

부유분진(PM₁₀) 측정기 상태 코드 분석을 통한 자동 품질검사 알고리즘 개선 및 평가

김미경 · 박영산 · 류상범 · 조정훈*

국립기상과학원 환경기상연구과

(접수일: 2019년 8월 9일, 수정일: 2019년 10월 21일, 게재확정일: 2019년 11월 12일)

Improvement and Evaluation of Automatic Quality Check Algorithm for Particulate Matter (PM₁₀) by Analysis of Instrument Status Code

Mi-Gyeong Kim, Young-San Park, Sang-Boom Ryoo, and Jeong Hoon Cho*

Environmental Meteorology Research Division, National Institute of Meteorological Sciences, Jeju, Korea

(Manuscript received 9 August 2019; revised 21 October 2019; accepted 12 November 2019)

Abstract Asian Dust is a meteorological phenomenon that sand particles are raised from the arid and semi-arid regions—Taklamakan Desert, Gobi Desert and Inner Mongolia in China—and transported by westerlies and deposited on the surface. Asian dust results in a negative effect on human health as well as environmental, social and economic aspects. For monitoring of Asian Dust, Korea Meteorological Administration operates 29 stations using a continuous ambient particulate monitor. Kim et al. (2016) developed an automatic quality check (AQC) algorithm for objective and systematic quality check of observed PM₁₀ concentration and evaluated AQC with results of a manual quality check (MQC). The results showed the AQC algorithm could detect abnormal observations efficiently but it also presented a large number of false alarms which result from valid error check. To complement the deficiency of AQC and to develop an AQC system which can be applied in real-time, AQC has been modulated. Based on the analysis of instrument status codes, valid error check process was revised and 6 status codes were further considered as normal. Also, time continuity check and spike check were modified so that posterior data was not referred at inspection time. Two-year observed PM₁₀ concentration data and corresponding MQC results were used to evaluate the modulated AQC compared to the original AQC algorithm. The results showed a false alarm ratio decreased from 0.44 to 0.09 and the accuracy and the probability of detection were conserved well in spite of the exclusion of posterior data at inspection time.

Key words: PM₁₀, quality check, instrument status code, flag, beta ray

1. 서 론

황사는 봄철에 동아시아 지역에서 발생하는 계절 현상으로 타클라마칸 사막, 고비 사막, 내몽골 고원

등 중국의 건조 및 반건조 지역에서 발원한 모래 먼지가 기류에 의해 수송되어 지표로 침강하는 현상이다(Lim and Chun, 2006; Lee and Sohn, 2011; Onish et al., 2012). 봄철이 되면 겨울철 동안 얼어있던 땅이 녹고 지상에는 저기압이 형성되며 기온, 기압, 토양 수분, 풍속 등의 기상 조건에 따라 모래 먼지 입자가 공기 중으로 부유한다. 이때 부유된 입자들이 서쪽에서 동쪽으로 이동하는 기류를 타고 이동하면서 우리나라와 일본, 러시아 등 주변국에 영향을 미치며

*Corresponding Author: Jeong Hoon Cho, Environmental Meteorology Research Division, National Institute of Meteorological Sciences, 33, Seohobuk-ro, Seogwipo-si, Jeju-do 63568, Korea.
Phone: +82-64-780-6584, Fax: +82-64-738-6512
E-mail: xecue@korea.kr

(Iwasaka et al., 1983; Chung et al., 1992; Murayama et al., 2001; Kim, 2008; Yoon et al., 2019), 태평양을 건너 미국 서부의 해안 지역까지 도달하기도 한다(Husar et al., 2001; Eguchi et al., 2009; Yu et al., 2019).

황사는 인체에 해로운 영향을 미치고 사회, 경제, 환경적 측면에 부정적인 영향을 미치기 때문에 지속적인 감시와 신속한 정보 제공을 통해 피해를 최소화하는 것이 중요하다. 대기 중에 부유하여 우리나라에 도달하는 황사 입자는 주로 직경 10 μm 이하의 모래 먼지이며 크기가 작기 때문에 호흡기로 유입되어 천식, 기관지염, 폐렴, 알레르기 비염과 같은 호흡기 질환을 유발할 수 있다(Lee et al., 2013; Goudie, 2014). 이 모래 입자는 시정을 감소시키고 대기층을 통과하는 태양 복사에 산란과 흡수의 영향을 미쳐 지구 복사수지에 영향을 주며, 식물의 기공을 막아 광합성을 억제하고 농작물의 생장을 저해한다. 그뿐만 아니라 조선, 전자, 반도체와 같은 정밀 산업의 생산을 감소시킨다.

