



시계열 분석을 이용한 가스사고 발생 예측 연구

이수경 · [†]허영택* · 신동일** · 송동우 · 김기성

서울과학기술대학교 에너지환경대학원, *한국가스안전공사, **명지대학교
(2013년 11월 26일 접수, 2014년 2월 10일 수정, 2014년 2월 11일 채택)

The Study of Prediction Model of Gas Accidents Using Time Series Analysis

Su-Kyung Lee · [†]Young-taeg Hur* · Dong-Il Shin · Dong-Woo Song · Ki-Sung Kim

Graduate School of Energy & Environment, Seoul National Univ. of Technology,
Seoul 139-743, Korea

*Korea Gas Safety Corporatio, Gyeonggi-do, 429-712, Korea

**Dept. of chemical Engineering, Myongji University, Yongin, Gyeonggido 449-728, Korea

(Received November 26, 2013; Revised February 10, 2014; Accepted February 11, 2014)

요약

본 연구에서는 국내에서 발생한 가스사고를 분석하여 가스사고의 건수예측모델에 대하여 제시하였다. 가스 사고 건수를 예측하기 위하여 단순이동평균법(3,4,5기간), 가중이동평균법 및 지수평활법을 적용해 본 결과, 4기간 이동평균법과 가중이동평균법에 의한 모델의 평균오차제곱합이 44.4와 43으로 가장 정확성이 높은 것으로 나타났다. 가스사고 발생건수 예측시스템을 개발함으로써 가스사고 예방활동에 적극 활용할 수 있을 것이다.

Abstract - In this study, the number of gas accidents prediction model was suggested by analyzing the gas accidents occurred in Korea. In order to predict the number of gas accidents, simple moving average method (3, 4, 5 period), weighted average method and exponential smoothing method were applied. Study results of the sum of mean-square error acquired by the models of moving average method for 4 periods and weighted moving average method showed the highest value of 44.4 and 43 respectively. By developing the number of gas accidents prediction model, it could be actively utilized for gas accident prevention activities.

Key words : data mining, accident warning and prediction, meteorological data, gas accident data

1. 서론

국내에서 연료가스로 사용 중인 액화석유가스(LPG)와 액화천연가스(LNG)의 사용량은 매년 증가하고 있으나 국내의 가스 사고는 감소하지 않고 지속적으로 발생하고 있다. 사고의 내용에서도 시기별, 날씨 등에 따라 가스사용형태가 변하고 있어 가스의 사용 환경과 가스 사고는 밀접할 수밖에 없고

가스의 사용에 따른 가스사고도 사용시기와 날씨 등에 크게 좌우될 것이다. 이와 같이 인간생활과 날씨와의 관계는 상호 의존적이다. 날씨가 변하면 인간의 생각과 행동도 그에 따라 변한다. 인류는 기상과 날씨에 민감하게 적응하여 문화를 형성하고 성장하였다. 이처럼 기상은 우리의 삶 속에서 바늘과 실의 관계와 같다. 따라서 기상 상황에 따라 가스의 사용 형태와 사고발생 형태도 가변적이라 할 수 있다. 과거와 비교할 때 기상이 가지고 있는 현대적 의미는 매우 깊고도 넓다. 기상은 국민 개개인의 건강, 레저 등 생활뿐만 아니라 생산·소비와 같은 산업 활동까

[†]Corresponding author:hyt@kgs.or.kr

Copyright © 2014 by The Korean Institute of Gas

지 영향을 미친다[1].

기상청에서는 기상정보를 지수화하여 불쾌지수, 자외선지수, 산불위험지수 등 일반국민들에게 제공하고 있다. 또한, 산림청에서는 산불위험예보시스템을 통하여 산불위험지수를 제공하고 있으며 소방방재청에서도 국가화재정보시스템을 통하여 화재위험지수를 제공하고 있다.

국내외 논문을 살펴보면 이동민 등은 도로의 주변 환경 및 주행조건을 고려하여 주행조건 동질구간을 결정하고, 각 동질구간에서의 도로 및 교통조건 등을 고려하여 사고예측 모형을 개발하였다[2]. Wu Wen-feng 등은 해양 사고발생의 빈도를 분석하여 collision 모델과 powered grounding 모델을 활용한 정량적인 분석을 통해 정량적 예측모형을 개발하였다[3].

