The Bayesian Forecasting of the Bridge Deflection Based on Constant Mean Discount Model

Shuangrui Chen and Quansheng Yan[*]

School of Civil Engineering and Transportation, South China University of Technology, Guangzhou, Guangdong, 510640, China

Abstract: Subject to various factors under loading, bridges appear to be discrete. Thus, it is unavoidable to take the practical bridge into consideration with regard to the bridge deflection forecasting. Given this, the Bayesian dynamic forecasting theory is introduced to forecast the bridge deflection. Since the bridge deflection can stay stable in a long term, create constant mean discount Bayesian conditional equation and observational equation and deduce the Bayesian posterior probability of the bridge deflection conditional parameters on the basis of the prior information of the parameters. With recursive deduction, the conditional parameters keep updating as observational data are imported. The results of Bayesian forecasting comprise values and confidence interval, which makes it more instructive. Finally, practical examples are adopted to examine the superior performance of Bayesian dynamic forecasting theory.

1. INTRODUCTION

The bridge deflection is an important parameter in terms of evaluating the safety performance of a bridge. It is closely related to the bridge load-bearing ability and too large deflection can directly influence the speed and comfort of traffic load on the bridge that is why it is important to be permanently monitored. Subjective factors such as external loading, internal prestress, concrete elasticity modulus and creep and shrinkage together with observational errors lead to discrete and uncertain examined deflection data.

Traditional means of forecasting including regression forecasting, parameters smoothing and linear time series forecasting create statistical models to obtain conventional forecasting, given the sample information and model information, and mechanically transform the input information into output information. The final result of these kinds of forecasting is just a specific value, which however, is quite uncertain and the limits of the uncertainty cannot be defined. Conventional means of forecasting treat the forward information and the recent information in the same way. However, as a matter of fact recent information is far more important than forward information [1].

Studying dynamic models, Bayesian dynamic forecasting is a time series forecasting means, the basic thought behind which is to integrate the forecaster's experience as known condition into practical model, create a dynamic model based on model, data and prior information (information on unknown parameters of the overall distribution) and deduce the Bayesian posterior probability of the state parameters [2].

This process is not about simply inputting data and calculating to get the final result; however it deduces from the obtained posterior distribution information, thus men are involved in the forecasting process, which makes the prior uncertain matters controllable. As a matter of fact the forecasting result will be more accurate.

This article samples finite observed deflection data, creates constant mean discount Bayesian model to forecast the deflection in the later stage and showcase the advantage of Bayesian dynamic forecasting theory regarding uncertain issues [3].

2. BAYESIAN FORECASTING THEORY

Bayesian forecasting does not only employ information of samples but also makes full use of the prior information of the samples. During the deduction, Bayesian forecasting requires that the prior information should be presented in the way of probability distribution of the unknown parameters (prior distribution), constantly correct prior distribution by importing sampling information to obtain posterior distribution and calculate the forecasting values. The forecasting mode is: prior distribution + sample information $\rightarrow$ posterior distribution [4]. See Bayesian forecasting recursive algorithm in Fig. (**1**). Here θ_t represents the state parameter at time t; y_t represents the observed data at time t; D_t represents all the effective information set before and at time t, $D_t=\{y_t, D_{t-1}\}$.

Starting with the basic theory of Bayesian statistics, Bayesian forecasting not only avails the prior data information but also incorporates forecaster's experience and judgment. Having integrated both subjective and objective factors, it is more flexible considering the unusual conditions. This forecasting can not only calculate the value but also the confidence interval, obtaining more instructive results [5-7].

