Reanalysis Precipitation Products over the Continental US: Are they Skillful for Century-scale Analyses?

Xiaodong Chen and Faisal Hossain

Department of Civil and Environmental Engineering, University of Washington, Seattle WA, USA

Submitted to:

Geophysical Research Letters

Corresponding Author:

Faisal Hossain
Department of Civil and Environmental Engineering
University of Washington, Seattle, WA 98195
Email: fhossain@uw.edu

Abstract:

Within the past decade, two long-term reanalysis products (i.e. 1915-2010) have been developed to enable 'century-scale' analyses. These twentieth century reanalysis (20CR) products are produced by: NOAA (20CRv2), and ECWMF (ERA-20C). Using daily precipitation observations of CONUS, we study the temporal variations ranging from annual to daily scale. Our evaluation shows that 20CRv2 overestimates the precipitation magnitude, while ERA-20C shows consistent underestimation. The reconstructions of precipitation are comparable between two datasets at coarser scales (annual and monthly scales), although 20CRv2 provides better estimates at finer temporal scales (weekly and daily scales). Both 20CRs correctly capture the precipitation variations in the west coast of US, and these variations are well captured at daily scale in 20CRv2 data. These results suggest the potential of using 20CR products to reconstruct storms events in the west coast area during 1915-1948 when there are currently no other reanalysis products available.

Keywords: reanalysis, precipitation, climate, numerical modeling.

1. Introduction

Twentieth century reanalysis (20CR) product is the model reconstruction of climate in the past century. They are built for climatological study, and they have demonstrated skill in capturing key climate variations such as North Atlantic Oscillation and El Nino [*Giese et al.*, 2010; *Compo et al.*, 2011]. They have also been used to evaluate other numerical simulations such as GCMs [*Sheffield et al.*, 2013]. Previous studies have shown their ability to capture the long-term statistics and variations in precipitation at annual to monthly scales in different regions [*Ferguson and Villarini*, 2012; *Misra et al.*, 2013; *Zhang et al.*, 2013; *Kong and Bi*, 2015]. Efforts have been made to evaluate the precipitation products in 20CR, but all of them have restricted the analysis to monthly scale due to the limitation of observed data. For example, *Ferguson and Villarini* (2012) evaluated the temperature and precipitation simulation in the central US using Climate Research Unit (CRU) monthly observation and concluded that the 20CR produced by NOAA overestimates the precipitation amount when compared with CRU observation. However, the quality of 20CR in event-scale simulations and over longer time period (such as a century) has not yet been investigated.

Modern atmospheric numerical models require initial and boundary conditions as model input, which are often obtained from atmospheric reanalysis datasets. Several reanalysis datasets have been developed for the post-1948 atmospheric conditions [*Kalnay et al.*, 1996; *Kanamitsu et al.*, 2002; *Mesinger et al.*, 2006]. However, 20CR remains the only source of boundary conditions for numerical modeling of events before 1948. Recent studies suggest that although in general 20CR is not a good option for precipitation event reconstruction, it may be usable in certain climate regions such as the west coast of US [*Chen and Hossain*, 2016]. From a modeling perspective, historical event reconstruction can also be viewed as dynamic downscaling of the reanalysis data used for boundary conditions. The difficulty of precipitation simulation has been recognized as a result of highly variable rain rates and large spatial variability [*Intergovernmental Panel on Climate Change et al.*, 2013]. Thus, it is possible that in areas/durations where reanalysis precipitation shows a good match with observation at coarser resolution, numerical model downscaling may produce skillful results. Therefore, it is important and useful to identify the areas/periods where the 20CR can serve for such event-scale reconstructions. Using quality-controlled ground-based precipitation observations as reference, this study aims to examine the spatial-temporal precipitation reconstruction quality over the continental US (CONUS) in the currently available 20CRs. The goal is to understand how skillful the currently available reanalysis products are for analyses of precipitation events over the last century (1915-2010).

2. Data and Analysis

Currently there are two 20CRs available: one is the Twentieth Century Reanalysis Product (20CRv2) produced through a collaboration between NOAA and University of Colorado Cooperative Institute for Research and Environmental Sciences; the other (ERA-20C) is produced by European Centre for Medium-Range Weather Forecasting (ECMWF), and it is an outcome of the ERA-CLIM project. Table 1 shows the details of these two 20CRs. The 20CRv2 is an ensemble product (56 ensemble members), and in this study only the mean values of ensembles are used.

