# Forward Intensity Model Monitoring Using Multivariate Exponential Weighted Moving Average Scheme

*Indicate Submission Type: Research-in-Progress*

**Linzi Liu**

**Xin Lai**

**Yuting Rong**

**Wei Huang**

## Abstract

*We propose a parameter monitoring method for the forward intensity model – the default probability prediction model of the Credit Research Initiative (CRI). We review the relative statistical process control scheme in the field of engineering. Based on this, we propose a new Multivariate Exponential Weighted Moving Average (MEWMA) scheme to monitor the forward intensity model monthly. This new chart might be applied to identify and diagnose the out-of-control (OC) parameters in real time as the data updating, which reduces the cost of recalculating all parameters and improve the operational and calculational efficiency of the default prediction models in practical application.*

**Keywords:** Default, forward intensity, Maximum pseudo-likelihood, MEWMA chart, Phase Ⅱ monitoring

## Introduction

The forward-intensity model of Duan et al. (2012) is a parsimonious and practical way for predicting corporate defaults over multiple horizons. It has highly prediction accurate for shorter horizons. And its accuracy deteriorates somewhat when the horizon is increased to two or three years, but the performance still remains reasonable. Furthermore, the forward-intensity model has been proved to perform well in comparison to the rating results proposed by S&P (Duan and Laere 2012). In recent years, the Credit Research Initiative (CRI) of the Risk Management Institute at the National University of Singapore has carried out practical application based on this model. The researchers recalculate the parameters with the data updated every month to guarantee the model's prediction accuracy.

In this way, the researchers have to calculate the parameters repeatedly for the updating data and the results might not be quite different, which takes a large amount of operating and calculating costs. Whereas in recent research, there are no one focus on the parameters detecting of default prediction models. Because most papers only use cross-section data for analysis, and there is no need to process data which is constantly updating in the long-term application process. Therefore, they usually ignore the significance of parameter variation and monitoring study.

Statistical process control (SPC) has been widely used to monitor various industrial processes. Multivariate Exponential Weighted Moving Average (MEWMA) scheme, one kind of SPC, is often applied to monitor linear profiles with multiple parameters, such as the forward-intensity model. The MEWMA chart was first proposed by Lowry et al. (1992), and the design of MEWMA charts was investigated by Prabhu and Runger (1997). Zou, Tsung and Wang (2007) propose a novel MEWMA scheme to monitor a general linear profile, they apply an MEWMA scheme to the transformations of estimated profile parameters as a single chart to monitor both the coefficients and variance of a general linear profile. Zou and Tsung (2008) proposes a directional MEWMA scheme integrating the EWMA scheme with the generalized likelihood ratio test (GLRT) scheme that can incorporate directional information based on the multistage state-space model and effectively monitor the process mean shift. In addition to industrial processes, MEWMA scheme has been used in economic and medical fields. Some papers provide an algorithm for evaluating the ARL values for the MEWMA chart through simulation and using the results to determine the optimal chart parameters given process cost and time information (Linderman and Love 2000; Niaki et al. 2010; Niaki et al. 2011). MEWMA also has been applied to detect an increase in the incidence rate of a given disease or other medical condition (Joner et al. 2010). From the above, MEWMA method can be used to solve problems which need to be controlled or detected in a time series and have multiple parameters. However, research on monitoring and diagnosis of default prediction models remains scanty. So we applied it to monitor default prediction models, like the forward intensity model.

In this article, we focus on studying the phase Ⅱ method for monitoring a nonlinear profile that can be applied to monitor the forward-intensity model, which expand the industrial technique to information system field. We contribute to the further research on default prediction model monitoring and propose a new MEWMA scheme (Zou et al. 2007) for the nonlinear profiles. Different than most of the existing multiple chart approaches, the complexity of the proposed single chart approach will not increase as the profile parameters increase. Also, this chart can be designed and constructed easily and has satisfactory performance. Based on this, we propose a novel MEWMA scheme for the nonlinear forward model. This scheme can reduce the computational cost and improve the operational efficiency of CRI. In addition, the proposed scheme can be applied to other default prediction model and determine which parameters in the profile have changed.

Our empirical analysis uses a subset sample of the US exchange listed companies (both industrial and financial) covering 2,000 firms and over 24000 firm-month observations for the period from 1991 to 2011. We set the observation window as 120 months to observe and monitor the changes of parameters.

The remainder of the article is organized as follows. In next section we briefly introduce the forward intensity model and the maximum likelihood estimation process such as those described by Duan et al. (2012), and propose the Multivariate Exponential Weighted Moving Average statistic for the forward model in Section 3. In Section 4 we present our current stage of research, including the result of maximum likelihood estimation of forward-intensity model. We provide a detailed plan for the next stage of the research in last section, basically focus on data processing and results analysis.

