Kim, D., & Han, D. (2018). Comparative study on long term climate data sources over South Korea. *Journal of Water and Climate Change*. Advance online publication. https://doi.org/10.2166/wcc.2018.032

Title : Comparative Study on Long Term Climate Data Sources over South Korea

Short Title : Long Term Climate Dataset Comparison over South Korea

By

Dong-Ik Kim[*] and Dawei Han

Water and Environment Research Group, Department of Civil Engineering,
University of Bristol, Bristol, UK

---

[*] : Queen's Building, University Walk, Clifton, Bristol, UK, BS8 1TR (Corresponding Author, E-mail : dikim0127@gmail.com)

Abstract

Long term climate data are vitally important in reliably assessing water resources and water related hazards, but in-situ observations are generally sparse in space and limited in time. Although there are several global datasets available as substitutes, there is a lack of comparative studies about their suitability at different parts of the world. In this study, to find out the reliable century-long climate dataset in South Korea, we first evaluate multi-decadal reanalyses (ERA-20cm, ERA-20c, ERA-40 and 20th century reanalysis (20CR)) and gridded observations (CRUv3.23 and GPCCv7) for monthly mean precipitation and temperature. In the temporal and statistical comparisons, CRUv3.23 and GPCCv7 for precipitation and ERA-40 for temperature perform the best, and ERA-20c and 20CR also indicate meaningful agreements. For ERA-20cm, it has only a statistical agreement, but the mean has the difficulty in representing its ensemble. This paper also shows that the applicability of each dataset may vary by region and all products should be locally adjusted before applied in climate impact assessments. These findings not only help to fill in the knowledge gaps about these datasets in South Korea but also provide an useful guideline to the readers on the comparative performance of the global datasets in this part of the world.

Keywords : climate data sources, interanuual variability, non-parametric trend test, skill score, reanalysis

Introduction

To adapt and mitigate climate change, it is essential to analyse the reliable long-term climate dataset. Although the gauged local data are generally considered as the best values, they are usually sparse and limited in the time range (Simmons et al. 2004; Becker et al. 2013). For this reason, the availability of the highly accessible and reliable gridded dataset has been developed since 1980s, and some research groups, such as the Climate Research Unit (CRU) and the Global Precipitation Climatology Centre (GPCC), have constructed the monthly precipitation or temperature dataset by applying their own interpolation methods based on observations worldwide (Chen et al. 2002; Becker et al. 2013; Harris et al. 2014). They have had an important role in trend analysis in areas lacking local observations and global climate change analysis (Nicholson et al. 2003; Fekete et al. 2004; Dinku et al. 2008; Zhang & Zhou 2011; Nikulin et al. 2012). The other surrogates for local observations are reanalysis products derived using modern data assimilation techniques, which have been increasingly applied in climate impact studies. Representatively, the European Centre for Medium-Range Weather Forecasts (ECMWF) and the National Oceanic and Atmospheric Administration (NOAA) have produced these kinds of products. Initially, most reanalysis datasets were only able to cover from the mid-twentieth century to present (Compo et al. 2011), e.g. the first National Centers for Environmental Prediction/National Center for Atmospheric Research reanalysis (NCEP/NCAR) : 1948-present (Kalnay et al. 1996); ECMWF 45-year reanalysis (ERA-40) : 1975-2002 (Uppala et al. 2005); Japan Meteorological Agencies Reanalysis (JRA-25) : 1979-present (Onogi et al. 2007); ECMWF reanalysis interim (ERA-interim) : 1989-present (Dee & Uppala 2009; Dee et al. 2011b). However, a few of recent reanalyses such as NOAA 20th century reanalysis v2c (20CR), ECMWF 20th century atmospheric model ensemble (ERA-20cm), and ECMWF 20th century assimilating surface observations only (ERA-20c) extended the data period up to the whole 20th century (Compo et al. 2011; Hersbach et al. 2015; Poli et al. 2016). Moreover, those reanalyses are able to perform on daily or sub-daily scales as well as monthly scales, while the interpolated datasets only provide monthly values. Nevertheless, because these datasets are not directly taken from observations, they have additional uncertainties (Dee et al. 2011a; Hersbach et al. 2015). Hence, it is essential to evaluate their qualities in order to use these products in the climate change study.

To examine the quality of these data sources, there have been a lot of global-, continental-, or local-scale studies. For instance, Simmons et al. (2004) compared ERA-40 and NCEP/NCAR reanalysis to CRU data for air temperature (CRUTEM2v) at 5°×5° resolution on global and continental scales and concluded that there was very similar variability between CRUTEM2v and ERA-40, especially in Northern Hemisphere from 1979 onward. In the global comparison between interpolated observations and reanalysis data with 3.75°×2.5°, Donat et al. (2014) showed that ERA-40 and ERA-interim had a better agreement than NCEP/NCAR and JRA-25 for the extreme temperature, and the reanalysis products for extreme precipitation performed with a low agreement but still correlated significantly. In the case of a national-scale evaluation, the performance over Iran was done by Raziei et al. (2011) by comparing GPCC Full Data Reanalysis Product Version 3 (GPCCv3) and NCEP/NCAR precipitation dataset, which showed that GPCCv3 could complement the observations but NCEP/NCAR had significant discrepancies before 1970s. A recent study over China by Gao et al. (2016) statistically evaluated ERA-20cm, the latest ECMWF twentieth-century reanalysis dataset. After comparing the each ensemble at 0.5°×0.5° grids for precipitation and temperature, it was concluded that generally all ensemble simulations were able to represent the real condition on a comparable level.

It is important that comparative studies should cover a wide range of locations around the world and gaps should be filled in for the sites lacking such studies so that a clear pattern could be understood. In South Korea, the long-term climate trend analysis on precipitation and temperature has generally been based on the observed values (Chung & Yoon 2000; Chung et al. 2004; Chang & Kwon 2007; Bae et al. 2008; Jung et al. 2011) and the time range of these studies were limited in the late 20th century. There were a few trials to apply the interpolated datasets or reanalysis products on the climate trend research over Korea, but these datasets were applied to estimate the features of the comparable region like East-Asia, not South Korea itself (Ho et al. 2003; Jeong et al. 2015; Choi et al. 2016). In other words, the climate datasets were used in Asian area to compare with the climate trend of Korea examined by the daily observation data. In South Korea, the number of stations over 50 years is less than 15, although there are hundreds of gauging stations that have been installed. For this reason, the time period for climate impact assessment has been limited up to the mid-twentieth century in South Korea. Thus, if the researchers would like to extend the study period, it is essential to attempt to find out the reliable long-term dataset with high resolution, which should be explored. However, as aforementioned, there has not been a lack of evaluation for the reliability and applicability of century-long reanalyses as well as observation-based global climate data over South Korea.

Given this background, this study has selected several century-long precipitation datasets (ERA-20cm, ERA-20c, 20CR, CRU TS v.3.23 (CRUv3.23) and GPCC Full Data Reanalysis Product Version 7 (GPCCv7)) and temperature datasets (ERA-20cm, ERA-20c, 20CR and CRUv3.23), covering the whole 20^{th} century. ERA-40 has also been considered as a benchmark for a half century reanalysis. By estimating the temporal variability, trend and statistical agreement for monthly values of each dataset in South Korea, this study focuses on the applicability, uncertainty and limitation of those multi-decadal datasets in the country-scale climate change study. For evaluation, we have assessed correlation coefficient r, the significance of trend by the Mann-Kendall test, and the skill score based on the probability density functions (PDFs). The specification of the datasets and methodology applied in this study are introduced at first and the main results for precipitation and temperature are followed. Finally, the discussion and conclusions are presented.

Data

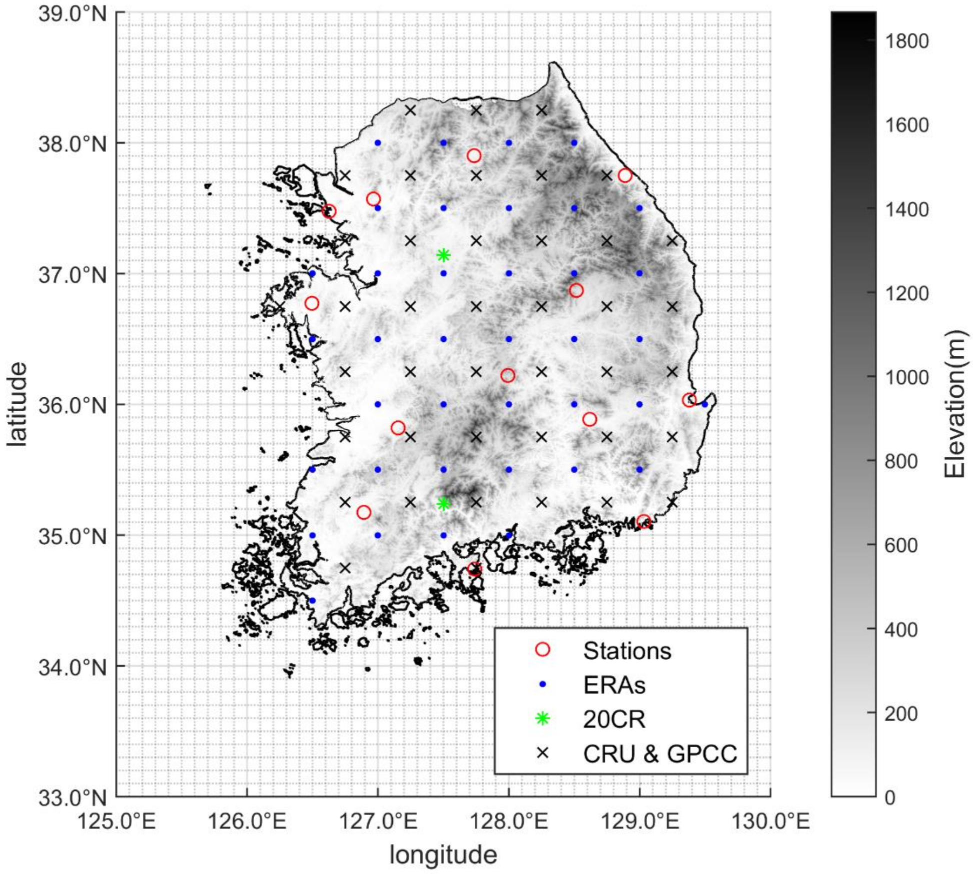
Observed Local Data

To analyse the precipitation and temperature change over the mainland of South Korea, daily total precipitations and daily mean 2-m air temperatures of 13 ground gauge stations, spanning 1961-2010, are taken from the data archive of Korea Meteorological Administration(KMA) (https://data.kma.go.kr/cmmn/main.do) and merged to the monthly values. In order to compare the datasets for the common period, the stations are evenly selected excluding islands of Korea from 1961 to 2001 with no empty values, although three of them are available from 1966, 1968 and 1973, separately. The quality of the observations is strictly controlled by KMA. The detailed information on the location and data period of the stations is given in Figure 1 and Table 1.

Table 1 Longitude, latitude and observation period of the selected stations

No.	Name	Longitude	Latitude	Observation Period	Elevation(m)
1	Seoul	126-57-56 E	37-34-17 N	1961-2010	11.1
2	Incheon	126-37-29 E	37-28-39 N	1961-2010	69.6
3	Seosan	126-29-45 E	36-46-25 N	1967-2010	30.3
4	Chuncheon	127-44-08 E	37-54-09 N	1966-2010	79.1
5	Gangneung	128-53-27 E	37-45-05 N	1961-2010	27.4
6	Jeonju	127-09-17 E	35-49-17 N	1961-2010	54.8
7	Chupungnyeong	127-59-40 E	36-13-11 N	1961-2010	246.1
8	Yeongju	128-31-00 E	36-52-18 N	1973-2010	212.2
9	Gwangju	126-53-29 E	35-10-22 N	1961-2010	73.8
10	Yeosu	127-44-26 E	34-44-21 N	1961-2010	66.0
11	Daegu	128-37-08 E	35-53-06 N	1961-2010	65.5
12	Pohang	129-22-46 E	36-01-57 N	1961-2010	3.7
13	Busan	129-01-55 E	35-06-16 N	1961-2010	71.0

Figure 1 Locations of 13 gauge stations shown in Table 1 and gridded points of ERAs (ERA-20cm, ERA-20c and ERA-40), 20CR, CRUv3.23 (CRU) and GPCCv7 (GPCC)



Reanalysis Data

ERA-20c is the first atmospheric 20th century reanalysis of the ECMWF. This dataset, covering 1900-2010, is produced by assimilating observations of surface pressure and surface marine winds only (Poli et al. 2016). Considering the data availability and resolution of other datasets, we extract total precipitation from the 24 hour accumulated forecasts and 2-m air temperature from 6 hourly analysis data with 0.5°×0.5° grid from January 1901 to December 2010 via the ECMWF web server. The products in South Korea are accumulated into monthly data and the values over the sea are excluded.

