Supplemental Material

Journal of Hydrometeorology

"It's Raining Bits": Patterns in Directional Precipitation Persistence across the United States

https://doi.org/10.1175/JHM-D-20-0134.1

Supplementary Information for

It's Raining Bits: Patterns in directional precipitation persistence across the U.S.

Allison E. Goodwell

allison.goodwell@ucdenver.edu

This PDF file includes:

Supplementary text
Figs. S1 to S2
Tables S1 to S3
References for SI reference citations

Supporting Information Text

Information Theory Methods

A. Dominant directions of predictability. For any given information measure, MI or TE, computed between a central cell and its neighboring lagged histories, we determine the measures of TE_{vect} and MI_{vect} by a vector resolution as follows:

$$TE_x = \sum_i \cos\theta_i \quad [1]$$

$$TE_y = \sum_i \sin\theta_i \quad [2]$$

where i corresponds to each of the 8 neighboring directions ($i =$ E, NE, N, NW, W, SW, S, SE), and θ is the angle associated with each direction (degrees from east). The magnitude of the resolved information value is then:

$$TE_{vect} = \sqrt{TE_y^2 + TE_x^2}, \quad [3]$$

which corresponds to an angular direction θ_{TEvect} as follows:

$$\theta_{TEvect} = \arctan\frac{TE_y}{TE_x}. \quad [4]$$

The same equations are used for computations of MI_{vect} and θ_{MIvect}. As mentioned in the paper, TE_{max} and MI_{max} and their associated angular directions are simply computed as the maximum of the eight individual values.

B. Information Partitioning. Redundant information is determined using a re-scaled redundancy measure (1) as follows:

$$R = R_{min} + \frac{I(X_{t-1,d1}; X_{t-1,d2})}{\min[H(X_{t-1,d1}), H(X_{t-1,d2})]}(R_{max} - R_{min}). \quad [5]$$

Here, R_{min} and R_{max} are minimum and maximum bounds defined by information theory as follows:

$$R_{\min} = \max(0, -I(X_{t-1,d1}; X_{t-1,d2}; X_{t,c})) \quad [6]$$

$$R_{max} = \min[I(X_{t-1,d1}; X_{t,c}), I(X_{t-1,d2}; X_{t,c})] \quad [7]$$

In this formulation, independent sources (where $I(X_{t-1,d1}; X_{t-1,d2}) = 0$) are minimally redundant, while strongly related sources are maximally redundant. For example, if precipitation in eastern and western directions were completely synchronized, their information contributions to current rainfall at location c would be redundant. On the other hand, if they were somewhat or completely independent, their information contributions would include synergistic and unique components. After Equation (5) is solved for R, U_{d1} and U_{d2} can be solved via the relationships $I(X_{t-1,d1}; X_{t,c}) = R + U_{d1}$ and $I(X_{t-1,d2}; X_{t,c}) = R + U_{d2}$, respectively. Finally, synergy, S can be computed as $S = I(X_{t-1,d1}, X_{t-1,d2}; X_{t,c} - U_{d1} - U_{d2} - R$.

C. Thresholds. For each grid cell and each season, we tested different threshold values between 0.3 and 3 mm to separate "wet" from "dry" states. For each timeseries, we computed the lag-1 mutual information, $I(X_{t-1,c}; X_{t,c})$, and chose the threshold that corresponds to the maximum mutual information. In other words, we define smaller precipitation values as "dry" as long as otherwise setting them as "wet" does not further reduce uncertainty in the current state. For most grid cells, this threshold was between 0.3 and 1 mm, indicating that even the lower magnitudes of precipitation are informative (Figure S1). However, for some grid cells, the threshold is higher. This mainly occurs in regional clusters in winter (DJF), spring (MAM), and summer (JJA).

