

고랭지배추 생육을 위한 유비쿼터스 센서 네트워크 품질관리 알고리즘 개발

조창제¹ · 황근보² · 윤상후^{2*}

¹경상북도 경산시 진량읍 대구대로 201, 대구대학교 일반대학원 통계학과,

²경상북도 경산시 진량읍 대구대로 201, 대구대학교 수리빅데이터학부
(2018년 9월 17일 접수; 2018년 10월 12일 수정; 2018년 11월 9일 수락)

Development of Ubiquitous Sensor Network Quality Control Algorithm for Highland Cabbage

Changje Cho¹, Guenbo Hwang² and Sanghoo Yoon^{2*}

¹Department of Statistics, Daegu University, Gyeongbuk 38453, Korea.

²Division of Mathematics and Big Data Science, Daegu University, Gyeongsan Gyeongbuk 38453, Korea.

(Received September 17, 2018; Revised October 12, 2018; Accepted November 9, 2018)

ABSTRACT

Weather causes much of the risk of agricultural activity. For efficient farming, we need to use weather information. Modern agriculture has been developed to create high added value through convergence with state-of-the-art Information and Communication Technology (ICT). This study deals with the quality control algorithms of weather monitoring equipment through Ubiquitous Sensor Network (USN) observational equipment for efficient cultivation of cabbage. Accurate weather observations are important. To achieve this goal, the Korea Meteorological Administration, for example, developed various quality control algorithms to determine regularity of the observation. The research data of this study were obtained from five USN stations, which were installed in Anbandegi and Gwinemi from 2015 to 2017. Quality control algorithms were developed for flat line check, temporal outliers check, time series consistency check and spatial outliers check. Finally, the quality control algorithms proposed in this study can also identify potential abnormal observations taking into account the temporal and spatial characteristics of weather data. It is expected to be useful for efficient management of highland cabbage production by providing quality-controlled weather data.

Key words: USN, Quality Control, Agro-meteorological data, Highland cabbage



* Corresponding Author : Sanghoo Yoon
(statstar@daegu.ac.kr)

I. 서 론

농산업을 기상에 직접적인 영향을 받는 대표적 경제활동 중 하나이다. 날씨의 모든 농산업활동에 영향을 미치고 있으므로 농사의 풍작여부는 기상상황에 달려있다. 따라서 효율적인 농산업활동을 견인하기 위해서는 기상정보를 적극 활용해야 한다. 이를 위해 기상정보의 농산업 활용을 위한 첨단 기술과 농업의 융합 연구가 진행되고 있다. 정보통신기술(Information and Communications Technology)의 발전은 무선인식(RFID) 및 유비쿼터스 센서 네트워크(Ubiquitous Sensor Network, USN) 등을 통해 농산물의 생육정보와 기상정보 간 상관성을 분석할 수 있는 환경을 조성하였다. Kim *et al.*(2013)은 농산물의 생산, 유통, 소비 전 과정에 걸친 자료를 실시간으로 생성하여 농산물의 생산성 및 효율, 품질 향상 등의 고부가가치를 창출하는 방안을 제시하였다. Kang *et al.*(2011)은 USN 및 RFID 기술을 기반으로 다양한 시범사업을 진행했으며, Yoo *et al.*(2009)는 USN을 활용한 농작물 재배관리와 농작물 이력추적 시스템을 제시하였다. 이와 같은 RFID/USN을 활용한 고부가가치 창출이 효율적으로 운영되기 위해서는 자료의 품질관리(Quality Control, QC)가 필요하다. 기상청은 기상관측표준위원회를 구성하여 기상관측자료의 품질 등급 기준을 수립하고 관측자료별 QC 절차를 개발하여 품질을 관리하고 있다. 기상관측자료의 품질관리 알고리즘은 단계 검사, 지속성검사, 허용범위검사 등으로 구성되어 있다. 하지만 USN을 통한 기상관측장비의 QC알고리즘은 아직 개발되어 있지 않아 돌발 상황에 따른 기기적 오류와 장비의 부식에 따른 관측자료의 품질 저하를 확인하기 어렵다.

고랭지는 해발고도가 높아질수록 기온이 낮아지는 기후특성을 지니고 있어 여름철 고온으로 재배가 어려운 각종 채소작물의 고품질 생산이 가능하다. 해발고도가 높은 강원도는 대표적인 고랭지 배추지역으로 고도에 따라 정식 및 수확시기가 다르다. 기온에 따른 작황변동은 고랭지배추 가격과 밀접한 관련(Ahn *et al.*, 2014)이 있으므로 이를 배추 재배전략에 활용한다면 병충해 및 이상기온에 의한 피해를 줄일 수 있다(Kim and Yun, 2015). 하지만 현재 기상청에서 운영 중인 기상관측장비의 경우 2018년 10월을 기준으로 1,000m 이상의 고도의 기상관측장비가 12개뿐이며, 현재 기상청에서 강원도에 설치운영 중인 중관기상관

측장비(Automated Synoptic Observing System, ASOS)의 최대고도는 842m(대관령)으로 해발고도 1,000m 이상의 고랭지 배추 재배지역의 기상 예측에 한계점이 있다. 높은 고도지역에서 재배되는 고랭지 배추의 효율적인 생육관리를 위해 농촌진흥청에서는 2015년부터 작황정보 고도화를 위해 USN 5대를 설치하여 토양수분, 토양온도, 일사량, 습도, 온도, 일조량, 풍속, 풍향, 강우량, 토양전도율의 기상자료를 센서로 수집하고 있다. USN은 다수의 센서에서 수집되는 정보를 싱크 노드로 전달하면 싱크 노드는 네트워크를 통해 사용자에게 전달한다. 본 연구에서는 온도, 습도, 풍속, 강우량을 사용하여 분석하였다.

