

# 산불 발생 피해 면적

2조 김세희, 문수인, 전동인, 전준호



---

시계열 분석 프로젝트

# 목차



## I. 서론

- 주제 선정 배경
- 데이터 설명/ 전처리
- 분석 목적/ Flow Chart

## II. 본론

- Model 1
- Model 2
- 기후 위험도 지수
- Model 3

## III. 결론

- 결과
- 시사점

## 우리나라 산불의 시계열 특성

김정환<sup>1</sup> · 권세명<sup>2</sup> · 강원석<sup>3\*</sup>

접수일(2020년 4월 21일), 수정일(2020년 6월 4일), 게재확정일(2022년 4월 26일)

### Time-Series Characteristics of Forest Fire Occurrence in Korea

Jeong-Hwan Kim<sup>1</sup> · Se-Myung Kwon<sup>2</sup> and Won-Seok Kang<sup>3\*</sup>

Received: APR. 21. 2020, Revised: JUN. 04. 2020, Accepted: APR. 26. 2022

50 ... Journal of Agriculture & Life Science 56(2)

경과에 따라 산불피해액이 급증하는 추세이다(Steel et al., 2015; Abatzoglou and Williams, 2016; Flannigan, et al., 2016).

우리나라는 산불로 인한 피해를 예방하기 위해 지형인자 또는 기상인자를 활용한 산불 발생 특성 확인 및 예측 연구를 비롯하여 (Lee & Jung, 2010; Lee, et al., 2011; Kim, 2014), 산불 확산 경로와 이동속도 예측 알고리즘 연구(Koo & Lee, 2016) 등 산불 피해의 실질적인 예측을 위한 연구가 다양하게 진행되었다(Lee & Jung, 2010; Lee, et al., 2011). 그러나 산불 발생에 관련한

시계열적 특성에 관련한 연구는 일부 산불 현황에 관련된 연구

결과에서 요일별 산불 발생 현황 혹은 발생 시간에 관련한 내용이 시대별로 어떻게 변화하는지에 관련하여 연구되었으나, 구체적인

시계열 자료의 분석은 거의 이루어지지 않았다(Lee et al., 2012).

캐나다, 러시아, 미국 등 세계 각국에서도 산불 발생 특성과 관련

한 다양한 연구가 진행되었는데(Herman & Paul, 2004; Williams et al., 2014), 특히 Black-carbon을 분석하여 구축한 수백 년에

달하는 산불 발생빈도에 관한 시계열적 데이터는 기상인자와 지형

인자와 연결하여 산불모형에 관한 고도화 등에 관련한 연구가 진행

된 바 있으며(Larsen, 1997; Bergeron et al., 2001; Wallenius et al., 2004), 나아가서 이를 활용하여 산불피해자를 다양한 관리지역으로 구분하는 모델을 정립하는 등 토지이용형태 혹은 산림관리

를 제안하는 단계에 이르러 있다(Bergeron, et al., 2004).

털에서, 2019년은 산림청의 산불발생현황에서 자료를 수집하였다. 한편, 통계청 국가통계포털은 월별, 지역별 산불발생 현황자료를 제공하지 않았고, 산림청에서 2003년부터 2019년까지의 구체적인 산불발생 현황자료를 제공하고 있어, 이를 활용하여 2003년부터 2019년까지의 월별 및 지역별 자료를 수집하여 분석에 활용하였다 (Korea Forest Service, 2016; 2018; 2020; Statistics Korea, 2020). 한편 제주도와 울릉도 등 섬의 경우에는 발생빈도가 적어 분석에서 제외하였다.

#### 2. 통계분석

2003년부터 2019년까지 17년간의 일별 산불발생자료를 활용하여 공휴일을 제외한 요일별 예측빈도를 설정하였고, 공휴일은 17년간 공휴일에 산불이 발생할 확률이 평일과 동일하다고 가정하여 예측빈도를 설정하고 분석에 활용하였다. 요일 및 휴일별 예측빈도와 발생빈도를 카이제곱검정(Chi-square test)을 활용하여 검정하였고, 이때, 귀부가설은 총 발생빈도가 공휴일과 요일별로 동일한 것으로, 대립가설은 발생빈도가 공휴일과 요일별로 차이가 있는 것으로 설정하였다. 그리고 일별 발생빈도가 Poisson 분포를 따르는지를 확인하였다. 카이제곱분석에 관한 가설설정을 나타내면 다음과 같다.

# 주제 선정 배경

- 올해 2월, 뉴스를 통해 이례적인 피해 면적을 기록한 산불
- 우리나라는 임목밀도가 높고 산림이 연속
- 산림이 대부분 산불 피해가 큰 침엽수림  
--> 산불 발생시 대형 산불로 이어질 위험이 높음
- 기존 산불 발생 연구는 시계열 자료의 분석의 거의 이루어지지 않았으나 산불 발생 모형에 관한 고도화 연구 진행

## 우리나라 산불의 시계열 특성

김정환<sup>1</sup> · 권세명<sup>2</sup> · 강원석<sup>3\*</sup>

접수일(2020년 4월 21일), 수정일(2020년 6월 4일), 게재확정일(2022년 4월 26일)

# 주제 선정 배경

Time Series Characteristics of Forest Fire Occurrence

50 ... Journal of Agric

경과에 따라 산불피해액  
Abatzoglou and Willia

우리나라는 산불로 인  
기상인자를 활용한 산불  
(Lee & Jung, 2010; L  
경로와 이동속도 예측 일  
피해의 실질적인 예측을  
& Jung, 2010; Lee, e

시계열적 특성에 관련된 연구는 일부 산불 현황에 관련된 연구  
결과에서 요일별 산불 발생 현황 혹은 발생 시간에 관련한 내용이  
시대별로 어떻게 변화하는지에 관련하여 연구되었으나, 구체적인  
시계열 자료의 분석은 거의 이루어지지 않았다(Lee et al., 2012).  
캐나다, 러시아, 미국 등 세계 각국에서도 산불 발생 특성과 관련  
한 다양한 연구가 진행되었는데(Herman & Paul, 2004; Williams  
et al., 2014), 특히 Black-carbon을 분석하여 구축한 수백 년에  
달하는 산불 발생빈도에 관한 시계열적 데이터는 기상인자와 지형  
인자와 연결하여 산불모형에 관한 고도화 등에 관련한 연구가 진행  
된 바 있으며(Larsen, 1997; Bergeron et al., 2001; Wallenius  
et al., 2004), 나아가서 이를 활용하여 산불피해자를 다양한 관리지  
역으로 구분하는 모델을 정립하는 등 토지이용형태 혹은 산림관리  
를 제안하는 단계에 이르러 있다(Bergeron, et al., 2004).

## 산불에 대한 데이터를 시계열적인 특성으로 분석해보면 산불 피해 면적의 특징을 알 수 있지 않을까?

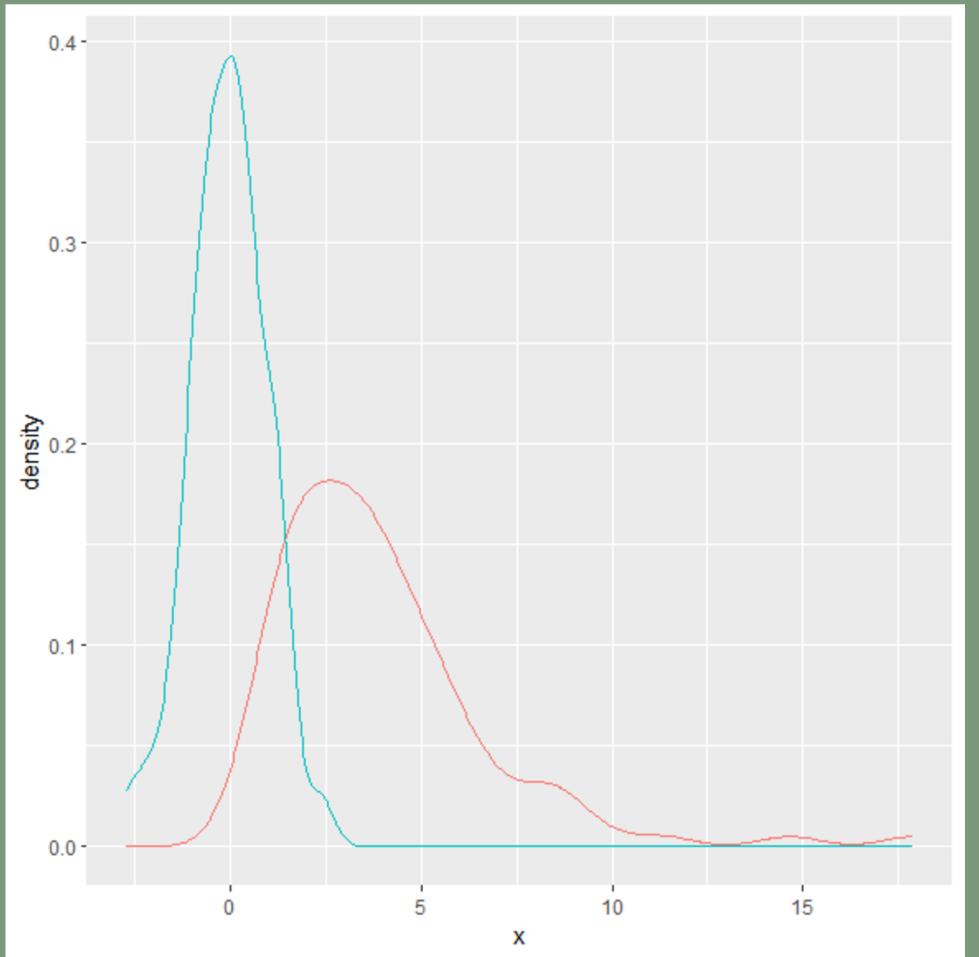
지지 않

### 있으나 산불 발생 모형에 관한 고노와 연구 신앙

#### 2. 통계분석

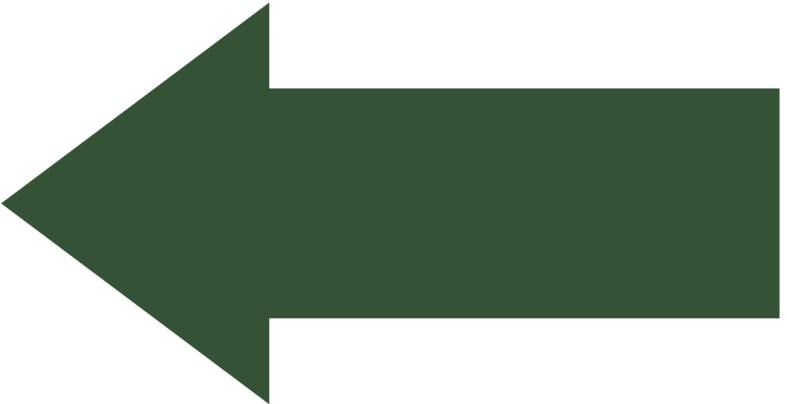
2003년부터 2019년까지 17년간의 일별 산불발생자료를 활용  
하여 공휴일을 제외한 요일별 예측빈도를 설정하였고, 공휴일은  
17년간 공휴일에 산불이 발생할 확률이 평일과 동일하다고 가정하  
여 예측빈도를 설정하고 분석에 활용하였다. 요일 및 휴일별 예측  
빈도와 발생빈도를 카이제곱검정(Chi-square test)을 활용하여 검  
정하였고, 이때, 귀부가설은 총 발생빈도가 공휴일과 요일별로 동  
일한 것으로, 대립가설은 발생빈도가 공휴일과 요일별로 차이가  
있는 것으로 설정하였다. 그리고 일별 발생빈도가 Poisson 분포를  
따로는지를 확인하였다. 카이제곱분석에 관한 가설설정을 나타내  
면 다음과 같다.

# 2차 분석 목적



"지역별 산불빈도 시기 차이"는 시계열 분석이라  
기보단 두 그룹의 차이

-->"피해면적이 시간에 따라 변하는지 분석"  
+ 기후 요인 고려



# 1차 분석 목적

## 목표

### CHECK 1

데이터 분석 시 매 달을 기준으로 산불의 전형적인 발생 빈도가 보인다. 이를 바탕으로 지역 별로 산불의 발생 시기에 차이가 있는지를 예측해보자.

ex) 산이 많은 지역은 산불이 더 자주 발생하는가?

### CHECK 2

피해면적이 시간에 따라 줄어드는 trend가 있는지 분석해보고 대처능력이 좋아졌는지 확인해보자.

ex) 피해면적과 날짜를 비교하면, 어느 시점의 산불이 더 위협적일까?

ex) 산불이 자주 발생하는 지역은 감소하는 추세가 존재하는가?

ex) 사람들의 안전불감증에 따라 언제 다시 피해가 증가하는가?

# 데이터설명

## 산불데이터

- 출처 : 산림청-산불발생현황
- 2013년-2023년 간의 데이터
- Colnames : "발생일시\_년", "발생일시\_월", "발생일시\_일", "발생일시\_시간", "발생일시\_요일", "진화종료시간\_년", "진화종료시간\_월", "진화종료시간\_일", "진화종료시간\_시간", "발생장소\_관서", "발생장소\_시도", "발생장소\_시군구", "발생장소\_읍면", "발생장소\_동리", "발생원인\_구분", "발생원인\_세부원인", "발생원인\_기타", "피해면적\_합계"

## 기후데이터

- 출처 : 기상청
- 2013년-2023년 간의 데이터
- Colnames : "지점", "지점명", "일시", "기온", "풍속", "습도", "현지기압", "온도"

## 산불데이터(2013~2023)

A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R
1	발생일시_년	발생일시_월	발생일시_일	발생일시_시간	도진화종료시	진화종료시	진화종료시	진화종료시	발생장소_관서	발생장소_시도	발생장소_시군구	발생장소_읍면	발생원인_구분	발생원인_세부원인	발생원인_기타	피해면적_합계	
2	2023	9	10 12:16	일	2023	9	10 12:50	부산	부산	기장	기장	만화	기	성묘객실화	구발생원인_구발생원인_피해면적_합계		
3	2023	9	8 15:07	금	2023	9	8 17:17	강원	강원	삼척	하장	숙암	입	성묘객실화 영정사진주	0.01		
4	2023	8	20 12:45	일	2023	8	20 18:00	전남	전남	해남	화원	기	기타(직접입	기타(직접입 미상	0.03		
5	2023	6	20 15:16	화	2023	6	20 19:00	북부	서울	관악		남현	기	방화	방화	0.02	
6	2023	6	19 14:45	월	2023	6	19 16:50	강원	강원	춘천	남산	산수	기	기타(직접입 불씨취급부	기타(직접입 태양광시설	0.13	
7	2023	6	19 14:13	월	2023	6	19 17:00	충북	충북	보은	내북	화전	기	작업장실화	예조기스파	0.1	
8	2023	6	19 13:35	월	2023	6	19 15:00	전북	전북	장수	계북	월현	기	작업장실화	산림사업장	0.05	
9	2023	6	18 16:39	일	2023	6	18 19:40	경남	경남	함안	군북	사존	입	기타(직접입 입산자실화	0.05		
10	2023	6	18 12:37	일	2023	6	18 13:39	강원	강원	양양	현남	주	쓰	기타(직접입 쓰레기소각	0.09		
11	2023	6	15 16:39	목	2023	6	15 18:25	전북	전북	남원	산동	대상	입	기타(직접입 입산자실화	0.05		
12	2023	6	14 13:52	수	2023	6	14 17:00	경남	경남	하동	화개	삼신	입	기타(직접입 입산자 담뱃	0.03		
13	2023	6	13 12:51	화	2023	6	13 17:00	충남	충남	충남	천안 동남	병천	도원	입	기타(직접입 입산자실화	0.06	
14	2023	6	12 15:39	월	2023	6	12 18:30	충남	충남	논산	양촌	산직	담	기타(직접입 담뱃불실화	0.02		
15	2023	6	3 16:57	토	2023	6	3 19:28	경기	경기	포천	이동	장암	입	기타(직접입 입산자실화	0.01		
16	2023	5	27 12:28	토	2023	5	27 13:33	강원	강원	정선	정선	봉양	입	기타(직접입 입산자실화	0.02		
17	2023	5	27 04:23	토	2023	5	27 11:00	경기	경기	화성	장지	입	기타(직접입 입산자 담뱃	0.4			
18	2023	5	26 22:16	금	2023	5	27 09:00	동부	강원	평창	대화	대화	입	기타(직접입 입산자실화	0.6		
19	2023	5	26 15:07	금	2023	5	26 16:20	강원	강원	원주	소조	평장	쓰	기타(직접입 쓰레기소각	0.09		
20	2023	5	25 14:37	목	2023	5	25 16:37	경기	경기	포천	영중	금주	담	기타(직접입 담뱃불실화	0.03		
21	2023	5	25 12:48	목	2023	5	25 14:25	전북	전북	장수	장계	기	기타(직접입 고속도로차	0.05			
22	2023	5	24 11:13	수	2023	5	24 13:00	경북	경북	문경	불정	입	기타(직접입 입산자실화	0.02			
23	2023	5	23 16:52	화	2023	5	23 20:00	북부	강원	횡성	서원	창준	기	기타(직접입 용접화재(죽	0.6		
24	2023	5	23 16:30	화	2023	5	23 16:45	경기	경기	의정부	의정부	민락	입	기타(직접입 입산자실화	0.02		
25	2023	5	23 14:43	화	2023	5	23 15:50	경기	경기	고양	덕양	용두	쓰	기타(직접입 쓰레기소각	0.1		
26	2023	5	23 13:39	화	2023	5	23 14:39	경북	경북	경북	청도	용선	기	기타(직접입 성묘객실화	0.06		
27	2023	5	22 19:39	월	2023	5	22 20:20	경기	경기	과천	막계	기	기타(직접입 미상	0.01			
28	2023	5	17 21:42	수	2023	5	18 02:01	경기	경기	화성	증	담	기타(직접입 담뱃불실화	0.13			
29	2023	5	17 16:36	수	2023	5	17 17:40	세종	세종	금남	용포	용포	기	기타(직접입 입산자실화	0.02		
30	2023	5	17 11:20	수	2023	5	17 13:00	강원	강원	평창	대화	담	기타(직접입 담뱃불실화	0.3			
31	2023	5	17 10:44	수	2023	5	17 14:30	충남	충남	당진	원당	입	기타(직접입 입산자실화	0.35			
32	2023	5	16 14:35	화	2023	5	16 17:00	충남	충남	청양	적곡	입	기타(직접입 입산자실화	0.03			
33	2023	5	16 13:44	화	2023	5	16 15:20	경북	경북	포항	봉	홍해	기타(직접입 입산자실화	0.1			
34	2023	5	16 13:39	화	2023	5	17 14:20	세종	세종	전의	원성	기	작업장실화	용접불씨부	0.17		
35	2023	5	12 13:48	금	2023	5	12 15:30	경남	경남	합천	가야	죽전	입	기타(직접입 입산자실화	0.01		