The evaluation of the products used the Livneh daily CONUS near-surface gridded meteorological dataset [*Livneh et al.*, 2013] as the reference observation. This precipitation dataset was generated from rain gauge records since 1915. Given its long-term temporal coverage, it has been used in other model evaluation studies [*Sheffield et al.*, 2013]. In this study, the data over CONUS (see Figure 1) was used, and all the data were first conservatively regridded to the global Gaussian T62 grids, the coarsest resolution among these three datasets, before the analysis was conducted. Variations in the reconstructed precipitation were evaluated using the spatial and temporal correlation between 20CR precipitation and Livneh precipitation datasets.

3. Results and Discussion

Figure 1 shows the climatology of precipitation during 1915-2010. The pattern of higher annual precipitation in both west and east coast area in the reference data is correctly captured in both reconstructions. Regarding the multiyear average annual precipitation amount, 20CRv2 shows a better match in the southeast area, though both failed to give out the observed large amount of precipitation in the Pacific Northwest area. Figure 1 also suggests that 20CRv2 tends to overestimate the precipitation in the center and the east US. ERA-20C provides a better matched spatial pattern, but it slightly underestimates the precipitation in the southeast US.

To understand the reconstructed spatial distribution of annual precipitation, annual spatial correlation coefficients were calculated. Figure 2 shows these correlations along with the factors that would influence these correlations. It is clear that the correlation coefficients (thus reconstruction of annual precipitation maps) are superior in skill in recent years, as one would expect. The two products not only agree with each other but also yield high correlation coefficients for the recent period of 1980-2010. As we go back in the past, such as pre-1950, the 20CRv2 outperforms ERA-20C. Also, for the wet years (larger circles in Figure 2), 20CRs tend to produce more accurate annual precipitation maps.

Figure 3 shows the temporal correlation of seasonal and monthly precipitation series. The large patterns are similar at seasonal and monthly scales, with the highest skill observed in the west coast, followed by the mid-west. For these two regions, two 20CRs show similar performance, with ERA-20C slightly better in central north. In the southeast area (i.e. Florida), however, 20CRv2 clearly outperforms the ERA-20C. The climate of the west coast can partly explain the high correlation in this region. Previous studies have pointed out that in this region (especially California), over 60% of annual precipitation is contributed by less than 20 events [*Sun et al.*, 2006], which are often the atmospheric rivers events [*Hagos et al.*, 2016]. Atmospheric river events are one of the large-scale meteorological events that numerical models can capture better than the more localized convective storms.

It is also important to note that the reanalysis products are capable of describing the monthly cycle of precipitations. Here we focus on the average monthly cycle, and all the monthly data were taken to calculate multiyear averaged monthly precipitation and standard deviation. The results are shown in Figure 4, and it is clear that for all 12 months, 20CRv2 usually overestimate, while ERA-20C tends to underestimate the precipitation amount. In the months except June-September, the precipitation in ERA-20C matches closer with the observation. In summer time, however, 20CRv2 gives out better estimates. This confirms the superiority of 20CR in capturing heavy rainfall events. Previous studies have suggested that taking 20CR and other reanalyses as “ensemble” could improve the reconstruction quality, such as the surface mass balance of Greenland ice sheet [*Hanna et al.*, 2011]. The comparison in Figure 4 also suggests that combining the information from 20CRv2 and ERA-20C may produce a better estimate on precipitation patterns.

To investigate the skill of the 20CRs to produce correct spatial-temporal structures in precipitation, the monthly precipitation data are decomposed using empirical orthogonal functions (EOF). The analysis was done on the detrended monthly data, in which the linear trends at each grid were taken out. The results of EOF analysis are shown in Figure 5. Panels (a-c) are the first mode in monthly precipitation, panel (d) shows the contribution of each mode in the monthly variation. Both 20CRv2 and ERA-20C produce similar first modes as seen in the observation: the transition of opposite phases from west to east coast. However, the first mode in 20CRv2 is closer to the observation, where the phase in the southeast inland area is less significant. ERA-20C overestimates the area of negative phase in the west US. Regarding the magnitude of variation, panel (d) shows that the contribution of the first mode in 20CRv2 is almost the same as that from observation, but ERA-20C gives out better estimates for the next a few modes (mode 2 to 5).

Figure 6 shows the spatial distribution of correlations on weekly and daily precipitation series. At weekly scale, 20CRv2 and ERA-20C show similar performance, with 20CRv2 being slightly more skillful in the eastern US. At the daily scale, however, 20CRv2 shows much better performance. In both datasets, the daily variations in the west coast are significantly better constructed, with 20CRv2 showing better overall correlation coefficients. The modeling efforts on several extreme rainstorms before 1948 also suggests that 20CRv2 can be used to reconstruct the west coast storms reasonably well, but the reconstruction at other areas are heavily biased [*Chen and Hossain*, 2016]. It is also important to note that the correlation of daily precipitation has different trends in two 20CRs, as shown in Figure 7. In 20CRv2, the correlation remains between 0.55-0.60 for the whole duration of 1915-2010, and the correlation is slightly higher in recent years. ERA-20C data shows a peak of correlation around 1960-1975, and the reconstruction in this period is comparable to 20CRv2. However, the correlation after 1975 has a systematic decrease, indicating three distinct epochs regarding the skill of daily precipitation: 1) 1915-1960, when the reconstruction shows the lowest quality; 2) 1960-1975, when the reconstruction is best; 3) 1975-2010, when the reconstruction quality is of medium quality.