고랭지배추의 생육에 날씨가 미치는 영향을 분석하기 위해선 고품질의 기상관측정보가 필요하다. 그러나 농림기상 관측자료는 지상기상관련 자료에 비해 QC가 원활하게 이루어지지 않고 있다(Oh *et al.*, 2015). Oh *et al.*(2015)는 기상청에서 제공하는 농업기상관측서 중 수원기상대에서 제공하는 2012년~2013년 자료를 사용하였으며, 크게 물리적 범위 검사, 단계 검사, 지속성검사, 지구물리학적 검사를 적용하였다. 하지만 농작물의 공간적 특성에 대하여 고려되지 않았다. 이에 본 연구에서는 공간분포 검사를 진행하였다. QC에 관한 선행연구로 Choi *et al.*(2015), Lee *et al.*(2010), Nam *et al.*(2014) 등이 있다. Choi *et al.*(2015)은 농림기상관측자료의 관측환경을 평가하고 개선방안을 제시하였고, Nam *et al.*(2014)는 자동 기상 관측소(Automatic Weather Station; AWS)의 QC를 조사하였다. Lee *et al.*(2010)은 기온자료의 극한값 검사, 내적일치성 검사, 기후범위검사, 공간분포 검사를 사용하여 QC를 개발하였다. 하지만 농림기상을 위한 QC 알고리즘은 미흡한 실정이다. 본 연구는 고랭지 배추 재배지역에 설치된 USN 기상관측자료의 품질을 평가하기 위한 QC알고리즘을 제안한다.

II. 재료 및 방법

2.1. 유비쿼터스 센서 네트워크(Ubiquitous Sensor Network, USN)

USN은 유·무선으로 연결된 다수의 센서들이 정보를 수집하고 이를 가공하여 활용하는 네트워크를 의미하며 센서노드, 싱크노드, 게이트웨이를 통해 언제 어디서나 사물을 및 환경 정보를 감시, 저장, 가공, 통합하여 무선으로 전송 할 수 있다(Fig. 1). 현재 USN은

ICT기술을 기반으로 국방, 제조, 건설, 교통 등 다양한 분야에 이용되고 있다. 본 연구에서 이용된 USN은 고랭지 배추의 생육과 관련이 있을 것이라 판단되는 토양수분, 토양온도, 일사량, 습도, 온도, 일조량, 풍속, 풍향, 강우량, 토양전도를 센서로 구성되어 수집된 정보를 네트워크를 통해 전달하여 실시간으로 모니터링 하도록 구성되어 있다.

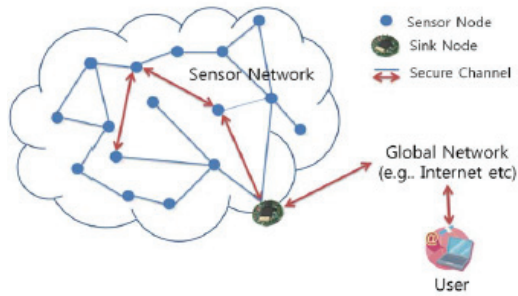


Fig. 1. The sensor network for USN.

Table 1. Location of USN & ASOS

No	Lat	lon	alt
USN 1	37.61474	128.7408	1057
USN 2	37.61768	128.7386	1052
USN 3	37.62078	128.7383	1006
USN 4	37.33685	129.0065	953
USN 5	37.2235	128.9654	1156
ASOS 100(Taebaek)	37.6771	128.7183	772
ASOS216 (Daegwallyeong)	37.1703	128.9893	712

2.2. 분석자료

배추는 생육 초기에는 온도에 대한 적응범위가 좁으며, 고온에 잘 견디나, 결구기에는 내고온성이 약한 채소로 서늘한 기후에 적합한 작물이다. 본 연구는 기후학적특성을 고려하여 고랭지배추 재배에 적합한 강원도 강릉, 태백 내 안반덕, 귀네미 인근의 5개지점(안반덕 3지점, 귀네미, 메봉산)에 설치된 USN을 선정하였다(Table 1). 연구자료는 2015, 2016, 2017년의 배추재배시기인 7월~10월 총 44,280개의 시간단위 기상자료이다. 관측된 USN의 기후범위검사를 위해 기상청에서 운영하는 ASOS의 해발고도가 772.57m와 712.82m인 대관령과 태백의 총 20년(1998년~2017년) 자료를 활용하였다(Fig. 2).

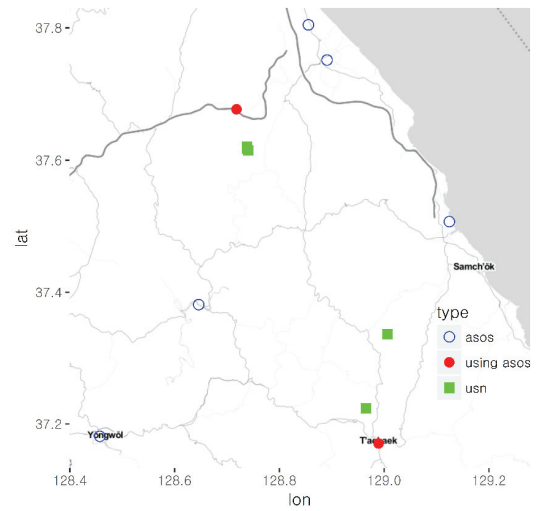


Fig. 2. The spatial distribution of equipments.

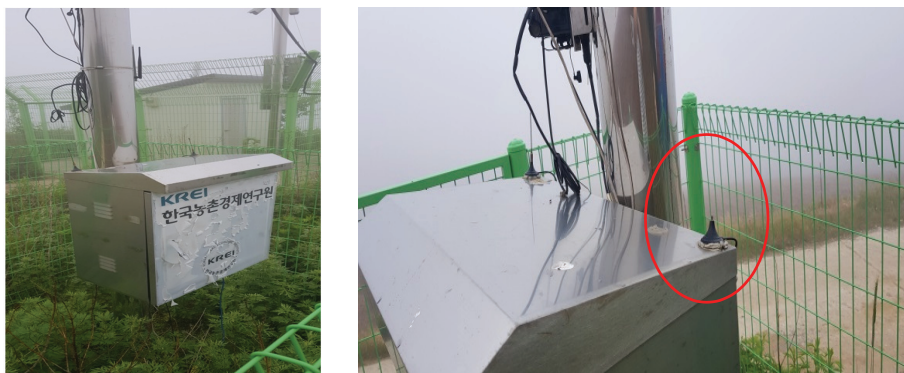


Fig. 3. An expected cause of error.