## 기후데이터(2013~2023)

A	B	C	D	E	F	G	H	I	J
1	지점	지점명	일시	기온..C.	풍속.m.s.	습도...	현지기압.h	지면온도..C.	
2	1	90 속초	2013/01/01 0:00	-4.9	6.1	43	1016.7	-4.9	
3	2	90 속초	2013/01/01 1:00	-4.6	5.5	42	1016.8	-4.7	
4	3	90 속초	2013/01/01 2:00	-4.4	4.3	41	1017.4	-4.7	
5	4	90 속초	2013/01/01 3:00	-5	3.5	45	1017.4	-4.9	
6	5	90 속초	2013/01/01 4:00	-5.9	2.7	53	1017	-5.6	
7	6	90 속초	2013/01/01 5:00	-5.3	2.1	50	1016.2	-5.8	
8	7	90 속초	2013/01/01 6:00	-5.6	1.5	52	1015.9	-6.1	
9	8	90 속초	2013/01/01 7:00	-6.2	2.1	56	1015.5	-5.7	
10	9	90 속초	2013/01/01 8:00	-6.1	0.8	56	1015.3	-5.1	
11	10	90 속초	2013/01/01 9:00	-4.6	1.1	56	1015	-3.8	
12	11	90 속초	2013/01/01 10:00	-2.8	0.7	52	1014.5	-2.7	
13	12	90 속초	2013/01/01 11:00	-0.7	1.1	58	1014.1	-0.6	
14	13	90 속초	2013/01/01 12:00	1.1	1.2	57	1012.5	1.3	
15	14	90 속초	2013/01/01 13:00	2	2.5	54	1012	2.8	
16	15	90 속초	2013/01/01 14:00	1.7	4.8	48	1012.2	2.6	
17	16	90 속초	2013/01/01 15:00	1.1	5.2	45	1012.7		

# 데이터 선정

## 선정 기준

- 산림청 및 기상청에 있는 데이터로 별도의 웹 크롤링 없이 다운로드 가능
- 원하는 열을 임의로 지정하여 저장 가능
- 기후 데이터의 컬럼은 산불과 연관이 있는 기상인자를 특정하여 구성

한국 방재 학회 논문집
제11권 4호 2011년 8월
pp. 123 ~ 130

## CCGIS를 활용한 강원도 산불의 기후변화 취약성 평가

The Vulnerability Assessment of Forest Fire in Gangwon Province Using CCGIS

채희문\* · 엄기증\*\* · 이시영\*\*\*

Chae, Hee Mmun · Um, Gi Jeung · Lee, Si-Young

### Abstract

Since forest sector is increasingly affected by climate change and a forest fire, one of forest disasters, makes a huge impact on a forest ecosystem and regional people's life in general, the vulnerability assessment of forest fires is extremely important. Gangwon province area was selected for the study. Eighty one percent of Gangwon province's land is covered by forest, and 8 cases of big fire, over 300 ha occurred from 2000 to 2009. The vulnerability assessment has been carried out with a help of CCGIS program developed by National Institute of Environmental Research. Climate Exposure Index, Sensitivity Index, Adaptation Capacity Index were employed to evaluate the vulnerability of Gangwon province's forest fire. It turned out that East Sea area of Yeongdong was highly vulnerable to forest fire and various meteorological factors could be involved. The future vulnerability of forest fire in Gangwon province would increase even though the regional difference could exist. The results of the study will be used to establish Gangwon province's adaptation policy for the forest sector.

**Key words :** climate change, forest fire, vulnerability assessment, adaptation, Climate Exposure Index, Sensitivity Index, Adaptation Capacity Index, meteorological factors

### 요지

기후변화가 산림부문 전반에 미치는 영향은 지대하며 그에 대한 평가는 여러가지 측면에서 면밀한 분석이 필요하다. 특히 산림재해 중 산불에 대한 영향 평가는 산림생태계 및 지역 주민의 삶에 미치는 영향이 크기 때문에 매우 중요하다. 본 연구는 국립환경과학원에서 개발한 CCGIS프로그램을 이용하여 강원도의 산불 취약성 평가를 실시하였다. 강원도는 면적의 81%가 산림으로 구성되어 있고 전국 산림면적의 21%를 차지하고 있어 기후변화에 의해 산림이 영향을 많이 받는 지역으로, 최근 10년(2000~2009)간 한국에서 발생한 300 ha이상의 대형 산불 10건 중 8건이 발생되었다. 기후노출지수, 민감도지수, 적응능력지수를 이용하여 산불에 대한 취약성을 평가하였다. 강원도에서 산불 취약성이 큰 지역은 영동 동해안이고 기상인자에 의해 크게 좌우하는 것으로 분석되었다. 미래 기후시나리오(A1B)하의 산불 취약성은 지역에 따라 다르지만 전체적으로 커 질것으로 평가되었다. 취약성 평가는 기후변화적용 대책을 수립하기 위한 중요한 기초자료로서 본 연구의 결과는 향후 강원도의 산림부문 적응 대책수립의 기초자료 활용될 것이다.

**핵심용어 :** 기후변화, 산불, 취약성평가, 적응, 기후노출지수, 민감도지수, 적응능력지수, 기상인자

### 1. 서 론

전 지구적인 기후변화에 의해 평균온도의 상승, 강우량의 지역적 감소 및 증가 추세 등 다양한 기상재해 현상이 우리나라를 비롯하여 세계적으로 만연되고 있고 기후변화가 초래하는 영향이 수자원, 생태계, 해양, 산림, 농업 등 여러 분야에 걸쳐 나타나고 있다. 산림분야에서 평균온도의 상승은 산불 조심기간 동안 산불위험률의 상승을 가져오게 될 것이고, 강우량의 변화 또한 산불 위험률에 막대한 영향을 미치게 될

것이다. 따라서 기상인자의(평균온도, 상대습도, 풍속, 강우량 등)변화는 산림화재의 변화량을 크게 좌우 할 수 있으며, 최근의 기후변화는 평균온도의 증가, 봄철 및 가을철의 강우량 감소 등을 유발하여 산림화재에 매우 민감한 영향을 미칠 것으로 보여지고 있다.

특히, 산불조심기간 동안 기상인자와 산림화재의 관계 분석에서 월별 평균기온의 상승과 상대습도의 변화에 따라 산불 발생 비율의 차이가 발생하며(박연희, 2011), 산불발생 경향 분석(1991-2008)을 통하여 기후변화가 산불 발생에 주요하게

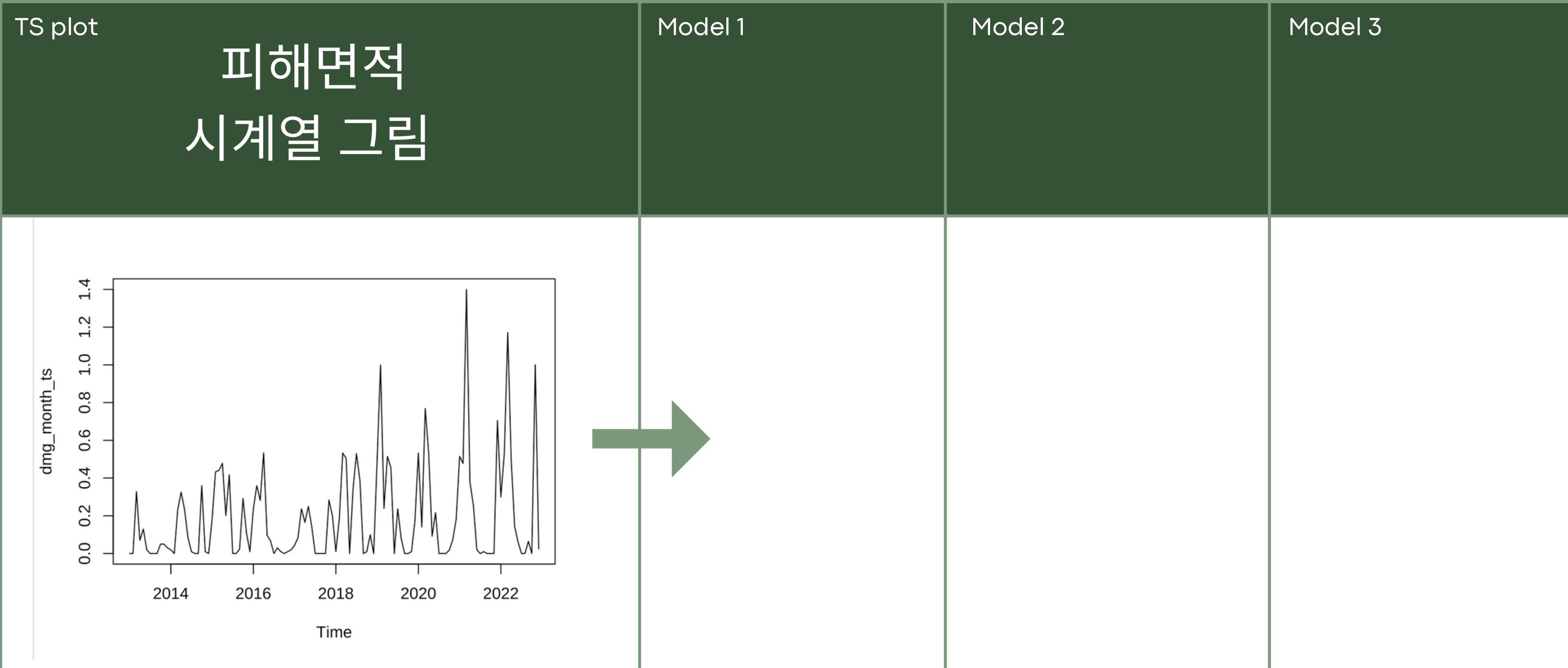
# 데이터 전처리

## 전처리

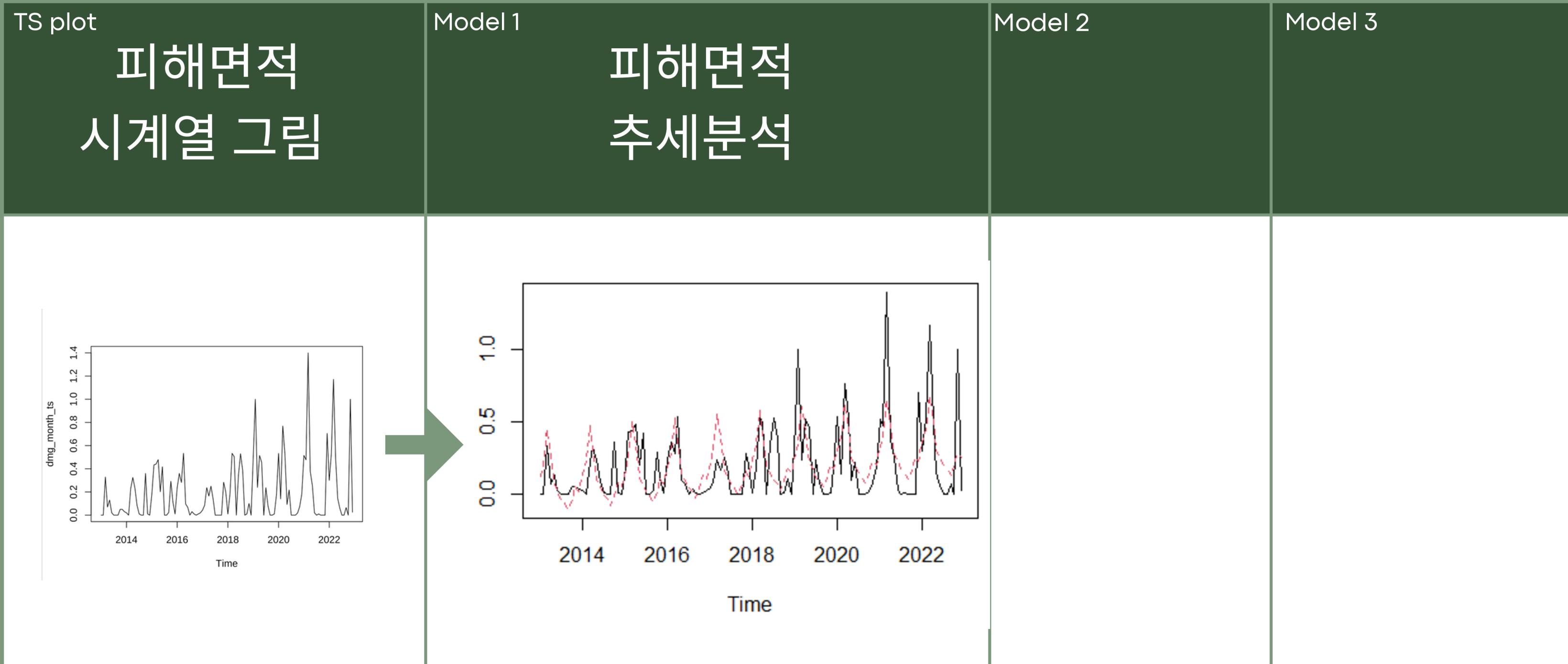
- 기후데이터와 산불데이터의 발생일을 조인시켜 하나의 데이터로 구성
- 피해면적(dmg)dmg의 결측값은 0으로 대치, 1.4이상은 1.4로 대치
- 월별 변화를 위해서 시간단위로 나누어져있는 기후인자 를 각 평균을 구하여 대치

A	B	C	D	E	F	G	H	I	J	K
1	date	dmg	loc	adj_celcius	adj_wind	adj_humid	adj_atm	adj_surf		
2 1	2022/12/25 21:07	0.03	강릉	0.83	3.53	42.116667	1017.5467	-0.735		
3 2	2022/12/05 13:32	0.02	강릉	5.7066667	1.5133333	27.533333	1022.02	9.3466667		
4 3	2022/11/25 23:43	1	강릉	15.971667	7.44	44.15	1005.0717	8.915		
5 4	2022/09/24 15:23	0.01	원주	21.791667	1.9933333	40.233333	998.56167	31.556667		
6 5	2022/09/22 14:48	0.12	홍천	23.2	0.96	44.8	997.88	37.5		
7 6	2022/06/04 11:16	0.06	강릉	23.326667	2.4733333	63	1009.7	42.04		
8 7	2022/05/26 8:40	0.02	강릉	21.4	2.5666667	40.333333	1001	27.766667		
9 8	2022/05/25 15:00	0.21	인제	26.6	2.3	35	982.7	41		
10 9	2022/05/21 17:18	0.01	원주	24.38	2.83	45.3	989.1	29.58		
11 10	2022/05/18 19:41	0.15	철원	20.396667	2.1433333	53.783333	995.55167	19.203333		
12 11	2022/05/18 4:43	0.2	원주	13.798333	0.2716667	73.3	999.24333	9.9283333		
13 12	2022/05/15 17:36	0.18	홍천	20.46	2.66	22.6	994.08	28.58		
14 13	2022/05/03 16:16	0.24	태백	17.753333	3.2333333	28.8	933.03333	22.8		
15 14	2022/04/25 16:57	0.05	영월	25.365	2.755	30.85	981.405	25.875		
16 15	2022/04/24 22:40	0.06	강릉	23.766667	4.0333333	37.333333	1005.3667	15.533333		
17 16	2022/04/24 9:38	0.01	강릉	23.466667	4.54	40.933333	1005.7733	30.393333		
18 17	2022/04/22 21:53	0.6	춘천	15.6	1.0333333	73.466667	1002.6833	14.041667		
19 18	2022/04/21 14:58	1.12	춘천	23.16	3.3966667	28.066667	1000.5267	34.486667		
20 19	2022/04/20 13:00	12.51	영월	21.2	2.6	14	986.9	36.8		
21 20	2022/04/17 12:31	0.2	인제	18.81	3.555	19.45	993.23167	31.446667		
22 21	2022/04/15 17:49	0.02	춘천	15.205	3.7533333	41.533333	1005.9517	18.075		
23 22	2022/04/12 15:14	0.3	홍천	26.653333	2.6	37.533333	991.55333	38.103333		
24 23	2022/04/10 12:41	13	인제	23.466667	3.8216667	21.9	992.01667	34.413333		
25 24	2022/04/04 12:37	0.3	강릉	18.838333	2.67	25.383333	1022.4917	24.323333		
26 25	2022/03/12 14:09	8	홍천	18.28	1.89	46.85	997.81	25.82		

# 분석 Flow Chart



# 분석 Flow Chart



# 단순 분해법 적합

01

전처리

선형 모델 적합.

