ISSN 1 746-7233, England, UK
World Journal of Modelling and Simulation
Vol. 6 (2010) No. 1, pp. 57-62

The research on a double forecasting model of port cargo throughput

Liang Sun*
School of Transportation and Vehicle Engineering, Shan Dong University of Technology, Zibo 255049, P. R. China

(*Received September 30 2008, Accepted November 5 2009*)

Abstract. A double forecasting model based on conditional expectation is proposed through probability distribution of port cargo throughput. A compound variable, being composed of the number of cargo ships arriving at the port and the operating efficiency of handling equipment, is constructed. The port cargo throughput is the sum of all compound variables which indicated the cargo throughput each time when cargo ships arrived at the port. Probability distribution of cargo throughput was acquired using throughout probability theory. In view of such a fact, future port cargo throughput can be viewed as an conditional mathematic expectation The forecasting model is proposed using growing function. The cargo volume of a port in Shandong province was taken as an example. Theoretical analysis and case study shows that model based on conditional expectation is better than other model available with respect to forecasting port cargo throughput.

Keywords: port handling capacity, conditional expectation, throughout probability theory

1 Introduction

Port handling capacity is an important index that can evaluate efficiency of a port. Besides, it has been becoming significant and difficult subject to construct a model of port handling capacity to describe the developing trends of it in port planning.

Recently, the research on the scale and developing trends of port handling capacity focused on the application of predicative methods. Chen[5] proposed that logarithm second-index flatness can predict some approximate value of the port handling capacity. Yang suggested least square fitting with orthogonal function that predict some approximate value of the developing trends of port handling capacity. Yang didn't indicate factors which affected developing trends of port handling capacity[10]. Yuan et al. proposed a forecasting model based on Probability and delphi's method. The efficiency of Yuan's model depended on many subjective factors, which is difficult to forecast future port handling capacity exactly. Wang[1] proposed a forecasting model using Unascertained Number Regressive in order to overcome shortcomings of Chen and Yuans'model. Xu[2] investigated the macro-factors which can affected port handling capacity and proposed system dynamics methods which can predict port handling capacity approximately. Yuan et al. proposed a forecasting model based on Probability and delphi's method. The efficiency of Yuan's model depended on many subjective factors, which is difficult to forecast future port handling capacity exactly[4]. Given the validity of the collected data, Sun et al.[9] proposed a model based on grey theory. Sun et al analysis the developing trends of port handling capacity. At the same time, he proposed possibility distribution of port handling capacity and analysis its charesristics[8]. Eddie indicated that multi-factors regression model can predict the port handling capacity He took the volume of the port handling capacity of Hong Kong as an example in order to testify the effectiveness of his method[3]. William H. K. Lam adopted ANN (artificial neural network) to predict the port handling capacity[7].

The models available is classified into 2 types: some established forecasting model using time series of port handling capacity, others analysis developing trends of port handling capacity based on some affected

* Corresponding author. *E-mail address*: liang_sun76@163.com.

factors. Model available based on some affected factors is still limited in that some affected factors is too difficult to quantity. The proposed model can integrate two kinds of models efficiently. Affected factors can be taken into considered using growing function. Meanwhile, time series can also be considered by conditional expectation.

2 Establishment of a forecasting model for port handling capacity

2.1 Problem statement

A port has finished m TEU for a perioid, operation times of cargo ships arrived at the port is N. It is a part of stochastic progress $\{N(t),\ t \geq 0\}$. X_N TEU was handled $N(t)$ times at a certain time t. It is assumed that the port will finish n TEU. Our aim is to acquire the future port handling capacity n. Meanwhile, the gross port handling capacity is acquired.

2.2 Background information

2.2.1 Some definitions related to the model

(1) Definition of growing function

$$S(x) = P(X > x) = 1 - F(x), S(0) = 1, S(\infty) = 0, \tag{1}$$

mathematical expectation of X is as follows:

$$E(X) = \int_0^\infty xf(x)\mathrm{d}x = -\int_0^\infty x\mathrm{d}S(x) = -xS(x)\Big|_0^\infty + \int_0^\infty S(x)\mathrm{d}x = \int_0^\infty S(x)\mathrm{d}x. \tag{2}$$

(2) Conditional mathematical expectation of X

On the assumption that $f(x, y)$ is a combine consistency function of (X, Y), we get:

$$f(x) = \int_{-\infty}^{+\infty} f(x, y)\mathrm{d}y,\ g(x) = \int_{-\infty}^{+\infty} f(x, y)\mathrm{d}x.$$

Thus,

$$g(y|x) = \frac{f(x, y)}{f(x)},\ f(y|x) = \frac{g(x, y)}{g(x)}.$$

Conditional mathematical expectation of Y when $X = x$ is: $E(Y|X = x) = \int_{-\infty}^{+\infty} yg(y|x)\mathrm{d}y$. Similarity, $E(X|Y = y) = \int_{-\infty}^{+\infty} xg(x|y)\mathrm{d}x$.

