

Investigation on Performance of Change Point Algorithm in Time Series Dynamical Regimes and Effect of Data Characteristics

Farhad Asadi, Mohammad Javad Mollakazemi

Abstract—In this paper, Bayesian online inference in models of data series are constructed by change-points algorithm, which separated the observed time series into independent series and study the change and variation of the regime of the data with related statistical characteristics. variation of statistical characteristics of time series data often represent separated phenomena in the some dynamical system, like a change in state of brain dynamical reflected in EEG signal data measurement or a change in important regime of data in many dynamical system. In this paper, prediction algorithm for studying change point location in some time series data is simulated. It is verified that pattern of proposed distribution of data has important factor on simpler and smother fluctuation of hazard rate parameter and also for better identification of change point locations. Finally, the conditions of how the time series distribution effect on factors in this approach are explained and validated with different time series databases for some dynamical system.

Keywords—Time series, fluctuation in statistical characteristics, optimal learning.

I. INTRODUCTION

GENERALLY making inferences about the state of the dynamical systems and time evolution state of system is one of the most important tasks in many computational methods like mechanical engineering or other major such as neuroscience. For instance, uncertainty processing in signals for perception and learning is important for modeling the cognition process. However, such a conclusion would be difficult in noisy and dynamical signals. Understanding of how such integration between events, information and sources of uncertainty is built, is the main objective of learning theory. In the field of neuroscience, it is common to fit the learning rate parameters that have been shown how to replace the old data to the new information. In other words, they try to determine the factors affecting the modeling [1].

Time-varying series data is the key important factor for modeling of complex dynamical systems. For example, the overall mechanism of learning rate should be set up so that the organism is conserving its ability to predict the future. Recently several neural network models are proposed to explain the adaptive behavior of theses mechanism [1].

Farhad Asadi is master student of mechanical engineering in dynamic and control, K. N. Toosi, Tehran, Iran (phone: +989357430949, e-mail: f.asadi@mail.kntu.ac.ir).

Mohammad Javad Mollakazemi is master student of mechanical engineering in dynamic and control, K. N. Toosi, Tehran (phone: +989192762663, e-mail: molakazemi.javad@yahoo.com).

Change-point prediction in different structure algorithms have been developed and applied to a variety of data series such as, process control [2], [3], disease demographics [4], DNA [5], [6], EEG signal processing [7]-[9]. However, most of this analysis has been applied and limited to offline inference, which uses the entire data series for prediction [10], [11]. Online methods also have practical limitations, such as unrealistic modeling in frequency of variation of structure of data, known as the hazard rate. In this paper, this limitation is modified and also the variation of hazard rate in complex time series is simulated. The paper is organized as follows. firstly change-point models have reviewed then simulation for achieving online inference of a constant hazard rate is simulated. And finally some numerical examples to show the approach and important features in this data are explained and investigated numerically.

II. STRUCTURE OF ALGORITHM

This paper is related with the problems of estimating and comparing time-series models. Suppose $Y_n = \{y_1, y_2, \dots, y_n\}$ is a time series data such that the density of variable $y(t)$ given $y(t-1)$ depends on a parameter m in our modelling, whose value changes at unknown time points. in some conditions in algorithm, $\tau_m = \{\tau_1, \dots, \tau_m\}$ Remains constant otherwise: where $\tau_1 > 1$ and $\tau_m < n$. the formulation of inference is related to the estimation of the parameters vector $(\theta_1, \dots, \theta_{m+1})$, so the detection of the unknown change points $Y_m = \{\tau_1, \dots, \tau_m\}$ is required.

This multiple change-point model has been generated in different articles such as [10] and [11]. In this article, we assumed first the joint distribution of the parameters $\{\theta_k\}$ is exchangeable in modeling and it is independent of the change points location.

For modeling change point algorithm there are two basic approaches. The first approaches based on the observation that the ability to detect a change-point location that depends critically on knowledge of the dynamical system before the change point is taken place. In Figs. 1, 3 and 5 the real and generated data for three different data series are shown.

In our simulation, data points are generated by adding Gaussian random noise to a mean value. The mean value varies in a continuous constant manner versus time, while changing suddenly at change-points but otherwise staying smooth constant. Thus, sample-by-sample fluctuations in the data series can reflect both noise and change-points to predict the position of the data point. In Figs. 2, 4 and 6 the run-length

distribution of our approach are plotted. In related figures these points show the run length, rt , which is defined as the number of time steps in modeling to most recent change-point. This parameter has a relatively simple time course, either increasing by 1 on time steps between change-points location or falling to 0 at a change-point. In general, the positions of change-points are not specified.