본 연구진은 선행 연구를 통하여 가스사고 예측모형을 개발하였다. 가스사고 예측모형은 통계적 기법을 활용하여 7가지 기상요소(평균기온, 최고기온, 최저기온, 상대습도, 운량, 강수량, 평균풍속)에 따라 회귀분석을 실시하여 가스사고의 영향도를 예측식을 도출하였다. 또한, 영향도를 지수화하여 5단계의 위험수준으로 도출될 수 있는 예측모형을 개발하였다[4]. 본 연구에서는 과거 가스사고사례 통계를 시계열분석기법을 적용하여 사고건수를 예측하였다. 선행연구를 통해 개발된 예측모형과 본연구에서 개발된 건수예측모형을 통해 가스사고를 사전에 예측하고 적절한 예방대책의 수립이 가능할 것이다.

II. 이론적 배경

가스사고의 발생가능성을 예측하기 위하여 시계열 예측기법을 적용하였다. 시계열 예측법이란 일정한 시간적 간격으로 나열된 과거의 자료에 나타난 수요의 패턴이 미래의 시간 간격에도 연속하여 적용된다는 가정에서 수요를 예측하는 기법이다. 이러한 시계열은 수년간의 과거 자료가 있으며 그 자료 속에는 추세가 분명하며 비교적 안정적일 때 적합한 방법이다. 이 시계열 자료는 시점과 양이 짝지어진 자료이다. 여기서 시점은 독립변수가 되며, 과거의 양은 종속변수가 된다. 이 예측기법은 환경조건이나 수량에 큰 변동이 있을 때 정확한 예측이 어렵기 때문에 수요패턴이 비교적 안정되어 있는 단기나 중기 예측에 적합하다[5,6]. 따라서, 시계열 분석을 통한 예측에서는 관측된 과거의 자료를 분석하여 규칙성을 발견하고, 이러한 규칙성이 미래에는 변화하지 않고 계속될 것이라는 가정 하에 이를 모형화 하여 추정하고, 추정된 모형을 이용하여 미래의 시계열을 예측하게 된다[7].

2.1. 이동평균법

이동평균법은 시계열 예측 기법 중 가장 쉽게 적용될 수 있는 방법으로 시계열 자료에 추세, 순환 변동, 계절적 변동이나 급격한 변화가 없고 우연변동만 존재하는 경우의 예측에 유용하게 적용될 수 있다. 또한 이동평균법은 시계열 변동 요인이 존재하는 경우에는 추세나 계절지수를 산정하는 데도 이용된다. 과거의 시계열 자료가 특정 추세 패턴을 나타낼 때 이동평균법에 의한 예측은 정확도가 떨어질 수 있지만 일반적으로 단순하고 응용하기 쉬워 관광 수요, 어획량예측 등 비교적 소규모의 수요나 실적 예측에 이용되고 있다[8].

2.1.1. 단순이동평균법

단순이동평균법(simple moving average method)은 과거의 자료에서 추세변동이나 계절적인 변동이 포함되지 않고, 설명할 수 없는 불규칙 변동만 크게 작용하는 경우의 예측에 적합한 기법이다. 이 때 이 불규칙 변동을 단순이동평균법에 의해 제거될 수 있다. 이 방법은 이동평균대상기간수에 따라 3기간이동평균, 4기간이동평균, 5기간이동평균 등으로 나눌 수 있다. 여기서 예측 값을 예측하려는 시점의 직전에 있는 일정기간의 실제건수 즉 예측하려고 하는 대상기간수의 실제건수를 동일한 비중으로 산술평균하여 계산한다. 단순이동평균법에 사용되는 식은 (1)과 같다.

$$F_t = \frac{\sum_{i=1}^n A_{t-i}}{n} \quad (1)$$

$F_t = t$ 기의 예측값
 $A_{t-i} = t-i$ 기의 실제량
 $n =$ 기간수

2.1.2. 가중이동평균법

가중이동평균법(weighted moving average method)은 이동 평균할 때 예측하려는 시점과 가까운 과거 자료에 더 가중치를 두어 예측하려는 방법이다. 단순이동평균법에서는 각 자료 값에 동일한 가중치가 부여되었다. 이에 비해 가중이동평균에서는 예측하려는 시점의 직전에서 n기간 동안 가중치의 합이 1이 되도록 부여한다.

예를 들어 시계열 자료에 추세 또는 패턴이 있을 때는 예측하려는 시점과 가까운 곳에 있는 자료가 더 미래의 예측 값을 잘 설명해 줄 수 있기 때문이다. 그래서 예측하려는 시점의 직전에 있는 자료에 높은

가중치를, 그리고 먼 과거의 실제 값에는 점점 낮은 가중치를 부여하는 방법이다.