In addition to ERA-20c, the ECMWF also released ERA-20cm data with 10-member ensemble from January 1900 to December 2010 (Hersbach et al. 2015). Comparing with ERA-20c, this dataset was produced with the same Integrated Forecasting System(IFS) version Cy38r1, but it includes no data assimilation (Donat et al. 2016). 3-hourly total precipitation and temperature data with 0.5°×0.5° grid from January 1901 to December 2010 are extracted from the web server and they are calculated as inland monthly datasets. In this study, to explore the general feature of ERA-20cm ensemble, we use the ensemble mean and the ensemble member 0 (hereafter "En0") only. A more detailed assessment on all ten ensemble members will be covered in another study.

To find out the difference among the ECMWF products, another data called ERA-40, the 45-year reanalysis data from September 1957 to August 2002 (Uppala et al. 2005), are extracted from the ECMWF archive in the same way as ERA-20c. We collect the 6-hourly convective precipitation data, large-scale precipitation data and 2-m air temperature data at 0.5°×0.5° grid from January 1961 to December 2001. The total precipitation is produced by the sum of convective and large-scale precipitation excluding the values on the sea and the products are aggregated into monthly data.

20CR is one of the long term reanalysis datasets provided by the NOAA. Its latest version 2c, spanning 1850 to 2014 with the 1.875°×1.9° resolution, is produced by assimilating only surface pressures and using Ensemble Kalman Filter technique to produce 56 ensemble members (Donat et al. 2016). Because each ensemble dataset is not available in the web server, we collect only 8-times daily ensemble means for total precipitation and 2m air temperature from 1901 to 2010 and accumulate them on a monthly basis. As with other datasets, the data over the sea are ignored.

Gridded observations by CRU and GPCC

CRU TS v.3.23 (CRUv3.23) is the recently updated time-series land-only dataset from 1901 to 2014, which covers all over the world except the Antarctic (Harris et al. 2014). This dataset constructed by using the Climate Anomaly Method based on the worldwide observations provides monthly total precipitation and monthly mean 2-m air temperature with its highest resolution (0.5°×0.5° latitude/longitude) (Harris et al. 2014). In this paper, for the comparison with the observations and reanalysis dataset, the data over South Korea from 1901 to 2010 are extracted.

GPCC has produced the global land-surface precipitation data, and its recent version, GPCC Full Data Reanalysis Version. 7.0 (GPCCv7), covers a 111-year analysis period from 1901 to 2013 based on the rain gauge database over 51,000 stations worldwide (Schneider et al. 2015). In this study, the monthly total precipitation product with its highest resolution of 0.5°×0.5° over South Korea from 1901 to 2010 is taken from this dataset.

Methodology

Evaluation of interannual variability

To explore the temporal strength of the linear relationship between the model products and the observed values, the Pearson's linear correlation coefficients(r) mean between the products and the observations of 13 stations from 1961 to 2001 are calculated. This method has been widely used to measure the degree of collinearity between the observed and the modelled data in the multi-decadal climate variability studies, although it is oversensitive to high extreme values and insensitive to proportional gaps between two variables (Legates & McCabe 1999; Deser et al. 2004; Herrmann et al. 2005; Dickinson et al. 2006; Wu et al. 2010; Gholami et al. 2015; Wang et al. 2015). Here, we focus on the variability between the observation and the modelled datasets using r, while the absolute differences between them are simply explored through figures on seasonal/annual change.

For this analysis, the seasonal/yearly total precipitation and mean temperature variables are derived from all the datasets. Every seasonal dataset is collected for Spring from March to May, Summer from June to August, Autumn from September to November, and Winter from December to February.

In case of r, considering the difference in coordinate and resolution between datasets (Figure 1), we have interpolated the data in each station point by using an inverse distance (ID) method, one of the most applied deterministic methods (Babak & Deutsch 2009). Compared with other preferred methods, kriging, the ID method is simple to calculate, more applicable to spatial estimation with small sized observation networks and does not require prior information like a semi-variogram model (Tomczak 1998; Lu & Wong 2008; Babak & Deutsch 2009). For this reason, this study has applied the ID method as follows :

$$\boldsymbol{w(x,y)} = \sum_{\boldsymbol{i=1}}^{\boldsymbol{N}} \boldsymbol{\alpha_i w_i}\,, \quad \boldsymbol{\alpha_i} = \frac{\left(\frac{\boldsymbol{1}}{\boldsymbol{d_i}}\right)^{\boldsymbol{p}}}{\sum_{\boldsymbol{i=1}}^{\boldsymbol{N}} \left(\frac{\boldsymbol{1}}{\boldsymbol{d_i}}\right)^{\boldsymbol{p}}} \tag{1}$$

where N is the number of the grids used in calculation, w is the evaluated value from the data product in each station point, w_i is the i-th data point among the selected values, d_i is the distance from the station to the i-th grid, and p is the specified weighting power. In this equation, the weighting parameter, p, can vary from 0 to infinite, and when the value increases, the estimated is less influenced by the further stations (Chang et al. 2006). In this analysis, all inland gridded values are used and the most common value, 2, is applied for the power p (Teegavarapu & Chandramouli 2005; Babak & Deutsch 2009). After calculating the r in 13 stations, the mean r values of them are compared.

Trend test

The correlation coefficient, r, does not represent the slope of the line of the best fit, although it shows the relationship between the observed and the model. Thus, to find out the significance of linear trends in each dataset, the Mann-Kendall test is applied for the reference period 1961 – 2001. The trends from 1901 to 2010 of several precipitation datasets (ERA-20cm, ERA-20c, 20CR, CRUv3.23 and GPCCv7) and temperature datasets (ERA-20cm, ERA-20c, 20CR and CRUv3.23) are also evaluated in order to assess the long-term patterns of them throughout the 20th century. The Mann-Kendall trend test created by Mann (1945) and Kendall (1955) is one of the widely used nonparametric tests for detecting the trend of environmental data such as precipitation, temperature and streamflow (Xu et al. 2005; Bae et al. 2008; Shadmani et al. 2012; Zang & Liu 2013). Compared with parametric tests like linear regression which require data normality as well as independence, this method only requires the independence of data (Hamed & Rao 1998; Xu et al. 2005). In the Mann-Kendall test, the test statistic S and the standardised test statistic Z are estimated by the related equations as follows :

$$\boldsymbol{S = \sum_{i=1}^{n-1} \sum_{j=i+1}^{n} sgn(x_j - x_i)} \tag{2}$$

$$\boldsymbol{sgn(x) = \begin{cases} 1 \ \ for\, x > 0 \\ 0 \ \ for\, x = 0 \\ -1 \ \ for\, x < 0 \end{cases}} \tag{3}$$

$$\boldsymbol{V(S) = \frac{n(n-1)(2n+S) - \sum_{i=1}^{m} t_i(t_i-1)(2t_i+5)}{18}} \tag{4}$$

$$\boldsymbol{Z = \begin{cases} \dfrac{S-1}{\sqrt{V(S)}} & for\, S > 0 \\ 0 & for\, S = 0 \\ \dfrac{S+1}{\sqrt{V(S)}} & for\, S < 0 \end{cases}} \tag{5}$$

where, $x_1, x_2, x_3, \dots, x_n$ are the time series of length *n*, *V(S)* is the variance of *S*, *m* is the number of tied groups, t_i is the number of ties for the *i*-th value, and *Z* follows a standard normal distribution (Xu et al. 2008). The significance of trends is evaluated by comparing *Z* with the standard normal variate at the desired significance (Hamed & Rao 1998). When $|Z| > Z_{1-\alpha/2}$, where $Z_{1-\alpha/2}$ is the standard normal deviates where the significance level is α, the null hypothesis is rejected and it means that there is a significant trend in the time series in the test. In this study, both 0.05 and 0.10, the most commonly used values, are applied for α, although significance levels can vary from 0.1 to 0.001 according to the study (Hamed & Rao 1998; Xu et al. 2005; Bae et al. 2008; Jung et al. 2011; Zang & Liu 2013). The magnitude of the linear trend for this method is estimated by Theil-Sen approach, sometimes referred to as “Kendall Slope Estimator”, defined by the median value of the ranked slope estimates as follows (Theil 1950; Sen 1968; Hirsch et al. 1982; Xu et al. 2008; Zang & Liu 2013) :

$$\boldsymbol{\beta = Median\left(\frac{x_j - x_i}{j - i}\right), \quad \forall i < j} \tag{6}$$

In this equation, the positive value of β represents the increasing trend over time, while the negative value means the opposite trend. The advantage of this method is that it is less sensitive to outliers or extreme values than the least-square method (Shadmani et al. 2012; Sayemuzzaman & Jha 2014).

PDF-based Evaluation method

To assess the statistical similarity between the observations and each dataset from 1961 to 2001, we have estimated the skill score based on the probability density function (PDF) suggested by Perkins et al. (2007). This method is very simple but powerful to capture the relative compatibility between observation and model distribution. In additional, compared with the traditional mean-based method, this performance shows more credible climate variations and it is flexible to collect data with different time periods from multiple stations (Perkins et al. 2007; Gao et al. 2016). By calculating the overlapped area between two distributions at each bin, this skill estimates how much the climate dataset distribution is similar to the observed. If a dataset matches the observed values perfectly in PDF, the skill score will be 1, which equals the sum of the probability. Otherwise, if the skill score is close to zero, it means that there is no common area between the model values and observations. In other words, the more overlapped the two curves, the closer to 1 this score is. The skill score is calculated as follows :

$$\boldsymbol{S_{score} = \sum_{1}^{n} minimum(P_m, P_0),} \tag{7}$$

where n is the number of bins for the calculation, P_m is the frequency of values in a given bin from a comparison target, and P_0 is the frequency values in a given bin from observations. In this study, the monthly variables are applied and the square root of 1mm month^{-1} for precipitation and 1°C for temperature are considered as the intervals of bins to effectively compare the PDFs like earlier studies (Perkins et al. 2007; Gao et al. 2016).

Results

Precipitation

Interannual variability

Table 2 quantitatively explains the seasonal/annual correlation between the observation and the simulated precipitation from 1961 to 2001. In the seasonal mean comparison, the r values for CRUv3.23 and GPCCv7 exceed 0.9 in every season, and ERA-20c, ERA-40 and 20CR performs moderate to high correlations($0.4 < r < 0.9$). Among seasonal values, spring and winter are more correlated than summer and autumn. However, the simulations for ERA-20cm mean and En0 are located between -0.149 and 0.313, which means that there is little temporal correlation with the observation for precipitation. The similar result is described in the annual mean comparison. CRUv3.23 and GPCCv7 perform very well with the r over 0.9 and ERA-20c follows with 0.621. 20CR and ERA-40 have the moderate correlations with 0.498 and 0.445, separately, but the r values for ERA-20cm mean and En0 are close to zero.

