

강릉 월별 강수량 시계열 모형과 조건부 저강수 월 탐지

Taenyong Lee

2025년 11월 24일

발표 개요

- 1 문제 설정과 자료
- 2 기본 SARIMA 모형
- 3 시계열 이상치 탐지
- 4 조건부 저장수 월 정의
- 5 계절더미 + ARMA 모형
- 6 논의 및 향후 과제

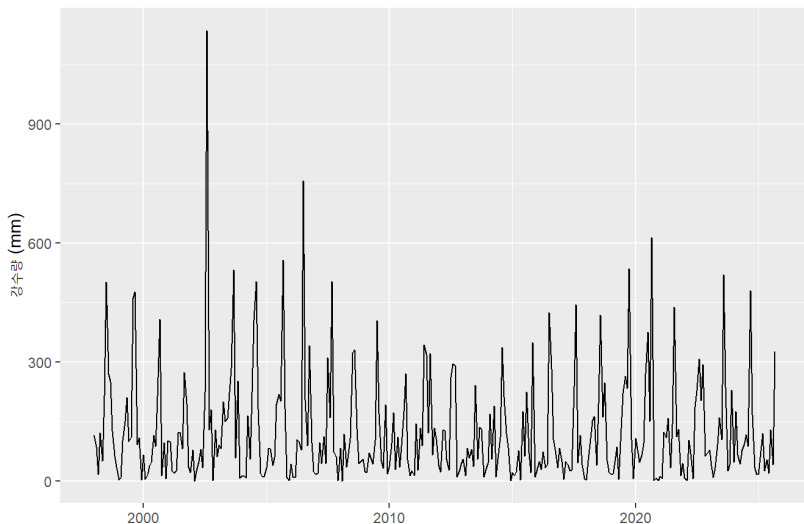
연구 질문과 아이디어

- 자료: 강릉 월별 강수량 (1998-01 ~ 2025-09, monthly)
- 목표
 - 계절성을 반영한 시계열 모형(SARIMA, ARMA + 계절더미) 적합
 - 시계열적인 이상치(outlier) 자동 탐지
 - 모형이 기대하는 값보다 유의하게 적게 내린 달을 조건부 저강수 월로 정의
- 오늘 발표에서 할 것
 - 1 기본 SARIMA 모형과 잔차 진단
 - 2 ARMA + 계절더미 모형과 잔차 진단
 - 3 tsoutliers를 이용한 이상치 탐지 결과
 - 4 SARIMA/ARMA+계절더미 기반 조건부 저강수 월 정의
 - 5 현재까지의 한계와 앞으로의 방향

자료 요약 및 원시계열

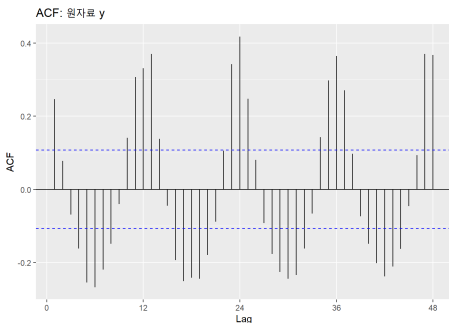
- 시작: 1998-01, 종료: 2025-09, 주기: 12 (월별 시계열)

강릉 월별 강수량(원자료)

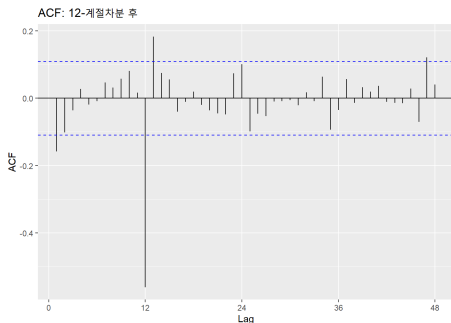


정상성 점검과 모형 후보 탐색

- 원자료 ACF/PACF에서 **강한 12개월 주기성** 확인
- 12-계절차분 $y_t^* = (1 - B^{12})y_t$ 적용 후 ACF/PACF로 계절 MA(1) 구조가 적당해 보임



ACF / PACF (원자료)



ACF / PACF (12-계절차분 후)

- AIC 그리드 탐색:

$p = 0 \dots 3, q = 0 \dots 3, P = 0 \dots 2, Q = 0 \dots 2, D = 1$ 에서

최종 SARIMA 모형과 식

선택된 기본 모형

$ARIMA(2, 0, 2)(0, 1, 1)_{12}$, (절편 없이)

