Journal of the Korea Academia-Industrial cooperation Society
Vol. 17, No. 2 pp. 87-95, 2016

http://dx.doi.org/10.5762/KAIS.2016.17.2.87
ISSN 1975-4701 / eISSN 2288-4688

국내 도시가스의 시간대별 수요 예측

한정희[1], 이근철[2*]
[1]강원대학교 경영학과, [2]건국대학교 경영대학

Forecasting Hourly Demand of City Gas in Korea

Jung-Hee Han[1], Geun-Cheol Lee[2*]
[1]Department of Business Administration, Kangwon National University
[2]College of Business Administration, Konkuk University

요 약 본 연구에서는 국내 도시가스 수요 데이터를 분석하여 시간대별 도시가스 수요의 특성을 파악하고 정확한 시간대별 도시가스 수요 예측을 위해 다중회귀모형(multiple regression model)을 개발하였다. 시간대별 도시가스 수요를 정확하게 예측하는 것은 공급자의 비용 절감뿐만 아니라 안정적인 배관망 관리 측면에서도 매우 중요하다. 수요 예측 오류로 인해 가스 공급이 부족한 상황이 발생하면 부족한 공급량을 빠른 시간내에 보충하기 위해 가스 배관망의 압력을 급격히 증가시켜야 하는 응급 상황이 전개될 수 있다. 반면, 시간대별 가스 생산량이 실제 수요보다 많은 경우에는 과다한 저장 시설 운용 및 불필요한 생산 비용이 발생하는 문제가 있다. 과거 시간대별 도시가스 수요 데이터를 분석한 결과 시간대별 도시 가스 수요는 직전 시간대(즉, 24시간 전) 수요와 매우 높은 상관관계를 보이며 24시간 수요 패턴은 1주일전 동일 요일(즉, 168시간 전)의 24시간 수요 패턴과 매우 높은 상관관계가 있음을 확인하였다. 또한, 외기 온도가 도시가스 수요에 영향을 주는 특수한 조건을 파악하였다. 즉, 시간대별 도시가스 수요와 시간대별 외기 온도는 평균적으로 0.853의 높은 상관계수 절대값을 보여주며, 상관관계 분석시 같은 요일에 속한 데이터만 분석하면 상관계수의 절대값은 최저 0.861 및 최고 0.965까지 증가한다. 이상의 분석 결과를 바탕으로 본 연구에서는 24시간 전 수요와 168시간 전 수요를 독립변수로 고려한 다중회귀모형 및 외기 온도를 추가한 두 번째 다중회귀모형을 제안하며, 제안한 예측모형의 성능을 확인하기 위해 2009년부터 2013년까지 5년간의 시간대별 수요 예측 결과를 평가하였다. 본 연구에서 제안한 24시간 전 수요와 168시간 전 수요를 독립변수로 고려한 다중회귀모형의 경우 과거 5년간의 수요 예측 오차율의 절대값 평균(mean absolute percentage error)은 4.5% 수준이며, 외기 온도를 추가한 모형의 경우 오차율의 절대값 평균은 5.13%임을 확인하였다.

Abstract This study examined the characteristics of the hourly demand of city gas in Korea and proposed multiple regression models to obtain precise estimates of the hourly demand of city gas. Forecasting the hourly demand of city gas with accuracy is essential in terms of safety and cost. If underestimated, the pipeline pressure needs to be increased sharply to meet the demand, when safety matters. In the opposite case, unnecessary inventory and operation costs are incurred. Data analysis showed that the hourly demand of city gas has a very high autocorrelation and that the 24-hour demand pattern of a day follows the previous 24-hour demand pattern of the same day. That is, there is a weekly cycle pattern. In addition, some conditions that temperature affects the hourly demand level were found. That is, the absolute value of the correlation coefficient between the hourly demand and temperature is about 0.853 on average, while the absolute value of the correlation coefficient on a specific day improves to 0.861 at worst and 0.965 at best. Based on this analysis, this paper proposes a multiple regression model incorporating the hourly demand ahead of 24 hours and the hourly demand ahead of 168 hours, and another multiple regression model with temperature as an additional independent variable. To show the performance of the proposed models, computational experiments were carried out using real data of the domestic city gas demand from 2009 to 2013. The test results showed that the first regression model exhibits a forecasting accuracy of MAPE (Mean Absolute Percentage Error) around 4.5% over the past five years from 2009 to 2013, while the second regression model exhibits 5.13% of MAPE for the same period.

Keywords : City gas, Forecasting, Hourly demand, Regression, Time series

*Corresponding Author : Geun-Cheol Lee (Konkuk University)
Tel: +82-2-450-4100 email: gclee@konkuk.ac.kr
Received November 23, 2015 Revised (1st January 6, 2016, 2nd January 11, 2016)
Accepted February 4, 2016 Published February 29, 2016

1. 서론

2013년 기준 국내 도시가스 시장규모는 약 22조원, 가정용 도시가스 보급률은 약 75%이며 도시가스 소비는 매년 약 4%의 성장을 지속해왔다[1]. 가정용 도시가스의 경우 대부분 난방과 취사에 사용되는 점을 고려하면 도시가스는 공공재임과 동시에 필수재 성격을 갖는다. 따라서, 국민의 편의와 공공의 이익 증진을 위해서는 도시가스의 안정적 공급과 정확한 수요 예측이 매우 중요하다. 연간 수요를 정확히 예측하는 것도 중요하지만, 하루 또는 시간 단위 수요 예측 또한 공공의 편의성 및 공급자의 비용 절감 측면에서 매우 중요하다. 잘못된 수요 예측으로 단시간 내에 도시가스 공급을 증가시키거나, 예측한 만큼 도시가스 소비가 이루어지지 않으면 배관망내 압력이 높게 유지되므로 사고의 위험이 증가한다[2]. 도시가스 공급소가 충분한 재고를 확보하고 있더라도 전력 공급과는 달리 가스 공급소의 공급량 조절 효과가 최종 소비자 배관망에 나타나기까지는 배관망의 길이 및 압력에 따라 최소 몇 시간이 소요된다. 따라서 시간대별 도시가스 수요의 정확한 예측을 전제로 경제적이고 안정적인 도시가스 공급 및 조달 계획 수립이 가능하다.

