Use of Time Series Analysis to Forecast Truck Accidents

SNEHAMAY KHASNABIS AND SEUNG HWA LYOO

The purpose of this paper is to test the feasibility of using the Box-Jenkins method of time series analysis for forecasting truck accidents. Time series analysis is a technique by which the autocorrelation between sequential observations is analyzed and models are developed to produce forecasts. The authors used the Auto-Regressive Integrated Moving Average Method (ARIMA) in an effort to incorporate seasonal fluctuations in the data base to develop the model. A total of 88 data points representing monthly accidents involving large trucks in Michigan, observed between January 1978 and April 1985, was used to develop the model. Two types of checks were used to test the goodness of fit of the model. First, diagnostic checks were conducted to test the degree of correspondence between the observed data (used for model development) and the model output. This test indicated that approximately 70 percent of the autocorrelations are accounted for in the model. Second, the model was used to forecast monthly accident data for the 20-month period between May 1985 and December 1986 (data base not used in model development). The forecast data were then compared with the actual truck accidents observed during the same period. This test showed excellent correspondence between the observed data and the model output. The authors recommend further studies to test the feasibility of using time series analysis as an accident prediction tool.

A time series is a set of observations generated sequentially in time, in either a continuous or a discrete form. The values of observations at different time points are not assumed independent. Rather it is assumed that there exists a pattern of autocorrelation between these sequential observations. In real life, a great deal of data in economics, business, engineering, and the natural and social sciences is found in the form of time series in which observations are dependent and the nature of the dependence itself is of great interest to researchers (*1*,*2*). Time series analysis is a technique by which the autocorrelation between these sequential observations is analyzed and models or mathematical formulations are developed to fit the data, which can be used to produce forecasts of time series that might be expected under various scenarios.

MODELS TO DESCRIBE TRAFFIC EVENTS

In accident analysis, forecasting accidents under alternative hypotheses has perplexed researchers for a long time. Accidents typically are considered random events, and the current literature indicates efforts by researchers to fit different types of mathematical distributions (ranging from stochastic to deterministic) to accident data in an effort to develop predictive models (*3*,*4*). The use of the Poisson distribution, for example, to describe the occurrence of accidents as random events is quite common. Indeed, researchers have applied the Poisson and the negative exponential distribution (as an outgrowth of the Poisson) in describing other traffic events as well, such as arrival of vehicles at isolated intersections and distribution of vehicular headways on rural highways.

The use of the Poisson function to describe traffic events requires the implicit assumption of randomness, in which the occurrence of any of the events is not likely to be influenced by the occurrence (or nonoccurrence) of the preceding event. In real life, however, many traffic events are not to be considered independent and thus the assumption of randomness becomes questionable. For example, the arrival of vehicles at a signalized intersection on a congested urban arterial is likely to be affected by arrivals upstream of the intersection. Similarly, as traffic volume on a rural intersection increases, vehicular headways are likely to be more dependent upon one another. The transition of traffic movement from a "free flow" regime to a "constrained" flow regime (as volume increases from light to heavy) was described originally by Schuhl through mathematical models (*5*). The Schuhl model, as it has been designated in the literature, treats the total flow in two separate components: one for the random flow and the other for the constrained flow. Similar efforts to describe traffic flow on two-lane highways in Indiana and North Carolina have been described by Grecco and Sword (*6*) and Khasnabis and Heimbach (*4*).

PURPOSE OF PAPER

The purpose of this paper is to demonstrate the applicability of time series analysis technique for forecasting traffic accidents. The purpose of the earlier discussion of random versus nonrandom events was simply to provide background information on the concept of dependence or independence of observation (whether of traffic accidents or vehicular headways) and how the phenomenon of dependence has been treated by researchers for the purpose of model development.

