

수원기상대 지중온도 및 토양수분 자료의 품질관리

오규림^{1*}, 이승재¹, 김 준¹, 이병렬^{1,2}, 김규랑³, 최병철³

¹국가농림기상센터, ²세계기상기구 농업기상위원회, ³국립기상연구소 응용기상연구과

Quality Control of Soil Temperature and Soil Moisture Data Measured at Suwon Weather Station

Gyu-Lim Oh^{1*}, Seung-Jae Lee¹, Joon Kim¹, Byong-Lyol Lee^{1,2}, Kyu-Rang Kim³,
Byoung-Choel Choi³

¹National Center for AgroMeteorology, ²World Meteorological Organization,

³National Institute of Meteorological Research

I. 서 언

농업기상은 기상, 수문학적 요인과 원예, 축산 및 산림을 포함하는 넓은 의미의 농업과의 상호관계에 관한 것이다. 농업기상의 목적은 기상과 농업의 상호관계를 파악하여 정의하고 대기에 관한 지식을 실제 농업경영에 적용하는 데 있다. 이러한 농업기상이 농업에 적절히 활용되기 위해서는 품질보증(QA, Quality Assurance)과 품질관리(QC, Quality Control)가 제대로 이루어진 농업기상관측자료가 필수적인 요건이다.

농업기상관측요소에 대한 QA에는 농업기상 관측장비의 운영과 감시 등이 있을 수 있다. 일반적으로 QA란, 제품을 개발하는 과정에서의 품질을 보장하기 위한 것으로, 처리 공정에 중점을 두며 결함을 방지하는 것을 목적으로 하는 사전 작업을 말한다. 이에 반해 QC는 실제 생성되는 제품의 결함 식별과 사후 수정에 초점을 맞춘다. 즉, 표출된 농업기상관측요소에 대하여 QC를 실시한다는 것은 관측요소의 오류(error)값을 걸러내고 정제된 데이터를 재표출하는 작업을 의미한다.

농업기상관측자료를 이루고 있는 요소는 연직 기온, 연직 습도, 연직 풍속 등의 대기요소와 토양수분 지중온도 등의 토양요소로 구분되어 지며, 특히 토양요소는 지상기상관측요소에는 존재하지 않는 요소이다. 그러나 국내에서 기상관측요소의 품질관리 연구는 대부분 지상기상관측요소의 기온과 같은 대기요소에 한하여 이루어지고 있으며(허복행 외, 2005; 이영섭 외, 2007; 박창용 외, 2012; 이영섭 외, 2013), 대기 외의 요소, 즉 토양요소에 대한 국내 연구는 국외 연구(Dorigo *et al.*, 2012)에 비해 매우 부족한 실정이다.

현재 기상청에서는 전국적으로 11곳에 농업기상관측관서를 운영하고 있는데 이 중 수원기상대가 기본관서로서의 역할을 하고 있다. 전국 농관지점에서 관측한 자료는 농림기상 유관기관의 웹사이트 등을 통해 제공되고 있지만 이에 대한 품질관리는 원만히 이루어지고 있다고 볼

* Correspondence to : sjlee@ncam.kr

수 없다.

따라서 본 연구에서는 수원기상대의 농업기상관측요소 중 토양요소를 중심으로 품질관리를 실시하여 토양요소에 대한 품질검사 표준을 만들고, 이를 전국 농업기상관측관서에서 참고할 수 있도록 하고자 한다.

II. 자료분석 및 사전처리

사용한 자료는 수원기상대의 총 2년 (2012년 1월~2013년 12월) 동안의 지중온도(5cm, 10cm, 20cm, 30cm) 및 토양수분(10cm, 20cm, 30cm, 50cm)이다. 대상 자료에 대하여 품질관리를 실시하기 이전에, 사전 자료 처리과정을 거쳤는데, 매일 19일 20시 자료부터 23시 자료가 97일 40시~43시로 표기되어 있었으므로 이에 대한 오류를 수정하였다.

지중온도와 토양수분에는 각각 다른 품질관리 알고리즘을 사용하였다. 지중온도의 경우, 시계열 양상이 연직 기온의 시계열 양상과 매우 유사하므로 지중온도의 품질관리는 허복행 등 (2005)에서 기술한 기상청 실시간 품질관리 시스템(RQMOD, Real-time Quality control system for Meteorological Observation Data)을 참고하여 품질관리를 실시하였다. 토양 수분의 경우에는 온도 요소와는 달리 강수량과 밀접한 연관이 있어 강수 유무에 따른 불규칙한 변동을 보이므로, 지중온도와는 다른 품질관리 방법이 필요하다. 따라서 토양수분은 Dorigo *et al.*(2012)이 제시한 토양수분의 QC 시스템을 참고하였다.