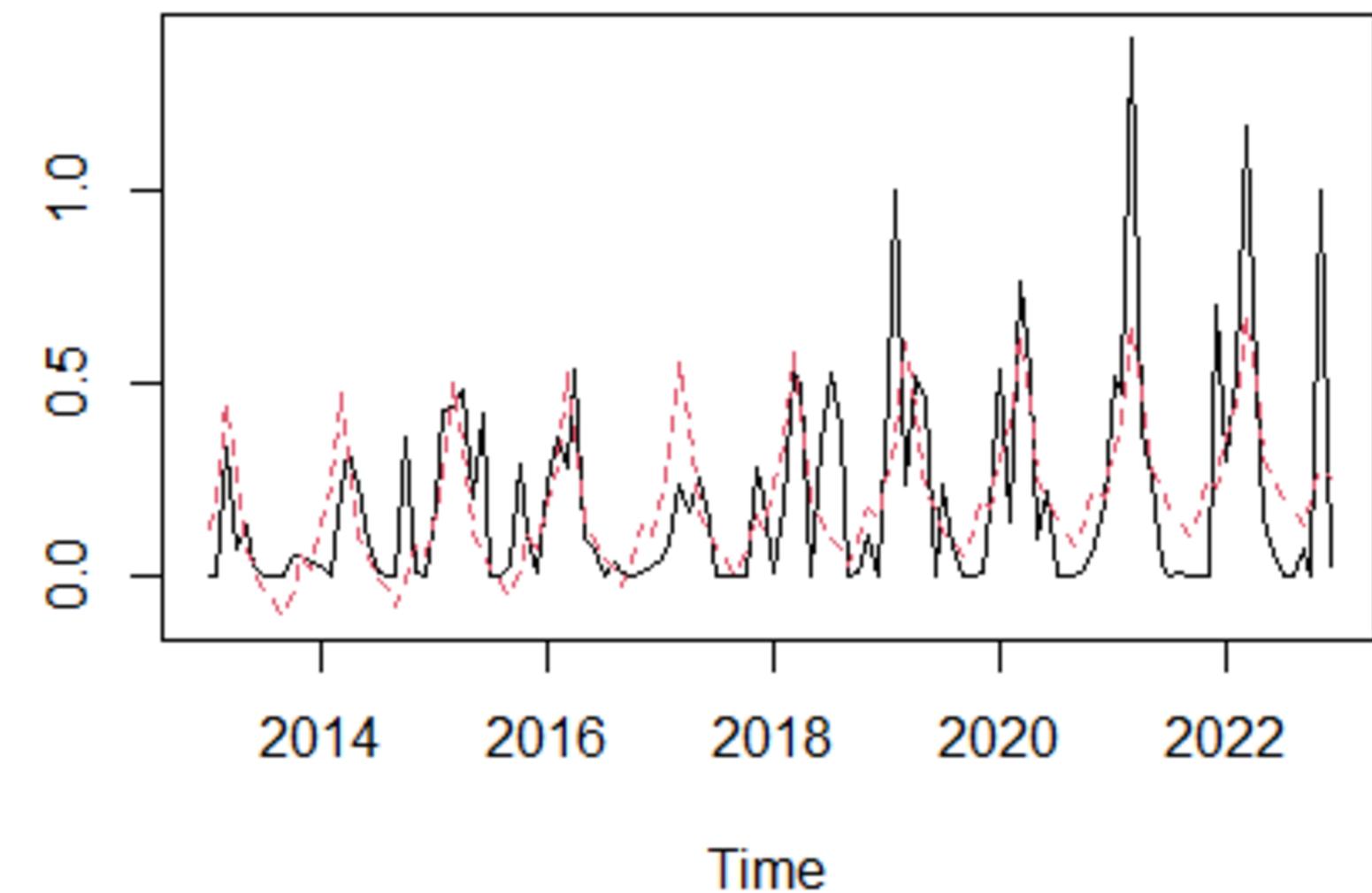
02

TS 모델링 분석 및 해석

적합 그래프 & 잔차 그래프

# Model 1: Trend + Seasonality

1)  $\text{dmg} \sim T + S + \text{error}$



# 1. 전처리

## 월별 정보를 포함하는 tibble 생성

- y는 연도, m은 월, ym은 연도와 월을 합친 것

## 기후 데이터 처리

- dmg 값이 1.4 이상인 경우 1.4로 대체 및 월별 평균 피해량 계산
- 중복 값 제거 및 dmg 열의 NA 값 0으로 대체

## 시계열 객체 생성 및 factor 생성

- dmg 열을 시계열 객체 dmg\_month\_ts로 변환
- 시계열의 계절 성분을 나타내는 factor l을 생성

## 선형 모델 적합

- 종속변수 : 시계열 객체 dmg\_month\_ts
- 독립변수 : 시간 지수 t\_ts, 계절 성분 lmat\_

# 2. TS 모델링 분석 및 해석

```
> summary(model_dmg_TS)

Call:
lm(formula = dmg_month_ts ~ t_ts + Iimat_ + 0)

Residuals:
    Min      1Q   Median      3Q     Max 
-0.36294 -0.13395 -0.02151  0.07796  0.74639 

Coefficients:
            Estimate Std. Error t value Pr(>|t|)    
t_ts        0.0021530  0.0005503  3.913  0.000161 ***
Iimat_I1    0.1144246  0.0723375  1.582  0.116643    
Iimat_I2    0.2002661  0.0725695  2.760  0.006809 **  
Iimat_I3    0.4404666  0.0728049  6.050  2.17e-08 ***
Iimat_I4    0.2747207  0.0730436  3.761  0.000276 ***
Iimat_I5    0.0588183  0.0732857  0.803  0.423991    
Iimat_I6    0.0075524  0.0735312  0.103  0.918385    
Iimat_I7   -0.0536655  0.0737799 -0.727  0.468586    
Iimat_I8   -0.0827352  0.0740318 -1.118  0.266257    
Iimat_I9   -0.1258048  0.0742870 -1.693  0.093270 .  
Iimat_I10  -0.0647911  0.0745454 -0.869  0.386711    
Iimat_I11   0.0247225  0.0748070  0.330  0.741680    
Iimat_I12  -0.0076471  0.0750716 -0.102  0.919055    
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.2078 on 107 degrees of freedom
Multiple R-squared:  0.6376,    Adjusted R-squared:  0.5935 
F-statistic: 14.48 on 13 and 107 DF,  p-value: < 2.2e-16
```

## 적합 통계량

- Trend와 2, 3, 4월에 대한 계수 p-value가 작다.
- 즉, Trend와 2, 3, 4월에 대해 적합 신뢰도가 높다.

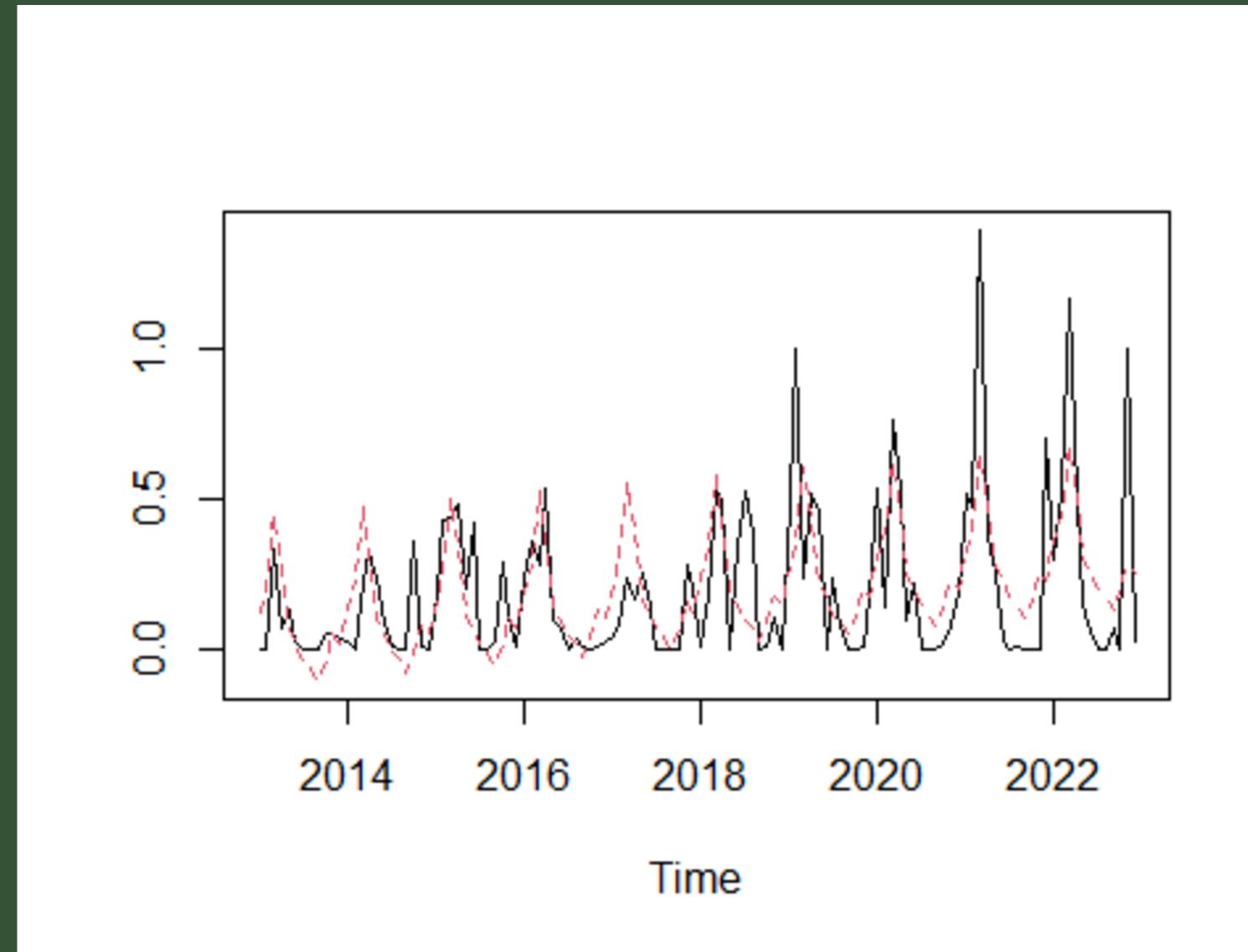
## 적합 그래프

- Seasonality 흐름은 잘 반영되었다.
- 그러나, 튜는 값들에 대한 적합은 잘 반영되지 않았다.

## 잔차 그래프

- normal이라고 보기 힘들다.
- t-test 결과, 평균이 0인 것은 확인되었다.
- 자기상관이 11에 있지만, 유의한 수준인지는 확인불가하다,

## 2. TS 모델링 분석 및 해석



### 적합 통계량

- Trend와 2, 3, 4월에 대한 계수 p-value가 작다.
- 즉, Trend와 2, 3, 4월에 대해 적합 신뢰도가 높다.

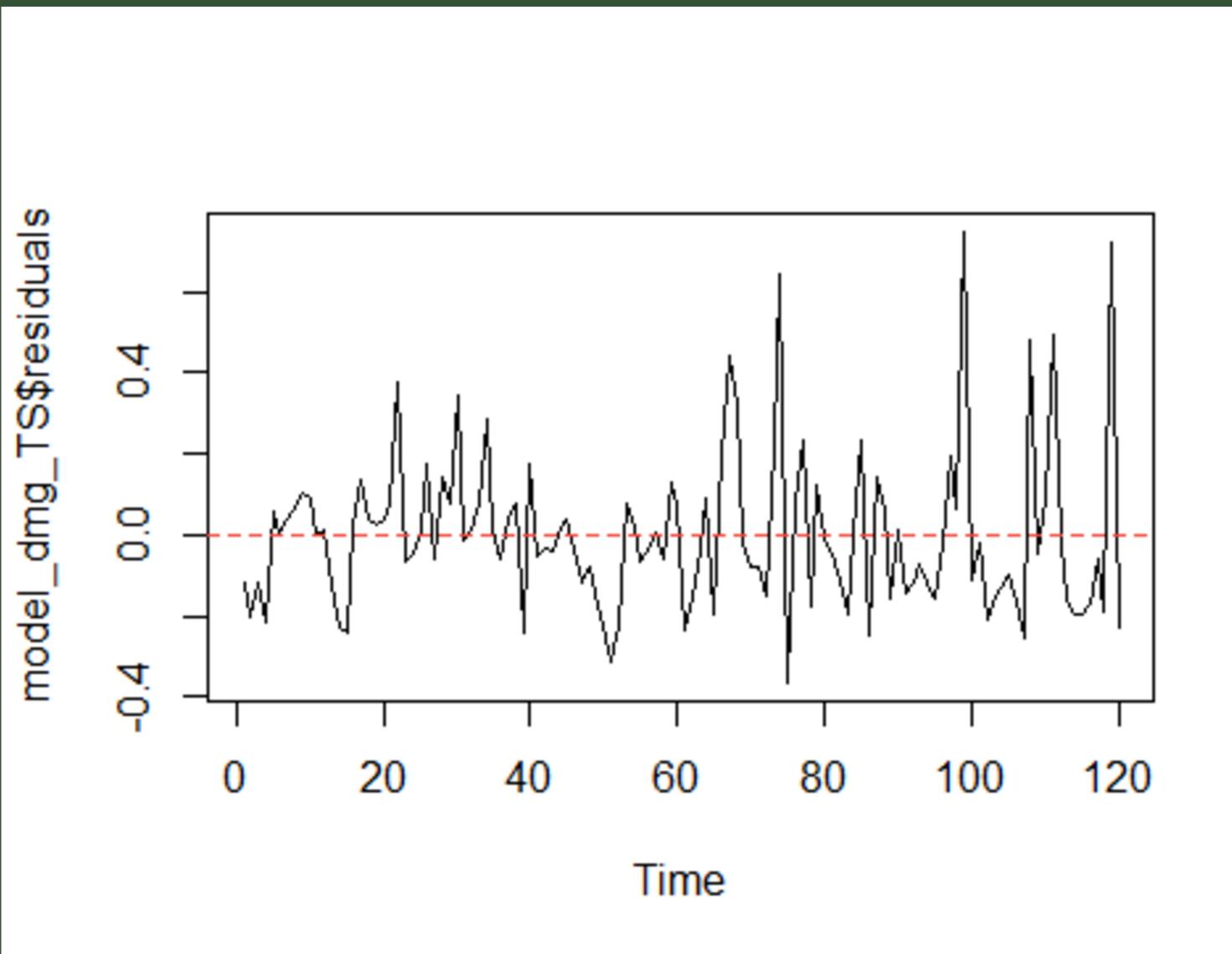
### 적합 그래프

- 시계열 `dmg_month_ts`와
- 모델의 적합 값 `model_dmg_TS$fitted.values`을 비교
- Seasonality 흐름은 잘 반영되었다.
- 그러나, 튜는 값들에 대한 적합은 잘 반영되지 않았다.

### 잔차 그래프

- normal이라고 보기 힘들다.
- t-test 결과, 평균이 0인 것은 확인되었다.
- 자기상관이 11에 있지만, 유의한 수준인지는 확인불가하다,

## 2.. TS 모델링 분석 및 해석



```
> nortest::ad.test(model_dmg_TS$residuals)
```

Anderson-Darling normality test

```
data: model_dmg_TS$residuals
```

```
A = 2.7181, p-value = 6.915e-07
```

```
> t.test(model_dmg_TS$residuals)
```

One Sample t-test

```
data: model_dmg_TS$residuals
```

```
t = -2.0564e-16, df = 119, p-value = 1
```

```
alternative hypothesis: true mean is not equal to 0
```

```
95 percent confidence interval:
```

```
-0.03561172 0.03561172
```

```
sample estimates:
```

```
mean of x
```

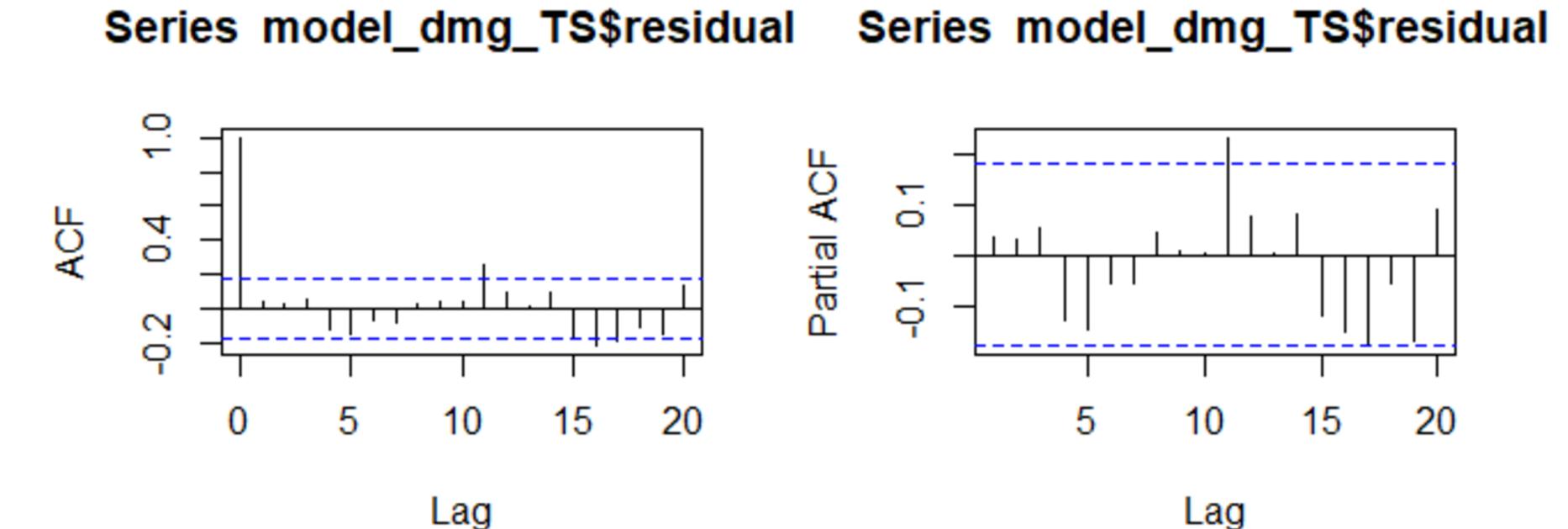
```
-3.698372e-18
```

### 잔차 그래프

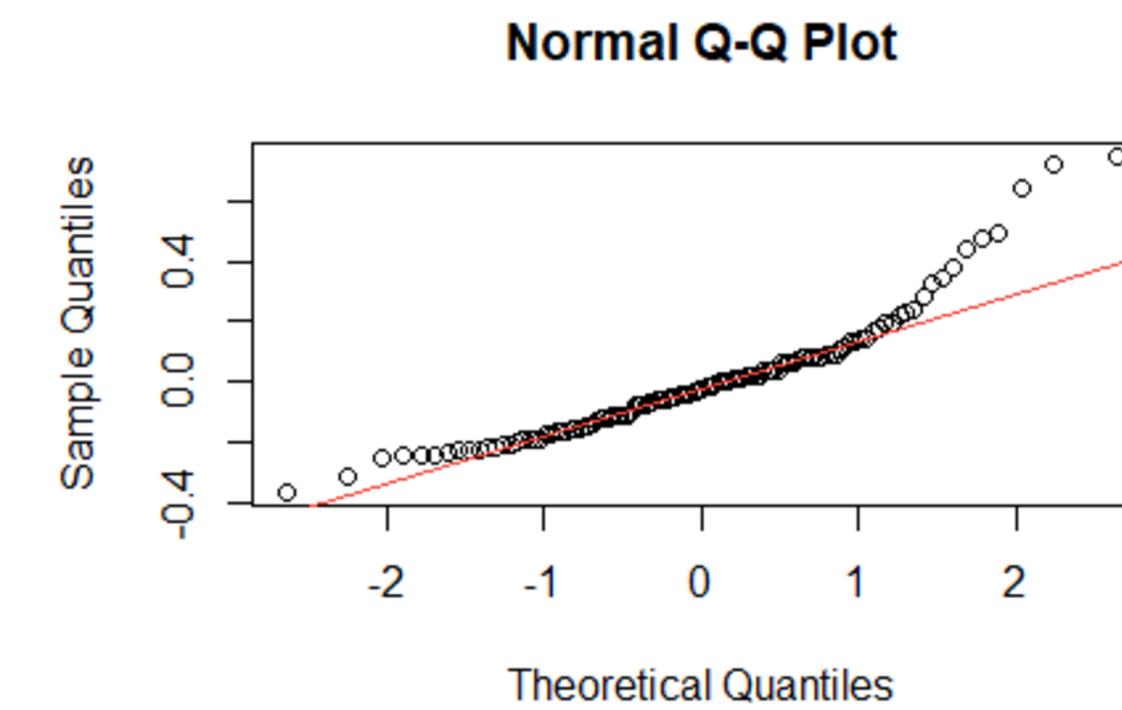
- normal하다고 보기 어렵다.
- t-test 결과, 평균이 0인 것은 확인되었다.
- 자기상관이 11에 있지만, 유의한 수준인지는 확인불가하다,

