



기후변화 시나리오 자료의 통계적 상세화를 위한 결측자료 추정 기법 개발

김광형 · 조재필[†]

APEC기후센터 기후사업부 선임연구원

Development of a Gap Filling Technique for Statistical Downscaling of Climate Change Scenario Data

Kim, Kwang Hyung and Cho, Jaepil[†]

Research Fellow Climate Services and Research Division, APEC Climate Center, Busan, Korea

ABSTRACT

Long-term observational data representing the climate characteristics of the target area are necessary for statistical downscaling of climate change scenarios or seasonal forecasts. However, it is difficult to obtain high-quality observation data in developing countries due to frequent sensor failure and lack of maintenance personnel and budget. In the study, we developed a gap filling scheme for missing data to enable estimation of the climatological characteristics of a target region. For this purpose, artificially generated observation data (precipitation) with various levels of missing rates (from 0% to 43.3%), mimicking actual precipitation data from a developing country, were examined. We found that monthly climatology of precipitation was affected by missing rates greater than 5%. To fill the missing gap, we adopted estimates from nearby weather stations or satellite/reanalysis data. The missing data were estimated and utilized based on the results of correlation with nearby weather stations and the grid data from selected satellite and reanalysis data. Five climate index-based evaluations indicated that the resulting gap-filled data showed better reproducibility for climate indices of total volume of precipitation than those associated with daily patterns of precipitation such as number of rainy days and extreme values of daily precipitation.

Key words: Missing data, Precipitation, Climate change scenario, Downscaling

1. 서 론

지역 및 전 지구 기후 변화와 변동을 파악하고 영향 및 적응을 평가하기 위해서는 30년 이상 되는 장기간의 기후자료가 필수적이지만, 전 세계 많은 지역에서 품질 높은 장기간의 관측자료를 확보하는 것이 어려운 실정이다. 이렇게 기후자료의 품질을 평가하고 관리하는 것이 기후 변화와 변동을 연구할 때 가장 핵심적인 분야 중 하나이지만 많은 연구자들에게 종종 간과된 측면이 있다. 기후변화 시나리오 및 계절예측 자료를 포함한 기후정보를 농수자원 분야에 활용하기 위해서는 기후정보의 시·공간적인 상세화 (downscaling)를 필요로 한다. 기후모형을 활용한 역학적 상세화 (dynamical downscaling)

와 달리, 통계학적 상세화 (statistical downscaling)를 위해서는 대상 지역의 기후특성을 대표할 수 있는 장기 기후자료의 확보가 중요하다. 국내의 경우에는 여러 기관의 자동기상 관측장비 (Automatic Weather Station, AWS)와 기상청의 중관기상관측장비 (Automatic Synoptic Observation System, ASOS)로부터 수집된 기후자료가 존재하며, 이 중 기후변화 시나리오의 통계적 상세화를 위해서는 30년 이상의 자료 기간을 포함하는 기상청 ASOS 자료가 적합하다. 하지만 개발도상국과 같이 기상관측 기반이 열악한 지역에서는 잦은 센서의 고장, 사람에 의한 관측 오류 등으로 인하여 품질이 좋은 장기간의 관측자료 획득이 어려운 상황이다. 기후자료에 대하여 품질관리 (Quality Control, QC)를 실시한다는 것은

[†] Corresponding author: jpcho89@gmail.com (APEC Climate Center, 12 Centum7-ro, Haeundae-gu, Busan, 612-020, Korea)
Received June 25, 2019 / Revised July 18, 2019 / Accepted November 20, 2019

저장되거나 표출된 관측자료에 존재하는 오류 (error)를 찾고 걸러내어 정제된 자료를 얻는 작업을 수행하는 것을 의미한다. 하지만 기후자료에 대한 광범위한 QC는 결측을 채우기 위한 추가적인 행위를 포함하기도 한다.

앞서 언급한 것처럼 기후자료에 존재하는 일정 비율 이상의 결측은 대상 지역의 기후특성을 파악하기 위한 통계학적 분석 (상세화)에 결정적인 장애 요인이다. 따라서 기존에 많은 연구들에서는 기후자료의 결측을 가장 근소한 추정값으로 대체하는 다양한 방법을 시도하고 있다 (Domonkos, 2013). 가까운 기상관측지점의 자료를 활용하는 것이 가장 보편적으로 받아들여지고 있는 결측값 추정 방법인데 개도국과 같이 기상관측망의 밀도가 낮은 지역에서는 관측지점 간의 낮은 상관관계로 인해서 결측자료의 대체를 어렵게 한다 (Gubler et al. 2017).

위성 및 재분석 자료는 기후분야에서 광범위하게 사용되고 있지만, 낮은 정확도, 상대적으로 짧은 자료 기간, 편이 문제, 검증을 위한 in situ 자료의 필요성 등의 제한으로 인해 지상 관측값을 대체하는 용도로 사용하는 것에 대한 다양한 견해가 존재한다 (Plummer et al. 2003; Mantas et al. 2015; Schmocker et al. 2016). 그럼에도 불구하고 이를 대체할 수 있는 추정값이 존재하지 않는 경우 위성 및 재분석 자료를 기반으로 결측치를 채울 수 밖에 없다. 또한 최근 가용한 자료의 종류나 해상도가 점차 늘어나고 자료동화 (data assimilation) 기법 자체도 점차 개선되고 있기 때문에 기존에 격자자료들이 가졌던 한계점이 점차 개선이 되고 있는 실정이다 (Golian et al. 2019; Sun et al. 2018; Urraca et al. 2018).

