

로지스틱 회귀분석과 판별분석을 활용한 광주광역시 폭염에 미치는 영향분석

김운수¹ · 공영선² · 장인홍^{3*}

조선대학교 컴퓨터통계학과^{1,3}
조선대학교 전산통계학과²

Analysis of the Impact of Heatwaves in Gwangju using Logistic Regression and Discriminant Analysis

Youn Su Kim, Yeong Seon Kong and In Hong Chang[†]

Department of Computer Science and Statistic, Chosun University, Gwangju, Korea

Abstract

Abnormal climate is a phenomenon in which meteorological factors such as temperature and precipitation are significantly higher or lower than normal, and is defined by the World Meteorological Organization as a 30-year period. However, over the past 30 years, abnormal climate phenomena have occurred more frequently around the world than in the past. In Korea, abnormal climate phenomena such as abnormally high temperatures on the Korean Peninsula, drought, heatwave and heavy rain in summer are occurring in March 2023. Among them, heatwaves are expected to increase in frequency compared to other abnormal climates. This suggests that heatwave should be recognised as a disaster rather than just another extreme weather event. According to several previous studies, greenhouse gases and meteorological factors are expected to affect heatwaves, so this paper uses logistic regression and discriminant analysis on meteorological element data and greenhouse gas data in Gwangju from 2008 to 2022. We analyzed the impact of heatwaves. As a result of the analysis, greenhouse gases were selected as effective variables for heatwaves compared to the past, and among them, chlorofluorocarbons were judged to have a stronger effect on heatwaves than other greenhouse gases. Since greenhouse gases have a significant impact on heatwaves, in order to overcome heatwaves and abnormal climates, greenhouse gases must be minimized to overcome heatwaves and abnormal climates.

Keywords: Heatwave, Logistic regression, Discriminant analysis, Abnormal climate, Disaster

(Received May 24, 2024; Revised June 10, 2024; Accepted June 13, 2024)

[†] Corresponding author: ihchang@chosun.ac.kr

1. 서론

IPCC(Intergovernmental Panel on Climate Change) 6차 보고서에 의하면, 전 세계의 평균 기온은 산업화 이전 세대(1900년)와 비교하면 1°C 이상 높아졌으며 2100년에는 2°C 이상 기온이 상승할 것으로 말하고 있다^[1]. 평균 기온 상승으로 인해 전 세계적으로 이상 기후 현상인 폭염의 문제에 시달리고 있다. 먼저 인명 피해를 알아보면, 2023년 여름 유럽 폭염 사망자는 총 6만 1000명 이상 사망한 것으로 조사되었고, 우리나라 또한 32명이 온열 질환으로 사망하였다. 질병 관리청에서는 역대 우리나라 최악의 폭염을 기록한 2018년은 전국에서 4500명 이상의 온열질환자가 발생하였고, 이 가운데 48명이 사망하였으며, 한국 환경연구원이 개발한 RCP 8.5 모형에 따르면 2100년에는 387명의 사망자가 발생할 것으로 예측된다. 폭염은 인명 피해뿐만 아니라 경제적 피해도 끼치는데, 사이언스 어드밴시스(Science Advances)는 식량 생산 감소 및 폭염에 따른 경제적 피해액이 무려 16조에 이른다고 발표했다^{[2]-[3]}. 폭염이 인명 피해 및 경제적 피해를 주는 것을 알 수 있으므로 폭염을 예측해 폭염 피해를 최소화하는 것이 중요하다. 기후정보센터에 의하면, 과거 48년 평균에 비해 최근 10년 동안 폭염이 3일 이상 증가하는 것으로 나타나, 폭염에 대한 대처가 절실하다. 또한, 단순히 폭염을 기온이 올라가는 현상이나 이상기후만의 문제로만 바라봐서는 안되며, 이는 우리가 식량이나 생활 등 삶과 직결되어 있기 때문에 재난적 차원으로 바라봐야 할 것이다.

이에 따라 폭염과 관련하여 많은 연구가 진행되고 있다. 장은세 등은 2007년부터 2016년까지 해발고도, 풍향, 습도, 풍속, 일 최고기온, 최저기온 기반으로 군집분석 중 하나인 와드 분석법을 이용해 폭염을 예측하였다^[4]. 김령은 등은 1973년부터 2018년까지 일 최저기온을 이용해 습도와 온도에 기초한 지수인 겔보기 온도를 사용하여 폭염의 한 가지 형태인 열대야를 예측하였다^[5]. 김성태 등은 영-과잉 회귀모형을 활용하여 지난 2009년부터 2018년까지 폭염주의보 및 폭염경보 발생일수를 적합하여 범용 크리깅을 이용하여 폭염 발생일수를 예

