## Detecting Inhomogeneities in the Twentieth Century Reanalysis over the Central United States

Craig R. Ferguson

Department of Hydrology and Water Resources Engineering, Institute of Industrial Science, The University of Tokyo, Tokyo, Japan

Gabriele Villarini Department of Civil and Environmental Engineering, Princeton University, Princeton, NJ, USA Willis Research Network, London, UK

A number of recent modeling studies suggest that, as our climate warms, the global hydrologic cycle will, and perhaps is, accelerating (i.e., increased precipitation, runoff, and evapotranspiration). For some regions, however, there is empirical evidence to the contrary. For instance, observed decreases (i.e., stilling) in mean winds over the U.S. and Australia and declines in global land evapotranspiration (due to soil-moisture constraints) indicate that the hydrologic cycle is decelerating. Improved quantification and attribution of changes in the hydrologic cycle are critical to our capacity to mitigate their socio-economic impacts.

Only long-term reanalyses offer a comprehensive and consistent observationally-based record of the climate system suitable for examining the complex interactions and feedbacks that effect hydrologic change. Unfortunately, their application in long-term trend assessment has historically been precluded due to two key shortcomings: (1) they suffer from observational "shocks", or unphysical time-varying biases associated with an evolving observational system [e.g., the introduction of Special Sensor Microwave Imager (SSM/I) in July 1987 and the Advanced Microwave Sounding Unit (AMSU-A) in November 1998] and (2) their temporal coverage (1979 to present) is insufficient for separating abrupt artificial shifts from natural climate variations. The recently released Twentieth Century Reanalysis Version 2 (henceforth, 20CR) was explicitly developed to address these limitations (data inhomogeneity and short record length) through the assimilation of surface synoptic observations only (surface- and sealevel pressure data) over an unprecedented 138-year period spanning 1871-2008.

In this study, we address two key questions: (1) Can the 20CR be used for the assessment of changes in different hydrometeorological variables over the 20th century? and (2) What would an experimental framework for differentiating between artificial and true climate signals comprise? A unique attribute of the 20CR is that it provides quantitative uncertainty estimates (i.e., the ensemble spread) for each variable field at every time step. These complementary data carry information on the assimilated observation distribution and density and likewise represent a key resource for separating climate-related changes from non-climate related artifacts.

Our analysis is focused on a  $15^{\circ} \times 15^{\circ}$  region ( $30^{\circ}-45^{\circ}N$ ,  $105^{\circ}-90^{\circ}W$ ) in the central U.S. where improved understanding of 20th century hydroclimatological trends and variability is a top priority to both scientific and economic stakeholders. The region is a well-known surface temperature "warming hole" and one of only a few Global Energy and Water Cycle Experiment (GEWEX) Global Land-Atmosphere Coupling Experiment (GLACE) land-atmosphere interaction "hot spots". We apply three statistical methods (Pettitt and Bai-Perron tests, segmented regression) to detect abrupt shifts in the monthly record of six primary land surface hydrological variables: 2 m air temperature, surface runoff, precipitation, surface latent heat flux, surface sensible heat flux, total column precipitable water vapor/ice; and three derived metrics descriptive of land-atmosphere interactions: lifting condensation level (*LCL*), low-level humidity index (*HI*), and convective triggering potential (*CTP*). We use the 3-hour forecast (first guess) fields, except in the case of *HI* and *CTP*, which are derived from the analysis (i.e., surface pressure and multi-level fields of geopotential height, specific humidity, and air temperature). For the primary variables, we analyze the associated monthly time-average of the 3-hourly uncertainty fields as well. *LCL*, *HI*, and *CTP* are not computed by 20CR directly and thus, no estimate of uncertainty is available. We use the monthly assimilated observation counts, available on a  $5^{\circ} \times 5^{\circ}$  geographic grid, to visually confirm correspondence between assimilated observation count and uncertainty. For surface air temperature and precipitation, we use the Climate Research Unit (CRU) time series dataset version 3.1 (TS3.1) for comparison.

The Pettitt and Bai-Perron tests serve to detect abrupt changes in the hydrometeorological variables because the inhomogeneities generally occurred as step changes. On the other hand, segmented regression is used to examine the presence of inhomogeneities in the associated uncertainty fields, owing to the broken-line nature of their time series. The use of these three tests concurrently, together with a visual assessment of the time series, provides a measure of the robustness of the results from any single test.

We find that for warm-season months, there is a consensus change-point among all variables between 1940 and 1950, which is not substantiated by the CRU record. While we cannot unequivocally conclude that the shift in the 20CR analysis fields is artificial, the fact that (1) the timing of abrupt change follows subsequent to dramatic shift in the density of the underlying observational record and (2) the change-point is not corroborated by the CRU datasets for air temperature and precipitation, makes the homogeneity of the 20CR in this region at the very least questionable. Our recommendation is therefore for users to restrict climate trend applications to the second half-century of the 20CR record, after observational density has stabilized.

Extended analyses performed over several large hydrologic basins in diverse climate regimes are discussed in light of our findings for the central U.S.

## **Corresponding Author:**

Name:	Craig R. Ferguson
Organization:	Institute of Industrial Science, The University of Tokyo
Address:	Ce407
	Institute of Industrial Science
	The University of Tokyo
	4-6-1 Komaba
	Meguro-ku, Tokyo, Japan 153-8505