WISE 복합기상센서 관측 자료 품질관리시스템

채정훈 · 박문수* · 최영진

기상기술개발원 차세대도시농림융합기상사업단

(접수일: 2014년 6월 20일, 게재확정일: 2014년 7월 29일)

The WISE Quality Control System for Integrated Meteorological Sensor Data

Jung-Hoon Chae, Moon-Soo Park*, and Young-Jean Choi

Weather Information Service Engine, Center for Atmospheric Science and Earthquake Research, Seoul, Korea

(Manuscript received 20 June 2014; accepted 29 July 2014)

Abstract A real-time quality control system for meteorological data (air temperature, air pressure, relative humidity, wind speed, wind direction, and precipitation) measured by an integrated meteorological sensor has been developed based on comparison of quality control procedures for meteorological data that were developed by the World Meteorological Organization and the Korea Meteorological Administration (KMA), using time series and statistical analysis of a 12-year meteorological data set observed from 2000 to 2011 at the Incheon site in Korea. The quality control system includes missing value, physical limit, step, internal consistency, persistence, and climate range tests. Flags indicating good, doubtful, erroneous, not checked, or missing values were added to the raw data after the quality control procedure. The climate range test was applied to the monthly data for air temperature and pressure, and its threshold values were modified from $\pm 2\sigma$ and $\pm 3\sigma$ to $\pm 3\sigma$ and $\pm 6\sigma$, respectively, in order to consider extreme phenomena such as heat waves and typhoons. In addition, the threshold values of the step test for air temperature, air pressure, relative humidity, and wind speed were modified to 0.7°C , 0.4 hPa , 5.9% , and 4.6 m s^{-1} , respectively, through standard deviation analysis of step difference according to their averaging period. The modified quality control system was applied to the meteorological data observed by the Weather Information Service Engine in March 2014 and exhibited improved performance compared to the KMA procedures.

Key words: AWS, climate range test, integrated meteorological sensor, quality control, WISE

1. 서 론

자동기상관측장비(Automatic Weather System, AWS)로 실시간 수집되는 관측 자료에는 기기고장, 교정편차, 편향오차 등의 관측 장비 자체의 오류와 전기 및 통신 장애 등에 의한 오류들이 포함되어 있다(Wade, 1987; WMO, 2004; Heo *et al.*, 2005). 미국품질협회

(American Society for Quality)에서는 품질에 대한 요구사항을 충족시키기 위하여 사용되는 관측 기술과 활동을 품질관리로 정의하고 있으며, 기상분야에서 품질관리란 사용자가 관측 자료를 최적으로 사용할 수 있도록 오류자료 검출과 오차 보정을 통하여 관측 자료의 정확도를 최대한 높이는 것을 말한다(WMO, 2004; KMA, 2006). 이러한 관측 자료의 품질관리는 사용자에게 보다 정확한 자료를 제공하기 위해 필수적이다.

세계기상기구(World Meteorological Organization, WMO)에서는 AWS 관측 자료의 품질관리에 대한 가이드라인을 제시하였으며(WMO, 2004), 기상 관측망

*Corresponding Author: Moon-Soo Park, Weather Information Service Engine, Center for Atmospheric Science and Earthquake Research, Sang-Am IT building 12th Fl., 434, World Cup buk-ro, Mapo-gu, Seoul 121-835, Korea.
Phone : +82-70-4617-3769, Fax : +82-2-6922-2930
E-mail : ngeograph2@gmail.com

을 운영하고 있는 세계 여러 기관들도 독자적인 품질 관리시스템을 개발하여 활용하고 있다. 예를 들면, 오 클라호마주와 오 클라호마 대학이 함께 구축한 메조넷(Mesonet)에서는 기후범위검사, 단계검사, 지속성검사, 공간검사, 동일장비비교검사 과정을 포함한 품질관리 시스템을 개발하였다(Shafer *et al.*, 2000; Fiebrich and Crawford, 2001). 그리고, 노르웨이 기상연구소에서는 관측소 수행 품질관리(QC0), 실시간 품질관리(QC1), 비실시간 품질관리(QC2), 관리자 품질관리(HQC)의 4 단계로 이루어진 품질관리시스템을 개발하였다(Vejen *et al.*, 2002). 이 중 실시간 품질관리(QC1) 과정에는 결측자료검사, 기후범위검사, 물리관계검사, 단계검사, 내적일치성검사, 구문검사, 모델비교검사가 포함되어 있다. 여기서 구문검사란 변수, 단위 등이 규칙에 맞게 작성되어 있는지를 점검하는 과정이다.

우리나라 기상청에서도 기상 관측 자료 실시간 품질관리시스템(Real-time Quality control system for Meteorological Observation Data, RQMOD)을 개발하여 실시간으로 수집되는 지상, 고층, 해양의 기상 관측 자료에 대한 품질관리를 수행하고 관측 자료에 품질 정보 플래그를 추가하여 사용자에게 제공한다. 이 중 AWS에 대한 품질관리는 물리관계검사, 기후범위검사, 단계검사, 지속성검사, 내적일치성검사, 중앙값 필터검사로 이루어져 있다(KMA, 2006).

최근 도시기상에 대한 관심이 높아지면서 복잡한 도시에서 관측이 용이한 복합기상센서의 사용이 증가하고 있다. 미국 오 클라호마 시티에서는 도심 신호등에 복합기상센서를 설치하여 마이크로넷(Micronet)이라는 관측망을 구축하였고(Winning *et al.*, 2007; Basara *et al.*, 2010), 국립기상연구소에서는 강남의 선릉 지역에 복합기상센서를 설치하여 도시에서의 녹지효과를 연구하였다(Kim *et al.*, 2014). 한편, 차세대도시농림융합기상사업단(Weather Information Service Engine, WISE)에서도 도시기상 관측기술 개발을 위해 인천 지역에 복합기상센서를 포함한 관측망을 구축하였다.