2.2.2 Some theory related to the model

Lemma 1. *If* X, Y *are random variable, then* $E[E(X \mid Y)] = E(X)$.

Proof. (1) Let X, Y be continue random variables.

$$\begin{aligned} E[E(X|Y)] &= \int_{-\infty}^{+\infty} E(X|Y = y) \cdot f_Y(y)\mathrm{d}y = \int_{-\infty}^{+\infty} \left[\int_{-\infty}^{+\infty} x \cdot f(x|y)\mathrm{d}x\right] f_Y(y)\mathrm{d}y \\ &= \int_{-\infty}^{+\infty} \left[\int_{-\infty}^{+\infty} x \cdot \frac{f(x, y)\mathrm{d}x}{f_Y(y)}\right] f_Y(y)\mathrm{d}x\mathrm{d}y = \int_{-\infty}^{+\infty} x \left[\int_{-\infty}^{+\infty} f(x, y)\mathrm{d}y\right] \\ &= \int_{-\infty}^{+\infty} x \cdot f_X(x)\mathrm{d}x = E(x). \end{aligned}$$

(2) As for X, Y are discrete random variables, the results follows by (1).

Lemma 2. *Let $S_N = X_1+X_2+\cdots+X_N$, N is a uncertain positive integer. $\{X_k\}$ is a sequence of independent identical distribution. We get $M_{S_N}(t) = M_N[\ln M_X(t)]$.*

Proof. By lemma 1, we get:

$$M_{S_N}(t) = E(e^{tS_N}) = E\left[E(e^{tS_N}|N)\right] = E\left\{E\left[e^{t(X_1+X_2+\cdots+X_N)}\right]\right\} = E\left\{M_X(t)^N\right\}$$
$$= E\left[e^{N\ln M_X(t)}\right] = M_N[\ln M_X(t)].$$

Theorem 1. *Let ξ, C is a random variable and constant, respectively. Then $D(\xi) \leq E(\xi - C)^2$ hold.*

Proof.

$$E(\xi - C)^2 = E(\xi - E\xi + E\xi + C)^2 = E(\xi - E\xi)^2 + 2(E\xi - C)E(\xi - E\xi) + (E\xi - C)^2$$
$$= D\xi + (E\xi - C)^2. \quad (3)$$

This proves that $D(\xi) \leq E(\xi - C)^2$. $D(\xi) \leq E(\xi - C)^2$ hold when $E\xi = C$.

2.3 Modeling method

The sum of the port handling capacity is as follows: $S = X_1 + X_2 + \cdots + X_N$. Some notation used in the model is defined:

S The gross port handling capacity;

N It is a part of stochastic progress$\{N(t), t \geq 0\}$, which indicate the operation times of cargo ships arri- -ved at the port;

X_N It indicates thatX_NUnit was handled$N(t)$times at a certain time t.

We can acquire distribution function of the port handling capacity using Formula of Total Probability:

$$F_s(s) = \sum_{n=0}^{\infty} P\{S \leq s|N = n\}P(N = n) = \sum_{n=0}^{\infty} P(X_1 + X_2 + \cdots + X_n \leq s)P(N = n)$$
$$= \sum_{n=0}^{\infty} F_nF_{n-1}\cdots F_1(s)P(N = n) = \sum_{n=0}^{\infty} F^{*n}(s)P(N = n). \quad (4)$$

At the same time, Distribution law of the port handling capacity is as follows: $f(s) = \sum_{n=0}^{\infty} p^{*n}(s)P(N = n)$.

The definition of $p^{*n}(s)$ is similar to $F^{*n}(s)$. Obviously, $S = S_1 + S_2$, where S_1 denotes that cargo volume in a port has been finished, S_2 denotes that cargo volume in a port will finish. The future port handling capacity is as follows:

$$E(S \mid S_1 = m) = E(m + n \mid S_1 = m) = m + \int_0^{\infty} S(x)\mathrm{d}x = m + \int_0^{\infty} S(x > m \mid x = m)\mathrm{d}x$$
$$= m + \int_0^{\infty} \frac{1 - F(x + m)}{1 - F(m)}\mathrm{d}x. \quad (5)$$

Thus,

$$n = \int_0^{\infty} \frac{1 - F(x + m)}{1 - F(m)}\mathrm{d}x. \quad (6)$$

The series of port handling capacity is as follows:

$$\mathrm{PHC}_{i+1} = \int_0^{\infty} \frac{1 - F(x + \mathrm{PHC}_i)}{1 - F(\mathrm{PHC}_i)}\mathrm{d}x. \quad (7)$$

Where PHC_{i+1} denotes port handling capacity in $i+1^{\mathrm{th}}$ year, PHC_i denotes port handling capacity in i^{th} year.