Time series analysis techniques were used during the past decade to evaluate the effectiveness of highway safety countermeasures. Wagenaar, Arnold, and others, for example, have used this technique to evaluate the effectiveness of mandatory seat belt laws and child restraint systems and the effect of a minimum-age law on alcohol-related accidents involving young drivers (*7*–*9*). In a 1985 study conducted for the National Highway Traffic Safety Administration, Arnold used time series techniques to demonstrate a reduction in fatal

Civil Engineering Department, Wayne State University, Detroit, Mich., 48202.

crash involvement among drivers affected by an increase in the drinking age in 13 states (*10*). Similarly, a time series technique was used by Wagenaar et al. in 1988 to evaluate the effect of mandatory seat belt laws in eight states (*9*). Wagenaar also used the Box-Jenkins intervention analysis method to assess the long-term effects of a raised drinking age on reducing motor-vehicle crash involvement among young drivers. He used multiple levels of comparison groups and multiyear time series designs to obtain accurate estimates of the effect of raising the drinking age and to ensure that the observed effects were, in fact, caused by changes in the drinking age (*8*).

Recent increases in fatalities resulting from truck accidents have caused researchers to question the relative role of trucks (particularly heavy trucks) in the incidence of traffic accidents. In addition, the passage of the 1982 Surface Transportation Assistance Act, which made it possible for heavier, longer, and wider trucks to operate on selected national highways, has raised concerns in the minds of many safety experts. During the coming decade trucking activity is likely to increase further because of increased application of the "just-in-time" concept of delivery techniques by the manufacturing industry and because of price competition, brought about by the deregulation of the trucking industry. As such there is considerable interest among researchers to develop forecasts of truck accidents over a specific geographic area. This paper is an effort to apply time series analysis techniques to forecast truck accidents, notwithstanding the argument of dependence versus randomness presented earlier.

TIME SERIES MODEL

Time series analysis postulates that future values have a probability distribution that is conditioned by a knowledge of past values; therefore, exact predictions are impossible. Moreover, the reliability of prediction values depends on the characteristics of the data, so the data may have to be modified appropriately.

A distinction has been made in the literature on time series analysis between stationary and nonstationary models (*1*). In stationary models, an assumption is made that the process remains in equilibrium about a constant mean. In this category three broad types of models can be identified. Nonstationary models, on the other hand, exhibit a pattern in which observations do not vary about a fixed mean.

Autoregressive Models

Autoregressive (AR) models are essentially stochastic models in which the current or future value of the process is expressed as a finite linear combination of previous values of the process and a shock (error term). Mathematically, the process (X_t) of order p can be represented as

$$X_t = m_1 X_{t-1} + m_2 X_{t-2} + \ldots + m_p X_{t-p} + A_t \qquad (1)$$

where

m_1, m_2 = model parameters,
X_{t-1}, X_{t-2} = observation at time $t - 1$, $t - 2$, etc., and
A_t = random shock (error).

In Equation 1, the variable X_t is regressed on previous values of itself, hence the name "autoregressive."

Moving Average Method

In the moving average (MA) method, the current value of the process is expressed as a linear combination of previous random shock values. Mathematically, the process (X_t) of order q can be expressed as

$$X_t = A_t - n_1 A_{t-1} - n_2 A_{t-2} - \ldots - n_q A_{t-q} \qquad (2)$$

where n_1, n_2 are model parameters and A_t, A_{t-1} are random shock at time t, $t - 1$, and so on.

Mixed ARMA Method

It is sometimes advantageous to combine the AR and MA processes to achieve greater flexibility in fitting actual time series. This leads to the ARMA model, which can be expressed for the process (X_t) of order p,q, mathematically as

$$X_t = (m_1 X_{t-1} + m_2 X_{t-2} + \ldots + m_p X_{t-p}) + (A_t - n_1 A_{t-1} - n_2 A_{t-2} - n_q A_{t-q}) \qquad (3)$$

In real life, representation of actually occurring stationary time series can be obtained by AR, MA, or ARMA methods in which p and q are not greater than 2 and often less than 2. However, in business, engineering, industry, and economics, where forecasting has been of particular importance, time series data can be better represented as nonstationary, having no natural mean yet exhibiting homogeneous behavior of a kind. It is possible in such cases that the general level about which fluctuations are occurring may be different at different times. However, the broad behavior of the series, when differences in level are accounted for, may be similar.