tsoutliers를 이용한 이상치 탐지

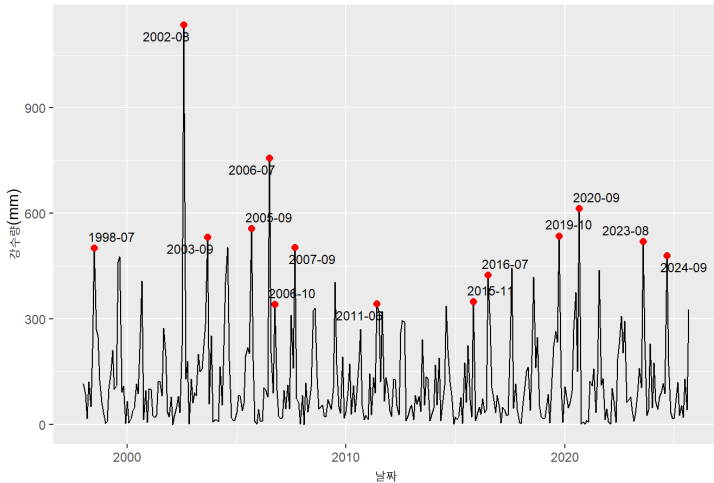
- `forecast::tsoutliers(y)`로 자동 이상치 탐지
- 결과: 시점 index

{56, 93, 103, 262, 273}

등 14개가 이상치로 탐지됨

- 월 단위로 환산하면
 - 56: 2002-08
 - 93: 2005-09
 - 103: 2006-07
 - 262: 2019-10
 - 273: 2020-09
- 모두 강수량이 비정상적으로 큰 이벤트로 해석 가능

강릉 월별 강수량 (이상치 라벨 포함)



빨간 점:

tsoutliers가 탐지한 이상치 월

아이디어: “조건부 저강수 월”

동기

- 단순히 “강수량이 적은 달” 이 아니라,
- 모형이 기대한 값보다 유의하게 적게 내린 달을 잡고 싶다.
- 계절성 + 자기상관 구조를 제거한 후 남은 “이상하게 적은 부분”에 관심.

제안

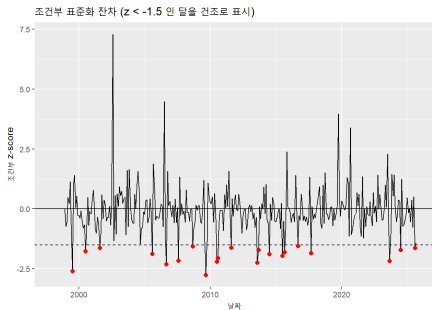
- 모형으로 얻은 조건부 기대값 $\hat{y}_t = \mathbb{E}(y_t \mid \text{과거})$
- 예측오차의 표준편차 $\hat{\sigma}$ 를 이용하여

$$z_t = \frac{y_t - \hat{y}_t}{\hat{\sigma}}$$

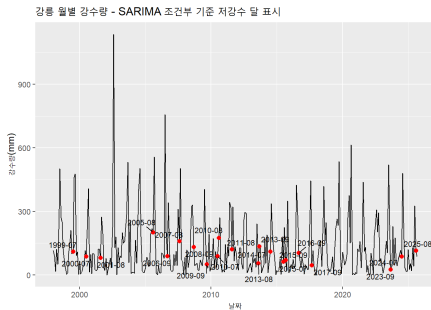
- 임계값 k 에 대해

$$z_t < -k \Rightarrow \text{“조건부 저강수 월”로 판단}$$

모형 기반 조건부 저강수 월 예시

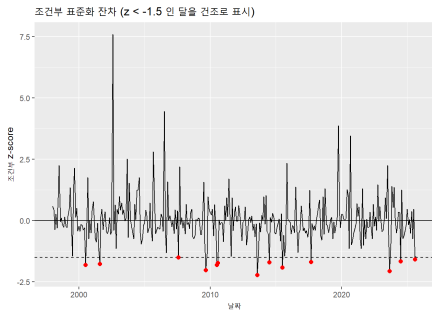


z_t 시계열과 $-k$ 기준선

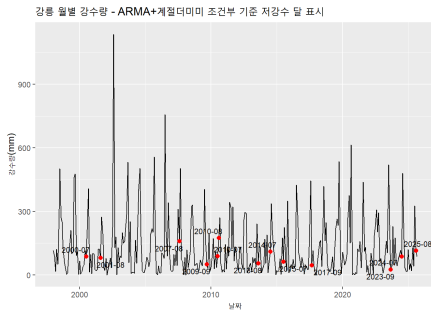


원시계열 + 조건부 저강수 월 표시

모형 기반 조건부 저강수 월 예시



z_t 시계열과 $-k$ 기준선



원시계열 + 조건부 저강수 월 표시

현재까지의 정리

- 기본 모형: $ARIMA(2,0,2)(0,1,1)[12]$ 가 AIC/BIC와 잔차진단을 기준으로 **무난한 시계열 모형**으로 보임
- tsoutliers를 통해 **5개의 시계열 이상치**를 자동 탐지
- 모형 기반으로 조건부 표준화 잔차 z_t 를 정의하고 **조건부 저장수 월** 개념을 도입

- 이상치 처리와 모형 재적합

- tsoutliers로 탐지된 이상치를 실제로 교정(replacement)한 후 모형을 다시 적합하는지 여부는 아직 실험 단계

- 분포 가정과 변환

- 강수량은 비대칭/heavy-tail 분포를 가질 수 있음