가중평균에는 기간을 길게 하면 불규칙한 시계열 변동을 평준화하는 효과를 낼 수 있으나, 가중치의 결정을 예측자의 경험적인 판단에 의존해야 한다는 점에서 예측자의 주관이 개입될 여지가 많다. 가중 이동평균법에 사용되는 식은 (2)와 같다.

$$F_t = \sum_{i=1}^n W_{t-i} \cdot A_{t-i} \quad (2)$$

$F_t = t$ 기의 예측값
 $W_{t-i} = t-i$ 기의 가중치
 $A_{t-i} = t-i$ 기의 실제량
 $n = \text{기간수} \sum_{i=1}^n W_{t-i} = 1$

2.2. 지수평활법

지수평활법(exponential smoothing)은 미래를 예측하는데 있어 가중이동평균법과 마찬가지로 과거로 올라갈수록 자료의 중요성이 감소한다는 가정에서 과거 자료에 점점 낮은 가중치를 부여하는 방법이다. 그러나 이 방법은 가중이동평균법과는 달리 최근의 실제 값에 가중치의 대부분을 부여하며 과거의 자료가 올라갈수록 지수적(exponentially)으로 감소하는 것으로 본다. 이 방법은 이동평균법과 마찬가지로 시계열자료에 추세나 계절적 변동이 적을 때 유용하며, 다른 이동평균법이나 가중이동평균법에 비해 과거의 자료수가 적을 때도 적용할 수 있다는 장점을 가지고 있다.

이 방법에 의한 당기 예측 값의 계산은 그 시점에서 전기 예측 값에 실제 값과 예측 값 차이의 일정 비율을 더함으로써 이루어진다. 여기서 사용하는 일

정 비율은 평활계수(smoothing constant)를 말하며, 현재 관측된 수요에 대한 가중치를 나타낸다. 이것을 수식으로 일반화시키면 식 (3)과 같다.

$$F_t = a \cdot A_{t-1} + (1-a) \cdot F_{t-1} \quad (3)$$

$F_t = t$ 기의 예측값
 $F_{t-i} = t-i$ 기의 예측값
 $A_{t-i} = t-i$ 기의 실제량
 $a = \text{평활계수}$

위의 수식에 의해 당기 예측 값을 계산하는 경우, 최초의 예측 값이 주어져야 되고, 평활계수 a가 결정되어야 한다. 초기 예측 값은 그 이전의 과거 자료가 있으면 단순이동평균이나 가중이동평균으로 최초의 예측 값을 계산하지만 과거 자료가 없으면 정적인 방법에 의한 예측 값을 적용할 수 있으며 최초의 실적 값을 예측 값으로 이용하는 것도 하나의 방법이다[9].

III. 가스사고 발생건수 예측

가스사고의 발생 건수를 예측하기 위하여 시계열 분석을 실시하였다. 본 연구에서 사용된 시계열 분석기법은 이동평균법(3기간, 4기간, 5기간), 가중이동평균법 및 지수평활법이다. 분석결과와 정확도는 예측오차에 의하여 결정되며 오차가 작으면 작을수록 정확도가 높아지게 된다. 본 연구에서는 기법별로 나타난 평균제곱오차를 이용하여 정확도를 평가하였다.

3.1. 단순이동평균법에 의한 예측

단순이동평균법(moving average method)에 의

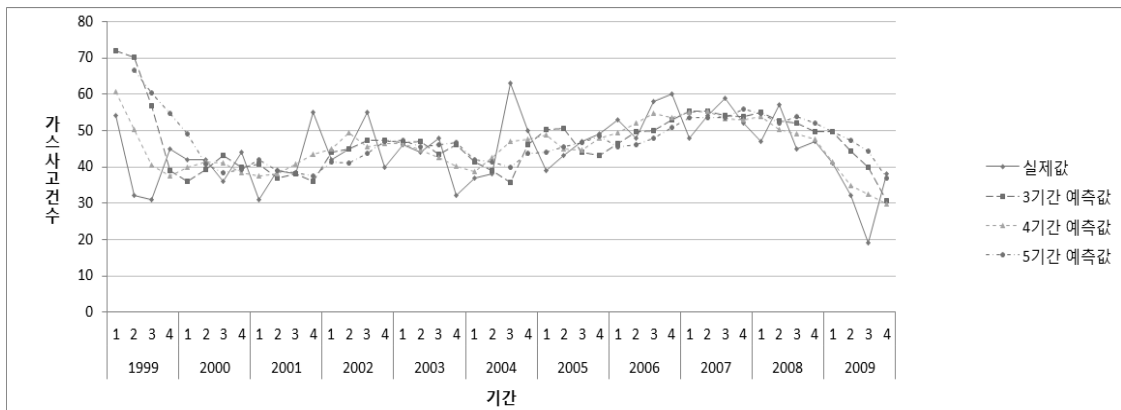


Fig. 1. Prediction of the number of accidents using moving average method.