Although 20CRs were originally built for climate reconstruction, our analysis suggests that it may also be suitable for event-scale precipitation simulations over the West Coast of US. This finding can potentially encourage science/engineering communities to extend analyses with confidence further back than the 1950s to re-evaluate storms of the early 20th century. For example, the Hydrometeorological Reports (HMR) published by NOAA [*Schreiner and Riedel*, 1978] has outlined a collection of extreme historical rainstorms for engineering safety design purposes. Previous studies have shown the capability of reanalysis products in rebuilding those rainstorms after 1948 using various numerical models [*Tan*, 2010; *Woldemichael et al.*, 2014; *Chen and Hossain*, 2016]. However, a significant portion of this HMR collection occurred before 1948, for which direct meteorological observations other than precipitation are limited. Therefore, numerical modeling is the only viable approach to obtain a more complete physical picture about these events. The good quality in 20CRs at least over the west coast suggests that we can now apply numerical modeling for extreme precipitation events that took place in the first half of the 20th century.

4. Conclusions

In this study we employed a gauge-based precipitation dataset to evaluate the precipitation simulated by two 20CR during 1915-2010 at various temporal scales. The major findings are:

(1) Both 20CRs show strengths in building the precipitation climatology, at scales from yearly to monthly. However, there is significant decrease in the reconstruction quality of weekly/daily precipitation in the whole CONUS;

(2) The skill of precipitation reconstruction varies in different parts of CONUS, with the west coast indicating good quality even in daily scale. Thus it is possible to use 20CR for the heavy storms simulation in the west coast area;

(3) 20CRv2 is better at capturing the statistics in heavy rainfall events/periods, while ERA-20C is better at describing light rainy periods;

(4) 20CRv2 tends to overestimate the precipitation climatology, while ERA-20C shows consistent underestimations. Using both 20CRs as "ensemble" for quantitative precipitation reconstruction is worthwhile.

Data Policy: All data are available from public sources. Specifically, 20CRv2 data was obtained from UCAR (University Corporation for Atmospheric Research) Research Data Archive at http://rda.ucar.edu as ds131.2 dataset; ERA-20C data was obtaineded from the same website as ds626.0 dataset. Livneh precipitation database was obtained from NOAA's Earth System Research Laboratory at http://www.esrl.noaa.gov/psd/data/gridded/data.livneh.html. The data that have been processed for the analysis are available upon request to the first author.

References

Chen, X., and F. Hossain (2016), Revisiting Extreme Storms of the Past 100 Years for Future Safety of Large Water Management Infrastructures, *Earth's Future*, doi:10.1002/2016EF000368.

Compo, G. P. et al. (2011), The Twentieth Century Reanalysis Project, *Q. J. R. Meteorol. Soc.*, *137*(654), 1–28, doi:10.1002/qj.776.

European Centre for Medium-Range Weather Forecasts (2014), ERA-20C Project (ECMWF Atmospheric Reanalysis of the 20th Century), http://dx.doi.org/10.5065/D6VQ30QG, Research Data Archive at the National Center for Atmospheric Research, Computational and Information Systems Laboratory, Boulder, Colo. (Updated daily.) Accessed 16 Mar 2016.

Ferguson, C. R., and G. Villarini (2012), Detecting inhomogeneities in the Twentieth Century Reanalysis over the central United States, *J. Geophys. Res. Atmos.*, *117*(5), 1–11, doi:10.1029/2011JD016988.

Giese, B. S., G. P. Compo, N. C. Slowey, P. D. Sardeshmukh, J. A. Carton, S. Ray, and J. S. Whitaker (2010), The 1918/19 El Niño, *Bull. Am. Meteorol. Soc.*, *91*(2), 177–183, doi:10.1175/2009BAMS2903.1.

Gilbert P. Compo et al. (2015), NOAA/CIRES Twentieth Century Global Reanalysis Version 2c, http://dx.doi.org/10.5065/D6N877TW, Research Data Archive at the National Center for Atmospheric Research, Computational and Information Systems Laboratory, Boulder, Colo. (Updated yearly.) Accessed 16 Mar 2016.