측하였다^[6]. 안중배 등은 PNU CGCM(Pusan National University Coupled general circulation model)을 이용해 대한민국 폭염의 장기 계절 예측하였다^[7]. 또한 허솔잎 등은 1991년부터 2010년 동안 기상청 기후예측 시스템(GloSea5) 자료를 이용해 극한예측지수를 생성하여 이상 기후를 예측하였다^[8]. 김인혜 등은 광주의 지리적 특성을 파악해 폭염에 대해 시계열적 특성을 이용해 분석한 결과 2000년부터 2020년간 폭염이 많이 관측된 8월에서 매년 4.4%만큼 증가하는 추세를 보여줬다^[9]. 1997년 교토의정서에서 평균 기온 상승에 있는 6대 온실가스를 이산화탄소(CO₂), 메탄(CH₄), 아산화질소(N₂O), 수소불화탄소(HFCs), 과불화탄소(PFCs), 육불화황(SF₆)로 규정하였다. 이처럼 폭염에 관한 많은 연구를 진행하였다. 이를 기반으로 본 연구에서는 과거 연구를 기반으로 폭염과 기온 상승에 직간접적으로 영향을 미칠 것으로 판단되는 여러 기후 정보와 온실가스의 정보를 통해 폭염에 영향을 미치는 주요 변수 변화를 로지스틱 회귀분석과 판별 분석을 통해 비교하고자 한다. 현재 기후변화가 급속도로 진행되고 있기 때문에 과거에 폭염에 미쳤던 영향 변수와 현재의 영향을 미치는 변수의 변화를 비교하기 위해 2003년부터 2022년까지의 기후 데이터와 온실가스 데이터를 함께 결합하여 5년 단위로 구분 지은 후 분석하고자 한다.

본 논문의 구성은 다음과 같다. 2장에서는 로지스틱 회귀모형, 선형판별분석에 대한 이론적 배경을 설명하고 3장에서는 본 논문에서 분석에 사용되는 데이터 소개를 하고, 4장에서는 분석을 통해 폭염에 미치는 기상 요소를 판단한 후, 마지막 5장에서는 결론을 제시하는 것으로 마무리한다.

2. 분석 방법

2.1. 로지스틱 회귀분석

로지스틱 회귀는 독립변수의 결합을 통해 사건을 예측하는 기법으로 독립변수가 범주형 결과에 미치는 경우를 사용된다. 일반적으로 결과의 유무, 0과

1로 분류되는 경우를 이분형이라고 하고, 종속변수가 이분형일 때 이를 이분형 로지스틱 모형이라고 한다. 1개의 독립변수 X를 사용할 때, 로지스틱 회귀분석은 수식 (1)로 나타난다.

$$\ln[\text{odds}(Y = 1)] = \beta_0 + \beta_1 X \quad (1)$$

여기서 Y는 결과로 Y는 사건이 발생할 때 1, 사건이 발생하지 않을 때 0을 의미한다. β_0 는 절편 항, β_1 은 회귀 계수, X는 독립변수를 의미하고 즉 로그의 변화를 나타낸다. 추가로 odds는 어떤 사건이 일어날 확률을 사건이 일어나지 않을 확률로 나누는 값으로, 수식 (2)로 표현한다.

$$\text{Odds} = \frac{P(\text{event occurring})}{1 - P(\text{event occurring})} \quad (2)$$

수식 (1)에 수식 (2)를 대입하면, 수식 (3)으로 나타난다.

$$\ln\left(\frac{P}{1 - P}\right) = \beta_0 + \beta_1 X \quad (3)$$

수식 (3)을 P에 대해서 정리하면, 수식 (4)로 나타난다.

$$P = \frac{e^{\beta_0 + \beta_1 X}}{1 + e^{\beta_0 + \beta_1 X}} \quad (4)$$

식 (4)를 로지스틱 곡선이라 부른다. 즉 로지스틱 곡선을 이용해 임계점 P를 기준으로 분류한다. 즉 0과 1만 가지는 종속변수를 로지스틱 변환을 해서 선형함수 형태로 만드는 것이 로지스틱 회귀모형이다.

2.2. 판별분석

판별분석은 두 개 또는 두 개 이상의 집단을 독립변수들을 이용해 분류하는 분석이며 이를 두 개 또는 두 개 이상의 집단은 판별 함수로 분류한다. 그중 가장 대중적으로 쓰이는 Fisher의 판별분석인 선형판별분석(LDA)을 소개를 하려고 한다. 먼저 선형판별분석은 두 개의 집단을 독립변수들을 이용해 분류하는 분석으로 선형판별분석을 사용하기 위해서는 등분산 가정을 먼저 시행한다. 만약 등분산이 다른 경우는 이차판별함수(QDA)를 진행한다. 본 연구

에서는 선형판별분석으로 진행한다^[10-11].

