International Symposium on Safety Science and Engineering in China, 2012 (ISSSE-2012)

Forecasting of Production Safety Situation by Combination Model

Xu Yabo\textsuperscript{a,}* , Wang Tong\textsuperscript{a}, Song Bingxue\textsuperscript{a}, Xie Yushua\textsuperscript{a}

\textsuperscript{a}Beijing Municipal Institute of Labour Protection, Beijing, 10054, China

Abstract

The forecast of production safety situation is a complicated non-linear problem. The developmental change possesses no obvious trend of change over time and random fluctuation. Taking construction industry data as an example, four forecast models are adopted separately, namely the back-propagation neural network model, the moving average model, the exponential smoothing model and the combination model. Estimated results show that the combination forecast model can overcome the shortcomings of the single prediction model, and solve the forecast difficulties caused by random changes of the number of safety indicators of system status. The combination model is feasible for the forecast of the construction industry production safety situation. In fact, the production safe situation is affected by time, policies and other related factors. Decision makers should dialectically use the forecast result in the actual application and may carry on the weight adjustment to the forecast value.

© 2012 Published by Elsevier Ltd. Selection and/or peer-review under responsibility of the Capital University of Economics and Business, China Academy of Safety Science and Technology. Open access under CC BY-NC-ND license.

Keywords: neural network; moving average model; exponential smoothing model; combination model; production safety indicator; forecast

1. Introduction

Safety production control assessment index are often used to measure the safety production status and work level of government and enterprises. The death toll and number of serious accident are often used as control index and issued to all regions through administration of work safety, which are some problems\cite{1,2}. Although the overall situation of production safety has a relation with the death toll and number of serious accident, accident has contingency, randomness and sudden, which is not the only factor to characterize the safety situation. Some factors such as potential accidents, safety culture, law enforcement license and emergency support capabilities also have an impact on safety condition. Therefore, using comprehensive indicator system to evaluate and compare quantitatively safety situation has become increasingly accepted and recognized\cite{3}. This method can only achieve ex-post evaluation of production safety, can’t forecast the future trend.

On basis of previous work, the trend prediction technology for production safety situation is explored, which can provide forward-looking analysis for industry safety situation and scientific support for active response measures to prevent accidents.

* Corresponding author. Tel.: 010 63514176; fax: 010 63551975.
\textit{E-mail address:} xuyabo931@126.com

© 2012 Published by Elsevier Ltd. Open access under CC BY-NC-ND license.
doi:10.1016/j.proeng.2012.08.023
2. Prediction model

The prediction model established is key factor to decide prediction accuracy. Some single prediction models are selected from the commonly used methods(regression analysis, time series decomposition, trend extrapolation, exponential smoothing, stationary time series prediction method, gray theory, neural networks, etc.). Based on these single prediction method selected, combination forecasting model are created and the prediction error are compared between single and combination models.

Combination forecasting method is the combination of a variety of forecasting methods in order to receive a relatively narrow range of projected value range for systems analysis and decision-making[4].

Assuming the same prediction, there are $k$ ($k \geq 2$) different forecasting methods. The $th$ actual observed value is $y_t$, forecasting value or error of the $ih$ method is $f_{it}$ or $e_{it} (e_{it} = y_t - f_{it}; i = 1,2,\ldots,k; t = 1,2,\ldots,n)$. Weights of the $ih$ method in combination forecast is $\omega_i (i = 1,2,\ldots,k; \sum_i \omega_i = 1)$. The forecasting value or error of the $th$ method is $f_{ct}$ and $e_{ct} (t=1, 2, \ldots, n)$.

\[
f_{ct} = \sum_i w_i f_{it}
\]

\[
e_{ct} = y_t - f_{ct} = \sum_{i=1}^{k} \omega_i e_{it}
\]

Prediction error sum of squares for combination prediction method $e_c^2$ is

\[
e_c^2 = \sum_{t=1}^{n} e_{ct}^2 = \sum_{t=1}^{n} \sum_{j=1}^{k} \left[ \omega_i \omega_j \left( \sum_{t=1}^{n} e_{it} e_{jt} \right) \right] = \omega^T E \omega
\]

Where, $\omega = (\omega_1, \omega_2, \ldots, \omega_k)^T$ is combination weight vector. $E = (c_{ij})_{k \times k}$ is prediction error information matrix, calculated by $k$ kinds of individual prediction error. $c_{ij} = \sum_{t=1}^{n} e_{it} e_{jt}$.