Hagos, S. M., L. R. Leung, J. Yoon, J. Lu, and Y. Gao (2016), A projection of changes in landfalling atmospheric river frequency and extreme precipitation over western North America from the Large Ensemble CESM simulations, *Geophys. Res. Lett.*, *43*, 1–7, doi:10.1002/2015GL067392.

Hanna, E. et al. (2011), Greenland Ice Sheet surface mass balance 1870 to 2010 based on Twentieth Century Reanalysis, and links with global climate forcing, *J. Geophys. Res. Atmos.*, *116*(24), 1–20, doi:10.1029/2011JD016387.

Intergovernmental Panel on Climate Change et al. (2013), *Climate Change 2013 - The Physical Science Basis*.

Kalnay, E. et al. (1996), The NCEP/NCAR 40-year reanalysis project, *Bull. Am. Meteorol. Soc.*, *77*(3), 437–471, doi:10.1175/1520-0477(1996)077<0437:TNYRP>2.0.CO;2.

Kanamitsu, M., W. Ebisuzaki, J. Woollen, S. K. Yang, J. J. Hnilo, M. Fiorino, and G. L. Potter (2002), NCEP-DOE AMIP-II reanalysis (R-2), *Bull. Am. Meteorol. Soc.*, *83*(11), 1631–1643+1559, doi:10.1175/BAMS-83-11-1631.

Kong, X., and X. Bi (2015), Dynamical Downscaling of the Twentieth Century Reanalysis for China: Climatic Means during 1981–2010, *Atmos. Ocean. Sci. Lett.*, *8*(3), 166–173, doi:10.3878/AOSL20140099.

Livneh, B., E. A. Rosenberg, C. Lin, B. Nijssen, V. Mishra, K. M. Andreadis, E. P. Maurer, and D. P. Lettenmaier (2013), A long-term hydrologically based dataset of land surface fluxes and states for the conterminous United States: Update and extensions, *J. Clim.*, *26*(23), 9384–9392, doi:10.1175/JCLI-D-12-00508.1.

Mesinger, F. et al. (2006), North American regional reanalysis, *Bull. Am. Meteorol. Soc.*, *87*(3), 343–360, doi:10.1175/BAMS-87-3-343.

Misra, V., S. M. DiNapoli, and S. Bastola (2013), Dynamic downscaling of the twentieth-century reanalysis over the southeastern United States, *Reg. Environ. Chang.*, *13*(1), 15–23, doi:10.1007/s10113-012-0372-8.

Schreiner, L. C., and J. T. Riedel (1978), *Probable maximum precipitation estimates, United States east of the 105th meridian*, Department of Commerce, National Oceanic and Atmospheric Administration.

Sheffield, J. et al. (2013), North American Climate in CMIP5 experiments. Part I: Evaluation of historical simulations of continental and regional climatology, *J. Clim.*, *26*(23), 9209–9245, doi:10.1175/JCLI-D-12-00592.1.

Sun, Y., S. Solomon, A. Dai, and R. W. Portmann (2006), How often does it rain?, *J. Clim.*, *19*(6), 916–934, doi:10.1175/JCLI3672.1.

Tan, E. (2010), Development of a Methodology for Probable Maximum Precipitation Estimation over the American River Watershed Using the WRF Model, University of California, Davis.

Woldemichael, A. T., F. Hossain, and R. Pielke (2014), Impacts of Postdam Land Use/Land Cover Changes on Modification of Extreme Precipitation in Contrasting Hydroclimate and Terrain Features, *J. Hydrometeorol.*, *15*(2), 777–800, doi:10.1175/jhm-d-13-085.1.

Zhang, Q., H. Körnich, and K. Holmgren (2013), How well do reanalyses represent the southern African precipitation?, *Clim. Dyn.*, *40*(3-4), 951–962, doi:10.1007/s00382-012-1423-z.

Tables

Table 1. Details of two 20CRs

Item	20CRv2	ERA-20C
Produced by	NOAA-CIRES	ECMWF
Temporal coverage	1850/12/31-2014/12/31	1900/1/1-2010/1/1
Temporal resolution	6-hourly	3-hourly
Spatial coverage	Global	Global (89.142S – 89.142N)
Spatial resolution	T62 grid (~210km)	T159 grid (~125km)
Vertical layers	28	91
Data assimilated	Surface observations of synoptic pressure	Surface and mean sea level pressure; Surface marine winds
Data assimilation method	Ensemble Kalman filter	4D-VAR