기상청은 황사의 실시간 감시를 위해 부유분진 측정기 이용한 황사 감시망을 운영하고 있다. 이러한 지상 관측 자료는 관측 지점의 이전이나 도시화 등 관측 환경의 변화, 장비 교체, 자료 처리 방법의 변화 등 다양한 원인으로 비균질성이 나타날 수 있다(Easterling and Peterson, 1995). 관측 자료의 비균질성은 관측 과정에서 오차를 발생시키며 자료의 신뢰도를 감소시킨다. 측정 과정에서 발생한 오차가 포함된 자료를 그대로 사용하게 될 경우 사용자에게 혼란을 일으킬 수 있다. 대표적으로 실시간 감시 목적으로 사용될 때 오경보를 높여 사회적 비용을 초래할 수 있으며, 수치예측모델의 입력 자료로 사용될 경우 초기 입력 자료의 품질을 저하시킨다. 그러므로 적절한 품질관리 과정을 통해 관측 자료의 신뢰도를 높이는 과정이 필요하다.

기상청은 기상 관측 자료의 품질관리를 위해 실시간, 준실시간, 비실시간 품질검사를 수행하고 있으나 부유분진 측정기로 관측된 PM₁₀ 농도는 준실시간 수동 품질검사만을 수행하고 있었다. 이를 보완하기 위해 객관적인 방법을 통해 자동화할 수 있는 품질검사 체계를 개발하였다(Kim et al., 2016). Kim et al. (2016)은 계측한계 검사, 기기예러 검사, 지속성 검사, 시간 연속성 검사, 튀는 값 검사를 순차적으로 적용하였다. Kim et al. (2016)의 자동 품질검사 검증 결과, 자동 품질검사에서 비정상적으로 판별한 자료 중 수동 품질검사에서도 비정상적으로 판정한 자료의 비율을 나타내는 탐지확률(Probability of detection, POD)과 수동 품질검사 결과와 동일한 검사 결과를 산출한 비율인 정확도(Accuracy, ACC)는 0.985, 0.979로 높게 나타났다. 자동 품질검사에서도 비정상적으로 판별하였으나 수동 품

질검사는 정상으로 판정한 비율을 의미하는 허위경보율(False alarm ratio, FAR)은 0.423로 높게 나타났고, 종합적인 숙련도 지수인 임계성공지수(Critical success index, CSI)는 0.572로 낮게 나타났다. Kim et al. (2016)은 허위경보율이 높게 나타난 원인으로 기기에러 검사 방법을 언급하였으며 기기 상태 코드에 대한 추가 분석을 통해 알고리즘 개선의 필요성을 제시하였다.

선행 연구에서 제기한 문제점을 해결하기 위하여 본 연구는 기기 상태 코드별 분석을 통해 상태 코드 처리 기준을 보완하여 허위경보율을 감소시키는 방법을 제안한다. 또한 기존에는 준실시간으로 적용 가능했던 자동 품질검사를 실시간으로 적용할 수 있도록 알고리즘을 수정하여 자동 품질검사 알고리즘을 개선하고 성능을 평가하고자 한다.

2. 자료 및 방법

2.1 자료

기상청은 황사 감시를 위해 전국 29개 지점에서 부유분진 측정기(FH62C14, Thermo Fisher Scientific Inc.)를 운영하고 있다. 본 연구는 황사 감시를 목적으로 관측하고 있는 PM₁₀ 질량 농도 5분 평균 측정값을 사용하였다. 품질검사 알고리즘의 기기에러 검사를 개선하기 위해 부유분진 측정기의 기기 상태 코드 자료를 사용하였다. 상태 코드는 16진수로 표현되는 여섯 자리의 코드이며, 앞에서부터 두 자리는 종합 오류 상태, 가운데 두 자리는 경고 상태, 마지막 두 자리는 관측기 작동 상태를 나타낸다. 오류 상태는 세부 오류 사항을 종합적으로 나타내지만 세부 오류 코드에 대한 정보가 부족하여 본 연구에서는 종합 오류 상태 자료만 분석하였다. 각 상태별로 사용되는 코드는 Table 1에 제시하였다. 또한 자동 품질검사 알고리즘의 성능을 평가하기 위해 관측자가 수행하는 수동 품질검사(Manual Quality Check, MQC) 정보를 사용하였다. 알고리즘 개선 후 선행 연구와 비교하기 위해 2013년부터 2014년까지 동일한 기간의 자료를 대상으로 분석하였으며, 현재는 관측이 종료된 지점을 제외한 총 27개 지점의 자료를 사용하였다.

2.2 자동 품질검사 알고리즘

황사 감시망의 PM₁₀ 농도 관측 자료의 객관적이고 체계적인 품질관리를 위해 Kim et al. (2016)은 자동 품질검사(Automatic Quality Check, AQC) 알고리즘을 개발하였다. Table 2에 선행 연구에서 개발한 품질검사 알고리즘의 단계별 설명을 제시하였다. 각 단계의 검사 기준을 만족하지 않는 경우 비정상 자료로 분류되며 품질검사 과정이 종료된다.

선행 연구의 분석에 사용된 유효 자료 수, 허위경

Table 1. Status codes for three modes (Thermo Fisher Scientific Inc., 2010).

