

도시가스 수요량 예측을 위한 시계열 모형 개발

최보승¹ · 강현철² · 이경윤³ · 한상태⁴

¹고려대학교 경제연구소, ²호서대학교 정보통계학과, ³고려대학교 경영학과

⁴호서대학교 정보통계학과

(2009년 8월 접수, 2009년 9월 채택)

요약

도시가스 수요량은 강한 계절성을 보이는 자료이다. 따라서 도시가스 수요량을 예측하기 위한 모형 구축에서 가장 중요한 요인은 계절성이다. 또한, 실제 도시가스 수요량에는 추가적인 여러 요인들에 의하여 영향을 받을 수 있는데, 온도, 요일효과, 명절효과, 유효일 수, 수용가수 등이 영향 요인들이다. 본 연구에서는 이와 같은 요인들이 도시가스 수요량에 미치는 영향력의 정도를 파악하고 효율적으로 향후 도시가스 수요량 예측을 위한 시계열 모형을 구축하였다. 적용된 모형은 오차항이 자기상관을 따르는 시계열 회귀모형을 이용하였으며 실제 자료를 이용한 예측결과 매우 우수한 예측력을 보였다.

주요용어: 도시가스, 수요량 예측, 오차자기상관 회귀모형, 유효일수, 온도 탄력성.

1. 서론

에너지 시장은 많은 부분이 국가에 의해서 관리되는 기간산업이다. 국가 전체의 경제상황에 따라 정부에 의하여 각종 규제 및 시장 개입이 빈번하게 이루어지는 산업이다. 이에 에너지 산업의 수요 및 공급량 예측은 국가기관으로부터 장기적인 측면에서 이루어지는 경우가 많이 있었다. 그러나 1998년 이후 국가에 의한 각종 규제가 완화되면서 이와 동시에 시장의 역할이 강조되었고 에너지 산업도 국가기관으로부터 일반 기업으로의 민영화가 진행되었다. 이에 공급자 중심에서 소비자 중심으로 변화가 이루어지게 되었고 (박광수 등, 2004), 이에 따라 장기적인 에너지 수급의 전망 뿐 만 아니라 중·단기적인 에너지 소비 및 공급량에 대한 연구가 부각되었다. 경제성장과 전반적인 생활수준의 향상과 더불어 에너지의 소비절약과 친 환경적인 에너지 소비가 강조되기 시작 하였으며 이에 도시가스의 수요는 전체 에너지 소비의 중요한 요인으로 자리 잡기 시작하였다.

본 연구는 이와 같이 전체 에너지 산업에서 점차적으로 중심적인 역할을 수행하고 있는 도시가스 수요량에 대한 중·단기 예측 모형을 구축하고자 한다. 특히, 본 연구는 정확하게 현실상황을 잘 적용할 수 있는 예측 모형의 구축에 초점을 맞추고자 한다. 수리적인 모형을 기반으로 하여 도시가스와 같은 에너지 산업의 예측 모형의 개발은 계량경제학 모형을 기반으로 하는 연구가 주로 이루어져 왔다. 박광수 등 (2004)과 나인강과 유지철 (2000)은 도시가스 뿐만 아니라 석유, 전력과 같은 우리나라에서 주요하게 소비되는 에너지원에 대한 중·단기 수급 예측 모형을 구축하였다. 그들이 이용한 방법은 각 에너지원의 수요량을 종속변수로 하고 설명변수는 종속변수에 대한 시차변수와 GDP와 같은 각종 거시경제지

⁴교신저자: (336-795) 충청남도 아산시 배방면 세출리, 호서대학교 정보통계학과, 교수.

표들을 소득 또는 가격과 관련된 변수로 구분하고 이와 함께 에너지 소비에 주요하게 영향을 미치는 날씨와 관련된 변수들을 가지고 선형모형을 구축한 것이다. 구체적으로 자기회귀시차분포(autoregressive distributed lag; ADL) 모형 (Pesaran 등, 1998)과 종속변수와 독립변수들 간의 공적분(cointegration) 관계를 이용한 공적분 회귀모형 (Engle과 Granger, 1987)을 이용하였다. 이와 같은 모형에 의한 예측 결과는 정확한 예측력을 보이며 모형의 설명력도 우수하였다. 그러나 미래 시점에 대한 예측을 수행하기 위해서는 모형 구축에 사용된 각종 거시경제 지표에 대한 예측값이 주어져야하는 제한점을 가지고 있다. 이와 같은 독립변수에 대한 제한점을 해결하기 위하여 이충렬과 강운경 (2007)은 자동차수의 변화와 배기량별 자동차 비율의 변화만을 고려하여 승용차의 주 에너지원인 석유의 수요량을 예측하고자 하였다. 구체적인 방법으로 Bass 모형 (Bass, 1969)을 기초로 확산모형으로 자동차수의 변화를 예측한 후 로짓 모형, 프로빗 모형, 이이산선택모형(discrete choice model)을 이용하여 자동차별 배기량 별 비율을 예측한 후 두 결과를 결합하여 최종적으로 에너지 수요량을 예측하고자 하였다. 에너지 예측 모형과 관련하여 보다 최근의 연구로는 양준모와 유상희 (2008)의 연구로 연도별 최대전력수요를 예측하고자 하였고 예측을 위한 모형으로 연도별 자료를 이용한 시차종속 회귀모형을 이용하였다. 그러나 이들은 연구는 상대적으로 복잡한 단계를 거치거나 (이충렬과 강운경, 2007) 오직 연도별 자료만을 이용한다 (양준모와 유상희, 2008)는 한계점을 지니고 있다.

본 연구에서는 도시가스 수요량을 종속변수로, 도시가스 수요량에 영향을 미칠 수 있는 여러 요인들을 독립변수로 하여 시계열 자료가 가지는 자기상관을 고려한 오차항이 자기회귀과정을 따르는 회귀모형을 이용하여 예측 모형을 구축하고자 하였다. 강한 계절성을 보이는 도시가스 수요량의 특징을 파악하고 독립변수에 대한 예측 부담을 최소화 할 수 있도록 모형 구축을 수행하고자 하였다. 또한, 도시가스 수요량에 가장 중요한 요인인 기온변화를 포함한 여러 요인들에 따른 수요량 변화에 대하여 알아보고 이를 적절히 모형화 하는 방법을 함께 고려하였다.

본 논문의 2절에서는 도시가스 수요량이 가지는 자료의 특성을 파악하고 이로부터 도시가스 수요량에 영향을 미칠 수 있는 여러 요인들을 함께 파악해 보고, 3절에서는 중단기 도시가스 예측을 위한 예측모형의 구축과정과 그 결과를 비교하고, 마지막으로 결론을 통해 본 연구의 시사점에 대하여 논의 하였다.

2. 도시가스 수요량 자료의 특성

간단한 시도표를 통하여 도시가스 수요량 자료가 가지는 특성을 살펴보았다. 그림 2.1은 2005년 9월 5일부터 2006년 2월 5일까지 국내 A회사 가스 수요량에 대한 일별 시도표를 나타낸다. 시도표를 살펴 보면 요일별 효과를 뚜렷하게 확인할 수 있다. 2005년 9월의 경우 4일, 11일, 18일, 25일이 일요일이며, 10월의 경우 2일 9일 16일 23일 30일이 일요일이다. 일요일의 가스 수요량은 눈에 띄게 감소해 있음을 볼 수 있다. 토요일 역시 상대적으로 적은 수요량을 보이고 있다. 이에 반하여 화요일부터 금요일까지의 일반적인 근무일에는 비교적 높은 가스 수요량을 보이고 있으며 비슷한 수치를 보여 주고 있다. 월요일의 경우(9월 5일, 12일, 19일, 26일) 근무일 임에도 불구하고 다른 근무일보다 상대적으로 적은 수요량을 보인다. 월요일 새벽시간은 일요일의 연장으로써 많은 산업 및 상업시설들이 가동을 하지 않기 때문에 다른 근무일에 비하여 상대적으로 적은 수요량을 보인다. 이와 같이 일별 가스 수요량은 요일에 따라서 매우 다른 패턴을 보이는 것을 확일 할 수 있으며 월별 수요량 예측을 위한 모형을 개발함에 있어 이와 같은 효과가 충분히 반영되어야 한다.