III. 지중온도의 품질관리

3.1. 물리한계 검사 (Physical range check)

물리한계 검사는 장비의 기계적인 오차에 따른 오류 값을 제거하기 위함이다. 물리적 한계를 벗어나는 관측 자료는 결측 자료로 처리한다. 물리한계범위는 WMO (2013)에서 권고한 사항을 참고하였다. 지중온도의 물리한계범위는 $-50^{\circ}\text{C} \sim +50^{\circ}\text{C}$ 이다.

3.2. 빈도분포 검사 (Frequency distribution check)

빈도 분포 검사에서는 극단이상치를 구하여 이를 결측으로 처리하였다. 극단이상치는 상위 25%에서 하위 25%를 뺀 값인 Q3-Q1(사분위범위, IQR)를 이용하여 Q1과 Q3로부터 사분위범위의 3배 밖에 있는 값을 말한다.

3.3. 단계 검사 (Step check)

단계검사는 변동량의 한계치를 산출하여 시간의 연속적인 자료에 대하여 이상 변동이 있는지

확인하는 것이다. 변동량의 한계치는 최대변동량이며, 이는 여러 해 자료에 대해 통계분석을 통해 산출하여야 하나, 본 연구에서는 2012, 2013년 자료만을 이용하므로 2012년을 기준으로 최대변동량을 산출하되, 5cm 지중온도, 30cm 지중온도의 요소에 대해서는 예외적으로 2013년을 기준으로 적용하였다. 그리하여 한 시간 간격 변동량이 최대변동량보다 크면 의심자료(doubtful data)로 처리하였다. 의심자료 중 그 자료를 제외한 일평균 자료에서 $\pm 10^{\circ}\text{C}$ 를 벗어나면 결측으로 처리하였다.

3.4. 지속성 검사 (Flat line check)

관측자료에 오류가 있음에도 불구하고 오랜 시간 오류가 지속되면 이상값이 추출되지 않는다. 따라서 지속성 검사는 특정기간 동안 누적변동량이 최소변동량 이내로 지속될 경우를 확인하는 검사로써, 변동량이 설정된 최소변동량보다 작으면 의심자료로 처리한다. 최소변동량은 1년 이상의 자료에 대하여 해당 관측요소별, 월별로 산출하여 적용하며 0보다는 큰 최소값으로 산출한다. 하지만 지중온도의 경우 요소 특성상 매우 안정적이기 때문에 최소변화량을 적용할 수 없으므로 지속성 검사에서 제외한다.

3.5. 중앙값 필터 검사 (Median filter check)

중앙값필터 검사는 연속적인 시계열 자료에서 짧은 시간동안 급격하게 일어나는 비정상 값들을 필터링하는 것으로, 해당하는 관측값을 포함하여 홀수횟수 관측값의 중앙값과 현재 관측값의 차이가 일정 한계치를 초과하는지 확인하는 방법이다. 여기서 홀수횟수는 3회 적용하였고, 한계값은 대상 관측값이 해당하는 달의 월 단위 표준편차를 적용하였다. 중앙값필터 검사는 모든 값을 대상으로 수행할 시 정상적인 값들도 필터링되는 경향이 있었으므로 이전 검사들에서 의심자료로 분류된 값들을 위주로 적용한다.

IV. 토양수분의 품질관리

4.1 물리한계 검사 (Physical range check)

지중온도와 마찬가지로 토양수분의 물리한계 값은 WMO (2013)에 제시되어 있다. 토양수분 물리한계 값에 대한 WMO 권고기준은 0 ~ 103 g/kg 의 질량단위를 사용하고 있지만 국내 토양수분 관측자료는 %를 사용하고 있어 단위를 질량비 %로 환산한다. 따라서 0% ~ 100%를 물리한계범위로 지정한다.

4.2 지구물리학적 검사 (Geophysical check)

4.2.1 강수량 이용

토양수분은 강수량과 매우 밀접한 관계가 있다. 따라서 토양수분이 상당히 증가할 때, 해당 시간으로부터 24시간 전에 강수량이 존재하지 않으면 의심 플래그를 부여한다. 여기서 상당한 증가는 증가 시점부터 역으로 24시간 동안의 토양수분 증가량이 그 시간동안의 토양수분 표준편차의 2배를 넘을 때로 정의한다. 하지만 강수가 있었다라도 토양수분 센서가 반응하지 않는 경우를 고려하여, 센서가 반응하는 최소 강수량(P_{min})을 정의한다. P_{min} 은 센서의 깊이, 센서의 정밀도, 그리고 토양의 공극율을 곱하여 산출해낸다. 여기서 센서의 정밀도는 0.0006, 토양의 공극율은 평균값인 0.5를 사용한다. 따라서 토양수분이 상당히 증가할 때 강수량이 존재했다라도, 강수량이 P_{min} 미만이면 역시 의심자료로 처리한다. 이 방법은 토양수분이 강수량에 즉각적으로 반응한다고 가정하기 때문에 10cm 토양수분 요소에만 적용한다.