Figures

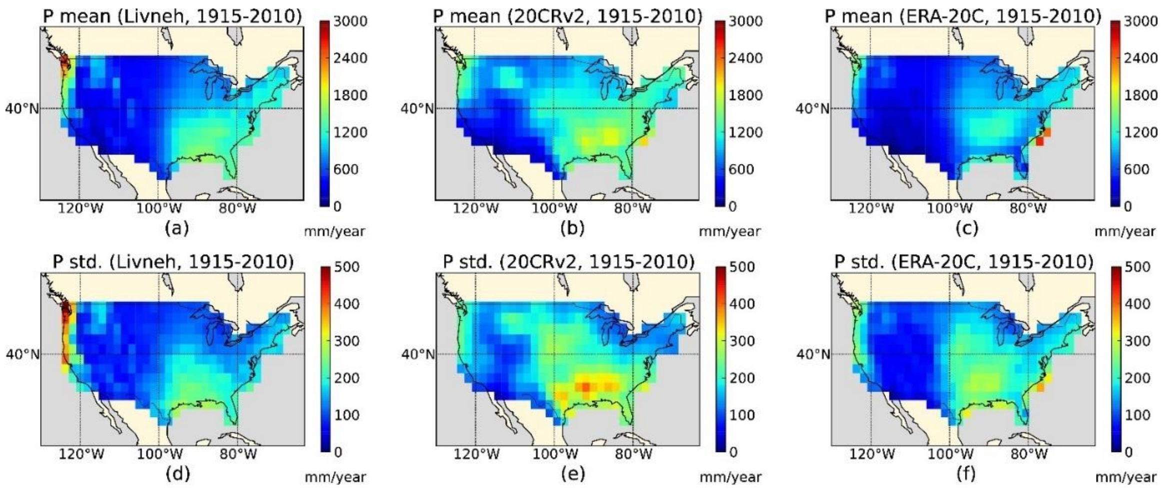


Figure 1. Precipitation climatology (1915-2010) from observation and model reconstructions. (a) multiyear average annual precipitation in the reference dataset; (b) and (c) multiyear average annual precipitation in 20CRv2 and ERA-20C. Panels (d-f) are the standard deviation of annual precipitation.

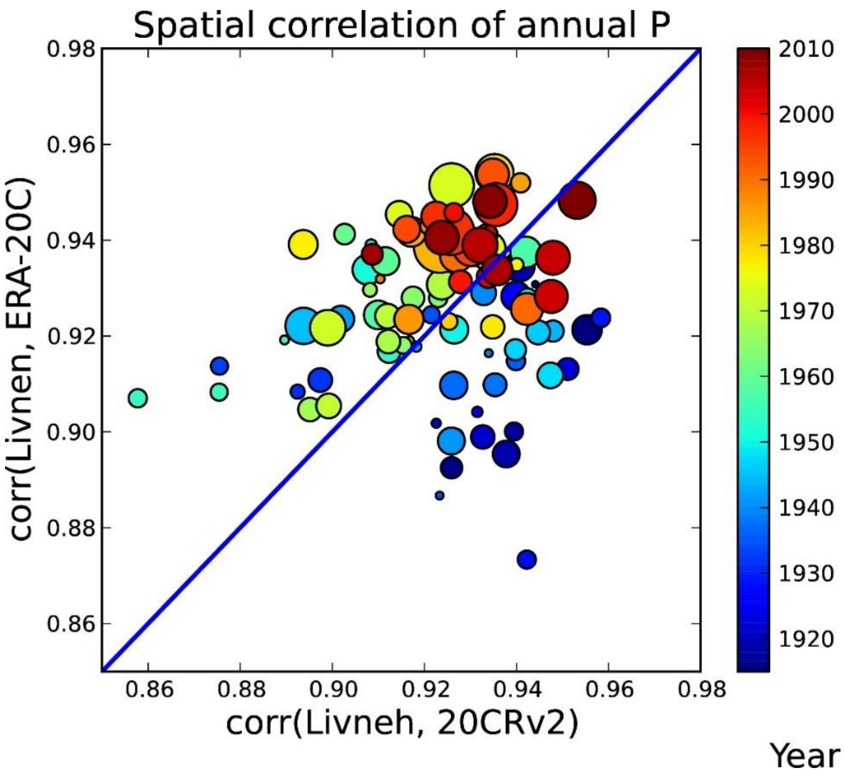


Figure 2. Spatial correlation between annual precipitation maps. Colors of points indicate different years, and sizes of points are the total rain from the given year (from Livneh dataset).