## Methodology

### *The Forward-Intensity Model*

The forward-intensity model is a new reduced-form approach based on a forward intensity construction to estimate a firm’s default probabilities for different periods ahead (Duan et al. 2012). It considers two common factors and six firm specific attributes to characterize the forward intensity functions. They also consider the influence of state variables in terms of both level and trend, there are altogether two common factors and 10 firm specific attributes.

Let $X_{it} = (x_{it,1},\ x_{it,2}, \dots, x_{it,k})$ be the set of the state variables (stochastic and/or deterministic) that affect the forward intensities for the $i$-th firm. These variables may include two types of variables: macroeconomic factors and firm-specific attributes. As presented in Duan et al. (2012), we have

$$f_{it}(\tau) = exp\big(\beta_0(\tau) + \beta_1(\tau)x_{it,1} + \beta_2(\tau)x_{it,2} + \cdots + \beta_k(\tau)x_{it,k}\big) \tag{1}$$

They define $f_{it}(\tau)$ as the forward intensity function of $X_{it}$, and the

$\beta(\tau) = (\beta_0(\tau), \beta_1(\tau), \beta_2(\tau), \dots, \beta_k(\tau))$ is the parameters of state variables respectively. For convenience, we write the formula as

$$f_{it}(\tau) = exp(\beta(\tau) * X_{it}{}') \tag{2}$$

Which has similar part with the general linear profile.

Duan et al. (2012) adopt two macroeconomic factors and six firm-specific attributes on companies' one-month forward default probabilities, which are supposed to be the most relevant factors to corporate default. In fact, these parameters change a little in the absence a severe economic crisis or market movements from month to month. Therefore, we only need to detect and identify the out-of-control parameters.

## *Parameter Estimates*

There are eight parameters in the forward intensity model to predict the default probability, including two common factors (SP500 and treasure rate) and six firm-specific attributes (Distance-to-Default, CASH/TA, NI/TA, Size, M/B, SIGMA). Therefore, we are going to monitor these parameters

Before we compute the MEWMA charting statistic, we need to obtain the estimator of the parameters. In this article, we estimate the parameter, $\beta(\tau)$, using the maximum likelihood method presented in Duan (2000) and Duan et al. (2012). Whereas they consider both default and other exits of company in the forward intensity prediction. For convenience, we adopt the pseud-likelihood function in Duan et al. (2012) but only deal with the case of default. The estimation and monitoring process for the intensity function corresponding to other exit can be derived similarly.

They classify all the firm-month observations into three categories:

$$X^0 = \left(x_1^0, x_2^0, \dots, x_{N_0}^0\right)',\ X^1 = (x_1^1, x_2^1, \dots, x_{N_1}^1)',\ X^2 = (x_1^2, x_2^2, \dots, x_{N_2}^2)' \tag{3}$$

$X^0$ contains all the surviving firm-month observations, $X^1$ contains all the default observations and $X^2$ has all the observations for other exit. There are altogether $N_0$, $N_1$ and $N_2$ observations for each category. Variables for each firm-month observation forms a row vector $\boldsymbol{x_i^j}$ and they assume the first variable is always a constant 1 corresponding to the intercept term in the intensity function. The pseudo log-likelihood function is expressed as follows:

$$L = -\sum_{i=1}^{N_0} \exp\big(\boldsymbol{x_i^0\beta}\big)\Delta t + \sum_{i=1}^{N_1} \log\big\{1 - \exp[-exp\big(\boldsymbol{x_i^1\beta}\big)\Delta t]\big\} - \sum_{i=1}^{N_2} exp\big(\boldsymbol{x_i^2\beta}\big)\Delta t \tag{4}$$

where

$$f_i^j = exp\Big(\boldsymbol{x_i^j\beta}\Big)\Delta t \tag{5}$$

And we have

$$L = -\sum_{i=1}^{N_0} f_i^0 + \sum_{i=1}^{N_1} \log\big\{1 - \exp\big[-f_i^1\big]\big\} - \sum_{i=1}^{N_2} f_i^2 \tag{6}$$

Obviously, the likelihood function is so complex that we cannot write the expression of estimator $\hat{\beta}$ directly. We estimate it using the Damped Newton method, and get the minimum value by iterating.

Hence, we can obtain the parameters estimator $\hat{\beta}$ by using MATLAB, which is directly substituted into the next step of the calculation.