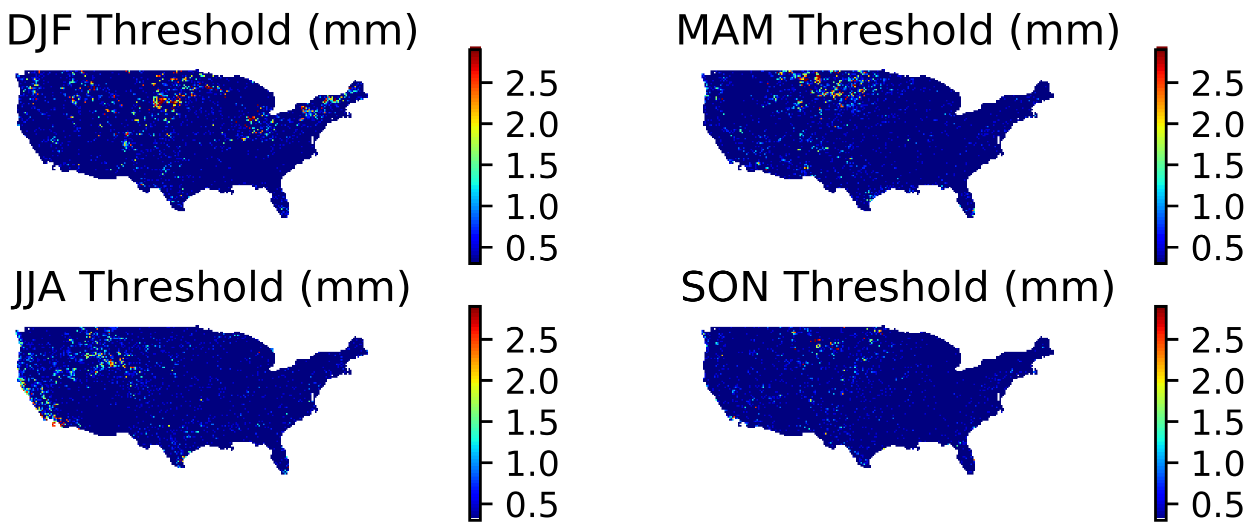

Fig. S1. Threshold precipitation magnitudes

Example Case

In each of the synthetic example cases discussed in the main text, 5000 data points are generated for three neighboring grid cells from east to west, W, C, and E. For every case, the W(t-1) state is generated as a Markov chain with transition probabilities $p(W(t) = 1|(W(t-1) = 1) = 0.9$ and $p(W(t) = 0|(W(t-1) = 1) = 0.1$. This results in precipitation time-series in which approximately, $p(W(t) = 1) = p(W(t) = 0) = 0.5$, and $H(W(t)) = 1$ *bit*. These synthetic cases were developed in Excel (see formulas in Table S1 for first several values of each case). Additionally in all cases, the uncertainty of the central grid current state, $C(t)$, is fully reduced by either or both of the 1-lagged states in either direction, i.e. $I(E(t-1), W(t-1); C(t) = H(C(t))$. Table S2 shows all information measures for the four case, in addition to the correlation between precipitation in both directions (east and west).

Case A: moving rainstorm		
West(t-1)	Center(t)	East(t-1)
=ROUND(RAND(),0)	=A3	0
=IF(A3=1,ROUND(RAND()+0.4,0),ROUND(RAND()-0.4,0))	=A4	0
=IF(A4=1,ROUND(RAND()+0.4,0),ROUND(RAND()-0.4,0))	=A5	=B3
=IF(A5=1,ROUND(RAND()+0.4,0),ROUND(RAND()-0.4,0))	=A6	=B4

CASE B: moving rainstorm from west to center, east random		
West(t-1)	Center(t)	East(t-1)
=ROUND(RAND(),0)	=A10	=ROUND(RAND(),0)
=IF(A10=1,ROUND(RAND()+0.4,0),ROUND(RAND()-0.4,0))	=A11	=ROUND(RAND(),0)
=IF(A11=1,ROUND(RAND()+0.4,0),ROUND(RAND()-0.4,0))	=A12	=ROUND(RAND(),0)
=IF(A12=1,ROUND(RAND()+0.4,0),ROUND(RAND()-0.4,0))	=A13	=ROUND(RAND(),0)

Case C: XOR - center rains if east and west in different states		
West(t-1)	Center(t)	East(t-1)
=ROUND(RAND(),0)	=IF(A17=C17,0,1)	=ROUND(RAND(),0)
=IF(A17=1,ROUND(RAND()+0.4,0),ROUND(RAND()-0.4,0))	=IF(A18=C18,0,1)	=IF(C17=1,ROUND(RAND()+0.4,0),ROUND(RAND()-0.4,0))
=IF(A18=1,ROUND(RAND()+0.4,0),ROUND(RAND()-0.4,0))	=IF(A19=C19,0,1)	=IF(C18=1,ROUND(RAND()+0.4,0),ROUND(RAND()-0.4,0))
=IF(A19=1,ROUND(RAND()+0.4,0),ROUND(RAND()-0.4,0))	=IF(A20=C20,0,1)	=IF(C19=1,ROUND(RAND()+0.4,0),ROUND(RAND()-0.4,0))