## 2. TS 모델링 분석 및 해석

잔차의 자기상관 및 부분자기상관함수



정규성 검정



# 기후 요소 적용과정

01

기후 요소 EDA

풍속, 기온, 습도 EDA

02

피해면적과 기후 요소간의  
관계 확인

피해면적과 상관관계 높은 기후요소  
선정

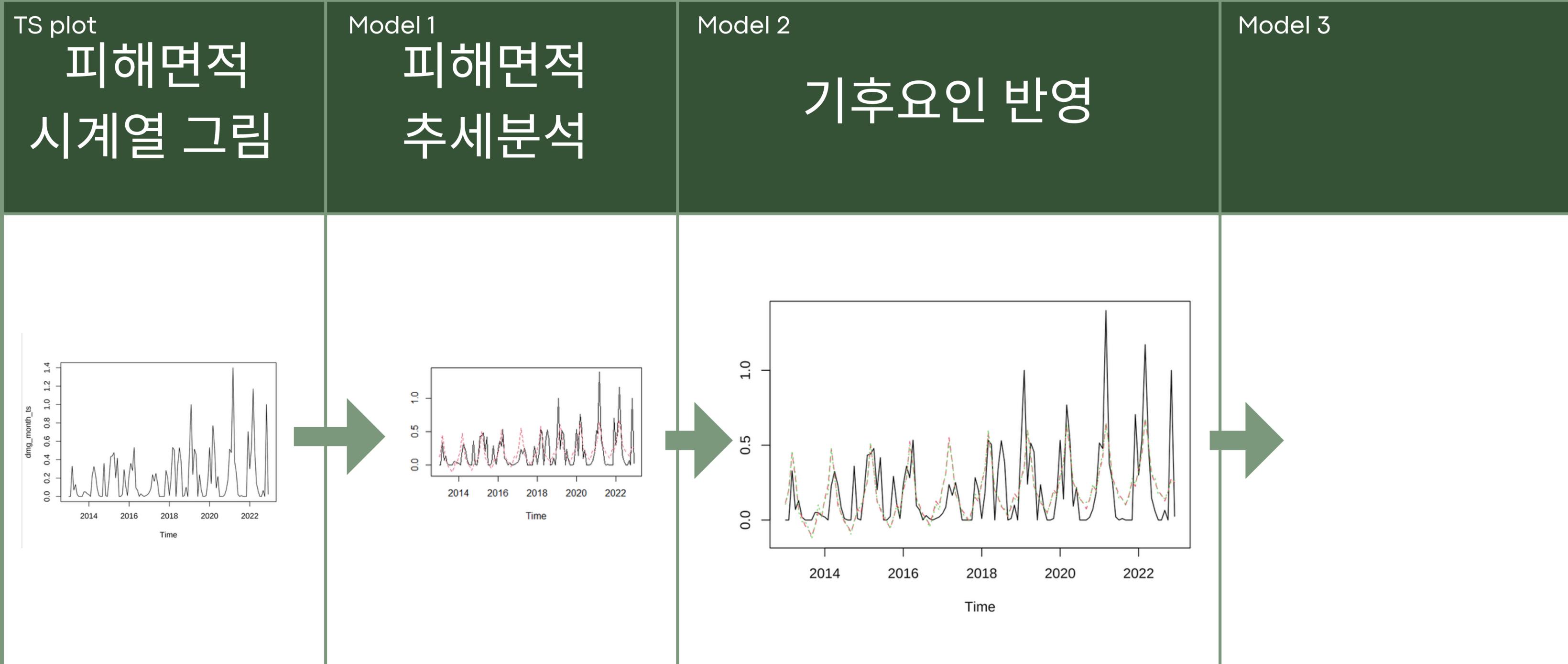
03

Trend + Seasonality +  
기후요인 모델 적용

model 2: T+S+wind+err

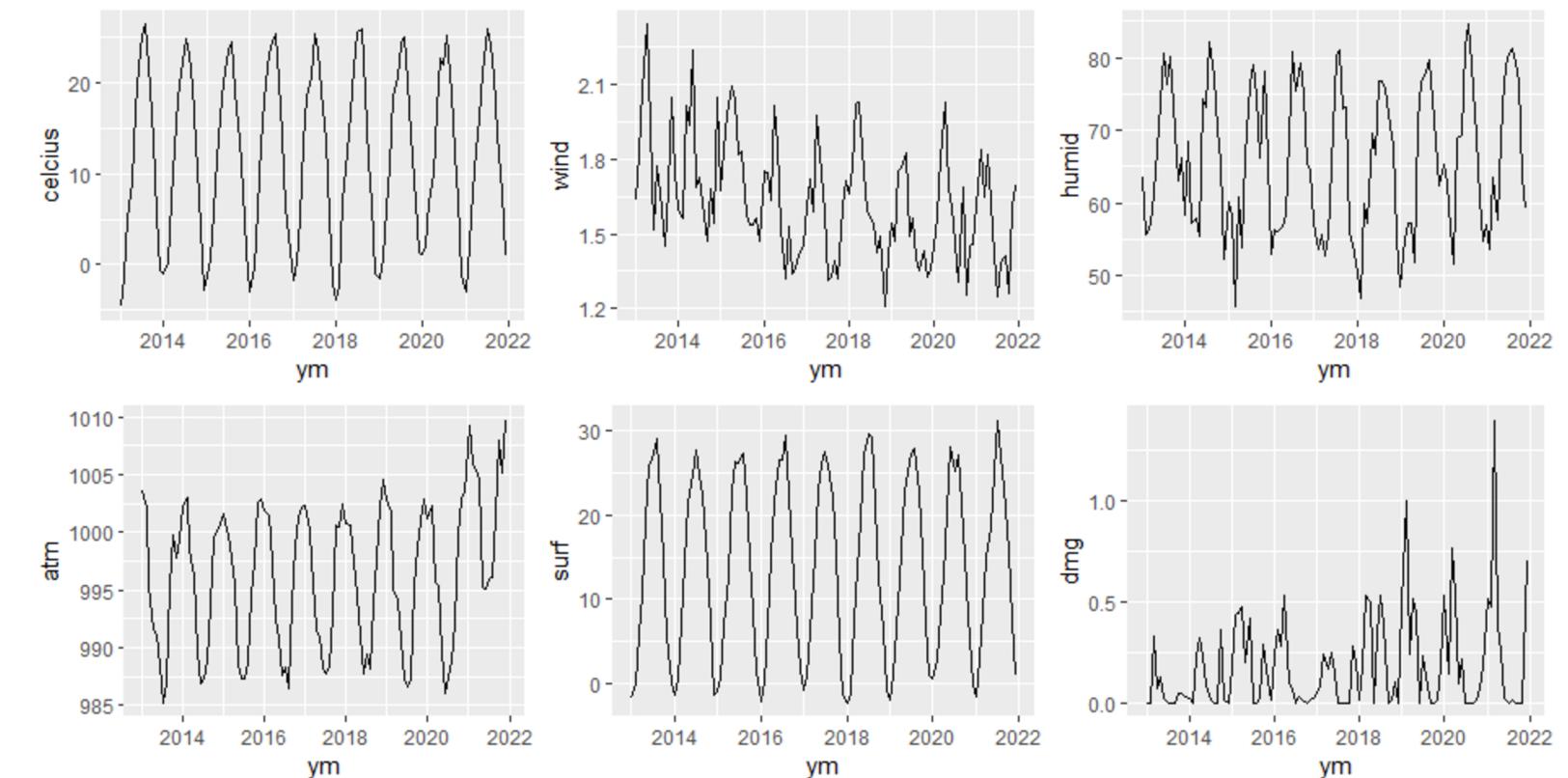
model 3: (T + S)\*danger + err

# 분석 Flow Chart

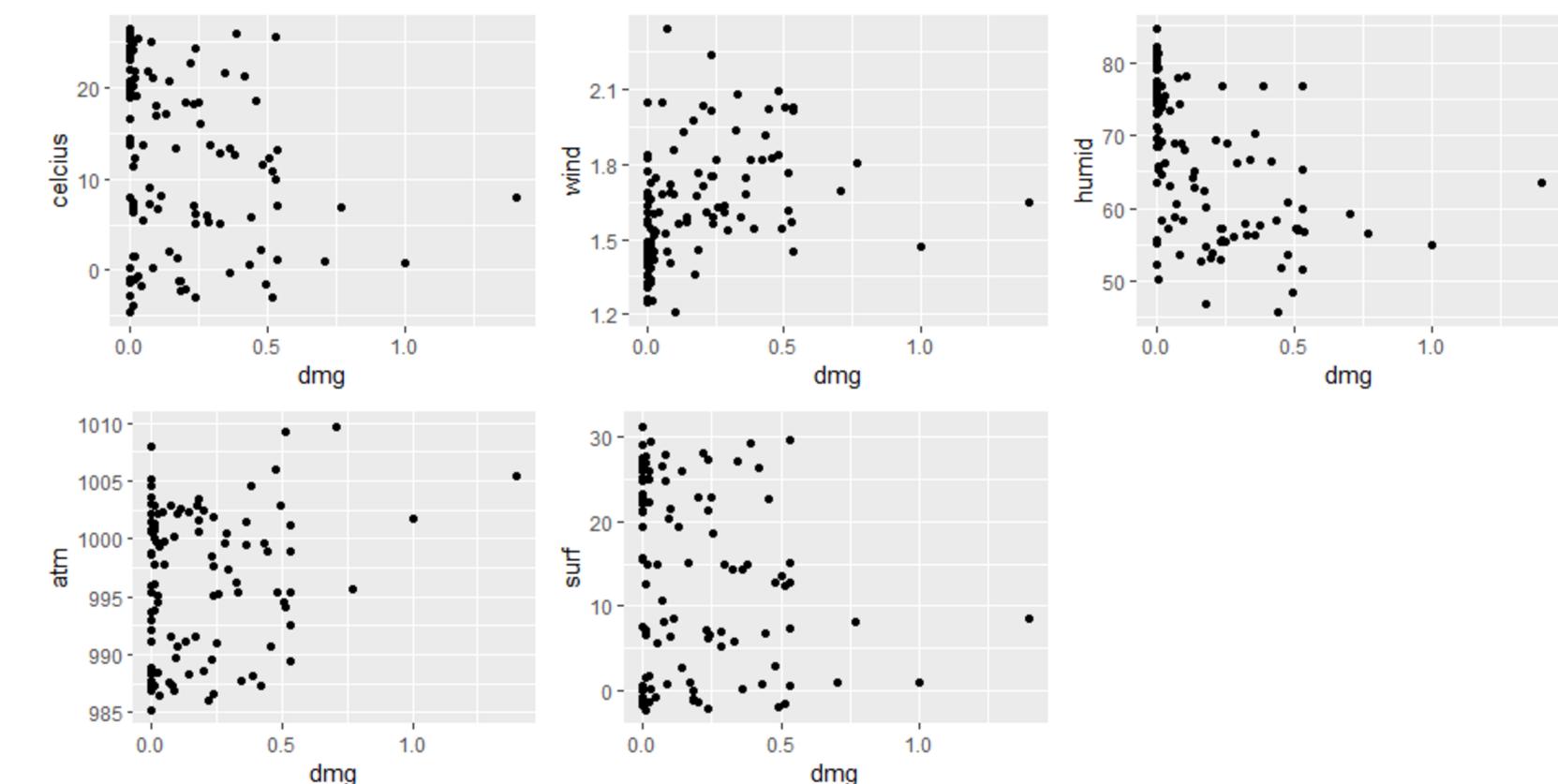


# 기후 요소 EDA

## 전체 연도 월별 피해 및 기상 데이터에 대한 그래프



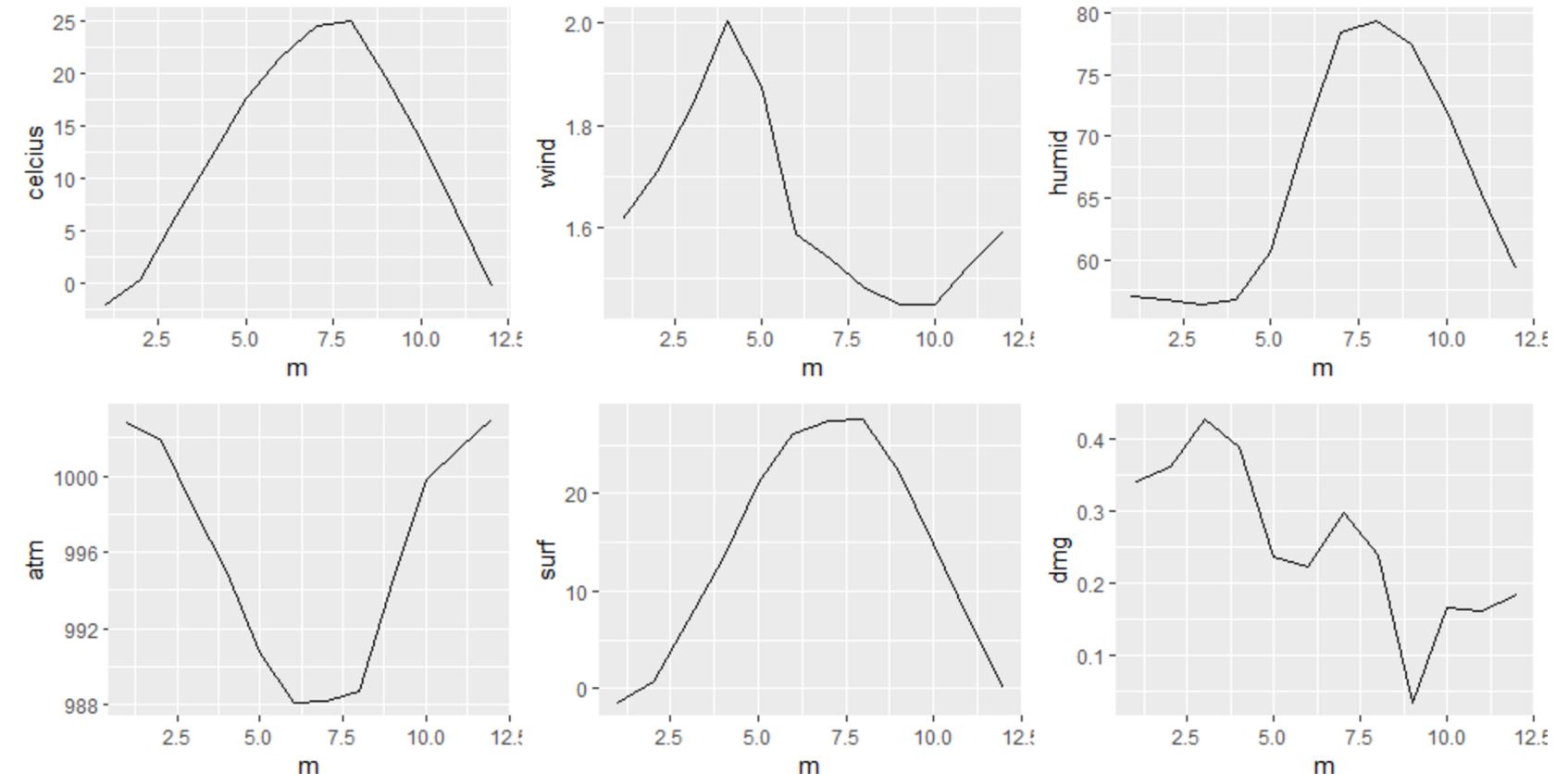
## dmg ~ climate plot by ym



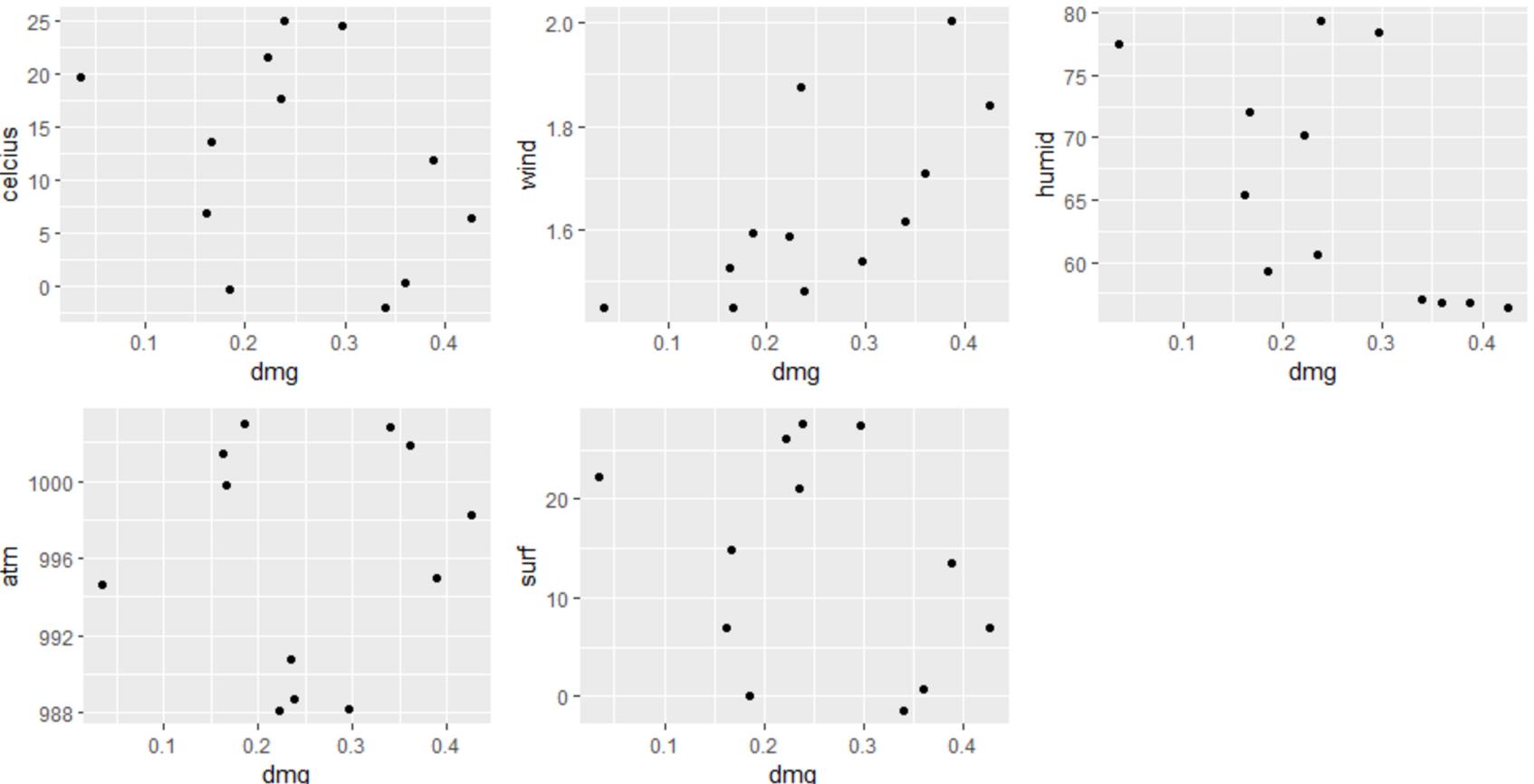
# 기후 요소 EDA

- dmg 데이터에서 500 이하의 값으로 필터링
- 1.4 이상의 값을 1.4로 대체

## 월별 climate



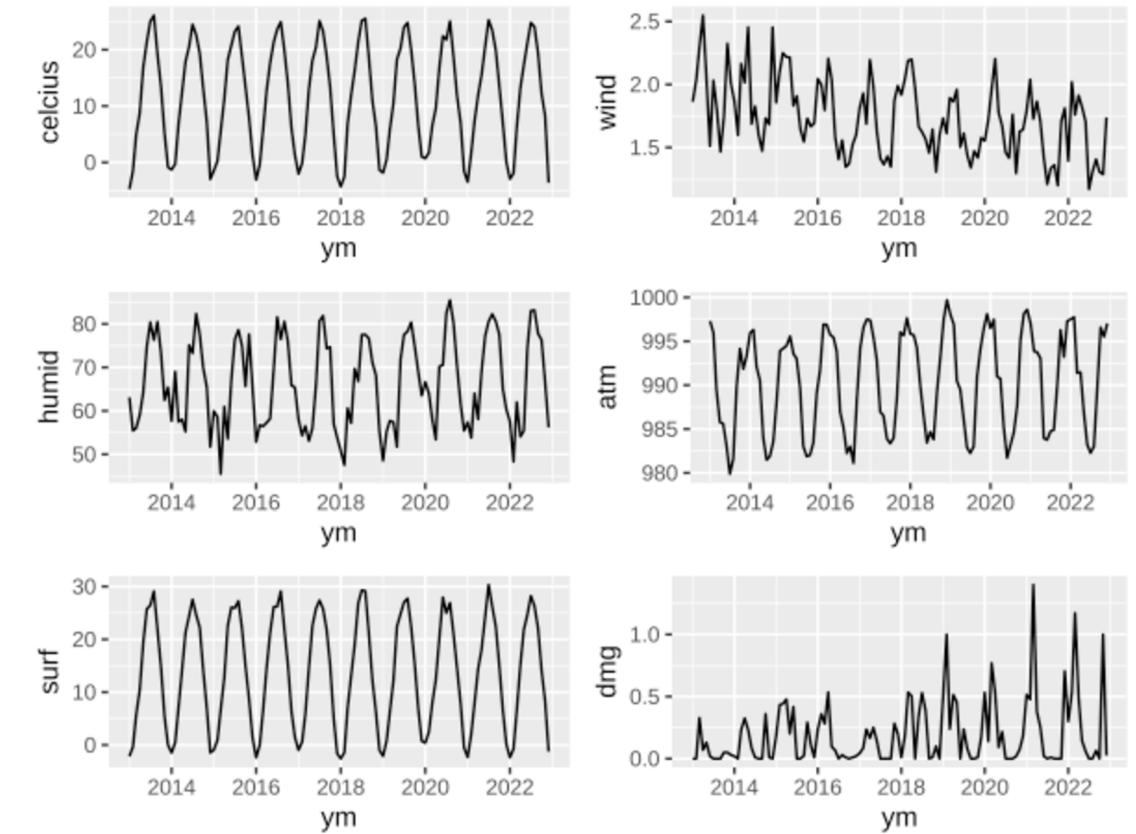
## dmg ~ climate plot



# 피해면적과 기후요소

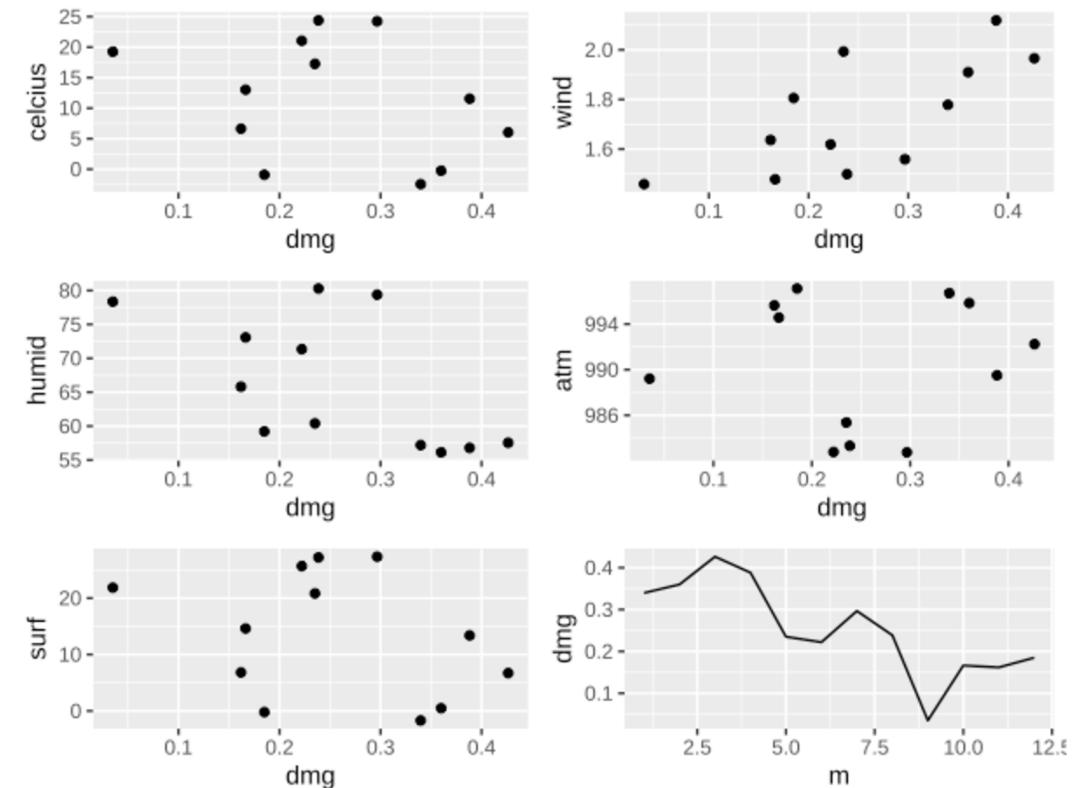
- 습도 : -0.446 / -0.592
- 풍속 : 0.276/ 0.704
- 기온 : -0.267/ -0.328

2013~2022



	celcius	wind	humid	atm	surf	dmg
celcius	1					
wind	-0.366	1				
humid	0.739	-0.709	1			
atm	-0.881	0.05	-0.511	1		
surf	0.996	-0.349	0.712	-0.895	1	
dmg	-0.267	0.276	-0.446	0.191	-0.272	1
> corr.visible(dmg_clim_onlymonth[,2:7])						
	dmg	celcius	wind	humid	atm	surf
dmg	1					
celcius	-0.328	1				
wind	0.704	-0.484	1			
humid	-0.592	0.817	-0.884	1		