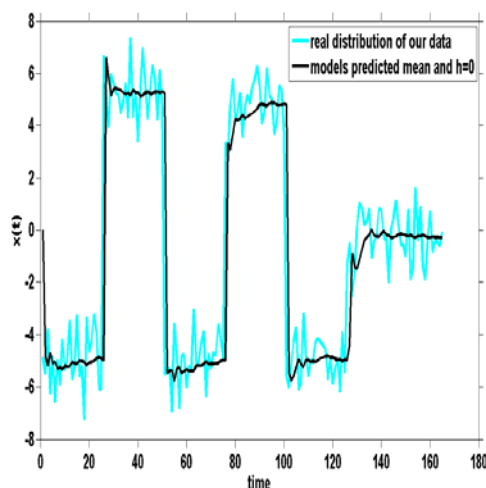


Fig. 1 Effect of changing the hazard rate on the mean value of the predictive distribution

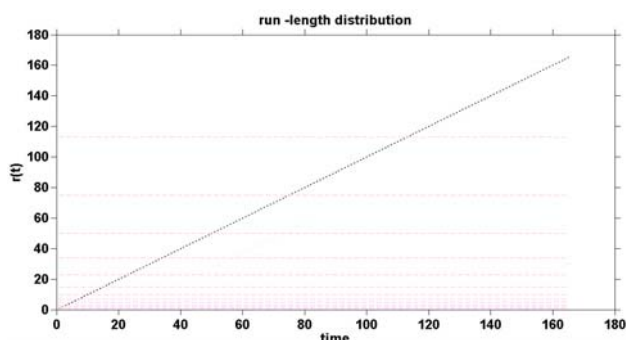


Fig. 2 The run-length distribution in data

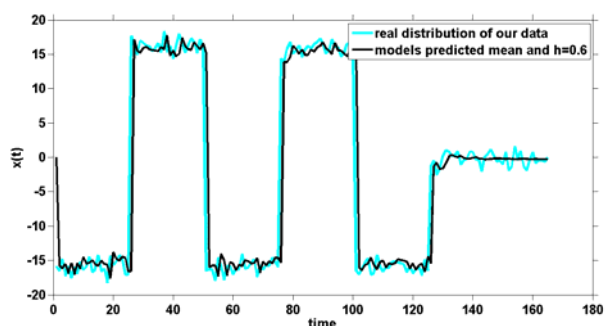


Fig. 3 Effect of changing the hazard rate on the mean of the predictive distribution

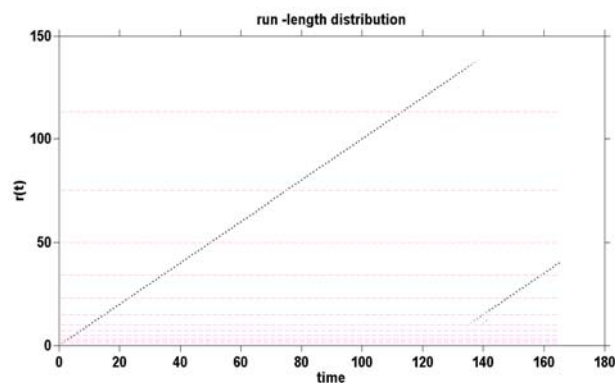


Fig. 4 The run-length distribution of data

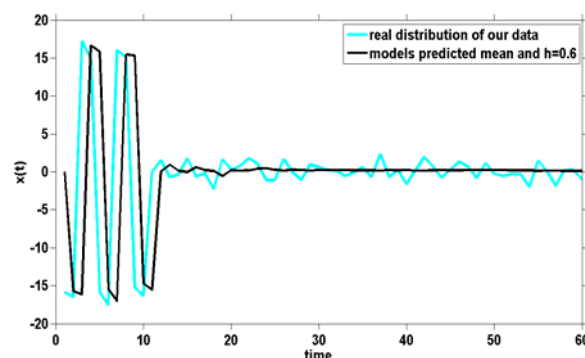


Fig. 5 Effect of changing the hazard rate on the mean of the predictive distribution