Case D: AND - center rains if east and west both raining		
West(t-1)	Center(t)	East(t-1)
=ROUND(RAND(),0)	=IF(A24+C24=2,1,0)	=ROUND(RAND(),0)
=IF(A24=1,ROUND(RAND()+0.4,0),ROUND(RAND()-0.4,0))	=IF(A25+C25=2,1,0)	=IF(A24=1,ROUND(RAND()+0.4,0),ROUND(RAND()-0.4,0))
=IF(A25=1,ROUND(RAND()+0.4,0),ROUND(RAND()-0.4,0))	=IF(A26+C26=2,1,0)	=IF(A25=1,ROUND(RAND()+0.4,0),ROUND(RAND()-0.4,0))
=IF(A26=1,ROUND(RAND()+0.4,0),ROUND(RAND()-0.4,0))	=IF(A27+C27=2,1,0)	=IF(A26=1,ROUND(RAND()+0.4,0),ROUND(RAND()-0.4,0))

Table S1. Formulas illustrating first several values of each of the 4 synthetic cases described in the main paper. Excel file with full synthetic examples can be found on Github site along with codes for precipitation analsysis.

measure	units	Case A	Case B	Case C	Case D
$H(C)$	*bits*	0.999	1.000	1.000	0.961
$I(E, W; C)$	*bits*	0.999	1.000	1.000	0.961
$I(E; C)$	*bits*	0.293	0.000	0.001	0.619
$I(W; C)$	*bits*	0.999	1.000	0.004	0.646
$I(E; W)$	*bits*	0.293	0.000	0.000	0.326
$Corr(E, W)$	n/a	0.655	-0.002	-0.020	0.633
$I_s : I(E; W)/H(C)$	*bits/bit*	0.293	0.000	0.000	0.339
R_{min}	*bits*	0.293	0.000	0.000	0.305
R_{max}	*bits*	0.293	0.000	0.001	0.619
R_{EW}	*bits*	0.29	0.00	0.00	0.41
U_E	*bits*	0.00	0.00	0.00	0.21
U_W	*bits*	0.71	1.00	0.00	0.23
S_{EW}	*bits*	0.00	0.00	0.99	0.11

Table S2. Full set of information theory measures computed for example cases described in main paper. In each case, precipitation at a central target cell is informed by precipitation in neighboring grid cells to the east and west.

Climate Indices

Monthly values for each of the 10 climate indices listed in Table S3 are averaged to seasonal measures to match the time-scale of the information-based measures of precipitation predictability. Before comparing with precipitation, we look for correspondence between the indices themselves, in the form of correlation between them over the study period (Figure S2). In general, there are more and stronger connections during the winter and spring (Figure S2) relative to summer and fall. As expected, the SOI, NINO3, and ONI indices are always strongly related as measures of the ENSO phenomenon, so we only show correlations between NINO3 and the other indices to reduce redundant linkages. The NP and PNA indices are negatively correlated for all seasons except the summer, and the AO and NAO indices are correlated in the winter and summer. Besides these persistent relationships between climate indices, there are several others that appear only in certain seasons. While these connections indicate some redundancies in climate indices, their variation in strengths and seasonalities could relate to precipitation patterns in different ways. The codes and datasets on the Github repository associated with this paper include the full analysis of correlations between all climate indices and precipitation information theory measures. The indices with the most frequent correlations with information measures, as presented in the main paper, were the AMO, PDO, EPNP, and NAO.

Climate Index (acronym, source)	Description
Nino 3.4 (NINO3, (2)	average of sea surface temperature anomalies in the Nino-3.4 region (5S to 5N; 170W to 120W), at a 5-month running mean.
Oceanic Nino Index (ONI, (3))	average of sea surface temperature anomalies in the Nino-3.4 region (5S to 5N; 170W to 120W), at a 3-month running mean
Southern Oscillation Index (SOI, (4))	sea level pressure differences between Tahiti and Darwin, Australia, measure of El Nino episodes
Pacific North American Index (PNA, (5))	anomalies in geopotential height fields over the United States, influential climate pattern in northern hemisphere mid-latitudes
North Pacific Pattern (NP, (6, 7))	geopotential height anomalies, prominent between March and July.
Eastern Pacific - North Pacific (EPNP, (8))	geopotential height anomalies
Northern Atlantic Oscillation (NAO, (9))	surface sea-level pressure difference between the Suptropical High and the Subpolar Low, associated with North Atlantic jet stream
Atlantic Multidecadal Oscillation (AMO, (10))	sea surface temperature anomalies in the North Atlantic, associated with Atlantic hurricane patterns and summer climate in North America.
Arctic Oscillation (AO, (11))	defined based on circulating winds in the Arctic around 55N latitude, largest variability during the cold season
Pacific Decadal Oscillation (PDO, (12))	sea surface temperature anomalies in the Northern Pacific, associated with climate in western North America.