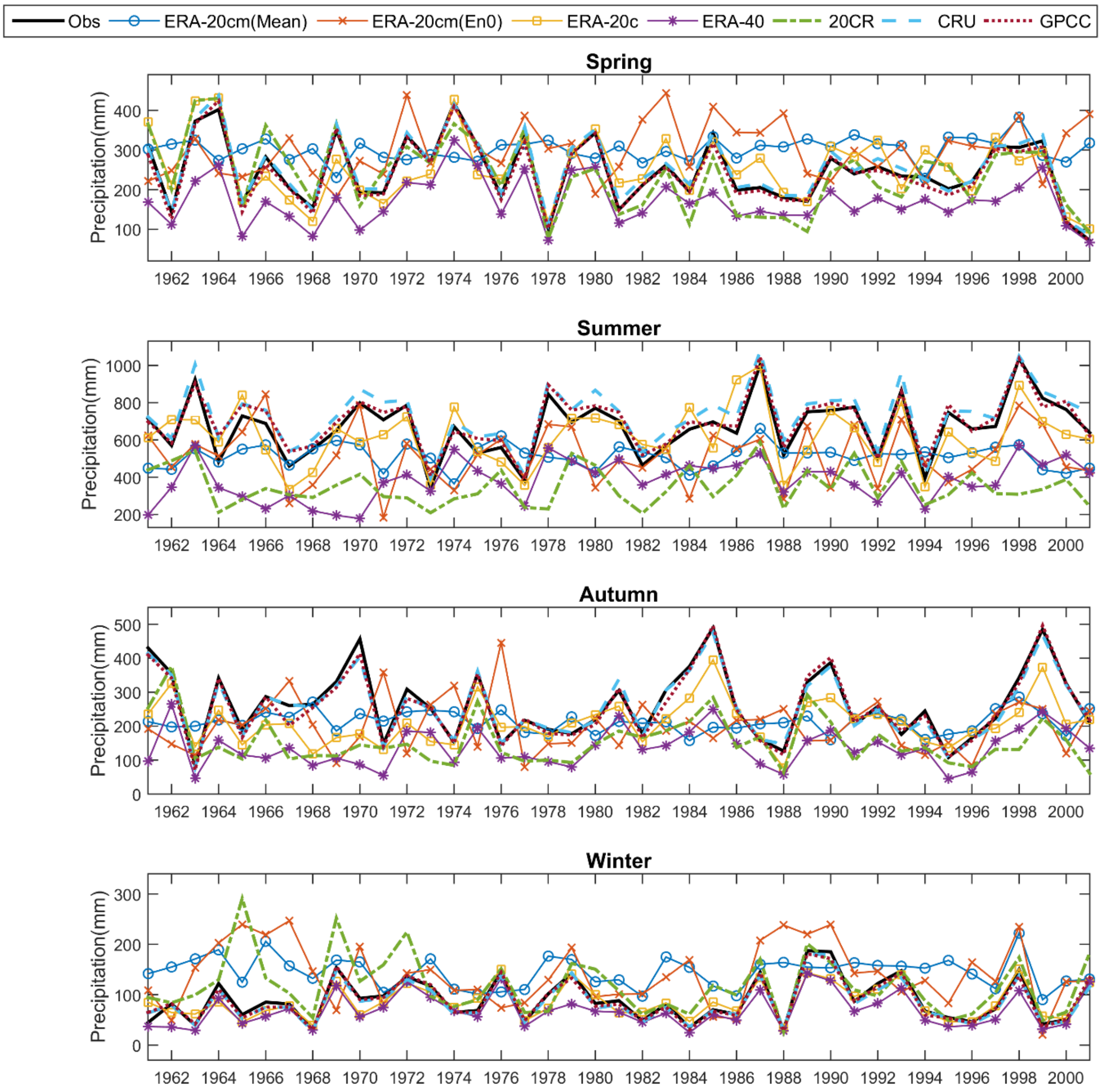
Table 2 Correlation coefficient(r) for seasonal and annual total precipitation for each dataset averaged over all regions from 1961 to 2001

Type	Seasonal comparison				Annual comparison
	Spring	Summer	Autumn	Winter	
ERA-20cm(Mean)	-0.110	0.070	0.014	0.284	0.091
ERA-20cm(En0)	0.042	0.313	-0.149	0.225	0.246
ERA-20c	0.762	0.600	0.665	0.829	0.621
ERA-40	0.821	0.466	0.647	0.883	0.445
20CR	0.744	0.407	0.562	0.638	0.498
CRUv3.23	0.963	0.922	0.942	0.960	0.929
GPCCv7	0.970	0.938	0.952	0.966	0.945

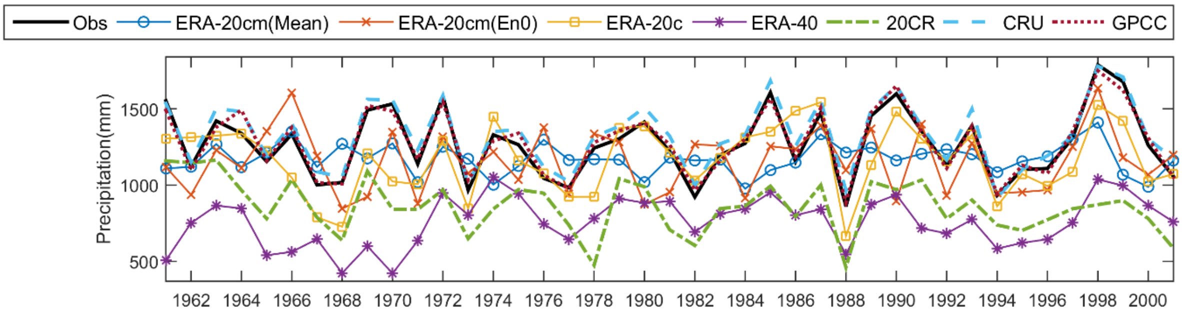
Figure 2 which illustrates the seasonal and annual precipitation change of each dataset from 1961 to 2001 supports this result. For the seasonal comparison, the fluctuations of ERA-20cm mean and En0 have little correlations with the observations in all seasons, while GPCCv7 and CRUv3.23 perform almost in similar movements with the observed values (Figure 2(a)). For ERA-20c, ERA-40 and 20CR, their movements have significant similarities to the observations, but the values themselves of each dataset are slightly different. For example, ERA-40 and 20CR have the lower rainfall than the observation, especially, in summer and autumn, whereas ERA-20c is relatively close to the observation (Figure 2(a)). This means that in terms of interannul variability, ERA-20c is less biased than ERA-40 and 20CR in South Korea. The annual change shows a similar result with the seasonal trend. The annul patterns of ERA-20cm mean and En0 are totally different from that of the observation, while CRUv3.23 and GPCCv7 perform very well (Figure 2(b)). For ERA-20c, ERA-40 and 20CR, they have the partial similarity to the observation in the annual comparison, but only ERA-20c has the equivalent value with the observed (Figure 2(b)). In other words, ERA-40 and 20CR are clearly underestimated.

Figure 2. Total precipitation change for observation (Obs), Mean of ERA-20cm (ERA-20cm (Mean)), En0 of ERA-20cm (ERA-20cm(En0)), ERA-20c, ERA-40, 20CR, CRUv3.23 (CRU) and GPCCv7 (GPCC) averaged over the whole region from 1961 to 2001

(a) The seasonal total precipitation change (from above, Spring, Summer, Autumn, and Winter)



(b) The annual total precipitation change



Long-term trend

Table 3 shows the long-term trends derived by the Mann-Kendall test. The standardised statistics

(Z) for the reference period 1961 to 2001 describe that there are no significant seasonal/annual trends at 90% or 95% confidence level for ERA-20cm, ERA-20c, CRUv3.23 and GPCCv7 as well as the observation. Only ERA-40 in summer and 20CR in spring have the increasing and the decreasing trend at 95% confidence level, separately. With 90% confidence level, a further declining trend is found in the annual trend for 20CR.

Table 3 Mann-Kendall test results for precipitation trend

Dataset	Spring		Summer		Autumn		Winter		Annual	
	Z	β	Z	β	Z	β	Z	β	Z	β
1961-2001										
Observation	- 1.00	- 1.38	1.13	2.78	- 0.51	- 0.74	0.30	0.14	- 0.08	- 0.39
ERA-20cm (Mean)	1.49	0.42	- 0.48	- 0.48	- 0.30	- 0.19	- 0.93	- 0.38	- 0.12	- 0.27
ERA-20cm (En0)	1.49	1.66	0.10	0.31	0.19	0.16	- 0.35	- 0.24	0.28	0.91
ERA-20c	- 0.33	- 0.43	0.19	0.45	1.20	1.06	0.62	0.26	0.39	1.43
ERA-40	- 0.21	- 0.16	2.01 [a]	3.44	1.43	1.02	0.39	0.16	1.47	3.95
20CR	- 1.99 [a]	- 2.83	0.15	0.21	- 0.46	- 0.27	- 0.86	- 0.75	- 1.83 [b]	- 5.11
CRUv3.23	- 1.00	- 1.43	1.20	2.36	- 0.55	- 0.78	- 0.01	- 0.02	- 0.10	- 0.42
GPCCv7	- 1.02	- 1.40	0.86	1.40	- 0.12	- 0.33	- 0.06	- 0.05	0.00	0.11
1901-2010										
ERA-20cm (Mean)	2.58 [a]	0.20	0.52	0.11	1.19	0.12	1.82 [b]	0.15	2.14 [a]	0.60
ERA-20cm (En0)	0.11	0.02	0.80	0.38	- 0.60	- 0.13	0.24	0.04	0.30	0.18
ERA-20c	4.95 [a]	1.21	3.97 [a]	2.04	4.53 [a]	1.01	5.97 [a]	0.55	6.09 [a]	5.00
20CR	- 0.11	- 0.02	- 2.19 [a]	- 0.76	- 2.32 [a]	- 0.48	1.84 [b]	0.25	- 1.76 [b]	- 1.13
CRUv3.23	0.80	0.19	3.00 [a]	1.70	1.51	0.46	- 0.68	- 0.07	2.72 [a]	2.13
GPCCv7	0.54	0.11	3.42 [a]	1.79	1.51	0.46	- 1.35	- 0.16	2.86 [a]	2.14

[a] : significant trend at the 0.05 significance level. [b] : significant trend at the 0.10 significance level.

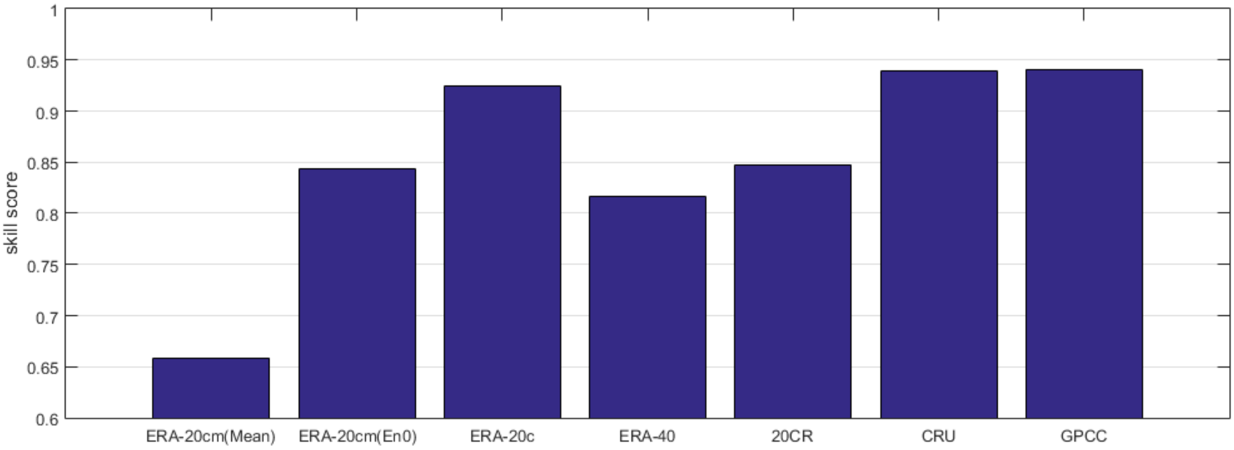
β(trends for precipitation) are in mm/yr.