Global Error Status		Warning Status		Operating Status	
00	Normal operation	00	Normal operation	00	Normal operation, no error
01	Sum status of the data and program memory	01	Maximum pump power	01	Mechanical filter change
02	Sum status of sampling and measuring system	02	Mass limit is exceeded; filter change will follow	02	Plateau check and zero adjustment of the mass (R ⁰)
04	Sum status of pressure sensors	04	Not used	04	Not used
08	Sum status of air flow regulation	08	Relative Humidity > RH Set point + 10%	08	Pump switched off
10	Sum status of the sample heater	10	Not used	10	Manual operation (off-line, keyboard enabled)
20	Sum status of temperature measurement	20	Compensation of the mass signal > ± 5%	20	Calibration enabled
40	Reserved	40	Two consecutive plateau checks are wrong during filter change: R _{ref} /R ₁ < 30% or > 70%	-	-
80	Not used	80	Not used	80	Power supply switched on

Table 2. Description of AQC algorithm.

Steps	Descriptions
I (Valid value check)	Check whether the observed PM ₁₀ mass concentration meets the requirements of instrumentation limit.
II (Valid error check)	Check whether the instrument status codes are equal to ‘Normal operation’ or ‘Not used’.
III (Persistence check)	Check if the PM ₁₀ mass concentration during the previous hour has no variability.
IV (Time continuity check)	Check if the difference of PM ₁₀ concentration between inspection time and the time before and after is out of a certain range.
V (Spike check)	Check whether the second-order differential coefficient of PM ₁₀ mass concentration has rapid fluctuation.

보 횟수, MQC, AQC 결과 및 관측 환경 정보는 Table 3에 제시된 바와 같다. AQC의 성능을 평가한 결과 허위경보는 0.1%에서 13.5%로 나타났으며 평균적으로 2.1%의 허위경보가 발생하였다. Kim et al. (2016)에 의하면 허위경보의 약 88%가 2단계(기기예러 검사)에서 비정상 자료로 분류되었기 때문이다.

2.3 개선된 자동 품질검사 알고리즘

연속된 관측 자료 확보가 필요한 황사 감시 자료에서 AQC의 허위경보율이 높다면 감시 효율 측면에서 AQC를 사용하기에 어려움이 있다. Table 4는 기기 상태 코드의 출현 빈도와 출현 시 허위경보 발생률을 분석한 결과이다. 결과를 보면, 특정 기기 상태 코드

에서 높은 발생 빈도와 허위경보 발생률이 집중적으로 나타났다. 이를 바탕으로 관측기의 정상 작동 중에 발생하는 작동 상태 코드를 제외한 출현 빈도가 5000회 이상이며 허위경보 발생률이 70% 이상인 6개의 기기 상태 코드의 특성을 분석하고 처리 기준을 조정하였다.

기존에 개발된 AQC 알고리즘은 4단계(시간연속성 검사)에서 Jiang et al. (2013)이 제시한 식(1), 5단계(튀는 값 검사)에서 Papale et al. (2006)이 제시한 식(2)를 사용하며, 각각 검사 시각(t)의 5분 후 자료(t+1)를 사용한다.

$$A = \min|PM_{10}(t) - PM_{10}(t \pm 1)| \geq 50 + 0.15 \times PM_{10}(t) \quad (1)$$

Table 3. Station information and number of significant data and percentage of false alarm, normal data considered at MQC and AQC, and result of AQC.

Station	Station type	Significant data	False alarm (%)	MQC passed (%)	AQC passed (%)	AQC (%)				
						Step I	Step II	Step III	Step IV	Step V
1	Ocean	205582	0.5	99.6	98.5	3.2	0.5	-	0.0	0.0
2	Mountain	206349	3.1	99.4	88.7	10.1	3.0	-	0.0	0.1
3	Mountain	203689	0.1	99.6	96.4	6.7	0.1	-	-	0.0
4	River	197675	4.4	99.6	93.6	8.3	4.4	-	0.0	0.0
5	Ocean	202519	0.8	99.7	96.2	6.7	0.5	-	0.0	0.3
6	Land	204883	0.4	99.6	98.1	4.1	0.4	-	0.0	0.0
7	Ocean	201701	1.3	99.6	89.4	13.4	1.2	-	0.0	0.1
8	Mountain	205098	13.5	99.1	81.7	7.1	13.6	-	0.0	0.1
9	River	207825	0.8	99.6	98.3	2.1	0.8	-	0.0	0.0
10	River	195007	0.3	99.5	98.3	9.2	0.2	-	0.0	0.1
11	Land	208430	0.1	99.6	98.5	2.3	0.1	-	0.0	0.1
12	Ocean	207159	0.1	99.6	99.2	2.1	0.1	-	-	0.0
13	Land	204682	0.3	99.5	97.0	5.2	0.4	-	0.0	0.0
14	Land	208329	0.2	99.6	99.1	1.7	0.1	-	-	0.0
15	Ocean	205097	0.2	99.5	99.0	3.3	0.2	-	0.0	0.0
16	Land	207317	7.1	99.5	91.1	3.1	7.1	-	-	0.1
17	Land	194355	5.0	99.9	93.8	9.4	5.0	-	0.0	0.0
18	Land	204964	1.1	97.6	97.8	3.6	1.1	-	0.0	0.0
19	Land	198327	0.7	99.6	95.4	9.8	0.7	-	0.0	0.0
20	Mountain	196393	1.1	99.5	90.0	15.9	1.0	-	0.0	0.1
21	Ocean	203891	0.3	99.6	97.8	4.9	0.3	-	0.0	0.1
22	Ocean	200704	3.1	99.6	93.3	8.3	3.1	-	0.0	0.0
23	Ocean	201748	0.6	99.4	95.1	8.5	0.5	-	0.0	0.1
24	River	206862	0.6	99.6	98.0	3.1	0.6	-	-	0.0
25	Ocean	205564	0.5	99.5	98.3	3.4	0.3	-	0.0	0.2
26	Ocean	197823	2.9	99.6	93.7	9.6	2.8	-	0.0	0.2
27	River	201757	6.1	99.6	91.7	6.4	6.1	-	-	0.0

