Journal of the Korean Data & Information Science Society
2015, **26**(6), 1409–1416

http://dx.doi.org/10.7465/jkdi.2015.26.6.1409
한국데이터정보과학회지

수원시 기온의 통계적 모형 연구[†]

이훈자[1]

[1]평택대학교 데이터정보학과

접수 2015년 10월 29일, 수정 2015년 11월 16일, 게재확정 2015년 11월 20일

요 약

기온의 변화는 인간의 건강뿐 아니라 동식물의 성장, 경제, 사회, 산업, 문화 등의 전 분야에 영향을 준다. 본 연구에서는 수원시 2003년-2012년 기온을 기상자료, 온실가스자료, 대기자료를 이용하여 자기회귀오차 (autoregressive error)모형으로 월별로 분석하였다. 기온을 위한 기상자료로는, 풍속, 강수량, 일사량, 운량, 습도를 사용했고, 온실가스자료는 이산화탄소 (CO_2), 메탄 (CH_4), 아산화질소 (N_2O), 염화불화탄소 (CFC_{11}), 대기자료는 미세먼지 (PM_{10}), 이산화황 (SO_2), 이산화질소 (NO_2), 오존 (O_3), 일산화탄소 (CO)을 사용하였다. 기온을 월별 분석한 결과 기상변수로는 일사량, 운량, 풍속이 영향을 많이 주는 것으로 분석되었다. 특히 일사량은 봄, 여름, 가을에 영향을 많이 주고 풍속은 겨울에 영향을 많이 주는 것으로 나타났다. 온실가스변수로는 염화불화탄소와 메탄이 기온에 영향을 많이 주고 대기변수로는 오존이 영향을 많이 주는 것으로 타났다. 자기회귀오차모형으로 월별 기온을 43%∼69% 정도 설명할 수 있다.

주요용어: 기상변수, 기온, 대기변수, 설명변수, 온실가스변수, 자기회귀오차모형.

1. 서 론

지구 기온변화는 세계 모든 국가가 공동으로 대처해야 할 중요한 과제중의 하나로 인류의 생활과 동식물의 성장에 큰 영향을 줄 수 있는 문제로 인식되고 있다. 기후변화 대응을 위한 정부 간 협약 (International governmental Panel on Climate Change (IPCC, 2007)) 제4차 보고서에 따르면 지난 100년간 (1906년-2005년 기준) 전 지구의 평균온도는 0.74^oC 상승한 것으로 예측되고 있다. 특히 우리나라 6대 도시의 평균 기온은 1.7^oC 상승 (1912년∼2008년 기준)하여 훨씬 높은 상승률을 나타냈다 (Korea Meteorological Administration, 2009). 이러한 기후 변화는 건강에 영향을 줄 뿐 아니라 경제, 사회, 산업, 문화 등에도 영향을 준다.

기온에 영향을 주는 요인은 크게 기상요인, 대기요인과 함께 온실가스의 증가로 알려져 있다. 기상요인으로는 일사량, 풍속, 운량, 강수량 등이 영향을 많이 주는 것으로 알려져 있고, 대기요인으로는 이산화황, 이산화질소, 일산화탄소, 오존 등이 영향을 주는 것으로 알려져 있다 (Song 등, 2011). 온실가스의 감축을 위하여 국제화사회는 1997년 쿄토 의정서를 채택하여 당시 이산화탄소, 메탄, 아산화질소, 수소불화탄소, 과불화탄소류, 육불화황 등 6가지 온실가스를 규제대상 물질로서 선정하였다 (Kim 등, 2008). 우리나라 기상청에서는 세계온실가스자료센터 (WDCGG; World Data Centre for Greenhouse Gases, http://gaw.kishou.go.jp) 사이트에 안면도 온실가스 자료인 이산화탄소, 메탄, 이산화질소 등을 1999년 1월1일부터 공개하고 있다.

[†] 이 논문은 2014학년도 평택대학교 학술연구비의 지원에 의하여 연구되었음.
[1] (450-701) 경기도 평택시 용이동 111, 평택대학교 데이터정보학과, 교수. E-mail: esther@ptu.ac.kr