두 개의 그룹을 G_1, G_2 라고 하면 G_1 에 속하는 n_1 개의 데이터를 X_i 라 두고 G_2 에 속하는 n_2 개의 데이터를 X_j 라고 했을 때 각각 데이터의 평균은 μ_1, μ_2 이고 공분산 행렬은 Σ 이다. 이때, G_1, G_2 를 최대한 떨어져 구분되게 하는 선형판별함수 $Y = l'X$ 을 구하려고 한다. G_1, G_2 의 평균, Y의 분산은 수식 (5)와 같다.

$$\begin{aligned} \mu_{1Y} &= E(Y/G_1) = l'\mu_1 \\ \mu_{2Y} &= E(Y/G_2) = l'\mu_2 \end{aligned} \quad (5)$$

$$\sigma_Y^2 = \text{Var}(Y) = \text{Var}(l'X) = l'\Sigma l$$

분산을 고려하여 두 개의 그룹의 평균차가 최대가 되는 함수를 구하며, 수식 (6)으로 표현된다.

$$\frac{(\mu_{1Y} - \mu_{2Y})^2}{\sigma_Y^2} = \frac{l'(\mu_1 - \mu_2)(\mu_1 - \mu_2)'l}{l'\Sigma l} \quad (6)$$

이 때 최대화 정리(Maximization theorem)를 이용하여 수식 (6)을 최대화 하는 l의 값은 수식 (7)과 같다.

$$l = c\Sigma^{-1}(\mu_1 - \mu_2) \quad (7)$$

이를 선형판별함수에 대입하면 수식 (8)로 나타난다.

$$y = (\mu_1 - \mu_2)'\Sigma^{-1}X = (\bar{X}_1 - \bar{X}_2)'\Sigma_{pl}^{-1}X \quad (8)$$

각 그룹의 평균점을 구하면 다음 수식 (9)와 같다.

$$\begin{aligned} \bar{y}_1 &= \tilde{l}'\bar{X}_1 \\ \bar{y}_2 &= \tilde{l}'\bar{X}_2 \end{aligned} \quad (9)$$

이를 이용하여 두 수식의 중간점을 계산하면 수식 (10)으로 나타난다.

$$Z = \frac{1}{2}(\bar{X}_1 - \bar{X}_2)'\Sigma_{pl}^{-1}(\bar{X}_1 + \bar{X}_2) \quad (10)$$

새로운 데이터 X_0 에 대해서 $Z > X_0$ 이면 G_1 에 속하고 $Z \leq X_0$ 이면 G_2 에 속한다.

3. 데이터 소개

본 연구에서는 기상청에서 제공하는 종관기상관측의 일별 기상요소와 관측된 온실가스의 대기농도를 기반으로 폭염이 미치는 영향을 알아보하고자 한다. 폭염이란 일 최고기온이 33°C 이상인 날의 수를 의미한다.^[9] 이는 하루 동안 지속적으로 33°C 이상을 의미하는 것이 아니기 때문에 기온이나 폭염에 영향을 미칠 수 있는 기상 요소를 추려 분석에 활용하고자 한다. 기온은 기압, 풍속, 습도 등 여러 기상 요소에 의해 공기가 이동하고, 변화하기 때문에 본 연구에서는 기상 요소 중 평균 풍속, 평균 상대습도, 현지기압과 해면기압의 차, 1시간 최다일사량을 활용한다. 지상관측지침에 따르면 평균 풍속은 1일간의 바람이 대기 상에서 한쪽 방향으로 움직인 총 거리를 24시간으로 나눈 값이다. 평균 상대습도는 대기 중에 실제로 포함되어 있는 수증기의 양과 그때의 기온으로 포함시킬 수 있는 최대 수증기량의 비율이다. 평균 현지 기압은 수은기압계에 의한 관측값으로 온도와 중력의 영향을 받으므로 이에 대한 보정을 한 값이다. 1시간당 최대 일사량은 1시간 동안 태양광선에 직각으로 놓은 1cm 넓이에 1분 동안 복사되는 에너지의 양이다.

또한, 단순 기상정보만 포함하는 것이 아닌 대기 중에 포함된 6대 온실가스까지 함께 포함하여 폭염에 미치는 영향을 분석하고자 한다. 6대 온실가스 중 아산화 질소, 수소불화탄소, 육불화황은 데이터가 충분하지 않아 신뢰성이 떨어져 변수로 채택하지 않고 나머지 6대 온실가스를 변수로 채택하였다. 그 중 이산화탄소는 인간의 화석연료 소비 증가로 배출되는 대표적 온실가스로 대기 중에 머무르는 시간이 100년 이상으로 전체 온실효과의 65%를 차지한다. 메탄은 이산화탄소 다음으로 중요한 온실가스 중 하나로 대기 중에 존재하고 농사, 화석연료 등으로 배출되고 대기 중 머무르는 기간은 가장 짧은 12년이다. 염화불화탄소는 오존층을 직접적으로 파괴하는 온실가스로 앞에 설명한 온실가스에 비해 위험성이 더 커 몬트리올의정서에 의해 사용이 금지되었다. 기상의 시계열적 특성을 활용하여 대기 중의 온실가스 농도를 관측하여 하루단위 변화량을 폭염에 미치

는 영향 요인으로 판단하고자 한다.