Knew from the above equation, $e_c^2$ is relate to $E$ and $\omega$. $E$ is decided by $k$ kinds of prediction methods. When $E$ is given, $\omega$ is attained by linear programming method.

\[
\begin{align*}
\text{Min} & \quad e_c^2 = \omega^T E \omega \\
\text{s.t.} & \quad R_k^T \omega = 1 \\
& \quad \omega \geq 0
\end{align*}
\]

Where, $R_k$ is k-dimensional vector with all elements are all 1.

As long as the error of various individual forecasting methods are knew, the most effective weight vector $\omega$ can be calculated. $\omega$ multiplies single prediction value is combing forecasting result.

The model represents a special convex quadratic programming problem, which has the only optimal solution in the feasible territory boundary. Under normal circumstances, the solution expression can’t be directly got. The model can be transformed into linear programming on the basis of the Kuhn-tucker condition and calculate the exact solution, but the solution process is very cumbersome. Small amount of calculation, faster convergence, non-negative weight iterative algorithm of optimal combination forecasting are adopted in the paper[5,6].
3. Application examples

The moving average method can eliminate the random fluctuations of historical statistical series and identify the major trends, but the method doesn’t consider the impact of long-term data. All the observed values are considered in exponential smoothing method and different weights are given according to the proximity of the period, which can make the predicted values closer to the actual observed values. But, single exponential smoothing method is only suitable for the analysis of time series with smooth fluctuation and larger error for bigger fluctuation. Neural network method can better reflect the volatility of time series, but which also has its own inherent shortcomings. Forecasting result has a lot of randomness and larger prediction error for small amounts of data.

In this paper, taking 2010 and 2011 construction industry safety index in a certain county as an example. The prediction accuracy among neural network, moving average and exponential smoothing model are compared. The principle for selection is the minimum mean square error (MSE), simultaneously taking into account variability coefficient (\( \delta \)) and correlation coefficient (R).

