Histogram Anomaly Time Series

A Compact Graphical Representation of Spatial Time Series Data Sets

Gerald L Potter, George J. Huffman, David T. Bolvin,
Michael G. Bosilovich, Judy Hertz, and Laura E. Carriere

ABSTRACT: We introduce a simple method for detecting changes, both transient and persistent, in reanalysis and merged satellite products due to both natural climate variability and changes to the data sources/analyses used as input. This note demonstrates this Histogram Anomaly Time Series (HATS) method using tropical ocean daily precipitation from MERRA-2 and from GPCP One-Degree Daily (1DD) precipitation estimates. Rather than averaging over space or time, we create a time series display of histograms for each increment of data (such as a day or month). Regional masks such as land–ocean can be used to isolate particular domains. While the histograms reveal subtle structures in the time series, we can amplify the signal by computing the histogram’s anomalies from its climatological seasonal cycle. The qualitative analysis provided by this scheme can then form the basis for more quantitative analyses of specific features, both real and analysis induced. As an example, in the tropical oceans the analysis clearly identifies changes in the time series of both reanalysis and observations that may be related to changing inputs.

KEYWORDS: Atmosphere; Precipitation; Satellite observations; Time series; Reanalysis data

<https://doi.org/10.1175/BAMS-D-20-0130.1>

Corresponding author: George J. Huffman, george.j.huffman@nasa.gov

In final form 12 September 2020

©2020 American Meteorological Society

For information regarding reuse of this content and general copyright information, consult the [AMS Copyright Policy](#).

AFFILIATIONS: **Potter, Hertz, and Carriere**—Center for Climate Simulation, NASA Goddard Space Flight Center, Greenbelt, Maryland; **Huffman**—Mesoscale Atmospheric Processes Laboratory, NASA Goddard Space Flight Center, Greenbelt, Maryland; **Bolvin**—Mesoscale Atmospheric Processes Laboratory, NASA Goddard Space Flight Center, and Science Systems and Applications Inc., Greenbelt, Maryland; **Bosilovich**—Global Modeling and Assimilation Office, NASA Goddard Space Flight Center, Greenbelt, Maryland

It is often said that reanalysis should not be used for trend analysis for such quantities as precipitation (Bosilovich et al. 2017; Robertson et al. 2011). This is largely due to the introduction of new observations or observational methods in the input data, and the same issue arises in merged satellite products that use a variety of inputs, such as those in the daily Global Precipitation Climatology Project One-Degree Daily Precipitation Dataset (GPCP 1DD) product (Huffman et al. 2001), despite the work done by the developers to intercalibrate the various inputs. It would be useful to identify transient and persistent changes in the actual distribution of precipitation intensity and separate them from the effects of changing satellite sensors at both short and long time scales (Adler et al. 2018). In this paper, we develop a new qualitative analysis that allows scientists to zero in on transient or persistent fluctuations, which can then be further analyzed to quantify the nature of these events. Such an analysis could make these reanalysis and observational datasets more useful for determining naturally occurring variations and trends in the time series. Here, we report on a simple method that utilizes a time series of histogram anomalies of precipitation that is sensitive to changes in the dataset, whether genuine or artificial, which we name the Histogram Anomaly Time Series (HATS).

Data

To demonstrate the usefulness of HATS, we use the repackaged Modern-Era Retrospective Analysis for Research and Applications, version 2 (MERRA-2) (Gelaro et al. 2017), precipitation from the Collaborative Reanalysis Technical Environment (CREATE) service (Potter et al. 2018) and the GPCP 1DD version 1.3 precipitation estimates as contained in the Earth System Grid Federation (ESGF) (Teixeira et al. 2014) and also available at www.ncei.noaa.gov/data/global-precipitation-climatology-project-gpcp-daily/access. The daily GPCP data are provided at the $1^\circ \times 1^\circ$ latitude–longitude resolution. The MERRA-2 daily precipitation data are aggregated from the 1-h reanalysis product and then regridded from the native $0.5^\circ \times 0.625^\circ$ latitude–longitude resolution to the daily GPCP resolution by using a standard bilinear regridding procedure.