본 연구에서는 위와 같은 여러 기상 요소를 통해 광주광역시의 폭염에 미치는 영향을 분석하고자 한다. 이를 관측하기 위해 2008년부터 2022년까지 기상청 기후자료관리에서 제공하는 기상 데이터를 활용한다. 기상 자료는 특정 해에 이상기후가 발생하거나 화산 폭발이나 지진 등 여러 기상 현상에 의해 변화될 수 있기 때문에 2008년부터 2022년까지 5년 주기로 구분하여 추세를 확인하고, 각 주기마다 폭염에 미치는 영향 변수 평균 풍속, 평균 상대습도, 기압 차이, 1시간 최다일사량, 이산화탄소 변화량, 메탄 변화량, 염화불화탄소 변화량을 로지스틱 회귀분석과 선형판별분석을 통해 확인하고자 한다.

4. 분석

4.1. 기술통계

표 1은 폭염의 여부를 연도별로 5월부터 9월까지 관측한 데이터이다. 폭염이 가장 많은 연도는 2018년도로 4개월 동안 총 43일의 폭염이 발생했고, 폭염 비율은 무려 28.1%에 달한다. 폭염이 가장 적은 연도는 2009년도로 4개월 동안 총 6일(3.9%)이 발생했다. 전체적으로 첫 번째 주기인 2008년부터 2012년에는 총 폭염일이 82일(12.0%), 두 번째 주기인 2013년부터 2017년에는 총 폭염일이 104일(15.7%), 세 번째 주기인 2018년부터 2022년에는 총 폭염일이 101일(15.2%)로 나타났으며, 이 중 두 번째 주기인 2013년부터 2017년에서 가장 많은 폭염이 발생하였다. 그림 1은 표 1에서 제시한 폭염 일수 그래프이다. 전반적으로 2018년, 2009년, 2014년을 제외한 다른 해는 비슷한 추이를 보였으며, 평균 기온 상승과는 별개로 발생하는 폭염 일수는 비슷한 수준으로 발생하였다.

Table 1. 폭염데이터 기술통계량 결과

주기	연도	폭염(일)	비폭염(일)	폭염비율(%)
1	2008	18	135	11.8
	2009	6	147	3.9
	2010	20	133	13.1
	2011	13	140	8.5
	2012	25	128	16.3
	소계	82	683	12.0
2	2013	22	131	14.4
	2014	8	145	5.2
	2015	14	139	9.2
	2016	31	122	20.3
	2017	29	124	19.0
	소계	104	661	15.7
3	2018	43	110	28.1
	2019	12	141	7.8
	2020	13	140	8.5
	2021	14	139	9.2
	2022	19	134	12.4
	소계	101	664	15.2

Table 2. 종관기상관측 데이터 기술통계량 결과

주기	연도	평균풍속(ms)		평균상대습도(mm)		기압 차이(hPa)		1시간 최대일사량 (MJ/m)	
		mean	sd	mean	sd	mean	sd	mean	sd
1	2008	2.12	0.80	69.72	10.52	10.16	1.08	2.49	0.81
	2009	2.14	0.88	69.98	13.44	10.55	1.01	2.50	0.93
	2010	2.24	0.78	74.27	11.79	10.61	1.08	2.52	0.84
	2011	2.42	1.02	74.51	11.20	10.37	0.91	2.43	0.88
	2012	2.21	1.03	71.10	11.08	10.47	1.43	2.39	0.82
2	2013	2.20	0.86	70.12	10.88	10.33	1.02	2.42	0.81
	2014	1.98	0.79	68.83	11.20	10.50	1.21	2.32	0.83
	2015	1.73	0.74	72.33	12.14	10.57	1.35	2.56	0.84
	2016	1.64	0.77	78.15	12.14	10.36	1.16	2.62	0.86
	2017	1.57	0.58	75.79	15.41	10.40	0.85	2.71	0.77
3	2018	1.66	0.64	76.53	12.46	10.52	1.35	2.74	0.86
	2019	1.58	0.83	76.55	14.80	10.47	1.21	2.68	0.91
	2020	1.71	0.78	83.47	12.16	10.82	1.66	2.50	0.97
	2021	1.40	0.55	79.51	13.77	10.35	1.17	2.61	0.91
	2022	1.59	0.80	78.94	15.25	10.20	1.32	2.63	0.86