Table 1 Foreasting of construction industry production safety indicator

| Month  | Safety indicator | Neural network | Moving average (t=3) | exponential smoothing (\( a=0.3 \)) |
|--------|-----------------|----------------|----------------------|-------------------------------------|
|        |                 | Prediction value | Error | Prediction value | Error | Prediction value | Error |
| 201001 | 0.8491          | 0.8491          | 0.0000  | 0.8491          | 0.0000  | 0.8643          | -0.0152 |
| 201002 | 0.8259          | 0.8259          | 0.0000  | 0.8259          | 0.0000  | 0.8597          | -0.0338 |
| 201003 | 0.9178          | 0.9369          | -0.0191 | 0.9178          | 0.0000  | 0.8496          | 0.0682  |
| 201004 | 0.9183          | 0.8127          | 0.1056  | 0.8643          | 0.0541  | 0.8700          | 0.0483  |
| 201005 | 0.7643          | 0.8127          | -0.0484 | 0.8874          | -0.1230 | 0.8845          | -0.1202 |
| 201006 | 0.7332          | 0.7691          | -0.0359 | 0.8668          | -0.1336 | 0.8485          | -0.1152 |
| 201007 | 0.9030          | 0.7926          | 0.1104  | 0.8053          | 0.0977  | 0.8139          | 0.0891  |
| 201008 | 0.6500          | 0.8127          | -0.1627 | 0.8002          | -0.1501 | 0.8406          | -0.1906 |
| 201009 | 0.7453          | 0.7960          | -0.0507 | 0.7621          | -0.0168 | 0.7835          | -0.0381 |
| 201010 | 0.7149          | 0.7664          | -0.0515 | 0.7661          | -0.0512 | 0.7720          | -0.0571 |
| 201011 | 0.7131          | 0.7954          | -0.0823 | 0.7034          | 0.0097  | 0.7549          | -0.0418 |
| 201012 | 0.7703          | 0.7954          | -0.0251 | 0.7244          | 0.0459  | 0.7423          | 0.0280  |
| 201101 | 0.6687          | 0.6712          | -0.0025 | 0.7328          | -0.0641 | 0.7507          | -0.0821 |
| 201102 | 0.8794          | 0.7960          | 0.0834  | 0.7173          | 0.1620  | 0.7261          | 0.1352  |
| 201103 | 0.8311          | 0.8127          | 0.0184  | 0.7728          | 0.0583  | 0.7721          | 0.0590  |
| 201104 | 0.8106          | 0.9369          | -0.1263 | 0.7930          | 0.0176  | 0.7898          | 0.0208  |
| 201105 | 0.9225          | 0.9369          | -0.0144 | 0.8403          | 0.0821  | 0.7960          | 0.1265  |
| 201106 | 0.8967          | 0.8127          | 0.0840  | 0.8547          | 0.0420  | 0.8340          | 0.0628  |
| 201107 | 0.8855          | 0.8127          | 0.0728  | 0.8766          | 0.0089  | 0.8528          | 0.0327  |
| 201108 | 0.7485          | 0.8127          | -0.0642 | 0.9016          | -0.1531 | 0.8626          | -0.1141 |
| 201109 | 0.7524          | 0.7893          | -0.0369 | 0.8436          | -0.0912 | 0.8284          | -0.0760 |
| 201110 | 0.7532          | 0.7673          | -0.0141 | 0.7954          | -0.0422 | 0.8056          | -0.0523 |
| 201111 | 0.7794          | 0.7655          | 0.0139  | 0.7514          | 0.0280  | 0.7899          | -0.0105 |
| 201112 | 0.9369          | 0.6691          | 0.2678  | 0.7617          | 0.1752  | 0.7867          | 0.1502  |

MSE 0.0075 0.0075 0.0078

\( \delta \) 0.1072 0.1075 0.1093

R 0.8933 0.893 0.8912
Three prediction results are shown in Table 1. Some things are need to explain. Two situation including two item and three item mobile are considered, however, calculation error of the three item is slightly higher, so only the calculation of the two item are given in table 1. Similarly, eight exponential smoothing coefficients including 0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7 and 0.8 are calculated respectively, and the predication value corresponding to minimum error are displayed in table 1.

Through iterative algorithm, weights of neural network, moving average model and exponential smoothing model are respectively 0.4878, 0.3778 and 0.1344. Calculation values and errors are shown in table 2 using combination model.

Seen from Table 1 and Table 2, MSE of combination model(0.006) is smaller than that of neural network(0.0075), moving averaging model(0.0075) and exponential smoothing model(0.0078). $\delta$ of combination model(0.099) is smaller than that of neural network(0.1072), moving averaging model(0.1075) and exponential smoothing model(0.1093). $R$ of combination model(0.9011) is larger than that of neural network(0.8933), moving averaging model(0.893) and exponential smoothing model(0.8912). Therefore, relative to the selected single models, the combination model is more applicable to industry trend forecast of production safety.