Table 4. Frequency for each status code and its ratio of false alarm from 2013 to 2014.

Code	Global error status			Warning status			Operating status		
	Total	False alarm	Ratio	Total	False alarm	Ratio	Total	False alarm	Ratio
00	5443329	164195	3.0	5327020	45879	0.9	5430227	173202	3.2
01	5	0	0.0	2	2	100.0	554	521	94.0
02	10660	1338	12.6	2	2	100.0	2482	2439	98.3
04	597	248	41.5	-	-	-	24187	624	2.6
08	10859	8299	76.4	33481	31703	94.7	11462	1972	17.2
10	15081	14209	94.2	-	-	-	16380	15539	94.9
20	7665	5532	72.2	45655	44804	98.1	4593	4081	88.9
40	-	-	-	77121	70329	91.2	-	-	-
80	-	-	-	-	-	-	5	0	0.0

$$d = (PM_{10}(t) - PM_{10}(t-1)) - (PM_{10}(t+1) - PM_{10}(t)) \quad (2)$$

제시된 식에 따르면 시간연속성 검사와 뒤는 값 검

사를 수행하기 위해서는 연속된 3개의 자료가 필요하다. 본 연구는 검사 시각 기준으로 미래 자료를 참조하지 않고 실시간으로 적용하기 위해 검사 시각(t) 이전 5분($t-1$)과 10분($t-2$) 자료를 사용하여 미래 자

Table 5. Status codes which frequency records high false alarm for each station.

Station	High false alarm status code					
	000800	002000	004000	080000	100000	200000
1	11	738	4	77	18	10
2	1354	4804	-	117	1376	-
3	8	98	-	31	3	-
4	4783	3941	1426	300	2740	-
5	26	162	4165	402	18	-
6	32	638	2386	48	31	-
7	6	2302	1909	187	6	-
8	12132	4764	3	9	10326	2657
9	8	1093	890	278	8	-
10	11	364	6	11	11	-
11	5	10	333	6	24	-
12	13	8	4	13	13	-
13	3	390	-	54	3	-
14	1	180	-	57	-	-
15	6	27	1	145	5	-
16	7	7309	4	2374	4	-
17	2415	6994	-	158	1	-
18	4	2085	-	213	-	-
19	6	1387	19311	2	2	-
20	1302	808	1626	185	43	-
21	234	1	1	41	-	-
22	1897	466	-	3409	3	-
23	5	672	31638	5	5	-
24	-	1068	-	51	-	-
25	15	347	1	35	15	-
26	28	1901	1	25	428	4989
27	9641	3179	13412	4	70	-

료를 사용하지 않도록 선행 연구에서 제시한 검사 식을 다음 식(3), (4)와 같이 변경하였다.

$$A = |PM_{10}(t) - PM_{10}(t-1)| \geq 50 + 0.15 \times PM_{10}(t) \quad (3)$$

$$d = \frac{PM_{10}(t-1) - PM_{10}(t-2)}{-(PM_{10}(t) - PM_{10}(t-1))} \quad (4)$$

3. 결 과

3.1 상태 코드 분석 및 자동 품질검사 기준 조정

허위경보가 높게 나타나는 6개의 상태 코드의 지점별 출현 빈도는 Table 5에 제시된 바와 같다. 상태 코드 처리 기준을 변경하기 위하여 허위경보가 높게 나타난 코드를 경고, 오류 상태의 범주 따라 살펴보았다. 각 상태 코드별 발생 빈도를 시간과 계절에 따라 분석하여 Fig. 1과 Fig. 2에 제시하였다.