본 연구에서 사용된 USN 자료는 해발고도가 높은 산악지역에서 수집되었다. Kim(2011)에 의하면 산악 지역은 전력공급이 원활하지 않은 문제로 기기의 오류가 발생할 수 있다고 하였다. 또한, 산악지역의 접근성이 좋지 않아 규칙적인 점검을 통해 양질의 자료를 수집하기 쉽지 않다. 예를 들어 본 연구의 USN은 잡초로 인한 센서의 인식불량 또는 안테나파손 등의 문제점이 있었다(Fig. 3). 이에 본 연구에서는 USN에서 관측된 농업기상정보의 품질을 향상시키기 위한 QC알고리즘을 제안하고자 한다.

QC알고리즘은 지속성검사, 기후범위검사, 시간변동성검사, 공간분포검사로 진행되며 기기정지 및 오류를 검출한 후, 시공간적 특성을 활용하여 기준치를 벗어나는 자료를 오류 의심자료로 검출하는 순서이다(Fig. 4).

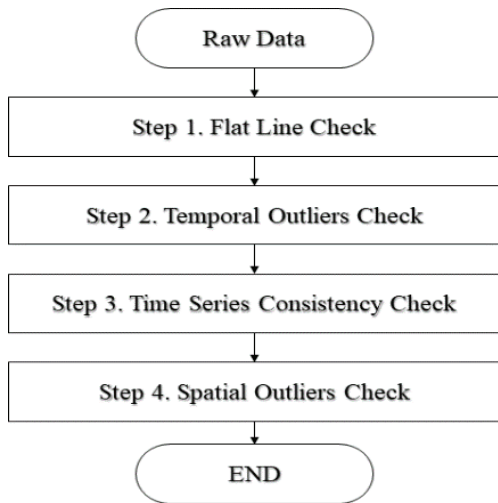


Fig. 4. The Flowchart of QC algorithm.

2.3. 지속성 검사(Flat line check)

전산 입력상의 오류 등으로 인해 같은 값을 연속으로 가지는 경우를 오류로 판단하는 검사이다(Reek et al., 1992). 본 연구에서는 모든 변수들의 값이 0으로 기록되는 경우를 기기 정지로 판단하였으며, 습도, 온도, 풍속이 5회 이상 같은 값으로 반복되는 경우 이를 오류 의심자료로 판단하였다. 습도의 경우 0으로 기록되는 경우 오류 의심자료로 판단하였다.

2.4. 기후범위 검사(Temporal outliers check)

기후범위 검사는 장기간 관측된 자료의 분포를 통

하여 일정 예상 범위를 벗어나는 관측값을 오류로 판정하는 검사이다(KMA, 2016). 본 연구에서는 USN 5개 지점 인근의 ASOS를 활용하였다. 대관령에 설치된 ASOS의 기후범위를 통해 USN 1, 2, 3지점의 기후범위를 설정하였고, 태백에 설치된 ASOS의 기후범위를 통해 USN 4, 5지점의 기후범위를 설정하였다.

기후범위를 위해선 자료의 대칭성이 중요하다. 관측값들의 분포가 대칭형 분포를 가지면 모수적 이상값 탐지방범(식 1) 활용하여 상한과 하한을 결정하였으며 k 값을 1부터 6까지 변화시키며 기준값을 산출하였다(Rousseeuw · Hubert: 2011).

$$mean \pm k(\text{standard deviation}) \quad k = 1, 2, \dots, 6. \quad (1)$$

관측값들의 분포가 비대칭적 분포를 이루는 경우 비모수적 방법으로 평균과 표준편차 대신 중앙값과 중앙값 절대편차(Median Absolute Deviation, MAD)를 이용하여(식 2) 기준값을 산출하였다(Wilcox, 1996). 단, 풍속처럼 최소값이 0인 경우 하한값은 의미가 없으므로 상한값만을 산출하였다.

$$Median \pm 1.4826 \times k \times (\text{median}(|X_i - \text{median}(X)|)) \quad k = 1, 2, \dots, 6. \quad (2)$$

Fig. 5는 USN으로부터 관측된 자료의 온도와 습도의 분포이다. 온도는 대칭적 분포이고 습도는 왼쪽으로 치우쳐진 비대칭 분포이다. 강수량과 풍속의 경우 20년간의 기후자료에서 최대값을 넘어가는 자료는 이상치로 판단하였고, 습도는 비정규분포를, 온도는 정규분포를 따르고 있다.

2.5. 시간변동성 검사(Time series consistency check)

시간변동성검사는 기상 요소가 일정 기간 내에 특정 값 이상 변화할 수 없다는 기준을 통해 최대 변동량의 범위를 산정하여 이를 벗어나면 오류로 판정하는 검사이다(KMA, 2016). USN 자료의 시간 변동량을 확인하기 위하여 전시점과 현시점 사이의 차이를 계산하였다(식 3).

$$diff_i = \frac{x_i - x_k}{NA+1}, \quad (3)$$

여기서, x_k 는 x_i 가 결측치가 아닌 값 중에서 x_i 의

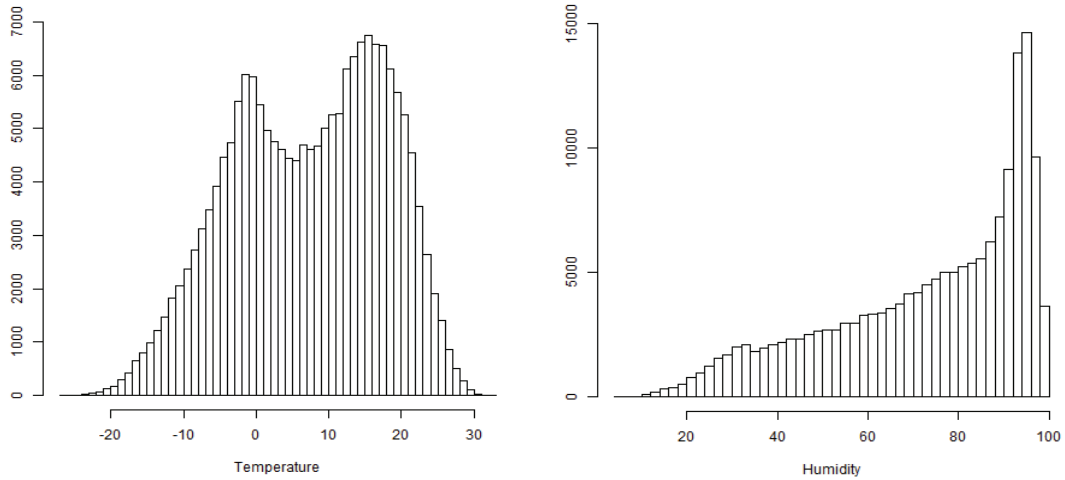


Fig. 5. The distribution of temperature and relative humidity.