The analysis from 1901 to 2010 shows more obvious trends. For ERA-20cm, the trends of the mean and En0 are different. ERA-20cm mean has the significant increasing trends in spring, winter and annual simulations, while En0 has no significant trends. Comparing ERA-20cm with CRUv3.23 and GPCCv7, they have no similarity in the seasonal trends and the magnitude of the slopes for ERA-20cm are generally lower than those of CRUv3.23 and GPCCv7 except winter. For instance, CRUv3.23 and GPCCv7 have the increasing trends in summer with the slopes of 1.70 and 1.79, but ERA-20cm mean has the upward trends in spring and winter with the slopes of 0.20 and 0.15. In case of ERA-20c, it performs the obvious increasing movement in every test and has the stronger increasing trend than CRUv3.23 and GPCCv7 in summer and annual tests. This shows that ERA-20c can exaggerate the long-term trend for precipitation than the other datasets. On the other hand, 20CR performs the downward trends in summer, autumn and annual test. That is to say, the long-term trend of 20CR is in contrast with the movements of other datasets.

Statistical comparability

Figure 3 describes the statistical agreement between the observation and each dataset from 1961 to 2001. CRUv3.23 and GPCCv7 perform the best simulations with the skill score of approximately 0.94, and ERA-20c follows them closely with 0.93. This indicates that ERA-20c has the statistical similarity with the observed at almost the same level as CRUv3.23 and GPCCv7. The scores for 20CR, ERA-40 and En0 are between 0.8 and 0.85 which shows significant agreements, whereas ERA-20cm mean has a clearly smaller value, 0.66.

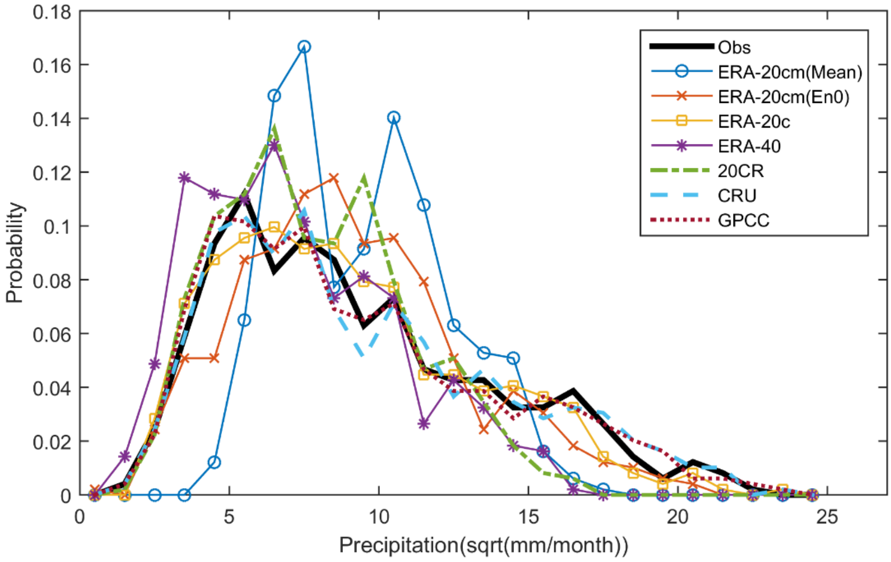
Figure 3 PDF-based skill score for monthly precipitation for the Mean of ERA-20cm (ERA-20cm (Mean)), En0 of ERA-20cm (ERA-20cm(En0)), ERA-20c, ERA-40, 20CR, CRUv3.23 (CRU) and GPCCv7 (GPCC) averaged over the whole region from 1961 to 2001



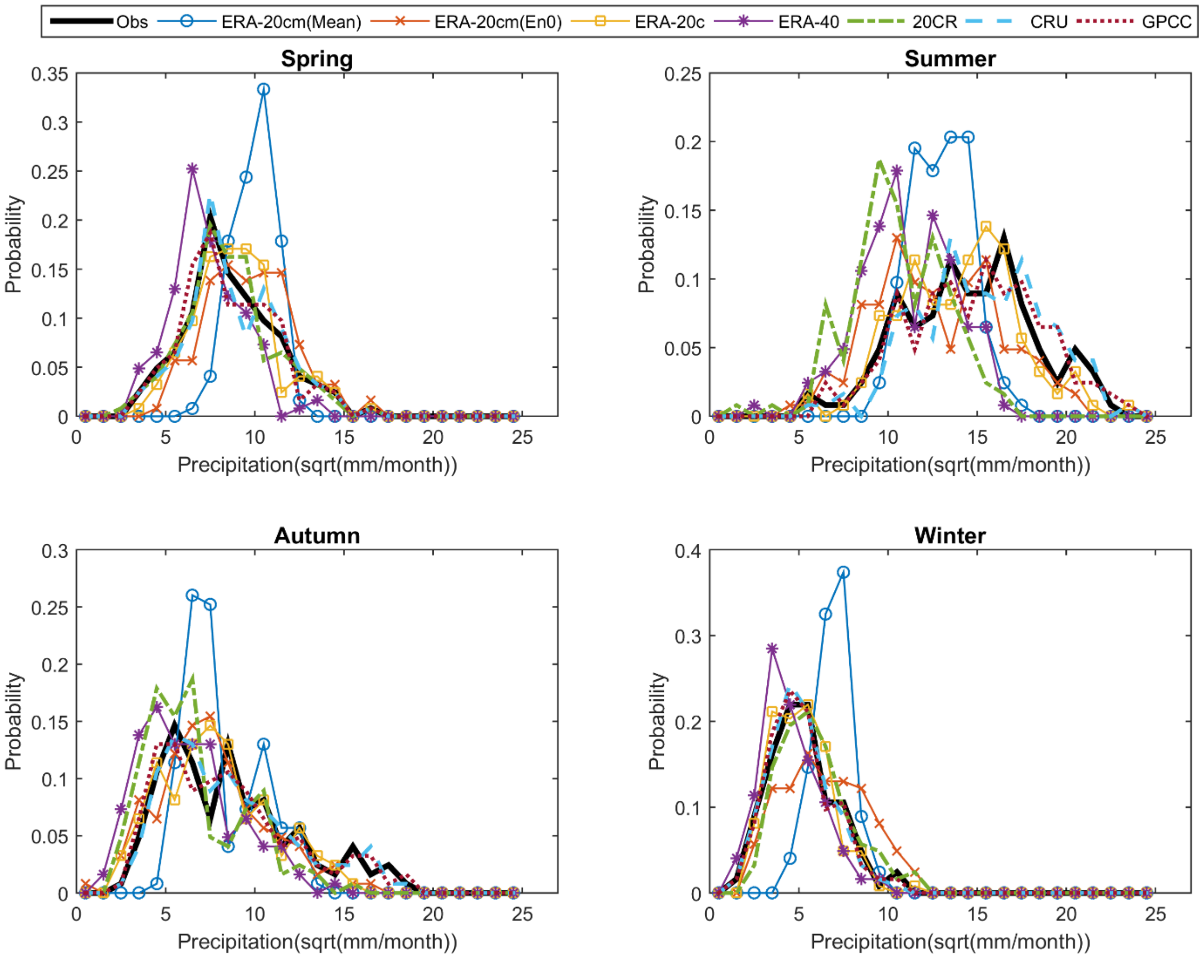
The specific discrepancies of each dataset are described in Figure 4(a) which illustrates the PDFs of the observation and each precipitation dataset over South Korea from 1961 to 2001 and Figure 4(b) which represents seasonally subdivided PDFs. It is obvious that ERA-20c as well as CRUv3.23 and GPCCv7 is one of the most fitted datasets to the observation with little discrepancies. However, the other datasets have partial gaps from the observation. For 20CR, the PDF in Figure 4(a) shows that it underestimates over 200mm month^{-1} and overestimates in the range of 25 to 100mm month^{-1}. This result is mainly due to the underestimated values in summer, as seen in Figure 4(b). The left-biased summer rainfalls lead to overestimation of moderate values and underestimation of intensive values. The PDF of ERA-40 in Figure 4(a) overall exaggerates the frequency under 50mm month^{-1} and underestimates over 200mm month^{-1}. It comes from the generally underestimated distributions in all seasons, especially in summer (Figure 4(b)). In case of ERA-20cm mean and En0, the dry months and intensive rainfall months are underestimated but the moderate months are overestimated in Figure 4(a). It is clear that the mean of ERA-20cm has this tendency more strongly than En0. This evaluation suggests that all datasets show the significant agreement with the observation, albeit some of them still need a cautious approach to use in the frequency analysis.

Figure 4 Probability density functions(PDFs) for monthly total precipitation for observation (Obs), Mean of ERA-20cm (ERA-20cm (Mean)), En0 of ERA-20cm (ERA-20cm(En0)), ERA-20c, ERA-40, 20CR, CRUv3.23 (CRU) and GPCCv7 (GPCC) over South Korea

(a) PDFs for monthly total precipitation from 1961 to 2001



(b) PDFs for seasonally subdivided monthly total precipitation from 1961 to 2001



Temperature

Interannual variability

Table 4 describes the *r* values between the gauged temperature and the model temperature from 1961 to 2001. In seasonal comparison, CRUv3.23 and ERA-40 have the highest values over 0.9 in every season and ERA-20c follows closely with 0.830 to 0.914. 20CR has the high correlations ($0.6 < r < 0.9$) and the values for ERA-20 mean and En0 are the lowest ones. To be more specific, ERA-20cm mean has the moderate correlations ($0.4 < r < 0.7$) in four seasons, while En0 has low correlations except spring. Of the four seasons, winter has the highest value except ERA-20cm mean and En0. Theses seasonal findings are similar to the annual simulations. In annual comparison, CRUv3.23 and ERA-40 show the most fitted correlations with the *r* values over 0.9 and ERA-20c closely follow them with 0.879. 20CR has the 0.808 and ERA-20cm mean (0.714) and En0 (0.523) have moderate to high correlations.

Table 4 Correlation coefficient(*r*) for seasonal and annual mean temperature for each dataset averaged over the whole region from 1961 to 2001