복합기상센서는 크기가 작고 다양한 구조물을 이용하여 설치할 수 있으며 기온, 기압, 습도, 풍향, 풍속, 강우량을 하나의 센서에서 관측하므로 단일장비임에도 AWS와 유사한 역할을 할 수 있다. 그러나 일부 기상 변수의 측정 방식이 기존의 AWS에 사용되는 관측 센서와 다르고 모든 기상 변수를 동일한 고도에서 측정하는 점에서 차이를 보인다.

본 연구에서는 이러한 복합기상센서의 특성을 고려하여 기존 연구에서 제시된 품질관리시스템의 알고리즘들을 선택적으로 적용하고 과거 기상 자료의 시계열분석 및 통계 분석을 통하여 적용된 알고리즘의 문턱값을 수정하여 복합기상센서 관측 자료에 대한 품질관리시스템(Quality Control system for Integrated

Fig. 1. Flow chart of the WISE quality control system for integrated meteorological sensor data.

Meteorological Sensor data, QCIMS)을 개발하고자 한다.

2. QCIMS의 시스템 구성

2.1 시스템 운용 기반

QCIMS는 Fig. 1에서 보여지는 것과 같이 품질관리에 필요한 변수를 정의하는 모듈, 현재시간 및 분석시간 추출모듈, 분석에 필요한 과거 자료 추출모듈, 과거 플래그 추출모듈, 분석자료 입출력 모듈, 기상 변수 별 품질관리 모듈들, 플래그 부여 및 품질관리 결과 저장모듈로 이루어져 있으며 주 프로그램을 통해서 모듈들을 순차적으로 구동시키고 전체 흐름을 제어하도록 구성되어 있다.

주 프로그램과 각각의 자료처리 모듈들은 IDL (Interactive Data language) 코드로 작성되어 있으며 일정한 시간마다 자동으로 실행시키는 데몬(crontab)에 등록된 쉘 스크립트를 통하여 1분마다 반복 실행되고 일시적으로 누락되었다가 과거 자료 수집에 의해 채워진 자료들을 위하여 매일 1시에 과거 24시간 동안의 자료에 대한 품질관리를 수행하도록 구성되어 있다. 이러한 품질관리시스템을 운용하기 위해서는 IDL 모듈을 컴파일하고 실행하기 위한 IDL 프로그램이 설치된 리눅스 운영체제가 필요하다. 그러므로 QCIMS는 리눅스 서버에 구축되어 있고 관측 자료가 수집되는 윈도우 서버의 원시자료 저장장치와는 서로 공유되어 있다.

2.2 관측 자료 수집

강수량은 1분 누적값과 일 누적값이 측정되고 강수량을 제외한 나머지 기상 변수들은 1분 동안 최대, 최소, 평균값이 측정되어 데이터로거에 저장되고 무선모뎀을 통하여 사업단에 위치한 관측 자료 수집 서버

Fig. 2. Flow chart of the quality check module for each meteorological variable.

로 1분마다 전송된다. 만약 통신 장애로 관측 자료의 전송이 원활하게 이루어지지 않을 경우 다음시간 관측 자료가 전송되기 전까지 반복해서 재전송 시도를 하게 되고, 통신 장애가 1분 이상 지속되어 해당시간 관측 자료가 누락된 경우에는 매 시간 그리고 매일 자정에 수행되는 과거 자료 수집을 통해 누락된 자료를 시간 순으로 다시 수집한다.

전송된 관측 자료는 윈도우 운영체제에서 구동되는 자료 수집 프로그램에 의해 실시간으로 관측소 별 폴더에 텍스트 파일 형태의 원시자료로 저장된다. 원시 자료는 하루에 한 개의 파일이 생성되며 각 파일은 00시부터 23시 59분까지 측정된 자료를 포함하고 있다.

2.3 품질관리 알고리즘

관측 자료의 품질관리에는 시·공간적으로 다양한 방법들이 존재하는데 QCIMS의 품질관리 알고리즘은 다음 두 가지 조건을 만족하는 방법들로 구성하였다. 첫째, 단일 지점의 관측 자료만으로 수행 가능하여야 한다. 둘째, 1분 간격으로 수집되는 관측 자료에 대하여 실시간으로 수행 가능하여야 한다. 기존의 연구에서 제시된 품질검사 방법들 중 위의 조건을 충족하는 방법에는 결측자료검사, 물리한계검사, 단계검사, 내적일치성검사, 지속성검사, 기후범위검사가 있다(Shafer *et al.*, 2000; Vejen, 2002; WMO, 2004; KMA, 2006). 알고리즘의 기본적인 흐름은 Fig. 2와 같이 결측자료

Fig. 3. Time series of 1-minute averaged (a) air temperature ($^{\circ}\text{C}$), (b) air pressure (hPa), (c) relative humidity (%), (d) wind speed (m s^{-1}) for the period from 1 January 2000 to 31 December 2011 at the Incheon Site.

검사를 가장 우선적으로 수행하고, 물리한계검사, 단계검사, 내적일치성검사, 지속성검사, 기후범위검사를 순차적으로 수행하도록 구성되어 있다. QCIMS의 품질관리 모듈은 기상 변수마다 독립적으로 구성되어 있다. 품질관리 과정 중 어느 단계에서든 관측 자료가 ‘결측’ 또는 ‘오류’로 판단되면 더 이상 다음단계의 품질관리 과정을 진행하지 않으며 이후의 품질관리 단계들은 ‘미검사’로 처리하고 해당 기상 변수에 대한 품질관리 수행을 종료하도록 알고리즘이 구성되어 있다. 또한 ‘의심’으로 판단된 자료에 대해서는 비실시간 관리자 품질관리를 통하여 ‘정상’ 또는 ‘오류’

Table 1. Quality control methods applied to each meteorological variable.