본 연구에서는 시간대별 도시가스 수요 예측을 위해 선형회귀모형을 개발하며, 도시가스 수요 예측에 관한 기존 연구를 소개한다. 거시경제 변수를 참고하여 중장기 도시가스 수요를 예측한 연구는 몇 건이 있으나[3, 4, 5], 하루 단위의 국내 도시가스 수요 총량을 예측한 연구는 1편[6] 외에는 찾을 수 없다. 특히, 국내 도시가스 공급 데이터를 분석하여 시간대별 도시가스 수요를 예측한 연구는 아직 발표된 바 없는 것으로 파악된다. 다만, 전력수요 예측 분야에서 초단기 수요예측 연구가 발표되고 있으나, 대부분 하루 중 최대 전력수요를 예측하는 내용이며[7 - 10], 전력수요 예측 분야에서도 시간대별 수요를 예측한 연구는 많지 않다[11, 12]. 전력수요 예측 분야에서 조차 시간대별 수요예측에 관한 연구가 드문 이유는 공급과 소비가 리드타임(lead time) 없이 거의 동시에 이루어지는 전력의 특성상 시간대별 수요예측보다는 블랙아웃(black- out)과 같은 사고를 막기 위해서는 1일 최대 전력수요 예측이 보다 중요하기 때문이다. 반면, 앞서 언급한 바와 같이 도시가스의 경우 공급소에서 조절한 도시가스 공급량이 최종 소비자 배관망에 전달되기까지는 배관망의 길이와 배관망 내 압력에 따라 최소 몇 시간의 리드타임이 존재한다. 도시가스의 공급과 수요가 전력과는 다른 방식으로 이루어지긴 하지만, 동일한 초단기 에너지 수요예측이라는 측면에서 기존의 1일 최대 전력수요 예측 연구를 참고하는 것은 도움이 될 수 있다. 평균 기온뿐만 아니라 15시 상대 습도를 고려한 회귀모형 개발[13], 예측일 오전 8시의 전력사용량을 당일 최대전력수요 예측에 사용하는 신경망(neural network) 알고리즘 개발[7], 예측일의 최고 기온과 최저 기온을 모두 참고하여 예측일의 최대 전력수요를 예측한 연구[10] 등이 1일 최대 전력수요 예측에 관한 몇 편의 기존 연구이다. 이 밖에도 일일 최대 전력수요를 예측한 대부분의 연구는 예측일 당일의 온도 정보를 참고하는 특징이 있다[8, 14, 15]. 한편, 시간대별 전력수요 예측에 관해서는 시간대별 과거 수요실적을 기준수요로 나눈 상대계수를 이용한 연구[11], 하루를 3개의 시간대로 나누어 구간별로 최고기온, 평균기온, 최고기온의 3일간 평균값을 독립변수로 고려한 회귀모형을 적용한 연구[12]가 있다.

본 연구에서는 향후 24시간 동안 발생하는 시간대별 도시가스 수요를 예측하기 위해 국내 도시가스 수요에 관한 수년간의 과거 데이터를 분석하며, 데이터 분석 결과를 바탕으로 직전일 동일 시간대 수요, 7일전 동일 시간대 수요 및 시간대별 외기온도를 고려한 선형회귀모형을 개발한다.

2. 도시가스 데이터 분석

본 연구에서는 도시가스 수요의 특성을 분석하기 위해 2009년 1월 1일부터 2013년 11월 30일까지 시간대별 전국적인 수요(공급량) 데이터를 사용한다. 먼저, 개괄적인 그래프를 이용하여 도시가스 수요 패턴에 대한 직관적인 이해를 돕는다. Fig. 1에는 시간대별 도시가스 수요의 연중 특성을 살펴보기 위해 2012년 12월 1일 ~ 2013년 11월 31일 사이의 시간대별 수요를 그래프로 나타낸다.

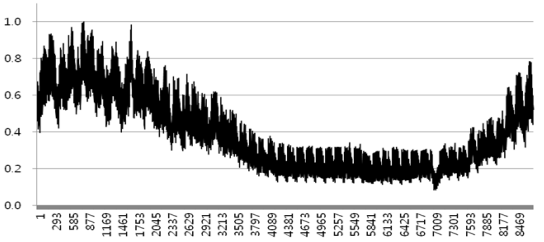


Fig. 1. Hourly Demand Pattern of City Gas Over One Year: 2012. 12. 1. ~ 2013. 11. 30

Fig. 1에서 세로축(Y-axis)은 해당 기간내 시간대별 수요의 최대값에 대한 비율을 나타낸다. 도시가스 수요는 겨울철에 가장 많고 여름철에 적으며 계절에 따라 도시가스 수요의 수준은 다르지만 주기적으로 유사한 패턴이 반복됨을 확인할 수 있다. 시간대별 도시가스 수요의 주기적인 패턴을 자세히 살펴보기 위해 Fig. 2에 겨울철과 여름철 1주일간의 시간대별 수요를 나타낸다. Fig. 2의 세로축 역시 해당 기간내 시간대별 수요의 최대값에 대한 비율을 나타낸다. Fig. 2를 살펴보면 겨울철에는 요일에 관계없이 아침과 저녁 하루에 두 번 식사 시간 즈음에만 수요가 급증함을 알 수 있다. 반면, 여름철에는 주로 아침에 극대점(peak)이 형성되며 하루 중 시간대별 수요 패턴이 전반적으로 완만하다. 동일한 계절만 살펴보면 시간대별 도시가스 수요는 일별로 조금씩 차이는 있지만 매우 유사한 패턴을 보인다.

식사 준비와 같은 생활 패턴 외에 시간대별 도시가스 수요에 영향을 주는 요인으로는 외기온도(℃)를 고려해 볼 수 있다. Fig. 3에는 Fig. 2와 동일한 기간 동안 관측된 외기온도를 나타낸다. Fig. 2와 Fig. 3의 패턴을 비교해보면 시간대별 도시가스 수요와 시간대별 외기온도의 상관성은 크지 않아 보인다. 시간대별 도시가스 수요와 외기온도의 추가적인 상관성 분석 결과는 뒤에서 다시 제시한다.