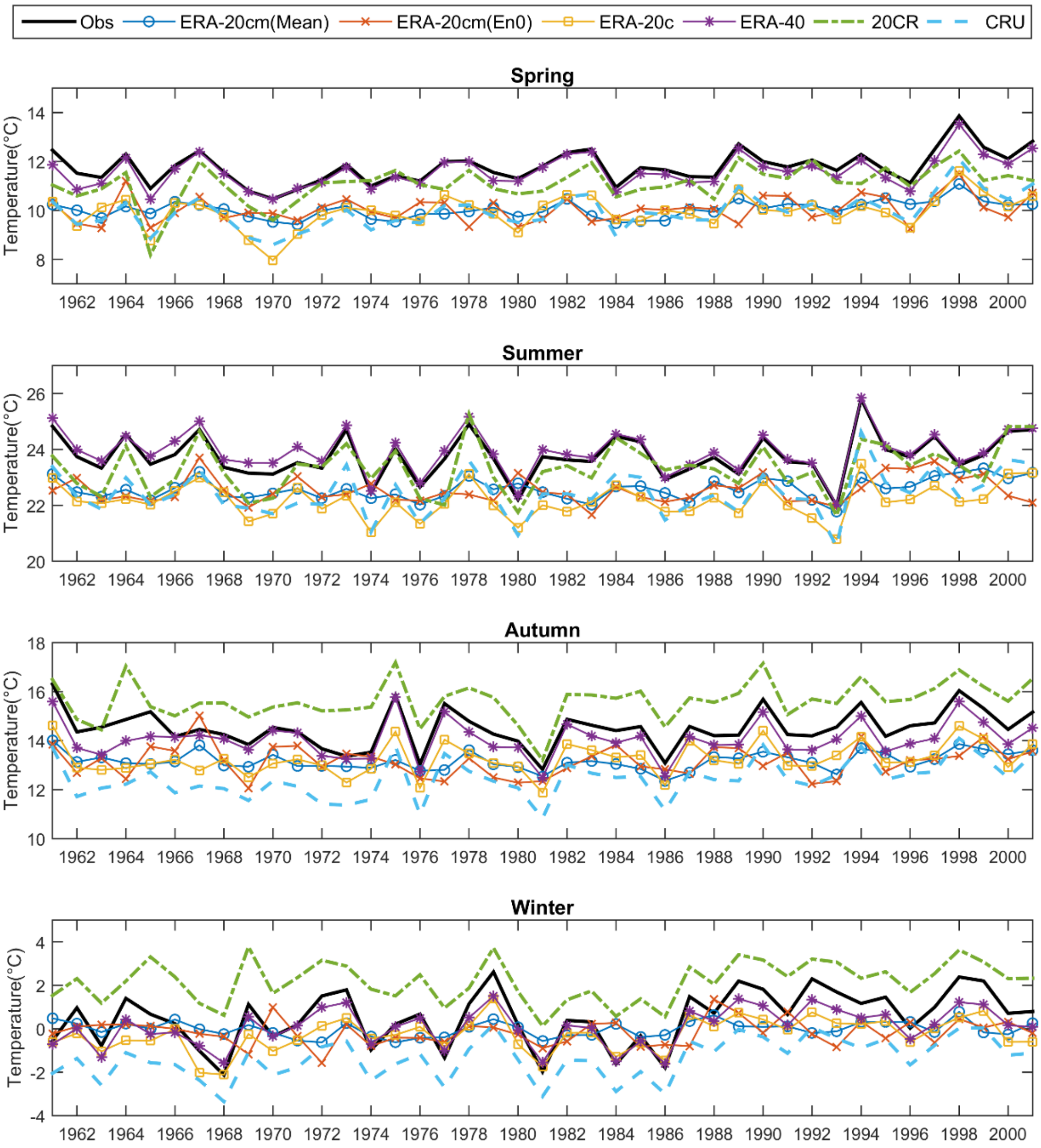
Type	Seasonal comparison				Annual comparison
	spring	summer	autumn	winter	
ERA-20cm(Mean)	0.671	0.597	0.578	0.407	0.714
ERA-20cm(En0)	0.493	0.194	0.251	0.161	0.523
ERA-20c	0.830	0.867	0.895	0.914	0.879
ERA-40	0.924	0.943	0.908	0.963	0.923
20CR	0.654	0.785	0.798	0.875	0.808
CRUv3.23	0.933	0.964	0.945	0.976	0.950

Figure 5 demonstrates the seasonal and annual mean temperature trends of each dataset over South Korea from 1961 to 2001. In Figure 5, we can see that each dataset performs the similar movements to the observations, but the values themselves are different depending on the dataset except ERA-40. For ERA-20cm and ERA-20c, the seasonal/annual variations seem to have partial correlations, but the model values are generally about 1 to 2 Celsius degrees lower than those of observations except winter season. In case of CRUv3.23, it is clear that the mean temperature for CRUv3.23 is about 2

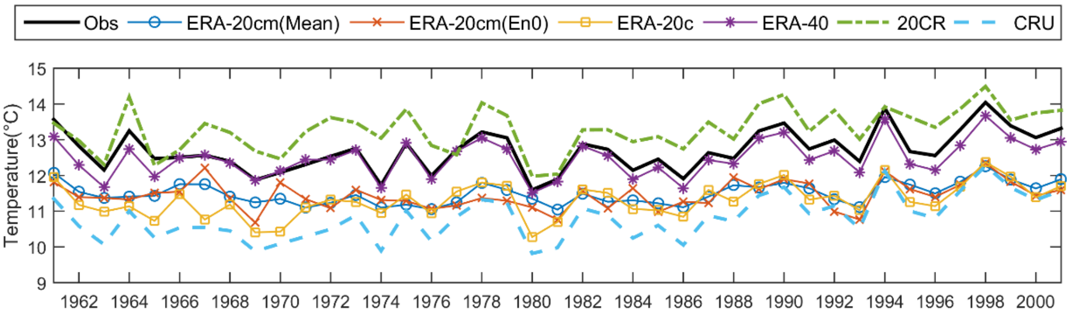
Celsius degrees lower than the observation in every comparison, although its variation trends have the similarity to the observations. On the other hand, 20CR has the higher values than the observation in annual comparison, affected by the autumn and winter temperature. Only ERA-40 is very well fitted to the observed values in every comparison. This result implies that, despite the significant correlations between the observation and each dataset, the bias correction should be considered before using them.

Figure 5 Mean temperature change for observation (Obs), Mean of ERA-20cm (ERA-20cm (Mean)), En0 of ERA-20cm (ERA-20cm(En0)), ERA-20c, ERA-40, 20CR, and CRUv3.23 (CRU) averaged over the whole region from 1961 to 2001

(a) The seasonal mean temperature change (from above, Spring, Summer, Autumn, and Winter)



(b) The annual mean temperature change



Long-term trend

Table 5 describes the seasonal and the annual patterns of the mean temperature by the Mann-Kendall approach. In the first analysis from 1961 to 2001, the result suggests that only ERA-40 has the increasing trends in spring, winter and annual simulations as well as the observations. Seasonally, the other datasets also have the upward trend in spring, but they show the different trends in other seasons. For CRUv3.23 and ERA-20c, there are significant increasing trends in spring, autumn and winter, whereas 20CR has the trends in spring and autumn. In case of ERA-20cm, the mean has the increasing trends in spring and summer at 95% confidence level, but En0 has it only in spring at 90% confidence level. In terms of annual analysis, all datasets except En0 show the upward trends. ERA-20c, 20CR and CRUv3.23 show the significant upward trends at 95% confidence level, and ERA-40 and ERA-20cm mean suggest them at 90% confidence level. For the slope of the annual comparison, those of CRUv3.23 and 20CR are higher than the observation's, while ERA-20c and ERA-40 are slightly smaller than the observed. In case of ERA-20cm mean, the slope shows less than the half of the observation's.

Table 5 Mann-Kendall test results for temperature trend

Dataset	Spring		Summer		Autumn		Winter		Annual	
	Z	*β*	*Z*	*β*	*Z*	*β*	*Z*	*β*	*Z*	*β*
1961-2001										
Observation	2.62 [a]	2.53	0.57	0.64	0.84	0.70	2.30 [a]	3.65	3.66 [a]	2.09
ERA-20cm (Mean)	2.62 [a]	1.17	2.12 [a]	1.31	0.78	0.59	0.75	0.33	1.81 [b]	0.86

ERA-20cm (En0)	1.85[b]	1.21	0.69	0.48	- 0.01	- 0.03	0.55	0.44	0.86	0.42
ERA-20c	1.67[b]	1.40	0.46	0.33	1.92[b]	1.30	2.62[a]	2.38	2.71[a]	1.80
ERA-40	2.55[a]	2.22	- 0.08	- 0.12	0.46	0.52	2.26[a]	2.41	1.81[b]	1.47
20CR	2.82[a]	2.20	1.54	1.86	2.41[a]	2.18	1.38	1.99	2.77[a]	2.34
CRUv3.23	3.36[a]	3.07	1.61	1.38	2.62[a]	2.16	2.86[a]	3.53	3.31[a]	2.91
1901-2010										
ERA-20cm (Mean)	8.69[a]	1.02	7.60[a]	0.91	6.61[a]	0.81	6.61[a]	0.79	9.38[a]	0.95
ERA-20cm (En0)	4.90[a]	1.02	4.02[a]	0.60	4.17[a]	0.80	3.19[a]	0.61	6.33[a]	0.80
ERA-20c	3.25[a]	0.66	2.94[a]	0.49	4.06[a]	0.86	6.15[a]	1.55	7.02[a]	1.04
20CR	6.73[a]	1.71	5.14[a]	1.35	6.11[a]	1.43	6.36[a]	2.09	8.15[a]	1.80
CRUv3.23	7.93[a]	1.99	3.77[a]	0.85	5.96[a]	1.51	5.15[a]	1.62	7.85[a]	1.61

[a] : significant trend at the 0.05 significance level. [b] : significant trend at the 0.10 significance level.

β(trends for temperature) are in $10^{-2}\,^{\circ}C/yr$

The second analysis for 20th century indicates the obvious increasing trends in all seasonal and annual simulations at 95% confidence level in Table 5. The only difference between datasets is the intensity of the slopes. As with the first analysis, the increasing magnitudes (β) of 20CR and CRUv3.23 are generally higher than those of the others. This result implies that the mean temperature in South Korea has been increased obviously over the past 100 years.

Statistical comparability

Figure 6 represents the skill score of the PDF of each dataset for monthly mean temperature from 1961 to 2001. The estimate of ERA-40 is approximately 0.90, and 20CR, En0, ERA-20c and CRUv3.23 follow with 0.74, 0.71, 0.69 and 0.69, separately. In other words, ERA-40 reanalysis has a probability density distribution approximately equal to the observed, and 20CR, En0, ERA-20c and CRU also have significant agreements with them. Reminding the high *r* values for ERA-20c and CRUv3.23 in annul comparison ($r > 0.87$), this result suggests that, despite the high correlation with the observation, ERA-20c and CRUv3.23 are clearly biased and they as well as other datasets need the bias correction for the application in the climate change study over South Korea. The mean of ERA-20cm also has the meaningful agreement, but not well as much as En0.

Figure 6 PDF-based skill score for monthly mean temperature for the Mean of ERA-20cm (ERA-20cm (Mean)), En0 of ERA-20cm (ERA-20cm(En0)), ERA-20c, ERA-40, 20CR, and CRUv3.23 (CRU) averaged over the whole region from 1961 to 2001

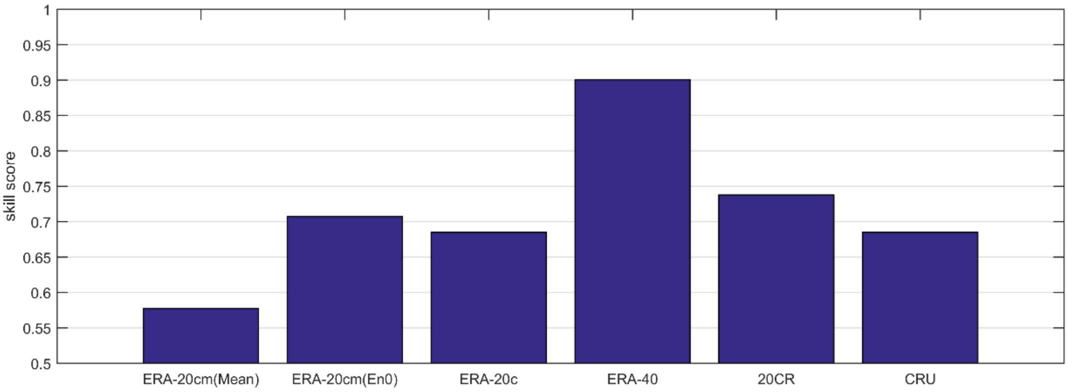
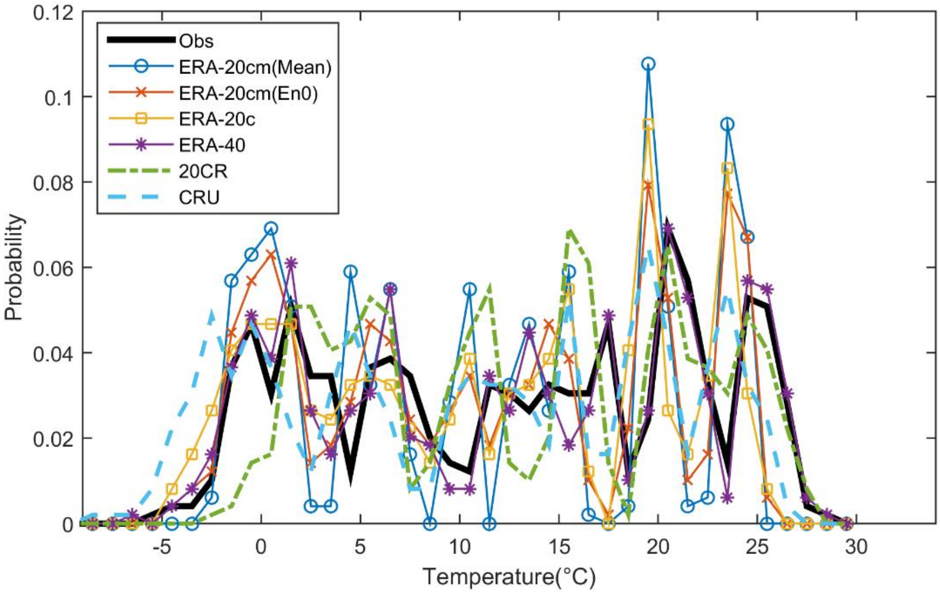