Meteorological variable	Missing value test	Physical limit test	Step test	Internal consistency test	Persistence test	Climate range test
Air Temperature	O	O	O	-	O	O
Air pressure	O	O	O	-	O	O
Relative Humidity	O	O	O	-	O	-
Wind Speed	O	O	O	O	O	-
Wind Direction	O	O	-	O	O	-
Precipitation	O	O	-	-	-	-

Fig. 4. Probability distribution of 1-minute averaged (a) air temperature ($^{\circ}\text{C}$), (b) air pressure (hPa), (c) relative humidity (%), (d) wind speed (m s^{-1}) for the period from 1 January 2000 to 31 December 2011 at the Incheon Site.

로 최종 판단한다.

기상 변수에 따라 시간적 연속성, 측정 범위 등의 특성이 다르기 때문에 품질관리 방법 역시 다르게 적용되었다. QCIMS에서 기상 변수별로 적용된 품질관리 방법을 Table 1에 나타내었다. 결측자료검사와 물리한계검사는 모든 기상 변수에 적용되었고 단계검사는 시간에 대하여 연속적이지 않은 강수량과 360° 와 0° 사이에 관측값의 변화가 불연속적인 풍향을 제외한 기온, 기압, 습도, 풍속에 적용되었다. 시계열 분석 결과를 나타낸 Fig. 3을 보면 네 가지 기상 변수 모두 시간적으로 연속성이 있음을 알 수 있다. 내적일치성검사는 서로 관련 있는 기상 변수인 풍향과 풍속에 적용되었으며 지속성검사는 시간에 대하여 연속적이지 않은 강수량을 제외한 나머지 기상 변수들에 적용되었다. 기후범위검사는 관측값이 일정한 범위 안에서 대칭 형태의 확률분포를 보이는 기상 변수에 적용할 수 있다. 따라서 기온, 기압, 습도, 풍속의 전체 기간에 대한 확률분포를 분석하고 한 쪽으로 치우친 분포를 보이는 습도와 풍속을 제외한 대칭적인 분포를 보이는 기온과 기압에 기후범위검사가 적용되었다 (Fig. 4). 여기서 확률변수의 계급간격은 모두 동일하게 1로 설정하였다. 기온의 확률분포는 두 개의 첨두를 가지고 있는데 이는 기온이 계절에 따라 관측 자료의 범위가 달라지는 연변동을 하며 특히 극값이 나타나는 여름과 겨울의 영향이 크기 때문으로 생각된다.

이렇게 만들어진 QCIMS의 품질관리 알고리즘의 구성과 문턱값을 품질관리단계별로 WMO (2004) 및 KMA (2006)와 비교하고 그 내용을 Table 2에 요약하여 정리하였다.

2.3.1 결측자료검사

WISE 복합기상센서 관측망에서는 관측 장비의 장애로 관측이 이루어지지 않았거나 통신 장애로 인해 관측 자료가 수집되지 않은 경우 해당 관측값은 결측코드(-999)로 저장된다. WMO (2004)에서는 결측코드에 대해서는 언급하지 않았으나 결측 발생 시 결측의 이유를 전송하도록 하고 있으며 KMA (2006)에서는 '9999' 또는 '99999'를 결측코드로 사용하고 있다. 본 시스템에서는 결측자료검사에서 KMA (2006)와 동일하게 관측값이 결측코드와 동일하면 '결측'으로 판단한다.

2.3.2 물리한계검사

모든 기상 관측 장비들은 물리적으로 측정 가능한 한계를 가지고 있다. WMO (2004)에서는 이러한 기상 관측 장비의 측정 범위를 물리한계검사의 문턱값으로 사용하여 문턱값 범위를 벗어나는 관측 자료들은 사용하지 않도록 권고하고 있다. 이에 따라 KMA (2006)에서는 AWS의 관측 범위를 고려하여 기온은 $-33\sim 40^{\circ}\text{C}$, 기압은 $886\sim 1042\text{ hPa}$, 습도는 $0.1\sim 100\%$, 풍속은 $0\sim 42\text{ m s}^{-1}$, 풍향은 $0.5\sim 360^{\circ}$, 강수량은 $0\sim 1000\text{ mm}$ 를 문턱값으로 사용하고 있다. 강수량은 한 시간 누적강수를 대상으로 하며 나머지 기상 변수들은 1분 평균값을 대상으로 한다. 이 중 풍향의 문턱값이 0.5° 부터 시작되는 것은 알고리즘 개발 당시 관측자료의 형태에 따른 것으로 기상청의 경우 북풍일 때 360° , 바람이 없을 때 0° 로 표기하였기 때문이다. 본 시스템에서는 WISE 복합기상센서 관측망의 관측 장비들의 측정 범위를 고려하여 기온은 $-50\sim 60^{\circ}\text{C}$, 기압은 $300\sim 1200\text{ hPa}$, 습도는 $0\sim 100\%$, 풍속은 $0\sim 75\text{ m s}^{-1}$, 풍향은 $0\sim 360^{\circ}$, 강수량은 $0\sim 30\text{ mm}$ 를 문턱값으로 사용하여 그 범위를 벗어나면 '오류'로 판단하고 범위에 포함되면 '정상'으로 판단한다. 강수량은 1분 누적강수를 대상으로 하고 있으며 나머지 기상 변수들은 1분 평균값을 대상으로 하고 있다.

2.3.3 단계검사

기상 관측 자료의 단위시간 변동량에는 한계가 있으며 기상 변수의 종류에 따라 서로 다른 단위시간 변동량을 가지고 있다. 단계검사에서는 과거 자료의 통계분석을 통해 결정된 최대 단위시간 변동량을 문턱값으로 사용하여 관측 자료의 정상 여부를 판단한다. WMO (2004)에서는 '의심'의 문턱값으로 기온은 3°C , 기압은 0.5 hPa , 습도는 10% , 풍속은 10 m s^{-1} 을,

Table 2. Comparison of quality control methods of WMO, KMA (RQMOD), and WISE (QCIMS). ‘T’, ‘P’, ‘RH’, ‘WS’, ‘WD’, and ‘PR’ stand for air temperature, air pressure, relative humidity, wind speed, wind direction, and precipitation rate, respectively.