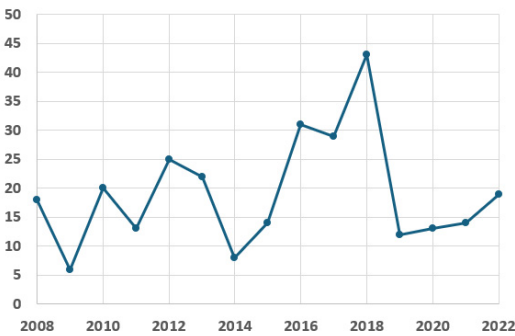


Fig. 1. 2008년부터 2022년까지 폭염 일수

표 2는 광주 지역의 데이터를 일일 단위로 측정된 종관기상관측 데이터에 대한 평균과 표준편차이다. 2011년에 평균 풍속이 2.42로 제일 높게 나타났고, 2021년에 1.40으로 가장 낮은 평균 풍속을 보였다. 또한, 시간이 지남에 따라 평균 풍속이 감소하고 있는 모습을 보이고 있다. 평균 상대습도는 2020년에 83.47로 가장 높게 나타났고, 2014년에 68.83으로 가장 낮게 나타났다. 기압차는 2020년에 10.82로 가장 높게 나타났고, 전반적으로 비슷한 추이를 보이고 있다. 마지막으로 1시간 최대일사량은 2008년부터 2012년에 비해 2013년부터 2022년이 큰 값을 가지고 있음을 알 수 있다.

표 3과 그림 2부터 그림 4까지는 안면도 지역의 데이터를 일일 단위로 측정된 온실가스 데이터에 대한 평균과 표준편차 데이터이다. 표 2랑 다르게 온실가스의 데이터는 결측치가 많기 때문에 데이터 수도 추가하였다. 이산화탄소, 메탄 배경대기농도는 연도가 지날수록 증가하는 추세를 보이고 있다. 염화불화탄소 배경대기농도는 몬트리올 의정서에 의한 사용규제로 인해 연도가 지남에 따라 감소하는 추세를 보이고 있다. 따라서 본 연구에서는 이산화탄소 대기농도, 메탄 대기농도, 염화불화탄소를 활용하며 시계열적 특성을 활용하기 위해 관측된 날짜의 이전 시점과의 농도 변화량을 비교하여 온실가스 대기농도 변화를 새로운 변수로 구성한 후, 분석에 활용하고자 한다.

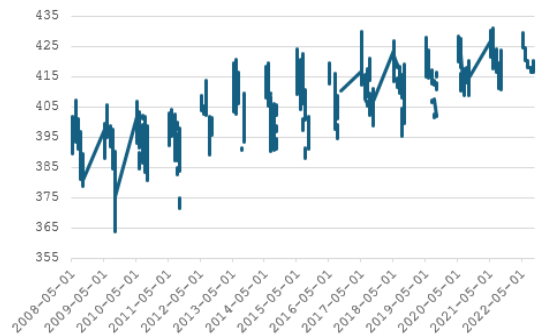


Fig. 2. 2008년부터 2022년까지 이산화탄소 농도

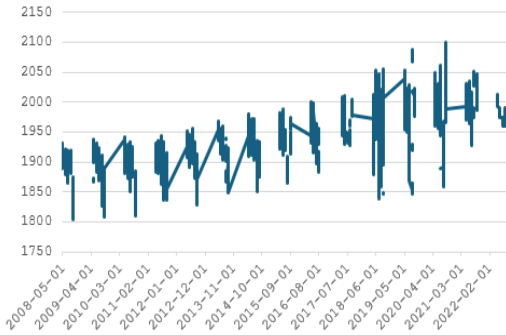


Fig. 3. 2008년부터 2022년까지 메탄 농도



Fig. 4. 2008년부터 2022년까지 염화불화탄소 농도

4.2. 로지스틱 회귀분석

먼저 표 1을 보면 알 수 있듯이 폭염 일수와 비 폭염 일수의 차이가 많이 나기에 폭염 일수 데이터를 2배로 오버샘플링 한 후 로지스틱 회귀분석을 진행하였다. 분석에 활용된 변수로는 평균상대습도, 기압 차이, 1시간 최대 일사량, 이산화탄소 변화량, 메탄 변화량, 염화불화탄소 변화량을 사용했다.

표 4를 보면 2008년부터 2012년은 기압 차이, 1시간 최대 일사량이 폭염에 유의한 영향을 미치는 것으로 판단되고, 2013년부터 2017년은 기압 차이, 1시간 최대 일사량이 유의한 영향을 미치는 것으로 판단된다. 마지막으로 2018년부터 2022년은 평균상대습도, 기압 차이, 1시간 최대 일사량, 메탄 변화량, 이산화탄소 변화량, 염화불화탄소 변화량이 유의한 영향을 미치는 것으로 판단된다. 2018년부터 2022년은 다른 주기에 비해 온실가스가 유의한 영향을 미치는 것으로 판단된다. 여기서 오즈비가 1보다 큰 경우는 폭염이 증가하는 영향을 미치고 1보다 작은 경우는 폭염이 감소하는 영향을 미친다. 2018년부터 2022년까지의 염화불화탄소 변화량은 다른 온실가스 변수에 비해 큰 오즈비를 가지고 있으므로 폭염이 증가하는 강한 영향을 미치는 것으로 나타났다. 표 4의 계수를 활용하여 주기별 폭염을 분류하기 위한 로지스틱 회귀분석의 모형은 다음과 같다.