본 연구에서는 다양한 수준의 결측이 포함된 장기 기후자료를 임의로 생산하고 이로부터 대상 지역의 기후특성을 재현할 수 있도록 결측 자료를 추정하여 보완할 수 있는 기법을 R 프로그램 (R Core Team, 2018)으로 개발하여 그 적용성을 평가하였다. 결측자료는 인근 관측지점의 기후자료나 대상 지점을 포함하는 위성 및 재분석 격자자료와의 상관성을 기반으로 추정하였다. 결측자료 추정기법을 통해 생산된 추정자료의 성능을 평가하기 위해 월별 평년값과 5개 극한기후지수를 기반으로 원자료와의 비교 평가를 수행하였고, 이를 통해 개발된 추정기법의 활용성과 한계를 제시하였다.

2. 자료 및 방법

2.1 자료

이 연구에 사용된 자료는 CMIP5 (Coupled Model

Intercomparison Project Phase 5) 기후변화 시나리오 자료에서 과거기간 (Historical)에 해당되는 30년 (1976년 1월~2005년 12월) 기간에 대한 기상청 중관기상관측지점 (ASOS)에서 관측한 강수량 자료이다. 우리나라 전체 60개 ASOS 중 10개를 무작위로 선택해 10개 ASOS 내에서 임의적으로 0%~43.3%의 순차적인 결측을 만들어 임의적인 결측자료를 구축하였다 (Fig. 1). 0%~43.3%의 결측률은 실제 APEC 기후센터의 기후자료 상세화 교육프로그램에 참여했던 개발도상국들의 기상관측자료의 불규칙한 결측 패턴을 그대로 적용한 결과이다. 본 연구에서 임의결측자료를 생성할 때는 0%~43.3%의 결측률을 가진 개발도상국 참고자료의 일단위 결측의 시간적 분포를 10개 ASOS에 동일하게 적용하여 생산하였다. 이렇게 만들어진 임의결측자료를 바탕으로 결측률이 해당 지역의 월별 기후패턴에 영향을 미치는 정도를 분석하기 위해 강수량에 대한 월별 평년값을 결측값이 전혀 없는 원본자료와 직접 비교하여 분석하였다.

2.2 결측자료 추정기법 개발 및 적용

세계기상기구 (Plummer et al. 2003; WMO 2007; 2008)에서 기상자료의 QC에 관한 지침을 제공하고 있지만, 결측자료를 추정하여 채우는 지침은 존재하지 않는다. 하지만 일부 QC 지침 중에서 의심이 되는 자료에 대해 인근 관측지점의 같은 시점의 자료를 바탕으로 품질을 평가하는 방법이 제시되어 있기 때문에, 이를 시작점으로 하여 결측자료 추정기법을 고안하였다.

앞서 기상청 10개 ASOS의 기상자료를 기반으로 만들어진 임의결측자료를 대상으로 해당 결측값을 추정하기 위해 총 3단계의 순차적인 방법을 사용하였다 (Fig. 3).

(단계1) 우선 결측값에 대해서 인근 ASOS와의 월별 일단위 상관성 분석 결과를 기반으로 결측값을 추정하여 채웠다. 대상이 되는 기상자료에 결측값이 검색되면 우선 사용자가 정한 거리 내에 있는 기상관측지점을 모두 찾게 된다. 해당 ASOS 지점들의 월별 기상변동성과 대상 지점의 기상변동성을 통계적으로 비교하여 상관관계가 일정 수준 이상이 되면 대체 기상값을 기반으로 결측값을 상관식에 따라 추정하여 채우게 된다. 계절에 따라 기상조건의 상관관계가 달라질 수 있기 때문에 월별 자료를 특정하여 상관성 분석을 수행한 후 그 결과를 기반으로 각 월의 결측값을 추정하게 하였다. 이렇게 인근 기상관측지점을 이용하는 첫번째 단계에서는 인근 기상청의 거리와 월별 상관성을 사용자가 결정하도록 했는데, 본 연구에서는 100 km 내에 있는 기상관측지점을 모두