전시점에 해당하는 값이며, NA 는 x_i 와 x_k 사이의 결측치의 개수를 의미한다. 이렇게 계산한 $diff_i$ 의 분포에서 상위 5% 분위수를 상한선으로 설정하여 이를 넘는 경우를 이상치로 판단하였다.

$$threshold = median(X) + 1.4826 \times k \times (median(|X_i - median(X)|)), (k = 1, 2, \dots, 6). \tag{4}$$

2.6. 공간분포검사(Spatial outliers check)

기상관측장비는 공간적 특성을 지니고 있다. 이 성질을 반영하여 공간분포검사를 실시하였다. 공간분포검사는 인근지역의 관측자료와 관심지역의 관측자료를 비교를 통해 오류를 판정한다. 본 연구에서는 기상청에서 강원도에 설치·운영 중인 ASOS 14개 지점과 AWS 72개 지점에서 관측된 기상자료로 USN 인근의 ASOS인 태백과 대관령의 기상관측값을 크리깅으로 예측하였다. 태백과 대관령의 기상관측값과 기상예측값 차이의 분포는 대칭형 분포이므로 아래와 같이 기준값 k 에 따른 상한과 하한을 구하였다(식 5).

$$Range = (\hat{A}_i - k(std_{\hat{A}_i}), \hat{A}_i + k(std_{\hat{A}_i})), k = 1, 2, \dots, 6. \tag{5}$$

공간분포검사에서 구해진 상한과 하한은 USN 지점별 일단위 기온, 상대습도, 강수량의 상한과 하한으로 사용된다.

2.7. 크리깅(Kriging)

크리깅은 지리통계학에서 미지의 위치의 관측값을 예측하는 모형이다. 크리깅은 관심 지점의 값을 예측하기 위해 위치정보가 있는 관측된 값들의 가중 선형 조합을 베리오그램(Variogram)으로 추정하고 이를 통해 관심 지점의 값을 예측한다. 본 연구에서는 단순크리깅(Simple kriging), 보통크리깅(Ordinary kriging), 일반크리깅(Universal kriging) 중 해발고도를 고려한 일반크리깅을 사용하였다.

III. 결과 및 분석

본 연구는 고랭지배추지역에 설치된 USN의 QC알고리즘을 통해 수집된 기상자료의 품질을 향상시키기 위한 방법을 제시한다. 먼저 기기의 오류 및 정지를 판단하는 지속성 검사를 통해 기기오류로 판단한 결과는 Table 2이다. 지속성 검사 결과 3지점이 다른 지점들에 비해 기기정지가 높게 나타났으며, 5지점이 기기오류가 높게 나타났다. 기상요소별로 살펴보면 습도의 기기오류가 다소 높은 편이다.

장기관측된 자료를 바탕으로 이상치를 판단하는 기후범위 검사의 검사결과는 Table 3와 Table 4이다. 대칭형 분포인 온도는 평균과 표준편차로 기준값을 산출하였고 비대칭형 분포인 상대습도는 중앙값과 MAD로 기준값을 계산하였다. 풍속과 강수량의 기후자료의 최대값을 통해 기준값이 산출되었다. k 의 크기에 따라 오류율이 다르게 나오고 있으며 k 가 1이면 너무 민감하게 반응하는 것으로 보여지며, k 가 4이상이면 대부

Table 2. Errors by stations for flat line check

Variable	all station (N=44,280)	USN 1 (N=8,856)	USN 2 (N=8,856)	USN 3 (N=8,856)	USN 4 (N=8,856)	USN 5 (N=8,856)
Sensor stop	6,634(15.0%)	1,249(14.1%)	1,264(14.3%)	3,182(35.9%)	497(5.6%)	442(5.0%)
Humidity error	10,606(24.0%)	2,220(25.1%)	1,608(18.2%)	1,618(18.3%)	2,624(29.6%)	2,536(28.6%)
Temperature error	1,664(3.8%)	379(4.3%)	337(3.8%)	0(0.0%)	503(5.7%)	445(5.0%)
Wind Speed error	3,891(8.8%)	345(3.9%)	344(3.9%)	121(1.4%)	592(6.7%)	2,489(28.1%)
Sensor stop or error	18,854(42.6%)	3,471(39.2%)	2,879(32.5%)	4,852(54.8%)	3,161(35.7%)	4,491(50.7%)

분의 지점에서 오류가 발생하지 않았다. 전반적으로 태백인근에 해당하는 USN 1, 2, 3지점과 대관령인근에 해당하는 USN 4, 5지점이 서로 유사한 특징을 보

이고 있었다. 따라서 지역적인 특수성을 가진다고 할 수 있다. 그러므로 지역별로 다른 기준값을 제시할 필요가 있다.