Table. 3. 온실가스 데이터 기술통계량 결과

주기	연도	이산화탄소			메탄			염화불화탄소		
		N	Mean	Sd	N	Mean	Sd	N	Mean	Sd
1	2008	105	393.32	5.29	106	1893.19	27.07	122	233.95	12.08
	2009	114	393.20	6.27	105	1895.12	23.47	139	246.29	7.22
	2010	97	395.14	5.12	106	1898.45	27.37	134	220.72	13.67
	2011	115	395.91	5.63	120	1899.44	21.14	124	234.93	5.56
	2012	79	402.61	4.61	126	1913.67	25.72	144	220.11	9.46
2	2013	75	405.73	6.31	82	1930.05	23.35	96	225.60	3.95
	2014	92	405.42	6.95	84	1933.85	23.97	117	241.89	11.16
	2015	85	406.20	7.77	80	1937.25	26.50	120	234.59	9.40
	2016	33	409.15	7.55	66	1956.12	26.97	90	237.72	4.78
3	2017	84	412.35	5.90	77	1961.67	22.22	123	232.10	2.32
	2018	82	412.53	6.48	100	1952.95	52.59	131	229.05	1.87
	2019	55	414.56	6.33	65	1973.99	55.44	126	228.15	2.24
	2020	104	417.62	4.93	105	1983.69	59.69	88	223.12	0.98
	2021	91	420.57	4.73	63	1994.11	26.92	124	221.77	0.80
	2022	153	420.70	3.94	153	1987.39	20.83	105	218.05	0.82

주기 1: 2008-2012 주기 2: 2013-2017 주기 3: 2018-2022

$$\ln\left(\frac{\hat{p}_y^{(1)}}{1-\hat{p}_y^{(1)}}\right) = 0.011x_1 - 2.111x_2 + 1.634x_3 - 0.021x_4 - 0.017x_5 + 0.003x_6 + 14.726$$

$$\ln\left(\frac{\hat{p}_y^{(2)}}{1-\hat{p}_y^{(2)}}\right) = 0.070x_1 - 1.267x_2 + 3.759x_3 - 0.143x_4 - 0.003x_5 - 0.016x_6 - 4.596$$

$$\ln\left(\frac{\hat{p}_y^{(3)}}{1-\hat{p}_y^{(3)}}\right) = 0.050x_1 - 1.111x_2 + 1.257x_3 - 0.175x_4 - 0.018x_5 + 0.838x_6 + 3.069$$

이때 x_1 은 평균상대습도, x_2 는 기압차, x_3 은 1시간 최대 일사량, x_4 는 이산화탄소 변화량, x_5 는 메탄 변화량, x_6 은 염화불화탄소 변화량이고

Table. 4. 로지스틱 회귀분석 결과

주기	변수	계수	표준오차	Wald	유의확률	오즈비
2008년~2012년	평균상대습도	0.011	0.019	0.333	0.564	1.011
	기압 차이	-2.111	0.362	33.991	0.000**	0.121
	1시간 최대 일사량	1.634	0.392	17.340	0.000**	5.124
	이산화탄소 변화량	0.021	0.062	0.110	0.741	1.021
	메탄 변화량	-0.017	0.011	2.347	0.126	0.984
	염화불화탄소 변화량	-0.003	0.045	0.005	0.941	0.997
	상수항	14.726	4.196	12.318	0.000**	
2013년~2017년	평균상대습도	0.070	0.028	6.229	0.013*	1.072
	기압 차이	-1.267	0.485	6.813	0.009**	0.282
	1시간 최대 일사량	3.759	0.944	15.857	0.000**	42.894
	이산화탄소 변화량	-0.143	0.088	2.625	0.105	0.867
	메탄 변화량	-0.003	0.017	0.027	0.870	0.997
	염화불화탄소 변화량	-0.016	0.049	0.103	0.748	0.984
	상수항	-4.596	6.407	0.514	0.473	
2018년~2022년	평균상대습도	0.050	0.014	13.324	0.000**	1.052
	기압 차이	-1.111	0.212	27.612	0.000**	0.329
	1시간 최대 일사량	1.257	0.289	18.873	0.000**	3.514
	이산화탄소 변화량	-0.175	0.087	4.048	0.044*	0.840
	메탄 변화량	-0.018	0.009	4.031	0.045*	0.982
	염화불화탄소 변화량	0.838	0.260	10.381	0.001**	2.311
	상수항	3.069	2.667	1.324	0.250	

$p < 0.01$ **, $p < 0.05$ *

$p_y^{(1)}, p_y^{(2)}, p_y^{(3)}$ 은 각각 2008년부터 2012년, 2013년부터 2017년, 2018년부터 2022년의 p_y 를 의미한다. $p_y > 0.5$ 이면 폭염으로 분류하고 $p_y < 0.5$ 이면 비폭염으로 분류한다.