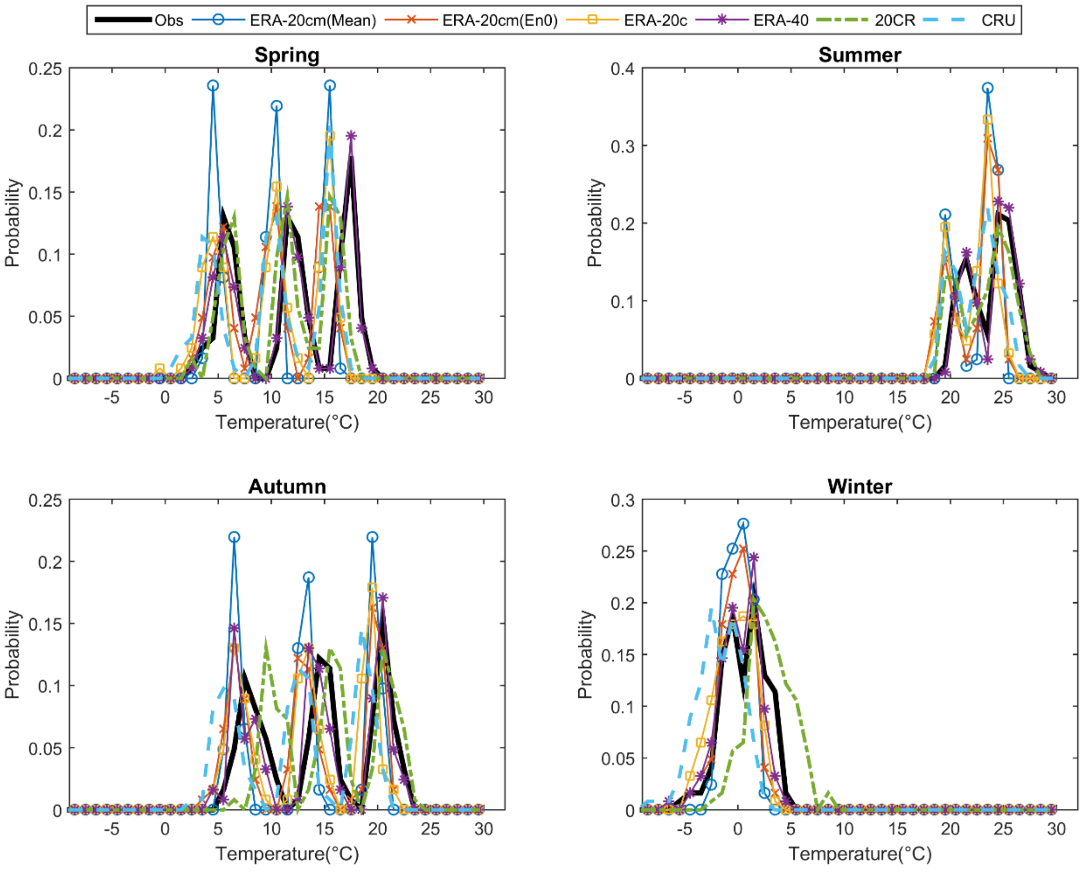
Figure 7 which illustrates the PDFs of the observed and the modelled dataset for temperature supports the skill score analysis. The performance of ERA-40 generally shows high agreements in all comparisons with the observations, but the other datasets have the seasonally biased distributions. The seasonally subdivided PDFs help to find out the difference of each dataset by comparing the peaks of them (Figure 7(b)). For the reference, the three peaks seen in spring and autumn are due to the rapid change in monthly mean temperature. For ERA-20cm mean, En0 and ERA-20c, the distributions are located in the left of the observations in the rest of the seasons except winter in Figure 7(b). Likewise, for CRUv3.23, the PDFs are located in the left side of the observation in every season (Figure 7(b)) and it causes the generally left-biased distribution in Figure 7(a). In case of 20CR, the PDF in Figure 7(a) seems to perform well except underestimation of the range of below 0°C and the partial discrepancies, but the seasonal PDFs imply that this result has been refined in the process of combining seasonal discrepancies (Figure 7(b)). For instance, the second and third peaks of 20CR in spring represent the lower temperature than the real, but the PDF for winter shows the warmer temperature than the observation. This suggests that statistical usage of 20CR without considering this seasonal deviation can distort the simulation.

Figure 7. Probability density functions for monthly mean temperature for observation (Obs), Mean of ERA-20cm (ERA-20cm (Mean)), En0 of ERA-20cm (ERA-20cm(En0)), ERA-20c, ERA-40, 20CR, and CRUv3.23 (CRU) over South Korea

(a) PDFs for monthly mean temperature from 1961 to 2001



(b) PDFs for seasonally subdivided monthly mean temperature from 1961 to 2001



Summary and Discussion

This study evaluates the multi-decadal reanalysis datasets, ERA-20cm, ERA-20c, ERA-40 and 20CR, and two century-long gridded observation datasets, CRUv3.23 and GPCCv7, over South Korea. The authors mainly focus on temporal and statistical applicability of monthly mean values for precipitation and temperature, which are the most commonly used data in climate change study (Gao et al. 2016).

In the temporal variability comparison for precipitation, the *r* values for ERA-20cm mean and En0 compared with the observation derived from the 13 gauged stations are closed to 0, while CRUv3.23 and GPCCv7 exceed 0.9 in every seasonal/annual comparison. This result reconfirms the well-known feature of ERA-20cm which cannot reproduce the actual synoptic situation for precipitation (Hersbach et al. 2015). On the other hand, the other reanalyses, ERA-20c, ERA-40 and 20CR, have moderate to high correlations ($0.4 < r < 0.9$) and, of them, ERA-40 and 20CR have the seasonal gaps comparing with the observations. This suggests that it is of importance to consider the local accuracy in national-scale studies using these datasets.

For the trend test on precipitation, there is no significant trend except ERA-40 for summer and 20CR in spring and annual trends for the reference period 1961 to 2001. However, the simulation from 1901 to 2010 shows the different trends depending on the dataset. CRUv3.23 and GPCCv7 have the identically increasing trends in summer and 12-month average simulations, whereas ERA-20c shows the upward tendencies in all tests and 20CR has the decreasing trends in summer, autumn and annual simulations. For ERA-20cm, the mean shows the increasing trends in spring, winter and annual tests, while En0 has no significant trends. It is clear that the result of the trend analysis can vary depending on the study period and regions in South Korea (Bae et al. 2008). Nevertheless, the previous long term trend researches have shown that summer precipitation observed in Korea has generally increased (Wang et al. 2006; Chang & Kwon 2007; Choi et al. 2009; Jung et al. 2011).

Chang & Kwon (2007) and Jung et al. (2011) suggested that all stations had increasing summer rainfalls since 1973. Choi et al. (2009) compared the gauged rainfalls of 10 Asian countries from 1955 to 2007 and described the significant increasing summer rainfall in South Korea at 95% confidence level. The longest trend analysis on Seoul, the capital of South Korea, also indicated a significant upward trend from 1778 to 2004, although the estimate for the pre-1950 period suggested no significant trend (Wang et al. 2006). Hence, the decreasing tendency of 20CR implies that despite the meaningful correlation with the observation, 20CR is able to provide distorted information in the long term trend analysis. In terms of the intensity of the annual trend, CRUv3.23 and GPCCv7 in the test from 1901 to 2010 have the significant increasing annual slopes (mm/yr), 2.13 and 2.14. These trends are different from Harris et al. (2014) which suggested 0.005 for CRU TS3.10 (CRUv3.10), the earlier version of CRUv3.23, and -0.019 for GPCC version 5, the earlier version of GPCCv7, in East Asia from 1901 to 2009. However, Choi et al. (2009), by evaluating the observations for the 1955-2007 period, showed that South Korea had a significant increasing trend (2.45) and it was much higher than China (0.33) and Japan (-1.75), the other East Asia countries. This supports that the slopes of CRUv3.23 and GPCCv7 have the reliability.

For statistical evaluation for precipitation, there are significant agreements between the monthly averaged observations derived from 13 gauged stations and each dataset. The skill scores for ERA-20c as well CRUv3.23 and GPCCv7, as the interpolated datasets, exceed 0.9, and the other reanalyses have over 0.8 except ERA-20cm mean which has 0.66. Gao et al. (2016) concluded that despite the spatial variability, all ten ensemble members of ERA-20cm for precipitation had the high skill scores, over 0.8, in China. However, the two obviously different values of ERA-20cm mean and En0 imply that the mean has difficulty in representing individual ensemble members. In other words, an ensemble member can describe the climate change than the mean in a certain area. Nevertheless, this evaluation suggests that all the reanalyses including ERA-20cm can be applied to the rainfall frequency analysis as the substitution of the observation after proper bias corrections.

For temperature, the *r* values of ERA-20cm mean have moderate correlations with the observations, whereas En0 performs with the low to moderate correlations. An interesting point is that the most fitted dataset is ERA-40 reanalysis, not CRUv3.23 which represents the interpolated observation dataset. In CRUv3.23 as well as ERA-20c and 20CR, there are the obvious gaps in the observations for temperature. The earlier comparative study for temperature from 1958 to 2001 described that CRUTEM2v, the earlier version of CRUv3.23, had the significantly lower temperature than ERA-40 in the northern hemispheres from 1958 to 1967 because of the limited availability of observations (Simmons et al. 2004). However, in this study, the annual discrepancy is shown in South Korea over the whole period 1961 to 2001 although it has been narrow.

In the temperature trend test, ERA-40 shows the identical tendencies to the observed which have upward trends in spring, winter and annual simulations for the 1961-2001 period and ERA-20c, 20CR and CRUv3.23 also have the similarity except autumn. On the other hand, ERA-20cm mean shows the significant increasing movements in spring, summer, and annual tests, whereas En0 has it only in spring. Although there are trend variations according to the study period and spatial distribution (Bae et al. 2008), the previous observed trends in South Korea for the late 20th century suggested that the winter and annual mean temperature had the significant upward trends but the summer trend was weak (Chung & Yoon 2000; Jung et al. 2002; Choi et al. 2009). Hence, it could be deduced that ERA-20cm mean which shows the strong summer and weak winter trends has little reliability in terms of long term trend. An interesting point is that the second trend assessment from 1901 to 2010 indicates the significant warming trends in all the simulations at the 0.95 confidence level, although the intensity of the slopes are different depending on the dataset. This trend has been shown in recent researches. Donat et al. (2016) suggested the warming trends over the world in their multi data sources analysis from 1901 to 2010, and Harris et al. (2014) showed the annual warming trend in East Asia, 0.11°C/decade, by using CRUv3.10 from 1901 to 2008. From this reason, the increasing trends over 100yr in this paper have the reliability, although the magnitudes of them have the uncertainty.

In the case of PDFs analysis, ERA-40 performs the best with the skill score of 0.90 and 20CR, ERA-20c and En0 as well as CRUv3.23 have the significant agreements to the observation with the values between 0.69 and 0.74. On the other hand, ERA-20cm mean has the lowest value, 0.58. This simulation indicates that these dataset have the significant reliability for the monthly frequency for temperature in Korea, but still it is challenging to apply them directly. In terms of ERA-20cm, Gao et al. (2016) showed that the skill scores of all ten ensemble members averaged over all regions in China for temperature exceed 0.9, but the skill scores for ERA-20cm mean and En0 in this study have the much lower values. This suggests that there may be a clear difference in applicability according to the region, and this gap should be explored before using the dataset in the regional scale study.