4.2.2 토양온도 이용

토양온도는 서리로 인해 토양수분이 잘못 측정됐을 경우를 걸러내기 위해 사용한다. 토양온도가 0°C미만일 경우, 이론적으로 토양수분은 존재할 수 없으므로 이때의 토양수분 값에는 의심 플래그를 부여한다.

4.3 단계 검사 (Step check)

4.3.1 튀는 값 검출 (Spike detection)

두 개의 연속한 시간에서 토양수분의 급격한 증가는 이전 값의 15%(혹은 센서 정확도의 최대값에 3배를 곱한 것)를 넘는 경우이다. 하지만 이 경우 튀는 값이 강수량의 영향인지 여부를 구분해내지 못하기 때문에 튀는 값 주변에 2차 미분을 실시한다. t 에서 급격한 증가가 일어났다면, $t-1$ 과 $t+1$ 에서의 토양수분은 큰 변화가 없었다고 가정하므로 해당 시간에서의 2차 도함수 비는 0.8과 1.2 사이의 값을 만족해야 한다. 하지만 이 또한 노이즈 자료에서는 효과가 없으므로, 튀는 값을 중심으로 24시간 동안의 평균과 분산의 비가 1보다 작아야 한다는 기준을 추가한다. 따라서 위의 세 가지 기준을 모두 충족할 경우, 튀는 값으로서 의심자료로 분류할 수 있다.

4.3.2 중단점 검출 (Break detection)

중단점은 갑작스런 점프나 드롭이 특징이다. 튀는 값과의 차이점은, 튀는 값은 그 후 원래의 값으로 돌아가지만 중단점의 경우 점프나 드롭이 일어나면 이전기간처럼 되돌아가지 않는다는 점이다. 검출방법은 튀는 값 검출과 비슷한 방법으로, 시계열을 보고 1차, 2차 도함수를 통해 잡아낸다. 중단점으로서 의심 플래그 처리하려면 다음의 세 가지 기준을 충족해야 한다. 첫째, 중단점이 일어나기 바로 전에는 최소한 10%내로 변화가 일어나야 한다. 하지만 토양수분 값 자체가 작을 경우 의심 플래그가 과부여될 수 있으므로 이를 고려하여 변화량의 절댓값은 0.01보다 커야한다는 기준을 추가한다. 둘째, 점프(드롭)은 1차 도함수의 강한 양의(음의) 변화로 나타

날 것이다. 따라서 24시간에서 중간 t 에서의 1차 도함수 값은 24시간의 모든 1차 도함수의 평균보다 10배는 커야(작아야) 한다. 셋째, 점프(드롭)은 매우 큰 양의(음의) 2차 도함수 값을 만들어 내며 또한 $t+1$ 에서 매우 큰 음의(양의) 2차 도함수 값을 만들어낸다. 따라서 t 와 $t+1$ 에서의 2차 도함수 값은 부호는 반대지만 크기는 거의 같으므로 그 비의 절댓값은 1이 되어야한다. 또한, $t+1$ 과 $t+2$ 에서의 2차 도함수는 매우 큰 차이가 나므로 그 비의 절댓값은 10을 넘어야 한다.

4.4 지속성 검사 (Flat line check)

토양수분 값이 일정하게 지속된다고 판단하기 위해서는 토양수분 값이 최소 12시간 동안 일정하게 지속되어야 하며 그 기간 동안의 분산은 센서 정밀도의 1%를 넘으면 안된다. 그 미만의 변화는 허용하기로 하는데, 왜냐하면 토양수분은 완전히 안정한 상태는 아니기 때문이다.

V. 결 과

지중온도에 품질검사 알고리즘을 적용한 결과, 튀는 값과 같은 비정상적인 값들이 제거되었으며 신뢰할 수 있는 시계열 자료와 통계값이 산출되었다(Fig. 1). 의심 자료들을 오류 자료로 분류하여 제거할 수 있었던 지중온도와는 달리, 토양수분은 의심자료로 플래그 처리하는 것에 그쳤다(Fig. 2). 왜냐하면 토양수분요소 특성상 관측값과 오류값의 확실한 구분이 어렵고 전체적으로 불규칙적이기 때문에 정상적인 값들이 필터링될 가능성이 있기 때문이다. 따라서 토양수분의 의심자료들을 오류값으로 분류하기 위해서는 더욱 확실한 근거가 필요하다. 수원기상대의 경우, 여러 해동안 한 곳에서 측정된 토양수분 관측자료를 분석하여 수원기상대에서의 토양수분의 양상이 어떻게 반복되고 있는지 파악하는 것이 가장 우선되어야 할 것이다.