Auto-Regressive Integrated Moving Average Method

A commonly used time series model is the Auto-Regressive Integrated Moving Average (ARIMA) Method, which provides analysts with a powerful tool for describing stationary and nonstationary processes. Mathematically, an ARIMA model of order p,d,q, can be represented as

$$W_t = \overline{V}^d X_t \qquad (4)$$

$$W_t = [m_1 W_{t-1} + m_2 W_{t-2} + \ldots + m_p W_{t-p}] + [A_t + n_1 A_{t-1} - n_2 A_{t-2} + \ldots - n_q A_{t-q}] \qquad (5)$$

where

W_t = set of time series removed from trends or seasonal effects,
$\overline{V}$ = differencing operator such that

$$\nabla X_t = X_t - X_{t-1} = (1 - B)X_t \tag{6}$$

B = backward shift operator such that $BX_t = X_{t-1}$, hence $B^{12}X_t = X_{t-12}$, and

∇^2 = second-order differencing such that

$$\nabla^2 X_t = \nabla(\nabla X_t) = \nabla(X_t - X_{t-1}), \tag{7}$$

hence

$$\begin{aligned}\nabla^2 X_t &= \nabla(\nabla X_t) = \nabla(X_t - X_{t-1}) \\ &= (X_t - X_{t-1}) - (X_{t-1} - X_{t-2}) \\ &= X_t - 2X_{t-1} - X_{t-2} = (1 - B)^2 X_t\end{aligned}$$

hence

$$\nabla^d X_t = (1 - B)^d X_t \tag{8}$$

Experience has shown that homogeneous nonstationary behavior can be represented by a model that calls for the dth difference of the process to be stationary. In most cases, d is usually 0, 1, or at most 2.

MICHIGAN CASE STUDY

Truck travel as well as truck-related accidents have grown steadily in Michigan and nationwide during the last 20 years. Further, many analyses predict that this trend is likely to continue. The purpose of the analysis presented in this paper is to develop forecasts of future truck accidents in Michigan based upon past observation through the use of time series analysis techniques.

The project from which this paper is developed used two separate data bases as follows:

Data 1: Number of large-truck accidents (gross vehicle weight > 10,000 lb) per month in Michigan for the 88-month period from January 1978 to April 1985.

Data 2: Number of trucks involved in accidents per year in Michigan between 1966 and 1986.

In this paper, only the analysis pertaining to Data 1 is reported. This is because the results of the forecast can be compared with actual monthly accident data following April 1985. On the other hand, sufficient time has not elapsed to allow a comparison of the annual accident forecast with actual data after 1986 for Data 2. Thus, a goodness-of-fit test cannot be compiled for Data 2. Also, the Box-Jenkins method that serves as the analytic tool for this study recommends the use of a minimum of 50 data points for model development; the only way this minimum data requirement could be met was through the use of monthly accident data (Data 1). In addition, the monthly accident data (Data 1) clearly showed a strong seasonal component, which would have otherwise been masked if the monthly data were combined to make yearly counts. This factor further justified the use of the monthly data so that the model can capture seasonal trends. The necessary accident data for model development were obtained from a recent report compiled by the Michigan Department of Transportation (MDOT) on truck safety, revenue, and taxation (*11*). The testing of the model output was conducted with data compiled by the Technical Services Unit of the MDOT Traffic and Safety Division.