Considering the improved assimilation and ensemble technique, it is easy to hypothesise that the higher the temporal and spatial resolutions, the more accurate the reanalysis dataset should be in terms of temporal and statistical variability. However, the results in this study indicates that each dataset has its own bias and the degree of the agreement of each data can vary in space and time as shown in previous studies (Simmons et al. 2004; Bosilovich et al. 2008; Ma et al. 2009; Bao & Zhang 2013). There may be some reasons for the data uncertainty. First of all, the inhomogeneity of input data for the simulated datasets can be one of the causes (Thorne & Vose 2010; Donat et al. 2016). In other words, the further from the present, the fewer number of stations are available and it is logical to reason the increase of uncertainty for the reanalyses as well as the interpolated observation data (Ferguson & Villarini 2012; Becker et al. 2013; Zhang et al. 2013; Harris et al. 2014). It is also known that altitude gap between the modelled data and actual terrain can be one of the reasons for the significant biases in a mountainous region like South Korea (Zhao & Fu 2006; Gao et al. 2012; Gao et al. 2014a; Gao et al. 2014b). Gao et al. (2014a) showed that the biases for ERA-interim temperature data were related with the elevation difference between ERA-interim grid points and gauging stations in complex terrains, but able to be reduced. Regional climate events like monsoons may explain the uncertainty of the modelled data (Shah & Mishra 2014; Gao et al. 2016). Shah & Mishra (2014) described that the reanalysis products like ERA-interim showed the clear bias in the monsoon season precipitation and temperature over India. The resolution of gridded points may also affect the uncertainty (Heikkilä et al. 2011). Heikkilä et al. (2011) compared the downscaled ERA-40 with different resolutions from 30 to 10km with observations and concluded that 10km resolution performed the best in complex terrains.

Likewise, there are numerous reasons for the uncertainty of the datasets and it is still challenging to reliably reproduce climate features of South Korea by directly using a single modelled data. Hence, it is necessary to evaluate the agreement between the datasets and observations and improve the quality of the products in order to apply them in the regional scale analysis. Nevertheless, due to the little attention on the global dataset in South Korea, this study can suggest the potentiality of the reanalysis data and interpolated data as an alternative data source supplementing the lack of long-term observations.

Conclusions

This study has firstly evaluated key century-long climate datasets for precipitation and temperature in South Korea. From the temporal and statistical comparisons, it could be concluded that GPCCv7 and CRUv3.23 for precipitation and ERA-40 for temperature perform the best among the compared datasets for the reference period 1961 to 2001. ERA-40, ERA-20c and 20CR for precipitation and CRUv3.23, ERA-20c and 20CR for temperature have the significant agreements with the observation, but they need to be improved for the application in Korea. ERA-20cm can be used for the frequency analysis over South Korea on a monthly basis after bias correction, but are not suitable for the temporal variability including the long-term trend. Moreover, ERA-20cm mean has difficulty in representing all ten ensemble members. This paper also shows that not only reanalyses but also the interpolated datasets such as CRUv3.23, which are generally accepted as the true values in the global climate change study, are able to be biased depending on the region. It means that no long-term dataset can be directly applied in climate impact analysis. These findings in this paper help to fill in the knowledge gaps about the applicability of these datasets in South Korea, and provide a useful guideline to readers from other countries on the comparative performance of the global datasets in different parts of the world.

This study has mainly explored the monthly/seasonal/annual mean change on the basis of the averaged dataset over the whole regions. This analysis is very useful for understanding the general pattern of each dataset in Korea, but it does not represent the extreme climate, which is one of the vital parameters in climate impact assessment. Spatial variations with the finer resolution such as daily or 10km scale should be highlighted in the future study and it is essential to correct biases of the model datasets. An advantage of the reanalysis data like ERA-20c and ERA-20cm by ECWMF is that they supply the daily datasets with 0.125° resolution without downscaling, while the others provide the coarser data. For ERA-20cm, it may be of importance to specifically assess the features by all ensemble members, which has simply been explored by just both of mean and En0 in this context. Hence, the bias correction for the reanalysis data with the higher spatio-temporal resolutions will be explored further in the future study as well as the features of the ERA-20cm ensemble.

Acknowledgements

The ERA-20c, ERA-20cm and ERA-40 data were collected via the ECMWF's public server (http://apps.ecmwf.int/datasets/). Support for the 20CR dataset was provided by the U.S. Department of Energy, Office of Science Biological and Environmental Research (BER), and by the NOAA Climate Program Office (http://www.esrl.noaa.gov/psd/). The CRU and GPCC datasets were supplied from their websites, https://crudata.uea.ac.uk/cru/data/hrg/ and http://gpcc.dwd.de, separately. The first author is grateful for the financial support from the Government of South Korea for carrying out his PhD studies at the University of Bristol.

References

Babak, O. & Deutsch, C. V. 2009 Statistical approach to inverse distance interpolation. *Stochastic Environmental Research and Risk Assessment,* **23**(5)**,** 543-553. DOI:10.1007/s00477-008-0226-6

Bae, D. H., Jung, I. W. & Chang, H. 2008 Long-term trend of precipitation and runoff in Korean river basins. *Hydrological processes,* **22**(14)**,** 2644-2656.

Bao, X. & Zhang, F. 2013 Evaluation of NCEP–CFSR, NCEP–NCAR, ERA-Interim, and ERA-40 reanalysis datasets against independent sounding observations over the Tibetan Plateau. *Journal of Climate,* **26**(1)**,** 206-214.

Becker, A., Finger, P., Meyer-Christoffer, A., Rudolf, B., Schamm, K., Schneider, U. & Ziese, M. 2013 A description of the global land-surface precipitation data products of the Global Precipitation Climatology Centre with sample applications including centennial (trend) analysis from 1901-present. *Earth System Science Data,* **5**(1)**,** 71-99. DOI:10.5194/essd-5-71-2013

Bosilovich, M. G., Chen, J., Robertson, F. R. & Adler, R. F. 2008 Evaluation of global precipitation in reanalyses. *Journal of applied meteorology and climatology,* **47**(9)**,** 2279-2299.

Chang, C.-L., Lo, S.-L. & Yu, S.-L. 2006 The parameter optimization in the inverse distance method by genetic algorithm for estimating precipitation. *Environmental monitoring and assessment,* **117**(1)**,** 145-155.

Chang, H. & Kwon, W.-T. 2007 Spatial variations of summer precipitation trends in South Korea, 1973–2005. *Environmental Research Letters,* **2**(4)**,** 045012.

Chen, M., Xie, P., Janowiak, J. E. & Arkin, P. A. 2002 Global land precipitation: A 50-yr monthly analysis based on gauge observations. *Journal of Hydrometeorology,* **3**(3)**,** 249-266.

Choi, G., Collins, D., Ren, G., Trewin, B., Baldi, M., Fukuda, Y., Afzaal, M., Pianmana, T., Gomboluudev, P. & Huong, P. T. T. 2009 Changes in means and extreme events of temperature and precipitation in the Asia-Pacific Network region, 1955–2007. *International Journal of Climatology,* **29**(13)**,** 1906-1925.

Choi, K.-S., Kim, B.-J., Zhang, R., Nam, J.-C., Park, K.-J., Kim, J.-Y. & Kim, D.-W. 2016 Possible influence of South Asian high on summer rainfall variability in Korea. *Climate Dynamics,* **46**(3)**,** 833-846.

Chung, U., Choi, J. & Yun, J. I. 2004 Urbanization effect on the observed change in mean monthly temperatures between 1951-1980 and 1971-2000 in Korea. *Climatic Change,* **66**(1)**,** 127-136. DOI:10.1023/B:CLIM.0000043136.58100.ce

Chung, Y. & Yoon, M. 2000 Interpretation of recent temperature and precipitationtrends observed in

Korea. *Theoretical and Applied Climatology,* **67**(3-4)**,** 171-180.

Compo, G. P., Whitaker, J. S., Sardeshmukh, P. D., Matsui, N., Allan, R. J., Yin, X., Gleason, B. E., Vose, R. S., Rutledge, G., Bessemoulin, P., Brönnimann, S., Brunet, M., Crouthamel, R. I., Grant, A. N., Groisman, P. Y., Jones, P. D., Kruk, M. C., Kruger, A. C., Marshall, G. J., Maugeri, M., Mok, H. Y., Nordli, Ø., Ross, T. F., Trigo, R. M., Wang, X. L., Woodruff, S. D. & Worley, S. J. 2011 The Twentieth Century Reanalysis Project. *Quarterly Journal of the Royal Meteorological Society,* **137**(654)**,** 1-28. DOI:10.1002/qj.776

Dee, D., Uppala, S., Simmons, A., Berrisford, P., Poli, P., Kobayashi, S., Andrae, U., Balmaseda, M., Balsamo, G. & Bauer, P. 2011a The ERA-Interim reanalysis: Configuration and performance of the data assimilation system. *Quarterly Journal of the royal meteorological society,* **137**(656)**,** 553-597.

Dee, D. P. & Uppala, S. 2009 Variational bias correction of satellite radiance data in the ERA-Interim reanalysis. *Quarterly Journal of the Royal Meteorological Society,* **135**(644)**,** 1830-1841.

Dee, D. P., Uppala, S. M., Simmons, A. J., Berrisford, P., Poli, P., Kobayashi, S., Andrae, U., Balmaseda, M. A., Balsamo, G., Bauer, P., Bechtold, P., Beljaars, A. C. M., van de Berg, L., Bidlot, J., Bormann, N., Delsol, C., Dragani, R., Fuentes, M., Geer, A. J., Haimberger, L., Healy, S. B., Hersbach, H., Hólm, E. V., Isaksen, L., Kållberg, P., Köhler, M., Matricardi, M., McNally, A. P., Monge-Sanz, B. M., Morcrette, J. J., Park, B. K., Peubey, C., de Rosnay, P., Tavolato, C., Thépaut, J. N. & Vitart, F. 2011b The ERA-Interim reanalysis: configuration and performance of the data assimilation system. *Quarterly Journal of the Royal Meteorological Society,* **137**(656)**,** 553-597. DOI:10.1002/qj.828

Deser, C., Phillips, A. S. & Hurrell, J. W. 2004 Pacific interdecadal climate variability: Linkages between the tropics and the North Pacific during boreal winter since 1900. *Journal of Climate,* **17**(16)**,** 3109-3124.

Dickinson, R. E., Oleson, K. W., Bonan, G., Hoffman, F., Thornton, P., Vertenstein, M., Yang, Z.-L. & Zeng, X. 2006 The Community Land Model and its climate statistics as a component of the Community Climate System Model. *Journal of Climate,* **19**(11)**,** 2302-2324.

Dinku, T., Connor, S. J., Ceccato, P. & Ropelewski, C. F. 2008 Comparison of global gridded precipitation products over a mountainous region of Africa. *International Journal of Climatology,* **28**(12)**,** 1627-1638.

Donat, M. G., Alexander, L. V., Herold, N. & Dittus, A. J. 2016 Temperature and precipitation extremes in century-long gridded observations, reanalyses, and atmospheric model simulations. *Journal of Geophysical Research: Atmospheres,* **121**(19).

Donat, M. G., Sillmann, J., Wild, S., Alexander, L. V., Lippmann, T. & Zwiers, F. W. 2014 Consistency of temperature and precipitation extremes across various global gridded in situ and reanalysis datasets. *Journal of Climate,* **27**(13)**,** 5019-5035.

Fekete, B. M., Vörösmarty, C. J., Roads, J. O. & Willmott, C. J. 2004 Uncertainties in precipitation and their impacts on runoff estimates. *Journal of Climate,* **17**(2)**,** 294-304.

Ferguson, C. R. & Villarini, G. 2012 Detecting inhomogeneities in the Twentieth Century Reanalysis over the central United States. *Journal of Geophysical Research: Atmospheres,* **117**(D5).