지금까지 기온 변화에 관해 다양한 각도에서 연구가 진행되어 왔다. 연구의 대부분이 서울, 부산, 울산 등 지역별로 기온변화 상태와 추세를 연구하고 있다. 그러나 기온에 영향을 주는 자료들인 기상자료, 온실가스자료, 대기자료를 이용한 기온 모형에 관한 연구는 충청남도 서산시 기온의 통계적 모형 연구 (Lee, 2014)외에는 찾아보기 힘들다. 기온변화에 관한 국내연구의 경향을 살펴보면, 우리나라의 겨울철 기온변화 및 한파 발생빈도 분석 (Jeong 등, 2015), 대한민국 40년간의 기온변이추이와 표준기상데이터 비교분석 (Yoo 등, 2010), 일일 최고기온의 변화에 대한 추정 (Ko, 2007), 일반화선현모형 (GLM) 날씨 발생기를 이용한 서울 일일기온 모형 (Kim 등, 2013), 한국의 기온변화와 기온변동성에 대한 통계적 연구 (Kim 등, 2011), 기후변화와 대기환경의 통합적 관리에 대한 고찰 (Song 등, 2011), 기후관측소 분석에 관한 연구 (Kim 등, 2013), 울산지역의 기상특성: 기온과 바람을 중심으로 (Oh 등, 2015), 텍스트 마이닝을 활용한 기후변화관련 식품 (Bae 등, 2013) 등 다양한 방법으로 기온변화에 관해 연구하였다.

본 연구에서는, 수원시 기온자료 2003년 1월 1일부터 2012년의 12월 31일까지의 10년간 일 평균자료를 시계열 모형에 적합하고자 한다. 정확한 기온의 분석과 예측을 위해서는 기온과 연관된 기상자료, 온실가스자료 및 대기자료를 포함시켜야 한다. 본 논문에서는 기온과 연관이 있는 5종류의 기상자료와 4종류 온실가스자료와 5종류 대기자료를 설명변수로 포함한 다변량 분석인 ARE (autoregressive error; 자기회귀오차) 모형으로 분석하였다.

본 연구의 목적은 수원지역의 기온변화에 관한 모형을 온실가스자료, 대기자료, 기상자료 들을 이용하여 월별로 분석하여 기온변화 대응전략을 수립할 수 있는 통계학적 기초자료로써 활용하고자 한다.

2. 수원 기온자료 및 설명변수 자료

기온자료는 기상청의 수원시 서둔동의 일 평균온도 자료를 사용하였다. 기간은 2003년부터 2012까지의 1월 1일부터 12월 31일까지 10년간 자료를 사용했다. 본 연구에서 1월에서 12월까지 자료를 사용한 이유는 월별로 기온 차이가 있기에 월별 분석이 중요할 것으로 생각되었기 때문이다.

설명변수로 사용한 기상자료, 온실가스자료, 대기자료는 기온과 같이 2003년~2012년 자료를 사용하였다. 기상자료도, 기온자료와 같은 곳인 기상청의 수원시 서둔동자료를 사용하였고, 5종류의 기상자료는 Song 등 (2011)을 참고로 하여 풍속 (wind speed), 상대습도 (relative humidity), 강수량 (rainfall), 일사량 (radiation), 운량 (amount of cloud)을 사용했으며 Table 2.1에 나타나 있다. 4종류의 온실가스자료는 충청남도 태안군 안면도에서 측정된 자료인 이산화탄소 (CO_2), 메탄 (CH_4), 아산화질소 (N_2O), 염화불화탄소 (CFC_{11})를 사용하였다. 참고로, 온실가스자료도 수원지역 자료를 사용하려 했지만 자료가 없고, 우리나라는 1999년 1월부터 충남 태안군 안면도의 해안 국립공원에서만 온실가스를 측정하여 GAW 세계온실가스센터로 자료를 송부하여 활용하고 있다. 5종류의 대기자료는 미세먼지 (PM_{10}), 이산화황 (SO_2), 이산화질소 (NO_2), 오존 (O_3), 일산화탄소 (CO)를 사용했고, 국립환경과학원의 수원시 영통동자료를 사용하였다. 사용된 온실가스자료, 대기자료도 Table 2.1에 나타나 있다. 각 설명변수의 하루 사용시점 등의 구체적인 설명은 Table 2.1에 나타나 있다. Table 2.1의 표기 (Notation)는 4절의 모형식 변수표시이다.

Table 2.1 List of explanatory variables

Variable	Contents	Unit	Notation
Temperature	Previous day 1 hour average surface temperature	^oC	Y_{t-1}
Wind	Daily maximum wind speed	m/s	W_t
Rainfall	Daily sum of precipitation	mm	R_t
Radiation	Daily sum of radiation	MJ/M^2	D_t
Cloud	Daily average cloud amount	-	C_t
Humidity	Daily average relative humidity	%	H_t
CO_2	Daily average CO_2	ppm	E_t
CH_4	Daily average CH_4	ppb	B_t
N_2O	Daily average N_2O	ppb	J_t
CFC_{11}	Daily average CFC_{11}	ppb	F_t
PM_{10}	Daily maximum PM_{10}	$\mu g/m^3$	P_t
SO_2	Daily maximum SO_2	ppb	S_t
NO_2	Daily maximum NO_2	ppb	N_t
O_3	Daily maximum O_3	ppb	O_t
CO	Daily maximum CO	0.1 ppm	T_t

