Atmospheric Research 183 (2017) 341–350

Contents lists available at ScienceDirect

Atmospheric Research

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

An investigation of predictability dynamics of temperature and precipitation in reanalysis datasets over the continental United States

C.T. Dhanya [a,*], Gabriele Villarini [b]

[a] Department of Civil Engineering, Indian Institute of Technology Delhi, New Delhi, India
[b] IIHR-Hydroscience & Engineering, The University of Iowa, Iowa City, IA, USA

ARTICLE INFO

Article history:
Received 11 May 2016
Received in revised form 18 September 2016
Accepted 23 September 2016
Available online 24 September 2016

Keywords:
Lyapunov exponent
Chaotic dynamics
NCEP reanalysis
MERRA
ERA-interim
JRA-55
Error and instability

ABSTRACT

Reanalysis datasets have been under critical scrutiny due to their widespread use in various climatic and hydrological modeling applications, in particular over many areas of the globe with limited or absent reliable observational data. Nevertheless, reanalysis products are in the process of continuous improvements reflecting the improved system knowledge, model physics and assimilation techniques. In addition, several internal model adjustments have also been adopted to minimize the bias in reanalysis datasets. Considering these factors, it is necessary to investigate the inherent chaotic dynamics of reanalyses and the possible discrepancies, if any, with respect to the observational data. Here we compare and contrast the chaotic dynamics of daily precipitation and daily mean surface temperature simulated by the reanalysis against observed data over the continental United States. Our focus is on four reanalysis products: the National Aeronautics and Space Administration's Modern Era Retrospective-Analysis for Research and Applications (MERRA), European Centre for Medium-Range Weather Forecasts' ERA-Interim, Japanese Meteorological Agency's Japanese 55-year Reanalysis (JRA-55), and National Center for Environmental Prediction/National Center for Atmospheric Research's Reanalysis I. The inherent chaotic dynamics measured in terms of three statistics (i.e., maximum predictability, predictive error and predictive instability) reveal the inconsistency among the four reanalysis products. ERA-Interim is capable of simulating the precipitation's chaotic dynamics over much of the study region, while MERRA is found to be superior in capturing the temperature's chaotic dynamics. Analyses on various aspects of daily precipitation and temperature indicate that the biases in precipitation's chaotic dynamics may be attributed to the inconsistencies in simulating the signal-to-noise ratio and non-rainy days, while biases in temperature's chaotic dynamics could be due to the failure in replicating the abrupt trends in the recent decades by the reanalyses products.

1. Introduction

Long consistent records of the global climate system and accurate information of its variability are essential for a number of research and societal applications. Moreover, reliable meteorological records based on observations are not always available for large areas of the world. These shortcomings are generally mitigated by using global retrospective analysis (or reanalysis) datasets. Reanalysis products provide a temporally and spatially consistent continuous record of the global state of the atmosphere by assimilating available observations within state-of-the-art numerical weather prediction (NWP) models. These products have been used for a wide range of applications, including climate monitoring, development and improvements of climate models, generation of meteorological datasets, and assessment of water budget. Due to their widespread use especially in climatologic and hydrologic analyses, the ability of reanalysis products to accurately simulate the global dynamics has always been under scrutiny (e.g., Rood and Bosilovich, 2010), with particular emphasis on the consistency in reproducing the variability in atmospheric moisture transport (e.g., Bosilovich et al., 2011; Trenberth et al., 2011; Zhao and Li, 2006), precipitation (e.g., Bosilovich et al., 2008; Decker et al., 2012; Lin et al., 2014) and temperature (e.g., Bosilovich, 2013; Chen et al., 2014; Decker et al., 2012).