Table 1. Prediction of the number of accidents using moving average method

연도	분기	사고 건수	3기간		4기간		5기간		연도	분기	사고 건수	3기간		4기간		5기간		
			예측치	예측 오차 자승	예측치	예측 오차 자승	예측치	예측 오차 자승				예측치	예측 오차 자승	예측치	예측 오차 자승	예측치	예측 오차 자승	
1998	1	63	#N/A	#N/A	#N/A	#N/A	#N/A	#N/A	2004	1	37	41.3	18.8	38.8	3.1	42.0	25.0	
	2	59	#N/A	#N/A	#N/A	#N/A	#N/A	#N/A		2	38	39.0	1.0	42.5	20.3	41.4	11.6	
	3	73	#N/A	#N/A	#N/A	#N/A	#N/A	#N/A		3	63	35.7	747.1	47.0	256.0	39.8	538.2	
	4	84	65.0	361.0	#N/A	#N/A	#N/A	#N/A		4	50	46.0	16.0	47.5	6.3	43.6	41.0	
1999	1	54	72.0	324.0	60.8	45.6	#N/A	#N/A	2005	1	39	50.3	128.4	48.8	95.1	44.0	25.0	
	2	32	70.3	1469.4	50.3	333.1	66.6	1197.2		2	43	50.7	58.8	44.8	3.1	45.4	5.8	
	3	31	56.7	658.8	40.5	90.3	60.4	864.4		3	47	44.0	9.0	44.5	6.3	46.6	0.2	
	4	45	39.0	36.0	37.5	56.3	54.8	96.0		4	49	43.0	36.0	48.0	1.0	48.4	0.4	
2000	1	42	36.0	36.0	40.0	4.0	49.2	51.8	2006	1	53	46.3	44.4	49.3	14.1	45.6	54.8	
	2	42	39.3	7.1	41.3	0.6	40.8	1.4		2	48	49.7	2.8	52.0	16.0	46.2	3.2	
	3	36	43.0	49.0	41.0	25.0	38.4	5.8		3	58	50.0	64.0	54.8	10.6	48.0	100.0	
	4	44	40.0	16.0	38.3	33.1	39.2	23.0		4	60	53.0	49.0	53.5	42.3	51.0	81.0	
2001	1	31	40.7	93.4	37.5	42.3	41.8	116.6	2007	1	48	55.3	53.8	55.0	49.0	53.6	31.4	
	2	39	37.0	4.0	38.0	1.0	39.0	0.0		2	54	55.3	1.8	55.3	1.6	53.4	0.4	
	3	38	38.0	0.0	40.8	7.6	38.4	0.2		3	59	54.0	25.0	53.3	33.1	53.6	29.2	
	4	55	36.0	361.0	43.5	132.3	37.6	302.8		4	52	53.7	2.8	53.0	1.0	55.8	14.4	
2002	1	42	44.0	4.0	45.0	9.0	41.4	0.4	2008	1	47	55.0	64.0	53.8	45.6	54.6	57.8	
	2	45	45.0	0.0	49.3	18.1	41.0	16.0		2	57	52.7	18.8	50.3	45.6	52.0	25.0	
	3	55	47.3	58.8	45.5	90.3	43.8	125.4		3	45	52.0	49.0	49.0	16.0	53.8	77.4	
	4	40	47.3	53.8	46.5	42.3	47.0	49.0		4	47	49.7	7.1	47.5	0.3	52.0	25.0	
2003	1	46	46.7	0.4	46.3	0.1	47.4	2.0	2009	1	41	49.7	75.1	41.3	0.1	49.6	74.0	
	2	44	47.0	9.0	44.5	0.3	45.6	2.6		2	32	44.3	152.1	34.8	7.6	47.4	237.2	
	3	48	43.3	21.8	42.5	30.3	46.0	4.0		3	19	40.0	441.0	32.5	182.3	44.4	645.2	
	4	32	46.0	196.0	40.3	68.1	46.6	213.2		4	38	30.7	53.8	29.7	69.4	36.8	1.4	
											MSE		130.6		44.4		120.4	

한 가스사고 건수 예측은 3기간, 4기간, 5기간을 나누어 실시하였다. 각각의 추이에 대한 그래프는 Fig. 1에 나타내었다.

Fig. 1은 기간별로 나누어 사고의 수를 예측한 것이다. 그 결과 3기간으로 나누어 예측한 결과 평균제곱오차(MSE)가 130.6으로 나타났다. 4기간으로 나누어 사고의 수를 예측한 결과 평균제곱오차(MSE)가 44.4로 나타났으며 5기간으로 나누어 사고의 수를 예측한 결과 평균제곱오차(MSE)가 120.4로 나타났다. 평균제곱오차가 가장 적은 4기간 이동평균에 의한 예측이 가장 정확도가 높은 것으로 나타났다. Table 1은 실제 사고건수와 예측치를 나타낸 것이다.