Address correspondence to this author at the School of Civil Engineering and Transportation, South China University of Technology, Guangzhou, 510640, China; E-mail: cvqshyan@scut.edu.cn.
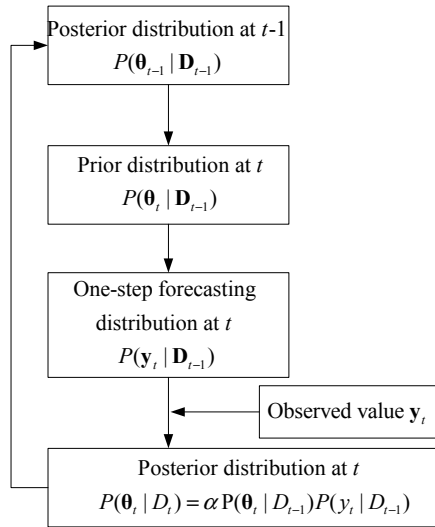
Fig. (1). Recursive steps of Bayesian dynamic model.

3. CONSTANT MEAN DISCOUNT MODEL

3.1. Bayesian Dynamic Linear Model

Same as conventional means of forecasting, Bayesian forecasting also needs to create a model. However, the difference lies in the fact that a dynamic model is required in the case of Bayesian forecasting. The dynamic linear model comprises a system of two equations, which can be described as: (1) how the observation of the process randomly relies on current state parameters; (2) how state parameters change as time increases, illuminating the dynamic changes and random disturbance of the interval system [1, 6]. Specifically, Bayesian dynamic linear model (BDLM) works on the basis of two hypothesis:
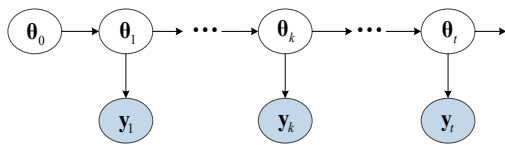

Fig. (2). Bayesian dynamic model.

(1) state variables $\boldsymbol{\theta}_t(t=1, 2,…,T)$ constitute a Markov chain, where the relationship between $\boldsymbol{\theta}_t$ and $\boldsymbol{\theta}_{t-1}$ is linear;

(2) observational variables $\mathbf{y}_t(t=1, 2,…,T)$ are independent of each other and $\mathbf{y}_t$ is only related to $\boldsymbol{\theta}_t$; the relationship between $\mathbf{y}_t$ and $\boldsymbol{\theta}_t$ is linear.

Fig. (**2**) is a BDLM sketch based on the above two hypothesis. For each $t(t=1,2,…, T)$, the model corresponds to a set of four elements: $\{\mathbf{F},\mathbf{G},\mathbf{V},\mathbf{W}\}_t=\{\mathbf{F}_t,\mathbf{G}_t,\mathbf{V}_t,\mathbf{W}_t\}$. Assume that both the observational variables and the state variables are random variables from the normal distribution, then for any time t [1]:

$$\begin{aligned}\mathbf{y}_t \mid \boldsymbol{\theta}_t &\sim N[\mathbf{F}_t\boldsymbol{\theta}_t,\mathbf{V}_t] \\ \boldsymbol{\theta}_t \mid \boldsymbol{\theta}_{t-1} &\sim N[\mathbf{G}_t\boldsymbol{\theta}_{t-1},\mathbf{W}_t]\end{aligned} \tag{1}$$

where $\mathbf{y}_t$ represents $r\times1$-dimension observational vector; $\boldsymbol{\theta}_t$ represents $n\times1$-dimension state parameter vector; $\mathbf{F}_t$ represents $n\times r$-dimension regression matrix; $\mathbf{G}_t$ represents $n\times n$-dimension state mobility matrices; $\mathbf{v}_t$ represents r-dimension normal zero mean observational error, the variance of which is $\mathbf{V}_t$; $\boldsymbol{\omega}_t$ represents n-dimension normal zero mean state error, the variance of which is $\mathbf{W}_t$. Equation (1) can be detailed as observational equation and state equation:

Observational equation:

$$\mathbf{y}_t= \mathbf{F}_t'\boldsymbol{\theta}_t+\mathbf{v}_t,\ \mathbf{v}_t \sim N[\mathbf{O},\mathbf{V}_t] \tag{2}$$