3. 기온의 ARE 모형과 적합성

반응변수에 영향을 주는 설명변수가 있을 때는 설명변수를 분석에 이용하는 것이 효율적이다. ARE 모형은 설명변수를 분석에 사용할 수 있는 모형으로 기온과 같은 시계열자료를 회귀모형에 적합 시킬 때 적합한 모형이다. 반응변수인 기온에 영향을 주는 일사량, 풍속, 수소불화탄소 등과 같은 설명변수를 사용하여 자료를 분석한다. ARE 모형은 전이함수 모형의 특수한 경우로, 오차항 ϵ_t가 독립이 아니라 서로 상관관계를 갖게 되며, 특히 AR (autoregressive) 형태를 갖는다. 일반적인 k차 ARE 모형은 다음과 같다.

$$\begin{gathered} Y_t = \beta_0 + \beta_1 X_{t1} + \beta_2 X_{t2} + \cdots + \beta_p X_{tp} + \epsilon_t,\ t = 1, 2, \cdots n, \\ \epsilon_t = e_t - \phi_1 \epsilon_{t-1} - \phi_2 \epsilon_{t-2} - \cdots - \phi_k \epsilon_{t-k} \text{이고 } e_t \sim i.i.d\ N(0, \sigma^2) \end{gathered} \tag{3.1}$$

식 (3.1)에서 오차항 ϵ_t는 서로 상관관계가 있고 AR(k) 형태를 갖는다. 또한 식 (3.1)에서 Y_t는 반응변수인 기온를 나타내며, $X_{t,j}$,$j = 1, 2, \cdots, p$는 Y_t에 영향을 주는 풍속이나 이산화탄소농도 등의 설명변수들을 나타낸다. ARE 모형에서는 반응변수의 과거시점 (Y_{t-1}, Y_{t-2} 등)이 설명변수로 종종 사용되기도 한다. 위의 ARE 모형과 회귀모형의 차이점은 ARE 모형은 식 (3.1)에서 오차항 ϵ_t가 서로 독립이 아니라 시차에 따른 자기 상관을 갖고 있으며 특히 AR (autoregressive)(k) 모형의 형태를 갖는다는 점이다. 따라서 식 (3.1)을 ARE (자기회귀오차)모형이라 부르며 시계열자료에 회귀모형을 적합 시킬 경우 많이 사용된다.

ARE 모형들의 적합성을 측정하는데 있어 결정계수 R^2와 잔차 분석의 RMSE (root mean square error)를 사용하였다. 결정계수 R^2은 기온 Y_t가 설명변수에 의해 어느 정도 설명되는지를 나타내는 지표로 값이 0이면 0% 설명되고 1 이면 100% 설명됨을 의미한다. 따라서 값이 1에 가까울수록 설명력이 높은 좋은 모형이다. 예측값과 실제값의 분산정도를 나타내는 RMSE의 공식은 다음과 같다.

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(o_i - p_i)^2}, \tag{3.2}$$

o_i와 p_i는 각각 i일의 관측기온과 추정기온을 의미하며, n은 통계량에 사용된 일수 (day)이다. RMSE는 값이 작을수록 좋은 모형이다.

4. 수원기온의 ARE 모형

수원지역의 기온 모형을 측정하기 위해 2003∼2012년의 10년간의 일별 평균기온을 사용했다. 월별로 10년간의 그래프를 그려 복잡해지는 것을 피하기 위해 2003년과 2012년에 대해서만 월 평균온도를 그래프 (Figure 4.1)로 살펴보았다. 봄, 가을인 3월∼5월, 9월∼10월에는 2003년과 2012년의 기온이 거의 비슷하게 나타났고, 겨울인 1월에는 2012년이 조금 높게 나타난 반면 2월, 11월, 12월에는 2012년이 조금 낮게 나타났다. 여름인 6월, 7월, 8월에는 2012년이 높게 나타났다. 여름의 변화를 좀 더 자세히 보기 위해 6월, 7월, 8월의 일 평균온도를 연도별 그래프 (Figure 4.2)로 살펴보았다. 전반적으로 일평균기온의 월평균이 8월, 7월, 6월 순으로 높게 나타났다. 년도에 따라 온도의 변화가 있지만 2012년의 온도는 2003년에 비해 크게 오른 것으로 나타났다. 6월을 살펴보면 2003년이 21.33^{o}C, 2012년에는 23.57^{o}C로 10년간 2.24^{o}C 상승했고 7월의 경우 2003년에는 23.81^{o}C, 2012년에는 25.54^{o}C로 나타나 1.73 ^{o}C 상승했다. 8월의 경우 2003년에는 24.18^{o}C, 2012년에는 27.27^{o}C로 10년간 무려 3.09^{o}C 상승한 것으로 나타났다. 이러한 추세는, 서론에서 언급했던, 지난 100년간 지구의 평균 상승값인 0.74^{o}C (IPCC, 2007)와 비교해 매우 높게 나타났고 또한 1912년∼2008년 사이 우리나라 6대 도시의 기온이

1.7도 상승 (Korea Meteorological Administration, 2009)한 것과 비교해도 수원의 6월, 8월의 온도 변화는 높게 상승한 것으로 나타났다.