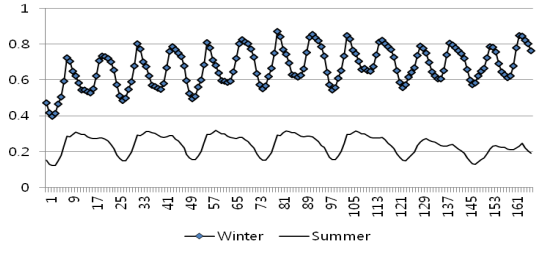


Fig. 2. Hourly Demand Pattern of City Gas: Winter, Summer (Winter: 2012. 12. 3(Mon) ~ 12. 9, Summer: 2013. 7. 1(Mon) ~ 7.7)

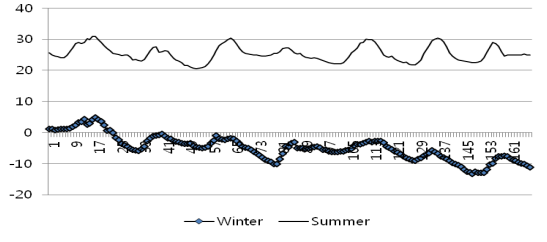


Fig. 3. Hourly Temperature: Winter, Summer (Winter: 2012. 12. 3(Mon) ~ 12. 9, Summer: 2013. 7. 1(Mon) ~ 7.7)

봄철과 가을철에도 시간대별 도시가스 수요와 시간대별 외기온도의 뚜렷한 상관성은 발견할 수 없으며, 봄철의 시간대별 수요 패턴은 겨울철 패턴과 유사하며 가을철 시간대별 수요 패턴은 여름철 패턴과 유사하다. Fig. 2, 3의 여름철 그래프에서 시간대별 도시가스 수요와 외기온도의 뚜렷한 상관성을 확인하지는 못하였다. 반면, 겨울철에는 기온이 완만히 내려감에 따라, 도시가스 수요가 전체적으로 증가하는 경향을 확인할 수 있다. 기존 연구에서도 1일 도시가스 수요 총량이 평균기온과 매우 높은 상관성을 가지며[9], 1일 최대 전력수요 예측시에도 외기온도가 중요한 요소임을 고려하면[1, 5, 10], 도시가스의 시간대별 수요와 시간대별 온도의 상관성 존재 여부를 보다 자세히 분석할 필요가 있다. 이후 데이터 분석에서는 시간대별 도시가스 수요의 자기상관성(auto-correlation)을 살펴본 후에 요일을 구분하여 시간대별 도시가스 수요와 외기온도의 상관성을 자세히 살펴본다.

2.1 자기상관성 분석

Table 1에는 임의의 t시간대 수요와 직전 1 ~ 218시간대 (t-1 ~ t-218) 수요의 상관계수(correlation coefficient)를, Fig. 4에는 Table 1의 상관계수 값을 그래프로 나타낸다. Table 1과 Fig. 4에서 파악할 수 있는 시간대별 도시가스 수요의 몇 가지 특징은 다음과 같다.

(1) t시간대의 수요는 12시간 단위로 과거 시간대의 수요와 높은 양의 상관관계를 보이며, 특히 24시간 단위로 매우 높은 상관계수 값을 갖는다.

(2) 상관계수 값은 과거로 거슬러 올라갈수록 감소하는 경향을 보이지만, 168시간 전 (즉, 7일전 동일 시간대) 수요와는 다시 증가된 상관계수 값을 갖는다. 시간대별 도시가스 수요가 7일전 동일 시간대 수요와 매우 높은 상관계수 값을 갖는다는 점은 시간대별 수요가 요일 패턴을 가짐을 의미한다.

(3) 시간대별 수요와 높은 상관계수를 갖는 주요한 직전 시간대를 순서대로 나열하면 다음과 같으며 괄호 안에는 상관계수 값을 나타낸다: 1시간 전 (0.982) → 24시간 전 (0.975) → 48시간 전 (0.95) → 168시간(1주일) 전 (0.942) → 12시간 전 (0.939) → 72시간 전 (0.937).

Table 1. Correlation between Hourly Demand of City Gas and Demand 1 ~ 218 Hours Ahead

Hourly Time Slot	Correlation	Hourly Time Slot	Correlation	Hourly Time Slot	Correlation	Hourly Time Slot	Correlation
t-1	0.982	t-24	0.975	t-47	0.935	t-142	0.878
t-2	0.943	t-25	0.958	t-48	0.950	t-143	0.915
t-3	0.895	t-26	0.920	t-49	0.934	t-144	0.932
t-4	0.848	t-27	0.873	t-50	0.897	t-145	0.916
t-5	0.812	t-28	0.826	…		t-146	0.879
t-6	0.795	t-29	0.790	t-70	0.886	…	
t-7	0.800	t-30	0.773	t-71	0.922	t-166	0.886
t-8	0.827	t-31	0.778	t-72	0.937	t-167	0.925
t-9	0.865	t-32	0.804	t-73	0.921	t-168	0.942
t-10	0.903	t-33	0.841	t-74	0.884	t-169	0.925
t-11	0.930	t-34	0.879	…		t-170	0.887
t-12	0.939	t-35	0.905	t-94	0.877	…	
t-13	0.929	t-36	0.914	t-95	0.913	t-190	0.875
t-14	0.900	t-37	0.904	t-96	0.928	t-191	0.912
t-15	0.860	t-38	0.875	t-97	0.913	t-192	0.927
t-16	0.820	t-39	0.835	t-98	0.876	t-193	0.911
t-17	0.792	t-40	0.796	…		t-194	0.874
t-18	0.785	t-41	0.769	t-118	0.872	…	
t-19	0.800	t-42	0.762	t-119	0.909	t-214	0.862
t-20	0.834	t-43	0.777	t-120	0.924	t-215	0.897
t-21	0.878	t-44	0.811	t-121	0.909	t-216	0.912
t-22	0.923	t-45	0.855	t-122	0.873	t-217	0.897
t-23	0.960	t-46	0.899	…		t-218	0.861

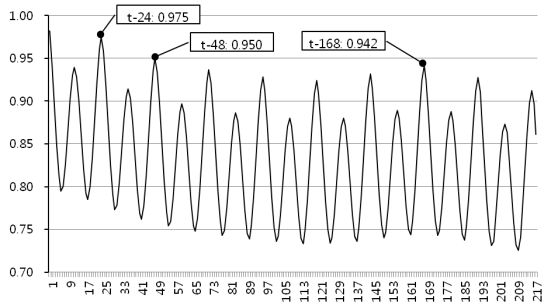


Fig. 4. Correlation between Hourly Demand of City Gas and Demand 1 ~ 218 Hours Ahead

2.2 외기온도와의 상관성 분석

Table 2에는 시간대별 도시가스 수요와 시간대별 외기온도와의 상관계수를 연도별로 정리하여 나타낸다.