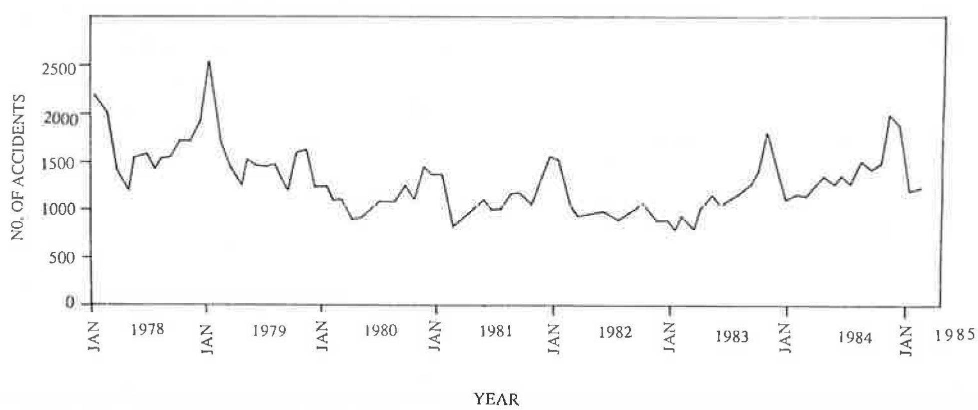
A five-step procedure was used to develop forecasts of truck accidents in Michigan. The well-known Box-Jenkins method of time series analysis was applied (*1*) and the software package MINITAB, developed by MiniTab, Inc. for developing the time series model (*12*), was used.

Data Plotting

Figure 1 from the above-mentioned MDOT study shows the trends in large-truck accidents during the 88-month period between January 1978 and April 1985. Figure 1 shows seasonal fluctuations; the number of accidents reached its highest level during the December-February period and the lowest level during the April-May period. This trend is referred to as the "seasonality factor" in the following text. Figure 1 also shows a long-term upward trend after 1982.

Model Identification

Considering the nonstationary nature of the data base, and in an effort to incorporate the seasonality factor, the authors identified the ARIMA method as the most appropriate approach. Further, following the Box-Jenkins method for data



SOURCE: REF. (2)

FIGURE 1 Trend in large truck accidents in Michigan.

sets not containing constant variance, a multiplicative (rather than an additive) model was used.

In Box-Jenkins, the variable W_t is formed from the original series X_t by differencing to remove both trend and seasonality. Because the data have monthly seasonality (Factor 12),

$$W_t = \nabla^d \nabla_{12}^D X_t \tag{9}$$

where for $d = D = 1$,

$$W_t = \nabla \nabla_{12} X_t = \nabla_{12} X_t - \nabla_{12} X_{t-1} \tag{10}$$

$$= (X_t - X_{t-12}) - (X_{t-1} - X_{t-13}) \tag{11}$$

(The values of the integers d and D do not usually need to exceed 2, according to the Box-Jenkins method.)

Therefore, the Box-Jenkins multiplicative seasonal models could be established by combining the following two processes:

For the seasonal process,

$$\oplus(B^s)\nabla_s^D X_t = \ominus(B^s) Z_t \tag{12}$$

where

s = seasonal period,
$\nabla_s = 1 - B^s$
$\oplus(B^s)$, $\ominus(B^s)$ = polynomials in B^s of degrees P and Q, and
Z_t = error.

For the whole process,

$$m(B)\, \nabla^d Z_t = n(B) A_t \tag{13}$$

so that, by combining above two equations, the following seasonal model, the ARIMA $(p,d,q) \times (P,D,Q)_s$ model, is obtained:

$$m_p(B)\oplus_P(B^s)\nabla^d \nabla_s^D X_t = n_q(B)\ominus_Q(B^s) A_t \tag{14}$$

Last, through the use of the MINITAB computer package and after more than 20 candidate models were reviewed, the ARIMA $(1, 1, 1) \times (1, 1, 1)_{12}$ model was identified as the most appropriate approach.