Table 3. The result of temporal outliers check given threshold

Variable	k	Mean / Median (sd) / (MAD)	Lower limits	Upper limits	USN1	USN2	USN3	
ASOS 100	Humidity	68(18)	1	41.31	94.69	798(14.8%)	630(10.5%)	613(15.1%)
			2	14.63	100.00	170(3.2%)	49(0.8%)	71(1.8%)
			3	-12.06	100.00	104(1.9%)	0(0.0%)	0(0.0%)
			4	-38.75	100.00	0(0.0%)	0(0.0%)	0(0.0%)
			5	-65.43	100.00	0(0.0%)	0(0.0%)	0(0.0%)
			6	-92.12	100.00	0(0.0%)	0(0.0%)	0(0.0%)
	Temperature	1	9.17 (10.39)	-1.22	19.56	3,317(45.9%)	3,284(45.3%)	2,850(50.2%)
		2		-11.61	29.96	155(2.1%)	112(1.5%)	205(3.6%)
		3		-22.01	40.35	15(0.2%)	0(0.0%)	0(0.0%)
		4		-32.41	50.75	13(0.2%)	0(0.0%)	0(0.0%)
		5		-42.80	61.14	2(0.0%)	0(0.0%)	0(0.0%)
		6		-53.20	71.53	2(0.0%)	0(0.0%)	0(0.0%)
Variable	k	Mean / Median (sd) / (MAD)	Lower limits	Upper Limits	USN4	USN5	All	
ASOS 216	Humidity	79 (15)	1	56.76	100.00	1,706(29.8%)	1,630(27.7%)	5,377(19.9%)
			2	34.52	100.00	0(0.0%)	2(0.0%)	292(1.1%)
			3	12.28	100.00	0(0.0%)	0(0.0%)	104(0.4%)
			4	-10.00	100.00	0(0.0%)	0(0.0%)	0(0.0%)
			5	-32.2	100.00	0(0.0%)	0(0.0%)	0(0.0%)
			6	-54.44	100.00	0(0.0%)	0(0.0%)	0(0.0%)
	Temperature	1	7.32 (10.55)	-3.23	17.87	2,570(32.7%)	2,614(32.8%)	5,377(19.9%)
		2		-13.79	28.42	45(0.6%)	32(0.4%)	292(1.1%)
		3		-24.34	38.98	0(0.0%)	10(0.1%)	104(0.4%)
		4		-34.90	49.53	0(0.0%)	7(0.1%)	0(0.0%)
		5		-45.44	60.08	0(0.0%)	0(0.0%)	0(0.0%)
		6		-56.00	70.64	0(0.0%)	0(0.0%)	0(0.0%)

Table 4. The result of temporal outliers check given threshold

Variable	Max		Errors(%)					
	ASOS 100	ASOS 216	all	USN1	USN2	USN3	USN4	USN5
Wind Speed	14.6m/s	24.3m/s	5 (0.01%)	0 (0%)	0 (0%)	0 (0%)	2 (0.03%)	3 (0.05%)
Precipitation	65.5mm	60.5mm	41 (0.11%)	0 (0%)	0 (0%)	37 (0.65%)	4 (0.05%)	0 (0%)

Fig. 6는 USN자료의 각각의 변수별로 전시점과 현시점의 차이를 분포로 나타낸 것이다. 편차 자료값들이 0에 해당되는 자료가 많아 산포도 추정이 되지 않아 상위 5%의 편차자료만을 사용하여 분포 특성을 파악하였다. USN자료의 편차에 대한 자료의 경우 비대칭형 분포를 따르므로 중앙값과 MAD를 이용하여 기준값을 산출하였다.

기기오류나 기기정지의 경우 지속성검사를 통해 확

인 할 수 있으나, 기기오류로 인해 관찰값이 급격한 변화를 보이는 경우 이를 오류로 판단할 수 없다. 따라서 전시점과 현시점의 차이를 이용한 시간변동성 검사를 실시하였고 그 결과는 Table 5와 같이 나타났다. 습도의 경우 온도나 풍속, 강수량에 비해 상대적으로 오류율이 높게 나타났다. 이는 분석에서 시간단위 자료를 일단위 자료로 변환함에 있어 평균을 사용할 때 문제가 된다. 시간단위 자료에 품질관리가 잘 되지 않

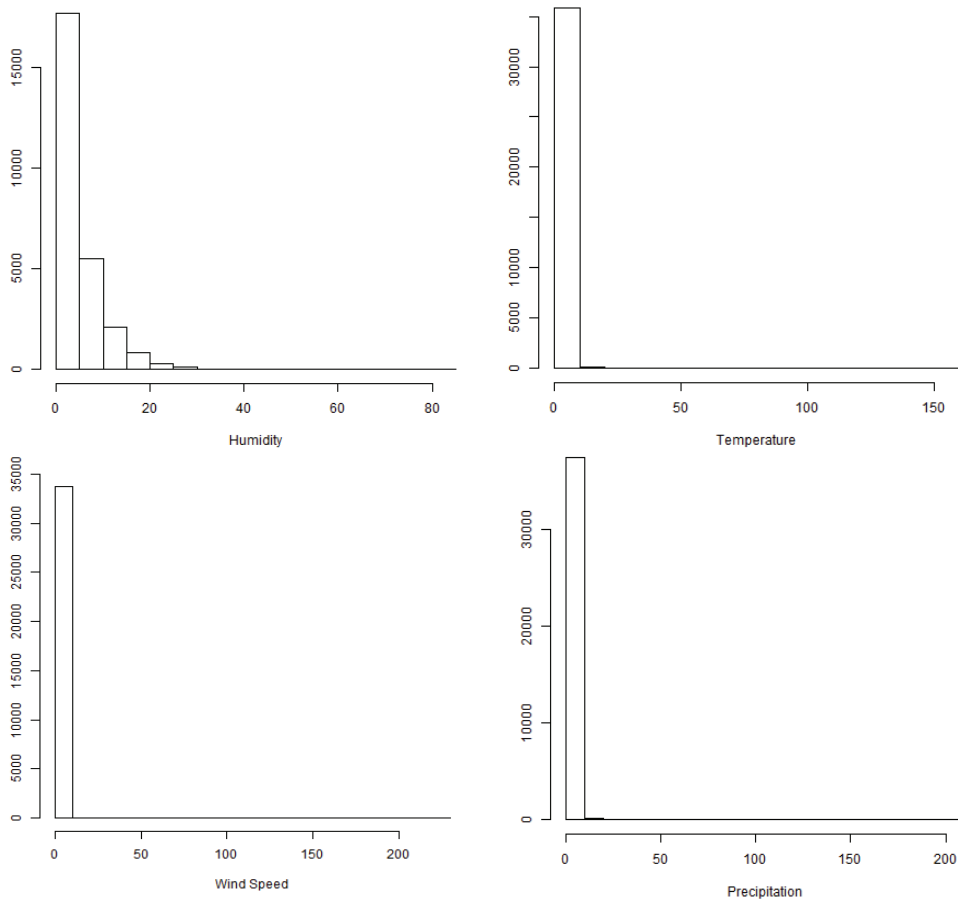


Fig. 6. Distribution of deviations by variables.