경고 상태에서는 상대습도 관련, 질량 신호 보정 관련, Plateau check 관련 코드에서 허위경보가 높게 나타났다. 상대습도와 관련된 경고 상태 코드는 '000800'

이며, 부유분진 측정기의 온도 센서를 통해 계산된 상대습도가 설정값의 10%를 초과하면 경고 상태가 발생한다(Thermo Fisher Scientific Inc., 2010). 상태 코드의 시간별 발생 빈도는 일 최고 기온이 나타나는 오후 2~3시에 가장 낮고 일출 전에 가장 높은 분포가 나타나 상대습도의 일 변화와 유사한 경향을 보였다(Fig. 1b). 하천 및 해안 지역 관측소 4, 27의 경우 집중적으로 상대습도 경고 상태가 발생하였으며 평균 상대습도 및 표준편차는 $73 \pm 6\%$ 로 나타났다. 그러나 관측 환경이 유사한 다른 관측소와 비교하였을 때 평균 상대습도와 표준편차($77 \pm 9\%$)는 큰 차이를 보이지 않는 것으로 나타났다. 이는 상대습도 경고 상태가 환경적 요소 외에도 관측기의 제습 과정 문제 등으로 발생할 수 있음을 의미한다. 질량 신호 보정이 $\pm 5\%$ 를 초과한 경우는 경고 상태 코드 '002000'이 발생한다. PM_{10} 질량 농도를 연속으로 측정할 때 이미 필터에 누적된 입자가 있으므로 클린 필터의 값을 참조하여 보정하며, 이때 보정 비율이 높아지면 질량 신호 보정 경고가 발생한다. 시간에 따른 발생 빈도는

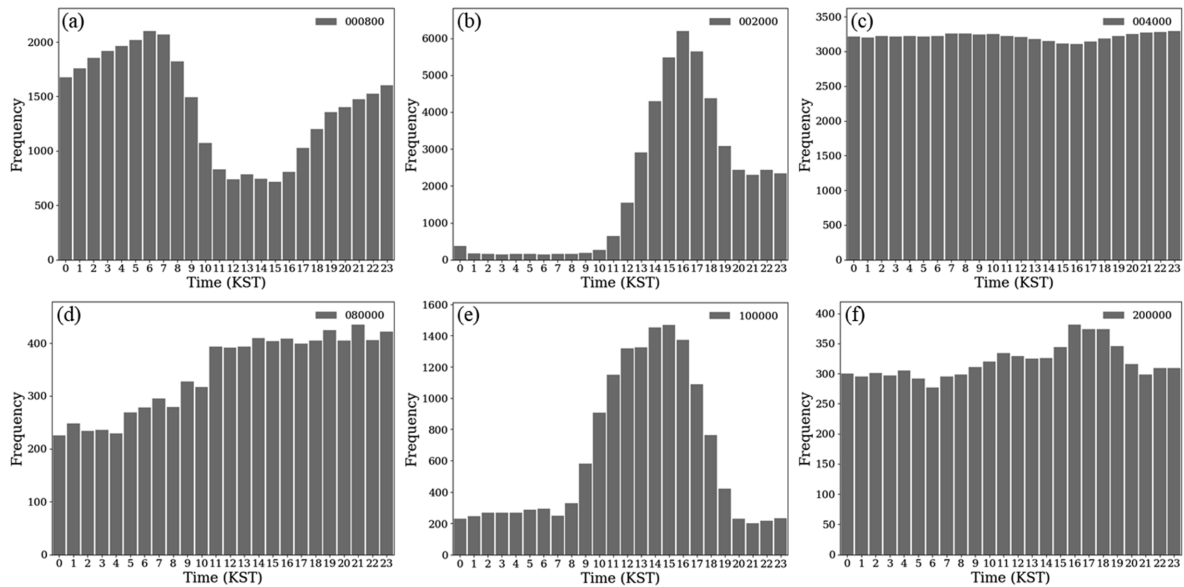


Fig. 1. Hourly frequencies of each status code; (a) 000800, (b) 002000, (c) 004000, (d) 080000, (e) 100000 and (f) 200000.

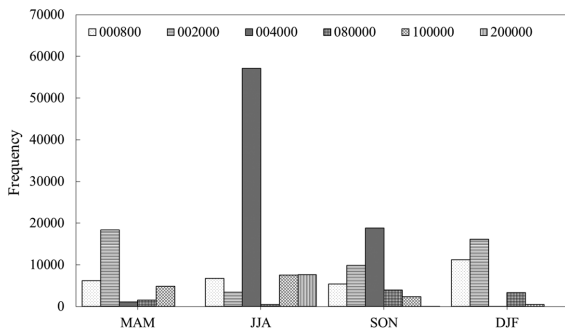


Fig. 2. Seasonal frequencies of 6 status codes.