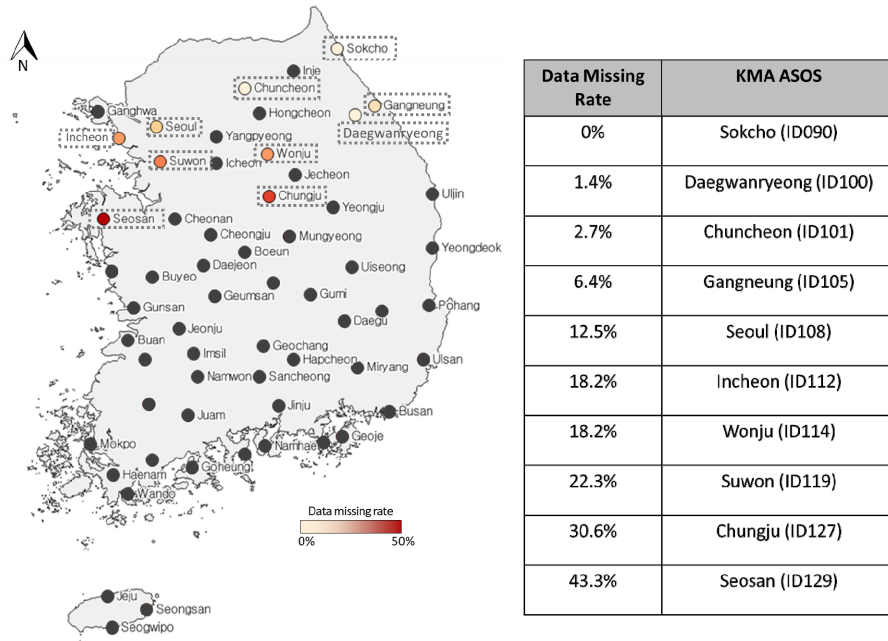


Fig. 1. Location of the 10 Automatic Synoptic Observation Systems (ASOS) for gap filling technique examination in the study (left map) and the artificially-generated missing rates for individual ASOSs (right table).

검색하여 강수량 자료의 상관계수 (Temporal Correlation Coefficient)가 0.7 이상이 되면 결측값을 추정하여 채우도록 했다. 이 방법은 주변에 가용한 기상관측지점이 존재할 때 가능한 방법인데, 우리나라와 같이 기상관측망이 밀집되어 있는 경우엔 그 가능성이 높아지지만 개도국과 같이 기상관측망이 흩어져 있는 경우에는 기상관측소 간 낮은 상관성으로 인해 적용하기 어려운 방법이다.

(단계2) 두번째 단계에서는 다양한 위성 및 재분석 자료 중에서 일단위 기후특성의 재현성 평가를 통해 선정된 격자형 자료와의 상관성 분석 결과를 기반으로 앞서 단계1에서 미처 채워지지 않은 결측값을 추정하도록 하였다. 한반도에 대한 강수량 격자자료를 제공하고 있는 위성 및 재분석 자료 중 이미 자료가 확보되거나 자료 접근성이 높은 CHIRPS (Climate Hazards group Infrared Precipitation combined with Station observations)와 PERSIANN (Precipitation Estimation from Remotely Sensed Information using Artificial Neural Networks) 등 위성기반 강수 추정자료와 MERRA2 (Modern-Era Retrospective analysis for Research and Applications version 2)와 ERA-Interim (European Centre for Medium-Range Weather Forecasts Interim Reanalysis) 등 재분석 강수량 자료를 사용하였다. 각 위성 및 재분석 자료의 자세

한 정보는 Table 1에 정리하였다. 임의결측자료의 결측값 추정을 위해서 위 4개의 위성 및 재분석 자료 중 일단위 또는 월단위 특성의 재현성 평가를 통해 가장 상관성이 높은 자료를 우선 선정하였다. 이렇게 선정된 격자자료와 임의결측자료의 일단위 상관성을 분석하여 일정 수준 이상의 상관성을 보인 경우 해당 결측값의 추정 여부를 결정하였다. 본 연구에서는 우선 대상 기상관측지점의 위경도 좌표를 바탕으로 선정된 격자자료의 겹치는 격자값을 특정하고 해당 격자값을 추출하여 관측값과의 상관성 분석을 수행한다. 두 자료 사이에 상관관계가 일정 수준 이상이 되면 해당 상관식을 바탕으로 결측값을 채우게 되는데, 이번 연구에서는 대상 기상관측지점으로부터 100Km를 영향반경으로 각 월별 강수량의 일변동에 대한 상관성 계수가 0.5 이상이 되는 경우 해당 격자값을 가져와 결측치를 채우도록 기준을 정하였다.

(단계3) 앞서 2단계를 거친 뒤에도 여전히 남은 결측값은 다시 인근 기상관측지점과의 일단위 상관성을 바탕으로 추정을 하되 첫번째 단계 보다 상관계수 임계치를 낮게 설정하는 방식으로 남은 결측값을 모두 채웠다. 본 연구에서는 상관계수 기준을 0.5로 낮춰 남은 모든 결측값이 채워지도록 하였다.