Table 2. Correlation between Hourly Demand of City Gas and Temperature

Year	2009	2010	2011	2012	2013	전체
Correlation	-0.849	-0.874	-0.875	-0.869	-0.855	-0.853

시간대별 도시가스 수요가 24, 48 및 168시간 전 수요와 0.942 이상의 높은 상관계수 값을 가지는 점을 고려하면(Table 1 참조), Table 2에 나타낸 시간대별 외기온도와의 상관계수 절대값은 상대적으로 크지 않다. Table 2의 결과만으로 시간대별 도시가스 수요와 시간대별 외기온도의 상관성을 판단하는 것은 무리가 있어 보인다. 왜냐하면, 외기온도뿐만 아니라 외기온도와 무관한 도시가스 사용자의 생활 패턴이 동시에 시간대별 도시가스 수요에 영향을 미치는 것으로 파악되기 때문이다 (Fig. 2, 3 참조). 예를 들면, 가정에서 식사 준비를 위해 사용하는 도시가스 수요는 외기온도와는 무관하게 동일한 시간대에 증가한다. 생활 패턴에 따른 도시가스 수요 증감 요소를 배제하고 외기온도에 의한 도시가스 수요 변화만을 살펴보기 위해서는 Table 3에 나타낸 바와 같이 시간대와 요일을 구분하여 시간대별 도시가스 수요와 외기온도와의 상관계수를 살펴볼 필요가 있다. 예를 들면, 월요일 오전 7 ~ 8시 수요와 해당 시간대 외기온도의 관계를 살펴보기 위해서는 월요일 오전 7 ~ 8시 데이터만을 따로 모아서 해당 시간대 도시가스 수요와 외기온도의 상관성을 분석한다.

시간대와 요일을 모두 구분하면 Table 3에서 나타낸 바와 같이 시간대별 수요와 외기온도의 상관계수 값이 크게 높아진다. Table 3에서 상관계수 절대값은 최대 0.965(수요일 23시)에 이르며, 최저값은 0.861(월요일 15시)로서 시간대와 요일을 구분하지 않고 계산한 상관계수 절대값 0.853(Table 2의 7열, 또는 Table 3의 마지막 행, 마지막 열)보다 높게 나타난다. 요일만 구분할 경우, 즉 시간대는 구분하지 않고 특정 요일의 시간대별 도시가스 수요와 외기온도의 상관계수(마지막 행, 2 ~ 8열)를 구하면 절대값 평균이 0.859 (= {0.817 + 0.844 + … + 0.885 + 0.880} / 7)로서 시간대와 요일을 구분하지 않고 계산한 상관계수 절대값 0.853와 크게 다르지 않다. 반면, 시간대만 구분할 경우, 즉 요일을 구분하지 않고 특정 시간대의 도시가스 수요와 해당 시간대 외기온도의 상관계수(2 ~ 25행, 마지막 열)를 구하면 절대값 평균이 0.927 (= {0.948 + 0.941 + … + 0.950 + 0.950} / 24)로서 시간대와 요일을 모두 구분할 때 얻은 최대값 0.965에는 미치지 못하지만 모든 시간대에 걸쳐서 전반적으로 높은 상관계수를 보인다. 이상 분석한 시간대별 도시가스 수요와 높은 상관성을 보이는 요인을 순서대로 정리하여 Table 4에 나타낸다.

Table 3. Correlation between Hourly Demand of City Gas and Temperature in Respect to Day and Hourly Time Slot

Time Slot	Mon	Tue	Wed	Thu	Fri	Sat	Sun	Total
1	-0.947	-0.947	-0.955	-0.958	-0.955	-0.954	-0.951	-0.948
2	-0.940	-0.940	-0.944	-0.953	-0.947	-0.949	-0.946	-0.941
3	-0.939	-0.938	-0.949	-0.952	-0.945	-0.947	-0.944	-0.942
4	-0.941	-0.943	-0.955	-0.955	-0.949	-0.952	-0.948	-0.946
5	-0.944	-0.947	-0.957	-0.957	-0.952	-0.954	-0.951	-0.949
6	-0.940	-0.944	-0.956	-0.956	-0.950	-0.951	-0.948	-0.945
7	-0.938	-0.943	-0.957	-0.959	-0.950	-0.953	-0.949	-0.931
8	-0.936	-0.942	-0.958	-0.960	-0.949	-0.957	-0.954	-0.920
9	-0.929	-0.932	-0.949	-0.953	-0.944	-0.951	-0.953	-0.932
10	-0.914	-0.917	-0.936	-0.938	-0.932	-0.937	-0.947	-0.927
11	-0.904	-0.907	-0.927	-0.931	-0.925	-0.933	-0.940	-0.919
12	-0.890	-0.897	-0.916	-0.919	-0.915	-0.923	-0.932	-0.906
13	-0.875	-0.887	-0.905	-0.910	-0.903	-0.913	-0.927	-0.895
14	-0.865	-0.880	-0.900	-0.904	-0.898	-0.906	-0.925	-0.887
15	-0.861	-0.876	-0.896	-0.896	-0.893	-0.903	-0.923	-0.883
16	-0.865	-0.879	-0.896	-0.897	-0.890	-0.904	-0.924	-0.885
17	-0.879	-0.898	-0.908	-0.906	-0.899	-0.915	-0.930	-0.898
18	-0.908	-0.923	-0.923	-0.929	-0.920	-0.928	-0.939	-0.920
19	-0.935	-0.948	-0.951	-0.948	-0.943	-0.944	-0.948	-0.941
20	-0.941	-0.955	-0.958	-0.952	-0.946	-0.947	-0.950	-0.944
21	-0.944	-0.958	-0.960	-0.955	-0.950	-0.950	-0.951	-0.946
22	-0.949	-0.960	-0.964	-0.959	-0.955	-0.953	-0.953	-0.948
23	-0.951	-0.961	-0.965	-0.961	-0.957	-0.954	-0.955	-0.950
24	-0.951	-0.959	-0.962	-0.960	-0.957	-0.953	-0.950	-0.950
Total	-0.817	-0.844	-0.860	-0.862	-0.862	-0.885	-0.880	-0.853