State equation:

$$\boldsymbol{\theta}_t=\mathbf{G}_t\boldsymbol{\theta}_{t-1}+\boldsymbol{\omega}_t,\boldsymbol{\omega}_t\sim N[\mathbf{O},\mathbf{W}_t] \tag{3}$$

Here both state variable $\boldsymbol{\theta}_t$ and observational variables $\mathbf{y}_t$ are univariates, thus,

$$\text{BDLM}\{\mathbf{F}_t,\mathbf{G}_t,\mathbf{V}_t,\mathbf{W}_t\}= \text{BDLM}\{F_t,G_t,V_t,W_t\}.$$

3.2. Constant Mean Discount Model

Constant mean discount model is a BDLM with $F_t=1,G_t=1$, an unknown W_t, and a discount factor δ. State observational equation and state equation are presented as [8, 9]:

Observational equation:

$$y_t=\theta_t+v_t,\ v_t \sim N[0,V_t] \tag{4}$$

State equation:

$$\theta_t=\theta_{t-1}+\omega_t,\ \omega_t\sim N[0,W_t] \tag{5}$$

Derived from objective uncertainty caused by the lack of knowledge, it is not easy to calculate W_t . To solve this problem, discount factor $\delta(0<\delta<1)$ is introduced in the model, $W_t=C_{t-1}(1/\delta-1)$, C_{t-1} represents the state variance at time t-1 [10].

As for unknown V_t, $\varphi_t= V_t^{-1}$ .Suppose $(\varphi_t|\theta_t)\sim\Gamma(n_t/2,d_t/2)$ then $\mathrm{E}(\varphi_t|\theta_t)= n_t/d_t=1/S_t$, for any large n_t , S_t nearly equals V_t. See the recursion of n_t and d_t in section 2.3. During the deduction, estimate and with regard to the examined sample information and correct state parameters with recursion equation [1, 11].

A remarkable feature of constant mean model is $\mathrm{E}[y_{t+k}|\theta_t]=\theta_t$, $\mathrm{E}[\theta_{t+k}|\theta_t]=\theta_t$, which is usually availed of in the forecasting of randomly fluctuating and steadily changing data. From the prior research experience, the author finds that the deflection can keep steady for a long time. Constant mean discount model forecasting is more accurate.

3.3. The Recursion and Correction of Constant Mean Discount Model

The recursion and correction process of constant mean discount model (observational variance V_t is unknown) is similar to Kalman filtering [12-14]. Below are the steps:

(1) the posterior distribution of state parameters , mean m_{t-1} , and variance C_{t-1} at time t-1.

$$\begin{cases}(\theta_{t-1} \mid \mathbf{D}_{t-1})\sim T_{n_{t-1}}[m_{t-1},C_{t-1}] \\ S_{t-1} = d_{t-1}/n_{t-1}\end{cases} \tag{6}$$

where $\mathbf{D}_{t-1}$ represents the set of all the effective information before and at time t-1. For large n_{t-1}, $T_{n_{t-1}}\left[m_{t-1}, C_{t-1}\right]$ nearly equals normal distribution $N\left[m_{t-1}, C_{t-1}\right]$, S_{t-1} nearly equals the observational variance V_{t-1}.

(2) the prior distribution of state parameters and one-step forward forecasting at time t.

$$\begin{cases} \left(\theta_{t-1} \mid D_{t-1}\right) \sim T_{n_{t-1}}\left[m_{t-1}, R_{t-1}\right] \\ \left(y_t \mid D_{t-1}\right) \sim T_{n_{t-1}}\left[f_{t-1}, Q_{t-1}\right] \\ n_t = n_{t-1} + 1 \\ d_t = d_{t-1} + S_{t-1} e_t^2 / Q_t \end{cases} \qquad (7)$$