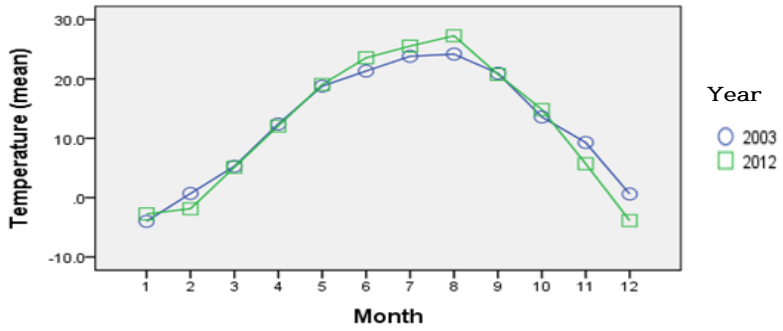

Figure 4.1 The line plots of monthly temperature in Suwon City

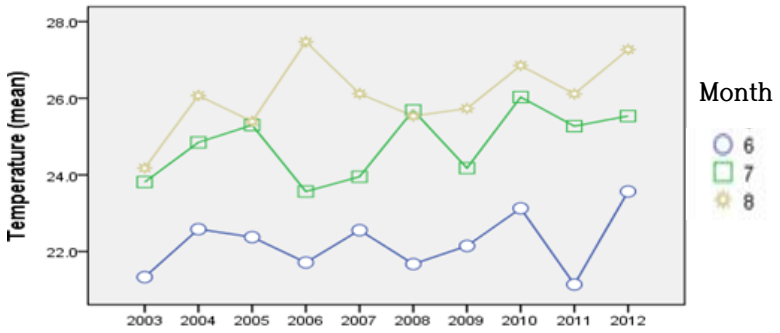

Figure 4.2 The line plots (June, July, August) of yearly temperature in Suwon

4.1. 월별기온에 관한 ARE 모형

월별 수원지역 기온모형을 분석하여 Table 4.1에 나타내었다. 월별 분석은 2003년-2012년의 10년간 일 평균자료를 모두 모형구축기간으로 사용했다. ARE 모형의 변수 선택은 14종류의 설명변수들 간의 다중공선성 (multi-collinearity)를 고려하면서 설명변수들 간의 다양한 모형을 분석한 후 RMSE 값이 작은 모형을 선택하였다. ARE 모형의 설명변수 차수 선택은 온실가스 변수인 CO_2, CH_4, N_2O, CFC_{11}와 대기변수인 SO_2, NO_2, CO, O_3는 당일자료 (t)를 사용하였고, 기상변수인 풍속, 습도, 강수량, 일사량, 운량은 기온과 더욱 밀접한 관계에 있을 것으로 사료되어 당일 자료 (t)와 전일자료 (t-1)를 교대로 사용한 후 적절한 값을 사용하였다.

ARE 모형의 모수 추정에는 YW (Yule-Walker), ULS (unconditional least square), MLE (maximum likelihood estimate) 등의 방법이 있지만, 본 연구에서는 기온모형과 같이 과거시점의 종속변수 (lagged dependent variable)인 Y_{t-1}이 설명변수로 사용되는 경우에 적합한 MLE 방법을 사용하였다. ARE모형의 적합성 검정으로는 기온 모형과 같이 과거시점의 종속변수가 설명변수로 사용되는 경우에 적합한 Durbin-h 검정을 사용해 검정하였고 또한 잔차의 ACF (auto-correlation function)와 PACF (partial auto-correlation Function)의 결과를 검정하였다 (Cho와 Lee, 1997).