2.4 The characteristic of the model

Corollary 1. *Probability Distribution of cargo throughput each time is $p(x_i) = \pi_i$, x_i indicates the volume of cargo throughput i^{th} times. π_i indicates the probability that x_i TEU cargo was handled in the gross cargo throughput. The times that x_i TEU cargo was handling is N_i. N_i is subjected to the poisson distribution whose parameter is λ_i. We draw a conclusion that S_N subjected to the compound poisson distribution.*

Proof. We get Eq. (8) obviously:

$$S = \sum_{i=1}^{N} x_i N_i. \tag{8}$$

On the assumption: $\lambda = \sum_{i=1}^{N} \lambda_i$. N_i is independent each other, in virtue of moment generating function, we get:

$$M_S(t) = E(e^{tS}) = E(\exp(t\sum_{i=1}^{m} x_i N_i)) = \exp(\lambda(\sum_{i=1}^{m} \frac{\lambda_i}{\lambda} e^{tx_i} - 1)) \tag{9}$$

$\sum_{i=1}^{m} \frac{\lambda_i}{\lambda} e^{tx_i}$ is the moment generating function of the discrete distribution as follows:

$$p(x) = \begin{cases} \frac{\lambda_i}{\lambda}, & x = x_i, i = 1, 2, \cdots, m \\ 0, & x \neq x_i \end{cases} \tag{10}$$

Therefore, S is subjected to the compound passion distribution whose parameter is λ.

Corollary 2. *Probability Distribution of cargo throughput each time is $P(x_i) = \pi_i$. x_i indicates the volume of cargo throughput i^{th} times; π_i indicates the probability that x_i TEU cargo was loaded or unloaded in the gross cargo throughput. The times of x_i TEU cargo being loaded or unloaded is N_i. N_i subjected to the binomial distribution whose parameter is p. Therefore, S is subjected to the compound binomial distribution.*

The proof follows immediately from corollary 1.

2.5 Error analysis

Eq. (6) shows that port handling capacity which will finish is a function of port handling capacity being finished. On the assumption that $g(S_1)$ is also a function for forecasting port handling capacity. S is a actual value when $S_1 = m$. Error is $|S - g(S_1)|$. According to characteristic of deviation, our aim is to minimize $E[S - g(S_1)]^2$. Thus,

$$\begin{aligned} E[S - g(S_1)]^2 &= \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} [S - g(S_1)]^2 p(S, S_1) \mathrm{d}S \mathrm{d}S_1 \\ &= \int_{-\infty}^{+\infty} p_{S_1}(y) \left(\int_{-\infty}^{+\infty} [x - g(y)]^2 p_{S|S_1}(x \mid y) \mathrm{d}x \right) \mathrm{d}y, \end{aligned} \tag{11}$$

where, $p(x, y)$ is a combine consistency function of (S, S_1).

By theorem 1, we get Eq. (11) minimized. Theorem 1 and Eq. (11) show that model based on conditional expectation is an appropriate approach to forecasting port handling capacity.

Table 1. Port handling capacity in the port (ten thousands TEU)

Year	1981	1982	1983	1984	1985	1986	1987	1988	1989
Throughput	4.94	6.61	8.04	11.5	20.16	20.38	22.42	31.29	35.38
Year	1990	1991	1992	1993	1994	1995	1996	1997	1998
Throughput	46.6	57.98	73.05	93.48	120.91	152.65	197.14	252.73	306.63
Year	1999	2000	2001	2002	2003	2004	2005	2006	2007
Throughput	422.66	516.78	566.88	626.97	916.73	977.32	1003.27	1104.28	1309.03

Table 2. The distribution law of operation times

N	6	7	12
P	0.4	0.5	0.1

Table 3. The distribution law of cargo volume per time (unit:Ton)

X	53	61	72
P	0.5	0.3	0.2

Table 4. The distribution law of port handling capacity in this time interval

(1) s	(2) $p^{*0}(s)$	(3) $p^{*1}(s)$	(4) $p^{*2}(s)$	(5) $p^{*3}(s)$	(6) $f(s)$	(7) $F(s)$
1	0	0	0.62	0	0.170	0.29
2	0	0	0.37	0	0.270	0.67
3	0	0	0.14	0.46	0.320	0.85
4	0	0	0	0.29	0.464	0.86

3 Numeric results

The port handling capacity in a port of Shandong province was taken as an example. The port handling capacity in the port is shown in Tab. 1. In a certain short time interval, the Distribution law of operation times and cargo volume per time are shown in Tab. 2 and Tab. 3. The distribution law of port handling capacity is acquired using Convolution as shown in Tab. 4.