Method	WMO	KMA (RQMOD)	WISE (QCIMS)
Missing value test	-	Missing value 9999 or 99999	Missing value -999
Physical limit test	Reject The measurement range of a pertinent sensor	Erroneous T (−33~40°C), P (886~1042 hPa), RH (0.1~100%), WS (0~42 m s ^{−1}), WD (0.5~360°), PR (0~1000 mm h ^{−1})	Erroneous T (−50~60°C), P (300~1200 hPa), RH (0~100%), WS (0~75 m s ^{−1}), WD (0~360°), PR (0~30 mm min ^{−1})
Step test	Doubtful T (3°C), P (0.5 hPa), RH (10%), WS (10 m s ^{−1})	-	Doubtful T (0.7°C), P (0.4 hPa), RH (5.9%), WS (4.6 m s ^{−1})
	Erroneous P (2 hPa), RH (15%), WS (20 m s ^{−1})	Erroneous T (3°C), P (0.5 hPa), RH (10%), WS (20 m s ^{−1})	Erroneous P (2 hPa), RH (15%), WS (20 m s ^{−1})
Internal consistency test	Erroneous WS = 0 m s ^{−1} and WD ≠ 0° WS ≠ 0 m s ^{−1} and WD = 0°	Erroneous WS = 0 m s ^{−1} and WD ≠ 0° WS ≠ 0 m s ^{−1} and WD = 0°	Erroneous WS = 0 m s ^{−1} and WD ≠ 0° WS ≠ 0 m s ^{−1} and WD = 0°
Persistence test	Doubtful T (0.1°C), P (0.1 hPa), RH (1%), WS (0.5 m s ^{−1}), WD (0.1°) ※ Apply when the humidity is less than 95% and 10-minute average wind speed is greater than 0.1 m s ^{−1}	-	Doubtful T (0.1°C), P (0.1 hPa), RH (1%), WS (0.5 m s ^{−1}), WD (0.1°) ※ Apply when the humidity is less than 95% and 10-minute average wind speed is greater than 0.1 m s ^{−1}
	-	Erroneous T (0.1°C), P (0.1 hPa), RH (1%), WS (0.2 m s ^{−1}), WD (10°)	-
Climate range test	-	Doubtful T (2σ), P (2σ)	Doubtful T (3σ), P (3σ)
	Erroneous T (−80~60°C), P (500~1100 hPa), RH (0~100%), WS (0~75 m s ^{−1}), WD (0~360°), PR (0~40 mm)	Erroneous T (3σ), P (3σ)	Erroneous T (6σ), P (6σ)

‘오류’의 문턱값으로 기압은 2 hPa, 습도는 15%, 풍속은 20 m s^{−1}을 제시하여 두 가지 기준을 사용하도록 권고하고 있다. KMA (2006)에서는 ‘오류’ 판단의 기준만 사용하고 있으며 기온은 3°C, 기압은 0.5 hPa, 습도는 10%, 풍속은 20 m s^{−1}을 문턱값으로 제시하고 있다. 본 시스템에서는 WMO (2004)의 권고에 따라 문턱값을 ‘의심’과 ‘오류’ 두 가지로 구분하였으며, 좁은 범위인 ‘의심’의 문턱값의 경우 WMO (2004)에서 제시한 값은 1분 평균자료를 고려했을 때 다소 큰 값으로 생각되어 과거 자료의 통계분석을 통해 산출한 값으로 수정하여 사용하였고 더 넓은 범위인 ‘오류’의 문턱값은 WMO (2004)에서 제시한 값을 그대로

사용하였다. 수정된 ‘의심’의 문턱값은 기온이 0.7°C, 기압은 0.4 hPa, 습도는 5.9%, 풍속은 4.6 m s^{−1}이며 이전 시간 관측값과 현재 관측값의 차이가 ‘오류’의 문턱값을 초과하면 ‘오류’로, ‘오류’의 문턱값 이하이고 ‘의심’의 문턱값을 초과하면 ‘의심’으로, ‘의심’의 문턱값 이하이면 ‘정상’으로 판단한다.

2.3.4 내적일치성검사

내적일치성검사에서는 물리적으로 서로 밀접하게 연관된 기상 변수 사이에서 자연적으로 발생할 수 없는 측정값을 다양한 검사기준을 통하여 판단한다. WMO (2004)와 KMA (2006)에서는 풍향과 풍속, 온

도와 이슬점온도, 강수량과 강수유무 등 서로 짝을 이루는 기상 변수들에 대하여 다양한 검사기준들을 제시하고 있으나 Table 2에서는 본 시스템에서 짝을 이루는 기상 변수인 풍향과 풍속에 대해서만 비교하여 정리하였다. 본 시스템에서 사용된 풍향과 풍속에 대한 내적일치성검사의 기준은 WMO (2004) 및 KMA (2006)와 동일하며 풍향이 '0'일 때 풍속이 '0'이 아니거나 풍향이 '0'이 아닐 때 풍속이 '0'이면 풍향과 풍속을 모두 '오류'로 판단한다.

2.3.5 지속성검사

기상 관측 장비에 장애가 발생하여 작동하지 않으면 일정기간 동안 관측값이 전혀 변화하지 않는 경우가 발생한다. 지속성검사에서는 최소변동량을 문턱값으로 사용하여 이러한 기상 관측 장비의 장애를 판단한다. WMO (2004)에서는 과거 60분에 대하여 기온은 0.1°C , 기압은 0.1 hPa , 습도는 1%, 풍속은 0.5 m s^{-1} , 풍향은 0.1° 를 문턱값으로 제시하고 있으며 '의심'의 판단 기준으로 사용하도록 권고하고 있다. 또한 습도의 경우는 95% 이하일 때, 풍속은 10분 평균 풍속이 0.1 m s^{-1} 이상일 때만 지속성검사를 수행하도록 권고하고 있다. KMA (2006)에서는 과거 60분에 대하여 기온은 0.1°C , 기압은 0.1 hPa , 습도는 1%, 풍속은 0.2 m s^{-1} , 풍향은 10° 를 문턱값으로 제시하고 있으며 '오류'의 판단 기준으로 사용하고 있다. 본 시스템에서는 WMO(2004)의 권고를 동일하게 적용하여 60분 동안 변동량의 총합이 문턱값 이하이면 '오류', 초과하면 '정상'으로 판단한다.