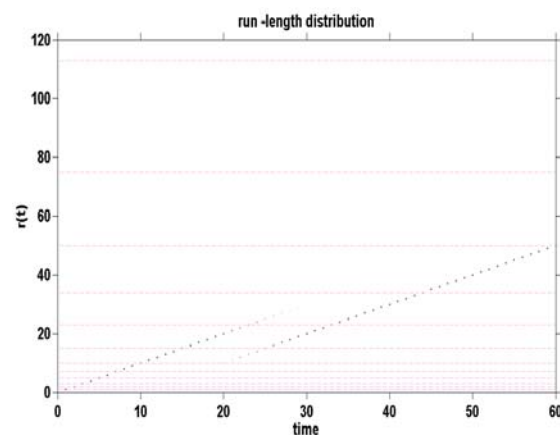


Fig. 6 The run-length distribution of data

Thus, for predictions of the next data point we should consider all possible run lengths in time series data and weight them by the probability of the run length [10]. For example, if $x(t)$ is written as $\{x_1, x_2, \dots, x_t\}$, then the problem of prediction is computing the predictive distribution, $p(x_{t+1} | x_{1:t})$. This distribution can be written in terms of run lengths by given the previous data series. These figures highlight the limitation and effect of choosing h to our data. In Fig. 1 by setting h to zero the algorithm cannot detect the possibility of finding any change-points. In this case, the predictive mean is same as before. So, the run-length distribution is 0 everywhere except at $rt = t - 1$, where is evident in Fig. 2.

Increasing h to 0.6 results in detection of the change-point at $t = 130$. Increasing h further to 0.66 results in more and more change-points is being identified. Thus, the performance of models that require a pre-specified hazard rate can depend critically on which value is chosen, but which value is best is not always obvious in advance in our modeling. So other method and their limitation are described on next section. This approach allows optimal inference in more demanding problems in which the hazard rate is not given and can vary (in a piecewise constant manner) over time [12], [13].

III. SIMULATION WITH CONSTANT HAZARD RATE

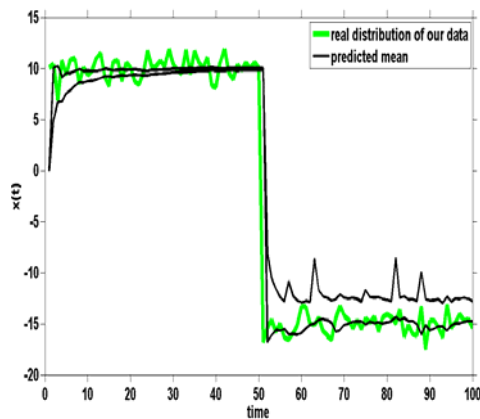


Fig. 7 The variation of actual data and the predicted mean value

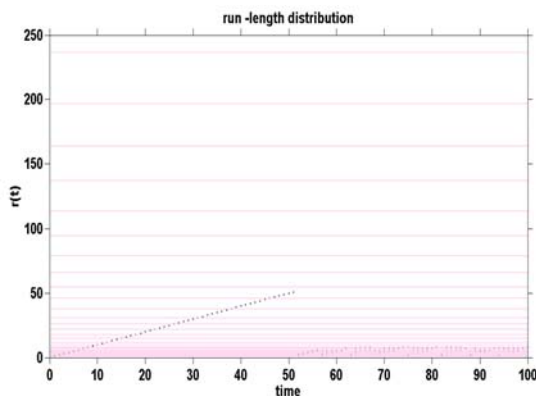


Fig. 8 The actual change-point locations from the generative process

In this section, change-point is modeled as independent distributed samples from a common distribution. We simulate our model by first considering that change-point locations are known, and then address the more challenging case in which change-point locations are changed suddenly. In Figs. 7 and 10 two different time series data are plotted and in Figs. 8 and 11 the run-length distribution of data are plotted and finally the estimation and convergence of hazard rate are plotted in Figs. 9 and 12, respectively. It is evident from figures with respect to the fluctuation of our data, the convergence of hazard rate in our algorithm is changed and also the identification of location in change point is different so it is very important factor that our algorithm can predict hazard rate for better identification. And also in Fig. 11 the variation

of change point is high because of the fast and long variation of amplitude in data series and convergence of hazard rate depend strongly to variation of data so if data has smoothing characteristics then more convergence and lower variation of change point is taken place that it is evident in figures.