where $R_t = C_{t-1}/\delta$ (state variance), $f_t = m_{t-1}$ (one-step forecasting result), $e_t = y_t - f_t$ (one-step forecasting error), $Q_t = S_{t-1} + R_t$ (one-step forecasting variance). According to the definition of HPD (the highest posterior density) confidence interval, the HPD confidence interval with 95% assurance rate is:

$$\left[f_t - 1.645\sqrt{Q_t}, f_t + 1.645\sqrt{Q_t}\right] \qquad (8)$$

(3) the posterior distribution of state parameters at time t

$$\begin{cases} \left(\theta_{t-1} \mid D_{t-1}\right) \sim T_{n_t}\left[m_t, C_t\right] \\ m_t = m_{t-1} + A_t e_t \\ C_t = \left(S_t / S_{t-1}\right)\left[R_t - A_t^2 Q_t\right] = A_t S_t \end{cases} \qquad (9)$$

where correction factor $A_t = R_t / Q_t$. It is easy to know that $0 < A_t < 1$.

3.4. Dereferencing of the Discount Factor

The discount factor δ is an important parameter. The larger the δ is, the more stable the model will be. Meanwhile the anti-intervention ability will be stronger. However, the state parameters are difficult to be corrected. The smaller the state parameters are, the easier it is to correct the state parameters. However at the same time the model is more sensitive to observational data and it is more disturbed by the noise. Thus, to obtain a reliable forecasting result it is important to select an appropriate δ. When it is not easy to determine the value of δ, conduct forecasting with models of different δ. Then compare the forecasting result and choose the most appropriate δ as the discount factor of the model. This article adopts mean absolute deviation *MAD* and mean square error *MSE* as the measurement of evaluating the forecasting performance.

Below are the calculating equations of *MAD* and *MSE* respectively:

$$MAD = \sum_{t=1}^{T} |e_t| / T \qquad (10)$$

$$MSE = \sum_{t=1}^{T} e_t^2 / T \qquad (11)$$

In the above equations, e_t represents one-step forecasting error, and T represents the data size. The smaller the MAD and MSE are, the better the forecasting performance will be. Assume the discount factors δ_1 and δ_2 can lead to the mini-

mum value then take the average of the two as the discount factor for the constant mean model.

4. EXAMPLES

Suppose Table **1** are the monitored deflection data of a middle-span in 24 successive months and the monitored frequency is once a month. Forecast the cumulative deflection when t=25.

Table 1.   Observed deflection data of middle-span.

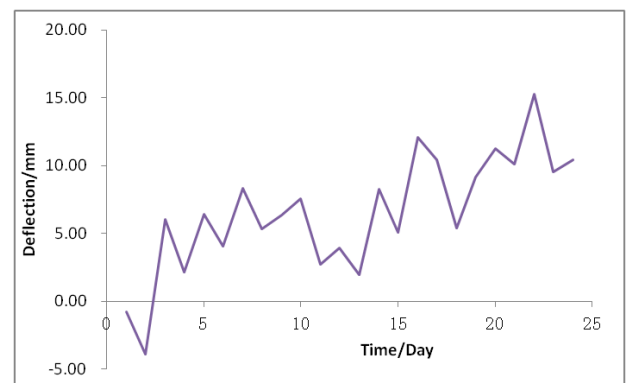| t (day)	Deflection (mm)	t (day)	Deflection (mm)
1	-0.78	13	1.93
2	-3.94	14	8.28
3	6.03	15	5.09
4	2.14	16	12.06
5	6.39	17	10.44
6	4.08	18	5.40
7	8.33	19	9.14
8	5.32	20	11.24
9	6.32	21	10.12
10	7.53	22	15.28
11	2.74	23	9.54
12	3.90	24	10.44

Fig. (3). Monitored deflection data of a middle-span.

