Journal of Hydrology 553 (2017) 559–573

Contents lists available at ScienceDirect

Journal of Hydrology

journal homepage: www.elsevier.com/locate/jhydrol

Research papers

High resolution decadal precipitation predictions over the continental United States for impacts assessment

Kaustubh Salvi [a,*], Gabriele Villarini [a], Gabriel A. Vecchi [b,1]

[a] IIHR-Hydroscience & Engineering, The University of Iowa, Iowa City, IA, USA
[b] NOAA/Geophysical Fluid Dynamics Laboratory, Princeton, NJ, USA

ARTICLE INFO

Article history:
Received 25 December 2016
Received in revised form 9 June 2017
Accepted 22 July 2017
Available online 24 July 2017
This manuscript was handled by A. Bardossy, Editor-in-Chief, with the assistance of Uwe Haberlandt, Associate Editor

Keywords:
Decadal predictions
Continental United States
Statistical downscaling
Precipitation

ABSTRACT

Unprecedented alterations in precipitation characteristics over the last century and especially in the last two decades have posed serious socio-economic problems to society in terms of hydro-meteorological extremes, in particular flooding and droughts. The origin of these alterations has its roots in changing climatic conditions; however, its threatening implications can only be dealt with through meticulous planning that is based on realistic and skillful decadal precipitation predictions (DPPs). Skillful DPPs represent a very challenging prospect because of the complexities associated with precipitation predictions. Because of the limited skill and coarse spatial resolution, the DPPs provided by General Circulation Models (GCMs) fail to be directly applicable for impact assessment. Here, we focus on nine GCMs and quantify the seasonally and regionally averaged skill in DPPs over the continental United States. We address the problems pertaining to the limited skill and resolution by applying linear and kernel regression-based statistical downscaling approaches. For both the approaches, statistical relationships established over the calibration period (1961–1990) are applied to the retrospective and near future decadal predictions by GCMs to obtain DPPs at ~4 km resolution. The skill is quantified across different metrics that evaluate potential skill, biases, long-term statistical properties, and uncertainty. Both the statistical approaches show improvements with respect to the raw GCM data, particularly in terms of the long-term statistical properties and uncertainty, irrespective of lead time. The outcome of the study is monthly DPPs from nine GCMs with 4-km spatial resolution, which can be used as a key input for impacts assessments.

1. Introduction

Water is the most abundant natural resource that manifests itself in diverse states and forms such as precipitation, snowfall, surface and ground water, rivers, lakes, oceans. Despite such a wide availability, most of our activities pertaining to water are constrained because of a lower percentage of available fresh water, upcoming problems concerning water usage, and even the solutions adopted to negotiate the problems. For instance, to address elevated food requirements of the rising population (population more than quadrupled in 100 years), measures such as doubled crop land area, sixfold increase in irrigated area (Freydank and Siebert, 2008) are implemented, leading to rise in global water use (i.e., withdrawal) by nearly 8 times with a steep increase at a rate of 15% per decade between 1960 and 2010. Such measures have put additional constraints on the available fresh water resources and the exacerbating situations demand careful planning and mastering of the available water resources. However, design and implementation of water resources planning measures is a difficult task because precipitation represents a key input. Precipitation is the most important and equally complex climate variable to understand and foresee mainly because of extreme variability, revealed by precipitation patterns at different spatio-temporal resolutions. The variability in precipitation might lead to high-intensity disasters such as flooding and mudslides (Trenberth et al., 2007), or lower-intensity, longer-duration events such as droughts (e.g., Gray, 2009; Henry et al., 2004; Hunter et al., 2011). Such catastrophes not only lead to heavy infrastructure damages and pose a serious threat to life but also lead to sensitive issues such as migration, which has been a common strategy to avert the consequences of weather events and/or a changing cli-

---

* Corresponding author at: IIHR-Hydroscience & Engineering, The University of Iowa, 100 C. Maxwell Stanley Hydraulics Laboratory, Iowa City 52242, IA, USA.
E-mail address: kaustubhanil-salvi@uiowa.edu (K. Salvi).

[1] Current affiliation: Geosciences Department, Princeton University, Princeton, NJ, USA.

http://dx.doi.org/10.1016/j.jhydrol.2017.07.043

mate (e.g., Nawrotzki et al., 2013; Kniveton et al., 2008; McLeman and Smit, 2006).