The uncertainties in reanalyses primarily arise from the use of NWP models in reproducing the observational records, and from the assimilation of various datasets with different densities and characteristics (e.g., satellite, weather stations). Moreover, there are uncertainties due to the deficiencies in the model's physics leading to incorrect representations of system dynamics. A number of major agencies and research centers, such as the National Centers for Environmental Prediction (NCEP), the European Centre for Medium-Range Weather Forecasts (ECMWF), the National Aeronautics and Space Administration (NASA) and the Japanese Meteorological Agency (JMA) have been developing and improving these products.

* Corresponding author at: Department of Civil Engineering, Indian Institute of Technology Delhi, New Delhi, India.
E-mail address: dhanya@civil.iitd.ac.in (C.T. Dhanya).

http://dx.doi.org/10.1016/j.atmosres.2016.09.017
0169-8095/

Despite recent advancements, including 4D data assimilation techniques, the strengths and weaknesses of reanalysis datasets in simulating different aspects of moisture transfer and various hydro-meteorological variables have been discussed in different studies (e.g., Bosilovich, 2013; Chen et al., 2014; Feng and Houser, 2014; Ferguson and Villarini, 2012; Hodges et al., 2011; Kang and Ahn, 2015; Lin et al., 2014; Thorne and Vose, 2010; Zhao et al., 2015). The ability of reanalyses in accurately simulating various components of the hydrological cycle, especially precipitation and temperature, has been critically examined, with particular emphasis on occurrences, decadal and seasonal variability, trends, just to cite a few. Overall, great caution is advised while using model diagnostic hydrologic variables from reanalyses (Bosilovich et al., 2011; Trenberth et al., 2011).

While the discrepancies in terms of consistency among reanalysis and predictive skill (in terms of the performance assessed using different skill metrics such as bias, root mean square error (RMSE), correlation between the model simulations and the observations) have been investigated (e.g., Bosilovich et al., 2007; Feng and Houser, 2014; Fu Xiouhua et al., 2011; Hodges et al., 2011; Hofer et al., 2012; Lee et al., 2010), the inherent predictability dynamics of the various reanalyses time series with respect to the observations is still unexplored. Unlike the common definition of the term “dynamics” used in any study in the NWP literature (e.g., Bauer et al., 2015; Kalnay, 2003; Li et al., 2013; Lorenz, 1965; Somerville, 1987), the use of the term “dynamics” in the present study indicates the sensitivity due to initial conditions, the divergence of trajectories and the uncertainty arising thereby in any chaotic time series. Hence, the term “dynamics” used in this study refers more to chaotic dynamics limiting the predictability of the time series, which can be measured by computing the maximum Lypaunov exponent of the series. The research questions we will address in this study are: (i) Are the reanalyses dynamically consistent with the observed variables? (ii) Are the trajectory dynamics and the sensitivity to initial conditions of the observations and reanalyses comparable? The investigation of these questions would reveal the similarities and dissimilarities in the inherent dynamics of reanalyses with respect to observations or reference datasets. The goal of this study lies in the analysis of the consistency in the dynamics of precipitation and temperature from various reanalyses datasets across the United States in terms of the inherent predictability of the system. Inherent predictability is estimated through the Lyapunov exponent, a commonly used measure of the exponential divergence of the nearby trajectories.

2. Data and study region

We use daily precipitation data with 0.25° × 0.25° resolution over the continental United States from the National Oceanic and Atmospheric Administration (NOAA) Climate Prediction Center (CPC) Unified Precipitation Project (http://www.esrl.noaa.gov/psd/) as precipitation reference data. The spatial coverage is 20°N–50°N, and 230°E–305°E and we focus on the period from 1 January 1979 to 31 December 2006 (to avoid the gaps in the CPC time series after 2006).

We use mean daily temperature from North American Land Data Assimilation System (NLDAS-2) as reference data. NLDAS-2 primary forcing is derived from North American Regional Reanalysis (NARR). NARR is reported to exhibit substantial improvement in the accuracy of (especially) temperatures throughout the troposphere, when compared with Global Reanalysis (Mesinger et al., 2006). The spatial resolution of NLDAS-2 is 0.125° × 0.125° and we focus on the period from 1 January 1979 to 31 December 2014.