Table 2. Weights of each quarter for applying weighted moving average method

분기	4분기	3분기	2분기	1분기	계
가중치	0.1	0.2	0.3	0.4	1

3.2. 가중이동평균법에 의한 예측

가중이동평균법(weighted moving average method)은 단순이동평균법의 단점을 보완한 방법으로 일종의 이동평균이라는 점에서 단순이동평균법과 유사하지만 각 분기별 가중치가 다르다는 점이 특이하다.

즉, 예측시기와 가까운 쪽에 더 큰 가중치를 부여하고 먼 쪽에 더 작은 가중치를 부여하는 방식으로

Table 3. Prediction of the number of accidents using weighted moving average method

연도	분기	기간	사고건수	예측치	예측오차	예측오차자승	연도	분기	기간	사고건수	예측치	예측오차	예측오차자승	
1998	1	1	63	#N/A	#N/A	#N/A	2004	1	25	37	38.4	-1.4	2.0	
	2	2	59	#N/A	#N/A	#N/A		2	26	38	37.5	0.5	0.3	
	3	3	73	#N/A	#N/A	#N/A		3	27	63	47.2	15.8	249.6	
	4	4	84	73.6	10.4	108.2		4	28	50	50.2	-0.2	0.0	
1999	1	5	54	67.3	-13.3	176.9	2005	1	29	39	47.0	-8.0	64.0	
	2	6	32	53.1	-21.1	445.2		2	30	43	45.2	-2.2	4.8	
	3	7	31	41.2	-10.2	104.0		3	31	47	44.5	2.5	6.3	
	4	8	45	39.1	5.9	34.8		4	32	49	46.2	2.8	7.8	
2000	1	9	42	39.7	2.3	5.3	2006	1	33	53	49.6	3.4	11.6	
	2	10	42	41.5	0.5	0.3		2	34	48	49.6	-1.6	2.6	
	3	11	36	39.9	-3.9	15.2		3	35	58	53.1	4.9	24.0	
	4	12	44	41.0	3.0	9.0		4	36	60	56.3	3.7	13.7	
2001	1	13	31	37.0	-6.0	36.0	2007	1	37	48	53.6	-5.6	31.4	
	2	14	39	37.3	1.7	2.9		2	38	54	53.8	0.2	0.0	
	3	15	38	37.5	0.5	0.3		3	39	59	55.4	3.6	13.0	
	4	16	55	44.3	10.7	114.5		4	40	52	54.1	-2.1	4.4	
2002	1	17	42	44.8	-2.8	7.8	2008	1	41	47	51.6	-4.6	21.2	
	2	18	45	45.4	-0.4	0.2		2	42	57	53.2	3.8	14.4	
	3	19	55	49.4	5.6	31.4		3	43	45	49.7	-4.7	22.1	
	4	20	40	45.7	-5.7	32.5		4	44	47	48.4	-1.4	2.0	
2003	1	21	46	45.9	0.1	0.0	2009	1	45	41	45.2	-4.2	17.6	
	2	22	44	44.9	-0.9	0.8		2	46	32	39.0	-7.0	49.0	
	3	23	48	45.6	2.4	5.8		3	47	19	30.1	-11.1	123.2	
	4	24	32	40.6	-8.6	74.0		4	48	38	31.4	6.6	43.6	
												MSE		43.0

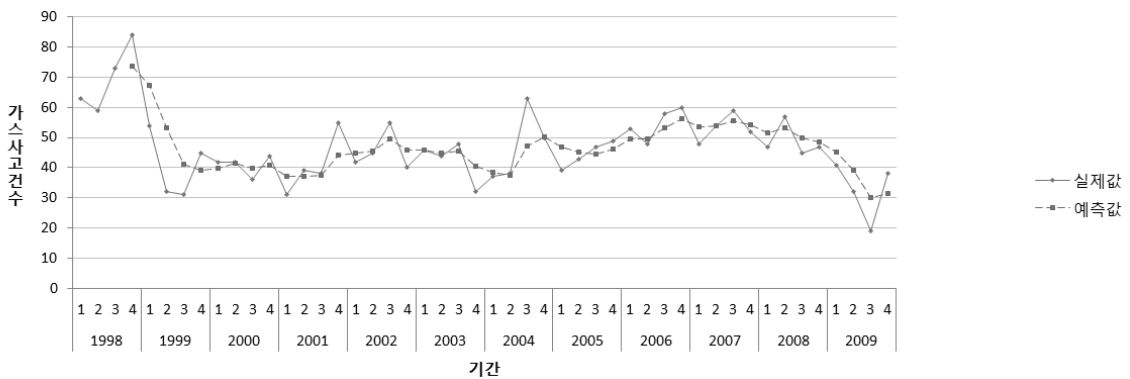


Fig. 2. Prediction of the number of accidents using weighted moving average method.