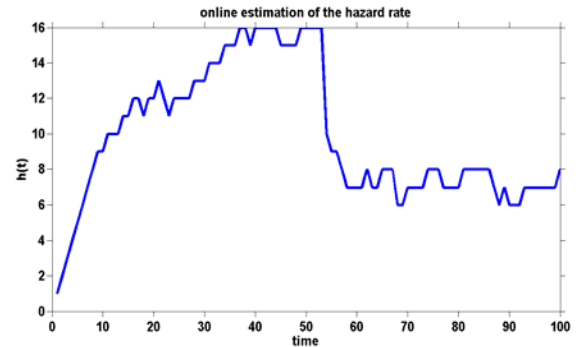


Fig. 9 Estimation of the hazard rate in our modeling

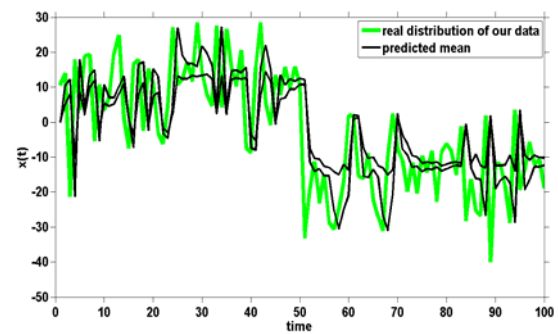


Fig. 10 The variation of actual data and the predicted mean value

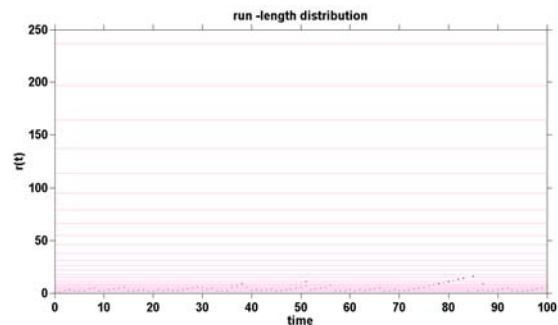


Fig. 11 The actual change-point locations from the generative process

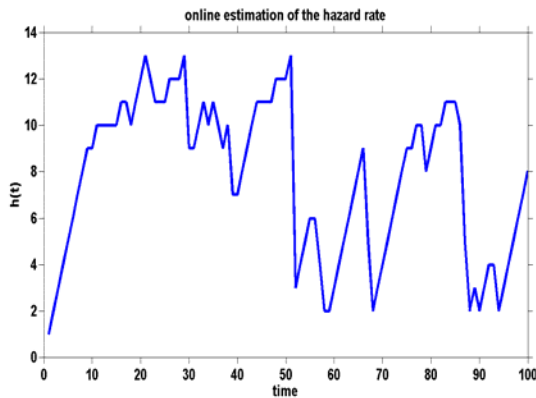


Fig. 12 Estimation of the hazard rate in our modeling

IV. VARIATION OF SPATIAL AND STATISTICAL CHARACTERISTICS OF DATA

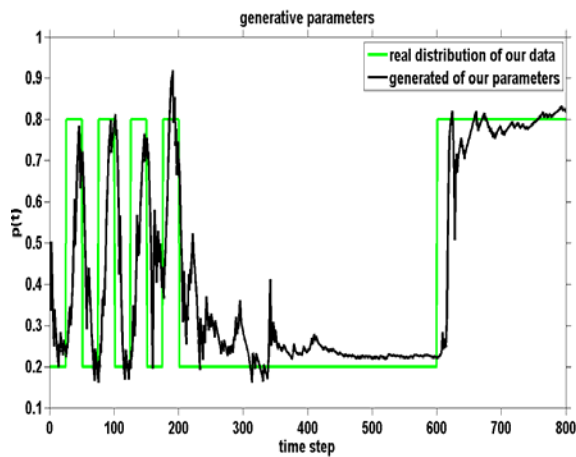


Fig. 13 Generative parameters and real data

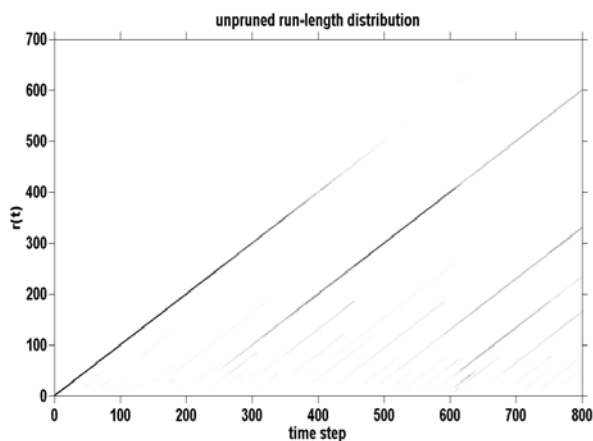


Fig. 14 Unpruned run-length distribution of data versus time

Generally, decomposition of a given time series for detecting phenomena is important. In process of modeling there are hidden or latent parameters in dynamical system when statistical change is taken place. And for this detection, there are different approaches in statistical science.