| Month | Safety indicator | Prediction value | Prediction error | Relative Error (%) |
|-------|------------------|------------------|------------------|--------------------|
| 201001 | 0.8491 | 0.8511 | -0.0020 | 0.2406 |
| 201002 | 0.8259 | 0.8305 | -0.0045 | -0.5499 |
| 201003 | 0.9178 | 0.9179 | -0.0001 | -0.0157 |
| 201004 | 0.9183 | 0.8399 | 0.0784 | 8.5418 |
| 201005 | 0.7643 | 0.8506 | -0.0863 | -11.2851 |
| 201006 | 0.7332 | 0.8167 | -0.0834 | -11.3809 |
| 201007 | 0.9030 | 0.8003 | 0.1027 | 11.3775 |
| 201008 | 0.6500 | 0.8117 | -0.1617 | -24.8712 |
| 201009 | 0.7453 | 0.7815 | -0.0362 | -4.8528 |
| 201010 | 0.7149 | 0.7670 | -0.0522 | -7.2978 |
| 201011 | 0.7131 | 0.7552 | -0.0421 | -5.9037 |
| 201012 | 0.7703 | 0.7615 | 0.0088 | 1.1467 |
| 201013 | 0.6687 | 0.7051 | -0.0365 | -5.4578 |
| 201014 | 0.8794 | 0.7569 | 0.1225 | 13.9267 |
| 201015 | 0.8311 | 0.7922 | 0.0389 | 4.6839 |
| 201016 | 0.8106 | 0.8628 | -0.0522 | -6.4383 |
| 201017 | 0.9225 | 0.8815 | 0.0410 | 4.4446 |
| 201018 | 0.8967 | 0.8314 | 0.0653 | 7.2828 |
| 201019 | 0.8855 | 0.8422 | 0.0432 | 4.8829 |
| 201020 | 0.7485 | 0.8530 | -0.1045 | -13.9621 |
| 201021 | 0.7524 | 0.8150 | -0.0627 | -8.3332 |
| 201022 | 0.7532 | 0.7831 | -0.0298 | -3.9602 |
| 201023 | 0.7794 | 0.7634 | 0.0159 | 2.0461 |
| 201024 | 0.9369 | 0.7199 | 0.2170 | 23.1616 |

Although overall prediction accuracy of combination model is better than single model, there are also larger errors for individual data. Seen from Table 2, relative error of the August 2010 is relatively high, reaching to 24.87%. The main reason is that the production safety indicator for the month has witnessed tremendous changes. For instance, production
safety indicator of July 2010 is 0.903 and that of August 2010 is 0.65. For the situation with larger data fluctuations, the model has some room for improvement. On the other hand, the decision-makers should dialectically use model output and adjust weight of prediction value according to the actual situation, such as considering the time, policies and other relevant factors.

4. Conclusions

By analyzing the structure and principle of combination forecast model, examples of the combined model are applied. The combination model can absorb the advantages of single models, which can well reflect the volatility of time series and maintain the stability of forecast performance. The prediction error of combination model has also been significantly improved than that of the single model. Combination forecasting model can be used to the actual forecast, in order to achieve the prior risk forecasting of production safety trends.

Acknowledgements

Supported by National key technology support program----City Safety Production Risk Dynamic Control Technology Research and demonstration (2011BAK01B00), china.

References

[1] Jing Yongqi. The suggestion for control target of safe production[J]. Labor protection, 2007(1):53-54.
[2] Zhi Tongxiang. Establishment and improvement of safe production control system[J]. Labor protection, 2005(1):38-39.
[3] LU Haiyan, Zhang Hongyuan, Wang Bianwen. Statistical indicators of reform in the production safety accident[J]. China Statistics, 2003(11): 17-19.
[4] LI Minlong, LI Jianping, HUANG Maoxing. Chinese provincial economic competitiveness evaluation and prediction research[M]. Beijing: social science Literature Press, 2007.
[5] YANG Guiyuan, TANG Xiaowo. A Superior Iterative Algorithm for optimal Combination Forecast of No-negative Weights[J].Operations Research and Management Science, 1996(2): 1-6.
[6] MA Yongkai, TANG Xiaowo, YANG Guiyuan. A study on basic theory of the optimal combined prediction method of non-negative weights[J]. Operations Research and Management Science, 1997(2):1-8.