The total time is divided into some short time interval i. Let S_i indicate the cargo volume in time interval i, thus is the Convolution of S_i. Probability distribution of port handling capacity can be illuminated as follows: Fig. 1 shows that Port handling capacity can't increase forever for ever using Probability theory. Port handling

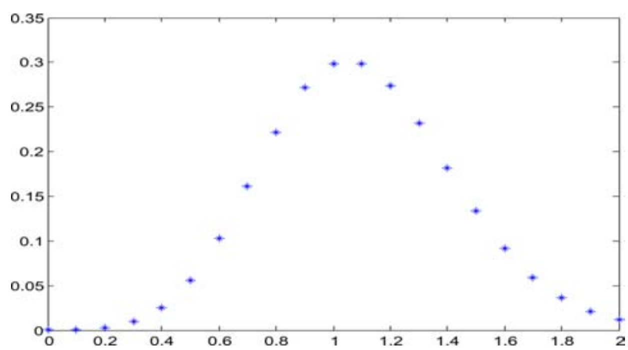

Fig. 1. Probability distribution of Port handling capacity for the port

capacity in 2002 is as follows:

$$\mathrm{PHC}_{2007} = \int_0^{\infty} \frac{1 - F(x + \mathrm{PHC}_{2001})}{1 - F(\mathrm{PHC}_{2001})} \mathrm{d}x = 1292.62 \,(\text{ten thousands TEU}).$$

Actual results in 2007 is 1309.03 ten thousands TEU, relative error is 1.28%. Aiming to further testify the efficiency of the model, We selected port handling capacity in the port from 2000 to 2006 to test using different models. The results is as follows:

Table 5. The comparison of the root mean PRESS

Year	LSI	LSF	SD	ANN	GM(1,1)	CME
2000 ~ 2006	2.97	3.33	3.26	2.86	1.96	1.02

Notes: LSI: logarithm second-index flatness;
LSF: least square fitting;
SD: system dynamics;
ANN: artificial neural network;
CME: conditional mathematic expectation.

Tab. 5 shows that model based on conditional mathematic expectation is an efficient one with respect to forecasting port handling capacity.

4 Conclusion

Here we may draw the following conclusions: (1) When the volume of cargo being handled each time is subjected to a discetetype distribution, at the same time, $N(t)$ is the poisson distribution or binomial distribution, then S is the compound poisson distribution or the compound binomial distribution; (2) When the volume of cargo being handled each time is subjected to a discetetype distribution, at the same time, $N(t)$ is the negative binomial distribution, then S is the compound negative binomial distribution; (3) If the expectation of $N(t)$ is very large, meanwhile, a distribution model can not be found to describe it, we hold that S is the normal distribution. The use of conditional mathematic expectation is an efficient approach for forecasting port handling capacity.

From the conclusion, it is shows that the curve of S and its probability is "bell type", that is, a peak value exists in the curve. As for future research, We'll focus on the relationship between the peak value and factors which affected it in order to promote the probability of the gross cargo throughput efficiently.

References

[1] C. Wang. Port Handling Capacity For ecasting Based on Unascer tained Number Regressive Model. *Communication Standardization*, 2006, **9**: 167–169. (In Chinese)

[2] C. Xu, Y. Yan, P. Zhang. Prediction Model of Port Throughput Based on System Dynamics. *Port & Waterway Engineering*, 2006, **32**(4): 26–29. (In Chinese)

[3] E. Hui, W. Seabrooke, K. Woong. Forecasting Cargo Throughput for the Port of Hong Kong: Error Correction Model Approach. *Journal of Urban Planning and Development*, 2004, **130**(4): 195–203.

[4] H. Yuan, Y. Xie. On Ports Throughput Probability For ecasting Model. *Port & Waterway Engineering*, 2007, **33**(5): 1028–1031. (In Chinese)

[5] N. Chen, M. Zhu, Z. Yu. Comparing Two Forecast Models of Ports Handling Capacity. *Journnal of Wuhan University of Technologuy*, 2005, **27**(9): 77–79. (In Chinese)

[6] L. Cheng, Y. Wong. *Port Facilities and Container Handling Services: The Hong Kong Economic Policy Studies*. City University of Hong Kong Press, 1997.

[7] W. Lam, L. Pan, et al. Forecasts and Reliability Analysis of Port Cargo Throughput in Hong Kong. *Journal of urban Planning and Development*, 2004, **130**(3): 133–144.

[8] L. Sun, Z. Wang. Studying the Probability Distribution of the port handling capacity in the port. *Proceedings of SSMSSD*, 2007, **6**: 1869–1873. (In Chinese)

[9] Y. Sun, G. Zheng. Research on Forecast of Port Throughput Capacity Based on Grey Theory. *China Water Transport*, 2007, **5**(4): 160–162.

[10] S. Yang, K. Guan. Application of Spreadsheet to Forecasting Port Throughput. *Journal of Wuhan University of Technology (Transportation Science & Engineering)*, 2005.