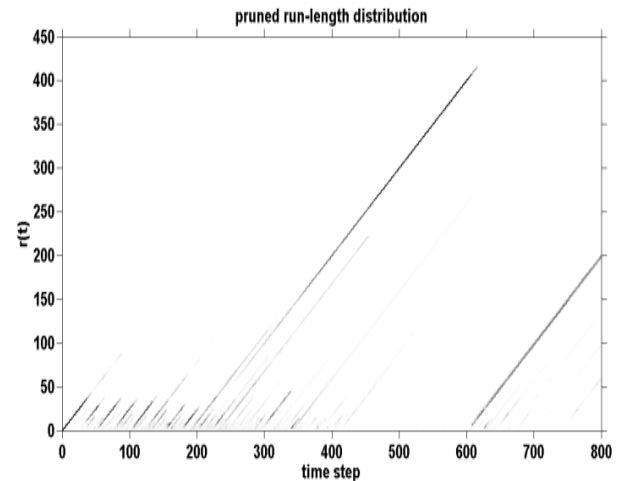


Fig. 15 Pruned run-length distributions of data versus time

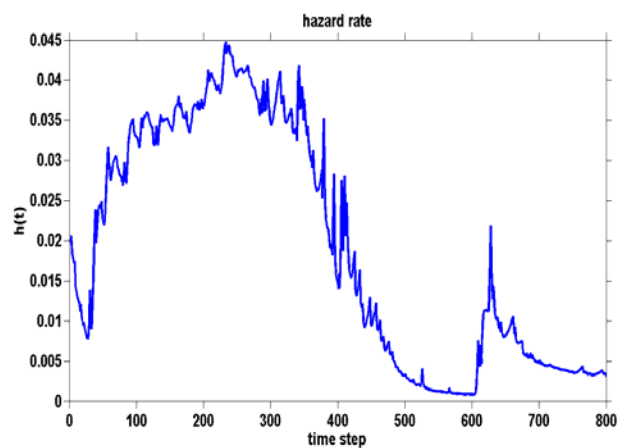


Fig. 16 Estimation and variation of hazard rate

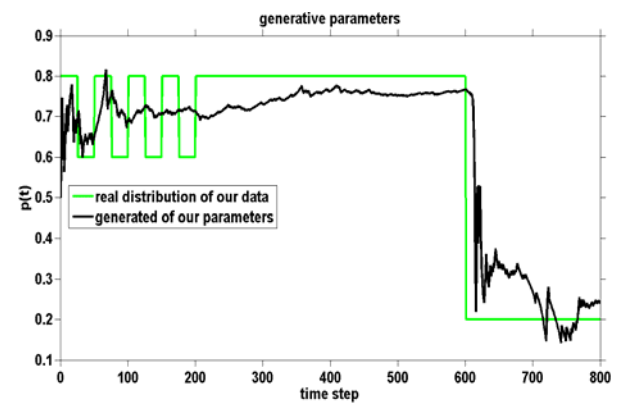


Fig. 17 Generative parameters and real data

In some algorithm when distance between real time data and prediction is larger than some threshold value; a change is detected [8]. In this section different time-series models that are subject to multiple of change points is simulated. In Figs. 13, 17, 21, and 25 four different time series with different features and pattern of variation of mean or variance of data are shown. In Figs. 14, 18, 22, and 26 the unpruned run-length distribution is plotted and pruned results are plotted in Figs.

15, 19, 23, and 27. Finally estimation and variation of hazard rate are plotted in Figs. 16, 20, 24, and 28.