Method

The ocean domain studied in this paper is shown in Fig. 1. Initially, we attempted to isolate changing precipitation distributions with a probability distribution function (PDF)

Fig. 1. Land–ocean mask. The ocean domain used for this paper is denoted by the blue shading inside the white border. The gray color is the masked land area.

calculated at different time intervals (Wehner et al. 2014). Figure 2 shows the PDF of tropical (20°S–20°N) ocean-only precipitation for MERRA-2 and GPCP for two different time periods (1996–2008 and 2009–18). The precipitation rate is only plotted on the range 0–100 mm day⁻¹ to emphasize lower precipitation rates. The plot shows that MERRA-2 has higher counts compared to GPCP at very low rates (<10 mm day⁻¹), but at rates from 10 to 40 mm day⁻¹, MERRA-2 has lower counts for both the 1996–2008 and the 2009–18 periods. Both MERRA-2 and GPCP indicate a reduction in precipitation events at most rates for the period 2009–18. Above 40 mm day⁻¹ the densities are very infrequent (low probability). While this type of representation is useful in indicating some fundamental differences between the two time periods, it does not help us to characterize the time(s) at which the dataset changed, or how persistent the change was.

A more informative technique combines the time series of precipitation histograms into a single plot, and like a traditional PDF, uses every data point in space and time. A new array was produced with time on the x axis and the histogram bins on the y axis, with bin density indicated by color shading (Fig. 3). Figure 3 shows the GPCP daily time series of histograms, which shows the structures of the variations in counts that can extend for weeks or even months. For each day in the data record, a histogram was calculated using latitudinally area-weighted counts of precipitation, which is necessary to account for changes in surface area on a constant latitude–longitude grid. Because of the large range in daily precipitation intensity, the logarithm of the density was used (Fig. 3). The white color indicates exact zeroes, for which logarithm is undefined. There is a clear shift in the tail of the distribution (i.e., near the top of the plot) starting around 2009.

Fig. 2. Probability density of MERRA-2 and GPCP daily ocean-only (20°N to 20°S) precipitation rates for two time periods. Blue line is MERRA-2 (1996–2008), orange line is MERRA-2 (2009–18), green is GPCP (1996–2008) and red is GPCP (2009–18). The range was limited to 100 mm day⁻¹ to enhance the detail at the lower precipitation rates.

Fig. 3. GPCP 1DD ocean-only time series (1996–2018) of daily precipitation histograms for the latitude band 20°N–20°S. The colors represent the density of the daily histogram. Red indicates a high occurrence of the daily space–time counts for each day. Blue indicates very low to near-zero counts. White indicates zero counts.

To tease out detailed events or periods that more clearly depict the increase or decrease of precipitation intensity with time, we calculate the corresponding daily anomaly from the annual cycle (Fig. 4). We do this by comparing each day with the mean of that day for all years. The same PDF/anomaly display can also be calculated for monthly data. This is a novel metric as it is the anomaly of the histogram, not of precipitation itself, which is a more traditional approach. To enhance the signal, we use a Savitzky–Golay (Savitzky and Golay 1964) filter

with a polynomial order of 3. We examined filter lengths of 5, 35, and 75 days (not shown), and chose 35 days to keep subseasonal detail while suppressing short-interval fluctuations.

As depicted in Fig. 4, a variety of transient anomalies occur, but as well, a persistent change occurs starting around 2009. Compared to the earlier part of the record, from 2009 to 2018, there appears to show more low precipitation amounts and fewer high precipitation amounts.

This persistent change is coincident with a switch in the Defense Meteorological Satellite Program (DMSP) sensor used in the daily GPCP record, which took place at the start of 2009, from the Special Sensor Microwave Imager (SSM/I) sensor series to the Special Sensor Microwave Imager/Sounder (SSMIS) sensor series. It was previously discovered that this change introduces a 0.04 mm day^{-1} decrease in estimated mean global precipitation (Adler et al. 2018), due to known differences between SSM/I and SSMIS in footprint sizes and navigation, sensor calibration, and channel frequency selections. The new HATS technique provides a more nuanced regional understanding of this subsequent change to the global precipitation distribution.

Figure 5 shows the time series of daily histograms using the MERRA-2 daily precipitation record. The lack of the white color (compared to Fig. 3) suggests that MERRA-2 produces persistent drizzle, which was observed in the previous MERRA (Rienecker et al. 2011) and is still present, even at native resolution (not shown). Figure 6 shows the daily anomaly of the

Fig. 4. GPCP 1DD histogram ocean-only daily time series anomaly (from 1996 to 2018) of daily precipitation for the latitude band 20°N – 20°S . Red (blue) indicates a positive (negative) deviation from the mean annual daily values for the selected time period. The color bar is linear from -1 to $+1$ and logarithmic beyond. The data were processed with a Savitzky–Golay filter (35 day window, order 3).

Fig. 5. As in Fig. 3, but for MERRA-2. Note that there are no exactly zero counts because of the persistent very light drizzle, a characteristic of several reanalyses.

time series for MERRA-2, which has similar characteristics to GPCP but shows an even larger increase in light precipitation counts starting in 2008–09. The MERRA-2 development team continues investigating this shift that the HATS technique illustrates so clearly.