오전 11시 이후 급격하게 증가하여 오후 4시에 가장 많이 발생하는 것으로 나타났다(Fig. 1c). 또한, 계절별 발생 빈도는 봄철 및 겨울철에 더 높았으며 이는 봄과 겨울철에 PM₁₀ 질량 농도가 높게 나타나므로 (Lee and Kim, 2018) 질량 신호 보정과 관련된 경고 상태가 발생할 확률이 높은 것으로 생각된다. Plateau check에서 연속해서 오류가 생기면 경고 상태 코드 '004000'이 나타난다. Plateau check는 부유분진 측정기의 입자 검출기가 베타선을 이용하여 PM₁₀을 측정하는 과정에서 알파선의 간섭을 최소화하도록 설계되어 있는지 확인하는 과정이며 자동으로 필터가 교체되는 과정에서 수행된다. 시간별 발생 빈도는 뚜렷한 특징이 없었으나 계절별로는 봄철과 겨울철에 거의 발생하지 않고 주로 여름철에 발생하는 것으로 나타

났다. 이러한 경고 상태 코드가 계속해서 나타난다면 장비 점검이 필요한 신호로 생각할 수 있다.

허위경보가 높았던 종합 오류 상태 코드는 유량 조절, 샘플 가열기, 온도 측정 상태이다. 관측 장비 내부로 유입되는 유량의 조절 상태에 오류가 있는 경우 '080000' 코드가 나타난다. 계절별 분석 결과 주로 가을철과 겨울철에 나타나는 것을 알 수 있다. 샘플 가열기의 종합 오류 상태를 나타내는 '100000'은 Table 5에 제시된 바와 같이 상대습도 경고와 같이 나타나는 경향이 있다. 이는 샘플 가열기에서 포집된 공기를 가열하여 상대습도를 설정된 값으로 맞추는데 이러한 과정에서 종종 오류가 발생하기 때문인 것으로 생각된다. 시간 및 계절별 발생 빈도 분석 결과 주로 봄철과 여름철의 낮 시간에 많이 발생하는 것을 알 수 있다(Fig. 1f). 온도 측정 상태의 종합적인 오류를 나타내는 '200000'은 주로 여름철에 발생하며 강한 일사와 높은 기온으로 인해 온도 측정 센서에 오류가 발생하는 것으로 추정된다. 전체 오류 상태 코드의 경우 모두 상대습도와 관련된 부분이지만 세부 오류 상태 코드에 대한 정보가 수집되지 않아 자세한 원인을 분석할 수 없었다.

6개의 상태 코드가 가장 많이 나타났던 지점을 각각 선택하여 상태 코드가 '정상'일 때와 허위경보가 높은 상태 코드가 발생했을 때의 관측 PM₁₀ 평균 농도와 표준편차를 비교하였다(Table 6). 경고 상태의 경우 지점별 평균 PM₁₀ 농도는 정상 상태일 때의 평균 농도의 1 표준편차 범위 내에 있으며, 표준

Table 6. Mean and standard deviation of observed PM₁₀ concentration for the normal status code and six status codes at the representative stations.

Station	8		16		23		22		8		26	
Status code	Normal	000800	Normal	002000	Normal	004000	Normal	080000	Normal	100000	Normal	200000
PM ₁₀ (μgm ⁻³)	46 ± 40	54 ± 40	38 ± 27	43 ± 26	33 ± 28	23 ± 26	34 ± 30	27 ± 25	46 ± 40	63 ± 37	39 ± 31	10 ± 11

Table 7. Dropout rates of 6 status codes on each AQC step in using revised algorithm.

Status code	Step I	Step II	Step III	Step IV	Step V	Normal
000800	1781	5	-	30	120	32017
002000	454	637	-	3	80	44562
004000	6618	133	-	10	311	70049
080000	254	2172	-	11	19	5781
100000	844	29	-	14	117	14149
200000	1764	249	-	-	136	5507

Table 8. Skill scores of AQC algorithms for all stations.

	Accuracy	Probability of detection	False alarm ratio	Critical success index
Kim et al. (2016)	0.979	0.985	0.423	0.572
Revised step II algorithm	0.997	0.974	0.088	0.891
Real-time algorithm	0.997	0.975	0.080	0.899

편차의 크기도 유사하게 나타났다. 오류 상태의 경우는 대체로 유사하게 나타났으나 온도 측정 오류(‘200000’)의 경우 정상 상태의 평균 농도와 차이가 나타났으며, 정상 관측에 비해 낮은 PM₁₀ 값이 관측되었다.

이를 바탕으로 6개의 상태 코드가 정상으로 처리되도록 알고리즘을 수정하였다. 6개의 상태 코드를 정상으로 처리했을 때 AQC 단계별 플래그를 조사한 결과 평균적으로 약 83% 자료가 정상으로 분류되고 나머지 17% 자료만이 비정상 자료로 검출되었다(Table 7). 유량 조절 오류(‘080000’)의 경우 오류 상태에서 비정상을 나타내는 다른 코드와 동시에 발생하여 2단계에서 비정상 자료로 분류되었다. 정상 관측에 비해 낮은 값이 측정된 온도 측정 오류의 경우 측정값이 관측 범위를 벗어나 1단계 관측한계 검사에서 비정상 자료로 분류되었다.

3.2 자동 품질검사 알고리즘 개선 결과

기기 상태 코드 분류 기준을 새롭게 적용한 AQC를 수행하고 전 지점에 대한 숙련도 평가 결과를 Table 8에 제시하였다. Kim et al. (2016)의 결과와 비교하였을 때, 전 지점 정확도는 0.979에서 0.997로 0.018 증가하였고 허위경보율은 0.423에서 0.088로 0.335 감소하였다.