Present the monitored data in the form of Fig. (**3**). As we can see the monitored data fluctuate randomly and change steadily. Thus, constant mean discount Bayesian model is suitable to be applied in the bridge deflection forecasting. The above has pointed out that state error variance W_t reflects the data's random change, and it can be expressed by means of discount factor δ. Choose 0.1, 0.2,…, 0.9 as the value of δ and suppose the initial prior distribution parameters m_0=0, C_0=100, n_0=1, d_0=100 (This supposition is quite blurry, but it can be seen from later recursion that the distribution parameters are corrected constantly). Calculate the

MAD and *MSE* of different δ values (see Fig. **2**), draw δ-*MAD* curve (Fig. (**4**)) and 和δ-*MSE* curve (Fig. (**5**)) and compare the forecasting performance of corresponding models. When δ_1=0.5, *MAD* has the minimum value; when δ_2=0.6, *MSE* has the minimum value. Then choose $\delta=(\delta_1+\delta_2)/2$=0.55 as the discount factor for the experimental model (Table **2**).

Table 2. **The comparison of the forecasting performance of different models.**

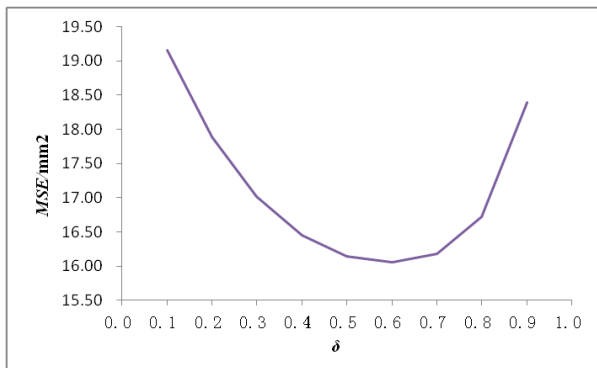| δ	MAD/mm	MSE/mm^2
0.1	3.50	19.15
0.2	3.27	17.89
0.3	3.12	17.02
0.4	3.04	16.46
0.5	2.99	16.15
0.6	3.01	16.05
0.7	3.08	16.18
0.8	3.22	16.72
0.9	3.54	18.39
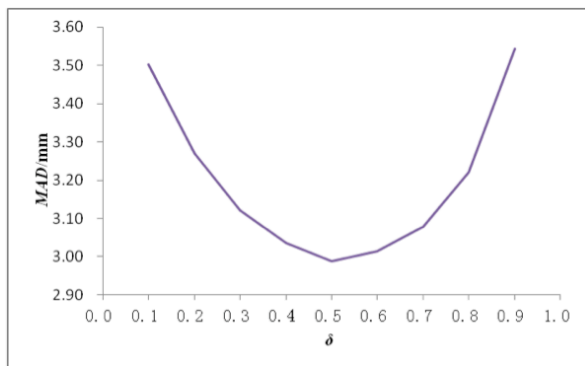
Table **3** is the state parameters' updating process when δ=0.55. As can be seen from the table that different from the regression coefficient in linear regression model where the coefficient is fixed, Bayesian state parameters (similar to regression coefficient) are wavy and keep updating, as new deflection data are imported. Meanwhile posterior variance C_t converges rapidly, which means that the importing of deflection information decreases the model's uncertainty. Also from Table **3**, when t=25, the forecasting result of cumulative deflection of the middle span of the bridge is 10.78, and the confidence interval of 95% HPD is [3.73, 17.82].