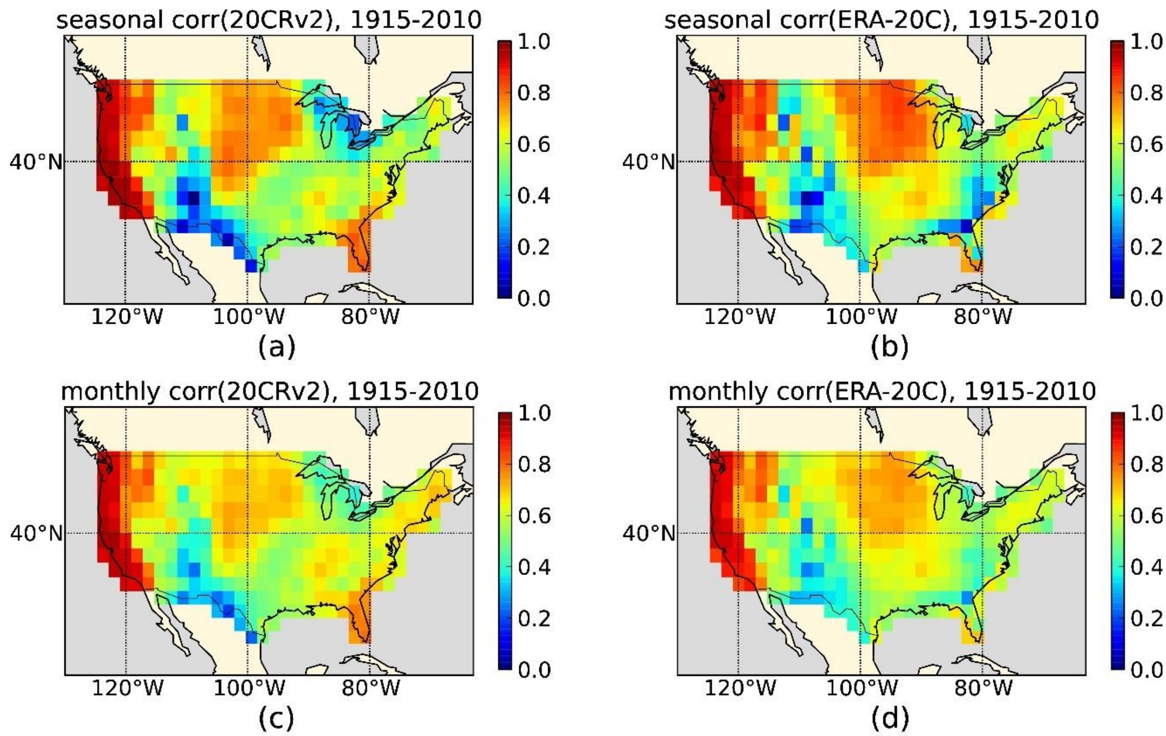


Figure 3. Spatial distribution of correlation of seasonal and monthly precipitation.