일별 수요량에서 파악할 수 있는 또 다른 특징은 명절효과이다. 2005년의 경우 추석연휴는 9월 17일, 18일, 19일이고, 2006년의 경우 설연휴는 1월 28일, 29일, 30일이다. 일반 공휴일(일요일을 제외한 공휴일)로는 10월 3일, 12월 25일 그리고 1월 1일이 있는데, 이 때 모두 가스 수요량에 급격한 감소가 있

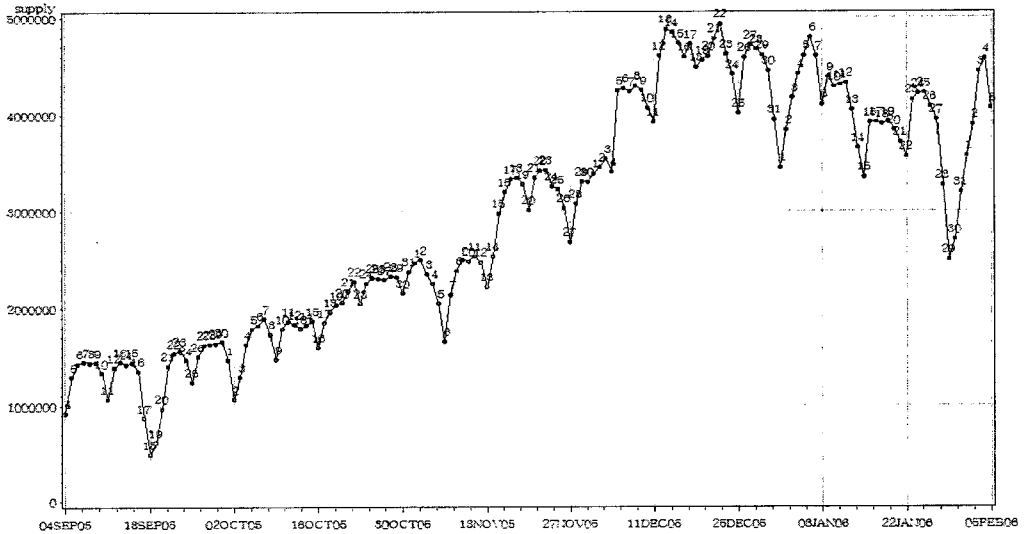


그림 2.1. 일별 가스 수요량(2005년 9월 4일 ~ 2006년 2월 5일)

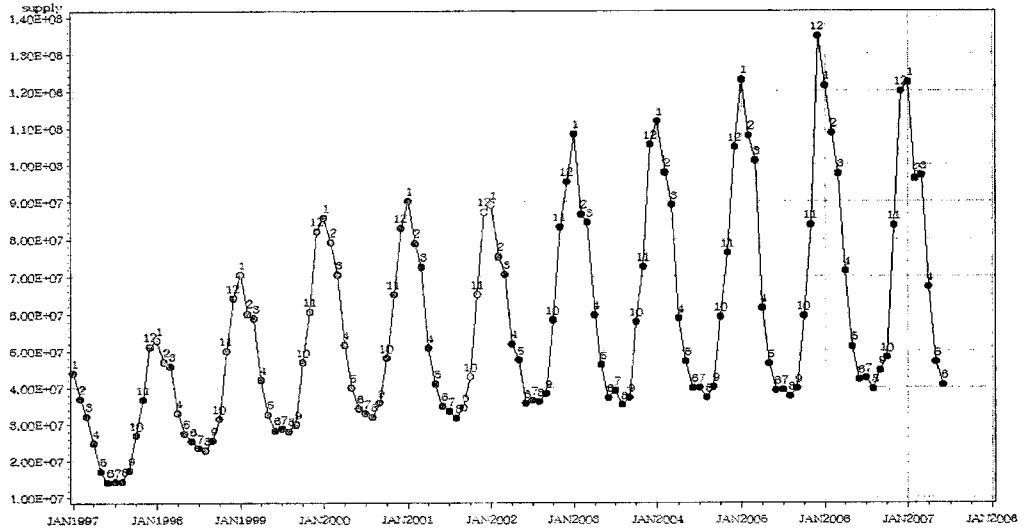


그림 2.2. 월별 가스 수요량(1997년 1월 ~ 2007년 6월)

음을 볼 수가 있다. 이와 더불어 연휴 마지막 날의 다음날 역시 가스 수요량의 감소가 있음을 확인할 수 있다. 따라서 요일효과와 더불어 명절 및 휴일효과 또한 분석을 수행하여 그 효과를 추정해서 모형에 반영해야 한다는 것을 알 수 있다. 그림 2.1에서 살펴볼 수 있는 또 다른 특징은 계절에 따른 효과이다. 시도표의 시작점인 9월부터 끝나는 시점인 2월은 가을에서 겨울로 기온이 꾸준히 감소하는 계절이다. 이에 따라 전반적인 가스 수요량의 추세가 증가하고 있음을 확인할 수 있다. 가스 수요량은 당연히 강한 계절성을 보이는 시계열 자료이다. 이와 같은 계절성은 그림 2.2 월별 가스 수요량의 시도표를 통하여

보다 자세하게 알 수 있다.

그림 2.2는 1997년 1월부터 2007년 6월까지의 월별 가스 수요량 시도표이다. 시도표를 살펴보면 뚜렷한 계절성을 보이고 있는 것을 알 수 있는데, 추운 계절인 12월과 1월을 고점으로 가스 수요량이 증가하며 여름철인 7월과 8월에 가스 수요량이 최저를 보이고 있다. 이와 함께 전반적으로 증가하는 추세를 보이고 있으며, 그 증가 정도는 2004년을 기점으로 하여 증가추세가 두드러지면서 정체되는 모습을 보이고 있다. 수요량 추세의 증가와 함께 살펴볼 수 있는 점은 전체적인 수요량의 증가와 함께 겨울철의 사용량이 상대적으로 크다는 것으로 수요량 자료의 분산이 시간의 흐름에 따라 증가하고 있다고 볼 수 있는 것을 볼 수 있다. 따라서 가스 수요량 예측 모형을 구축하는데 있어서 추세와 분산의 증가 특성을 고려한 모형을 구축해야 함을 알 수 있다.

3. 도시가스 수요량 예측 모형

2절에서 살펴본 도시가스 수요량 자료의 특성에 따라 수요량에 영향을 미칠 수 있는 각 요인들을 파악해 보고 월별 도시가스 수요량 예측을 위한 모형 구축 과정을 소개하도록 하겠다.

3.1. 유효일수의 산정

그림 2.1의 시도표에서 살펴본 바와 같이 도시가스 수요량은 요일에 따라 큰 편차를 보인다. 요일을 포함한 공휴일이 많고 적음에 따라 수요량의 차이는 다르게 나타날 것이다. 특히 설과 추석과 같은 명절이 포함되는 경우 수요량의 감소는 더욱 심할 것이다. 따라서 월별 도시가스 수요량을 예측하는데 있어서 이와 같은 요일 구성의 차이는 수요량의 영향을 줄 수 있는 중요한 요인 가운데 하나가 될 수 있다. 본 연구에서는 월별로 상이한 근무일수를 수요량 예측 모형에 반영할 수 있도록 모형화 하고자 하였는데, 이를 유효일수라 칭하였고 이를 통하여 월별 유효일 수를 계산하고자 하였다. 적절한 유효일 수의 산정은 월별 수요량을 측정하는데 보다 효과적인 역할을 수행할 수 있다.