atm	0.055	-0.919	0.222	-0.603	1	
surf	-0.323	0.998	-0.465	0.798	-0.934	1

1~12 월별



[ 연도+월 기준 ]

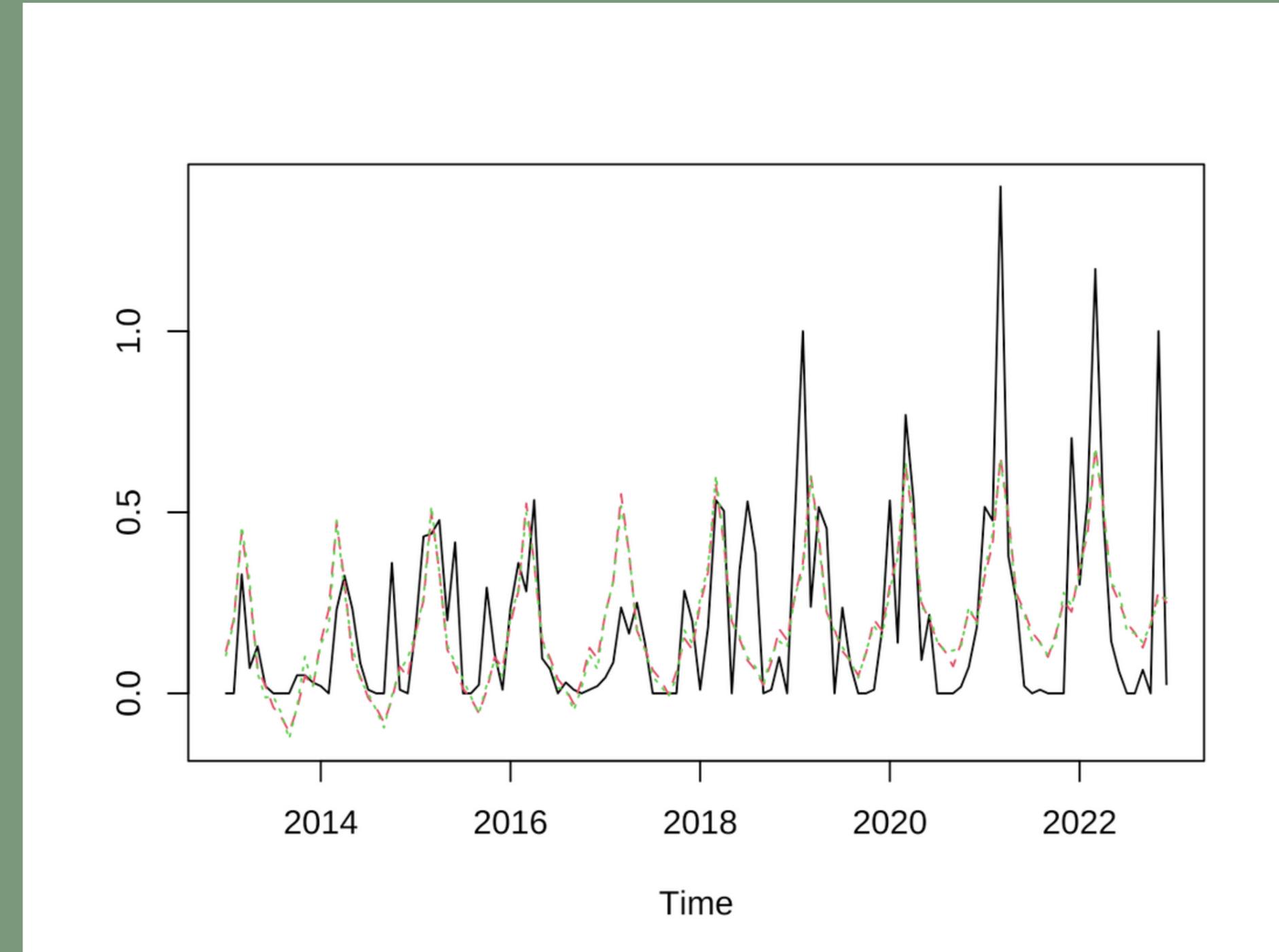
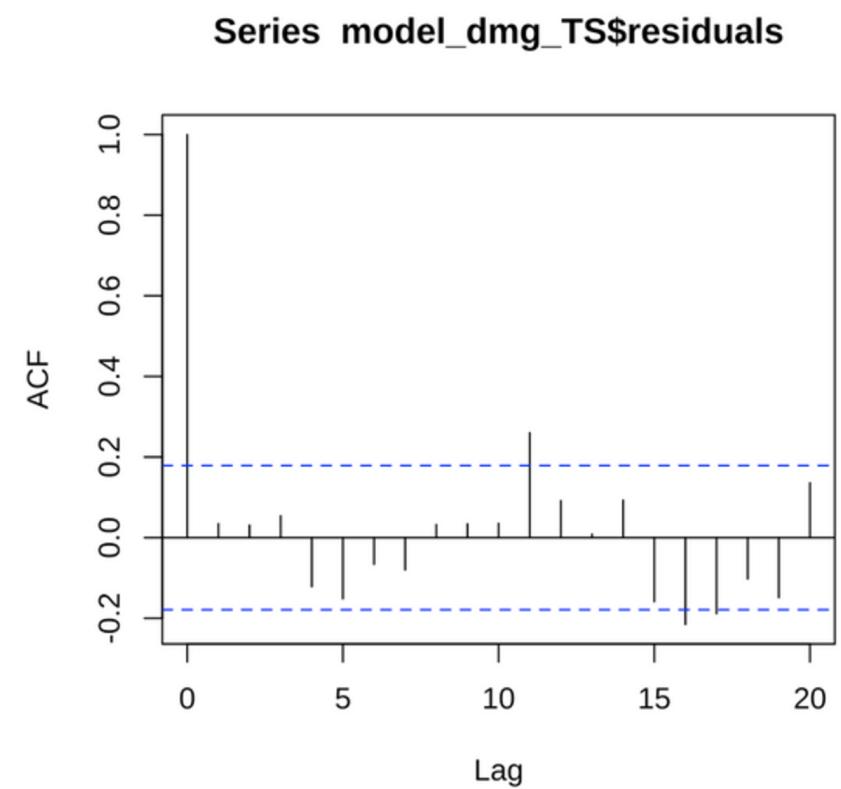
[ 월 기준 ]

## Model 2:

### Trend + Seasonality + 기후요인

$dmg \sim T + S + \text{Wind} + \text{error}$

- 통계량을 보았을때 유의미한 변화가 없는것으로 확인
- 시계열 그림을 보았을때 실제값과 비슷한 경향성이 보임
- 잔차의 ACF 결과, 정상성을 따르는 것을 확인



## Model 2:

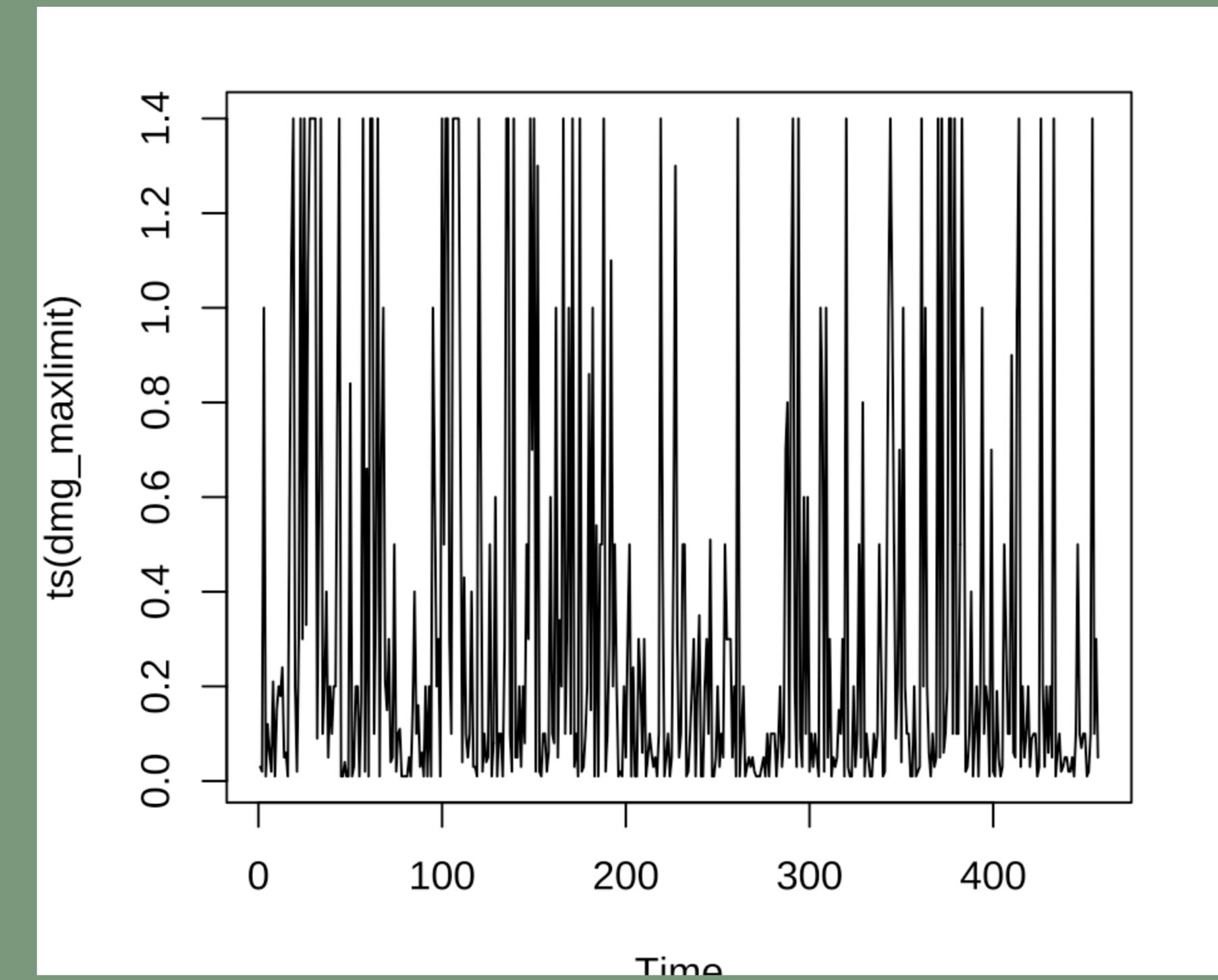
### Trend + Seasonality + 기후요인

$dmg \sim T + S + \text{Wind} + \text{error}$

문제점:

1. 일별기준 0.5이상의 피해량 상위 20%  
월별기준 0.5이상의 피해량은 상위 5%  
---> 피해면적이 **robust**하지 않음
2. 기후 데이터가 연속적인 값일때 적합을  
잘 못시킴

[시간단위 피해량]



## Model 2:

**Trend + Seasonality + 기후요인**

dmg ~ T + S + Wind + error

solution:

1. robust문제 해결위해 index 설정

--->상위 90% 피해량(1.4) 기준

2.기후요인 group으로 나누어 분석

->기후 위험도 파생지수 \*

\*Fisher test/ Kruskal test

3. 월별 데이터로 통일

기후 위험도 지수

$$= I(\text{wind} >= 3.7) + I(\text{humid} <= 29)$$

Def

기후 요소가 피해면적에 악조건인  
기준치를 설정하여 나타낸 지수

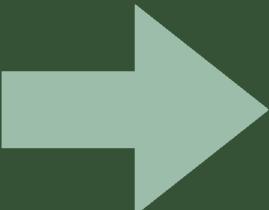
```
quantile(dat_clim$dmg, seq(0, 1, by=0.1))
 0%    10%    20%    30%    40%    50%    60%    70%    80%    90%   100%
0.010  0.010  0.020  0.050  0.064  0.100  0.200  0.300  0.508  1.408 4190.380
```