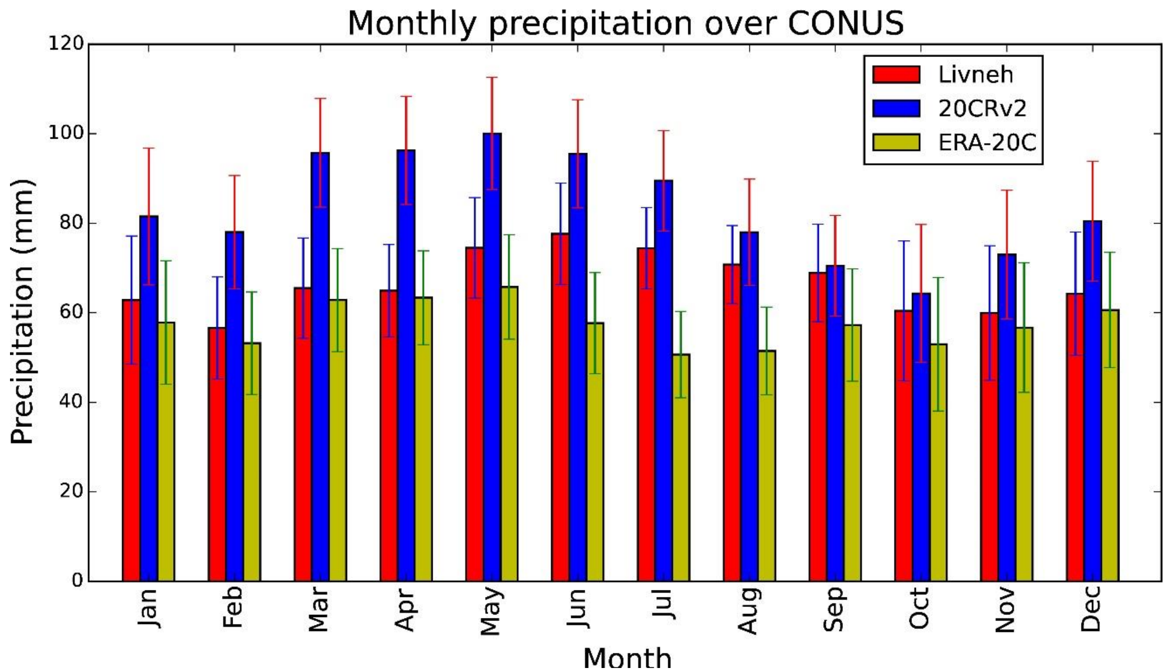
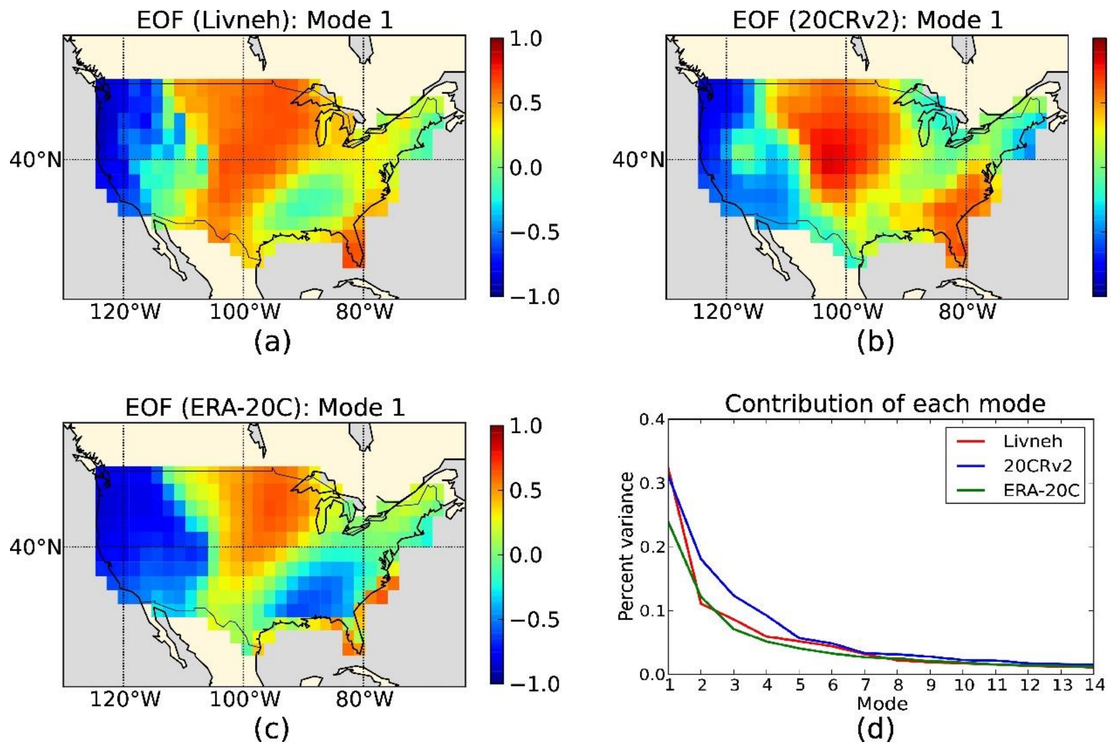


Figure 4. Statistics of monthly precipitation over CONUS. Error bars show the standard deviation.

