

Innovative utilization of internet search volume data to understand public awareness and perception of air quality

Young-Hee Ryu and
Seung-Ki Min
POSTECH, South Korea

0. Question

Why is the public so concerned about air quality in recent years?; even though mean PM_{10} decreases

2. Method

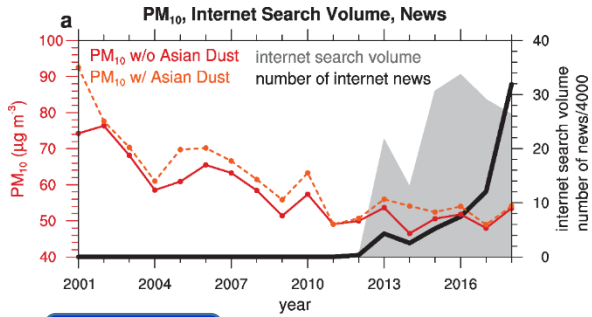
- Decay theory of forgetting: memory decays with time following a power-law function
 - $S = A_i t^{-d}$: strength of memory trace for a single event
 - d = decay exponent
- Assumption: Strength of memory trace (A) depends on PM_{10} and Visibility
 - $A_i = PM_{10_i}^p \text{Visibility}_i^v$
- Air Quality Perception Index (AQPI): Sum of memory trace during the last 6 days
 - $AQPI = \sum_{i=-6}^0 PM_{10_i}^p \text{Visibility}_i^v t^{-d}$, where $t = |i| + 1$
- Internet search volume data = AQPI, we determine the coefficients of d, p, v by minimizing root-mean-square-error

1. Data

18-year winter-spring mean PM_{10} data

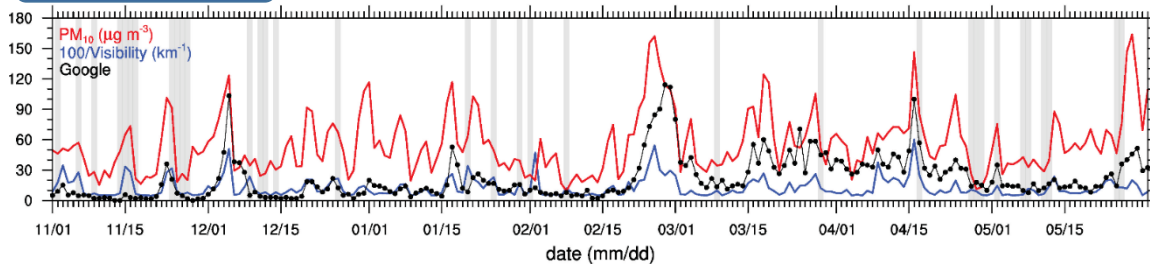
Internet search volume data from Google Trends and NAVER Datalab with a keyword of "air quality"

Number of Internet news about air quality

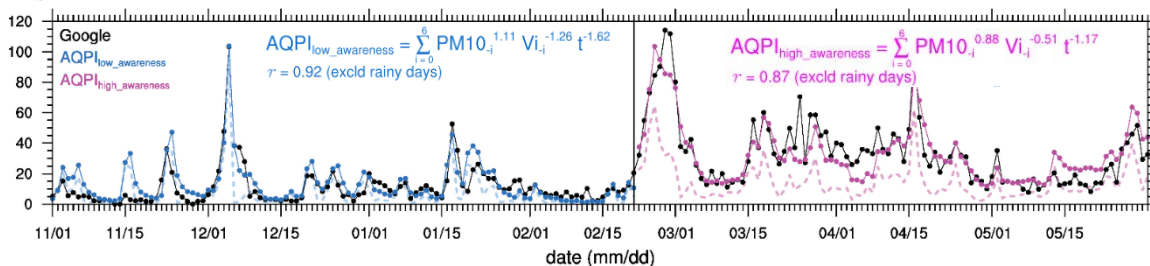


3. Example

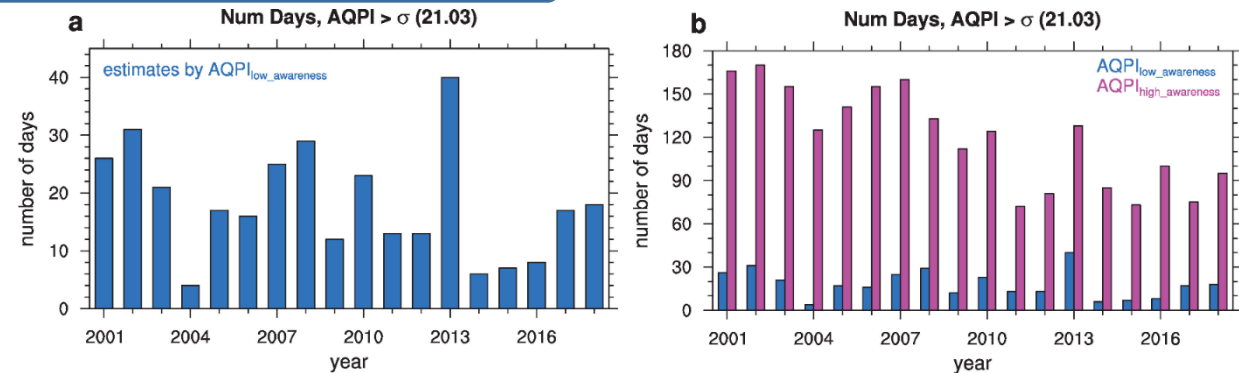
PM_{10} , Visibility, Google Trends, 2013-2014



Google Trends, AQPI models, 2013-2014



4. What matters?: Awareness



Low awareness: 2013 is the worst year in terms of poor perceived air quality.
High awareness: The public would have perceived bad air quality in the past. Raised public awareness caused by severe haze episode explains the elevated public concern.