Estimation of Parameters

From Equation 14 and for $p = 1$, $P = 1$, $w = 1$, $Q = 1$, $s = 12$, $d = 1$, $D = 1$ [because the model is ARIMA $(1, 1, 1) \times (1, 1, 1)_{12}$],

$$(1 - m_1 B)(1 - \oplus_1 B^{12})\nabla \nabla_{12} X_t = (1 - n_1 B)(1 - \ominus_1 B^{12}) A_t \tag{15}$$

that is,

$$(1 - m_1 B)(1 - \oplus_1 B^{12})\nabla (X_t - X_{t-12}) = (1 - n_1 B)(1 - \ominus_1 B^{12}) A_t \tag{16}$$

Consequently, the equation could be modified as follows:

$$\begin{aligned} X_t = {} & (1 + m_1)X_{t-1} - m_1 X_{t-2} + (1 + m_1)X_{t-12} \\ & - (1 + m_1 + \oplus_1 + m_1\oplus_1)X_{t-13} \\ & + (m_1 + m_1\oplus_1)X_{t-14} - m_1 X_{t-24} \\ & + (m_1 + m_1\oplus_1)X_{t-25} + m_1\oplus_1 X_{t-26} \\ & + \bar{\mu} + A_t - n_1 A_{t-1} - \ominus_1 A_{t-12} + n_1\ominus_1 A_{t-13} \end{aligned} \tag{17}$$

The method of least squares was used in the MINITAB package to develop the following parameters: $m_1 = 0.2256$, $\oplus_1 = -0.3063$, $\bar{\mu} = \bar{\bar{x}} = 8.0876$, $n_1 = 0.9671$, $\ominus_1 = 0.8697$.

Diagnostic Checking

The purpose of the diagnostic check is to assess the degree of correspondence between the model output and the observed data for the 88-month period. This is generally done by examining the residuals, which are the differences between the observations and the fitted values. The following procedure suggested by Box-Jenkins was used:

$$Q = N \sum_{k=1}^{K} r_{k'}^{\,2} \tag{18}$$

where N is the number of data in the difference series and $r_{k'}$ is the autocorrelation function of the residual.

If the fitted model is appropriate, then Q should be approximately distributed as chi-square with $(K - p - q)$ degrees of freedom (df), where p, q are the number of orders in the AR and MA models, respectively. The values of Q for lags 12, 24, 36, and 48 were obtained as follows:

Lag	*Chi-Square (calc.)*	*df*
12	7.1	8
24	19.7	20
36	29.4	32
48	36.6	44

A comparison of the above-calculated chi-square values with critical chi-square values indicates that, overall, the model was built on 70 percent residual autocorrelation. In other words, approximately 70 percent of the autocorrelations are accounted for in this model. Thus the diagnostic check shows a reasonable fit to the observed data.

Forecasting

By substituting the estimated parameters, the following forecasting equation is developed:

$$\begin{aligned} X_{t-1} = {} & 1.2256X_{t-1} - 0.2256X_{t-2} + 1.2256X_{t-12} \\ & - 0.8502X_{t-13} + 0.1565X_{t-14} \\ & - 0.2256X_{t-24} + 0.1565X_{t-25} \\ & - 0.0691X_{t-26} + 8.0876A_t + 0.9671A_{t-1} \\ & - 0.8697A_{t-12} + 0.8411A_{t-13} \end{aligned} \tag{19}$$

By taking the conditional expectation approach, at time $t + h$ (h being the lead time), the model is rewritten as follows:

$$
\begin{aligned}
X_{t+h} = {} & 1.2256X_{t+h-1} - 0.2256X_{t+h-2} \\
& + 1.2256X_{t+h-12} - 0.8502X_{t+h-13} \\
& + 0.1565X_{t+h-14} - 0.2256X_{t+h-24} \\
& + 0.1565X_{t+h-25} - 0.0691X_{t+h-26} \\
& + 8.0876A_{t+h} - 0.9671A_{t+h-1} \\
& - 0.8697A_{t+h-12} + 0.8411A_{t+h-13}
\end{aligned} \tag{20}
$$

In Table 1, the expected values of the forecast data for 20 months starting in May 1985 through December 1986, along with the lower and upper 95 percent values, are presented. Also presented are the actual truck accident data experienced in Michigan during the same period and obtained from the records of MDOT. Figure 2 also shows the actual observations and the forecast data, along with the confidence band.