272



273 **Figure 5. EOF analysis of monthly precipitation data.**

274

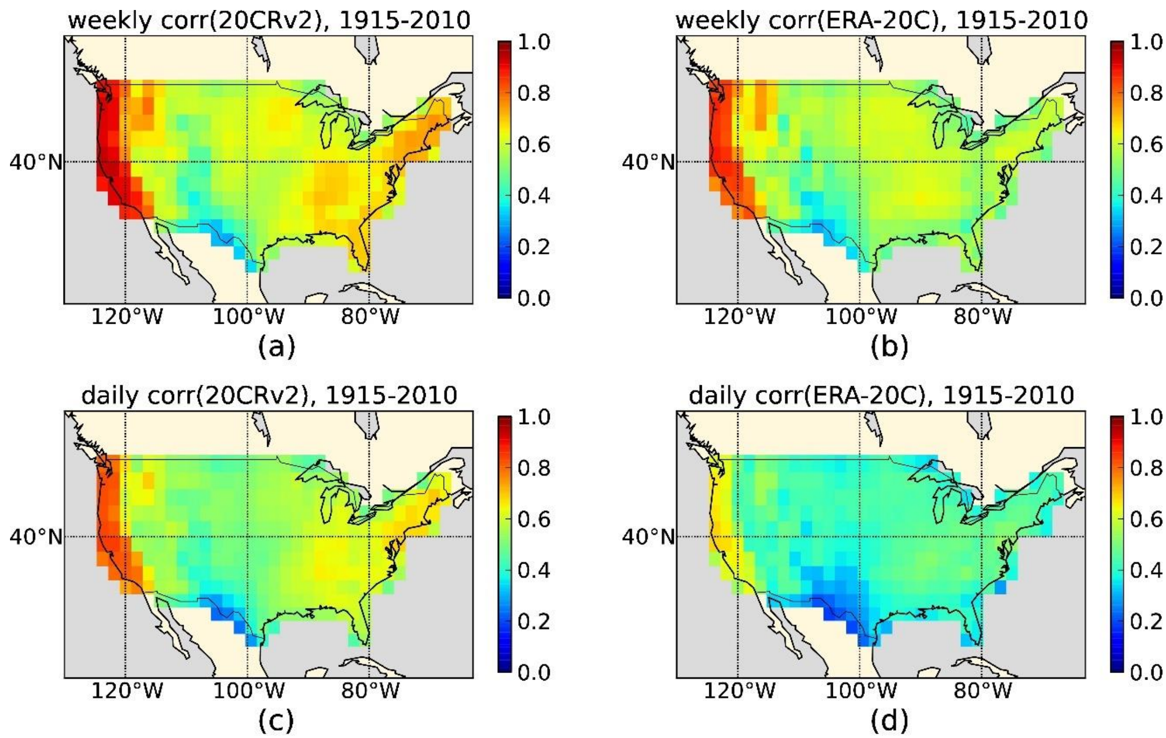


Figure 6. Spatial distribution of correlation of weekly and daily precipitation.

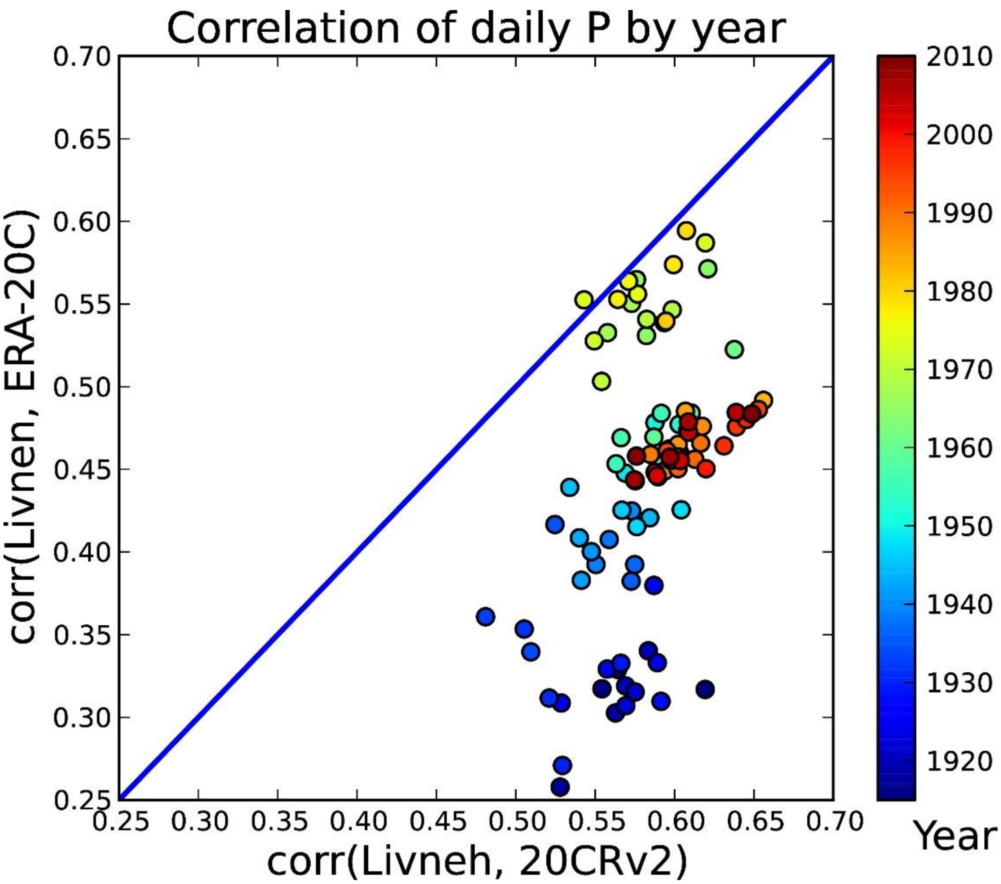


Figure 7. Correlation of daily precipitation as a function of year.