유효일수를 산출하는 과정은 다음과 같다. 유효일수의 산정은 일별 가스 수요량 자료를 이용하였다. 그림 2.1에서 살펴본 바와 같이 일별 수요량의 특징은 일주일을 단위로 하는 계절성이고 이러한 계절성이 추세에 따라 증가 또는 감소함을 볼 수 있다. 이와 더불어 빈번하게 발생하는 공휴일과 명절에 따른 불규칙 효과도 함께 볼 수 있고 이에 승법분해모형을 적용하여 계절요인과 불규칙요인만을 분해하는 것이 적절하다 할 수 있다 (박유성과 김기환, 2002). 먼저 시점 t 에서의 일별 수요량을 day_t 라 하고 일별 수요량 자료가 가지고 있는 요일에 따른 주기를 반영하여 중심화 된 14개항 이동평균을 계산하였다. 요일 주기를 고려하면 7개항 이동평균을 이용하는 것이 보다 타당하지만 공휴일과 명절에 따른 불규칙 효과를 보다 잘 반영하기 위하여 14개항 이동평균을 적용하였다. 이동평균을 계산한 후 다음과 같이 일별 수요량에서 나누어 줌으로써 계절성분과 불규칙 성분만을 가지고 있는 시계열 z_t 를 계산할 수 있다.

$$z_t = \frac{day_t}{\frac{1}{2} \left(\sum_{k=-7}^6 \frac{day_{t+k-1}}{14} + \sum_{k=-6}^7 \frac{day_{t+k}}{14} \right)}$$

이제 분해된 시계열 z_t 가 1보다 크면 그날의 가스 수요량이 주변일 보다 크다고 할 수 있고 1보다 작으면 수요량이 최근 주변일보다 작다고 할 수 있다. 본 연구에서는 이 분해 시계열 z_t 를 표준화된 일별 수요량으로 고려하였고 이를 이용하여 가스 수요에 대한 유효일수 산정에 이용하였다. 유효일수의 모형 적합과 유효일수에 대한 예측치를 계산하기 위하여 선형 회귀모형을 이용하였다. z_t 를 반응변수로 하고 일요일을 참조변수로 하여 요일별 가변수, 일요일을 제외한 기타 공휴일, 추석과 설을 가변수로 하여 선

표 3.1. 월별 유효일수

| | 2001년 | 2002년 | 2003년 | 2004년 | 2005년 | 2006년 | 2007년 |
|-----|-------------|-------------|-------------|-------------|-------------|-------------|-------------|
| 1월 | 30.4 | 31.3 | 31.1 | 30.4 | 31.1 | 30.1 | 30.4 |
| 2월 | 28.2 | 27.4 | 27.6 | 29.1 | 27.4 | 28.2 | 28.2 |
| 3월 | 31.3 | 31.0 | 31.0 | 31.3 | 31.3 | 31.3 | 31.3 |
| 4월 | 30.0 | 30.2 | 30.2 | 30.3 | 30.2 | 30.1 | 30.0 |
| 5월 | 31.2 | 31.2 | 31.2 | 30.9 | 30.9 | 31.2 | 31.2 |
| 6월 | 30.2 | 30.0 | 30.0 | 30.3 | 30.2 | 30.2 | 30.2 |
| 7월 | 31.0 | 31.3 | 31.3 | 31.3 | 31.0 | 31.0 | 31.0 |
| 8월 | 31.3 | 31.3 | 31.0 | 31.0 | 31.3 | 31.3 | 31.3 |
| 9월 | 29.6 | 28.4 | 28.6 | 28.7 | 28.7 | 30.3 | 28.4 |
| 10월 | 30.0 | 31.3 | 31.3 | 31.0 | 31.0 | 29.3 | 31.3 |
| 11월 | 30.3 | 30.3 | 30.1 | 30.3 | 30.3 | 30.3 | 30.3 |
| 12월 | 31.0 | 31.0 | 31.2 | 31.3 | 31.3 | 31.0 | 31.0 |

형 회귀모형에 적합시킨다. 추석과 설 명절을 고려할 때는 각 명절기간의 바로 다음날도 명절에 포함하였다. 식 (3.1)은 모형적합 결과이다.

$$\hat{z}_t = 0.8230 + 0.1724D_{1,t} + 0.2274D_{2,t} + 0.2378D_{3,t} + 0.2348D_{4,t} + 0.2364D_{5,t} + 0.1860D_{6,t} - 0.0783H_t - 0.4208CHU_t - 0.2087SUL_t. \quad (3.1)$$

여기서 $D_{1,t} \sim D_{6,t}$ 는 요일효과, H_t 는 기타 공휴일 효과, CHU_t 와 SUL_t 는 추석과 설 명절효과를 나타낸다.

다음 단계로 회귀모형 적합 후 반응변수의 적합값을 월별로 합하여 이를 월별 유효일수로 계산하였다. 월별로 계산된 유효일수 가운데 일부가 표 3.1에 정리되어 있다. 표 3.1의 결과를 살펴보면, 1월과 2월, 그리고 9월과 10월에서 명절의 포함여부에 따라 상이한 유효일수가 계산되었음을 볼 수 있다. 설 연휴가 1월 이었던 경우는 2001년, 2004년, 2006년, 2007년인데, 다른 연도의 1월보다 작은 유효일수를 가진다. 추석의 경우 2001년과 2006년 10월에 추석이 있었다. 그리고 2004년 2월은 윤달로써 다른 연도의 2월보다 더 큰 유효일수를 가짐을 볼 수 있다. 본 모형에서 독립변수들은 모두 시간의 흐름으로부터 계산된 가변수들이다. 따라서 미래시점에서의 각 가변수와 적합된 모형을 이용하여 미래시점에 대한 유효일수 예측치를 계산할 수 있으며 이를 가스 수요량의 예측치를 구하는데 이용할 수 있다.

3.2. 추세효과의 모형화

그림 2.2의 시도표에서 살펴본 특징 가운데 하나는 도시가스 수요량이 일정정도의 선형추세를 가지고 있다는 점이다. 가스 수요량 예측 모형을 구축하는데 있어서 이러한 추세효과도 중요한 요인 가운데 하나가 될 수 있다. 추세효과를 반영하는 방법은 여러 가지가 있을 수 있겠으나 본 연구에서는 외부 변수 가운데 하나인 수용가수를 이용하여 추세효과를 반영하고자 하였다. 수용가수는 일종의 가구단위로써 도시가스를 공급받은 가구, 산업체, 상업시설 등을 의미한다. 그림 3.1의 시도표에서 검은 점으로 표시된 부분이 월별 수용가수를 나타낸다. 시도표에서 볼 수 있는 바와 같이 수용가수는 뚜렷한 증가 추세를 가지고 있으며 이러한 수용가수를 적절히 모형화 한다면 이를 가스 수요량 예측모형에서 추세효과로 이용할 수 있다.