월별 ARE 모형 (Table 4.1)을 보면, 기온을 설명하는 기상변수, 온실가스변수, 대기변수가 1월에서 12월까지 월 마다 조금씩 다르게 나타났다. 모든 달에 전일 기온 (Y_{t-1})이 포함되었다, 먼저 계절별로

살펴보면, 겨울인 1월, 2월, 11월, 12월에는 기상변수로 운량 (C_t)와 풍속 (W_t 혹은 W_{t-1})이 공통적으로 포함되어 있고 대기변수로는 오존 (O_t)과 온실가스변수인 수소불화탄소 (F_t)가 포함돼 기온에 영향을 주는 것으로 나타났다. 봄인 3월, 4월, 5월에는 일사량 (D_t), 운량 (C_t), 오존 (O_t)이 포함되어 기온에 영향을 주는 것으로 나타났다. 여름인 6월, 7월, 8월에는 일사량과 메탄 (B_t)이 공통적으로 포함되어 있고, 습도 (H_t), 풍속 (W_t 혹은 W_{t-1}), 이산화탄소 (E_t), 메탄 (B_t), 아산화질소 (J_t), 가 7월, 8월에 포함되어 있어 영향을 주는 것으로 나타나났다. 여름기온의 특이한 사항은 비록 대기변수인 일산화탄소 (6월), 미세먼지 (7월)가 각각 모형에 포함되어 있지만, 기상변수인 일사량, 습도, 운량, 풍속과 온실가스변수인 이산화탄소, 메탄, 아산화질소로 기온 설명이 많이 된다는 사실이다. 가을인 9월, 10월에는 습도, 운량, 일사량, 수소불화탄소가 모형에 공통적으로 들어가 있어 기온에 영향을 주는 것으로 나타났다.

전체적인 기상변수를 살펴보면, 특히 일사량과 운량이 포함된 달이 많아 기온에 영향을 주는 것으로 나타났다. 3월∼11월에는 일사량 (D_t)이 모두 포함되어 있어 겨울을 제외한, 봄, 여름, 가을 기온에 영향을 주는 것으로 나타났다. 운량 (C_t)은 1월∼6월, 9월∼11월에 포함되어 있어, 7월, 8월을 제외한 모든 계절에 영향을 주는 것으로 나타났다. 풍속(W_t 혹은 W_{t-1})는 1월∼3월, 5월, 7월∼8월, 10월∼12월에는 포함되어 있다. 일사량과 운량의 계수는 양수로 풍속의 계수는 음수로 나타나, 잘 알려진 바와 같이, 일사량이 많을수록, 운량이 많을수록, 풍속이 적을수록 기온은 올라가는 것으로 나타났다. 습도 (H_t)는 5월, 7월∼8월, 9월∼12월에 영향을 주는 것으로 나타났고, 5월, 7월∼8월에는 습도의 계수가 음수로 나머지 달에는 양수로 나타나 5월, 7월∼8월에는 습도가 높을수록 기온은 내려가고, 나머지 달에는, 습도가 높을수록 기온이 올라가는 것으로 나타났다. 온실가스 변수는 염화불화탄소 (F_t)가 1∼5월, 9월∼12월에 나타나 영향을 많이 주는 것으로 나타났고, 이산화탄소 (E_t)는 7월∼8월에 포함되어 있고, 메탄 (B_t)는 3월, 6월∼8월, 10월에 포함되어 있고 아산화질소 (J_t)는 5월, 7월∼8월 11월에 영향을 주는 것으로 나타났다. 대기변수 중에는 오존 (O_t)이 1월∼5월, 10∼12월에 나타나 영향을 많이 주는 것으로 분석되었다. 이산화질소 (N_t)가 2월-4월과 11월∼12월에 미세먼지 (P_t)는 7월, 9월에 영향을 주는 것으로 나타났다.