## A Multivariate Exponential Weighted Moving Average Chart for Forward Intensity Model

In this section we propose the MEWMA scheme for the forward intensity model. In the last section, we have obtained the estimator of parameter vector $\beta$ through maximum likelihood estimation, $\hat{\beta}$. In this step, we get the transformation of estimator $\hat{\beta}$ based on the central limit theorem. Note that this type of variance transformation with the use of an EWMA chart have been presented by Quesenberry (1995) and Chen, Cheng and Xie (2001). We define

$$Z_j = \sqrt{n}\,\frac{\hat{\beta} - E(\hat{\beta})}{\sigma} \tag{7}$$

This transformation has considerably nice properties. The distribution of $Z_j$ can be regard as a standard normal distribution as $n$ approaches infinity. And $\sigma$ represents the standard deviation of estimator $\hat{\beta}$. Here the MEWMA charting statistic is defined as

$$W_j = \lambda Z_j + (1-\lambda)W_{j-1}, \qquad \text{j=1,2,…,} \tag{8}$$

Where $W_0$ is a $(p+1)$-dimensional starting vector and $\lambda$ is a parameter (chosed such that $0 < \lambda \leq 1$) that regulates the magnitude of the smoothing. The staring vector $W_0$ is chosen to be the zero vector. We have

$$W_j = \lambda \sum_{n=1}^{j-1}(1-\lambda)^n\, Z_{j-n} + (1-\lambda)^j W_0 \tag{9}$$

namely

$$W_j = \lambda \sum_{n=1}^{j-1}(1-\lambda)^n\, Z_{j-n} \tag{10}$$

When we compute $W_j$ and obtain a new sample. Consequently, compute the plot statistic, $U_j$, and compare it with control limit $L\frac{\lambda}{2-\lambda}$. The chart signals are

$$U_j = W_j{}'\Sigma^{-1}W_j \tag{11}$$

and

$$\Sigma = \{\lambda[1-(1-\lambda)^{2n}]/(2-\lambda)\}\Sigma \tag{12}$$

In this article, the control limit, $L$, and the smoothing constant, $\lambda$, are fixed according to variation and distrinution of estimator $\hat{\beta}$. The analysis part has not been completely finished, so the subsequent results cannot be report yet. We will introduce our ongoing data processing work in the plans for completion.

## Current Stage of the Research

In this section, we introduce our research framework and the current stage of the research.

### *Framework*

Same as the computational procedure of general statistical process control chart, there are four steps completing the whole monitoring processes. Detailed implementation steps are as follows.

*Step 1*. Estimate the parameters of the forward intensity model using a moving window approach, which ensure each estimator are independent. We use the maximum likelihood method proposed in Duan et al. (2012) obtain the parameter estimator $\hat{\beta}$ with the case of default.

*Step 2*. Chose the smoothing constant, $\lambda$, and determine the control limit, $L$, based on the etimator distribution.

*Step 3*. Start monitoring the process, obtain the new transformed sample, $W_j$, and compute the plot statistic consequently, $U_j$. Then compare it with the control limit (including the upper control limit (UCL) and lower control limit (LCL)) for each parameter.

*Step 4*. Compare the performance of our new MEWMA chart with other Control chart, such as MaxEWMA chart (Chen, Cheng and Xie 2001). Verify its efficiency and significance.

### *Data*

Our data set is a subset of the US public firms over the period from 1991 to 2011 in Duan, Sun and Wang (2012), which is completely public on the website of Credit Research Initiative. The sample data contains two common factors and 10 firm specific attributes data for 2000 companies over 251 months.

### *Current Stage of Research*

We have completed the theoretical derivation and preliminary data analysis results up to now. That is the parameter estimation in step 1 and the statistics derivation in step 3. The estimation and derivation process have been presented in Section 2 and Section 3. As for the preliminary data analysis results, there are some error in the results caused by mistakes in the process of analysis, some of which is presented in the Appendix.

## Plans for Completion

Our further research plan is completing the rest part of the research framework. It is mainly the part of data analysis and the statistics computation. Detailed plans are as follows.

First, we will re-estimate the parameters of forward intensity model using the moving window method (Babcock et al. 2002). We set the window length as 120 months and estimate the parameters every time when the window moves. And we will obtain a parameter variation time series among 132 times (or windows), which can be used to the next observation.

Following this, we can calculate the mean value and variance of parameter time series. Based on this, we will compute the new series, $Z_j$, and then get the MEWMA charting statistic, $W_j$. Finally obtain the chart signals $U_j$ and draw the corresponding MEWMA control chart of forward intensity model.

Lastly, we will get the chart signal function and the control limit for each parameter. Accordingly, we can monitor each parameter of the forward intensity model monthly and identify the out-of-control parameter swiftly and effectively.