Further analyze the model. Fig. (**6**) is the curve of constant mean discount Bayesian model one-step forecasting variance Q_t as time changes with δ=0.55. Fig. (**7**) is constant mean discount Bayesian model one-step forward forecasting series and the confidence interval of 95% HPD (including the one-step forward forecasting when t=25). Fig. (**6**) shows that Q_t converges rapidly and tends to stabilize as time increases; Fig. (**7**) shows that the confidence interval of 95% HPD used to be wide and as time increases it converges rapidly and tends to stabilize; Table **3** shows that correction coefficient A_t converges rapidly and tends to stabilize in the recursion process. Generally speaking, the primary importing and updating of data have obviously improves the model's performance; finite monitoring includes enough update and define the model; the information updating has obviously improved the forecasting performance.
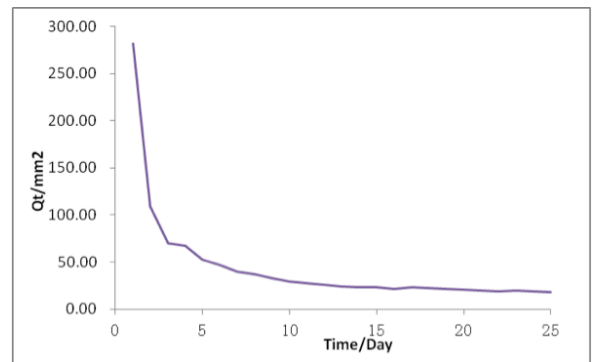

Fig. (4). δ-*MAD* curve.

Fig. (6). One-step forward forecasting variance $Q_t(\delta$=0.55).
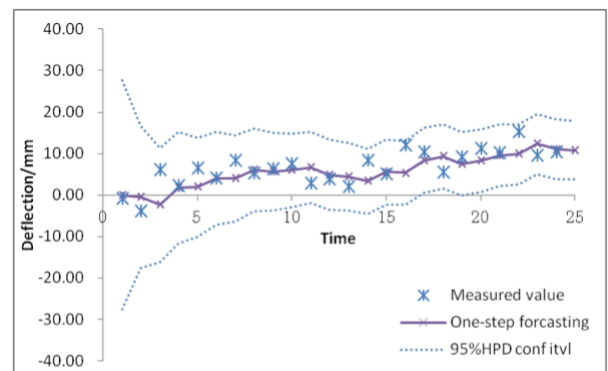

Fig. (5). δ-*MSE* Curve.

Fig. (7). One-step forward forecasting and the confidence interval of 95% HPD.

Table 3.   The updating process of state parameters (δ=0.55).

t (day)	Distribution parameters of one-step prediction			Correction factor A_t	Obs val y_t/mm	Error e_t/mm	Post information	
	Error Q_t/mm²	Mean f_t/mm	95%HPD conf itvl				m_t/mm	C_t/mm²
0	-	-	-	-	-	-	0	100
1	281.82	0.00	[-27.62,27.62]	0.65	-0.78	-0.78	-0.50	32.33
2	108.89	-0.50	[-17.67,16.66]	0.54	-3.94	-3.44	-2.36	19.01
3	69.78	-2.36	[-16.10,11.38]	0.50	6.03	8.39	1.80	17.48
4	67.07	1.80	[-11.68,15.27]	0.47	2.14	0.34	1.96	13.38
5	52.58	1.96	[-9.97,13.89]	0.46	6.39	4.43	4.01	11.71
6	46.58	4.01	[-7.22,15.24]	0.46	4.08	0.07	4.04	9.91
7	39.70	4.04	[-6.32,14.41]	0.45	8.33	4.29	5.99	9.18
8	36.92	5.99	[-4.01,15.98]	0.45	5.32	-0.67	5.69	8.14
9	32.81	5.69	[-3.74,15.11]	0.45	6.32	0.63	5.97	7.32
10	29.54	5.97	[-2.97,14.91]	0.45	7.53	1.56	6.67	6.70
11	27.06	6.67	[-1.88,15.23]	0.45	2.74	-3.93	4.90	6.46
12	26.09	4.90	[-3.50,13.30]	0.45	3.90	-1.00	4.45	5.98
13	24.15	4.45	[-3.63,12.54]	0.45	1.93	-2.52	3.32	5.66
14	22.88	3.32	[-4.55,11.19]	0.45	8.28	4.96	5.55	5.69
15	23.00	5.55	[-2.34,13.44]	0.45	5.09	-0.46	5.34	5.34
16	21.57	5.34	[-2.30,12.98]	0.45	12.06	6.72	8.37	5.68
17	22.96	8.37	[0.48,16.25]	0.45	10.44	2.07	9.30	5.43
18	21.92	9.30	[1.60,17.00]	0.45	5.40	-3.90	7.54	5.34
19	21.57	7.54	[-0.09,15.18]	0.45	9.14	1.60	8.26	5.10
20	20.62	8.26	[0.79,15.73]	0.45	11.24	2.98	9.60	4.96
21	20.06	9.60	[2.24,16.97]	0.45	10.12	0.52	9.84	4.74
22	19.16	9.84	[2.64,17.04]	0.45	15.28	5.44	12.29	4.85
23	19.61	12.29	[5.00,19.57]	0.45	9.54	-2.75	11.05	4.73
24	19.11	11.05	[3.86,18.24]	0.45	10.44	-0.61	10.78	4.54
25	18.36	10.78	[3.73,17.82]	0.45	-	-	-	-