수용가수 자료 또한 기본적으로 시계열 자료이므로 자기상관이 존재할 수 있다. 자기상관이 존재하면서 시계열 자료의 변동이 느린 경우 다항 추세를 이용한 시계열 회귀모형을 통하여 예측 모형을 구축하

표 3.2. 수용가수(추세) 예측모형 적합결과

| Variable | Estimate | Standard Err. | t-value |
|-----------|----------|---------------|-----------|
| Intercept | 207,319 | 7,379 | 28.10**** |
| t | 1,664 | 467.5 | 3.56*** |
| t^2 | 19.04 | 7.9 | 2.41* |
| t^3 | -0.0603 | 0.0373 | -1.62 |

* < 0.05, ** < 0.01, *** < 0.001, **** < 0.0001

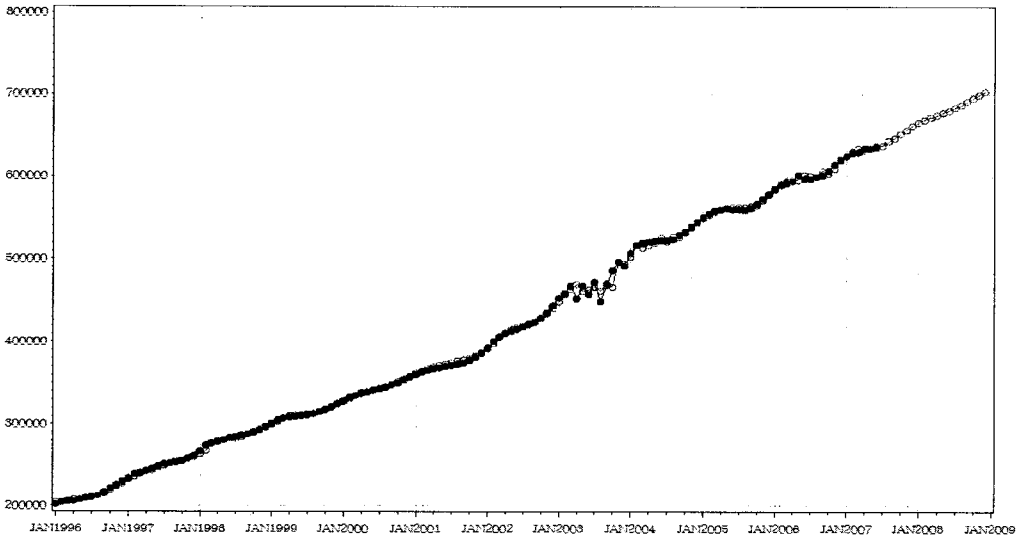


그림 3.1. 수용가수 예측 시도표

는 것이 효율적이라 할 수 있다 (박유성과 김기환, 2002). 본 연구에서는 3차 다항추세모형을 이용하여 모형 적합을 수행하였고 시계열자료의 자기상관을 반영하기 위하여 오차자기회모형을 이용하여 모형 적합을 수행하였다. 다음 표 3.2는 다항회귀모형 적합 결과이며 시차에 대한 3차항을 제외하고 모두 유의하다. 오차항의 자기상관을 반영하기 위하여 AR(5) 모형이 적용되었으며 적합모형의 R^2 는 99.6%이다. DW 통계량 값은 1.90이고 이에 대한 p -value는 0.20으로 추가적인 자기상관은 남아있지 않다고 할 수 있다. 그림 3.1은 예측 시도표이다. 그림 3.1에서 동그라미로 표시된 부분이 다항추세 모형에 의하여 예측된 결과이다. 실제값과 예측값의 차이가 거의 없고 매우 정확한 예측이 수행되었음을 볼 수 있다. 이후 본 연구에서는 3차 다항추세모형에 의하여 예측된 수용가수 값을 가지고 도시가스 수요량 자료의 추세를 반영하였고 수용가수 예측값을 이용하여 도시가스 수요량 예측 모형에 이용하고자 하였다.

3.3. 온도효과의 반영

기온변화는 가스수요량의 변화에 가장 중요한 요인이다. 그림 2.2의 월별 수요량 시도표에서 보이는 뚜렷한 주기는 월별 기온변화에 기인한 것이라 할 수 있다. 기온의 하강은 가스 수요량의 증가를 가져올 것이고 기온 상승은 반대로 가스수요량의 감소를 초래할 것이다. 그러나 이러한 온도변화에 따른 가스 수요량의 탄력성은 계절에 따라 다른 특성을 보일 수 있다. 본 연구에서는 온도변화에 따른 가스 수요량의 탄력성을 구하기 위하여 회귀모형을 이용한 가격탄력성을 계산하는 방법을 이용하였다 (유병철,

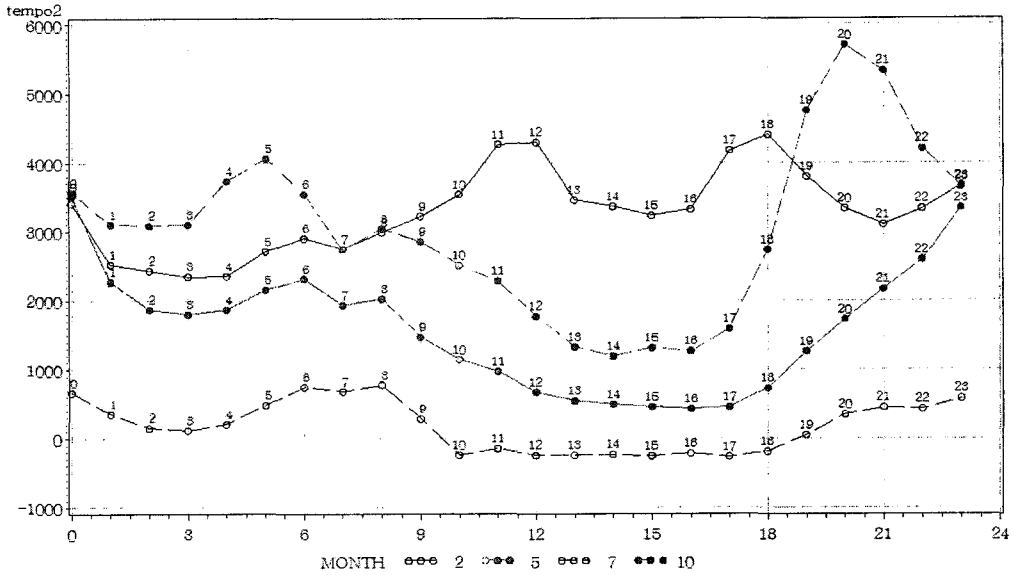


그림 3.2. 월별 온도 탄력성 도표

1996; 나인강과 유지철, 2000). 먼저 한 시간 단위로 측정된 가스 수요량 자료와 온도 자료를 이용하여 전체 자료를 시간별, 월별로 구분하였다. 즉 전체 자료를 288개의 자료로 구분한 후 각 구분된 자료로부터 가스 수요량을 종속변수로 하고 같은 시간에 측정된 기온을 독립변수로 하는 단순 선형회귀모형을 적합하였다. 회귀모형 적합 후 계산된 독립변수에 대한 회귀계수 값을 월별, 시간대별 온도변화가 가스 수요량에 미치는 탄력성으로 정의 하였다. 다음 그림 3.2는 12개월 가운데 계절별로 대표적인 2월, 5월, 7월, 10월에 대하여 시간대별 탄력성을 나타낸 도표이다. 여기서는 분석의 편의를 위하여 각 계수의 부호를 반대로 하여 정의하였다. 이와 같이 반대 부호를 정의 한 것은 다음 표 3.3도 마찬가지이다. 도표에서 ‘실선-동그라미’는 2월을 나타내며, ‘실선-점’은 5월, ‘점선-동그라미’는 7월, ‘점선-점’은 10월을 나타낸다.