Table 4.1 The monthly ARE models for Suwon temperature

Month	ARE Model	R^2	RMSE
January	$Y_t = -29.41 + 0.29Y_{t-1} + 0.47C_t - 0.40W_{t-1} + 0.10F_t + 73.64S_t + 62.37O_t + \epsilon_t$, $\epsilon_t = 0.55\epsilon_{t-1} + e_t$	0.50	1.95
February	$Y_t = -32.89 + 0.27Y_{t-1} + 0.43C_t - 0.14W_{t-1} + 0.11F_t + 81.72O_t + 54.76N_t + \epsilon_t$, $\epsilon_t = 0.51\epsilon_{t-1} + e_t$	0.53	2.05
March	$Y_t = 0.39Y_{t-1} + 0.41C_t + 0.10D_t - 0.27W_{t-1} - 0.05R_{t-1} - 0.02B_t + 0.10F_t + 81.35O_t + 65.68N_t + \epsilon_t$, $\epsilon_t = 0.38\epsilon_{t-1} - 0.11\epsilon_{t-14} + e_t$	0.56	1.93
April	$Y_t = -12.26 + 0.40Y_{t-1} + 0.47C_t + 0.22D_t - 0.04R_{t-1} + 0.04F_t + 56.26O_t + 26.02N_t + \epsilon_t$, $\epsilon_t = 0.41\epsilon_{t-1} + 0.13\epsilon_{t-9} - 0.14\epsilon_{t-16} + e_t$	0.48	1.89
May	$Y_t = 0.38Y_{t-1} - 0.06H_t + 0.16C_t + 0.05D_t - 0.22W_t + 0.02J_t + 0.05F_t + 35.30O_t + \epsilon_t$, $\epsilon_t = 0.49\epsilon_{t-1} + 0.10\epsilon_{t-9} - 0.08\epsilon_{t-18} + e_t$	0.43	1.50
June	$Y_t = 0.59Y_{t-1} + 0.25C_t + 0.18D_t + 0.004B_t + 1.47T_t + \epsilon_t$, $\epsilon_t = 0.15\epsilon_{t-4} - 0.14\epsilon_{t-15} - 0.16\epsilon_{t-17} + 0.13\epsilon_{t-19} + e_t$	0.58	1.35
July	$Y_t = 0.25Y_{t-1} - 0.10H_t + 0.06D_t - 0.11W_t + 0.06E_t - 0.003B_t + 0.05J_t + 0.005P_t + \epsilon_t$, $\epsilon_t = 0.57\epsilon_{t-1} - 0.08\epsilon_{t-11} + e_t$	0.53	1.03
August	$Y_t = -19.44 + 0.37Y_{t-1} - 0.02H_t + 0.13D_t - 0.13W_{t-1} - 0.01R_t + 0.08E_t - 0.005B_t + 0.05J_t + \epsilon_t$, $\epsilon_t = 0.43\epsilon_{t-1} + 0.15\epsilon_{t-5} + e_t$	0.63	1.03
September	$Y_t = -10.91 + 0.60Y_{t-1} + 0.06H_t + 0.42C_t + 0.37D_t + 0.03F_t + 0.01P_t + \epsilon_t$, $\epsilon_t = -0.13\epsilon_{t-9} - 0.16\epsilon_{t-18} + e_t$	0.69	1.41
October	$Y_t = -18.81 + 0.50Y_{t-1} + 0.05H_t + 0.42C_t + 0.23D_t - 0.21W_t - 0.05R_t + 0.01B_t + 0.03F_t - 107.84S_t + 67.61O_t + \epsilon_t$, $\epsilon_t = 0.26\epsilon_{t-1} + 0.13\epsilon_{t-6} + e_t$	0.68	1.50
November	$Y_t = 0.45Y_{t-1} + 0.05H_t + 0.58C_t + 0.22D_t - 0.39W_{t-1} - 0.07R_t - 0.04J_t + 0.06F_t + 100.37O_t + 30.09N_t + \epsilon_t$, $\epsilon_t = 0.34\epsilon_{t-1} + e_t$	0.68	1.92
December	$Y_t = -26.32 + 0.32Y_{t-1} + 0.06H_t + 0.53C_t - 0.27W_{t-1} + 0.07F_t + 91.02O_t + 51.94N_t + \epsilon_t$, $\epsilon_t = 0.53\epsilon_{t-1} + 0.19\epsilon_{t-6} + 0.06\epsilon_{t-15} + e_t$	0.60	2.10

적합도 척도인 결정계수 R^2는 0.43~0.69로 나타났다. 4월, 5월이 각각 0.48, 0.43으로 비교적 낮게 나타났고 9월, 10월, 11월이 각각 0.69, 0.68, 0.68로 비교적 높게 나타났다. 9월의 예를 들면 일 최고 온도 Y_t는 Y_{t-1}, 습도 (H_t), 운량 (C_t), 일사량 (D_t), 염화불화탄소 (F_t), 미세먼지 (P_t) 에 의해 69% 설명되는 것으로 나타났다. 적합도 척도로 실제값과 예측값의 차이를 나타내는 RMSE 값은 1.03에서 2.10으로 나타났다.

본 연구에서 사용한 기상변수, 온실가스변수, 대기변수는 각각 기상청, 세계온실가스자료센터, 국립환경과학원에서 구할 수 있는 자료이므로, Table 4.1을 이용하면, 월별 기온를 예측하는데 도움이 될 것으로 생각된다.

4.2. ARE 모형의 실제값과 추정치의 산점도

수원기온 실제값과 ARE 모형 추정치의 산점도를 결정계수가 가장 낮은 5월과 가장 높은 9월을 Figure 4.3에 나타내었다. 5월 모형의 산점도를 보면, 기온에 따라 과대추정 혹은 과소추정이 된 경향이 있다. 특히 기온이 16^{o}C 이하인 경우에 약간의 과대 추정된 경향이 있고, 16^{o}C와 20^{o}C 사이에는 과소 추정이 되었다. 반면에 9월 모형은 기온이 20^{o}C와 25^{o}C사이에 약간의 과소 추정된 값이 있으나, 그 외의 경우에는 추정이 비교적 괜찮은 것으로 나타났다. ARE 모형은 기온에 영향을 주는 설명변수를 수식으로 표현하는 장점이 있으므로, 모형에 기온에 영향을 주는 기상변수, 온실가스변수와 대기변수를 좀 더 포함시켜 개발하면, 실제값에 좀 더 가까이 접근하는 좋은 결과를 가질 수 있을 것으로 기대한다.