Table 1. Satellite and reanalysis data used in the study

Data type		Descriptions				Reference
		Spatial scale	Temporal scale	Data period	Update interval	
Satellite	CHIRPS	25 km	daily	1981~	monthly update with 1-monthdelay	Funk et al. (2015)
	PERSIANN	25 km	daily	1983~	Annual update	Ashouri et al. (2015)
Reanalysis	MERRA2	50 km	hourly	1980~	monthly update with 1-monthdelay	Gelaro et al. (2017)
	ERA-Interim	75 km	hourly	1979~	monthly update with 1-monthdelay	Dee et al. (2011)

Table 2. Precipitation-related climate indices used in the study

Climate Index	Description	Unit
PRCPTOP	Annual total precipitation	mm
Rx5day	Maximum 5-day precipitation	mm
R80mm	Number of days that precipitation amount is greater than 80 mm	days
CDD	Consecutive dry days	days
Intensity	Annual total precipitation / number of wet days	mm/day

2.3 결측추정 기법 적용 결과 분석

앞서 결측추정 기법을 순차적으로 적용하게 되면 인근 기상관측지점이나 위성 또는 재분석 자료와의 상관성을 기반으로 임의결측자료의 모든 결측값이 채워지게 된다. 이렇게 결측값이 모두 채워진 최종 복원자료를 원본자료와 직접 비교하여 자료의 품질이 의도한대로 향상되었는지를 판단하였다. 결측자료 추정기법의 적용성 평가는 기후변화 상세화를 위한 입력자료로 활용하기 위한 관점에서 자료의 평년기후값 재현성을 분석하였다.

먼저 기후변화 상세화 과정에서 수행되는 기후모형 자료의 편이보정 (bias-correction)을 고려하여, 최종 복원자료가 원본자료의 30년 월평균값을 얼마나 재현하는지를 평가하였다.

정량적인 개선 정도를 평가하기 위해 원본자료와 결측추정 이전의 임의결측자료 또는 최종 복원자료의 월별 평년값 사이의 Root Mean Square Error (RMSE)를 계산하였다. 사용자 중심의 기후변화 상세화의 경우 여러 개의 기후모형 중에서 지역 기후의 시공간적인 특성 및 사용자의 목적을 잘 설명하는 기후모형을 선정하기 위해 특정 기후지수의 재현성 평가를 한다 (Cho et al. 2018a; 2018b). 본 연구에서는 결측추정 기법이 대상 기후자료의 강수 총량과 일단위 패턴을 효과

적으로 복원하는지 확인하기 위해 이와 관계된 대표 극한기후지수 5개를 선택하여 최종 복원자료의 재현성을 평가하였다 (Table 2).

3. 결과 및 고찰

3.1 임의결측자료 분석 결과

10개 기상청 ASOS 지점의 강수량 관측자료에 대해 0%~43.3%의 임의결측을 적용하였다. 결측률에 따라 월별 누적강수량 패턴이 어떻게 변화하는지를 알아보기 위해 각 처리 지점의 원본자료로부터의 평년값과 임의결측자료의 평년값을 그래프로 비교하였다 (Fig. 2). 2.7%의 결측률까지는 원본자료와 별다른 차이를 보이지 않다가 6.4%의 결측률부터 임의결측자료의 강수량 월별 패턴이 원본자료와 달라지는 것을 알 수 있었다. 같은 방법으로 최고기온과 최저기온의 월평균값을 원본자료와 임의결측자료를 비교했지만 강수량과는 달리 43.3%의 결측률을 가진 자료에서도 평년값을 그대로 재현하는 것을 확인하였다 (Data not shown). 강수량과 달리 기온은 매년 상당히 유사한 계절변동을 보이기 때문에 결측률이 높아지더라도 평년값 자체에는 영향을 미치지 않기 때문이다.

3.2 결측자료 추정기법 적용 결과

본 연구에서는 결측자료 추정기법을 적용해 임의결측자료를 대상으로 인근 기상관측망이나 위성 또는 재분석 자료로부터 총 3단계의 순차적인 결측값 추정을 수행하였다.

Fig. 3에서는 임의결측자료 중 43.3% 결측률을 가진 서산 ASOS 지점 (ID129)에 대한 총 3 단계의 결측값 추정 과정과 각 단계 전과 후의 일단위 결측의 시간적 분포의 변화를 결측률 (missing rate)과 함께 그래프로 나타냈다. 결측값을 일괄적으로 -99로 표시하였기 때문에 그래프 상에서 -99를 나타낸 값이 각 단계를 거치면서 결측률이 변화함에 함께 줄어들