본 연구에서는 Table 2와 같이 n(기간 수)를 4가지로 분류하여 가중치를 주어 예측 값을 산정하였다.

그 결과, Table 3에서와 같이 평균제곱오차(MSE)가 43.0으로 이동평균법 중에서 가장 낮은 값을 나타내어 정확도가 가장 높은 것으로 나타났으며 그 추이 그래프는 Fig. 2와 같다.

3.3. 지수평활법에 의한 예측

지수평활법(exponential smoothing)의 적용을 위해서는 평활계수(smoothing constant) a의 값을 결정해야 하며, 일반적으로 시행오차(trial and error)를 통해 예측오차를 최소화 할 수 있도록 하는 방법이 이용되고 있다. 본 연구에서는 시행오차를 사용하여 예측오차를 가장 최소화 할 수 있는 a를 결정하였다. Table 4에서 평균제곱오차가 최소가 되는

평활계수 (a) 는 0.54이다.

지수평활법에 의한 가스사고 건수 예측결과는 Table 5와 같다. 분석결과를 요약하면 평균제곱오차는 109.43으로 4기간 이동평균이나 가중이동평균보다는 높은 것으로 나타났다. 추이 그래프는 Fig. 3와 같다.

Table 4. Smoothing constant in minimum mean square error

평활계수 (a)	-	0	0.2	0.4	0.54	0.6	0.8	1
평균제곱오차 (MSE)	0.00	403.8	126.4	111.2	109.4	109.7	113.2	120.3

Table 5. Prediction of the number of accidents using exponential smoothing method

연도	분기	기간	사고건수	예측치	예측오차	예측오차자승	연도	분기	기간	사고건수	예측치	예측오차	예측오차자승
1998	1	1	63	63.00	#N/A	#N/A	2004	1	25	37	38.65	-1.65	2.73
	2	2	59	63.00	-4.00	16.00		2	26	38	37.76	0.24	0.06
	3	3	73	60.84	12.16	147.87		3	27	63	37.89	25.11	630.55
	4	4	84	67.41	16.59	275.35		4	28	50	51.45	-1.45	2.10
1999	1	5	54	76.37	-22.37	500.28	2005	1	29	39	50.67	-11.67	136.11
	2	6	32	64.29	-32.29	1042.57		2	30	43	44.37	-1.37	1.87
	3	7	31	46.85	-15.85	251.31		3	31	47	43.63	3.37	11.37
	4	8	45	38.29	6.71	44.99		4	32	49	45.45	3.55	12.61
2000	1	9	42	41.91	0.09	0.01	2006	1	33	53	47.37	5.63	31.73
	2	10	42	41.96	0.04	0.00		2	34	48	50.41	-2.41	5.80
	3	11	36	41.98	-5.98	35.78		3	35	58	49.11	8.89	79.07
	4	12	44	38.75	5.25	27.54		4	36	60	53.91	6.09	37.09
2001	1	13	31	41.59	-10.59	112.06	2007	1	37	48	57.20	-9.20	84.61
	2	14	39	35.87	3.13	9.80		2	38	54	52.23	1.77	3.13
	3	15	38	37.56	0.44	0.19		3	39	59	53.19	5.81	33.80
	4	16	55	37.80	17.20	295.92		4	40	52	56.33	-4.33	18.71
2002	1	17	42	47.09	-5.09	25.88	2008	1	41	47	53.99	-6.99	48.86
	2	18	45	44.34	0.66	0.44		2	42	57	50.22	6.78	46.03
	3	19	55	44.70	10.30	106.16		3	43	45	53.88	-8.88	78.84
	4	20	40	50.26	-10.26	105.27		4	44	47	49.08	-2.08	4.34
2003	1	21	46	44.72	1.28	1.64	2009	1	45	41	47.96	-6.96	48.42
	2	22	44	45.41	-1.41	1.99		2	46	32	44.20	-12.20	148.87
	3	23	48	44.65	3.35	11.23		3	47	19	37.61	-18.61	346.42
	4	24	32	46.46	-14.46	209.05		4	48	38	27.56	10.44	108.96
MSE													109.43