Table S3. Climate indices for comparisons with information measures. Bold highlighted indices are those discussed in main paper.

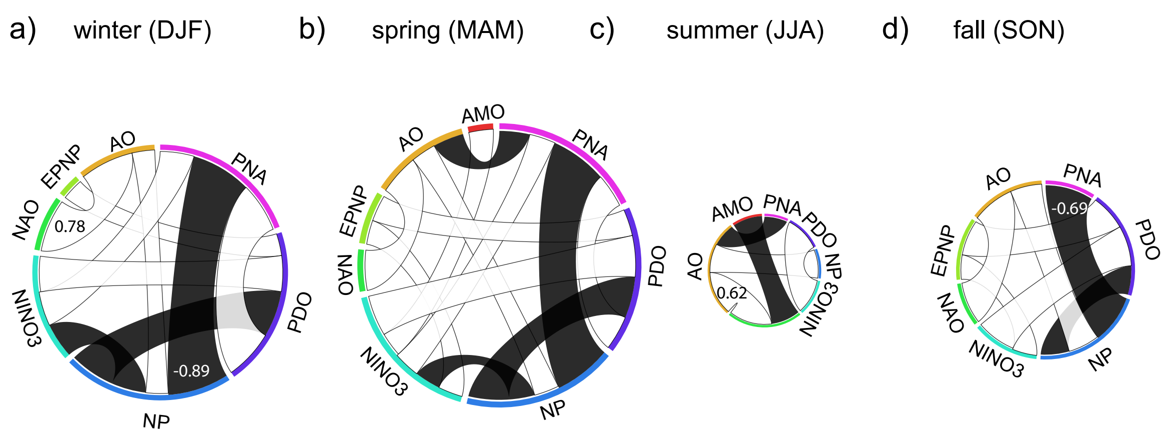

Fig. S2. Correlations between seasonal climate indices over 70 year period for *(a)* winter (DJF), *(b)* spring (MAM), *(c)* summer (JJA), and *(d)* fall (SON). Monthly climate indices were averaged to obtain seasonal values. Circle diameters are scaled to match correlation totals for all pairs, black links indicate negatively correlated indices, and white links indicate positively correlated indices.

References

1. Goodwell AE, Kumar P (2017) Temporal Information Partitioning: Characterizing synergy, uniqueness, and redundancy in interacting environmental variables. *Water Resources Research* 53(7):5920–5942.
2. NOAA (2019) Nino-3.4: East central tropical pacific sst (https://www.cpc.ncep.noaa.gov/data/indices/). Accessed: 2019-05-01.
3. NOAA (2019) Oceanic niño index (https://origin.cpc.ncep.noaa.gov/products/analysis_monitoring/ensostuff/ONI_v5.php). Accessed: 2019-05-01.
4. NOAA (2019) Southern oscillation index (https://www.cpc.ncep.noaa.gov/data/indices/). Accessed: 2019-05-01.
5. NOAA (2019) Pacific north american monthly index (https://www.cpc.ncep.noaa.gov/data/teledoc/pna.shtml). Accessed: 2019-05-01.
6. NOAA (2019) North pacific pattern (https://www.esrl.noaa.gov/psd/data/climateindices/list/). Accessed: 2019-05-01.
7. Trenberth KE, Hurrell JW (1994) Decadal atmosphere-ocean variations in the pacific. *Climate Dynamics* 9(6):303–319.
8. NOAA (2019) East pacific/north pacific oscillation (https://www.cpc.ncep.noaa.gov/data/teledoc/ep.shtml). Accessed: 2019-05-01.
9. NOAA (2019) North atlantic oscillation (https://www.cpc.ncep.noaa.gov/data/teledoc/nao.shtml). Accessed: 2019-05-01.
10. NOAA (2019) Atlantic multidecadal oscillation (https://www.esrl.noaa.gov/psd/data/timeseries/AMO/). Accessed: 2019-05-01.
11. NOAA (2019) Arctic oscillation (https://www.cpc.ncep.noaa.gov/products/precip/CWlink/daily_ao_index/ao.shtml). Accessed: 2019-05-01.
12. NOAA (2019) Pacific decadal oscillation (http://research.jisao.washington.edu/pdo/PDO.latest). Accessed: 2019-05-01.