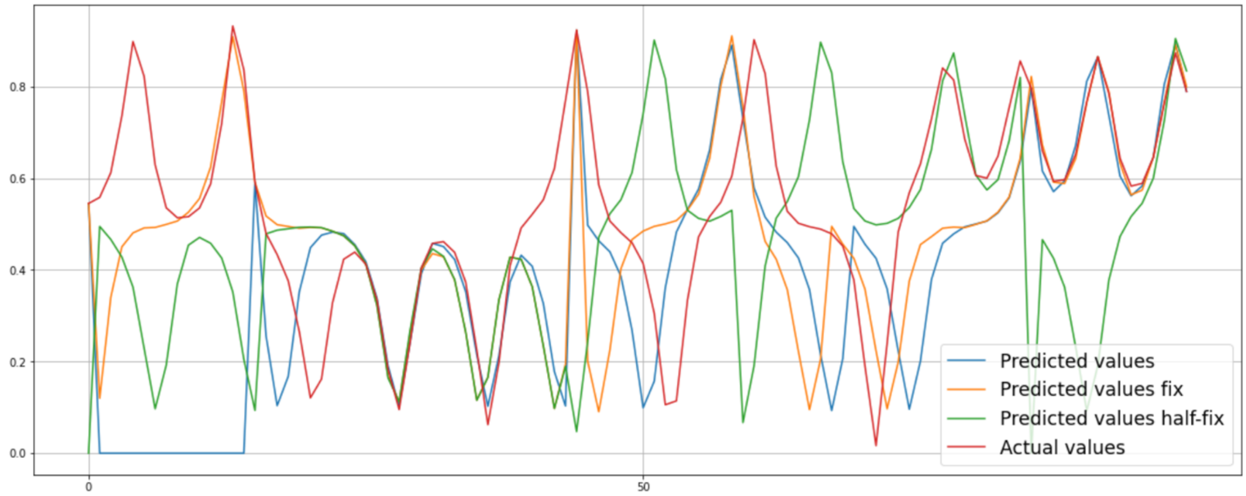
10%:



50%:



90%:



Appendix 5. The link to the code

https://github.com/lysanyikkseniia/chaotictimeseries