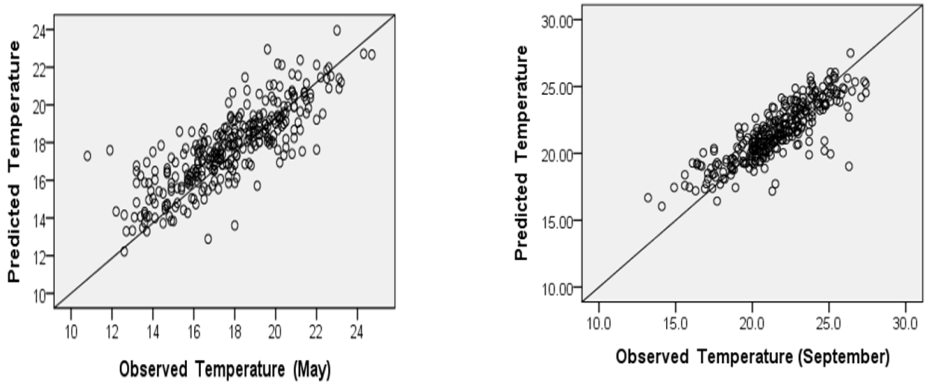

Figure 4.3 The scatter plots (May and September) between actual value and ARE estimated value of temperature

5. 결론

수원 기온을 월별로 시계열 모형인 ARE 방법으로 분석하였다. ARE 모형은 설명변수를 분석에 이용할 수 있는 모형으로 시계열자료를 회귀모형 형태로 적합 시킬 때 적절한 모형이다. 본 연구에서는, 기온에 영향을 주는 5종류의 기상자료, 4종류의 온실가스자료, 5종류의 대기자료를 설명변수로 사용하였다. 5종류의 기상자료는 수원시 기상청 자료로 풍속, 강수량, 일사량, 운량, 습도이며 4종류의 온실가스자료는 충청남도 태안군 안면도에서 측정된 자료인 CO_2, CH_4, N_2O, CFC_{11}를 사용하였고 5종류의 대기자료는 국립환경과학원의 수원시 PM_{10}, SO_2, NO_2, CO, O_3이다.

연구기간인 2003년부터 2012년의 1월 1일에서 12월 31일까지의 10년 일 평균자료를 사용하여 월별 ARE 모형으로 분석하였고 비교적 잘 설명되는 것으로 나타났다. 월별 기온을 분석한 결과 기상변수로는 일사량, 운량, 풍속이 영향을 많이 주는 것으로 나타났다. 3월-11월에는 일사량이 모두 포함되어 있어 겨울을 제외한 봄, 여름, 가을에 영향을 주는 것으로 나타났다. 운량은 7월, 8월을 제외한 1월-6월, 9월-12월에 포함되어 있다. 풍속은 1월∼3월, 5월, 7월∼8월, 10월∼12월에 포함되어 있어, 특히 겨울기온에 영향을 주는 것으로 나타났다. 여름기온의 특이한 사항은 비록 대기변수인 일산화탄소 (6월), 미세먼지 (7월)가 각각 모형에 포함되어 있지만, 기상변수인 일사량, 습도, 운량, 풍속과 온실가스변수인 이산화탄소, 메탄, 아산화질소로 기온 설명이 많이 된다는 사실이다. 온실가스변수는 CFC_{11}와 CH_4가 영향을 많이 주며 대기변수 SO_2, NO_2, CO, O_3는 월별로 각각 다르게 영향을 주는 것으로 나타났고 오존 O_3가 영향을 많이 주는 것으로 나타났다. 결정계수 R^2는 0.43∼0.69로 나타났고 특히 9월이 0.69, 10월과 11월은 모두 0.68로 비교적 높게 나타났다. RMSE 값은 1.03에서 2.10으로 비교적 낮게 나타났다.

기온 변화는 인간의 건강과 생활 뿐 아니라 한 국가의 경제, 사회, 산업, 문화 등의 많은 분야에 영향을 주는 중요한 과제로 앞으로 계속 관심 가져야 할 분야이다. 본 연구에서 사용한 기상변수, 온실가스변수, 대기변수는 각각 기상청, 세계온실가스자료센터, 국립환경과학원에서 구할 수 있는 자료이므로, Table 4.1을 이용하면, 수원시 월별 기온를 예측하는데 연구의 활용도가 있을 것으로 기대한다. 또한 ARE 모형은 Table 4.1과 같이 기온에 어떤 설명변수들이 영향을 주는지 등의 수식으로 표현할 수 있는 장점이 있음으로, 본 연구에서 사용하지 못한 설명변수 들을 좀 더 첨가하여 분석하면 좀 더 좋은 모형이 되리라 기대한다. 앞으로, 기온에 사용되는 설명변수에 관한 다양한 연구와 함께 다양한 모형에 관한 연구가 우리나라 많은 지역에서 활발히 이루어져야 된다고 생각된다.