한편, 기존 알고리즘은 4, 5단계 검사 과정에서 검

사 시각으로부터 5분 후의 자료를 참조하도록 설계되어 있었다. 본 연구의 분석 과정에서는 AQC 수행 시 미래 자료를 참조하지 않고 실시간으로 수행할 수 있도록 알고리즘을 변경하였다. 실시간 AQC 알고리즘을 적용한 결과, 허위경보율은 0.088에서 0.080으로 감소하고 임계성공지수는 0.891에서 0.899로 증가하여 실시간 알고리즘을 적용하여도 성능 저하가 일어나지 않음을 확인하였다.

4. 요약 및 결론

Kim et al. (2016)은 기상청에서 황사 감시를 목적으로 운영하고 있는 부유분진 측정기 PM₁₀ 농도 관측 자료의 체계적인 품질관리를 위해 AQC 알고리즘을 개발하였다. 그러나 선행 연구는 기기 상태 코드에 대한 추가 분석을 통해 AQC의 성능 개선이 필요함을 제시한 바 있다. 본 연구는 기기 상태 코드를 분석하고 기기여러 검사의 처리 기준을 개선하여 AQC의 높은 허위경보율을 개선하였다. 또한, 준실시간으로 적용할 수 있던 기존의 알고리즘을 실시간으로 적용할 수 있도록 수정하고 선행 연구의 성능 평가 결과와 비교하였다.

전 지점에서 발생하는 기기 상태 코드 중 발생 빈도가 5000회 이상이며 허위경보 발생률이 70% 이상인 6개의 코드(Table 4)를 대상으로 지점별 발생

빈도와 코드별 시간 당 발생 횟수를 분석하였다. 각 코드별 PM₁₀ 평균 농도 및 표준편차는 대부분 정상 자료의 범위 내에 있는 것으로 나타났다. 예외적으로 온도 측정 오류 상태의 경우 정상 자료의 범위보다 낮은 값이 관측 되었으나, 황사 현상 감시는 고농도 PM₁₀의 관측 여부가 중요하므로 정상 자료 범위보다 낮은 측정값은 정상 자료로 간주하였다. 다음으로 시간연속성 검사와 튀는 값 검사 단계에서 검사 기준 시각 5분 후의 미래 자료를 참조하는 부분을 수정하여 실시간으로 품질검사가 가능하도록 하였다.

AQC의 기기 상태 코드 처리 기준을 개선한 결과, 전 지점 정확도는 0.997, 허위경보율은 0.088로 나타났으며 정상 자료 판별률은 MQC의 정상 자료 판별률(96%)에 근접한 94%로 나타났다. 또한 실시간 품질검사를 수행할 수 있도록 수정한 알고리즘을 적용한 결과, 허위경보율은 0.080으로 감소하였고 임계성 공지수는 0.899로 증가하였다. 이를 통해 실시간 AQC 알고리즘을 적용하여도 성능 저하가 일어나지 않음을 확인하였다.

관측 환경의 변화로 인해 정상 작동 외의 다른 상태 코드가 발생할 수 있으며, 이는 전체 자료 중 약 4%로 나타났다. 정상 작동이 아닌 다른 코드가 계속해서 발생하는 경우 기기 점검 및 유지 보수가 이루어지며, 관측 장비 점검과 유지 보수가 완료되기까지 소요되는 시간이 있기 때문에 얼마 간은 이러한 상태 코드가 발생한다. 장비 점검이나 수리가 완료되면 다시 정상 작동을 의미하는 상태 코드가 표출된다. 환경 기상 감시를 목적으로 연속적인 자료의 확보가 중요한 경우에는 기기의 작동, 경고, 오류 상태에 따른 상태 코드가 발생한 자료를 무조건 비정상 자료로 처리하기 보다 자료의 용도에 따라 선별적으로 처리하고 사용하여야 한다.

본 연구는 상태 코드 분석을 통해 허위경보가 높게 나타났던 6개의 상태 코드를 정상 자료로 처리함으로써 AQC를 사용할 때 실시간 감시 효율을 높일 수 있는 방안을 제시하였다. 관측 장비의 상태 코드 분석 자료는 관측 시기별 주요 점검 사항 제작 등 효율적인 장비 운영 지침을 작성하기 위한 기반 자료로 활용될 수 있다. 추후 장비 상태 코드의 주기적인 통계 분석을 통해 체계적인 장비 관리 지침을 작성할 수 있을 것으로 기대한다.

감사의 글

이 연구는 기상청 국립기상과학원 「황사·연무 감시 및 예보기술 개발」(1365003013)의 지원으로 수행되었습니다.