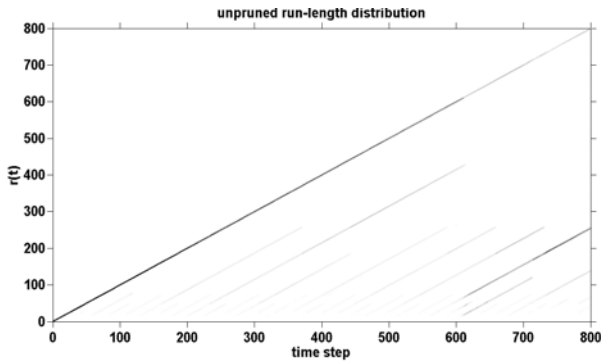


Fig. 18 Run-length distribution of data versus time

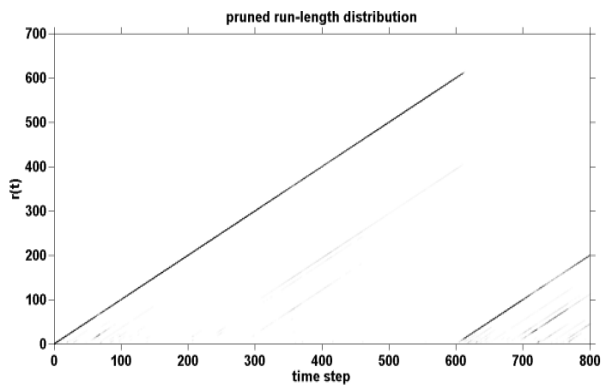


Fig. 19 run-length distributions of data versus time

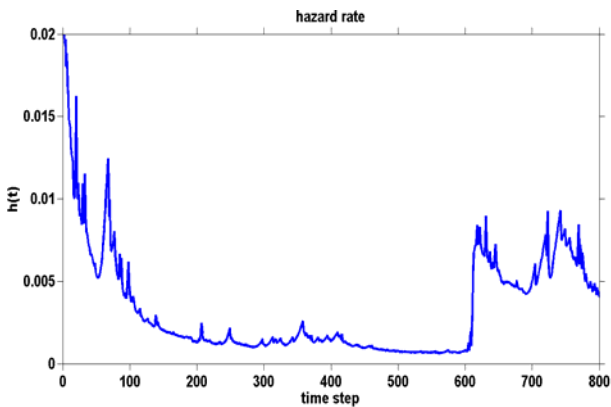


Fig. 20 Estimation and variation of hazard rate

In simulation results, at locations with large changes between data series, the final change point probability is quite high. At other locations, the true change in the mean value is very small, and the model is less likely to put in a change point condition. For instance in Fig. 13 when the variation in change point is larger than run-length distribution then this quantity is low in this area but in all simulations when the variation of our data before the large change point is taken place then estimation of hazard rate has more difficult pattern with more fluctuation.

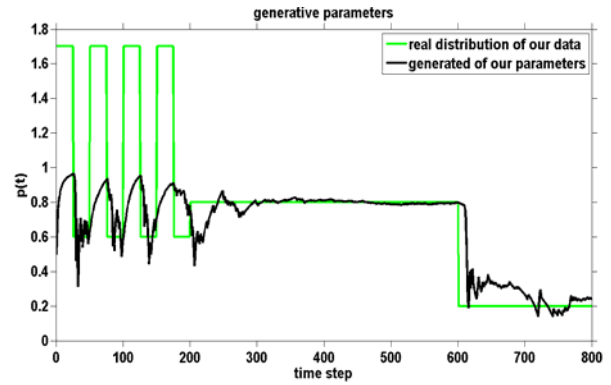


Fig. 21 Generative parameters and real data

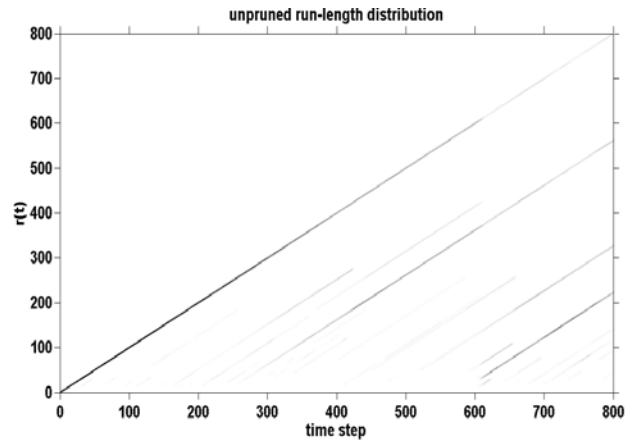


Fig. 22 Run-length distribution of data versus time

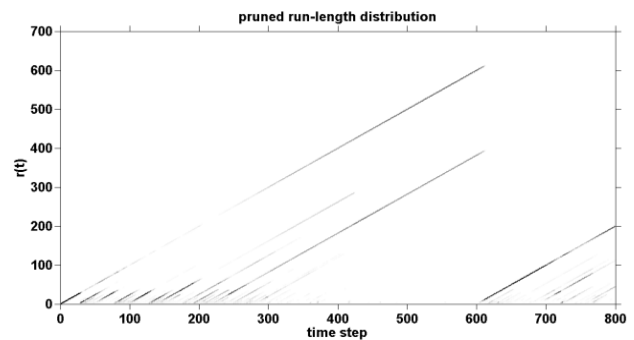


Fig. 23 Run-length distributions versus time

Generally variation of regime in data is important for convergence and distribution of hazard rate and change point distribution. Obtaining posterior distributions of our data over the all the parameters in the model is vital in order to better quantify uncertainty. And it should be noted here when the estimation of our parameters has more density in some region of our data or when the length of proposed distribution is designed in suitable pattern then the whole distribution of our model is better and it has little fluctuation characteristics.