2.3.6 기후범위검사

기상 관측 자료는 관측소 위치에 따른 기후조건이나 계절 변화 등에 따라서 일정한 범위를 가지는데 기후범위검사에서는 관측 지점의 기후에 따른 관측값의 범위를 문턱값으로 사용하여 관측 자료의 정상 여부를 판단한다. WMO (2004)에서는 '오류'의 문턱값만 권고하였으며 전 지구적으로 사용할 수 있도록 기온은 $-80\sim 60^{\circ}\text{C}$, 기압은 $500\sim 1100\text{ hPa}$, 습도는 0~100%, 풍속은 $0\sim 75\text{ m s}^{-1}$, 풍향은 $0\sim 360^{\circ}$, 강수량은 $0\sim 40\text{ mm}$ 를 제시하였다. KMA (2006)에서는 기후범위검사의 대상을 기온으로만 한정하고 있으며 지난 3년간의 관측 자료의 평균과 표준편차(σ)를 이용한 통계분석을 바탕으로 지점별, 월별 평균값의 $\pm 2\sigma$ 이내이면 '정상', $\pm 2\sigma$ 를 초과하면 '의심', $\pm 3\sigma$ 를 초과하면 '오류'로 판단한다. 본 시스템에서는 기온과 기압에 대하여 기후범위검사를 수행하며 관측 자료의 통계분석을 통해 KMA (2006)의 문턱값을 수정하여 월별 평균값의 $\pm 3\sigma$ 이내이면 '정상', $\pm 3\sigma$ 를 초과하면 '의심', $\pm 6\sigma$ 를 초과하면 '오류'로 판단한다.

2.4 품질관리 플래그

품질관리 결과는 숫자 형태의 플래그로 부여되는데 플래그 종류는 모두 5가지로 '0'은 정상, '1'은 의심, '2'는 오류, '8'은 미검사, '9'는 결측의 의미를 가지며 각 기상 변수 별로 5자리 숫자로 저장된다. 첫 번째 자리는 결측자료검사와 물리한계검사에 대한 정보를 담고 있으며 정상(0), 오류(2), 결측(9) 플래그로 구분된다. 두 번째 자리는 단계검사에 대한 정보를 담고 있으며 정상(0), 의심(1), 오류(2), 미검사(8) 플래그로 구분된다. 세 번째 자리는 내적일치성검사에 대한 정보를 담고 있으며 정상(0), 오류(2), 미검사(8) 플래그로 구분된다. 네 번째 자리는 지속성검사에 대한 정보를 담고 있으며 정상(0), 의심(1), 미검사(8) 플래그로 구분된다. 마지막 다섯 번째 자리는 기후범위검사에 대한 정보를 담고 있으며 정상(0), 의심(1), 오류(2), 미검사(8) 플래그로 구분된다.

3. 문턱값 산출을 위한 통계분석

WISE 복합기상센서 관측망은 지난 2014년 3월 중 설치가 완료되어 통계분석을 수행하기에는 과거 자료 수집 기간이 너무 짧기 때문에 관측망이 설치된 인천의 대표 관측소인 인천기상대 관측 자료를 사용하여 통계분석을 수행하였다. 분석에는 2000년부터 2011년까지 12년 동안의 AWS 1분 관측 자료를 사용하였고 보다 정확한 통계분석을 위하여 분석을 수행하기 전 관측 자료에 대하여 물리한계검사와 단계검사를 이용하여 기본적인 품질관리를 수행하였다.

3.1 단계검사 문턱값 산출

관측 자료의 평균시간 변화에 따라 단위시간 변동량이 어떻게 변화하는지 분석하기 위하여 단계검사를 수행하는 기상 변수들에 대하여 평균시간을 1분에서 60분까지 변화시키면서 평균시간에 따른 단위시간 변동량의 표준편차(σ)를 산출하고 둘 사이의 상관계수를 구하여 Table 3에 나타내었다. 풍속의 경우는 평균시간과 단위시간 변동량의 표준편차(σ)의 상관관계가 0.58로 낮지만 나머지 기상 변수들은 0.99의 높은 상관성을 보이고 있다. 평균시간에 따른 단위시간 변동량의 표준편차(σ)를 정규화시켜 나타낸 Fig. 5를 보면 풍속의 경우 평균시간의 변화에 따른 단위시간 변동량의 표준편차(σ)의 변화가 거의 나타나지 않지만 나머지 기상 변수들은 평균시간의 증가에 따라 단위시간 변동량의 표준편차(σ)가 선형적으로 증가하며 서로 밀접한 관련을 가지고 있음을 알 수 있다.

따라서 1분의 관측 간격을 가지는 복합기상센서의 단계검사 문턱값을 산출하기 위하여 2000년부터 2011년까지 기온, 기압, 습도, 풍속의 1분 평균 자료를 사

Table 3. Standard deviation of ΔT , ΔP , ΔRH , ΔWS according to each averaging interval for the period from 1 January 2000 to 31 December 2011 at the Incheon Site. Correlation coefficients between the averaging interval and standard deviation are listed.

Average time (minute)	Air temperature ($^{\circ}\text{C}$)	Air pressure (hPa)	Relative humidity (%)	Wind speed (m s^{-1})
1	0.06	0.04	0.59	0.64
2	0.09	0.07	0.70	0.59
5	0.16	0.08	1.02	0.53
10	0.23	0.12	1.52	0.52
30	0.47	0.25	3.06	0.58
60	0.80	0.43	4.79	0.68
Correlation coefficient	0.99	0.99	0.99	0.58

Fig. 5. Standard deviation of ΔT , ΔP , ΔRH , ΔWS according to each averaging interval normalized by those of 60-minute average value for the period from 1 January 2000 to 31 December 2011 at the Incheon Site.