4.3. 판별분석

판별분석 또한 로지스틱 회귀분석과 동일한 변수로 진행하였다. 정준판별함수 계수가 클수록 폭염에 미치는 영향이 크다. 표 5를 보면 2008년부터 2012년은 기압 차이, 1시간 최대 일사량, 평균상대습도, 메탄 변화량, 염화불화탄소 변화량, 이산화탄소 변화량 순으로 폭염에 영향을 미친다. 2013년부터 2017년은 1시간 최대 일사량, 기압 차이, 이산화탄소 변화량, 염화불화탄소 변화량, 메탄 변화량, 평균상대습도 순으로 폭염에 영향을 미친다. 2018년부터 2022년은 기압 차이, 1시간 최대 일사량, 메탄 변화량, 이산화탄소 변화량, 평균상대습도, 염화불화탄소 변화량 순으로 폭염에 영향을 미친다. 이처럼 주기가 증가함에 따라 온실가스의 정준판별함수 계수가 커지는 걸 알 수 있다. 즉, 과거에 비해 온실

가스가 폭염의 여부가 판단되는 것을 알 수 있다. 특히 2018년부터 2022년은 로지스틱 회귀모형과 동일하게 염화불화탄소 변화량이 폭염에 유의미한 변수로 채택되었고 폭염과 강한 영향을 끼친다고 분석할 수 있다.

또한, 표 5의 계수를 활용하여 각 주기별 폭염을 분류하기 위한 선형판별분석의 모형은 다음과 같다.

$$Z_0^{(1)} = 1.321x_1 + 10.363x_2 + 21.351x_3 - 0.229x_4 + 0.019x_5 + 0.307x_6 - 128.759$$

$$Z_1^{(1)} = 1.344x_1 + 9.874x_2 + 22.344x_3 - 0.218x_4 + 0.013x_5 + 0.292x_6 - 128.176$$

$$Z_0^{(2)} = 1.075x_1 + 9.130x_2 + 18.496x_3 - 0.016x_4 - 0.081x_5 - 0.250x_6 - 107.201$$

$$Z_1^{(2)} = 1.145x_1 + 8.564x_2 + 20.298x_3 - 0.101x_4 - 0.084x_5 - 0.238x_6 - 114.351$$

$$Z_0^{(3)} = 0.827x_1 + 9.130x_2 + 14.274x_3 - 0.094x_4 + 0.000x_5 - 0.222x_6 - 99.690$$

$$Z_1^{(3)} = 0.880x_1 + 8.564x_2 + 15.297x_3 - 0.234x_4$$

Table 5. 판별 분석 결과

주기	변수	정준판별함수 계수	분류함수 계수(비폭염)	분류함수 계수(폭염)
2008년~2012년	평균상대습도	0.318	1.321	1.344
	기압 차이	-0.584	10.363	9.874
	1시간 최대 일사량	0.855	21.351	22.344
	이산화탄소 변화량	0.038	-0.229	-0.218
	메탄 변화량	-0.114	0.019	0.013
	염화불화탄소 변화량	-0.057	0.307	0.292
	상수항		-128.759	-128.176
2013년~2017년	평균상대습도	0.711	1.075	1.145
	기압 차이	-0.271	9.130	8.564
	1시간 최대 일사량	1.082	18.496	20.298
	이산화탄소 변화량	-0.230	-0.016	-0.101
	메탄 변화량	-0.050	-0.081	-0.084
	염화불화탄소 변화량	0.064	-0.250	-0.238
	상수항		-107.201	-114.351
2018년~2022년	평균상대습도	0.620	0.827	0.880
	기압 차이	-0.559	9.130	8.564
	1시간 최대 일사량	0.681	14.274	15.297
	이산화탄소 변화량	-0.259	-0.094	-0.234
	메탄 변화량	-0.249	0.000	-0.015
	염화불화탄소 변화량	0.399	-0.222	0.442
	상수항		-99.690	-101.122

$$-0.015x_5 + 0.442x_6 - 101.122$$

이 때, x_1 은 평균상대습도, x_2 는 기압 차이, x_3 은 1시간 최대 일사량, x_4 는 이산화탄소 변화량, x_5 는 메탄 변화량, x_6 은 염화불화탄소 변화량이고 $Z^{(1)}, Z^{(2)}, Z^{(3)}$ 는 각각 2008년부터 2012년, 2013년부터 2017년, 2018년부터 2022년의 Z 를 의미한다. Z_0 은 비폭염의 판별점수, Z_1 은 폭염의 판별점수로 Z_0, Z_1 을 비교하여 더 높은 점수로 분류한다.