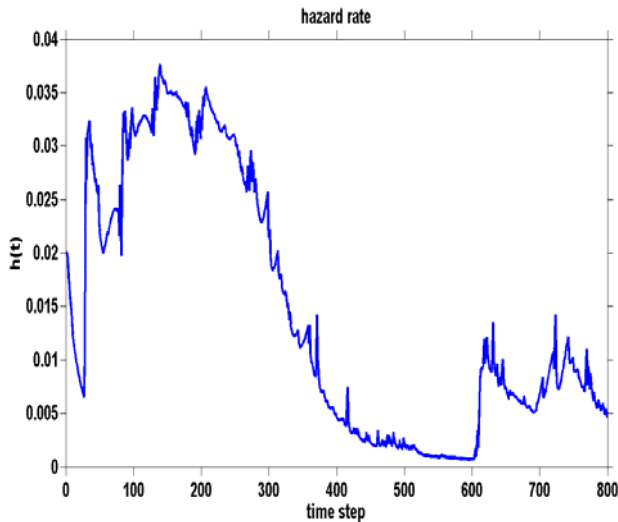


Fig. 24 Variation of hazard rate in our modeling

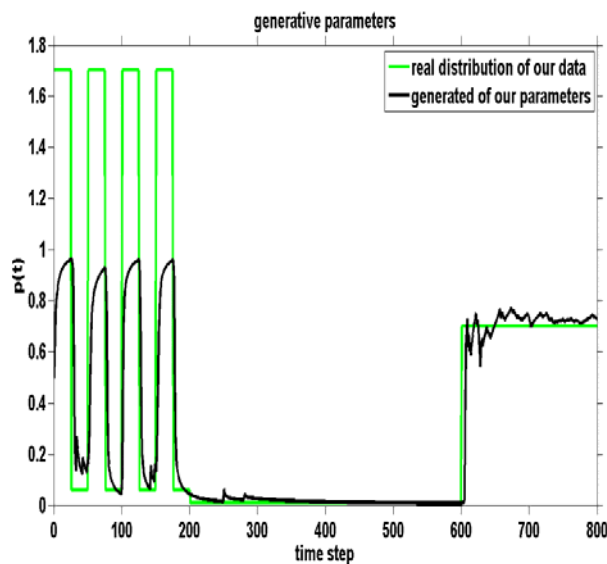


Fig. 25 Generative parameters and real distribution data

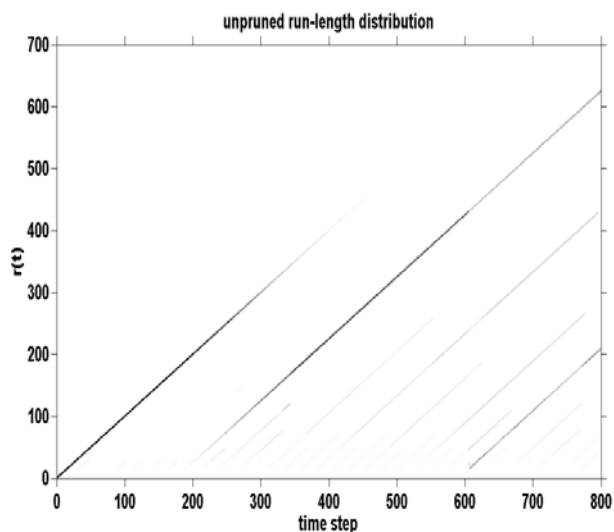


Fig. 26 Run-length distribution of data versus time

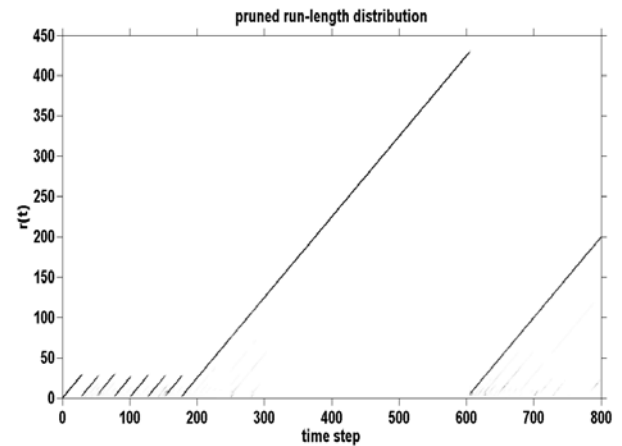


Fig. 27 Run-length distributions of data versus time

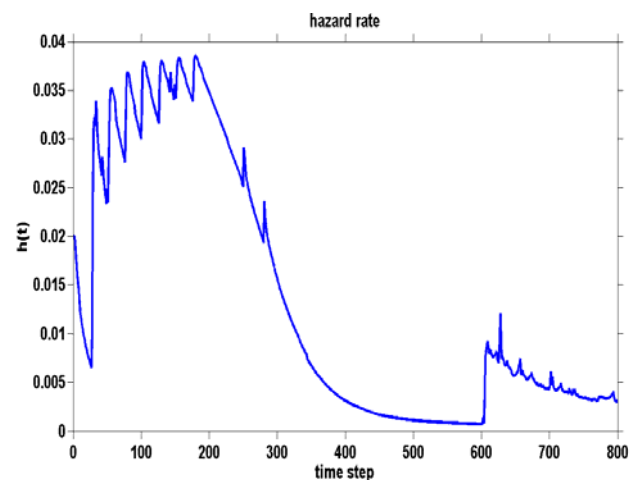


Fig. 28 Variation of hazard rate in our modeling

Another important contribution of this article is simulation results is verified estimation of hazard rate and identification of change –point location is depend to regions of high density in data and algorithm must be designed better in this condition for obtaining these parameters.

V. CONCLUSION

We used an online prediction task to study change point variation. We evaluated algorithm by different time series for showing the importance of time series distribution on change point location and convergence of our model. Finally, we mentioned the length of our proposed distribution for estimating of our data is important factor that effect on simpler and smother fluctuation of hazard rate or better identification of change point.

REFERENCES

- [1] Yu AJ, Dayan P, "Uncertainty, neuromodulation, and attention, *National Center for Biotechnology Information*", 19;46(4):526-8,2005 May.
- [2] Timothy E J Behrens, Mark W Woolrich," *Learning the value of information in an uncertain world*". *Nature Neuroscience* **10**, 1214 - 1221, 2007.

- [3] Xiang Xuan, “*Modeling changing dependency structure in multivariate time series*”, Proceedings of the 24th international conference on Machine learning, Pages 1055-1062, 2004.