Table 1 provides a comparison between observed data and the model output and shows that in 15 out of 20 cases compared, the expected value of the forecast data is within 10 percent of the actual observation. Further, it is also seen that in only 2 out of the 20 cases the actual observations are beyond the limit of the 95 percent confidence interval of the forecast values.

Two other nonparametric tests were conducted to assess the goodness of fit of the model output. First, in Table 2, the results of the chi-square test are presented for comparing the distribution of the accident data as obtained from the two sources. The calculated value of the chi-square of 0.89 is much smaller than the critical chi-square value of 5.99, at a 5 percent level of significance for 2 df. This implies the absence of any significant difference between the two distributions. Second, in Table 3, the percent root-mean-square (RMS) errors of the estimated values from actual values (computed as a function of the deviations) are also presented. The range of the RMS error is between 4.8 and 11.3 percent, thus indicating excellent correspondence between the observed data and the model output.

Last, the confidence intervals of forecast values are expected to expand gradually as these values move further away from the time of the last real data value available. This is because one becomes less confident that the forecast value will approach the true value as one moves further into the future. The width of the confidence interval (difference between the upper and

TABLE 1 COMPARISON BETWEEN FORECASTED ACCIDENTS AND ACTUAL OBSERVATION

Year (Month)	Forecasted Number of Accidents: Expected Value	95 Percent Limit: Lower	95 Percent Limit: Upper	Width of Confidence Interval	Actual Number of Accidents
1985 (5)	1455	1108	1802	694	1494+
(6)	1566	1208	1924	716	1542+
(7)	1531	1171	1891	720	1585+
(8)	1595	1234	1955	721	1665+
(9)	1646	1285	2006	721	1612+
(10)	1818	1457	2179	722	2038+
(11)	1888	1527	2249	722	2010+
(12)	2076	1715	2437	722	2726*
1986 (1)	2098	1737	2460	723	2104+
(2)	1788	1426	2150	724	2084
(3)	1755	1393	2117	724	1585
(4)	1650	1288	2013	725	1619+
(5)	1836	1471	2202	731	1840+
(6)	1958	1592	2323	731	2006+
(7)	1930	1565	2296	731	1956+
(8)	2010	1644	2376	732	1883+
(9)	2035	1669	2401	732	2087+
(10)	2232	1866	2598	732	2251+
(11)	2263	1897	2629	732	2081
(12)	2419	2053	2786	733	1915*

* - In 2 out of 20 cases, actual observations are beyond the limit of the 95 percent confidence interval.
\+ - In 15 out of the 20 cases the expected value of the Forecasted Data is within 10% of the actual observation.