Table 4. Key Factors Affecting Hourly Demand of City Gas

Order	Factor	Absolute Value of Correlation Coefficient
1	Demand of 1 hour ahead	0.982
2	Demand of 24 hours ahead	0.975
3	Demand of 48 hours ahead	0.950
4	Demand of 168 hours ahead	0.942
5	Demand of 12 hours ahead	0.939
6	Demand of 72 hours ahead	0.937
7	Temperature measured by hour (regardless of day and hourly time slot)	0.927
8	Temperature measured by hour (in respect of day and hourly time slot)	0.861 ~ 0.965

3. 회귀모형 분석

본 장에서는 도시가스 시간대별 수요 예측을 위한 모형을 제시하고, 계산 실험을 통해 제안 모형의 성능을 분석한다.

본 연구에서는 시간대별 도시가스 수요 예측시 다음과 같은 상황을 가정한다.

- 임의의 시점 d일(오늘), t = 24 시점(자정)에 d+1일(내일) 모든 시간대(t = 1, 2, …, 24)의 도시가스 수요를 예측한다.
- d+1일의 모든 시간대에 대해서 24시간 전 시간대별 도시가스 수요(실제 공급량) 데이터가 존재한다.

3.1 예측 모형

본 연구에서 제시하는 회귀모형(regression model)을 자세히 설명하기 위해 기호 및 변수를 정의한다.

기호 및 변수 정의

- Y_t: d+1일 t시간대의 예측 수요 (종속변수)
- X_{t-24}: d+1일 t시간대의 24시간 전 수요
- X_{t-168}: d+1일 t시간대의 168시간 전(1주일 전 동일 시간대) 수요
- X_{t-336}: d+1일 t시간대의 336시간 전(2주일 전 동일 시간대) 수요
- Temp: d+1일 t시간대의 외기온도
- β_{24}, β_{168}: 독립변수 X_{24}, X_{168}의 회귀계수
- β_{Temp}: 독립변수 Temp의 회귀계수
- β_0: 회귀 상수

본 연구에서 제시하는 회귀모형은 자기상관성만을 고려한 T24_168 모형과 자기상관성과 외기온도 요인을 동시에 고려한 T24_168+Temp 모형으로 다음과 같이 표현할 수 있다.

회귀모형

T24_168: $$Y_t = \beta_{24} X_{t-24} + \beta_{168} X_{t-168} + \beta_0$$

T24_168+Temp: $$Y_t = \beta_{24} X_{t-24} + \beta_{168} X_{t-168} + \beta_{Temp} Temp + \beta_0$$

회귀모형 T24_168 및 T24_168+Temp의 적합도 검증을 위해 분산분석(ANOVA)을 실시한 결과는 Table 5와 같다. 두 종류의 회귀모형 모두 회귀모형의 적합도 검정에서 유의한 F-value가 0이 나왔으므로 유의수준 0.01% 수준에서도 두 회귀모형은 적합하다고 판단할 수 있다. 또한, 회귀계수 적합도 검정에서도 가장 높은 p-value가 1.23E-12이므로 두 종류의 회귀모형에 표현된 모든 독립변수 및 회귀상수 β_0를 포함한 회귀모형 구

성은 유의수준 0.01% 수준에서도 올바른 선택이었음을 확인할 수 있다.

Table 5. Analysis of Variance (ANOVA) of Regression Model T24_168(+Temp)

Regression Model	Test of Regression Model (Significant F-value)	Test of Regression Coefficient (p-value)			
		β_{24}	β_{168}	β_{Temp}	β_0
T24_168	0	0	0	-	1.23E-12
T24_168 +Temp	0	0	0	0	0

회귀모형 T24_168 및 T24_168+Temp의 수행도를 평가하기 위해 본 연구에서는 이동평균법(moving average)과 지수평활법(exponential smoothing)에 의한 수행도를 함께 제시한다. 시간대별 수요가 24시간 단위의 이전 시간대 수요와 매우 높은 자기상관성을 가짐을 고려하여 이동평균법은 다음과 같이 2가지 방법을 평가한다. 이동평균법 적용시 α는 0과 1사이의 사전에 결정된 값으로 두 관측치에 대한 가중치를 정의한다.

이동평균법

MA24_168: $Y_t = \alpha\ X_{t-24} + (1-\alpha)\ X_{t-168}$
MA168_336: $Y_t = \alpha\ X_{t-168} + (1-\alpha)\ X_{t-336}$

본 연구에서 수행도 비교를 위해 테스트하는 지수평활법 모형은 아래에 제시하며, τ는 0과 1사이의 사전에 결정된 값으로 직전 기간의 관측값에 대한 가중치를 정의한다. 또한, Y_{t-24} 및 Y_{t-168}는 수요를 예측하고자 하는 d+1일 t시간대를 기준으로 24시간 및 168시간 전 시간대의 수요를 예측한 과거 데이터를 의미한다.