Table 5. The result of time series consistency check given threshold

Variable	k	Median (MAD)	Threshold values	error(%)	Variable	Median (MAD)	Threshold values	error(%)
Humidity	1	18.24 (2.3)	21.65	846(3.1%)	Wind Speed	2.33 (0.34)	2.83	523(1.6%)
	2		25.06	687(2.5%)			3.34	330(1.0%)
	3		28.47	599(2.2%)			3.84	228(0.7%)
	4		31.88	561(2.1%)			4.35	178(0.5%)
	5		35.29	543(2.0%)			4.85	151(0.5%)
	6		38.70	527(2.0%)			5.35	130(0.4%)
Temperature	1	3.98 (0.61)	4.88	495(1.4%)	Precipitation	2.24 (1.12)	3.90	635(1.7%)
	2		5.79	312(0.9%)			5.56	495(1.3%)
	3		6.69	255(0.7%)			7.22	418(1.1%)
	4		7.60	224(0.6%)			8.88	363(1.0%)
	5		8.50	198(0.6%)			10.54	294(0.8%)
	6		9.41	187(0.5%)			12.20	263(0.7%)

은 자료가 존재한다면 평균은 이상치에 영향을 많이 받
으므로 일단위로 변환한 자료를 신뢰할 수 없게 된다.

기상자료는 공간특성을 지니고 있어 미관측지역의
기상은 인근 관측지역의 기상과 유사하다. 이 특성을

반영한 공간분포검사를 진행한 결과는 Table 6이다.
기상청에서 운영중인 ASOS와 AWS의 위치정보와 관
측값을 활용하여 관심지점의 기상 예측값을 생성하고
상한과 하한을 규정하여 범위를 벗어나면 이상치로 판

Table 6. The result of spatial outlier check given threshold

Variable	k	All	USN 1	USN 2	USN 3	USN 4	USN 5
Humidity	1	5,855(21.7%)	905(16.8%)	1,006(16.8%)	744(18.3%)	1,546(27.0%)	1,654(28.1%)
	2	1943(7.2%)	272(5.1%)	263(4.4%)	246(6.1%)	515(9.0%)	647(11.0%)
	3	627(2.3%)	127(2.4%)	50(0.8%)	116(2.9%)	129(2.3%)	205(3.5%)
	4	274(1.0%)	74(1.4%)	12(0.2%)	57(1.4%)	49(0.9%)	82(1.4%)
	5	165(0.6%)	61(1.1%)	0(0.0%)	35(0.9%)	22(0.4%)	47(0.8%)
	6	105(0.4%)	55(1.0%)	0(0.0%)	17(0.4%)	12(0.2%)	21(0.4%)
Temperature	1	15,164(42.1%)	2,868(39.7%)	3,069(42.3%)	2355(41.5%)	3,214(40.9%)	3,658(45.9%)
	2	9,973(27.7%)	1,785(24.7%)	1,962(27.0%)	1,408(24.8%)	2,073(26.4%)	2,745(34.5%)
	3	6,105(17.0%)	1,070(14.8%)	1,120(15.4%)	799(14.1%)	1,210(15.4%)	1,906(23.9%)
	4	3,746(10.4%)	674(9.3%)	664(9.2%)	532(9.4%)	716(9.1%)	1,160(14.6%)
	5	2,538(7.1%)	500(6.9%)	494(6.8%)	397(7.0%)	467(5.9%)	680(8.5%)
	6	1,858(5.2%)	395(5.5%)	387(5.3%)	316(5.6%)	350(4.5%)	410(5.1%)
Precipitation	1	604(1.6%)	141(1.9%)	125(1.7%)	94(1.7%)	133(1.6%)	111(1.3%)
	2	298(0.8%)	71(0.9%)	61(0.8%)	45(0.8%)	68(0.8%)	53(0.6%)
	3	179(0.5%)	47(0.6%)	41(0.5%)	31(0.6%)	35(0.4%)	25(0.3%)
	4	116(0.3%)	35(0.5%)	27(0.4%)	22(0.4%)	23(0.3%)	9(0.1%)
	5	95(0.3%)	30(0.4%)	22(0.3%)	18(0.3%)	19(0.2%)	6(0.1%)
	6	73(0.2%)	21(0.3%)	14(0.2%)	15(0.3%)	18(0.2%)	5(0.1%)