5. 결론

본 연구에서는 2008년부터 2022년까지 광주광역시 기상 요소와 온실가스 데이터를 활용하여 로지스틱 회귀분석과 선형판별분석을 통해 주기별 폭염에 미치는 영향을 비교 분석하였다. 기상데이터의 폭염 일수는 비폭염 일수에 비해 적게 발생한 불균형데이터이기 때문에 오버샘플링하여 폭염 일수 데

이터를 증대시켜 분석하였고, 과거 선행연구를 기반으로 폭염에 영향을 미칠 것으로 판단되는 평균상대습도, 기압 차이, 1시간 최대 일사량, 이산화탄소 변화량, 메탄 변화량, 염화불화탄소 변화량을 통해 비교분석하였다. 그 결과, 본 연구를 통해 폭염에 미치는 변화되는 요인을 잘 살펴볼 필요가 있다. 2008년부터 2012년, 2013년부터 2017년까지, 2018년부터 2022년까지의 주기에 따라 미치는 변수가 다름을 확인하였다. 특히, 2018년부터 2022년까지는 이전에 영향을 많이 미치지 않던 온실가스가 영향을 미치는 것으로 나타났고, 이 중, 염화불화탄소의 변화량에 대한 영향이 매우 크게 나타났다. 당연히 온실가스가 폭염에 영향을 많이 미친다고 판단할 수 있으나, 증가폭이 컸던 2013년부터 2017년까지의 주기에는 온실가스가 영향을 미치지 않았기 때문에 이는 그림 2부터 4의 결과처럼 단순하게 온실가스가 증가한다고 폭염에 영향을 미친다고 해석할 수만은 없다. 따라서 우리는 본 연구를 통해 과거보다 현재에 온실가스로부터 폭염에 대한 피해를 입고 있으며 점점 더 온실가스가 증가하는 추세에 폭염에 대한 많은 논의가 필요할 것으로 보인다. 또한, 본

연구는 15년동안의 기상자료를 기반으로 연구를 시행했다면, 2018년도의 폭염이 급격히 많이 발생한 2018년도를 변화점의 기준으로 이전과 이후의 변화 차이를 비교해보고자 한다.

최근 전 세계적으로 기후위기를 극복하기 위해 탄소중립 정책을 시행하여 2050년까지 탄소배출량을 제로로 만들기 위한 노력을 시행하고 있다. 이는 광주광역시도 마찬가지이며, 2030년까지 기업이 필요로 하는 전력을 친환경 신재생에너지로 대체하는 2030 기업 RE100부터 2045년 탄소 중립 달성까지 기후위기를 극복하기 위한 정책을 시행하고 있다. 해당 정책들을 잘 시행할 수 있도록 우리도 온실가스나 탄소배출을 최대한 줄여 폭염으로부터 이상기후로부터 더 나아가 지구온난화로부터 우리를 지켜야 할 것으로 보인다.

Acknowledgements

이 논문은 2019년 대한민국 교육부와 한국연구재단의 지원을 받아 수행된 연구임 (NRF- 2019S1A6A3A01059888).

참고문헌

- [1] 한국., 기후변화에 관한 정부 간 협의체, 제 6차 평가보고서. 외교부, 2023.
- [2] 노복진, 박상훈., 국내외 기후변화의 경제적 영향 평가 연구에 대한 고찰. 한국지역지리학회지, Vol. 29, No. 2, pp. 401-414, 2023.
- [3] 김도우, 정재학, 이종설, 이지선., 우리나라 폭염 인명피해 발생특징. 한국기상학회, Vol. 24, No. 2, pp. 225-234, 2014.
- [4] 장은세, 김진아, 조미영, 김태희, 김준형, 고수완, 최주권., 부산, 울산 경남지역의 군집분석 기법을 활용한 폭염특성 연구. 2017년 한국기상학회 가을학술대회 논문집, pp. 563-564, 2017.
- [5] 김령은, 원정은, 이정민, 최정현, 김상단., 일 최저 기온을 이용한 한국의 폭염사상 추세. 한국습지학회, Vol. 21, No. 4, pp. 344-353, 2019.
- [6] 김성태, 박만식., 영-과잉 회귀모형을 활용한 폭염자료분석. 한국자료분석학회, Vol. 20, No. 6, pp. 2829-2840, 2018.
- [7] 안중배, 이준리, 조세라., PNU CGCM 앙상블 예보 시스템의 겨울철 남한 기온 예측 성능 평가. 한국기상학회, Vol. 28, No. 4, pp. 509-520, 2018.
- [8] 허솔잎, 현유경, 류영, 강현석, 임운진, 김운재., 기상청 현업 기후예측시스템(GloSea5)에서의 극한예측지수를 이용한 여름철 폭염 예측 성능 평가. 한국기상학회, Vol. 29, No. 3, pp. 257-267, 2019.
- [9] 김인혜, 김재민, 이그림, 임철희, 이윤곤., 최근 한반도 폭염 및 열대야의 도시별 특성 분석. 한국기상학회 학술대회 논문집, pp. 385-385, 2021.
- [10] Wright, R.E., "Logistic regression.", 1995.
- [11] Peng, C. Y. J., Lee, K. L., and Ingersoll, G. M., "An introduction to logistic regression analysis and reporting.", The journal of educational research, Vol. 96, No. 1, pp. 3-14, 2002.