시도표상에 가장 낮은 위치를 차지하고 있는 7월의 경우 다른 계절에 비하여 온도변화에 따른 수요량이 전반적으로 비탄력적이라 할 수 있다. 특히 낮 시간의 경우 계수의 값이 거의 0에 가깝고 유의성 검정 결과 유의하지 않았다. 즉 온도변화에 대한 탄력성에 있어서 전혀 탄력적이지 않으며 온도변화에 따른 수요량의 변화가 없다고 할 수 있다. 2월의 경우 전반적으로 계수값이 큰 값을 가지고 있으며 주목할 점은 다른 계절과 달리 한 낮의 경우도 높은 계수값을 유지하고 있음을 볼 수 있다. 추운 계절의 경우 낮 시간의 활동량에 따라 이 시간대의 원활한 활동을 보장하기 위하여 난방용 가스 수요량이 증가하고 이에 따른 탄력성이 증가한다고 볼 수 있다. 5월과 10월의 경우 전체적인 시간대별 패턴은 비슷한 양상을 보이고 있으며 두 계절 모두 시간대에 따라 매우 상이한 탄력성을 보이고 있다. 낮 시간에 상대적으로 가장 낮은 값을 보이고 있으며 저녁 시간과 새벽 시간에 높은 값을 보이고 있다. 이 시기는 또한 낮과 밤의 일교차가 매우 큰 계절이라 할 수 있다. 따라서 이와 같은 일교차의 변화에 따른 탄력성이 다른 계절에 비하여 훨씬 크게 작용하게 될 것이다. 다음 표 3.3은 추정된 탄력성을 월별로 정리하여 탄력성의 평균(Mean)과 범위(Range)를 정리한 것이다. 그림 3.2에서 살펴본 바와 같이 겨울철(12월, 1월, 2월) 그리고 환절기인 3월과 10월, 11월에서 높은 탄력성을 보인다. 특히 10월의 경우 가장 큰 탄력성의 편차를 보이고 있다.

표 3.3. 월별 탄력성 비교

| 통계량 | 1월 | 2월 | 3월 | 4월 | 5월 | 6월 | 7월 | 8월 | 9월 | 10월 | 11월 | 12월 |
|-------|------|------|------|------|------|------|------|------|------|------|------|------|
| Mean | 2849 | 3288 | 3195 | 2335 | 1593 | 202 | 182 | 775 | 657 | 3019 | 3246 | 3344 |
| Range | 1067 | 2056 | 1917 | 3093 | 3077 | 1105 | 1038 | 1287 | 3230 | 4513 | 2801 | 2537 |

결론적으로 계절별로 온도변화가 수요량의 변화에 미치는 효과는 상당히 다르다고 할 수 있으며 이와 같은 효과가 온도를 이용한 수요량의 예측에 반영되어야 보다 정확한 예측모형의 구축이 가능하다 할 수 있다. 이를 위하여 본 연구에서는 가스 수요량 예측 모형을 구축하는데 있어서 계절별로 다른 온도의 효과를 반영하기 위하여 온도효과와 월별 가변수의 상호작용 요인을 모형에 추가하여 계절별 온도효과를 반영하고자 하였다.

3.4. 오차자기회귀모형을 이용한 가스 수요량 예측 모형 구축

앞 절에서 살펴본 바와 같이 도시가스 수요량 자료는 온도 뿐만 아니라 여러 요인들에 의해서 영향을 받는다. 가스 수요량에 영향을 미치는 여러 요인들을 적절히 반영하고 시계열자료의 특성을 가지고 있는 가스 수요량의 특징을 반영하여 본 연구에서 제안하는 월별 가스 수요량 예측모형은 다음과 같다.

$$y_t = TR_t + ZM_t + SN_t + TEMPO_t + \epsilon_t,$$

$$\epsilon_t = \phi_1\epsilon_{t-1} + \phi_2\epsilon_{t-2} + \phi_3\epsilon_{t-3} + \dots + \nu_t, \quad \nu_t \sim N(0, \sigma^2),$$

여기서, 종속변수 y_t 는 월별 가스수요량을 나타낸다. 또한, 요인 TR_t 는 시점 t 에서 반응변수에 영향을 주는 추세요인을 나타낸다. 추세는 시간의 흐름에 전체 시계열에 영향을 주는 장기변동으로써 고정적인 효과를 나타내며 많은 경우에 시간 t 가 추세 요인으로 이용되는 경우가 많다. 그러나 본 연구에서는 위에서 언급한 바와 같이 추세를 반영하기 위하여 수용가수의 예측값을 이용하였다. 그림 3.1로부터 실제 시간의 흐름과 수용가수의 관계를 살펴보면 매우 직선적인 관계를 보이고 있음을 보았고, 그림 2.2로부터 월별 가스 수요량은 서서히 증가하다가 그 증가추세가 줄어드는 경향을 볼 수 있었다. 이를 반영하기 위하여 월별 수용가수를 x_t 라 할 때 이를 로그변환 시킨 후 $TR_t = \log x_t$ 를 이용하였다. 다음으로 ZM_t 는 월별로 계산된 유효인수를 나타낸다.

SN_t 는 계절 요인을 나타내는데, 그림 2.2의 월별 시도표에서와 같이 월별 가스 수요량은 계절에 따라 뚜렷한 주기를 보이고 있으며 시간이 지남에 따라 계절효과의 변동량이 증가 추세를 보이는 확산 계절변동(increasing season variation)을 보인다. 확산 계절변동과 모형의 간소화를 위하여 삼각함수(trigonometric function)를 이용하여 계절효과를 반영하도록 하였다. 즉 12개월을 주기로 하는 삼각함수와 시점의 상호작용 효과인 $t \times \sin(2\pi t/12)$ 와 $t \times \cos(2\pi t/12)$ 를 설명변수로 추가하였다.

$TEMPO_t$ 는 가스 수요량에 가장 중요한 요인인 온도효과를 나타낸다. 온도효과를 반영하기 위하여 기본적으로 월별 평균 기온을 계산하여 이를 온도효과로 사용하였다. 그리고 3.3절에서 확인한 바와 같이 온도효과가 수요량에 영향을 미치는데 계절에 따라 상대적으로 다른 탄력성을 보이고 있는 것을 반영할 수 있도록 다음과 같이 온도효과를 세분화 하였다.

$$TEMPO_t = \beta_m MTEMPO_t + \beta_{D_1} TEMPO_t D_{1,t} + \beta_{D_2} TEMPO_t D_{2,t} + \beta_{D_3} TEMPO_t D_{3,t}$$

$$+ \beta_{D_4} TEMPO_t D_{4,t} + \beta_{D_5} TEMPO_t D_{5,t} + \beta_{D_6} TEMPO_t D_{6,t} + \beta_{D_7} TEMPO_t D_{7,t}$$

$$+ TEMPO_t D_{8,t} + \beta_{D_9} TEMPO_t D_{9,t} + \beta_{D_{10}} TEMPO_t D_{10,t} + \beta_{D_{11}} TEMPO_t D_{11,t},$$

여기서 $MTEMPO_t$ 는 t 월의 평균온도를 나타내고, $TEMPO_t D_{1,t}$ 에서 $TEMPO_t D_{11,t}$ 는 월별로 온도

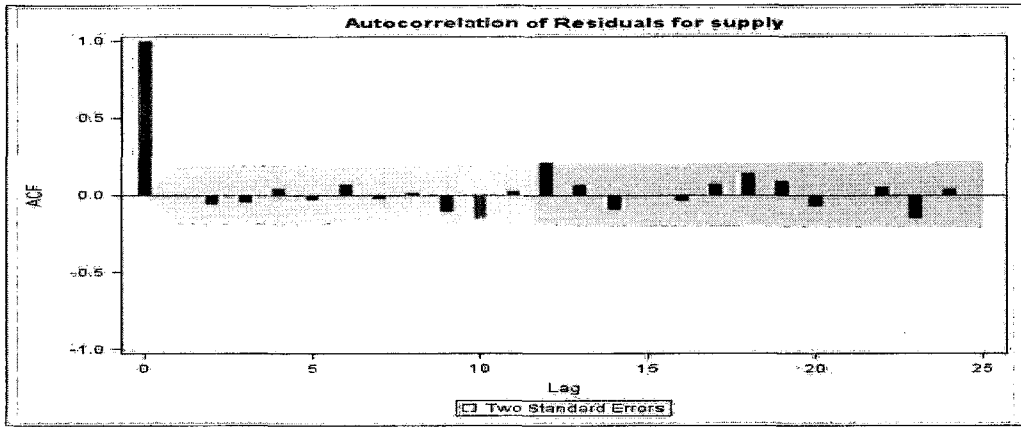


그림 3.3. 잔차에 대한 자기상관도표(ACF)