용하여 단위시간 변동량을 산출하고 이 변동량을 확률변수로 하는 확률분포를 분석하여 Fig. 6에 나타내었다. 각 기상변수별 변동량의 범위를 고려하여 기온과 기압은 0.1, 습도와 풍속은 0.5로 확률변수의 계급간격을 설정하였다. 네 가지 기상 변수 모두 변동량의 평균 값인 '0'에서 확률분포의 첨두가 나타나며 첨두에서 멀어질수록 급격히 감소하는 형태의 분포를 보이고 첨두에서의 확률밀도는 정규분포에 비해 매우 높은 값을 보이므로 단위시간에 대한 변동이 거의 없음을 알 수 있었다.

본 시스템에서는 99.99% 이상의 관측 자료가 포함되는 범위를 단계검사의 문턱값으로 결정하였으며 이에 따라 기온은 0.7°C , 기압은 0.4 hPa , 습도는 5.9% , 풍속은 4.6 m s^{-1} 로 문턱값이 최종 결정되었다.

3.2 기후범위검사 문턱값 산출

기후범위검사의 문턱값 산출을 위하여 먼저 기온과

Fig. 6. Probability distribution of differences between the values at $t+1$ minute and those at t minute of (a) air temperature ($^{\circ}\text{C}$), (b) air pressure (hPa), (c) relative humidity (%), (d) wind speed (m s^{-1}) for the period from 1 January 2000 to 31 December 2011 at the Incheon Site.

기압의 월별 확률분포를 분석하였고 이를 Fig. 7과 Fig. 8에 각각 나타내었다. 확률변수의 계급간격은 두 변수 모두 동일하게 1로 설정하였다. 기온의 경우 전체기간에 대한 확률분포 분석에서는 두 개의 첨두를 보이며 확률변수의 범위가 넓게 나타나지만, 월별 확률분포에서는 첨두가 하나로 나타나고 확률변수의 분포 범위가 좁아졌다(Figs. 4a, 7). 또한 기온과 기압 모두 겨울철보다는 여름철에 표준편차가 작아지는 것을 알 수 있는데 이는 고온 다습하고 기압경도력이 약한 북태평양 기단의 영향을 받는 우리나라 여름철의 기후 특성 때문으로 생각된다. 이를 통하여 기후범위검

Fig. 7. Monthly probability distribution of 1-minute averaged air temperature (°C) for the period from 1 January 2000 to 31 December 2011 at the Incheon Site. 'm' is mean of data and 'σ' is standard deviation of data.

Fig. 8. Monthly probability distribution of 1-minute averaged air pressure (hPa) for the period from 1 January 2000 to 31 December 2011 at the Incheon Site. 'm' is mean of data and 'σ' is standard deviation of data.

Fig. 9. Time series of 1-minute averaged air temperature ($^{\circ}\text{C}$) on April in 2002 (thin solid line) and 2005 (thick solid line) at the Incheon Site. 3σ , 4σ , 5σ , and 6σ during the 12 years from 2000 to 2011 are denoted.

Fig. 10. Time series of 1-minute averaged air pressure (hPa) on September in 2000 (thin solid line), 2003 (thick solid line), and 2004 (dotted line) at the Incheon Site. 3σ , 4σ , 5σ , and 6σ during the 12 years from 2000 to 2011 are denoted.

사의 문턱값 산출을 위한 통계분석은 전체기간을 대상으로 하는 것보다는 월별로 이루어지는 것이 타당하다는 결론을 내렸다. 이에 따라 본 시스템에서는 월별 기온과 기압의 평균과 표준편차(σ)를 각각 산출하고 시계열분석을 통한 정상적인 분석과 확률밀도 분석을 통한 정량적인 분석을 통하여 기후범위검사의

문턱값을 산출하였다.

기온의 월별 누적 시계열분포를 분석한 결과 대부분의 관측값이 $\pm 3\sigma$ 범위 안에 들지만 Fig. 9에 나타난 것과 같이 이상고온이 발생하였을 경우 $\pm 3\sigma$ 를 초과하는 값이 나타나기도 하였는데 2005년 4월 30일 인천 기상대는 평균기온 23.5°C 최고기온 32.7°C 로 전

Table 4. Monthly threshold values and their probability included in the climate doubtful and erroneous threshold for air temperature.

Month	Doubtful Threshold (3σ)			Erroneous Threshold (6σ)		
	Min ($^{\circ}\text{C}$)	Max ($^{\circ}\text{C}$)	Probability	Min ($^{\circ}\text{C}$)	Max ($^{\circ}\text{C}$)	Probability
1	-15.7	12.4	0.9968	-29.8	26.5	1
2	-12.2	14.6	0.9949	-25.6	28.0	1
3	-6.8	18.2	0.9985	-19.3	30.7	1
4	-0.4	23.7	0.9941	-12.5	35.7	1
5	5.7	28.3	0.9964	-5.6	39.6	1
6	11.7	30.7	0.9986	2.2	40.3	1
7	16.0	32.2	0.9945	7.9	40.3	1
8	16.9	33.7	0.9957	8.5	42.1	1
9	11.4	31.8	0.9978	1.2	42.1	1
10	3.3	28.0	0.9959	-9.0	40.3	1
11	-6.3	22.9	0.9998	-20.9	37.5	1
12	-14.1	16.0	0.9995	-29.2	31.1	1

Table 5. Monthly threshold values and their probability included in the climate doubtful and erroneous threshold for air pressure.