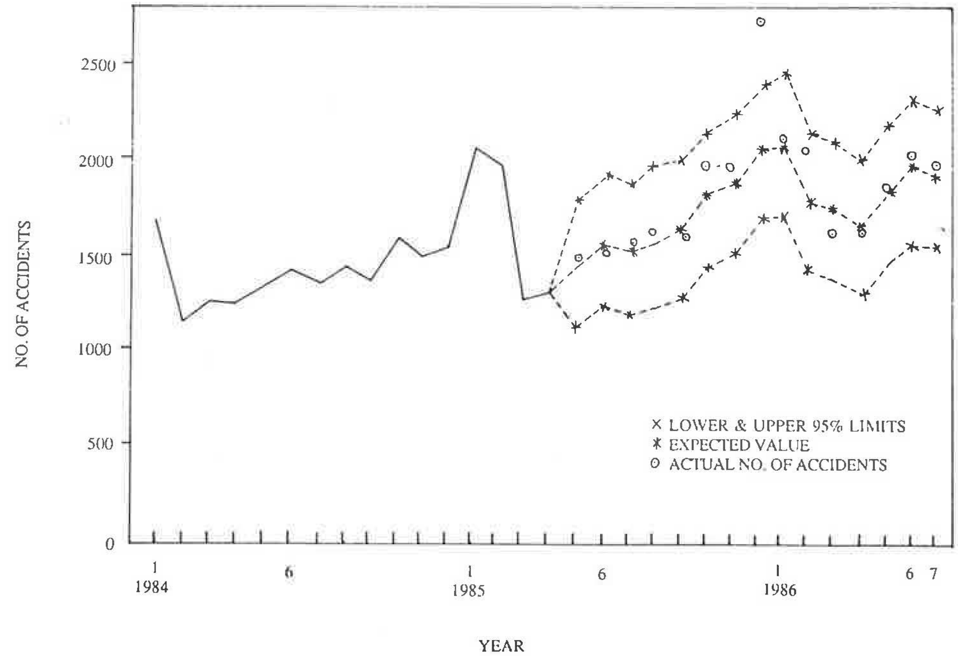


FIGURE 2 Forecasted trend of large truck accidents.

TABLE 2 COMPARISON OF DISTRIBUTION OF ACCIDENT DATA—CHI SQUARE TEST

Range (Accidents/Month	Number of Observations	
	Actual	Model
1450 - 1750	7	6
1751 - 2050	7	9
2051 and above	6	5

Calculated Chi-Sq Value = 0.89

Critical Chi-Sq Value for

2 df at 5% level of significance = 5.991

TABLE 3 COMPARISON OF ACTUAL VS. FORECASTED ACCIDENT DATA—RMS TEST

Range (Accident/Month)	n (# of observations	$\sum(\text{Deviation})^2$	RMS Error = $\sqrt{\sum(\text{Deviation})^2/n}$	Percent RMS Error
1450 - 1750	7	40,930	76.46	4.8%
1751 - 2050	7	321,434	214.28	11.3%
2051 and above	6	497,141	287.80	8.0%

lower percentile values) is also shown in Table 1. Although the increase in the width is not significant, there is a general trend toward an expansion of the confidence interval, as would normally be expected.

CONCLUSIONS

The purpose of this paper is to test the feasibility of using the Box-Jenkins method of time series analysis for forecasting truck accidents. Time series analysis is a technique by which the autocorrelation between sequential observations is analyzed and models are developed to produce forecasts. The authors used the ARIMA Method to develop the model because of the nonstationary nature of the data base. Also, the accident data used for developing the model reflected a strong seasonal component. This feature served as a strong motivation for using monthly accident data.

A total of 88 data points representing 88 monthly observations between January 1978 and April 1985 was used to develop the model. Two types of checks are presented in the paper to test the goodness of fit of the model. First, diagnostic checks were conducted to test the degree of correspondence between the observed data (used for model development) and the model output. This test indicated that approximately 70 percent of the autocorrelations are accounted for in the model. Second, the model was used to forecast monthly accident data for the 20-month period between May 1985 and December 1986 (data base not used in model development). The forecast data were then compared with the actual truck accidents observed during the same period. This test showed that in 18 of the 20 cases analyzed, the actual observations lie within the 95 percent confidence interval of the forecast values produced by the model. Further, in 15 of 20 cases, the expected values and the actual values are within 10 percent of each other.

In addition, two other nonparametric tests (chi-square and RMS) indicated excellent correspondence between the observed data and the model output. Because dependence of sequential observations is a basic assumption in times series analysis, an argument can be made against the use of this technique for accident prediction problems, because accidents are considered random events. However, when the data base reflects seasonal peaking (as in the case study presented), the feature of autocorrelation, although not necessarily reflecting dependence, may be effectively utilized in fitting a time series model. In the case study presented, the model developed appears to indicate a statistically significant correspondence between the observed data and the forecast values. This could be partly because of the seasonal peaking of the accident observations, indicating some degree of seasonal correlation, which may not necessarily be construed as dependence. The authors recommend further studies to test the feasibility of using time series techniques for accident forecasting problems.