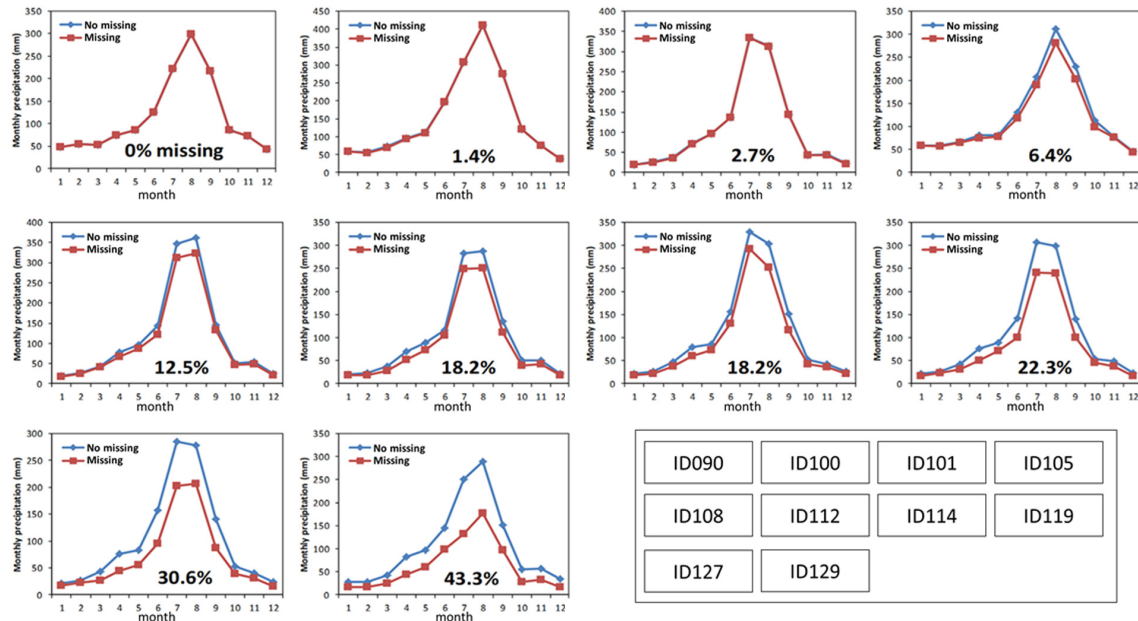


Fig. 2. Comparison of the monthly climatologies for 30 years (1976-2005) of the artificial missing data (Missing) and original data (No missing) of 10 individual ASOSs.

고 있음을 확인할 수 있다. 43.3%의 결측률을 가졌던 서산 지점의 경우 인근 기상관측지점과의 높은 강수 상관성을 기반으로 대부분의 결측값이 채워져 결과적으로 1.1%의 결측률만 남은 상태가 되었다. 우리나라의 경우 기상관측망의 밀도가 높기 때문에 첫번째 단계에서 대부분의 결측률이 채워질 수 있었지만, 개발도상국의 경우 상대적으로 가용한 기상관측지점이 적어 유사한 결과를 얻기는 어려울 것이다. 따라서 이런 결점을 보완할 수 있는 전지구 규모의 관측 추정자료인 위성 및 재분석 자료를 활용할 필요가 있다.

두번째 단계인 위성 및 재분석 자료 기반 결측자료 추정방법을 적용하기 위해 4개의 위성 및 재분석 자료 중 일단위 또는 월단위 특성의 재현성 평가를 통해 가장 상관성이 높은 자료를 우선 선정하였다. Figure 4와 같이 MERRA2가 다른 위성 및 재분석 자료에 비해서 상대적으로 높은 일단위 재현성을 보였다. 월단위 재현성은 MERRA2가 CHIRPS와 비슷한 성능을 보였지만, 본 연구에서는 일부 일단위 결측값을 추정하여 최종 복원자료의 월평균을 원본자료와 근접하게 만드는 것이 주 목적이므로 일단위 재현성 결과에서 더 좋은 결과를 보인 MERRA2 자료를 최종 선정하였다. MERRA2를 기반으로 서산 지점의 결측값을 추정한 결과, 결측률이 1.1%에서 0.5%로 줄어들었다. 여전히 남은 0.5%의 결측값은 인근 기상

관측지점과의 상관계수가 0.5이상인 경우에 결측값을 채우도록 하여 최종적으로 결측률이 0%인 복원자료를 생산하였다.

3.3 결측추정 결과 분석

결측추정 품질관리의 모든 단계를 거친 자료가 이전 임의 결측자료에 비해 얼마나 개선이 되었는지 평가하기 위해 원본자료와 결측추정 이전 또는 이후의 임의결측자료의 월별 평년값 사이의 RMSE를 통해 분석하였다. 서산 ASOS 지점의 경우, 결측추정 전에는 RMSE가 57.2 mm 정도로 크게 나타났다. 이후에는 1.23 mm로 RMSE가 크게 개선되었다. 원본자료와 최종 복원자료의 월별 평년값의 직접 비교에서도 43.3% 결측값에 의해서 심각하게 변형되었던 월별 평년값의 패턴 (Fig. 2)이 원본관측자료의 그것과 거의 유사하게 개선되는 것을 확인하였다 (Fig. 5). 43.3% 보다 적은 결측률을 가졌던 다른 ASOS 지점의 임의결측자료 역시 원본자료와 거의 유사한 월별 평년값 패턴으로 복원되었음을 확인하였다.

강수와 관계된 극한기후지수 5개를 바탕으로 최종 복원자료의 재현성을 평가한 결과, 기후지수의 성격에 따라 다른 결과를 보이는 것을 확인하였다 (Fig. 6). 예를 들어, 강수의 총량과 관련된 기후지수들 (PRCPTOP, Rx5day, Intensity)에서는 원본자료와 유사하게 복원이 된 반면, 최대무강우일수