The background discussed up to this point highlights the importance of precipitation and the associated problems, and leads us to concluding that such problems can be mitigated only through robust planning policies. The success of such measures would be highly dependent on the knowledge of future precipitation, which itself represents an extremely challenging task. Our current best knowledge of future precipitation and the associated uncertainties can be obtained from projections from General Circulation models (GCMs) produced in support of the fifth generation of Coupled Model Intercomparison Project (CMIP5) framework. This includes results from a standard set of experiments by state-of-the-art climate models developed by different teams of experts worldwide. This "ensemble of opportunity" (commonly used terminology to represent the suite of GCM data; e.g., Sanderson et al., 2008) is the best available source of information on the range of physically plausible climate evolutions over the next century. Representative Concentration Pathways (RCPs) from the CMIP5 suite provide us with future projections of climate variables as a possible response to different radiative forcings. These projections solve our problem to a certain extent and provide us with the climate projections as a possible response to these forcings. However, there is a three-fold problem associated with the use of GCM projections in the planning process: (1) the RCP scenarios are driven by end-conditions, where the trajectories of different climate variables are plausible responses to the corresponding radiative forcing at the end of the 21st century. Such projections form a better way of understanding the long-term future of climate variables, but they cannot be used for short-term planning policies. This is because realistic weather conditions that we encounter in our day-to-day life are driven by their initial rather than their boundary conditions. Also, it is established that long-term behaviors of climate variables may differ from their short term properties (e.g., Cane, 2010; van Oldenborgh et al., 2012). (2) A complex variable such as precipitation is difficult to predict and the skill exhibited by the GCM in projecting them is limited. (3) The coarse spatial resolution at which the GCM projections are available cannot be used for impacts assessment over limited areas. CMIP5 simulations themselves tend to resolve the first problem by providing decadal scale forecasts that are initialized every year or intermittently to provide predictions for the next 10–30 years. These predictions are known as decadal predictions, which range from 10 to 30 year lead times (Meehl et al., 2009) and are forced with observed initial and boundary conditions. Initialized decadal predictions, while still in their infancy, have been subject to increasing attention by the scientific community. These are called upon by different names, including "a new kid on the block," "the fascinating baby that all wish to talk about" (Goddard et al., 2012), "high profile predictions" (Goddard et al., 2013). Several scientists are currently working on evaluating: (i) different aspects of this new product such as initialization strategies (Meehl et al., 2014), assessment of skill in decadal predictions (Meehl et al., 2009), gap between decadal climate predictability and predictions, impacts of initialization (Branstator and Teng, 2012), opportunities and challenges, importance of ocean observations for successful decadal predictions (Hurrell et al., 2009), positive phase of the inter-decadal Pacific oscillations (Meehl et al., 2016), role of sea ice, land surface, stratosphere, and aerosols in decadal-scale predictability (Bellucci et al., 2015a), and comparison between initialized and non-initialized predictions (Fyfe et al., 2011); (ii) performances of individual GCM with respect to the decadal predictions e.g. Flexible Global Ocean-Atmosphere-Land System model, Grid-point Version 2 (FGOALS-g2) (Bin et al., 2013), coupled Earth System model of the Max Planck Institute for Meteorology (MPI-ESM) (Müller et al., 2012), MIROC4h and MIROC5 (Mochizuki et al., 2011); and (iii) the skill of decadal predictions for different climate variables and quantities, including surface temperature (Smith et al., 2007; Choi et al., 2016; Salvi et al. 2017), Sahelian precipitation (Gaetani and Mohino, 2013), hurricane activity (Smith et al., 2010; Vecchi et al., 2013), and regional surface climate (Kim et al., 2012; van Oldenborgh et al., 2012; Doblas-Reyes et al., 2013; Goddard et al., 2013; Caron et al., 2014; Meehl et al., 2014; Bellucci et al., 2015b). However, despite recent progress, we still have to comprehensively quantify the quality of GCM precipitation predictions, including over the United States; moreover, even though CMIP5 simulations have outperformed CMIP3 in terms of spatial resolution, the resolution is still not high enough for impacts assessment. With this background, the goal of this work is to obtain high-resolution precipitation decadal predictions over the continental United States (CONUS).

The specific objectives of this study involve: 1) the evaluation of raw GCM precipitation skill over CONUS; 2) the enhancement of the skill using two data-driven approaches; and 3) the development of a dataset of decadal predictions of precipitation over CONUS at a spatial resolution of ~4 km. The two data-driven approaches that are implemented in this study are transfer function based statistical downscaling (SD) methodologies. Transfer function based downscaling methodologies rely upon establishing statistical relationships between coarse resolution climate variables (predictors) that are relatively well understood in terms of underlying physics and hence, well simulated, and the fine resolution climate variable of interest that needs to be downscaled, i.e. predictand (precipitation in this study). The established relationships are applied to GCM predictions to obtain downscaled data. We use two SD approaches: (1) linear regression (LR) based, which assumes a parametric form of the relationship, with the variation between predictors and predictand that is assumed to be linear; and (2) kernel regression (KR) based, which is a non-parametric form of regression. The premise behind the selection of two methodologies that are categorized under the same cluster (transfer function approach) is explained in section 3. These methodologies are applied to the predictors by nine GCMs to obtain improved decadal precipitation predictions at fine spatial resolution (~4 km). The manuscript is organized as follows. Details about the study region and the data used are in Section 2. Section 3 provides details about the data-driven approaches, which are applied here to enhance the prediction skills and different evaluation metrics. The results of the evaluation of raw and processed GCMs' skill are discussed in Section 4. We conclude this study with summary and discussion in Section 5, followed by concluding remarks.

2. Study region and data

Fig. 1 shows the spatial variations of precipitation over the study region for the period 1961–2014, obtained using the Parameter–Elevation Regressions on Independent Slopes Model (PRISM) data product (additional information about the PRISM data product can be found in Section 2.1). The figure also provides the details about the seven regions we focus on: Northeast (NE), Midwest (MW), Southeast (SE), Great Plains North (GPN), Great Plains South (GPS), Northwest (NW), and Southwest (SW). These regional delineations are largely based on the National Climate Assessment Report (Karl et al., 2009) and then implemented in different studies (e.g., Pryor and Schoof, 2008). Similar to the recent studies, which use slightly modified delineations (Schoof et al., 2010; Kunkel et al., 2013; Mutiibwa et al., 2015), we divide the Great Plains into two regions GPN and GPS. The division of the Great Plains into two visually homogeneous regions (homogeneous in terms of spatial variation of average precipitation) is an additional advantage for the improvement in performance of SD. In this study, we focus on the seven regions in Fig. 1 and consider each one as an individ-