Summary

We have introduced a simple analysis technique, referred to as HATS, to examine the time series of precipitation anomaly histograms that identifies transient and persistent changes in Earth science data records, as shown using precipitation data from GPCP 1DD and MERRA-2. The analysis was conducted over the ocean in the latitude band 20°N–20°S, but can be equally applied to other regions, such as North America or eastern Asia. These regions could be examined to determine similar transient and persistent changes in the distribution of the precipitation. In addition to the apparent sensor-related precipitation histograms changes revealed in this paper, application at the regional scale should reveal interannual variations in the histograms due to El Niño–Southern Oscillation (ENSO) events, the Pacific decadal oscillation (PDO), and other large-scale climatic fluctuations. We plan to apply this technique to look at other variables, time ranges, and spatial domains, and we would expect similarly useful results to emerge.

Acknowledgments. The authors would like to acknowledge support for GLP, LEC, JH, from the NASA High-End Computing (HEC) Program through the NASA Center for Climate Simulation (NCCS) at Goddard Space Flight Center; also GLP, the Goddard Earth Sciences, Technology, and Research cooperative agreement between NASA and the Universities Space Research Association; for GJH and DTB, the GPM mission, WBS 573945.04.80.01.01; for MGB the Global Modeling and Assimilation Office, NASA Goddard Space Flight Center. We also thank the reviewers for their valuable comments.

For further reading

- Adler, R., and Coauthors, 2018: The Global Precipitation Climatology Project (GPCP) monthly analysis (new version 2.3) and a review of 2017 global precipitation. *Atmosphere*, **9**, 138, <https://doi.org/10.3390/atmos9040138>.
- Bosilovich, M. G., F. R. Robertson, L. Takacs, A. Molod, and D. Mocko, 2017: Atmospheric water balance and variability in the MERRA-2 reanalysis. *J. Climate*, **30**, 1177–1196, <https://doi.org/10.1175/JCLI-D-16-0338.1>.
- Gelaro, R., and Coauthors, 2017: The Modern-Era Retrospective Analysis for Research and Applications, version 2 (MERRA-2). *J. Climate*, **30**, 5419–5454, <https://doi.org/10.1175/JCLI-D-16-0758.1>.
- Huffman, G. J., R. F. Adler, M. M. Morrissey, D. T. Bolvin, S. Curtis, R. Joyce, B. McGavock, and J. Susskind, 2001: Global precipitation at one-degree daily resolution from multisatellite observations. *J. Hydrometeorol.*, **2**, 36–50, [https://doi.org/10.1175/1525-7541\(2001\)002<0036:GPAODD>2.0.CO;2](https://doi.org/10.1175/1525-7541(2001)002<0036:GPAODD>2.0.CO;2).
- Potter, G. L., L. Carriere, J. Hertz, M. Bosilovich, D. Duffy, T. Lee, and D. N. Williams, 2018: Enabling reanalysis research using the Collaborative Reanalysis Technical Environment (CREATE). *Bull. Amer. Meteor. Soc.*, **99**, 677–687, <https://doi.org/10.1175/BAMS-D-17-0174.1>.
- Rienecker, M. M., and Coauthors, 2011: MERRA: NASA's Modern-Era Retrospective Analysis for Research and Applications. *J. Climate*, **24**, 3624–3648, <https://doi.org/10.1175/JCLI-D-11-00015.1>.
- Robertson, F. R., M. G. Bosilovich, J. Chen, and T. L. Miller, 2011: The effect of satellite observing system changes on MERRA water and energy fluxes. *J. Climate*, **24**, 5197–5217, <https://doi.org/10.1175/2011JCLI4227.1>.
- Savitzky, A., and M. J. E. Golay, 1964: Smoothing and differentiation of data by simplified least squares procedures. *Anal. Chem.*, **36**, 1627–1639, <https://doi.org/10.1021/ac60214a047>.
- Teixeira, J., D. Waliser, R. Ferraro, P. Gleckler, T. Lee, and G. Potter, 2014: Satellite observations for CMIP5: The genesis of Obs4MIPs. *Bull. Amer. Meteor. Soc.*, **95**, 1329–1334, <https://doi.org/10.1175/BAMS-D-12-00204.1>.
- Wehner, M. F., and Coauthors, 2014: The effect of horizontal resolution on simulation quality in the Community Atmospheric Model, CAM5.1. *J. Adv. Model. Earth Syst.*, **6**, 980–997, <https://doi.org/10.1002/2013MS000276>.

Fig. 6. As in Fig. 4, but for MERRA-2.