Month	Doubtful Threshold (3σ)			Erroneous Threshold (6σ)		
	Min (hPa)	Max (hPa)	Probability	Min (hPa)	Max (hPa)	Probability
1	1007.5	1039.4	0.9995	991.5	1055.4	1
2	1004.8	1038.6	0.9927	987.8	1055.5	1
3	999.6	1035.8	0.9951	981.4	1054.0	1
4	996.5	1030.7	0.9977	979.4	1047.8	1
5	994.5	1025.1	0.9940	979.2	1040.5	1
6	992.9	1020.1	0.9967	979.3	1033.7	1
7	991.6	1018.5	0.9959	978.2	1032.0	1
8	995.0	1020.1	0.9954	982.4	1032.6	1
9	998.8	1026.5	0.9905	984.9	1040.3	1
10	1004.3	1031.3	0.9954	990.8	1044.9	1
11	1005.9	1036.3	0.9971	990.7	1051.5	1
12	1007.2	1039.3	0.9968	991.1	1055.3	1

날인 29일 보다 평균기온은 6.9°C , 최고기온은 9.5°C 더 높게 나타났다(KMA, 2005). 기압의 경우도 대부분의 관측값이 $\pm 3\sigma$ 범위 안에 들지만 Fig. 10에 나타난 것과 같이 여름철 태풍이 발생하였을 경우 $\pm 3\sigma$ 를 초과한 값이 나타나기도 하였다. 그림에서 기압이 급감한 시기는 각각 2000년 ‘PRAPIROON’과 ‘SAOMAI’, 2003년 ‘MAEMI’, 2004년 ‘SONGDA’의 발생 시점이다.

본 시스템에서는 이러한 이상기온이나 태풍을 고려하여 월별 평균값의 $\pm 3\sigma$ 와 $\pm 6\sigma$ 를 각각 ‘의심’과 ‘오류’의 문턱값의 기준으로 결정하였다. Table 4와 Table 5는 각각 기온과 기압에 대한 월별 의심과 오류의 문턱값을 나타낸 것이다. 기온과 기압의 문턱값 범위를 분석해 보면 여름에는 좁고 겨울에는 넓게 나타나는

것을 알 수 있는데 이는 문턱값 범위가 Fig. 7 과 Fig. 8에서 분석한 평균과 표준편차에 의해 결정되기 때문이다. 기온과 기압의 관측값이 ‘의심’의 문턱값 범위에 포함될 확률의 평균값은 각각 0.997과 0.996으로 나타나고 ‘오류’의 문턱값 범위에 포함될 확률의 평균값은 두 기상 변수 모두 1이다.

4. 품질관리 수행 결과

WISE 복합기상센서 관측망은 지난 2014년 3월에 설치가 완료되었으며 지난 4월 한 달 동안 관측된 자료에 대하여 Table 2에 제시한 RQMOD와 QCIMS의 알고리즘을 각각 적용하여 품질관리를 수행하였다. 또한, 두 기관의 품질관리 과정에서 ‘의심’ 또는 ‘오류’

로 판정을 받은 비율(R_p)을 산출하여 비교하였다.

분석기간 동안 관측 자료 수집률은 98.53%를 기록했으며 RQMOD의 알고리즘을 적용한 경우 각 기상 변수별 R_p 는 기온이 7.27%, 기압은 0.04%, 습도는 2.87%, 풍속은 0.1%, 풍향은 0.19%, 강수량은 0%로 나타났고 QCIMS의 알고리즘을 적용한 경우는 기온이 0.29%, 기압은 0.04%, 습도는 2.87%, 풍속은 0.1%, 풍향은 0.07%, 강수량은 0%로 나타났다.

두 가지 알고리즘 모두 기압, 습도, 풍속, 강수량에 대해서는 동일한 결과를 보이고 있다. 풍향의 경우 약간의 차이를 보였는데 RQMOD의 알고리즘에서는 물리한계검사와 내적일치성검사에서 각각 0.15%와 0.04%가 '오류' 판정을 받아 R_p 가 총 0.19%로 나타났고, QCIMS의 알고리즘의 경우 내적일치성검사에서 0.07%가 '의심' 판정을 받아 R_p 가 0.7%로 나타났다. 물리한계 검사에서 차이를 보이는 것은 문턱값의 차이로 판단되며, 내적일치성검사에서 RQMOD가 QCIMS보다 오류가 적은 이유는 물리한계검사에서 '오류'로 판단된 자료에 대해서는 내적일치성검사를 수행하지 않기 때문이다. 하지만 기온의 경우는 두 알고리즘에 의한 차이가 크게 나타났다. RQMOD의 알고리즘에서는 기후범위검사에서 6.98%가 '의심', 0.26%가 '오류' 판정을 받았고 단계검사에서 0.03%가 '미검사' 판정을 받아 R_p 가 총 7.27%로 높게 나타난 반면, QCIMS의 알고리즘의 경우 기후범위검사에서 0.26%가 '의심' 판정을 받고 단계검사에서 0.03%가 '미검사' 판정을 받아 총 0.29%로 낮은 값을 보였다.

위의 결과를 분석해보면 풍향의 경우는 센서의 측정 범위의 차이에 의해 물리한계검사에서 차이를 보인 것이다. 하지만 기온의 경우 RQMOD의 알고리즘에서 사용하고 있는 기후범위검사 '의심'의 문턱값인 $\pm 2\sigma$ 가 기후범위를 충분히 고려하지 못하고 있으며 통계분석을 통하여 수정된 문턱값을 적용한 본 시스템에서는 이러한 부분을 고려할 수 있도록 개선되었다.

5. 요약 및 결론

복잡한 대도시에서 기상 관측을 위하여 유용한 장비인 복합기상센서의 관측 자료에 대한 품질관리를 위하여 기존의 AWS 관측 자료 품질관리시스템들의 알고리즘을 분석하고 기상 변수들의 시계열 및 통계분석을 통하여 품질관리 알고리즘들을 수정 및 보완하여 QCIMS를 개발하였다. 본 시스템의 알고리즘 개발을 위해서 WMO(2004)와 우리나라 기상청의 품질관리 알고리즘을 비교하였으며(KMA, 2006), 기상 변수들의 시계열분석 및 통계분석을 위해서 2000년부터 2011년까지 인천기상대의 기온, 기압, 습도, 풍속, 풍향, 강수량 자료를 사용하였다.