Prior information is of paramount importance to model one-step forecasting at an early stage. However, as time increases prior information decays rapidly. For example, when t=1, correction coefficient A_1=0.65, m_1=0.65e_1+m_0=0.65y_1+0.35m_0; when t=2, A_2=0.41, m_2=0.41y_2+0.59m_1=0.41y_2+0.38y_1+0.21m_0. From t=1 to t=2, the weight of prior mean m_0 in posterior mean m_t has decayed from 0.35 to 0.21. See m_0 weight curve in Fig. (**8**). It can be seen that the weight of m_0 in m_t is monotonically decreasing as time increases. As a matter of fact, it is easy to get the weight $w_m(t)$=(1-A_1) (1-A_2) $\cdots$ (1-A_t) of m_0 in m_t from m_t recursion equation, $w_m(25)\approx0$ when t=25. As time increases, the prior information has few-

er and fewer influence over the forecasting and at a certain point the influence is close to zero.

We can also see from the above analysis that the weight of y_1 in m_1 is 0.65, however the weight of y_1 in m_2 decreases to 0.38. Fig. (**9**) is the weight curve of y_i in m_t (for the purpose of clearness only draw the weight curve with i=1,6,11,16,21). Fig. (**9**) shows that the weight of y_i monotonically decreases as time increases. It is easy to prove that the importance of the y_i weight $w_i(t)$=A_i(1-A_{i+1}) (1-A_{i+2}) $\cdots$ $\cdot$(1-A_t) in m_t ($i<t$)decreases as time increases. Thus, even if several unusual data might occur and intervene the model

accuracy in a short term as time increases the accuracy of the model can be restored. Compared with conventional forecasting method, Bayesian forecasting has the advantage of coping with unusual data which conventional static model fails to do and it also has better robustness.
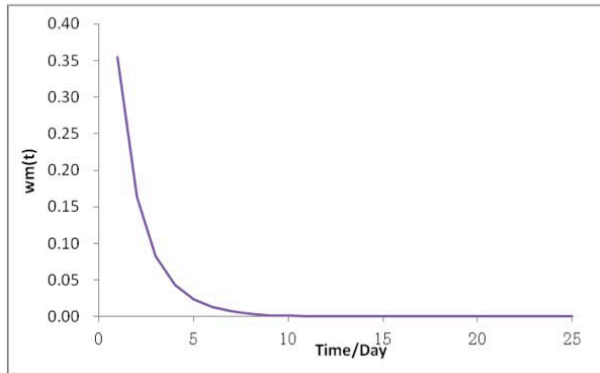

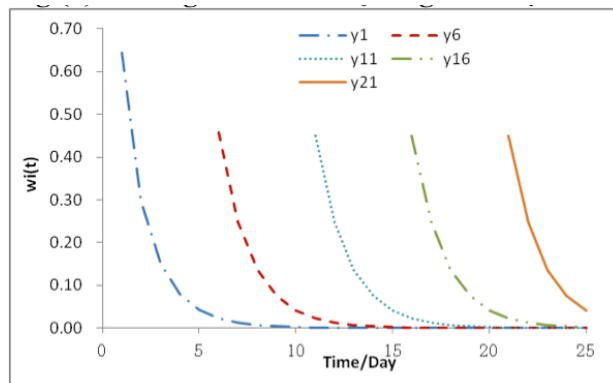**Fig. (8).** Change curve of m_0 weight in m_t