한국기상학회 대기 제29권 4호 (2019)

REFERENCES

- Chung, Y.-S., T.-K. Kim, and J.-S. Chung, 1992: On relationship of low visibility to air pollution in cities. *J. Korean Soc. Atmos. Environ.*, **8**, 1-6 (in Korean with English abstract).
- Easterling, D. R., and T. C. Peterson, 1995: A new method for detecting undocumented discontinuities in climatological time series. *Int. J. Climatol.*, **15**, 369-377.
- Eguchi, K., I. Uno, K. Yumimoto, T. Takemura, A. Shimizu, N. Sugimoto, and Z. Liu, 2009: Trans-pacific dust transport: integrated analysis of NASA/CALIPSO and a global aerosol transport model, *Atmos. Chem. Phys.*, **9**, 3137-3145.
- Goudie, A. S., 2014: Desert dust and human health disorders. *Environ. Int.*, **63**, 101-113, doi:10.1016/j.envint.2013.10.011.
- Husar, R. B., and Coauthors, 2001: Asian dust events of April 1998. *J. Geophys. Res.*, **106**, 18317-18330.
- Iwasaka, Y., H. Minoura, and K. Nagaya, 1983: The transport and spacial scale of Asian dust-storm clouds: a case study of the dust-storm event of April 1979. *Tellus B*, **35B**, 189-196.
- Jiang, Z., Z. Liu, T. Wang, C. S. Schwartz, H.-C. Lin, and F. Jiang, 2013: Probing into the impact of 3DVAR assimilation of surface PM₁₀ observations over China using process analysis. *J. Geophys. Res.*, **118**, 6738-6749, doi:10.1002/jgrd.50495.
- Kim, J., 2008. Transport routes and source regions of Asian dust observed in Korea during the past 40 years (1965~2004). *Atmos. Environ.*, **42**, 4778-4789.
- Kim, S., H. C. Lee, and S.-B. Ryoo, 2016: Development and assessment of real-time quality control algorithm for PM₁₀ data observed by continuous ambient particulate monitor. *Atmosphere*, **26**, 541-551 (in Korean with English abstract).
- Lee, E.-H., and B.-J. Sohn, 2011: Recent increasing trend in dust frequency over Mongolia and Inner Mongolia regions and its association with climate and surface condition change. *Atmos. Environ.*, **45**, 4611-4616, doi:10.1016/j.atmosenv.2011.05.065.
- Lee, H., H. Kim, Y. Honda, Y.-H. Lim, and S. Yi, 2013: Effect of Asian dust storms on daily mortality in seven metropolitan cities of Korea. *Atmos. Environ.*, **79**, 510-517, doi:10.1016/j.atmosenv.2013.06.046.
- Lee, J., and K.-Y. Kim, 2018: Analysis of source regions and meteorological factors for the variability of spring PM₁₀ concentrations in Seoul, Korea. *Atmos. Environ.*, **175**, 199-209, doi:10.1016/j.atmosenv.2017.12.013.

- Lim, J.-Y., and Y. Chun, 2006: The characteristics of Asian dust events in Northeast Asia during the springtime from 1993 to 2004. *Global Planet. Change*, **52**, 231-247.
- Murayama, T., and Coauthors, 2001: Ground-based network observation of Asian dust events of April 1998 in east Asia. *J. Geophys. Res.*, **106**, 18345-18359.
- Onishi, K., Y. Kurosaki, S. Otani, A. Yoshida, N. Sugimoto, and Y. Kurozawa, 2012: Atmospheric transport route determines components of Asian dust and health effects in Japan. *Atmos. Environ.*, **49**, 94-102, doi: 10.1016/j.atmosenv.2011.12.018.
- Papale, D., and Coauthors, 2006: Towards a standardized processing of Net Ecosystem Exchange measured with eddy covariance technique: algorithms and uncertainty estimation. *Biogeosciences*, **3**, 571-583.
- Thermo Fisher Scientific Inc., 2010: FH62C14 Continuous Ambient Particulate Monitor Instruction Manual. Part N. 100230-00, 136 pp.
- Yoon, J.-E., J.-H. Lim, J.-M. Shim, J.-I. Kwon and I.-N. Kim, 2019: Spring 2018 Asian Dust Events: Sources, Transportation, and Potential Biogeochemical Implications. *Atmosphere*, **10**, 276-293, doi:10.3390/atmos10050276.
- Yu, Y., O. V. Kalashnikova, M. J. Garay, and M. Notaro, 2019: Climatology of Asian dust activation and transport potential based on MISR satellite observations and trajectory analysis. *Atmos. Chem. Phys.*, **19**, 363-378, doi: 10.5194/acp-19-363-2019.
- Zahumenský, I., 2004: Guidelines on quality control procedures for data from automatic weather stations. 10 pp [Available online at [https://www.wmo.int/pages/prog/www/IMOP/publications/IOM-82-TECO_2005/Papers/3\(14\)_Slovakia_2_Zahumensky.pdf](https://www.wmo.int/pages/prog/www/IMOP/publications/IOM-82-TECO_2005/Papers/3(14)_Slovakia_2_Zahumensky.pdf)].