지수평활법

ES24: $Y_t = \tau\ X_{t-24} + (1-\tau)\ Y_{t-24}$
ES168: $Y_t = \tau\ X_{t-168} + (1-\tau)\ Y_{t-168}$

3.2 샘플링 및 오차율 측정 방법

d+1일 t시간대의 도시가스 수요를 예측하는 경우, d+1일의 요일과 시간대(t)가 모두 일치하는 가장 최근의 데이터 N개를 사용하여 회귀모형 T24_168(+Temp)의 회귀계수를 최소자승법(least square method)으로 구한다. 즉, 회귀계수를 구하기 위한 샘플로 요일과 시간대가 동일한 데이터만 사용하므로 한 주에 속한 168개의 시간대별 데이터 중에서 단 한 개의 샘플만 추출한다. 따라서, 예측 시간대마다 사용되는 샘플 데이터가 달라지므로, 예측 시 사용되는 회귀 모형의 계수 값도 달라진다. 반면, 이동평균법과 지수평활법 적용시에는 수요를 예측하고자 하는 d+1일의 시간대(t)와 동일한 시간대의 과거 관측치 중에서 24시간 전 관측치 X_{t-24}, 1주일 전 관측치 X_{t-168}및 2주일 전 관측치 X_{t-336}를 모형에 맞게 선택적으로 추출하여 사용한다.

회귀모형 T24_168(+Temp)의 시간대별 도시가스 수요 예측 오차율을 평가하기 위해 2009년 1월 1일 ~ 2013년 11월 30일 사이의 43,080개 (= 1,795일×24시간)의 과거 데이터 중에서 샘플 크기 N을 제외한 시간대별 실제 수요와 예측값의 차이를 실제 수요로 나눈 APE(Absolute Percentage Error)의 연도별 또는 전체 평균값 MAPE(Mean APE)을 계산한다. 예를 들어, d+1일 t (= 1,⋯,24) 시간대의 수요를 예측하는 경우, N = 10이면 {(d+1) - 7}일 t시간대의 수요가 첫 번째 관측치로 샘플에 포함되며, 마지막 샘플은 {(d+1) - 7×10}일 t시간대의 수요가 된다.

3.3 계산 실험 결과

회귀모형 T24_168(+Temp)의 오차율을 측정한 결과와 해당 모형의 수정된 결정계수(adjusted determinant coefficient) R^2를 Table 6과 Table 7에 나타낸다. 회귀모형 T24_168의 오차율을 나타낸 Table 6에서 다음 내용을 확인할 수 있다.

(1) 회귀모형 T24_168의 R2는 0.959로서 매우 높은 설명력을 가진다.

(2) 샘플크기 N이 증가할수록 전반적으로 오차율이 감소하며, N = 160일 때 오차율이 가장 낮게 나타난다 (MAPE = 4.5%).

Table 6. MAPE(%) of Regression T24_168: R^2 = 0.959

N	5	10	20	40	60	80	100	120	140	160
2009	9.61	6.96	5.99	6.99	-	-	-	-	-	-
2010	9.47	6.75	6.16	5.88	5.07	5.25	5.33	-	-	-
2011	8.48	6.34	5.85	5.54	4.95	4.97	4.98	4.68	5.07	-
2012	7.91	5.88	5.39	4.88	4.49	4.56	4.56	4.49	4.54	4.43
2013	8.32	6.06	5.46	5.15	4.60	4.64	4.66	4.60	4.61	4.56
Total	8.74	6.38	5.75	5.44	4.77	4.79	4.75	4.58	4.64	4.50

회귀모형 T24_168+Temp의 결과를 정리한 Table 7에서 다음 내용을 확인할 수 있다.

(1) 회귀모형 T24_168+Temp의 R^2 = 0.964로서 회귀모형 T24_168의 R^2 = 0.959보다 개선되었다.

(2) 반면, 외기온도를 회귀모형에 표현한 T24_168+Temp의 MAPE은 회귀모형 T24_168의 MAPE(Table 6 참조)과 비교할 때 약 1% 증가하였다. Table 3에 나타낸 바와 같이 외기온도의 경우 시간대에 따라서 도시가스 수요와의 상관계수 절대값이 큰 차이를 보이며(0.861 ~ 0.965), 특히 오전 7시 ~ 오후 7시 사이의 상관계수 절대값은 24시간 전 수요와의 상관계수 0.975 및 168시간 전 수요와의 상관계수 0.942보다 모두 낮게 나타난다. 즉, 시간대별 도시가스 수요는 외기온도보다는 자기상관성 요인이 강하게 나타나므로 외기온도를 독립변수로 추가하는 것은 바람직하지 않아 보인다.

Table 7. MAPE(%) of Regression T24_168+Temp: R^2 = 0.964

N	5	10	20	40	60	80	100	120	140	160
2009	8.95	5.80	6.60	5.01	-	-	-	-	-	-
2010	9.32	5.17	5.71	7.06	6.22	6.40	3.75	-	-	-
2011	8.40	5.06	5.60	6.76	4.96	5.22	5.45	6.01	3.54	-
2012	7.87	4.91	5.52	6.12	4.78	5.13	5.32	5.07	5.34	5.56
2013	8.94	5.09	5.54	6.50	4.80	5.11	5.36	5.11	5.39	5.23
Total	8.68	5.18	5.72	6.54	5.15	5.31	5.36	5.33	5.13	5.40

이동평균 모형 MA24_168 및 MA168_336을 이용한 시간대별 예측 오차율 결과는 Table 8과 Table 9에 각각 나타낸다. 이동평균 모형 MA24_168는 α값이 증가함에 따라 전반적으로 오차율이 감소하며, α = 0.7일 때 오차율이 가장 낮게 나타난다 (MAPE = 7.39%). 즉, 동일한 요일인 168시간 이전 시간대 보다 요일은 다르지만 24시간 이전 시간대의 수요가 시간대별 수요에 더 큰 영향을 주는 것으로 판단할 수 있다. 이동평균 모형 MA24_168이 회귀모형 T24_168과 동일한 독립변수를 참조하고 있음에도 불구하고 오차율 측면에서 큰 차이를 보이는 이유는 이동평균 모형과는 달리 회귀모형 T24_168에서는 (1) 회귀상수 β_0가 기저 수요를 표현할 수 있으며, (2) 최소자승법으로 구한 회귀계수 β_{24}와 β_{168}를 각 시간대별로 최적화하기 때문이다.