피해량 1.4 이상 --> 심각  
피해량 1.4 이하 --> 경미

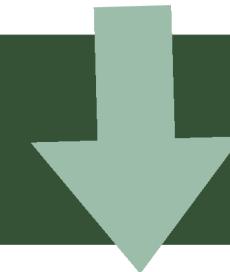
# 기후위험도지수

범주를 나누는 과정에서  
dmg->90백분위수로 설정

기상요인은 확실한 근거가  
필요하다 판단



Fisher test의 p.value로 어떤 분위수를 선택해야 가장  
dmg high/low가 설명 되는지 판단



위 결과를 통해 나온 기상요인을 통해 dmg\_high와  
dmg\_low로 나누고 각 결과의 위험도 분포가 일치하는  
지 다른지는 kruskal.test로 검정

# 기후위험도지수

## Fisher.test

	dmg_little	dmg_lot
wind_little	372	34
wind_lot	34	13

- 피해량 '심각' 47에서 '강풍'은 13, '약풍'은 34
- 피해량 '경미' 406에서 '강풍'은 34 '약풍'은 372

Fisher's Exact Test for Count Data

```
data: contin_wind
p-value = 0.0003201
alternative hypothesis: true odds ratio is not equal to 1
95 percent confidence interval:
1.837642 9.057897
sample estimates:
odds ratio
4.163628
```

- 검정 결과 : 귀무가설 기각
- 바람과 피해량이 관계가 있다

# 기후위험도지수

## Fisher.test

습도

```
[[2]]  
Fisher's Exact Test for Count Data  
  
data: wip_summary_matrix  
p-value = 0.006328  
alternative hypothesis: true odds ratio is not equal to 1  
95 percent confidence interval:  
 0.138369 0.782596  
sample estimates:  
odds ratio  
 0.3176041  
  
[[3]]  
Fisher's Exact Test for Count Data  
  
data: wip_summary_matrix  
p-value = 0.00059  
alternative hypothesis: true odds ratio is not equal to 1  
95 percent confidence interval:  
 0.1530229 0.6297275  
sample estimates:  
odds ratio  
 0.3078114  
  
[[4]]  
Fisher's Exact Test for Count Data  
  
data: wip_summary_matrix  
p-value = 0.0004237  
alternative hypothesis: true odds ratio is not equal to 1  
95 percent confidence interval:  
 0.1591867 0.6246009  
sample estimates:  
odds ratio  
 0.3169388
```

바람

```
[[2]]  
Fisher's Exact Test for Count Data  
  
data: wip_summary_matrix  
p-value = 0.2918  
alternative hypothesis: true odds ratio is not equal to 1  
95 percent confidence interval:  
 0.6157168 22.1648387  
sample estimates:  
odds ratio  
 2.513973  
  
[[3]]  
Fisher's Exact Test for Count Data  
  
data: wip_summary_matrix  
p-value = 0.1657  
alternative hypothesis: true odds ratio is not equal to 1  
95 percent confidence interval:  
 0.7901093 7.0035255  
sample estimates:  
odds ratio  
 2.091684  
  
[[4]]  
Fisher's Exact Test for Count Data  
  
data: wip_summary_matrix  
p-value = 0.004918  
alternative hypothesis: true odds ratio is not equal to 1  
95 percent confidence interval:  
 1.388992 12.106631  
sample estimates:  
odds ratio  
 3.641298
```

온도

```
[[2]]  
Fisher's Exact Test for Count Data  
  
data: wip_summary_matrix  
p-value = 0.7945  
alternative hypothesis: true odds ratio is not equal to 1  
95 percent confidence interval:  
 0.3241079 3.0441012  
sample estimates:  
odds ratio  
 0.8880761  
  
[[3]]  
Fisher's Exact Test for Count Data  
  
data: wip_summary_matrix  
p-value = 1  
alternative hypothesis: true odds ratio is not equal to 1  
95 percent confidence interval:  
 0.4553007 2.4936641  
sample estimates:  
odds ratio  
 1.013947  
  
[[4]]  
Fisher's Exact Test for Count Data  
  
data: wip_summary_matrix  
p-value = 0.7314  
alternative hypothesis: true odds ratio is not equal to 1  
95 percent confidence interval:  
 0.5748461 2.6282692  
sample estimates:  
odds ratio  
 1.192069
```

# 기후위험도지수

## Fisher.test

습도

```
[[2]]  
Fisher's Exact Test for Count Data  
  
data: wip_summary_matrix  
p-value = 0.006328  
alternative hypothesis: true odds ratio is not equal to 1  
95 percent confidence interval:  
0.138369 0.782596  
sample estimates:  
odds ratio  
0.3176041  
  
[[3]]  
Fisher's Exact Test for Count Data  
  
data: wip_summary_matrix  
p-value = 0.00059  
alternative hypothesis: true odds ratio is not equal to 1  
95 percent confidence interval:  
0.1530229 0.6297275  
sample estimates:  
odds ratio  
0.3078114  
  
[[4]]  
Fisher's Exact Test for Count Data
```

바람

```
[[2]]  
Fisher's Exact Test for Count Data  
  
data: wip_summary_matrix  
p-value = 0.2918  
alternative hypothesis: true odds ratio is not equal to 1  
95 percent confidence interval:  
0.138369 0.782596  
sample estimates:  
odds ratio  
0.3176041  
  
[[3]]  
Fisher's Exact Test for Count Data  
  
data: wip_summary_matrix  
p-value = 0.004918  
alternative hypothesis: true odds ratio is not equal to 1  
95 percent confidence interval:  
1.388992 12.106631  
sample estimates:  
odds ratio  
3.641298
```

온도

```
[[2]]  
Fisher's Exact Test for Count Data  
  
data: wip_summary_matrix  
p-value = 0.7945  
alternative hypothesis: true odds ratio is not equal to 1  
95 percent confidence interval:  
0.5748461 2.6282692  
sample estimates:  
odds ratio  
1.192069
```

dmg의 high/low 기준을 1.4로 고정

wind와 humidity의 10~90백분위수로 변경  
-> p.value가 작은 값 발견

wind=3.7, humidity=29로 선정

# 기후위험도지수

## Fisher.test

습도

```
[[2]]  
Fisher's Exact Test for Count Data  
  
data: wip_summary_matrix  
p-value = 0.006328  
alternative hypothesis: true odds ratio is not equal to 1  
95 percent confidence interval:  
0.138369 0.782596  
sample estimates:  
odds ratio  
0.3176041  
  
[[3]]  
Fisher's Exact Test for Count Data  
  
data: wip_summary_matrix  
p-value = 0.00059  
alternative hypothesis: true odds ratio is not equal to 1  
95 percent confidence interval:  
0.1530229 0.6297275  
sample estimates:  
odds ratio  
0.3078114  
  
[[4]]  
Fisher's Exact Test for Count Data
```

바람

```
[[2]]  
Fisher's Exact Test for Count Data  
  
data: wip_summary_matrix  
p-value = 0.2918  
alternative hypothesis: true odds ratio is not equal to 1  
95 percent confidence interval:  
0.138369 0.782596  
sample estimates:  
odds ratio  
0.3176041  
  
[[3]]  
Fisher's Exact Test for Count Data  
  
data: wip_summary_matrix  
p-value = 0.00059  
alternative hypothesis: true odds ratio is not equal to 1  
95 percent confidence interval:  
0.1530229 0.6297275  
sample estimates:  
odds ratio  
0.3078114  
  
[[4]]  
Fisher's Exact Test for Count Data
```

온도

```
[[2]]  
Fisher's Exact Test for Count Data  
  
data: wip_summary_matrix  
p-value = 0.7945  
alternative hypothesis: true odds ratio is not equal to 1  
95 percent confidence interval:  
0.5748461 2.6282692  
sample estimates:  
odds ratio  
1.192069  
  
[[3]]  
Fisher's Exact Test for Count Data  
  
data: wip_summary_matrix  
p-value = 0.0004237  
alternative hypothesis: true odds ratio is not equal to 1  
95 percent confidence interval:  
0.1591867 0.6246009  
sample estimates:  
odds ratio  
0.3169388  
  
[[4]]  
Fisher's Exact Test for Count Data
```

wind 3.7이상은 강풍(1), 미만은 약풍(0)

humidity 29이하는 건조(1), 미만은 습함(0)

->기후 위험도 = I(wind>=3.7) + I(humid<=29)로 설정

# 기후위험도지수

## Kruskal test

1

	date <dttm>	dmg <dbl>	cel <dbl>	wind <dbl>	humid <dbl>	atm <dbl>	surf <dbl>	dmg_ind <dbl>	ind <dbl>
1	2022-12-25 21:07:00	0.03	0.83	3.53	42.1	1018.	-0.735	0	0
2	2022-12-05 13:32:00	0.02	5.71	1.51	27.5	1022.	9.35	0	1
3	2022-11-25 23:43:00	1	16.0	7.44	44.2	1005.	8.91	0	1
4	2022-09-24 15:23:00	0.01	21.8	1.99	40.2	999.	31.6	0	0
5	2022-09-22 14:48:00	0.12	23.2	0.96	44.8	998.	37.5	0	0
6	2022-06-04 11:16:00	0.06	23.3	2.47	63	1010.	42.0	0	0
7	2022-05-26 08:40:00	0.02	21.4	2.57	40.3	1001	27.8	0	0
8	2022-05-25 15:00:00	0.21	26.6	2.3	35	983.	41	0	0
9	2022-05-21 17:18:00	0.01	24.4	2.83	45.3	989.	29.6	0	0
10	2022-05-18 19:41:00	0.15	20.4	2.14	53.8	996.	19.2	0	0

	ind <dbl>	dmg_ind <dbl>	n <int>
1	0	0	178
2	0	1	9
3	1	0	187
4	1	1	24
5	2	0	46
6	2	1	13

강풍/약풍, 습함/건조함 기준으로  
강풍/건조를 위험도+1로 설정

앞서 정한 기상요인을 **dmg\_low, dmg\_high**로  
나누어 **dmg\_group**으로 만들어 검정 실시

2

Kruskal-Wallis rank sum test

```
data: value by group
Kruskal-Wallis chi-squared = 14.475, df = 1, p-value = 0.000142
```

# 기후위험도지수

## Kruskal test

1

	date	dmg	cel	wind	humid	atm	surf	dmg_ind	ind
	<dttm>	<dbl>	<dbl>	<dbl>	<dbl>	<dbl>	<dbl>	<dbl>	<dbl>
1	2022-12-25 21:07:00	0.03	0.83	3.53	42.1	1018.	-0.735	0	0
2	2022-12-05 13:32:00	0.02	5.71	1.51	27.5	1022.	9.35	0	1
3	2022-11-25 23:43:00	1	16.0	7.44	44.2	1005.	8.91	0	1
4	2022-09-24 15:23:00	0.01	21.8	1.99	40.2	999.	31.6	0	0
5	2022-09-22 14:48:00	0.12	23.2	0.96	44.8	998	37.5	0	0
6	2022-06-04								
7	2022-05-26								
8	2022-05-25								
9	2022-05-21								
10	2022-05-18								
	# i 447 more n								

ind	dmg_ind	n
<dbl>	<dbl>	<int>
1	0	0
2	0	1
3	1	9
4	0	177

강풍/약풍, 습함/건조함 기준으로  
가포/거조로 의하도나로 설정

low 와 high는 명확히 다르다

앞서 정한 기상요인을 dmg\_low,dmg\_high로  
나누어 dmg\_group으로 만들어 검정 실시

2

Kruskal-Wallis rank sum test

```
data: value by group
Kruskal-Wallis chi-squared = 14.475, df = 1, p-value = 0.000142
```

# 기후위험도지수

## Kruskal test

```
group_A <- c(rep(0, 178), rep(1, 9))  
group_B <- c(rep(0, 184), rep(1, 24))  
group_C <- c(rep(0, 45), rep(1, 13))
```



각 그룹의 범위를 다르게  
설정하여 검정 진행

```
Kruskal-Wallis rank sum test
```

```
data: Value by Group  
Kruskal-Wallis chi-squared = 15.805, df = 2, p-value = 0.0003699
```

1

value별로 group 분포

```
Kruskal-Wallis rank sum test
```

```
data: value by group  
Kruskal-Wallis chi-squared = 14.475, df = 1, p-value = 0.000142
```

2

group별로 value 분포

```
Kruskal-Wallis rank sum test
```

```
data: Value by Group  
Kruskal-Wallis chi-squared = 15.805, df = 2, p-value = 0.0003699
```

# 기후위험도지수

## Kruskal test

```
group_A <- c(rep(0, 178), rep(1, 9))  
group_B <- c(rep(0, 184), rep(1, 24))  
group_C <- c(rep(0, 45), rep(1, 13))
```

각 그룹의 범위를 다르게

```
data: Value by Group  
Kruskal-Wallis chi-squared = 15.805, df = 2, p-value = 0.0003699
```

=> value별로 group 분포가 같다는 귀무가설을 기각

=> group별로 value 분포가 같다는 귀무가설을 기각

즉, 기후 위험지표로 피해량의 심각도를 구분하는게 유효하다.

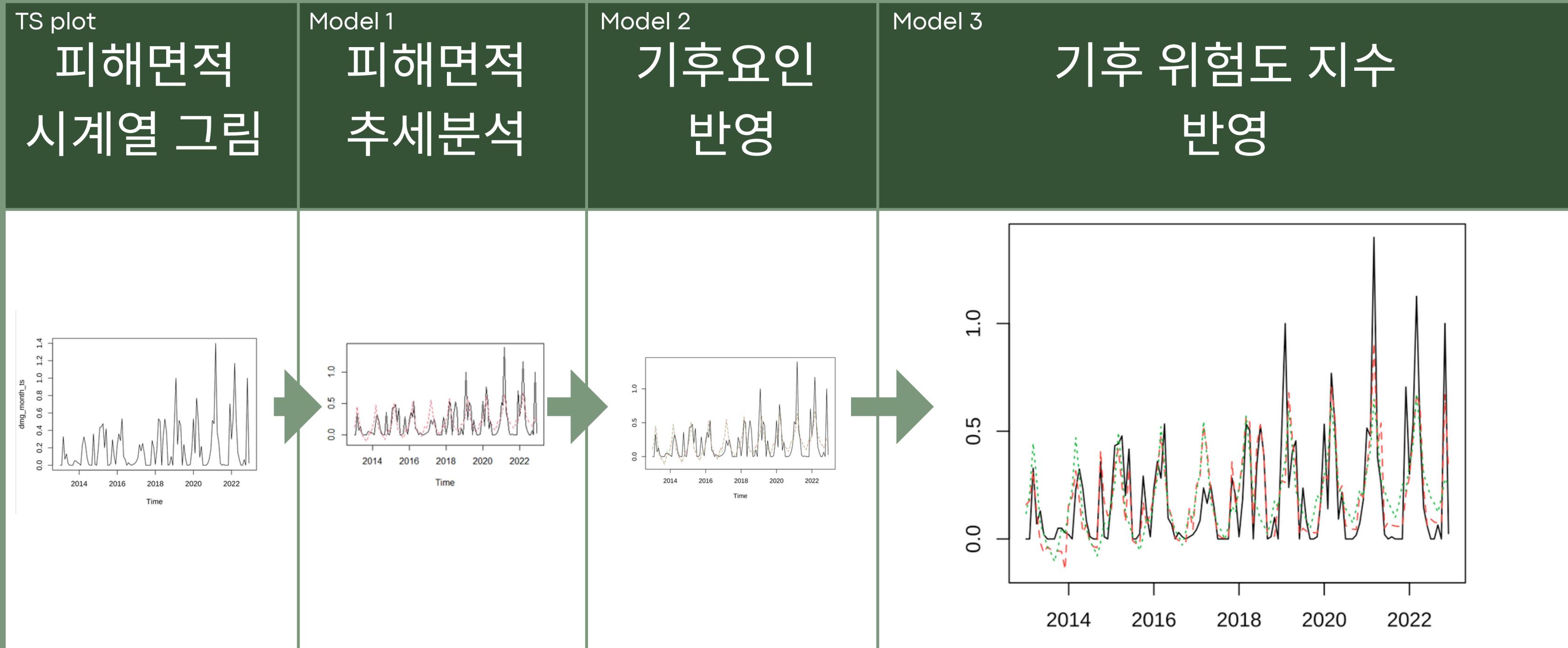
1

```
value  
data:  
Kruska
```

2

```
group  
Kruskal-Wallis rank sum test  
data: Value by Group  
Kruskal-Wallis chi-squared = 15.805, df = 2, p-value = 0.0003699
```

# 분석 Flow Chart



# Model 3: 피해량 index & 기후 위험도 지수

fit1:  $dmg\_mean.ts \sim t + err$

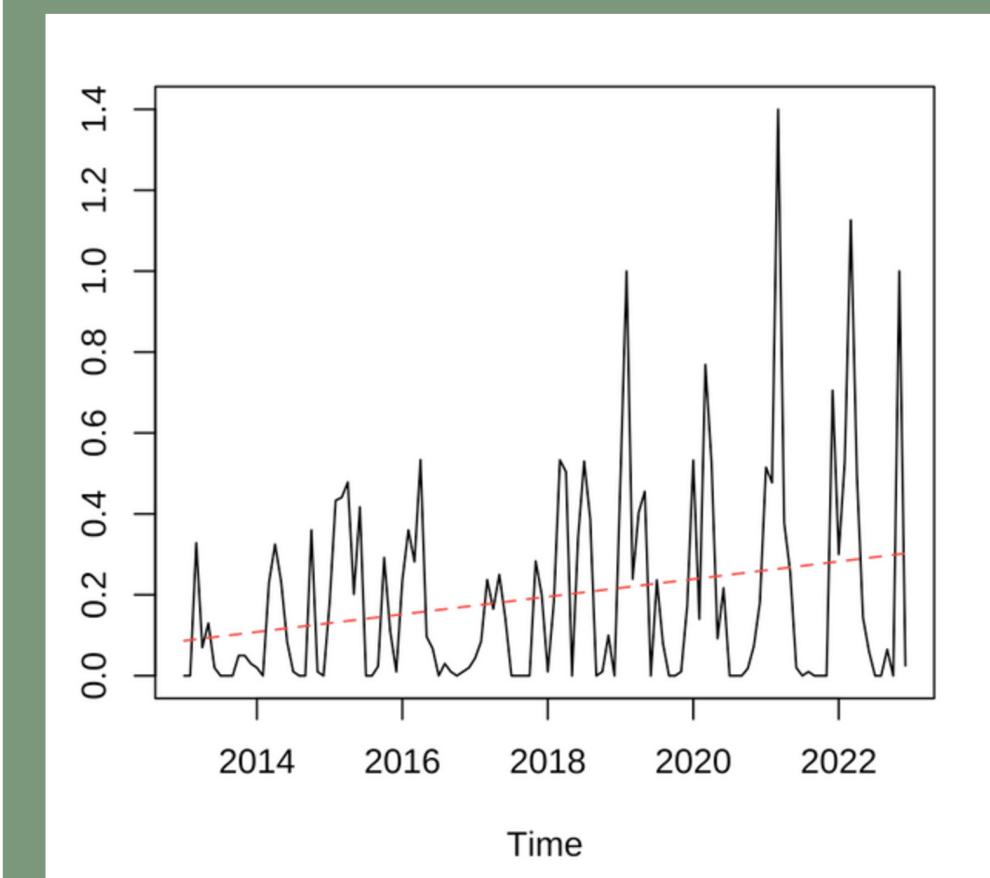
- 단순 추세 분석 진행
- 피해지수 (dmg\_ind)  
1.4 이상 = 심각/ 1.4 이하 = 경미
- 피해면적 월평균 전처리 진행  
피해 면적의 상한 1.4f로 처리 후 => 월 평균
- 잔차의 ACF를 통해 모형진단을 했을때 더 적합한 모형이 필요해보임