변화가 수요량에 미치는 탄력성이 서로 다르게 미치는 정도를 반영하기 위한 변수로써 계절효과에서 12월을 참조변수로 하는 월별 가변수와 해당 월의 월 평균 온도의 상호작용 효과를 나타낸다. 따라서 각 월별로 온도변화에 따른 수요량의 변화 정도는 월별로 다르게 계산되게 된다. 예를 들어 1월의 경우 온도변화에 따른 수요량의 변화는 $\beta_m + \beta_{D_1}$ 으로 계산되며 2월은 $\beta_m + \beta_{D_2}$ 로 계산된다. 유사한 방법에 따라 11월의 효과는 $\beta_m + \beta_{D_{11}}$ 로 계산된다. 마지막으로 ϵ_t 는 회귀모형의 오차항을 나타내는데 시계열 자료가 가지는 자기상관성을 반영하기 위하여 오차항에 대한 자기회귀모형을 고려하였다. 자료분석 결과 오차항이 시차 1, 5, 9의 자기회귀모형을 따르는 것을 확인하였다. 최종 잔차인 ν_t 는 백색잡음항을 나타낸다.

이와 같이 오차항이 자기회귀과정을 따르는 회귀모형을 이용하여 모형을 적합한 이후에 예측 정확도를 측정하기 위하여 실제값 대비 절대오차 비율을 사용하였다. 이 값들에 대한 평균치인 평균절대오차를 예측모형의 정확도로 사용할 수 있다. 절대오차비율은 다음과 같으며

$$AE_t = \frac{|y_t - f_t|}{y_t} \times 100,$$

여기서 y_t 와 f_t 는 각각 월별 가스 수요량과 그에 대한 예측값을 나타낸다. 이 오차비율의 평균치인 평균절대오차는 다음과 같다.

$$MAE = \frac{\sum_{t=1}^T AE_t}{T}.$$

이제 실제 가스 수요량 자료를 이용한 분석 결과를 살펴보도록 하겠다. 전체 자료기간 가운데 1997년 이전 자료는 분석 기간에서 제외시키고 실제 분석을 위한 자료는 1998년 1월부터 2007년 6월까지의 자료를 이용하였다. 월별 도시가스 수요량에 영향을 미칠 수 있는 모든 요인을 고려한 후에 구축된 모형에 대한 각 설명변수에 대한 분석 결과는 표 3.4와 같다. 전체 모형의 설명력은 약 99.4%로써 설명력이 매우 높게 나타났으며, 잔차에 남아있는 자기상관성을 측정하기 위한 DW 통계량 값은 약 1.88이고 이에 대한 p -value는 0.27이다. 다음 그림 3.3과 3.4는 잔차에 대한 자기상관도표와 편자기상관도표로써 추가적인 자기상관은 남아있지 않았다고 판단할 수 있다 (박유성과 김기환, 2002).

표 3.4의 결과를 살펴보자, 먼저 유효일수(ZM_t), 수용가수(log x_t), 확산계절주기($t \times \sin(2\pi t/12)$), $t \times \sin(2\pi t/12)$)에 대한 검정결과 모두 통계적으로 유의하다. 즉 이러한 효과들이 모두 가스 수요량에 중

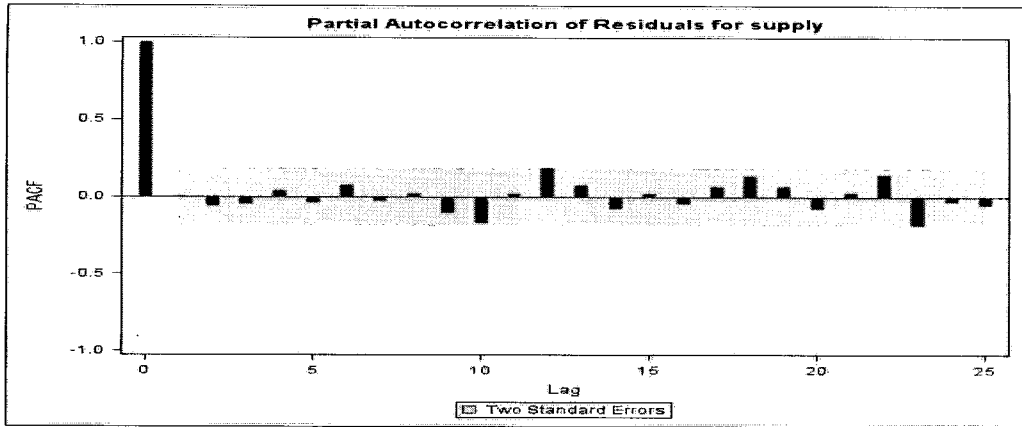


그림 3.4. 잔차에 대한 편자기상관도표(PACF)

표 3.4. 추정된 회귀계수

| Variable | Estimate | Standard Err. | t-value |
|----------------------------|---------------------|-------------------|------------|
| Intercept | -5.30×10^8 | 3.5×10^7 | -15.29**** |
| ZM_t | 2,568,397 | 456,324 | 5.63**** |
| $\log x_t$ | 41,419,366 | 2,485,858 | 16.66**** |
| $t \times \sin(2\pi t/12)$ | 133,789 | 11,900 | 11.25**** |
| $t \times \cos(2\pi t/12)$ | 202,491 | 11,744 | 17.24**** |
| β_m | -2,919,900 | 391,101 | -7.46**** |
| β_{D_1} | -1,019,094 | 634,044 | -1.61 |
| β_{D_2} | -151,537 | 540,633 | -0.28 |
| β_{D_3} | 712,215 | 418,745 | 1.70 |
| β_{D_4} | 801,134 | 383,212 | 2.07* |
| β_{D_5} | 1,058,756 | 369,972 | 2.86** |
| β_{D_6} | 1,529,374 | 374,480 | 4.08**** |
| β_{D_7} | 1,706,134 | 375,752 | 4.54**** |
| β_{D_8} | 1,585,121 | 380,634 | 4.16**** |
| β_{D_9} | 1,365,561 | 377,381 | 3.62*** |
| $\beta_{D_{10}}$ | 753,018 | 352,827 | 2.13* |
| $\beta_{D_{11}}$ | 388,064 | 314,039 | 1.24 |

* < 0.05, ** < 0.01, *** < 0.001, **** < 0.0001

요한 요인이라 할 수 있다. 그리고 각 계수의 부호를 살펴보았을 때 이 또한 가스 수요량에 미치는 효과와 같은 방향으로 움직이고 있음을 볼 수 있다. 모두 적절한 방향성을 보이고 있다. 다음으로 온도 효과를 살펴보자. 월평균 온도효과에 대한 계수값이 2,919,900이므로 12월은 온도가 1도 하강함에 따라 2,919,900만큼 가스 수요량이 증가한다 할 수 있다. 1월의 경우 온도가 1도 감소할 때 $2,919,900 + 1,019,094 = 3,938,944$ 만큼 가스 수요량이 증가한다 할 수 있고 2월의 경우 $2,919,900 + 151,537 = 3,071,437$ 만큼 증가한다 할 수 있고 7월의 경우는 $2,919,900 - 1,706,134 = 1,213,766$ 만큼 증가한다 할 수 있다. β_{D_2} 는 2월의 온도 효과를 나타내는데 추정치 값이 -151,537로써 12월에 비하여 온도하강에 따른 수요량의 증가가 더 큰 것으로 볼 수 있다. 이는 그림 2.2와 다른 결과를 보인다. 월별 가스 수요량을 살펴보면 12월의 수요량이 2월의 수요량보다 더 큰 것을 볼 수 있다. 이와 같은 현상이 발생하게