Table 7. The result for quality control of USN data

Variable	k	All	USN1	USN2	USN3	USN4	USN5
Humidity	1	28,417(64.2%)	5,149(58.1%)	4,538(51.2%)	6,200(70.0%)	6,340(71.6%)	6,190(69.9%)
	2	20,032(45.2%)	3,970(44.8%)	3,297(37.2%)	5,206(58.8%)	3,800(42.9%)	3,759(42.5%)
	3	18,487(41.8%)	3,740(42.2%)	3,018(34.1%)	5,006(56.5%)	3,408(38.5%)	3,315(37.4%)
	4	18,064(40.8%)	3,641(41.1%)	2,975(33.6%)	4,946(55.9%)	3,316(37.4%)	3,186(35.6%)
	5	17,942(40.5%)	3,625(40.9%)	2,960(33.4%)	4,922(55.6%)	3,287(37.1%)	3,148(35.6%)
	6	17,868(40.4%)	3,617(40.8%)	2,958(33.4%)	4,900(55.3%)	3,271(36.9%)	3,122(35.3%)
Temperature	1	32,084(72.5%)	6,553(74.0%)	6,613(74.7%)	7,252(81.9%)	5,748(64.9%)	5,918(66.8%)
	2	18,830(42.5%)	3,568(40.3%)	3,703(41.8%)	4,712(53.2%)	3,160(35.7%)	3,687(41.6%)
	3	14,583(32.9%)	2,750(31.1%)	2,764(31.2%)	3,996(45.1%)	2,248(25.4%)	2,825(31.9%)
	4	12,206(27.6%)	2,350(26.5%)	2,305(26.0%)	3,729(42.1%)	1,745(19.7%)	2,077(23.5%)
	5	10,978(24.8%)	2,174(24.6%)	2,131(24.1%)	3,590(40.5%)	1,493(16.9%)	1,590(18.0%)
	6	10,293(23.3%)	2,069(23.4%)	2,022(22.8%)	3,508(39.6%)	1,373(15.5%)	1,321(14.9%)
Wind Speed	1	11,051(25.0%)	1,702(19.2%)	1,703(19.2%)	3,337(37.7%)	1,259(14.2%)	3,050(34.4%)
	2	10,864(24.5%)	1,663(18.8%)	1,676(18.9%)	3,326(37.6%)	1,190(13.4%)	3,009(34.0%)
	3	10,755(24.3%)	1,638(18.5%)	1,650(18.6%)	3,322(37.5%)	1,157(13.1%)	2,988(33.7%)
	4	10,704(24.2%)	1,629(18.4%)	1,638(18.5%)	3,320(37.5%)	1,142(12.9%)	2,975(33.6%)
	5	10,677(24.1%)	1,624(18.3%)	1,631(18.4%)	3,320(37.5%)	1,132(12.8%)	2,970(33.5%)
	6	10,656(24.1%)	1,619(18.3%)	1,628(18.4%)	3,320(37.5%)	1,127(12.7%)	2,962(33.5%)
Precipitation	1	7,738(17.5%)	1,490(16.8%)	1,440(16.3%)	3,460(39.1%)	762(8.6%)	586(6.6%)
	2	7,349(16.6%)	1,395(15.8%)	1,368(15.5%)	3,395(38.4%)	671(7.6%)	520(5.9%)
	3	7,181(16.2%)	1,359(15.4%)	1,339(15.1%)	3,372(38.1%)	624(7.1%)	487(5.5%)
	4	7,088(16.0%)	1,337(15.1%)	1,324(15.0%)	3,359(37.9%)	597(6.7%)	471(5.3%)
	5	7,011(15.8%)	1,316(14.9%)	1,309(14.8%)	3,343(37.8%)	582(6.6%)	461(5.2%)
	6	6,965(15.7%)	1,302(14.7%)	1,299(14.7%)	3,330(37.6%)	576(6.5%)	458(5.2%)

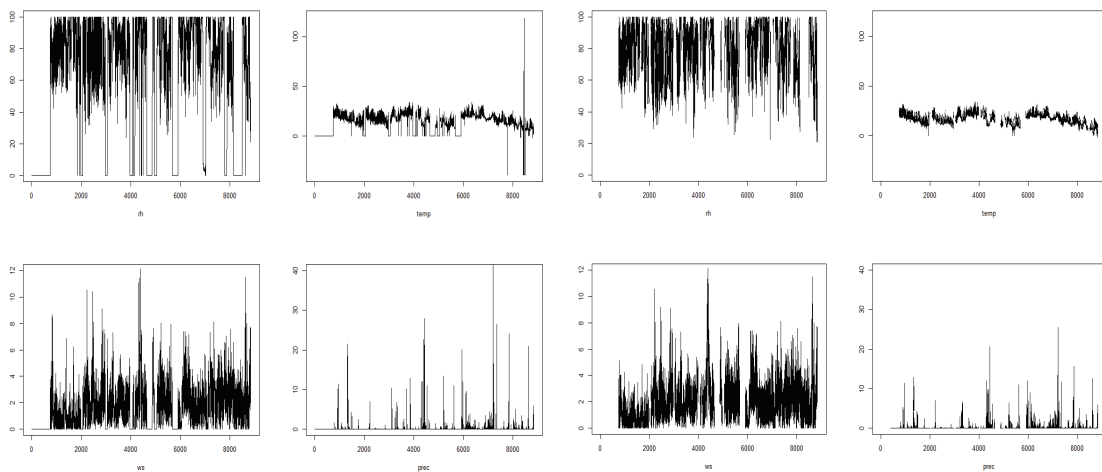


Fig. 7. Before applying QC algorithm.

Fig. 8. After applying QC algorithm.

단하였다. 습도의 경우 대관령에 해당하는 1, 2, 3지점에 비하여 태백에 해당하는 4, 5지점이 이상치로 판단되는 지점이 더 많이 나타났으며, 강수량의 경우 3지점이 다소 낮게 나타났다.

Table 7은 지속성검사, 기후범위검사, 시간변동성검사, 공간분포검사를 모두 적용한 품질검정의 결과 비정상적으로 분류된 자료의 비율이다. k에 따라 가장 변동폭이 큰 변수는 온도로 나타났으며, 습도가 다른 변수보다 높은 오류율을 보인다. 품질검사 알고리즘 결과 3지점의 오류율이 가장 높게 나타났다. 이는 3지점이 기기정지로 판정되는 자료가 제일 많이 나타났기 때문이다.

Fig. 7은 QC알고리즘을 적용하기 전 원자료를 시간순서대로 나타낸 자료이고, Fig. 8은 k를 3으로 한 QC알고리즘을 적용한 후 시간순서대로 나타낸 자료이다. QC알고리즘을 적용하기전에는 이상치로 판단되는 급변하는 자료들이 많았으나 QC알고리즘 적용한 후 다소 안정된 값을 가지는 것을 볼 수 있다.

IV. 결 론

본 연구에서는 해발고도가 높은 고랭지배추 재배지역에 설치된 USN자료의 QC알고리즘을 개발하고, 기상청의 AWS와 ASOS자료를 이용하여 USN 관측자료의 품질을 검사하였다. QC알고리즘은 지속성 검사, 기후범위 검사, 시간변동성 검사, 공간분포 검사로 구성되어 있다. USN 기상관측자료의 품질은 지속성 검사를 먼저 수행한 후, 기후범위 검사, 시간변동성 검사, 공간분포검사를 독립적으로 수행하였다. 시간변동성 검사에서 습도는 다른 기상요소에 비해 오류율이 다소 높으므로 설치지역에 방문하여 습도 관측센서를 점검할 필요성이 있다. 공간분포 검사에서는 온도, 습도, 강수량 순으로 오류율이 높게 나타났으며, 태백지역이 대관령지역보다 오류율이 높게 나타났다. 품질관리 알고리즘을 모두 적용한 결과 온도, 습도, 풍속, 강수량 순으로 오류율이 높게 나타났으며, 습도, 온도, 풍속은 대관령지역이 대체로 오류율이 높게 나타났고, 강수량은 태백지역이 상대적으로 높은 오류율을 보였다. 품질 관리 결과 기기 오류 및 오작동을 제외하더라도 상대적으로 다소 높은 오류율을 보인다.