Fig. (9). Change curve of y_i in m_t

CONCLUSION

Starting with blurry prior information, Bayesian dynamic model keeps updating recursive data. The updating results have shown that the model performance has been greatly improved after primarily importing the deflection data for several times and parameters such as state variance, one-step forecasting variance, correction coefficient and the confidence interval of 95% HPD converge rapidly and tend to stabilize during the recursion process, resulting in more subjective forecasting results. The forecasting results include deflection value and confidence interval. This method of forecasting can quantify the forecasting uncertainty, which makes it more instructive.

Different from the conventional forecasting methods, which treat all the data in the same way, Bayesian forecasting is based on the fact that the importance of data decrease as time increases. Thus even if one or two unusual data may intervene the model's accuracy, as data cumulate the accuracy can be restored. It has better ability of self-correction and better robustness. As for the case with many unusual data, great error might occur if we only rely on the model's self-correction ability. In this case, we need to create Bayesian model monitoring mechanism to conduct feedback intervention. This matter needs further studies.

CONFLICT OF INTEREST

The authors confirm that this article content has no conflict of interest.

ACKNOWLEDGEMENTS

REFERENCES

[1]   X.L. Zhang, F.S. Liu, C.J. Zhang, and Y.X. Ge, "Bayesian dynamic model and forecasting," Shandong, pp. 136-152, 1992.

[2]   S.L. Ding, J. Z. Ding, and Y. Ji, "Bayesian statistical analysis of time-series forecasting model," *China Public Health*. vol. 18, pp. 1139-1140, 2002.

[3]   X. Chen, X. Zhang, and J. Yang, "Forecasting of Deflection for Prestressed Concrete Girders Using Bayesian Approach," *Journal of Wuhan University of Technology*, vol. 26, pp. 32-34. 2004.

[4]   W. Jian, L. Xila, "Life-cycle reliability management of structures". *Chinese Journal of Rock Mechanics and Engineering*, vol 24, pp. 3125-3130, 2005.

[5]   M. West, and P. J. Harrison, "*Bayesian Forecasting and Dynamic Models*," 2nd ed. New York: Springer-Verlag, 1997.

[6]   S. Russell, and N. Peter, "Artificial Intelligence: A Modern Approach," 3rd. Prentice Hall, 2009.

[7]   S. Mao, "Bayesian statistics," Beijing: China Statistics Press, 1999.

[8]   J.F. Jiang, "Different interval first-order polynomial model and Bayesian forecasting," *Statistics and Decision*, vol. 3, pp. 21-22, 2004.

[9]   L.W. Zhang, and H.P. Guo, *Introduction of Bayesian Network*, Beijing, 2006.

[10]   H.M. Zhu, Y.Q. Han, *Bayesian Multivariate Statistical Inference Theory*, Beijing, 2005.

[11]   X.P. Fan, and D.G. Lü, "Real-Time Reliability Forecast of Bridge Structures Based on Multiple BDLMs". *Journal of South China University of Technology (Natural Science Edition)*. 2013, vol. 41, pp. 70-75, 2013.

[12]   X.P. Fan, "Time-Dependent Reliability Assessment of Concrete Continuous Beam Bridge Based On Real-Time Monitoring Information," Harbin: Harbin Institute of Technology, 2010.

[13]   G.J. Wei, Y.M. Dang, and C.Y. Zhang, "Foundation settlement probability analysis and forecasting based on Bayes dynamic linear model," *Science of Surveying and Mapping*, vol. 37, pp.52- 53, 2012.

[14]   L.W. Zhang, H.P. Guo, *Introduction of Bayesian Network*, Beijing, 2006.