Table 8. MAPE(%) of Moving Average MA24_168

α	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
2009	12.28	11.27	10.41	9.67	9.04	8.54	8.20	8.06	8.19
2010	10.49	9.68	9.01	8.47	8.07	7.84	7.77	7.87	8.17
2011	10.23	9.40	8.68	8.07	7.60	7.27	7.13	7.16	7.40
2012	8.72	8.09	7.58	7.20	6.93	6.80	6.82	7.01	7.34
2013	8.53	7.98	7.55	7.22	7.01	6.94	7.01	7.23	7.62
Total	10.06	9.29	8.65	8.13	7.73	7.48	7.39	7.46	7.74

Table 9에 나타낸 이동평균 모형 MA168_336의 최저 오차율 MAPE = 10.8% (α = 0.8)은 모형 MA24_168의 최저 오차율 MAPE = 7.39% (α = 0.7)를 크게 웃도는 수준이다. 즉, 시간대별 도시가스 수요 예측시 가장 최근의 관측 데이터를 참조하는 것이 바람직해 보인다.

Table 9. MAPE(%) of Moving Average MA168_336

α	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
2009	15.40	14.68	14.07	13.58	13.22	13.00	12.91	12.94	13.11
2010	13.56	12.93	12.36	11.87	11.48	11.21	11.08	11.06	11.18
2011	13.53	12.93	12.38	11.91	11.52	11.25	11.07	10.98	11.01
2012	12.26	11.69	11.20	10.76	10.38	10.05	9.78	9.59	9.49
2013	12.06	11.54	11.06	10.63	10.24	9.90	9.61	9.39	9.26
Total	13.37	12.76	12.22	11.75	11.37	11.09	10.89	10.80	10.82

지수평활 모형 ES24 및 ES168를 이용한 시간대별 예측 오차율 결과는 Table 10과 Table 11에 각각 나타낸다. 두 모형 모두 τ = 0.9에서 가장 낮은 오차율 MAPE = 8.41% 및 8.22%를 보이는데, 이동평균 모형 MA24_168의 최저 오차율 MAPE = 7.39% 및 회귀모형 T24_168의 최저 오차율 MAPE = 4.5%와는 큰 차이가 있다. 다만, 이동평균 법과는 달리 지수평활법의 경우 모형 ES24보다 1주일 전 동일 시간대의 예측 결과를 반영하는 ES168의 오차율이 더 낮게 나타나는 점은 주목할 필요가 있다. 이 같은 현상은 24시간 전 수요 데이터뿐만 아니라 요일과 시간대가 동일한 168시간 전 수요 데이터 역시 시간대별 도시가스 수요 예측시 중요한 자기상관 요인임을 설명하는 것으로 판단할 수 있다.

Table 10. MAPE(%) of Exponential Smoothing ES24

τ	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
2009	13.31	11.68	11.07	10.62	10.25	9.89	9.53	9.18	8.87
2010	12.60	10.95	10.49	10.21	9.94	9.66	9.38	9.12	8.88
2011	11.71	9.96	9.45	9.14	8.89	8.65	8.44	8.23	8.03
2012	11.53	9.77	9.28	9.02	8.82	8.61	8.40	8.19	8.00
2013	11.39	9.84	9.43	9.24	9.06	8.86	8.66	8.48	8.31
Total	12.11	10.44	9.94	9.64	9.39	9.13	8.88	8.64	8.41

Table 11. MAPE(%) of Exponential Smoothing ES168

τ	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
2009	40.83	22.91	16.27	13.10	11.36	10.23	9.45	8.92	8.65
2010	35.01	21.66	15.61	12.56	10.87	9.83	9.19	8.81	8.63
2011	32.89	20.72	14.90	11.94	10.23	9.18	8.51	8.08	7.87
2012	32.63	20.30	14.55	11.64	9.98	8.99	8.39	8.02	7.85
2013	31.57	19.49	14.24	11.62	10.10	9.17	8.57	8.21	8.10
Total	34.56	21.02	15.12	12.17	10.50	9.48	8.82	8.41	8.22

4. 결론

본 연구에서는 국내 도시가스 과거 공급량 실적 데이터(2009년 ~ 2013년)를 분석하여 시간대별 도시가스 수요의 특성을 조사하였다. 시간대별 도시가스 수요는 매우 높은 자기상관성을 갖는다. 즉, 수요를 예측하려는 시간대의 24시간, 48시간 및 168시간 전 수요 데이터와 0.942 이상의 매우 높은 상관계수 값을 보인다. 또한 수요를 예측하려는 시간대와 동일한 요일, 동일한 시간대의 도시가스 수요는 외기온도와도 밀접한 상관관계를 갖는다(상관계수 절대값은 요일과 시간대에 따라서 0.861 ~ 0.965임).

이 같은 데이터 분석 결과를 바탕으로 예측 시간대의 24시간 전 및 168시간 전 도시가스 수요 데이터와 외기온도를 독립변수로 표현하는 선형회귀모형을 개발하였다. 2009년 ~ 2013년 사이의 시간대별 과거 실적 데이터를 이용하여 회귀모형의 오차율을 평가한 결과 예측 시간대의 24시간 전 실적 데이터가 주어지는 경우 약 4.5%의 평균절대오차율을 보인다. 특히, 시간대별 외기온도는 시간대별 도시가스 수요와 비교적 높은 상관관계를 보이지만 선형회귀식을 이용한 시간대별 도시가스 수요 예측에는 그다지 도움이 되지 않는 것으로 나타났다. 즉, 시간대별 도시가스 수요 예측에는 외기온도보다는 예측하려는 시간대의 수요와 자기상관성이 높은 24시간 전 및 168시간 전 과거 수요 데이터를 활용하는 것이 바람직하다.