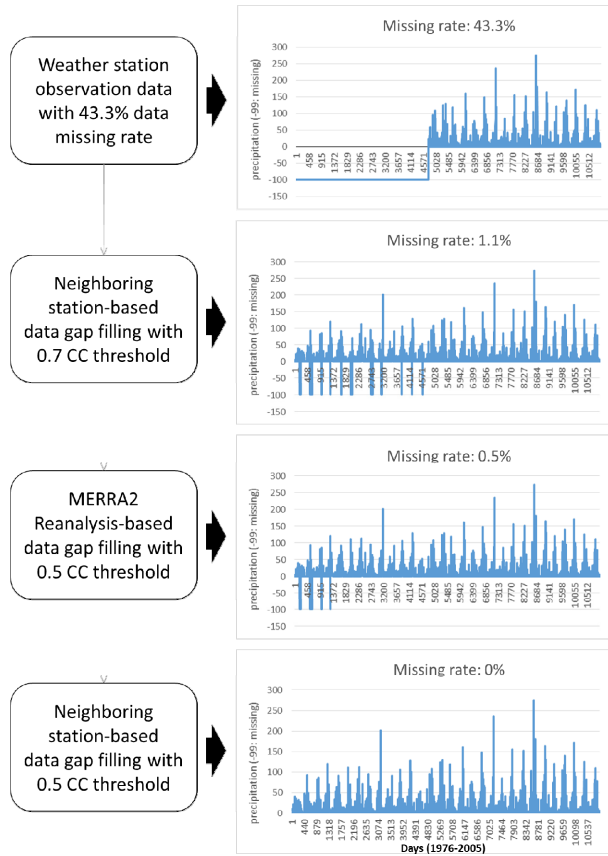


Fig. 3. Stepwise application of the gap filling technique for the artificially-generated missing data of Seosan ASOS (ID129) with 43.3% missing rate. Note that daily missing data (-99) are gradually filled after each step of the gap filling technique, resulting in missing rate change from 43.3% to 0%.

(CDD) 및 일 강수량 (R80mm)과 같이 강수의 일단위 패턴과 관련된 기후지수들에서는 최종 복원된 자료가 원본자료를 제대로 복원하지 못하는 결과를 보였다. 80 mm 이상 강수량 (R80 mm)의 경우 복원자료가 원본자료에 비해 과소 추정을 하고 (Fig. 6b and 6g), 최대무강우일수 (CDD)의 경우 원본자료에 비해 과대 추정을 하고 있었다 (Fig. 6c and 6h). 이는 복잡한 지형으로 인해 국지적으로 비가 집중해서 내리는 우리나라의 강우 패턴을 인근 기상관측지점의 관측값이나 위성 또는 재분석 자료로 복원하는데 한계가 있음을 나타낸다. 실제로 많은 위성 기반 강수량 추정자료나 재분석 자료가 기상요소 극값에 대한 과소 모의 한계를 가지고 있음이 보고된 바 있다 (Funk et al. 2015; Gelaro et al. 2017). 그럼에도 불구하고 월별 평년값에 좀더 영향을 미치는 나머지 3개의 기후지

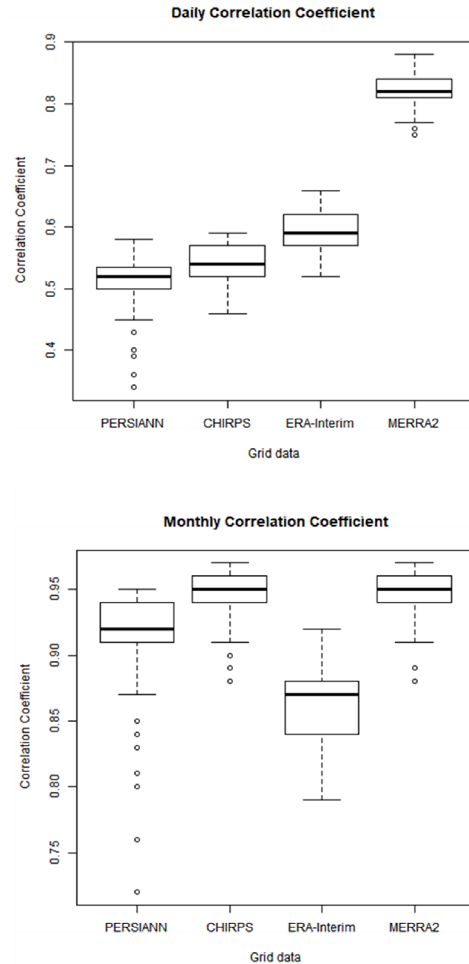


Fig. 4. Daily and monthly correlation coefficients to evaluate the reproducibility of four satellite and reanalysis data (PERSIANN, CHIRPS, ERA-Interim, MERRA2) for the climate observation data of Seosan ASOS (ID129).

수에서 높은 복원력을 보였다라는 점에서 본 결측자료 추정 방법이 기후변화 상세화 과정에서 필요한 통계적 편이보정에 긍정적인 영향을 미칠 것이라는 결론을 얻을 수 있었다.