Gao, L., Bernhardt, M. & Schulz, K. 2012 Elevation correction of ERA-Interim temperature data in complex terrain. *Hydrology and Earth System Sciences,* **16**(12)**,** 4661.

Gao, L., Bernhardt, M., Schulz, K., Chen, X. W., Chen, Y. & Liu, M. B. 2016 A First Evaluation of ERA-20CM over China. *Monthly Weather Review,* **144**(1)**,** 45-57. DOI:10.1175/Mwr-D-15-0195.1

Gao, L., Hao, L. & Chen, X.-w. 2014a Evaluation of ERA-interim monthly temperature data over the Tibetan Plateau. *Journal of Mountain Science,* **11**(5)**,** 1154-1168.

Gao, L., Schulz, K. & Bernhardt, M. 2014b Statistical downscaling of ERA-interim forecast precipitation data in complex terrain using lasso algorithm. *Advances in Meteorology,* **2014**.

Gholami, V., Chau, K., Fadaee, F., Torkaman, J. & Ghaffari, A. 2015 Modeling of groundwater level fluctuations using dendrochronology in alluvial aquifers. *Journal of hydrology,* **529,** 1060-1069.

Hamed, K. H. & Rao, A. R. 1998 A modified Mann-Kendall trend test for autocorrelated data. *Journal of Hydrology,* **204**(1)**,** 182-196.

Harris, I., Jones, P. D., Osborn, T. J. & Lister, D. H. 2014 Updated high-resolution grids of monthly climatic observations - the CRU TS3.10 Dataset. *International Journal of Climatology,* **34**(3)**,** 623-642. DOI:10.1002/joc.3711

Heikkilä, U., Sandvik, A. & Sorteberg, A. 2011 Dynamical downscaling of ERA-40 in complex terrain using the WRF regional climate model. *Climate dynamics,* **37**(7-8)**,** 1551-1564.

Herrmann, S. M., Anyamba, A. & Tucker, C. J. 2005 Recent trends in vegetation dynamics in the African Sahel and their relationship to climate. *Global Environmental Change,* **15**(4)**,** 394-404.

Hersbach, H., Peubey, C., Simmons, A., Berrisford, P., Poli, P. & Dee, D. 2015 ERA-20CM: a twentieth-century atmospheric model ensemble. *Quarterly Journal of the Royal Meteorological Society,* **141**(691)**,** 2350-2375.

Hirsch, R. M., Slack, J. R. & Smith, R. A. 1982 Techniques of trend analysis for monthly water quality data. *Water resources research,* **18**(1)**,** 107-121.

Ho, C. H., Lee, J. Y., Ahn, M. H. & Lee, H. S. 2003 A sudden change in summer rainfall characteristics in Korea during the late 1970s. *International Journal of Climatology,* **23**(1)**,** 117-

128.

Jeong, M. S., Kim, J. S., Oh, S. M. & Moon, Y. I. 2015 Sensitivity of summer precipitation over the Korean Peninsula to temperature gradients. *International Journal of Climatology,* **35**(6)**,** 836-845.

Jung, H. S., Choi, Y., Oh, J. H. & Lim, G. H. 2002 Recent trends in temperature and precipitation over South Korea. *International Journal of Climatology,* **22**(11)**,** 1327-1337.

Jung, I. W., Bae, D. H. & Kim, G. 2011 Recent trends of mean and extreme precipitation in Korea. *International Journal of Climatology,* **31**(3)**,** 359-370. DOI:10.1002/joc.2068

Kalnay, E., Kanamitsu, M., Kistler, R., Collins, W., Deaven, D., Gandin, L., Iredell, M., Saha, S., White, G., Woollen, J., Zhu, Y., Leetmaa, A., Reynolds, R., Chelliah, M., Ebisuzaki, W., Higgins, W., Janowiak, J., Mo, K. C., Ropelewski, C., Wang, J., Jenne, R. & Joseph, D. 1996 The NCEP/NCAR 40-Year Reanalysis Project. *Bulletin of the American Meteorological Society,* **77**(3)**,** 437-471. DOI:10.1175/1520-0477(1996)077<0437:tnyrp>2.0.co;2

Kendall, M. G. 1955 *Rank correlation methods.* Griffin, London.

Legates, D. R. & McCabe, G. J. 1999 Evaluating the use of “goodness-of-fit” Measures in hydrologic and hydroclimatic model validation. *Water Resources Research,* **35**(1)**,** 233-241. DOI:10.1029/1998WR900018

Lu, G. Y. & Wong, D. W. 2008 An adaptive inverse-distance weighting spatial interpolation technique. *Computers & Geosciences,* **34**(9)**,** 1044-1055.

Ma, L., Zhang, T., Frauenfeld, O. W., Ye, B., Yang, D. & Qin, D. 2009 Evaluation of precipitation from the ERA-40, NCEP-1, and NCEP-2 Reanalyses and CMAP-1, CMAP-2, and GPCP-2 with ground-based measurements in China. *Journal of Geophysical Research: Atmospheres,* **114**(D9).

Mann, H. B. 1945 Nonparametric tests against trend. *Econometrica: Journal of the Econometric Society***,** 245-259.

Nicholson, S. E., Some, B., McCollum, J., Nelkin, E., Klotter, D., Berte, Y., Diallo, B., Gaye, I., Kpabeba, G. & Ndiaye, O. 2003 Validation of TRMM and other rainfall estimates with a high-density gauge dataset for West Africa. Part I: Validation of GPCC rainfall product and pre-TRMM satellite and blended products. *Journal of Applied Meteorology,* **42**(10)**,** 1337-1354.

Nikulin, G., Jones, C., Giorgi, F., Asrar, G., Büchner, M., Cerezo-Mota, R., Christensen, O. B., Déqué, M., Fernandez, J. & Hänsler, A. 2012 Precipitation climatology in an ensemble of CORDEX-Africa regional climate simulations. *Journal of Climate,* **25**(18)**,** 6057-6078.

Onogi, K., Tsutsui, J., Koide, H., Sakamoto, M., Kobayashi, S., Hatsushika, H., Matsumoto, T., Yamazaki, N., Kamahori, H. & Takahashi, K. 2007 The JRA-25 reanalysis. *Journal of the Meteorological Society of Japan. Ser. II,* **85**(3)**,** 369-432.

Perkins, S. E., Pitman, A. J., Holbrook, N. J. & McAneney, J. 2007 Evaluation of the AR4 climate models' simulated daily maximum temperature, minimum temperature, and precipitation over Australia using probability density functions. *Journal of Climate,* **20**(17)**,** 4356-4376. DOI:10.1175/Jcli4253.1

Poli, P., Hersbach, H., Dee, D. P., Berrisford, P., Simmons, A. J., Vitart, F., Laloyaux, P., Tan, D. G., Peubey, C. & Thépaut, J.-N. 2016 ERA-20C: An Atmospheric Reanalysis of the Twentieth Century. *Journal of Climate,* **29**(11)**,** 4083-4097.

Raziei, T., Bordi, I. & Pereira, L. S. 2011 An application of GPCC and NCEP/NCAR datasets for drought variability analysis in Iran. *Water Resources Management,* **25**(4)**,** 1075-1086.

Sayemuzzaman, M. & Jha, M. K. 2014 Seasonal and annual precipitation time series trend analysis in North Carolina, United States. *Atmospheric Research,* **137,** 183-194.

Schneider, U., Becker, A., Finger, P., Meyer-Christoffer, A., Rudolf, B. & Ziese, M. 2015. GPCC Full Data Reanalysis Version 7.0 at 0.5°: Monthly Land-Surface Precipitation from Rain-Gauges built on GTS-based and Historic Data. ftp://ftp.dwd.de/pub/data/gpcc/html/fulldata_v7_doi_download.html (Accessed 10 November 2016)

Sen, P. K. 1968 Estimates of the regression coefficient based on Kendall's tau. *Journal of the American Statistical Association,* **63**(324)**,** 1379-1389.

Shadmani, M., Marofi, S. & Roknian, M. 2012 Trend analysis in reference evapotranspiration using Mann-Kendall and Spearman’s Rho tests in arid regions of Iran. *Water resources management,* **26**(1)**,** 211-224.

Shah, R. & Mishra, V. 2014 Evaluation of the reanalysis products for the monsoon season droughts in India. *Journal of Hydrometeorology,* **15**(4)**,** 1575-1591.

Simmons, A. J., Jones, P. D., Bechtold, V. D., Beljaars, A. C. M., Kallberg, P. W., Saarinen, S., Uppala, S. M., Viterbo, P. & Wedi, N. 2004 Comparison of trends and low-frequency variability in CRU, ERA-40, and NCEP/NCAR analyses of surface air temperature. *Journal of Geophysical Research: Atmospheres,* **109**(D24). DOI:10.1029/2004jd005306

Teegavarapu, R. S. & Chandramouli, V. 2005 Improved weighting methods, deterministic and stochastic data-driven models for estimation of missing precipitation records. *Journal of Hydrology,* **312**(1)**,** 191-206.

Theil, H. 1950 A rank-invariant method of linear and polynomial regression analysis. I, II, III. *Proceedings of the Koninklijke Nederlandse Akademie van Wetenschappen,* **53,** 386–392, 521–525, 1397–1412.

Thorne, P. & Vose, R. 2010 Reanalyses suitable for characterizing long-term trends: are they really achievable? *Bulletin of the American Meteorological Society,* **91**(3)**,** 353-361.

Tomczak, M. 1998 Spatial interpolation and its uncertainty using automated anisotropic inverse distance weighting (IDW)-cross-validation/jackknife approach. *Journal of Geographic Information and Decision Analysis,* **2**(2)**,** 18-30.

Uppala, S. M., Kållberg, P., Simmons, A., Andrae, U., Bechtold, V. d., Fiorino, M., Gibson, J., Haseler, J., Hernandez, A. & Kelly, G. 2005 The ERA-40 re-analysis. *Quarterly Journal of the Royal Meteorological Society,* **131**(612)**,** 2961-3012.

Wang, B., Ding, Q. & Jhun, J. G. 2006 Trends in Seoul (1778–2004) summer precipitation. *Geophysical research letters,* **33**(15).

Wang, W.-c., Chau, K.-w., Xu, D.-m. & Chen, X.-Y. 2015 Improving Forecasting Accuracy of Annual Runoff Time Series Using ARIMA Based on EEMD Decomposition. *Water Resources Management,* **29**(8)**,** 2655-2675. DOI:10.1007/s11269-015-0962-6

Wu, C., Chau, K. & Fan, C. 2010 Prediction of rainfall time series using modular artificial neural networks coupled with data-preprocessing techniques. *Journal of Hydrology,* **389**(1)**,** 146-167.

Xu, Z., Gong, T. & Li, J. 2008 Decadal trend of climate in the Tibetan Plateau—regional temperature and precipitation. *Hydrological Processes,* **22**(16)**,** 3056-3065.

Xu, Z., Takeuchi, K., Ishidaira, H. & Li, J. 2005 Long-term trend analysis for precipitation in Asian Pacific FRIEND river basins. *Hydrological Processes,* **19**(18)**,** 3517-3532.

Zang, C. & Liu, J. 2013 Trend analysis for the flows of green and blue water in the Heihe River basin, northwestern China. *Journal of Hydrology,* **502,** 27-36.

Zhang, L. & Zhou, T. 2011 An assessment of monsoon precipitation changes during 1901–2001. *Climate Dynamics,* **37**(1)**,** 279-296.

Zhang, Z., Guo, W., Gong, D. & Kim, S.-J. 2013 Evaluation of the twentieth century reanalysis dataset in describing East Asian winter monsoon variability. *Advances in Atmospheric Sciences,* **30**(6)**,** 1645.

Zhao, T. & Fu, C. 2006 Comparison of products from ERA-40, NCEP-2, and CRU with station data for summer precipitation over China. *Advances in Atmospheric sciences,* **23**(4)**,** 593-604.