본 연구에서 제안한 QC알고리즘의 한계점은 다음과 같다. 사용된 알고리즘의 기후검사는 근처의 ASOS 자료의 기후자료가 사용되었으나 ASOS의 해발고도

보다 USN의 해발고도가 높다. 장기간 관측이 이루어진다면 기상청의 기후검사와 비슷하게 산악지역의 기후를 토대로 품질관리를 할 수 있을 것이다. 또한 온도, 강수량, 습도, 풍속 외에 토양수분, 토양온도, 토양전도율, 일조량 등의 다른 기상요인의 QC알고리즘에도 적용 시킬 수 있다.

농산업을 기상에 많은 영향을 받으므로 농작물의 효율적인 관리를 위해선 기상과 농작물 생육 간 관계를 규명해야 한다. 품질이 높은 기상관측자료는 농작물의 생육과 수확량을 예측하는데 필수적이다. 본 연구에서 제안한 QC알고리즘은 기상청에서 설치 운영 중인 기상관측자료를 활용하여 고랭지배추 지역에 설치된 USN 기상관측자료의 품질을 관리하기 위해 개발되었으며, 연구결과는 농가의 효율적인 운영 및 농업발전에 도움이 될 것이라 기대한다.

적 요

농업활동의 위험은 대부분은 기상에 의해 발생한다. 효율적인 농작업을 위해선 기상정보를 활용해야 한다. 현대 농업은 첨단 기술인 ICT와 융합을 통해 고부가가치를 창출하는 방향으로 발전하고 있다. 본 연구에서는 고랭지배추의 효율적인 재배를 위한 USN 관측장비를 통한 기상관측장비의 품질관리 알고리즘을 다룬다. 기상관측에서 정확한 관측이 중요하다. 이를 위해서 기상청에서는 기상관측 장비별로 품질관리 알고리즘을 개발하여 기상정보의 정확성 검증을 통해 정상자료 여부를 판정한 후 이를 활용한다. 연구자료는 2015년부터 2017년까지 3년간 대표적인 고랭지배추 재배지인 안반덕, 귀네미에 설치된 5개 USN 자료이다. 품질관리 알고리즘은 지속성검사, 기후범위검사, 시간변동성검사, 공간분포검사로 구성되어 있다. 마지막으로 본 연구에서 제안하는 품질관리 알고리즘은 기상자료의 공간적 특성을 고려한 잠재적 이상관측 여부도 확인할 수 있다. 또한 품질관리를 거친 자료를 토대로 고랭지배추와 기상관측자료의 상관성을 분석함으로써 효율적 농산업 관리에 도움이 될 것으로 보여진다.

REFERENCES

- Ahn, J. H., K. D. Kim, and J. T. Lee, 2014: Growth modeling of Chinese cabbage in an Alpine area. *Korean Journal of Agricultural and Forest Meteorology* **16**(4), 309-315.
- Choi, S. W., J. S. Lee, J. Kim, B. L. Lee, K. R. Kim, and B. C. Choi, 2015: Agrometeorological observation environment and periodic report of Korea Meteorological Administration: Current Status and Suggestions. *Korean Journal of Agricultural and Forest Meteorology* **17**(2), 144-155.
- Kang, S. S., S. H. Kim, J. W. Lee, and H. J. Kang, 2011: USN based agricultural IT convergence technology trends. *Electronics and Telecommunications Trends* **26**(6), 97-107.
- Kim, D. J., 2011: An implementation of meteorological observation network based on RFID/USN. Master Thesis, Soongsil University.
- Kim, J. H., and J. I. Yun, 2015: A thermal time - Based phenology estimation in Kimchi cabbage (*Brassica campestris* L. ssp. *pekinensis*). *Korean Journal of Agricultural and Forest Meteorology* **17**(4), 333-339.
- Kim, Y. J., S. Y. Gouk, Y. R. Kim, M. G. Lee, J. S. Kim, Y. H. Kim, K. T. Min, I. B. Ji, and J. H. Sim, 2013: The present status and development direction of smart agriculture. *Korea Rural Economic Institute*, 1-159.
- KMA (Korea Meteorological Administration), 2016: Guideline for comprehensive quality management of meteorological agency data.
- Lee, J. D., Y. E. Choi, and C. Y. Park, 2010: Developments of quality control algorithms for Korean temperature data. *Journal of climate research* **5**(2), 162-174.
- Nam, Y. U., D. W. Kim, Y. H. Lee, and Y. H. Kim, 2014: A survey on quality control of automatic weather Station. *Korea Information Science Society* **2014**(6), 16-18.
- Oh, G. L., S. J. Lee, B. C. Choi, J. Kim, K. R. Kim, S. W. Choi, and B. L. Lee, 2015: Quality control of Agro-meteorological data measured at Suwon weather station of Korea Meteorological Administration. *The Korean Society of Agricultural and Forest Meteorology* **17**(1), 25-34.
- Reek, T., S. R. Doty, and T. W. Owen, 1992: A deterministic approach to the validation of historical daily temperature and precipitation data from the cooperative network. *Bulletin of the American Meteorological Society* **73**(6), 753-762.
- Wilcox, R. R., 1996: *Statistics for the social sciences*. Academic Press.
- Yoo, N. H., G. J. Song, J. H. Yoo, S. Y. Yang, C. S. Son, J. G. Koh, and W. J. Kim, 2009: Design and implementation of the management system of cultivation and tracking for agricultural products using USN. *Korea Information Science Society* **15**(9), 661-674.