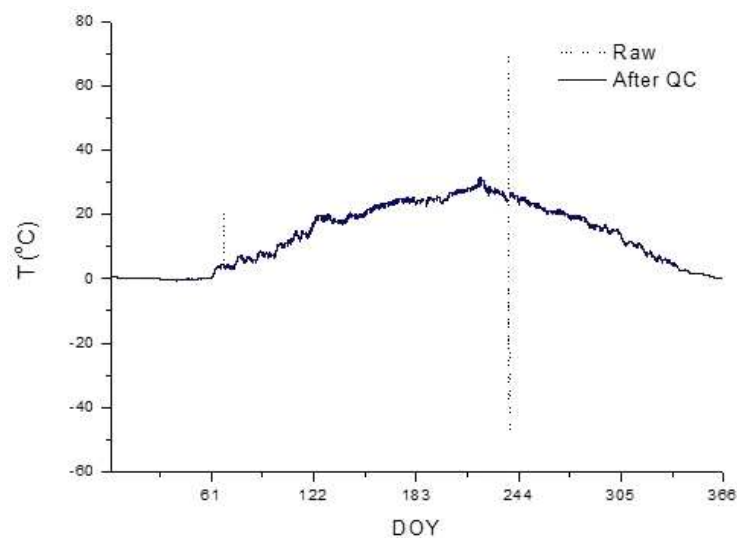


Fig. 1. Applying QC methods on 30cm soil temperature measured at Suwon weather station in 2012.

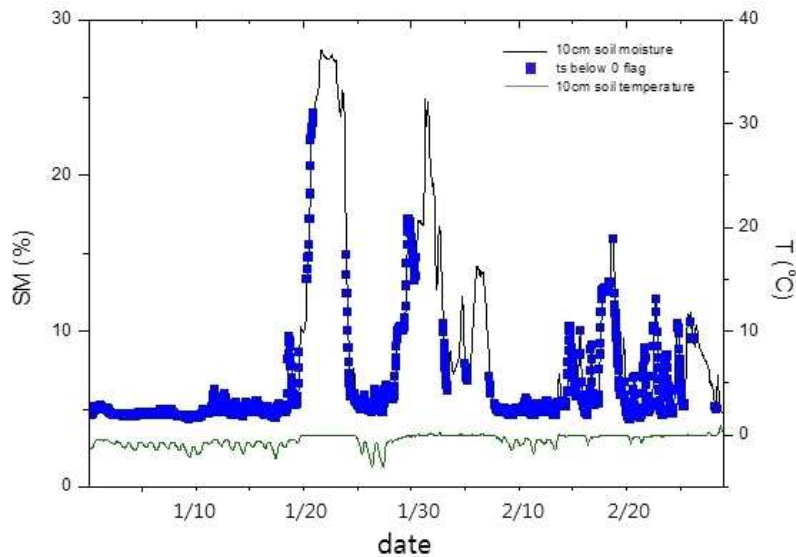


Fig. 2. Applying QC methods on 10cm soil moisture measured at Suwon weather station from January to February in 2013. It shows raw data (black line) and flagged data (filled square) through geophysical check.

사 사

이 연구는 기상청 국립기상연구소 “응용기상기술개발연구”의 지원으로 수행되었습니다.

인용문헌

- 기상청, 2011: 국가기후자료 품질관리 매뉴얼.
- 박창용과 최영은, 2012: 한국의 기온자료 품질관리 알고리즘의 검증, 2012년 한국기상학회 학술대회 논문집 22(3), 299-307.
- 오규림, 이승재, 김준, 이병렬, 김규량, 최병철, 2014: 기상청 농업기상관측요소의 품질관리, 2014년 한국기상학회 학술대회 논문집.
- 이영섭, 심우석, 오승준, 2007: 스펙트럼 분석을 이용한 기온자료 QC 기법에 관한 연구, 2007년 한국기상학회 학술대회 논문집, 300-301.
- 이영섭, 김동수, 김근현, 이재원, 김희경, 정세영, 2013: AWS 관측자료 품질관리, 2013년 한국기상학회 학술대회 논문집 2013(10), 28-29.
- 허복행, 이진아, 추영옥, 김준형, 박남철, 조주영, 오승준, 노맹석, 이영조, 2005: 기상관측자료 실시간 품질관리시스템(RQMOD)의 품질검사 문턱값을 산출하기 위한 AWS 관측자료 통계처리, 2005년 한국기상학회 학술대회 논문집, 390-391.
- Dorigo W.A., A. Xaver, M. Vreugdenhil, A. Gruber, A. Hegyiova, A.D. Sanchis-Dufau, D. Zamojski, C. Cordes, W. Wagner, M. Druch, 2012: Global Automated Quality Control of In Situ Soil Moisture Data from the International Soil Moisture Network, *Vadose Zone Journal* 12(3).
- WMO, 2013: *Guide to the Global Observing System*(WMO No. 488), Geneva.