References

Bae, K., Park, J., Kim, J. and Lee, Y. (2013). Analysis of the abstracts of research articles in food related to climate change using a text-mining algorithm. *Journal of the Korean Data & Information Science Society*, **24**, 1429-1437.

Cho, S. and Lee, J. (1997). *Analysis of economic time series analysis using SAS/ETS,* Freedom Academy, Seoul.

IPCC (2007). *Contribution of working group II to the fourth assessment report of the intergovernmental panel on climate change*, Cambridge Press, Cambridge, U.K. 976.

Jeong, M. J. and Cho, Y. (2015). An analysis of a winter-time temperature change and an extream cold waves frequency in Korea. *Journal of Climate Change Research*, **6**, 87-84.

Kim, H. C., Choi, S. K. and Yun, B. R. (2011). A statistical analysis on temperature change and climate variability in Korea. *Communications of the Korean Statistical Society*, **18**, 1-12.

Kim, H., Do, H. Y. and Kim, Y. (2013). A modeling of daily temperature in Seoul using GLM weather generator. *The Korean Journal of Applied Statistics*, **26**, 413-420.

Kim, H. K. and Lee, Y. (2013). A study on the density analysis of climatological stations using the correlation integral method in the fractal dimension. *Journal of the Korean Data & Information Science Society*, **24**, 53-62.

Ko, W. K. (2007). Estimation for change of daily maxima temperature. T*he Korean Journal of Applied Statistics*, **20**, 1-9.

Korea Meteorological Administration (2009). *Report of global atmosphere watch 2008*, Korea Meteorological Administration, Korea, 178.

Lee, H. (2014). Analysis of statistical models on temperature at the Seosan city in Korea. *Journal of the Korean Data & Information Science Society*, **25**, 1293-1300.

Oh, I., Bang, J. and Kim, Y. (2015). Meteorological characteristics in the Ulsan metropolitan region: Focus on air temperature and winds. *Journal of Korean Society for Atmospheric Environment*, **31**, 181-194.

Song, C. K., Lee, S. and Yoon, J. S. (2011). A review of the integrated strategy for climate change and air pollution management. *Journal of Korean Society for Atmospheric Environment*, **27**, 805-818.

Yoo, H. C. and Kang, H. G. (2010). Comparative analysis of temperature change trend and standard meteorological data in Korea over 40 years. *Journal of KIAEBS*, **4**, 97-103.

Journal of the Korean Data & Information Science Society
2015, **26**(6), 1409–1416

http://dx.doi.org/10.7465/jkdi.2015.26.6.1409
한국데이터정보과학회지

Analysis of statistical models on temperature at the Suwon city in Korea[†]

Hoonja Lee[1]

[1]Department of Data Information, Pyeongtaek University

Received 29 October 2015, revised 16 November 2015, accepted 20 November 2015

Abstract

The change of temperature influences on the various aspect, especially human health, plant and animal's growth, economics, industry, and culture of the country. In this article, the autoregressive error (ARE) model has been considered for analyzing the monthly temperature data at the Suwon monitoring site in Korea. In the ARE model, five meteorological variables, four greenhouse gas variables and five pollution variables are used as the explanatory variables for the temperature data set. The five meteorological variables are wind speed, rainfall, radiation, amount of cloud, and relative humidity. The four greenhouse gas variables are carbon dioxide (CO_2), methane (CH_4), nitrous oxide (N_2O), and chlorofluorocarbon (CFC_{11}). And the five air pollution explanatory variables are particulate matter (PM_{10}), sulfur dioxide (SO_2), nitrogen dioxide (NO_2), ozone (O_3), and carbon monoxide (CO). Among five meteorological variables, radiation, amount of cloud, and wind speed are more influence on the temperature. The radiation influences during spring, summer and fall, whereas wind speed influences for the winter time. Also, among four greenhouse gas variables and five pollution variables, chlorofluorocarbon, methane, and ozone are more influence on the temperature. The monthly ARE model explained about 43-69% for describing the temperature.

Keywords: Autoregressive error model, explanatory variable, greenhouse gas variables, meteorological variables, pollution variable, temperature.

[†] This paper was supported by the Research Fund, 2014, Pyeongtaek University in Korea.
[1] (450-701) Professor, Department of Data Information, Pyeongtaek University, Pyeongtaek 450-701, Korea. E-mail: esther@ptu.ac.kr