4. 결론

기후자료의 품질 문제는 기후정보를 바탕으로 수행하는 다양한 연구 결과의 불확실성에 기여한다. 따라서 가장 기본적인 중요한 문제임에도 불구하고 지금까지 간과된 연구 분야 중 하나이다. 특히 기후정보의 통계적인 상세화를 위해서는 장기간의 끊임없는 기후자료가 필요한데, 이런 자료의

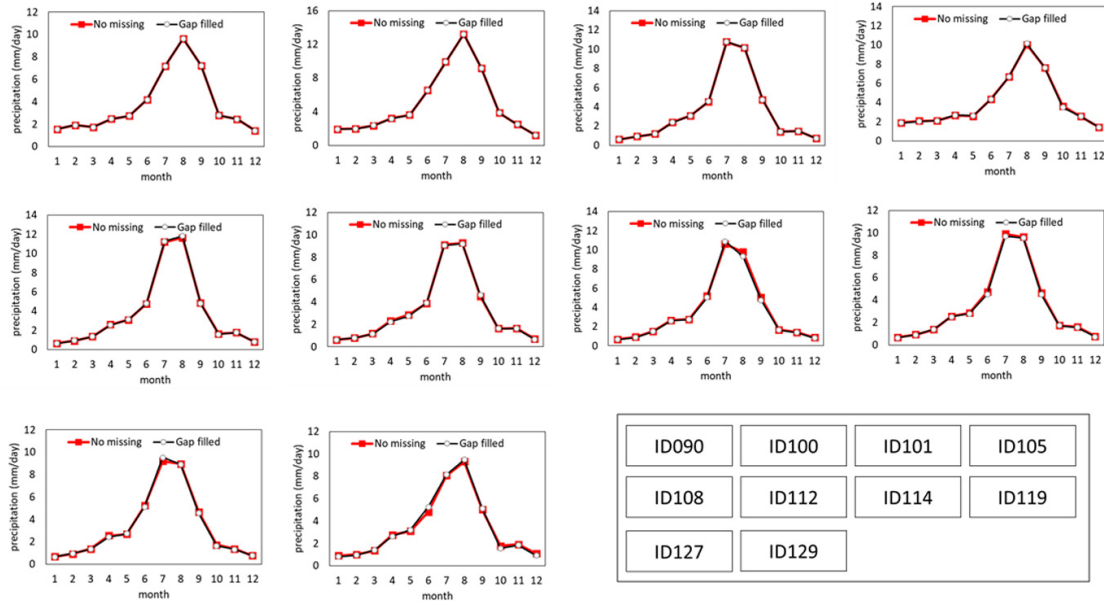


Fig. 5. Comparison of the monthly climatologies for 30 years (1976-2005) of the artificial missing data (Gap filled) and original data (No missing) of 10 individual ASOSs after the application of gap filling technique.

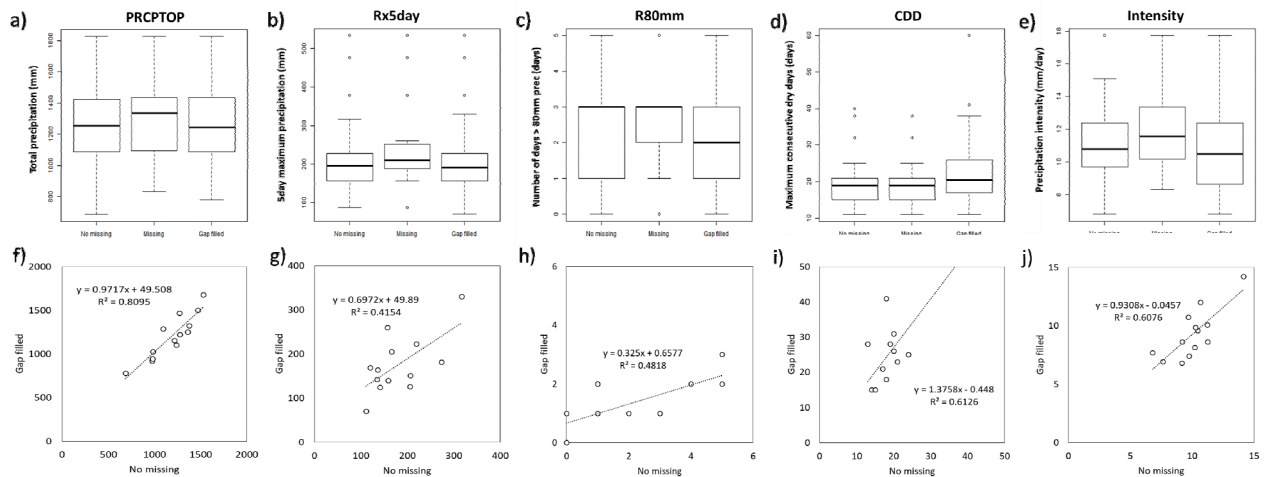


Fig. 6. Box plots of five precipitation-related climate indices (a:PRCPTOP, b:Rx5day, c:R80mm, d:CDD, e:Intensity) of original climate data (No missing), 43.3% missing data (Missing), and gap filled data (Gap filled), and regression analyses (f:PRCPTOP, g:Rx5day, h:R80mm, i:CDD, j:Intensity) of the original climate data (No missing) and the gap filled data (Gap filled) of Seosan ASOS (ID129) for the 30 years of 1976-2005.