본 시스템에 사용된 품질관리는 기본적으로 결측자료검사, 물리한계검사, 단계검사, 내적일치성검사, 지속성검사, 기후범위검사 순으로 진행되고 품질관리 과정 중 어느 단계에서든 관측값이 '오류' 판정을 받게 되면 이후 과정에서는 모두 '미검사' 판정을 받게 되고 해당 관측값에 대한 품질관리를 종료하게 된다. 또한, 기상 변수의 특성에 따라 적용되는 품질관리 과정이 다르게 적용되었다. 즉, 결측자료검사와 물리한계검사는 모든 기상 변수에 적용되었고 단계검사는 기온, 기압, 습도, 풍속 내적일치성검사는 풍향과 풍속 지속성검사는 기온, 기압, 습도, 풍속, 풍향 기후범위검사는 기온과 기압에 각각 적용되었다.

이 중 단계검사와 기후범위검사의 문턱값을 결정하기 위하여 기온, 기압, 습도, 풍속의 1분 평균 자료를 분석하였다. 풍속을 제외한 나머지 기상 변수들은 관측 간격 즉, 관측 자료의 평균시간과 변동량의 표준편차(σ)가 선형적으로 0.99의 높은 상관성을 보인다. 따라서 복합기상센서의 관측 간격을 고려하여 1분 평균 자료에 대한 최대변동량을 산출하여 단계검사 '의심'의 문턱값으로 기온 0.7°C, 기압 0.4 hPa, 습도 5.9%, 풍속 4.6 m s⁻¹가 결정되었으며, 이에 따라 기존의 연구(WMO, 2004; KMA, 2006)에서 제시된 단계검사 문턱값에 비해 작아졌다. 관측 자료의 시계열 및 확률분포 분석을 통해 기온과 기압을 기후범위검사 대상으로 결정하고 월별 평균, 표준편차(σ), 누적 시계열분포를 분석하여 월별 평균값의 $\pm 3\sigma$ 와 $\pm 6\sigma$ 를 각각 '의심'과 '오류'의 문턱값의 기준으로 결정하였으며, 이에 따라 봄철 이상고온과 태풍 등으로 인한 급격한 관측 자료의 변화를 고려할 수 있도록 기존의 연구(KMA, 2006)에서 제시된 문턱값에 비해 그 범위가 넓어졌다.

이상에서 기존의 품질관리 알고리즘들과 과거 AWS 자료의 분석을 통해서 수정 및 보완된 복합기상센서 관측 자료에 대한 품질관리시스템을 개발하였다. 하지만 본 시스템은 단일 지점에서 수행 가능한 시간의 흐름에 따른 1차원적인 품질관리 알고리즘만으로 구성되어 있으며 문턱값 산출을 위해 관측지점들과 인접해 있는 인천기상대의 관측자료를 사용했다는 한계를 가지고 있다. 따라서 추후에 공간성 검사 등의 2차원적인 품질관리 알고리즘을 추가하고 각 지점별 관측 자료의 통계분석을 통하여 품질검사 문턱값을 최적화할 예정이다. 본 품질관리시스템을 통해 개발된 품질관리 프로그램은 한국저작권위원회(프로그램명: 복합기상센서 관측자료 품질관리 프로그램, 등록번호: C-2014-013696)에 등록되어 있으며 공동 활용이 가능하다.

감사의 글

본 연구는 기상청 차세대도시농림융합스마트기상서

비스개발(WISE) 사업의 지원으로 수행되었습니다(153-3100-3133-302-350).

REFERENCES

- Basara, J. B., and Coauthors, 2010: The Oklahoma city micronet. *Meteor. Appl.*, **18**, 252-261.
- Fiebrich, C. A., and K. C. Crawford, 2001: The impact of unique meteorological phenomena detected by the Oklahoma mesonet and ARS micronet on automated quality control. *Bull. Amer. Meteor. Soc.*, **82**, 2173-2188.
- Heo, B.-H., J.-A. Lee, Y.-O. Chu, J.-H. Kim, N.-C. Park, J.-Y. Cho, S.-J. Oh, M.-S. Noh, and Y.-J. Lee, 2005: Statistical procedure of AWS observation data to determine the threshold value for the RQMOD(real-time quality control system for meteorological observation data). *Proceedings of the autumn meeting of Korean Meteorological Society*, 390-391.
- Kim, G.-H., D.-G. Lee, S.-S. Shin, K.-Y. Nam, and B.-C. Choi, 2014: Effects of urban green using the integrated meteorological sensors. *Proceedings of the spring meeting of Korean Meteorological Society*, 120-121.
- KMA, 2005: Monthly weather report 2005. 4., 11-1360000-000002-06, 54 pp.
- _____, 2006: Real-time quality control system for meteorological observation data (I) Application, 11-1360000-000206-01 (Tech. Note 2006-2), 157 pp.
- Shafer, M. A., C. A. Fiebrich, D. S. Arndt, S. E. Fredrickson, and T. W. Hughes, 2000: Quality assurance procedures in the Oklahoma Mesonet network. *J. Atmos. Oceanic Technol.*, **17**, 474-494.
- Vejen, F., C. Jacobsson, U. Fredriksson, M. Moe, L. Andresen, E. Hellsten, P. Rissanen, T. Palsdottir, and T. Arson, 2002: Quality control of meteorological observations Automatic methods used in the Nordic countries. *KLIMA Report 8/2002*, 109 pp.
- Wade, C. G., 1987: A quality control program for surface mesometeorological data. *J. Atmos. Oceanic Technol.*, **4**, 435-453.
- Winning, Jr. T. E., B. G. Illston, M. M. Ferris, and J. B. Basara, 2007 : An overview of the Oklahoma city urban Micronet test facility. *14th Symposium on Meteorological Observation and Instrumentation*, San Antonio, TX, Amer. Meteor. Soc., 5.5
- WMO, 2004: Guidelines on quality control procedures for data from automatic weather stations. *CIMO/OPAG-SURFACE/ET ST&MT-1/Doc. 6.1(2)*, 10 pp.