본 연구에서는 최초로 국내 도시가스의 시간대별 수요의 특성을 분석하고 회귀모형을 제안하였으며, 이 연구에서 제시한 데이터 분석 및 회귀모형이 도시가스 시간대별 수요 예측에 관한 보다 심도 있는 후속 연구의 참고자료가 되길 기대한다. 본 연구의 한계점으로는 일반적으로 에너지 수요가 특수일(구정, 추석 등)에는 일반일과는 다른 패턴을 보인다는 점을 시간대별 가스수요 예측모형에 반영하지 못한 점을 들 수 있다. 따라서, 국내 도시가스의 시간대별 수요 예측에 관한 후속 연구에서는 예측 정확도를 높이기 위해 특수일 보정 방법을 개발할 필요가 있다. 또한, 회귀모형을 이용한 예측에서는 일반적으로 샘플크기 및 샘플링 방법에 따라 예측 정확도가 달라질 수 있으므로 샘플크기 및 샘플링 방법에 대한 보다 광범위한 실험을 실시하여 최적의 샘플크기 및 샘플링 방법을 찾을 필요가 있다. 회귀모형을 이용한 예측 방법 외에도 패턴 분석을 이용한 시간대별 수요 예측 방법의 성능도 확인할 필요가 있다. 가령, 계절과 요일에 따라 조금씩 다른 24시간 수요 패턴을 몇 개의 그룹으로 분류하여 샘플 크기를 줄이는 방법을 개발하고 오차율 개선을 위한 다양한 예측 기법을 테스트할 필요가 있다.

References

[1] C. H. Kim, S. M. Lee, S. Y. Lee, D. O. Lim, B. R. Nam, B. H. Lee, KEEI View of Energy Demand, Spring Report of Korea Energy Economics Institute, Vol.17, No.1, 2015.

[2] H. Y. Oh, Forecasting of the Short-Term Demand for the Natural Gas Using Time Series Analysis and Artificial Neural Networks, Master Thesis, Graduate School of Management, KAIST, 1997.

[3] J. J. Her, H. J. Lim, “An analysis of growth factors on the city-gas industry by input-output structural decomposition analysis”, Journal of Energy Engineering, Vol.21, No.2, 158-167, 2012.
DOI: http://dx.doi.org/10.5855/ENERGY.2012.21.2.158

[4] J. S. Kim, C. S. Yang, J. G. Park, “An Empirical Study on the Consumption Function of Korean Natural Gas for City Gas”, Journal of Energy Engineering, Vol.20, No.4, 318-329, 2011.
DOI: http://dx.doi.org/10.5855/ENERGY.2011.20.4.318

[5] S. J. Lee, S. S. Euh, S. H. Yoo, “Estimation of city gas demand function using time series data”, Journal of Energy Engineering, Vol.22, No.4, 370-375, 2013.

DOI: http://dx.doi.org/10.5855/ENERGY.2013.22.4.370

[6] J. S. Park, Y. B. Kim, C. W. Jung, "Short-Term Forecasting of City Gas Daily Demand", Journal of the Korean Institute of Industrial Engineers, Vol.39, No.4, 247-252, 2013.
DOI: http://dx.doi.org/10.7232/JKIIE.2013.39.4.247

[7] N. H. Choi, K. M. Son, T. G. Lee, "Daily Peak Load Forecasting Considering the Load Trend and Temperature", Journal of the Korean Institute of Illuminating and Electrical Installation Engineers, Vol.15, No.6, 35-42, 2001.

[8] S. I. Kong, Y. S. Baek, K. B. Song, J. H. Park, "The Daily Peak Load Forecasting in Summer with the Sensitivity of Temperature", The Transactions of the Korean Institute of Electrical Engineers, Vol.53A, No.6, 358-363, 2004.

[9] S. I. Kong, Y. S. Baek, K. B. Song, "The Peak Load Forecasting During Summer Considering Sensitivity to Temperature", Electronic Technology Reports, Vol.23, 133-140, 2003.

[10] K. B. Song, "Development of Short-Term Load Forecasting Algorithm Using Hourly Temperature", The Transactions of the Korean Institute of Electrical Engineers, Vol.63, No.4, 451-454, 2014.
DOI: http://dx.doi.org/10.5370/KIEE.2014.63.4.451

[11] D. H. Ahn, S. J. Lee, "Daily Load Forecasting Including Special Days Using Hourly Relative Factors", Journal of the Korean Institute of Illuminating and Electrical Installation Engineers, Vol.19, No.5, 94-102, 2005.
DOI: http://dx.doi.org/10.5207/JIEIE.2005.19.5.094

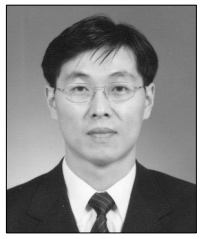
[12] H. S. Koh, C. W. Lee, C. S. Lee, "Power Demand Forecast Classified by Time Period for Optimum Operation of Power Load Equipment", Journal of the Korean Institute of Illuminating and Electrical Installation Engineers, Vol.10, No.1, 75-79, 1996.

[13] H. S. Koh, C. S. Lee, J. K. Choi, B. H. Ji, "Short-term Peak Power Demand Forecasting Using Model in Consideration of Weather Variable", Journal of Korea Institute of Signal Processing and Systems, Vol.2, No.3, 73-78, 2001.

[14] K. B. Song, B. S. Ku, Y. S. Baek, "An Improved Algorithm of the Daily Peak Load Forecasting for the Holidays", The Transactions of the Korean Institute of Electrical Engineers, Vol.51, No.3, 109-117, 2002.

[15] Y. M. Wi, K. B. Song, S. K. Joo, "Data Mining Technique Using the Coefficient of Determination in Holiday Load Forecasting", The Transactions of the Korean Institute of Electrical Engineers, Vol.58, No.1, 18-22, 2009.

한 정 희(Jung-Hee Han) **[정회원]**



- 1996년 2월 : 고려대학교 산업공학과 (공학석사)
- 1999년 8월 : 고려대학교 산업공학과 (공학박사)
- 2000년 3월 ~ 2004년 2월 : LG전자 이동통신연구소 선임연구원
- 2004년 3월 ~ 현재 : 강원대학교 경영학과 교수

<관심분야>
최적화, 응용통계

이 근 철(Geun-Cheol Lee) **[정회원]**

- 1997년 2월 : 한국과학기술원 산업공학과 (공학석사)
- 2003년 8월 : 한국과학기술원 산업공학과 (공학박사)
- 2003년 11월 ~ 2006년 2월 : UC Davis 박사후연구원
- 2006년 3월 ~ 현재 : 건국대학교 경영대학 교수

<관심분야>
제조시스템 운영 및 설계, 에너지 수요예측