We intend to finish the above plans in one month and complete the whole paper as early as possible.

## References


Babcock, B., Datar, M., and Motwani, R. 2002, January. “Sampling from a moving window over streaming data,” *In Proceedings of the thirteenth annual ACM-SIAM symposium on Discrete algorithms. Society for Industrial and Applied Mathematics,* pp. 633-634.

Chen, G., Cheng, S. W., and Xie, H. 2001. “Monitoring process mean and variability with one EWMA chart,” *Journal of Quality Technology* (33:2), pp. 223-233.

Duan, J. C. 1994. “Maximum likelihood estimation using price data of the derivative contract,” *Mathematical Finance* (4:2), pp. 155-167.

Duan, J. C., and Van Laere, E. 2012. “A public good approach to credit ratings–From concept to reality,” *Journal of Banking & Finance* (36:12), pp. 3239-3247.

Duan, J. C., Sun, J., and Wang, T. 2012. "Multiperiod corporate default prediction—A forward intensity approach," *Journal of Econometrics* (170:1), pp. 191-209.

Joner, M. D., Woodall, W. H., Reynolds, M. R., and Fricker, R. D. 2010. "A one-sided MEWMA chart for health surveillance," *Quality & Reliability Engineering International* (24:5), pp. 503-518.

Kim, K., Mahmoud, M. A., and Woodall, W. H. 2003. "On the monitoring of linear profiles," *Journal of Quality Technology* (35:3), pp. 317-328.

Lowry, C. A., Woodall, W. H., Champ, C. W., and Rigdon, S. E. 1992. "A multivariate exponentially weighted moving average control chart," *Technometrics* (34:1), pp. 8.

Linderman, K., and Love, T. E. 2000. "Economic and economic statistical designs for MEWMA control charts," *Journal of Quality Technology* (32:4), pp. 410-417.

Niaki, S. T. A., Ershadi, M. J., and Malaki, M. 2010. "Economic and economic-statistical designs of MEWMA control charts—a hybrid Taguchi loss, Markov chain, and genetic algorithm approach," *International Journal of Advanced Manufacturing Technology* (48:1-4), pp. 283-296.

Niaki, S. T. A., Malaki, M., and Ershadi, M. J. 2011. "A particle swarm optimization approach on economic and economic-statistical designs of MEWMA control charts," *Scientia Iranica* (18:6), pp. 1529-1536.

Prabhu, S. S., and Runger, G. C. 1997. "Designing a multivariate EWMA control chart," *Journal of Quality Technology* (29:1), pp. 8-15.

Quesenberry, C. P. 2010. "On properties of Q charts for variables," *Journal of Quality technology* (27:3), pp. 184-203.

Zou, C., and Tsung, F. 2008. "Directional MEWMA schemes for multistage process monitoring and diagnosis," *Journal of Quality Technology* (40:4), pp. 407-427.

Zou, C., Tsung, F., and Wang, Z. 2007. "Monitoring general linear profiles using multivariate exponentially weighted moving average schemes," *Technometrics* (49:4), pp. 395-408.

## Appendix

Table 1. The preliminary data analysis results of parameter variation time series

| Windows / Factors | 1-120 | 2-121 | 3-122 | 4-123 | 5-124 | … | 132-251 |
|---|---|---|---|---|---|---|---|
| Intercept | -2.980 | -2.218 | -1.124 | -0.700 | -0.072 | | -3.431 |
| SP500 | -7.399 | -5.171 | -7.360 | -7.783 | -5.588 | | 0.661 |
| Treasury rate | 0.257 | 0.068 | -0.079 | -0.172 | -0.367 | | 0.074 |
| DTD_level | -0.788 | -0.802 | -0.833 | -0.892 | -0.866 | | -1.243 |
| DTD_trend | -0.462 | -0.441 | -0.536 | -0.637 | -0.595 | | -0.971 |
| CASH/TA_level | -1.781 | -1.589 | -1.972 | -2.374 | -2.156 | | 0.311 |
| CASH/TA_trend | -2.054 | -2.605 | -2.363 | -1.463 | -1.419 | | 0.346 |
| NI/TA_level | -0.618 | -0.683 | -1.286 | -1.179 | -1.591 | | 0.884 |
| NI/TA_trend | -1.291 | -1.281 | -1.058 | -1.630 | -1.594 | | -2.705 |
| SIZE_level | 0.128 | 0.169 | 0.148 | 0.117 | 0.112 | | 0.085 |
| SIZE_trend | -1.900 | -1.420 | -1.302 | -1.224 | -1.297 | | -1.383 |
| M/B | -0.125 | -0.119 | -0.215 | -0.142 | -0.176 | | 0.003 |