오류와 빈번한 결측은 다음 단계의 분석을 어렵게 만든다 (WMO 2008). 따라서 관측자료의 품질관리의 여러 단계 중 하나로 자료의 결측값을 가장 근소한 추정값으로 대체하는 다양한 방법이 시도된 바 있다 (Domonkos 2013; Gubler et al. 2017). 본 연구에서는 가장 보편적으로 받아들여지고 있

는 상관성을 바탕으로 인근 기상관측지점의 자료를 활용하는 기법에 더해, 개도국과 같이 기상관측망의 밀도가 낮은 지역에서의 활용성을 고려하여 위성 및 재분석 자료와의 상관성을 바탕으로 대상 관측자료의 결측값을 채우는 방법을 시도 하였다. 결과적으로 결측률이 상대적으로 큰 케이스에서도

기존 관측자료의 월별 평년값 분포를 근사하게 복원하였다. 극한기후지수를 바탕으로 한 평가에서는 강수의 총량은 잘 복원하였지만 강수의 일단위 패턴을 복원하는데 한계가 있음을 확인하였다. 본 연구에서 사용된 결측자료 추정방법이 지속적으로 성능이 향상되고 있는 위성 및 재분석 자료 기반 격자자료와 결합되어 개도국의 관측기후자료를 복원하고 결과적으로 기후변화와 관련된 통계적 상세화의 신뢰도를 높이는 데 기여를 하기를 기대한다.

사 사

본 연구는 환경부/한국환경산업기술원의 지원으로 수행되었음 (과제번호 83076)

REFERENCES

- Ashouri H, Hsu K-L, Sorooshian S, Braithwaite DK, Knapp KR, Cecil LD, Nelson BR, Prat OP. 2015. PERSIANN-CDR: daily precipitation climate data record from multi-satellite observations for hydrological and climate studies, *Bulletin of the American Meteorological Society*, 96, p.69-83.
- Dee DP, Uppala SM, Simmons AJ, Berrisford P, Poli P, Kobayashi S, ... & Bechtold P. 2011. The ERA Interim reanalysis: Configuration and performance of the data assimilation system, *Quarterly Journal of the Royal Meteorological Society*, 137 (656), p.553-597.
- Eischeid JK, Pasteris PA, Diaz HF, Plantico MR, Lott NJ. 2000. Creating a serially complete, national daily timeseries of temperature and precipitation for the western United States, *Journal of Applied Meteorology and Climatology*, 39, p.1580-1591.
- Funk C, Peterson P, Landsfeld M, Pedreros D, Verdin J, Shukla S, Husak G, Rowland J, Harrison L, Hoell A. 2015. The climate hazards infrared precipitation with stations—a new environmental record for monitoring extremes, *Scientific Data*, 2, p.150066.
- Gelaro R, McCarty W, Suárez MJ, Todling R, Molod A, Takacs L, ... & Wargan K. 2017. The modern-era retrospective analysis for research and applications, version 2 (MERRA-2), *Journal of Climate*, 30 (14), p.5419-5454.
- Golian S, Javadian M, Behrangi A. 2019. On the use of satellite, gauge, and reanalysis precipitation products for drought studies, *Environmental Research Letters*, 14, p.075005.
- Hubbard KG. 2001. Multiple station quality control procedures. In *Proceedings of Automated Weather stations for Applications in Agriculture and Water Resources Management: Current Use and Future Perspectives*, Hubbard KG, Sivakumar MVK (eds), Lincoln, Nebraska, USA, 6 - 10 March 2000, p.133-136.
- Korea Meteorological Administration. 2011. *Quality Control Manual for National Climate Data*, p.44.
- R Core Team. 2018. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL <https://www.R-project.org/>.
- Reek T, Doty SR, Owen TW. 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, p.753-762.
- Sun Q, Miao C, Duan Q, Ashouri H, Sorooshian S, Hsu KL. 2018. A review of global precipitation data sets: data sources, estimation, and intercomparisons, *Reviews of Geophysics*, 56 (1), p.79-107.
- Urraca R, Huld T, Gracia-Amillo A, Martinez-de-Pison FJ, Kaspar F, Sanz-Garcia A. 2018. Evaluation of global horizontal irradiance estimates from ERA5 and COSMO-REA6 reanalyses using ground and satellite-based data, *Solar Energy*, 164, p.339-354.
- World Meteorological Organization (WMO). 2007. *Climate Data Management Systems*, WMO/TD No. 1176. World Meteorological Organization, Geneva, Switzerland.
- WMO. 2008. *Guide to Meteorological Instruments and Methods of Observation (7th edn)*, WMO No. 8. World Meteorological Organization, Geneva, Switzerland.
- Mantas VM, Liu Z, Caro C, Pereira AJSC. 2015. Validation of TRMM multi-satellite precipitation analysis (TMPA) products in the Peruvian Andes, *Atmospheric Research*, 163, p.132-145.
- Plummer N, Allsopp T, Lopez JA. 2003. *Guidelines on climate observations networks and systems*, WCDMP No 53, WMO/TD No 1186. World Meteorological Organization, Geneva, Switzerland.
- Schmocker J, Liniger HP, Ngeru JN, Brugnara Y,

Auchmann R, Brönnimann S. 2016. Trends in mean and extreme precipitation in the Mount Kenya region from observations and reanalyses, *International Journal of Climatology*, 36, p.1500-1514.