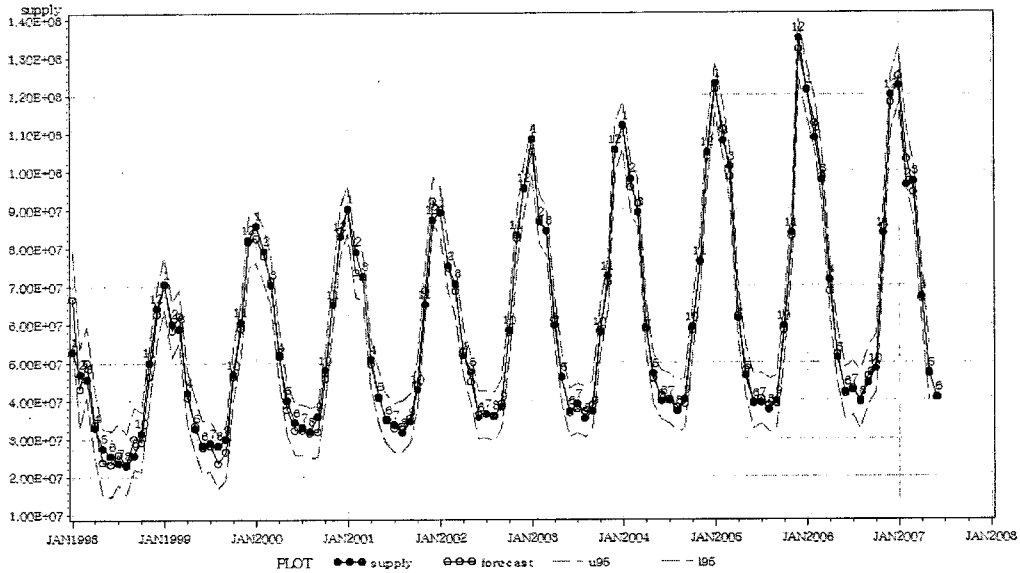


그림 3.5. 월별 온예측결과 시도표(1998년 1월 ~ 2007년 6월)

된 이유는 실제 평균기온을 살펴보면 12월이 평균기온이 2월의 평균 기온보다 더 낮음을 볼 수 있고 유효일수에서 12월의 유효일수가 2월의 유효일수 보다 항상 큰 것을 고려한다면 β_{D_2} 계수값의 부호가 타당하다 할 수 있다.

그림 3.5는 실제 수요량과 예측결과를 함께 표시한 시도표이다. 검은 점은 실제 월별 수요량을 나타내고 동그라미는 예측값을 나타낸다. 점선으로 표시된 부분은 예측값에 대한 95% 신뢰구간을 나타낸다. 예측값과 실제값을 구분하기 힘들 정도로 매우 정확한 예측이 이루어진 것을 볼 수 있다. 또한, 전체 자료에 대한 MAE는 3.2%로서 이 또한 높은 정확도를 나타내 주고 있다.

다음 표 3.5는 분석시점 가운데 마지막 1년인 2006년 7월부터 2007년 6월까지의 실제 수요량과 예측치를 정리한 것이다. 그리고 분석 이후 시점인 2007년 7월부터 2008년 6월까지의 예측치를 함께 정리한 것이다. 여기서 예측 모형을 이용하여 미래시점의 예측치를 계산하기 위해서는 독립변수들에 대한 예측치가 필요하게 된다. 예측모형에 사용된 독립변수들 가운데 온도효과를 제외한 다른 독립변수들은 모두 시간의 흐름에 따라서 계산될 수 있는 값이므로 적절한 미래시점 값을 구할 수 있으며 3.1절과 3.2절에서 자세히 설명하였다. 그러나 온도에 대한 예측치를 구하는 것은 결코 수월한 문제가 아니다. 본 연구에서는 온도에 대한 미래시점의 예측치를 구하기 위하여 분석기간 가운데 최근 3년간의 온도의 평균값을 계산하여 온도 효과에 대한 예측치로 이용하였다. 표 3.5에서 두 번째 열은 실제 수요량을 나타낸다. 세 번째부터 여섯 번째 열은 본 연구에서 제안하고 있는 모형으로부터 계산된 예측치, 95% 신뢰구간 하한, 95% 신뢰구간 상한, MAE를 나타낸다. 마지막 두 열은 모형 비교를 위하여 가스 수요량에 가장 큰 영향을 미치는 온도효과만을 독립변수로 고려한 단순 회귀모형으로부터 적합한 예측치와 이에 대한 MAE를 나타낸다. 이 온도효과 하나만을 고려한 모형에서 회귀계수는 매우 유의하였고 ($p\text{-value} < 0.0001$) 결정계수는 약 65% 나타났다. 하지만 전반적인 예측결과는 매우 부정확한 것을 확인할 수 있다. 이에 반하여 본 연구의 예측모형은 2007년 2월을 제외하고는 모두 3% 이하의 MAE를 보이며 매우 정확한 예측을 수행하였다 할 수 있다. 2007년 2월의 경우 평균 기온이 약 6.2도로써 예년

표 3.5. 예측 결과 비교

| Time | 오차자기회귀예측모형 결과 | | | | | 온도회귀모형 결과 | |
|--------|---------------|-------------|-------------|-------------|-------|------------|--------|
| | 수요량 | 예측치 | 95%하한 | 95%상한 | MAE | 예측치2 | MAE2 |
| Aug-06 | 39,674,806 | 39,970,286 | 32,603,134 | 47,337,438 | 0.74% | 21,359,969 | 46.16% |
| Sep-06 | 44,773,844 | 46,266,495 | 38,942,967 | 53,590,023 | 3.33% | 40,909,444 | 8.63% |
| Oct-06 | 48,377,770 | 49,713,163 | 42,138,643 | 57,287,683 | 2.76% | 45,975,714 | 4.97% |
| Nov-06 | 83,760,585 | 85,459,504 | 77,999,281 | 92,919,727 | 2.03% | 67,688,795 | 19.19% |
| Dec-06 | 120,038,635 | 117,987,237 | 110,720,488 | 125,253,985 | 1.71% | 84,650,920 | 29.48% |
| Jan-07 | 122,460,406 | 124,856,267 | 117,396,009 | 132,316,525 | 1.96% | 85,581,939 | 30.11% |
| Feb-07 | 96,408,116 | 103,057,133 | 95,532,879 | 110,581,387 | 6.90% | 78,108,388 | 18.98% |
| Mar-07 | 97,223,604 | 94,336,096 | 87,166,633 | 101,505,560 | 2.97% | 71,242,029 | 26.72% |
| Apr-07 | 67,260,983 | 66,617,430 | 59,452,972 | 73,781,887 | 0.96% | 57,515,504 | 14.49% |
| May-07 | 46,996,468 | 47,391,010 | 39,928,866 | 54,853,155 | 0.84% | 42,737,747 | 9.06% |
| Jun-07 | 40,784,211 | 41,680,096 | 34,468,446 | 48,891,746 | 2.20% | 34,758,601 | 14.77% |
| Jul-07 | . | 42,755,435 | 35,543,934 | 49,966,936 | . | . | . |
| Aug-07 | . | 40,461,312 | 32,427,803 | 48,494,821 | . | . | . |
| Sep-07 | . | 40,609,824 | 32,242,774 | 48,976,873 | . | . | . |
| Oct-07 | . | 61,041,504 | 52,462,354 | 69,620,653 | . | . | . |
| Nov-07 | . | 90,753,629 | 81,932,901 | 99,574,356 | . | . | . |
| Dec-07 | . | 127,903,292 | 119,232,791 | 136,573,794 | . | . | . |
| Jan-08 | . | 133,055,875 | 124,394,403 | 141,717,348 | . | . | . |
| Feb-08 | . | 119,058,339 | 110,259,538 | 127,857,140 | . | . | . |
| Mar-08 | . | 105,037,660 | 96,236,796 | 113,838,525 | . | . | . |
| Apr-08 | . | 68,714,179 | 59,879,675 | 77,548,682 | . | . | . |
| May-08 | . | 48,974,818 | 39,811,393 | 58,138,243 | . | . | . |
| Jun-08 | . | 40,811,446 | 31,759,003 | 49,863,890 | . | . | . |

에 비하여 2~3도 이상 높은 이상기온을 보이고 있다. 이러한 이상기온에 의하여 상대적으로 큰 예측오차를 보였다 할 수 있다.