- [4] Kiyohito Iigaya, “*Dynamical Regimes in Neural Network Models of Matching Behavior*” neural computation, Vol. 25, No. 12, Pages 3093-3112, December 2013.
- [5] M Steyvers, “*prediction and change detection*”: Springer-Verlag, 2006.
- [6] David J.C. MacKay, “*Ensemble Learning for Hidden Markov Models*”, 1997.
- [7] Adams, R. P., & MacKay, D. J., “*Bayesian online changepoint detection*.” (Tech. Rep.). Cambridge: Cambridge University, 2007.
- [8] Fearnhead, P., & Liu, Z “*On-line inference for multiple changepoint problems*”. Journal of the Royal Statistical Society: Series B (Statistical Methodology), 69(4), 589–605.
- [9] Xuan, X., & Murphy, K., Rocky, “*Modeling changing dependency structure in multivariate time series*. In Proceedings of the 24th International Conference on Machine Learning. San Francisco: Morgan Kaufmann, 2007.
- [10] Wilson, R. C, “*Parallel Hopfield networks*.” Neural Computation, 21(3), 831-850, 2009.
- [11] Wilson, R. C., Nassar, M. R., & Gold, J. I, *Bayesian online learning of the hazard rate in change-point problems*. Neural Computation, 22(9), 2452-2476, 2010.
- [12] Farhad Asadi, Mohammad javad Mollakazemi et al, “*The influence of parameters of modeling and data distribution for optimal condition on locally weighted projection regression method*”.. Accepted and oral presentation in ICMSE 2014: XII International Conference on Mathematics and Statistical Engineering, October, 27-28, 2014, Istanbul, turkey.
- [13] Mohammad javad Mollakazemi, Farhad Asadi., “*Real-time adaptive obstacle avoidance in dynamic environments with different D-S*”. Accepted and oral presentation in ICARM 2014: XII International Conference on Automation, Robotics and Mechatronics, October, 27-28, 2014, Istanbul, turkey.