Finally, a general comment is in order about the application of the ARIMA method in forecasting truck accidents. ARIMA models are designed to explain the stochastic autocorrelation structure of the series and to filter out any variance in the variable that is explainable on the basis of past history. The ARIMA method thus presents an advantage over standard regression techniques by implicitly taking into account autocorrelations within each variable rather than assuming that the error terms are independent (as is customarily done with ordinary least-squares regression) or by treating only the first-order autoregression. Wagenaar, in his article on the effect of macroeconomic conditions on motor vehicle accidents, has illustrated this point and has described how the above feature of the ARIMA method provides (*13*) "information on the time-ordered structure of the relationship, further increasing the degree of confidence in interpreting observed relationships in causal terms."

ACKNOWLEDGMENT

This paper is the result of research conducted at the Civil Engineering Department, Wayne State University, during 1987–1988. The authors thank the Michigan Department of Transportation and the Michigan Department of State Police for providing the necessary data base.

REFERENCES

1. G. E. P. Box and G. M. Jenkins. *Time Series Analysis, Forecasting and Control*. Holden Day, Inc., Oakland, Calif., 1970.
2. C. Chatfield. *The Analysis of Time Series: An Introduction*. Chapman & Hall, Ltd., London, 1985.
3. W. J. Campbell. Highway Safety Program Evaluation and Research. *Traffic Digest and Review*, Jan. 1970.
4. S. Khasnabis and C. L. Heimbach. Headway Distribution Models for Two-Lane Rural Highways. In *Transportation Research Record 772*, TRB, National Research Council, Washington, D.C., 1980, pp. 44–50.
5. A. Schuhl. The Probability Theory Applied to Distribution of Vehicles on Two-Lane Highways. In *Poison and Traffic*, Eno Foundation, Westport, Conn., 1955.
6. W. L. Grecco and E. C. Sword. Prediction Parameters for Schuhl's Headway Distribution. *Traffic Engineering*, Feb. 1968.
7. A. C. Wagenaar. *Alcohol, Young Drivers and Traffic Accidents: Effects of Minimum Age Laws*. D. C. Heath, Lexington, Mass., 1983.
8. A. C. Wagenaar. Preventing Highway Crashes by Raising the Legal Minimum Age for Drinking: The Michigan Experience 6 Years Later. *Journal of Safety Research*, Vol. 17, 1988, pp. 101–109.
9. A. C. Wagenaar, R. G. Maybee, and K. P. Sullivan. Mandatory Seat Belt Laws in Eight States: A Time Series Evaluation. *Journal of Safety Research*, Vol. 19, 1988, pp. 51–59.
10. R. Arnold. *Effect of Raising the Legal Drinking Age on Driver Inducement in Fatal Crashes: The Experience of Thirteen States*. National Highway Traffic Safety Administration, U.S. Department of Transportation, 1985.
11. *Background Report to the State Transportation Commission on Truck Safety Revenue and Taxation, and Truck Services and Highway Facilities*. Michigan Department of Transportation, Lansing, 1986.
12. *MiniTab Reference Manual: A General Purpose Data Analysis System*. MiniTab Inc., State College, Pa., 1988.
13. A. C. Wagenaar. Effects of Macroeconomic Conditions on the Incidence of Motor Vehicle Accidents. *Accident Analysis and Prevention*, Vol. 16, 1984, pp. 191–205.

The opinions and comments expressed in this paper are entirely those of the authors and not necessarily those of the Michigan Department of Transportation or the State Police.

Publication of this paper sponsored by Committee on Traffic Records and Accident Analysis.