\*

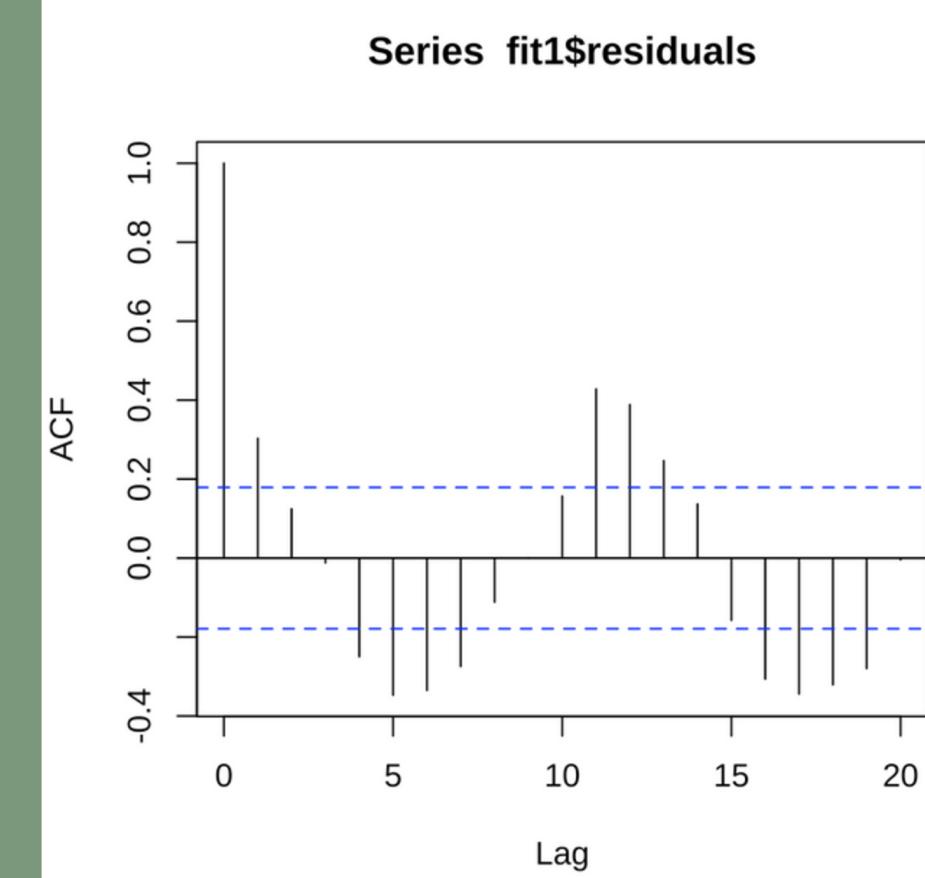
[index 반영 데이터 ]

dmg	cel	wind	humid	atm	surf	dmg_ind	ind
0.03	0.8300000	3.5300000	42.116667	1017.5467	-0.73500000	0	0
0.02	5.7066667	1.5133333	27.533333	1022.0200	9.34666667	0	1
1.00	15.9716667	7.4400000	44.150000	1005.0717	8.91500000	0	1
0.01	21.7916667	1.9933333	40.233333	998.5617	31.55666667	0	0

[fit1 ts plot]



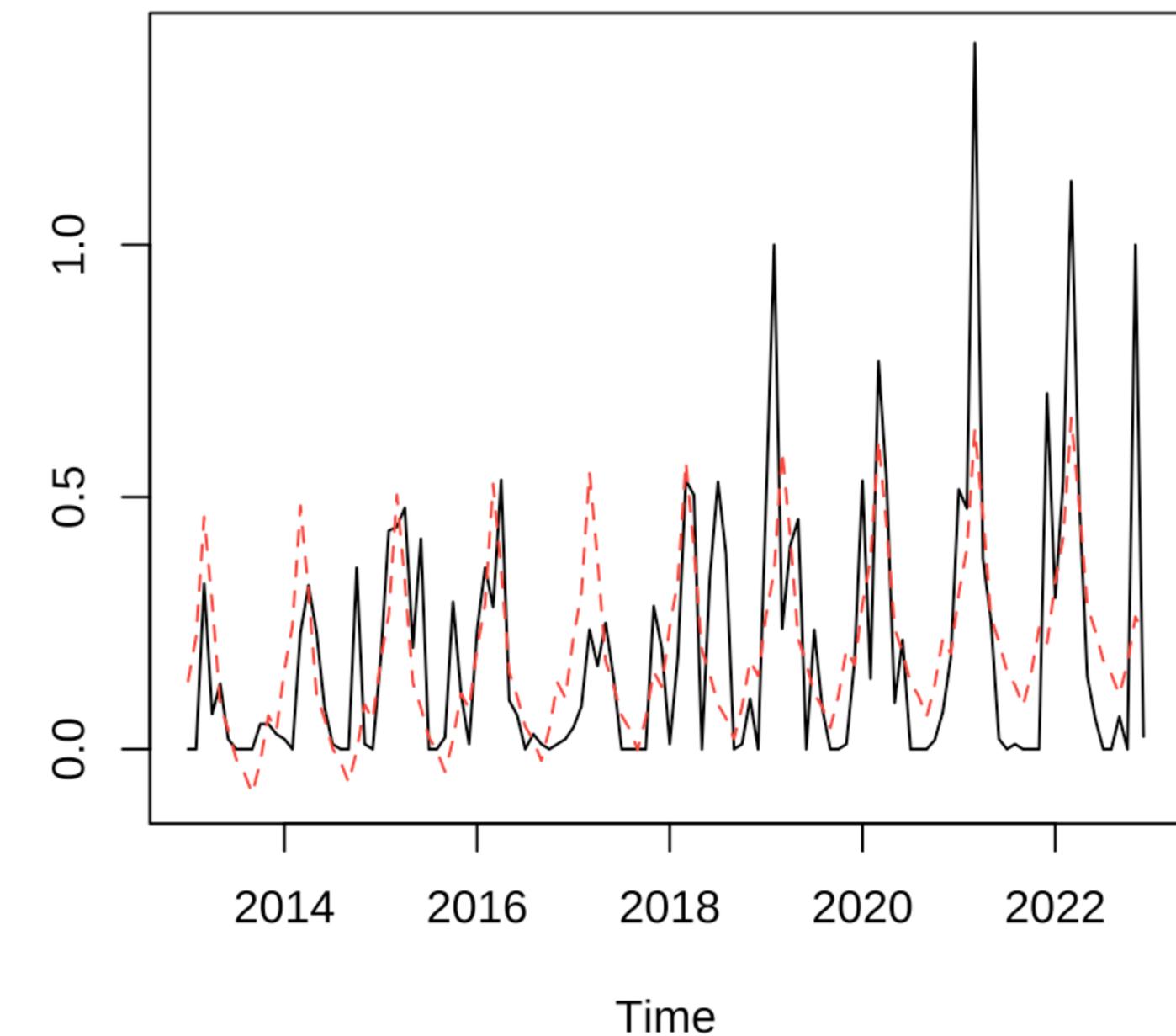
[residual의 ACF ]



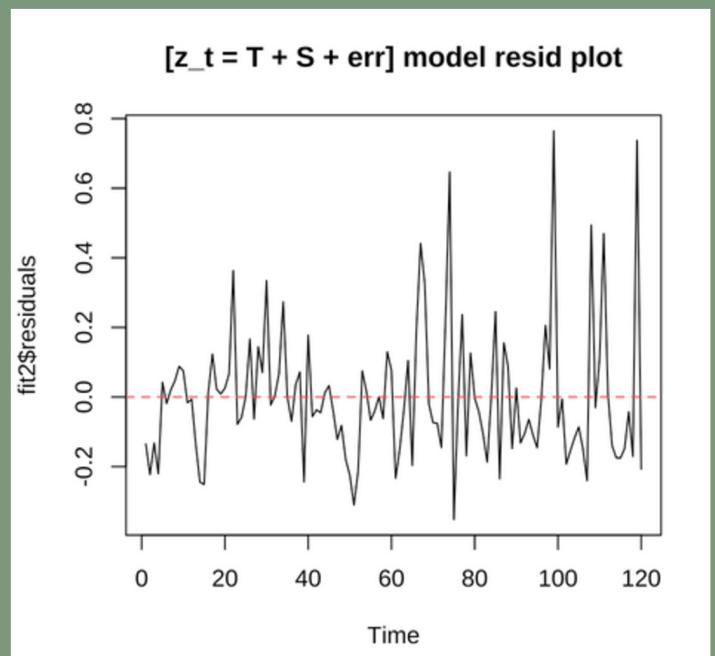
# Model 3: 피해량 index & 기후 위험도 지수

fit2: `dmg_mean.ts ~T + S + err`

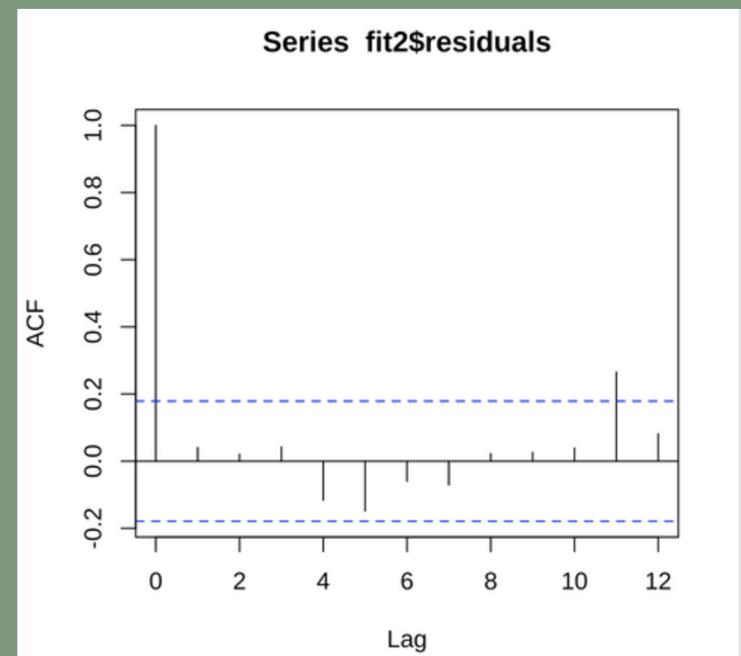
- Trend + Seasonality 모델링
- fit1 보다 더 안정된 stationary data로 보임
- residual plot도 모형이 적합함을 보여줌
- `mse = 4.583347`



[residual plot ]



[residual의 ACF ]



# Model 3: Trend + Seasonality + 기후 위험도

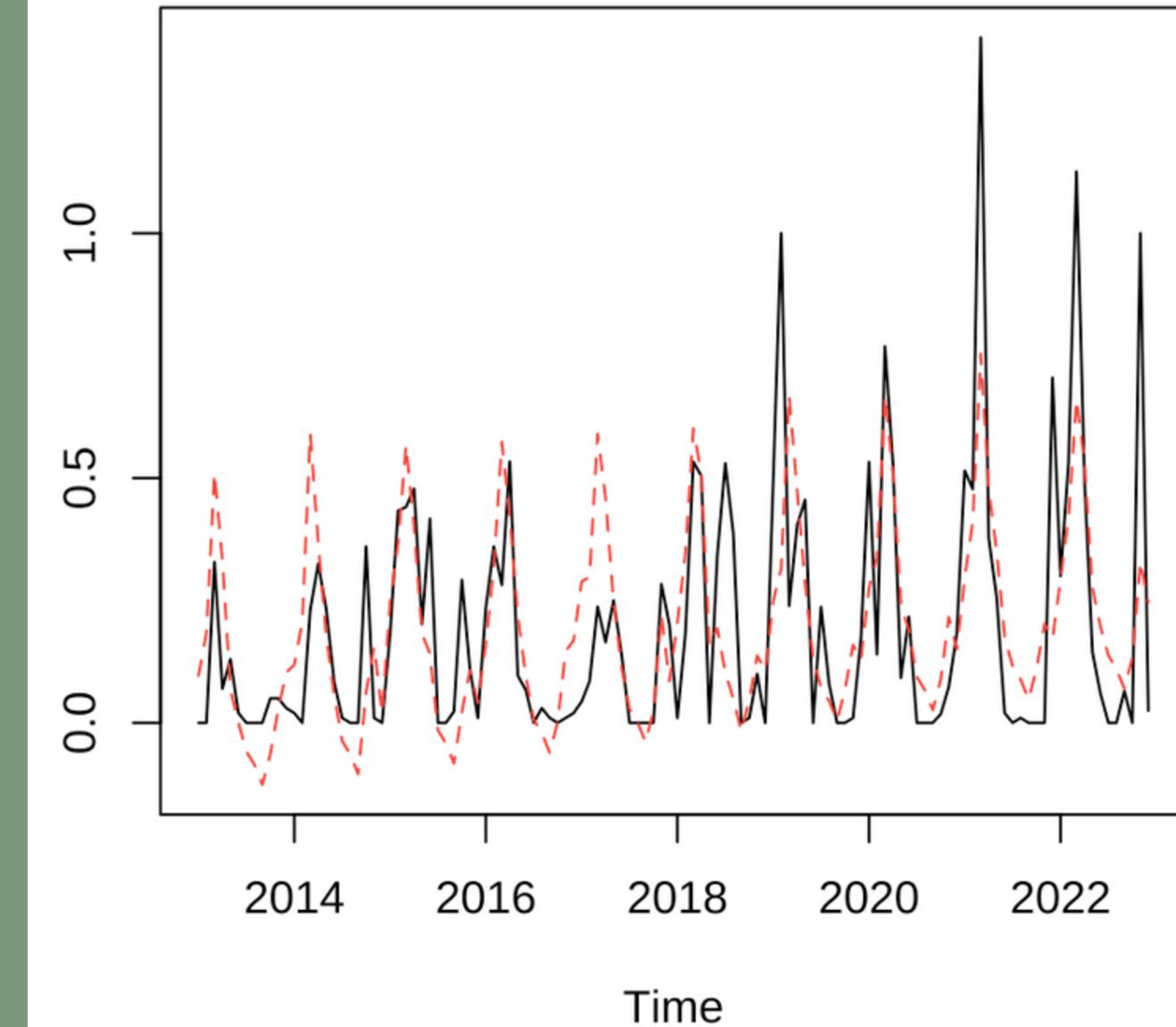
\*기후 위험도(CI\_index) =  $I(\text{wind} \geq 3.7) + I(\text{humid} \leq 29)$

**fit3:  $\text{dmg} \sim T + \text{CI\_index} + \text{err}$**

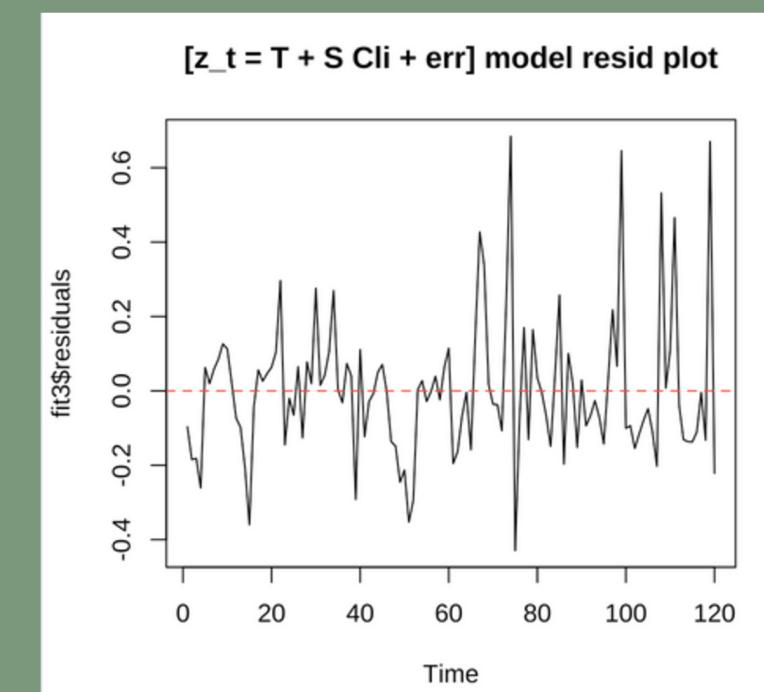
- Trend + 기후 위험도 지수 모델링
- residual ACF 결과, 안정된 stationary data로 보임
- residual plot도 모형이 적합함을 보여줌
- mse가 fit2 보다 감소

```
> mse_model2 ; mse_model3  
[1] 4.583347  
[1] 4.321959
```

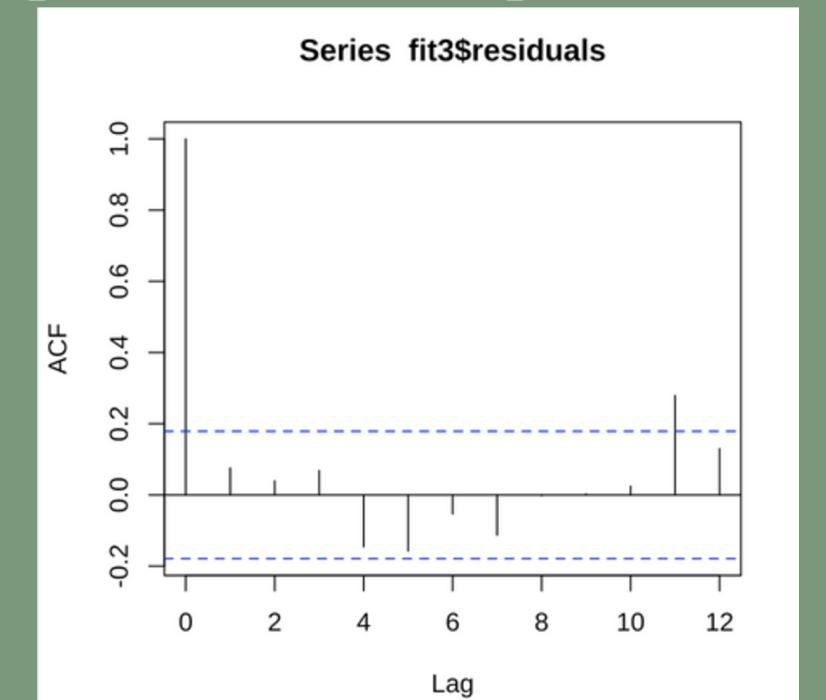
\*



[residual plot ]