We consider four of the most widely used reanalysis products: NCEP/National Center for Atmospheric Research (NCAR) Reanalysis I (Kalnay et al., 1996), ECMWF's ERA-Interim (Dee et al., 2011b), NASA's Modern-Era Retrospective Analysis for Research and Applications (MERRA; Rienecker et al., 2011), and JMA's Japanese 55-year Reanalysis (JRA-55; Ebita et al., 2011; Kobayashi et al., 2015). We restrict our analyses to these four reanalysis products because they are widely used in climatologic and hydrologic studies.

The spatial resolution of these reanalyses differs from the reference data. Because the main objective of this study is to check the consistency in internal dynamics through predictability, the reference datasets are interpolated to the grid resolution of the different reanalyses through a nearest neighbour method.

3. Methodology

In any chaotic system, the limit in predictability arises due to the dependence on initial conditions and subsequent exponential divergence of nearby trajectories, thereby limiting the predictability after a few time steps. The Lyapunov spectrum measures the divergence of trajectories with respect to all dimensions and hence, gives a limit of predictability. The rate of divergence may vary for different dimensions, depending upon the total dimension of the system considered. The divergence of trajectories with respect to all dimensions is referred as Lyapunov spectrum. The vast majority of the studies concentrate on maximal Lyapunov exponent instead of computing the whole spectrum because the maximal Lyapunov exponent itself represents a fair approximation of the total uncertainty exhibited by the system. Many algorithms have been developed to calculate the maximal Lyapunov exponent (e.g., Kantz, 1994; Rosenstein et al., 1993; Wolf et al., 1985). More details on the computation of maximal Lyapunov exponent and its application in daily hydrological time series can be found in Dhanya and Nagesh Kumar (2011, 2013). The algorithms developed to estimate the Lyapunov exponent or Lyapunov spectrum either neglect the local nature of the divergence or dilute the computations with linearity assumption for divergence. However, the assessment of uncertainty of highly complex nonlinear systems (e.g., atmosphere) using these algorithms may be inappropriate. In light of these limitations, a nonlinear finite time Lyapunov exponent (FTLE) was introduced by Ding and Li (2007) to estimate the predictability limit of geopotential height. In this study, we employ FTLE to measure the predictability characteristics of temperature and precipitation over the continental United States.

FTLE takes into account the expected variations in the initial error growth at different places in the attractor, which complex atmospheric systems usually exhibit (e.g., Nese, 1989; Yoden and Nomura, 1993; Ziehmann et al., 2000). The average error growth of trajectories with respect to the initial position is measured, and the time at which the initial error growth reaches a saturation value is taken as the predictability limit. The detailed steps to calculate FTLE and predictability limit can be found in Ding and Li (2007) and Ding et al. (2008).

Any longer-lead predictions are meaningless after this limit, because the system is assumed to have achieved a stochastic state at this point.

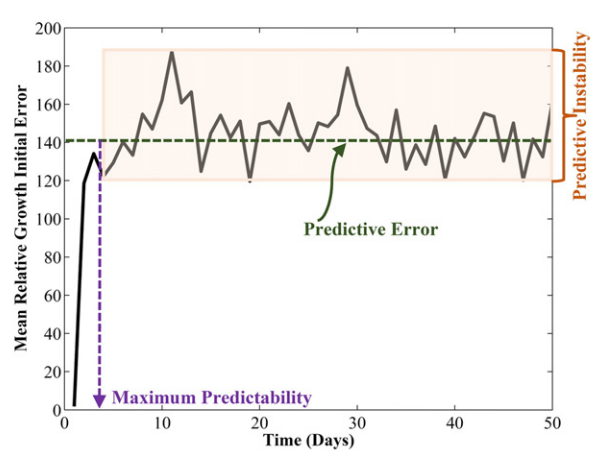

Fig. 1. Illustration of the typical evolution of the relative growth of the initial error with time.