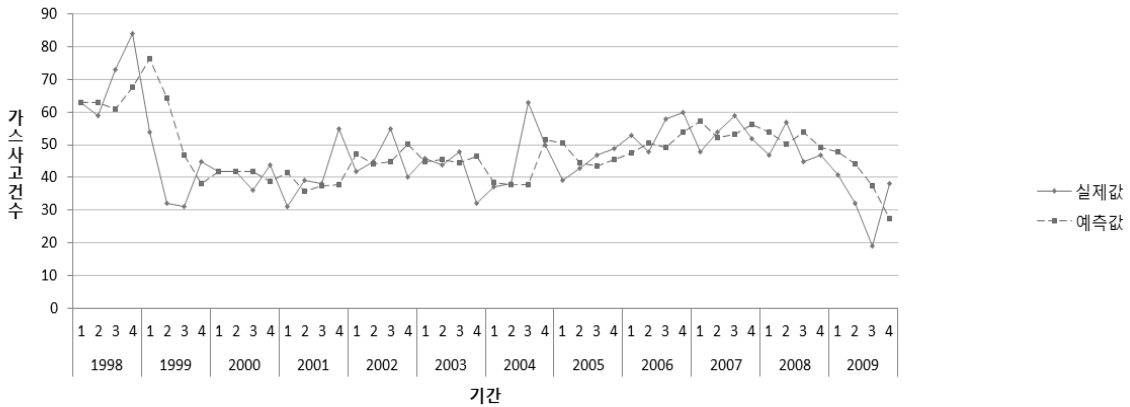


Fig. 3. Prediction of the number of accidents using exponential smoothing method.

Table 6. Mean square error by each technique

예측기법	이동평균 (3기간)	이동평균 (4기간)	이동평균 (5기간)	가중이동 평균	지수평활
평균오차 제곱합	130.6	44.4	120.4	43	109.43

Table 7. Seasonal index for each quarter

분기	1	2	3	4	
98년	-	-	1.064	1.31	
99년	0.973	0.705	0.795	1.161	
00년	1.034	1.021	0.909	1.162	
01년	0.821	0.99	0.902	1.243	
02년	0.891	0.95	1.196	0.863	
03년	1.014	1.011	1.16	0.81	
04년	0.911	0.849	1.333	1.039	
05년	0.834	0.964	1.016	1.008	
06년	1.047	0.899	1.072	1.106	
07년	0.871	0.995	1.111	0.974	
08년	0.904	1.149	0.933	1.059	
09년	1.079	0.952	-	-	
계절지수	평균	0.943	0.953	1.043	1.042
계절지수	수정	0.948	0.958	1.047	1.047
계	평균	3,982			
	수정	4			

3.4. 예측방법의 선택

예측 결과 Table 6에서와 같이 4기간 이동평균법과 가중이동평균법이 가장 예측력이 높은 것으로 나타났다. 이 두 가지 기법을 이용하여 가스사고 건수를 예측하는 것이 가장 좋을 것으로 판단된다.

또한, 가스사고의 경우 계절적인 요인도 있을 것으로 생각되므로 승법모형(multiplicative model)에 의한 계절지수(seasonal index)를 구하여 이를 활용하면 예측의 정확도를 높일 수 있다. 이동평균중심치란 이동평균을 구하는 과정에서 예측대상기간의 값을 중심으로 과거와 미래의 데이터를 활용하여 계산된 값이다. 승법모형의 계절지수는 추세변동의 백분율로 표시되며 각분기별 하나의 계절지수가 존재하게 된다. 따라서 시계열 자료가 분기별로 주어진 경우에는 4개의 분기별 계절지수가 있게 된다. 산출된 계절지수는 Table 7과 같이 1분기는 0.948, 2분기는 0.958, 3분기 및 4분기는 1.047로 나타났다. Table 8은 분기별 시계열자료를 사용하여 계절지수를 산출한 표이다. 이 계절지수를 각 기법에서 도출된 예측 값에 곱하여 보정하면 된다. 따라서 향후 가스사고의 데이터가 축적되면 매년 기법 적용에 따른 계절지수의 보정이 필요하다.