[residual의 ACF ]



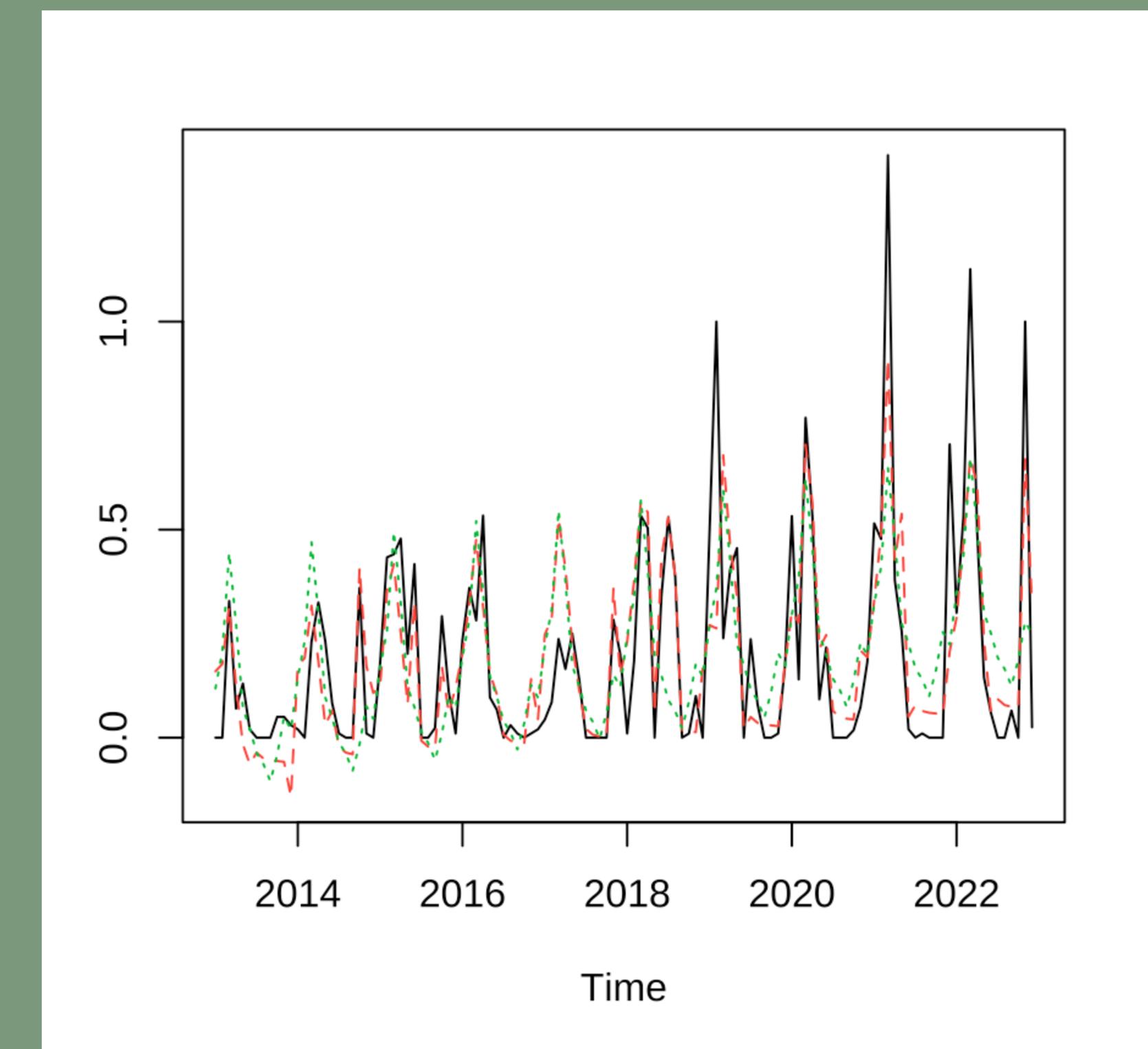
# Model 3: Trend + Seasonality + 기후 위험도

\*기후 위험도(Cl\_index) =  $I(\text{wind} >= 3.7) + I(\text{humid} <= 29)$

final:  $\text{dmg} \sim (\text{T} + \text{S})\text{Cl\_index} + \text{error}$

- 기후 위험도는 0~2 사이 값 --> coeff로 사용되기 적합  
--> (T+S) \* 기후 위험도 모델링 시도
- TS plot을 확인했을 때 가장 잘 적합하는 것을 확인  
--> final 모델인 빨간선이 실제 값에 가장 가까움

[final ts plot]



# Model 3: Trend + Seasonality + 기후 위험도

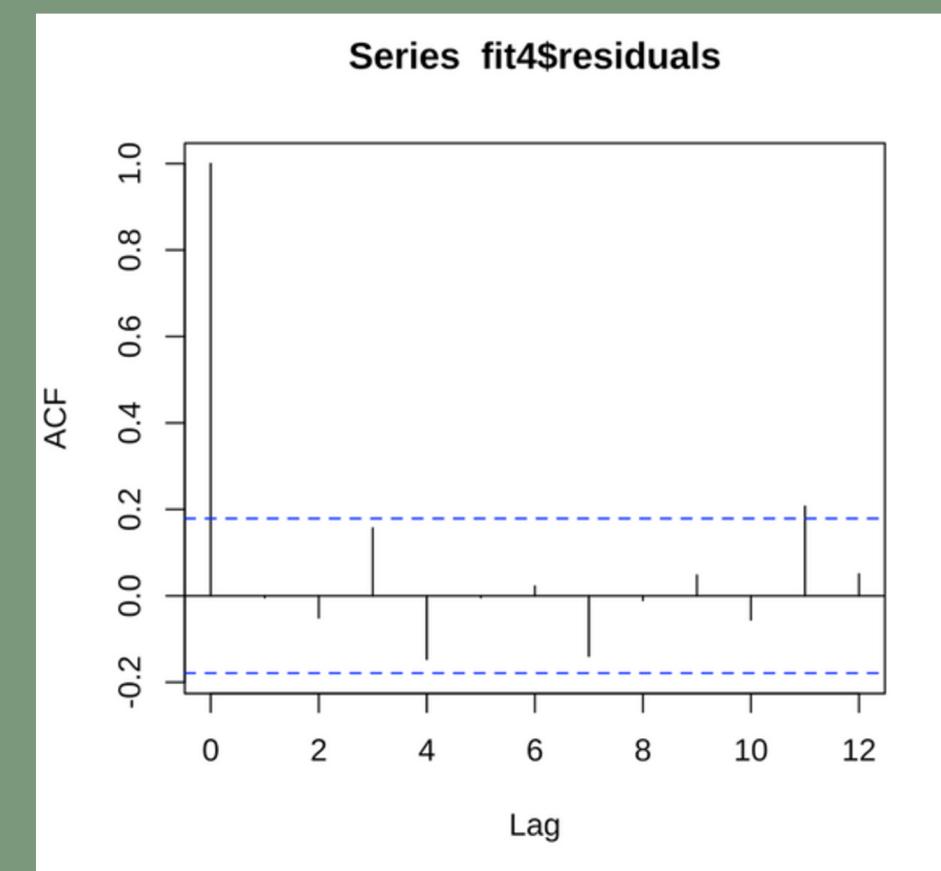
\*기후 위험도(CL\_index) = I(wind>=3.7) + I(humid<=29)

final:  $dmg \sim (T + S)CL\_index + error$

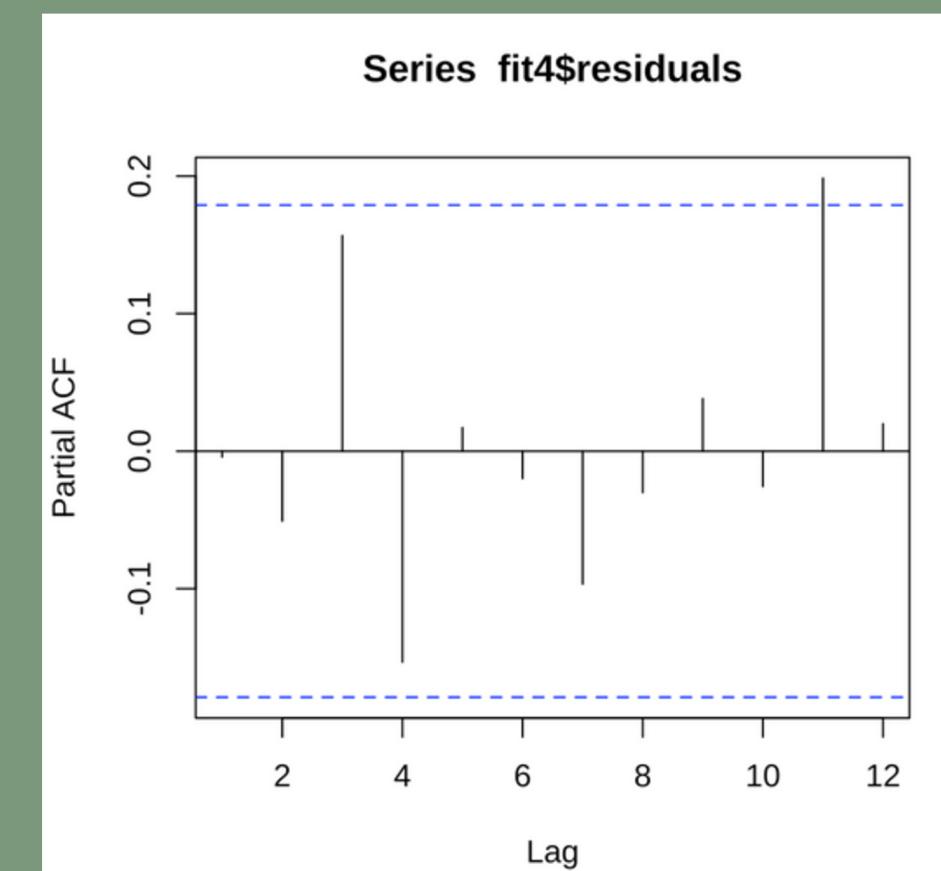
- residual ACF, PACF 결과,  
안정된 stationary data로 보임
- residual plot도 모형이 적합함을 보여줌
- mse가 2.95로 전 모델에 비해 상대적으로 작음

```
> sum(resid(fit4)^2)
[1] 2.949547
> sum(resid(fit_nocli)^2)
[1] 4.569878
> mse_model2
[1] 4.583347
```

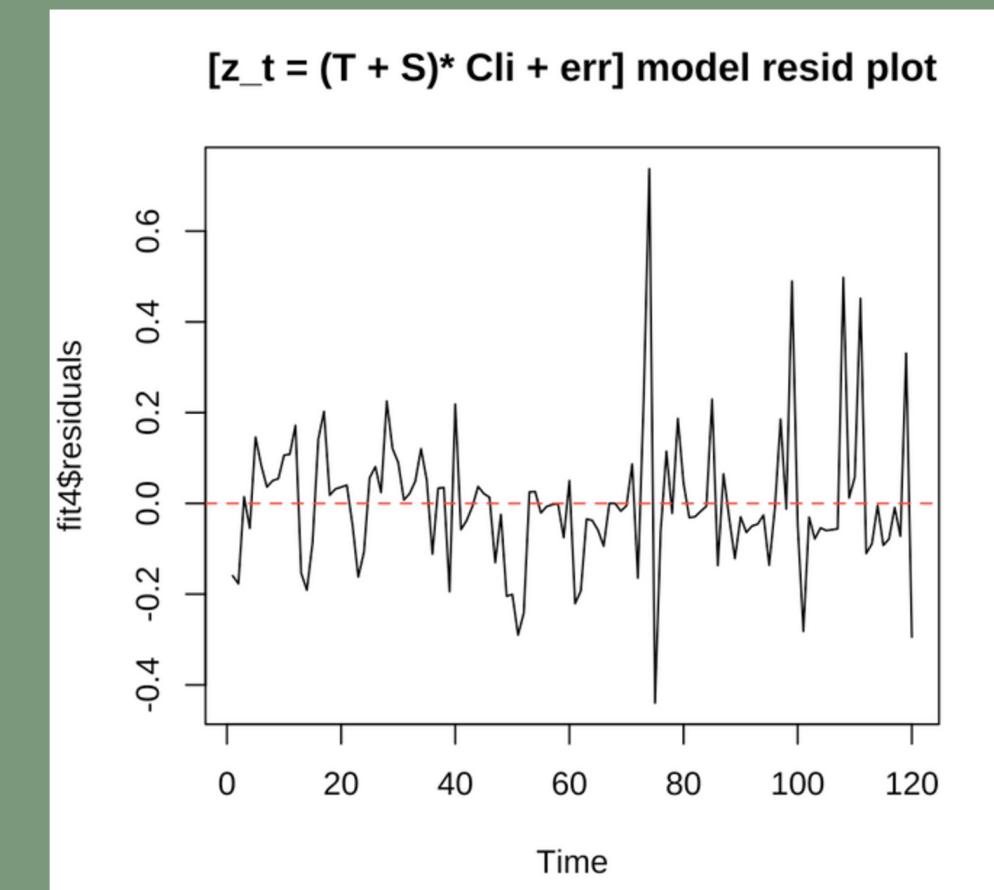
[residual의 ACF ]



[residual의 PACF ]



[residual plot ]



# 결론

## 1. 산불 발생 피해면적 모델의 선택

기후 위험도 지수를 도입한

$$\text{dmg} \sim (T + S)\text{CI\_index} + \text{error}$$

모델이 가장 피해면적 시계열 데이터  
를 잘 적합

## 2. 추후 연구 사항

오른쪽 논문 참고,

시기별 지역별 산불 발생 위도가

점점 올라가는 것을 확인

-->피해면적이 가장 클 것 같은

위도 모델링

## 3. 적용

피해면적이 클 것 같은 시기를 추정할

수 있어 산불 피해 대비에 활용될 가능성

## 한국의 지역별 산불 발생 원인 특성 분석

Regional Characteristics of Forest Fire Occurrences in Korea  
from 1990 to 2018

배미란\* · 채희문\*\*

Bae, Miran\*, and Chae, Heemun\*\*

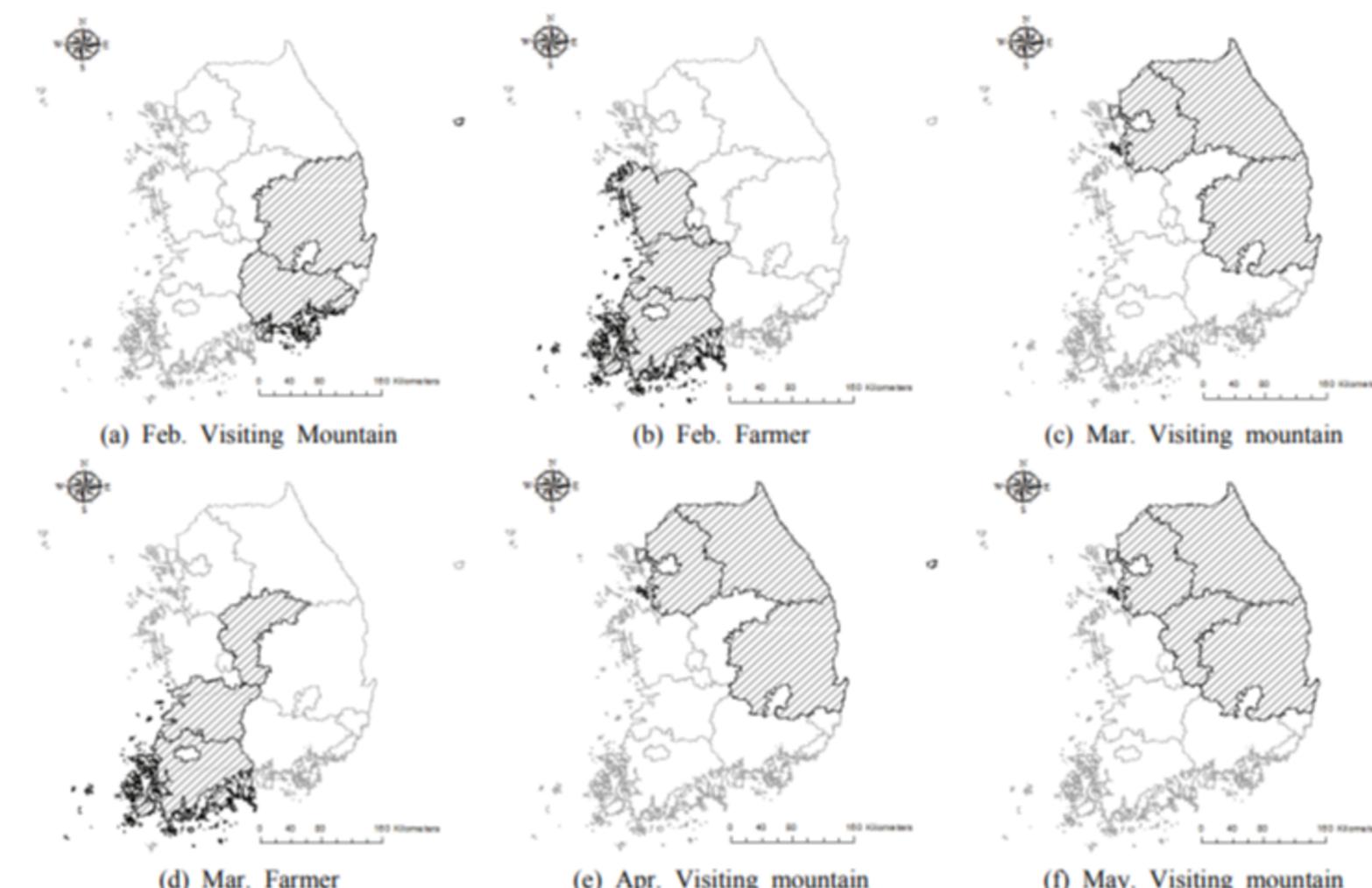


Fig. 3. The Regional Distribution of the Most Occurred of Forest Fire Occurrence Causes During Spring Fire Season

Table 4. Regional Differences of Most Causes of Forest Fires Occurrence During the Spring Forest Fires (February to May)

month & causes rank	February		Mararch		April	May.
	Visiting mountain	Farmer	Visiting mountain	Farmer	Visiting mountain	Visiting mountain
1	G.N.	J.N.	G.B.	J.N.	G.B.	K.W.
2	G.B.	J.B.	G.G.	J.B.	K.W.	G.B.
3	B.S.	C.N.	K.W.	C.B.	G.G.	G.G., C.B.

\* K.W.: Kangwon-do, G.G.: Gyeonggi-do, G.N.: Gyeongsangnam-do, G.B.: Gyeongsangbuk-do, B.S.: Busan, J.N.: Jeollanam-do, J.B.: Jeollabuk-do, C.N.: Chungcheongnam-do, C.B.: Chungcheongbuk-do

# Q & A