4. 결론

본 연구의 목적은 월별 가스 수요량에 대한 예측 모형의 개발에 있다. 가스 수요량에 영향을 미치는 요인으로는 온도 효과와 확산 계절 효과, 유효일수 효과, 수용가수를 이용한 추세효과를 고려하였으며 온도효과를 구축하는데 있어서 온도가 가지는 월별 탄력성을 반영하는 효과를 함께 고려하였다. 이와 더불어 시계열 자료가 가지는 자기상관성을 반영하기 위하여 오차항이 자기회귀모형을 따르는 회귀모형을 모형 구축의 기본 모형으로 이용하였다. 구축된 모형의 설명력은 약 99.4%로서 그 설명력이 매우 높음을 확인할 수 있었고 모형이 가지는 예측능력을 평가하는 절대오차비율도 대부분이 3.2%로 안정적으로 예측을 수행하고 있음을 확인할 수 있었다. 또한 고려하였던 모든 효과들이 통계적으로 매우 유의하였고, 방향성 또한 의미 있는 결과를 보였다. 모형 비교를 위하여 온도만을 이용하여 가스 수요량을 예측하는 모형과의 비교를 수행하였는데, 본 연구에서 제안하고 있는 모형이 보다 높은 예측력을 보였다. 본 연구에서 모형을 구축하는데 또 다른 고려 사항은 상대적으로 구현이 쉬운 모형의 구축에 있었다. 구현된 결과는 MicroSoft사의 Excel과 같은 상용화 프로그램을 이용하여 구현이 가능하도록 하였으며, 이를 위하여 구축된 모형이 복잡한 계량경제학적 모형보다는 보다 간결한 통계학 모형을 이용하고자 하였다.

본 연구에 대한 한계점 및 추후 과제로는 크게 두 가지를 고려할 수 있다. 이전 절의 마지막 예측 부분에

서 언급한 바와 같이 본 연구의 모형이 가스 수요량에 영향을 미치는 다양한 효과들을 적절히 설명하고 있으나 온도효과를 반영함에 있어서 보다 정교한 모형이 고려하여야 하겠다. 특정 시점에서 이상 고온이나 이상 저온 현상이 발생한 경우 상대적으로 예측오차가 증가함을 볼 수 있다. 이는 표 3.5에서 제시된 값 뿐만 아니라 이전 시점에서도 유사한 현상이 발생하고 있음을 확인할 수 있었다. 두 번째로 본 모형을 이용하여 미래 시점의 예측치를 구하기 위해서는 모든 독립변수들에 대한 예측치가 필요하게 된다. 본 모형에서 이용한 독립변수들 가운데 온도효과를 제외한 모든 효과들은 모두 시간의 흐름에 따라 적절한 모형을 통하여 구현될 수 있는 값들이고 독립변수들에 대한 모형 및 수식(확산 계절변동을 위한 삼각함수의 이용)을 통하여 구현되었다. 그러나 온도 효과에 대한 부분은 적절한 모형을 고려하지 않았고 단순히 과거 값들에 대한 평균값으로 온도에 대한 예측값을 대체하였다. 이를 개선하기 위해서는 보다 정교하고 적절한 온도효과에 대한 모형의 구현이 추가적으로 요구된다 할 수 있겠다. 이에 대한 한 가지 대안으로써 공식통계에서 주로 이용되는 계절조정 방법의 대표적인 방법인 X-12 ARIMA (Findley 등, 1998) 방법을 고려할 수 있다. X-12 ARIMA는 시계열 분해기법에 기초한 예측방법으로써 별도의 설명변수를 요구하지 않는다. 김희영과 박유성 (2007), 송명숙 등 (2008)은 시계열자료를 대상으로 하여 X-12 ARIMA 방법을 이용한 예측을 수행하였으며 본 연구에서 이용한 도시가스 수요량 자료에도 적용이 가능할 것이다. 이와 관련된 연구를 추후로 남겨둔다.

참고문헌

- 김희영, 박유성 (2007). 제조업을 중심으로한 한국 산업별 GDP 예측모형연구, *Journal of the Korean Data Analysis Society*, **9**, 1229-2354.
- 나인강, 유지철 (2000). <에너지 수요 분석 및 전망 - 2001년 수요전망과 정책 이슈 ->, 에너지 경제 연구원, 서울.
- 박광수, 김태현, 최도영, 박호정, 정창봉, 김남일, 김영덕, 김선정 (2004). <중단기 에너지수급 전망 연구(II)>, 에너지 경제 연구원, 서울.
- 박유성, 김기환 (2002). <시계열 자료분석 I>, 자유아카데미, 서울.
- 송명숙, 최보승, 김성용, 라울모레티 (2008). 계절조정 방법을 이용한 지적업무량 예측, *Journal of the Korean Data Analysis Society*, **10**, 1579-1590.
- 양준모, 유상희 (2008). 전력수요관리를 감안한 합리적 전력수요예측, *Journal of the Korean Data Analysis Society*, **10**, 2755-2765.
- 유병철 (1996). <전력수요의 가격탄력성과 요금조정방안>, 에너지 경제 연구원, 서울.
- 이충열, 강운영 (2007). 배기량별 승용차 등급 변화를 고려한 승용차 에너지 수요 예측, *Journal of the Korean Data Analysis Society*, **9**, 1179-1195.
- Bass, F. M. (1969). A new product growth model for consumer durables, *Management Science*, **15**, 215-227.
- Engle, R. F. and Granger, C. W. J. (1987). Co-integration and error correlation: Representation, estimation, and testing, *Econometrica*, **55**, 251-276.
- Findley, D. F., Monsell, B. C., Bell, W. R., Otto, M. C. and Chen, B. C. (1998). New capabilities and methods of the X-12-ARIMA seasonal adjustment program (with discussion), *Biometrika*, **61**, 485-492.
- Pesaran, H., Smith, R. and Akiyama, T. (1998). *Energy Demand in Asian Developing Economies*, A World Bank Study.

A Development of Time-Series Model for City Gas Demand Forecasting

Boseung Choi¹ · Hyuncheol Kang² · Kyung-Yun Lee³ · Sang Tae Han⁴

¹Institute of Economics, Korea University

²Department Informational Statistics, Hoseo University

³Department Business Management, Korea University

⁴Department Informational Statistics, Hoseo University

(Received August 2009; accepted September 2009)

Abstract

The city gas demand data has strong seasonality. Thus, the seasonality factor is the majority for the development of forecasting model for city gas supply amounts. Also, real city gas demand amounts can be affected by other factors; weekday effect, holiday effect, the number of validity day, and the number of consumptions. We examined the degree of effective power of these factors for the city gas demand and proposed a time-series model for efficient forecasting of city gas supply. We utilize the liner regression model with autoregressive regression errors and we have excellent forecasting results using real data.

Keywords: City gas, demand forecast, regression with autoregressive errors, validity day effect, elasticity of temperature.

⁴Corresponding author. Professor, Department Informational Statistics, Hoseo University, Asan 336-795, Korea. E-mail: sthan@hoseo.edu