Table 8. Calculation of seasonal index using multiplicative model

연도	분기	기간	사고건수	이동평균 중심치	특정계절지수
1998	1	1	63		
	2	2	59		
	3	3	73	68.625	1.064
	4	4	84	64.125	1.310
1999	1	5	54	55.500	0.973
	2	6	32	45.375	0.705
	3	7	31	39.000	0.795
	4	8	45	38.750	1.161
2000	1	9	42	40.625	1.034
	2	10	42	41.125	1.021
	3	11	36	39.625	0.909
	4	12	44	37.875	1.162
2001	1	13	31	37.750	0.821
	2	14	39	39.375	0.990
	3	15	38	42.125	0.902
	4	16	55	44.250	1.243
2002	1	17	42	47.125	0.891
	2	18	45	47.375	0.950
	3	19	55	46.000	1.196
	4	20	40	46.375	0.863
2003	1	21	46	45.375	1.014
	2	22	44	43.500	1.011
	3	23	48	41.375	1.160
	4	24	32	39.500	0.810
2004	1	25	37	40.625	0.911
	2	26	38	44.750	0.849
	3	27	63	47.250	1.333
	4	28	50	48.125	1.039
2005	1	29	39	46.750	0.834
	2	30	43	44.625	0.964
	3	31	47	46.250	1.016
	4	32	49	48.625	1.008
2006	1	33	53	50.625	1.047
	2	34	48	53.375	0.899
	3	35	58	54.125	1.072
	4	36	60	54.250	1.106
2007	1	37	48	55.125	0.871
	2	38	54	54.250	0.995
	3	39	59	53.125	1.111
	4	40	52	53.375	0.974
2008	1	41	47	52.000	0.904
	2	42	57	49.625	1.149
	3	43	45	48.250	0.933
	4	44	47	44.375	1.059
2009	1	45	41	38.000	1.079
	2	46	32	33.625	0.952
	3	47	19		
	4	48	38		

IV. 결론

본 연구에서는 국내 가스 사고의 발생 건수를 분석하기 위하여 1998년부터 2009년까지 12년간 2,244건의 가스 사고를 이동평균법, 가중이동평균법, 지수평활법을 활용하여 분석 결과에 대해 모델의 활용 가능성 여부를 제시하였다.

가스사고 발생 건수를 예측하기 위하여 단순이동평균법(3기간, 4기간, 5기간), 가중이동평균법 및 지수평활법을 적용해 본 결과, 각각의 평균오차제곱이 3기간 이동평균법의 경우 130.6, 4기간 이동평균법은 44.4, 5기간 이동평균법은 120.4로 나타났으며, 가중이동평균법은 43.0으로, 지수평활법의 경우는 109.43으로 나타나, 4기간 이동평균법과 가중이동평균법에 의한 모델이 가장 정확도가 높은 것으로 나타났다.

가스사고 예측을 통해 가스사고 예방활동에 적극 활용할 수 있을 것으로 사료되고, 제시된 모델식을 활용하여 가스사고 발생건수를 시각적으로 지도상에 지역별로 구현되도록 함으로써 지역별 특성에 맞는 사고예방활동이 가능하도록 지원이 가능할 것으로 보인다. 또한 일반 국민들에게 예측값을 제공하여 가스사고 발생을 감소시키는데 기여할 수 있을 것이다.

감사의 글

본 연구는 소방방재청 ‘차세대 핵심소방안전기술 개발지원 사업(과제번호2012-NEMA14-009-01020000-2013)’의 지원에 의하여 수행된 것입니다.

참고문헌

- [1] Heo, In-Hye and Lee, Seung-Ho., *The Impact of Climate Change on Ski Industries in South Korea*, Journal of the Korean Geographical Society, Vol. 43(5), 715-727, (2008)
- [2] Wu Wen-feng, Wu Wan-qing, Feng Xing, Wang Zhen-mao, "Study on the Frequency Prediction of Marine Accidents for Ship on the Sea", Bioinformatics and Biomedical Engineering, (iCBBE) 2011 5th International Conference on, 1-4, (2011)
- [3] Lee, Dong-Min, Kim, Do-Hoon, Sung, Nakmoon, "Development of a Accident Frequency Prediction Model at Rural Multi-Lane Highways", *KOR-KST*, 27(4), 207-215, (2009)
- [4] Heo, Young-Taeg, Shin, Dong-Il, Lee, Su-Kyung, "Data Mining of Gas Accident and

- Meteorological Data in Korea for a Prediction Model of Gas Accidents", *KIGAS*, 16(1), 33-38, (2012)
- [5] Harvey, Andrew C. "Time Series Models", Cambridge, MA, MIT Press, (1993)
- [6] Zou, H., Yang, Y. H., "Combining Time Series Model for Forecasting", *International Journal of Forecasting*, 20(1), 69-84, (2004)
- [7] Godfrey, L. "Testing the Adequacy of a Time Series Model", *Biometrika*, 66(0), 67-72, (1979)
- [8] Lim, C., McAleer, Michael, "Time Series Forecasts of International Travel Demand for Australia", *Tourism Management*, 23(4), 389-396, (2002)
- [9] Box, George E. P., Gwilym M. J., "Time Series Analysis : Forecasting and Control", Revised Edition, Oakland, CA., Holden-Day, (1976)