AGUPUBLICATIONS

Supporting Information for

Evaluating the effects of precipitation and evapotranspiration on soil moisture variability

Xuan Xi[1], Pierre Gentine[2], Qianlai Zhuang[1,3], Seungbum Kim[4]

[1]Department of Earth, Atmospheric, and Planetary Sciences, Purdue University, West Lafayette, IN, 47907

[2]Department of Earth of Environmental Engineering, Columbia University, New York, NY, 10027

[3]Department of Agronomy, Purdue University, West Lafayette, IN 47907

[4]NASA Jet Propulsion Laboratory, Pasadena, CA 91109

Contents of this file

Figures S1 to S5
Tables S1 to S6
Text S1 to S2

Additional Supporting Information

Captions for Figures S1 to S5
Captions for Table S1 to S6

Introduction

This supplement includes additional figures, tables, and texts to provide more information about the contents shown in the main text.

Specifically, Figure S1 gives a detailed version of Figure 1 shown in the main text. Figure S2 gives diagrams of two conceptual linear time-invariant (LTI) systems supporting Figure 2 shown in the main text. Figure S3 shows the global mean surface soil moisture content based on the SMAP product (Entekhabi et al., 2010) after spatiotemporal normalization. Figure S4 shows the global distribution of the coefficient of variation (CV) for H_{SEP_n} and H_{EEP_n} across CMIP5 models in the three frequency bands. Figure S5 shows the display of the regions that have potential errors in the "uncertainty analysis" discussed in the main text.

Table S1 provides specific information on the models from CMIP5 (Taylor et al., 2012) used in this study. Table S2 and S3 give additional specific information on the Fourier transform provided in Text S1. Table S4 gives the observational value of H_{SEP_n} and H_{EEP_n} defined in the main text in the three frequency bands. Table S5 gives quantitative differences of H_{SEP_n} and H_{EEP_n} between CMIP5 models and observation-based data. Table S6 gives the quantitative coefficient of variation (CV) of H_{SEP_n} and H_{EEP_n} across the models within CMIP5.

Text S1 provides more detailed information on Fourier transform, including an overview, descriptions of Discrete Fourier Transform (DFT), Fast Fourier Transform (FFT), and spectrum analysis. Text S2 provides the background of the color of noise and its application based on the spectral slope.
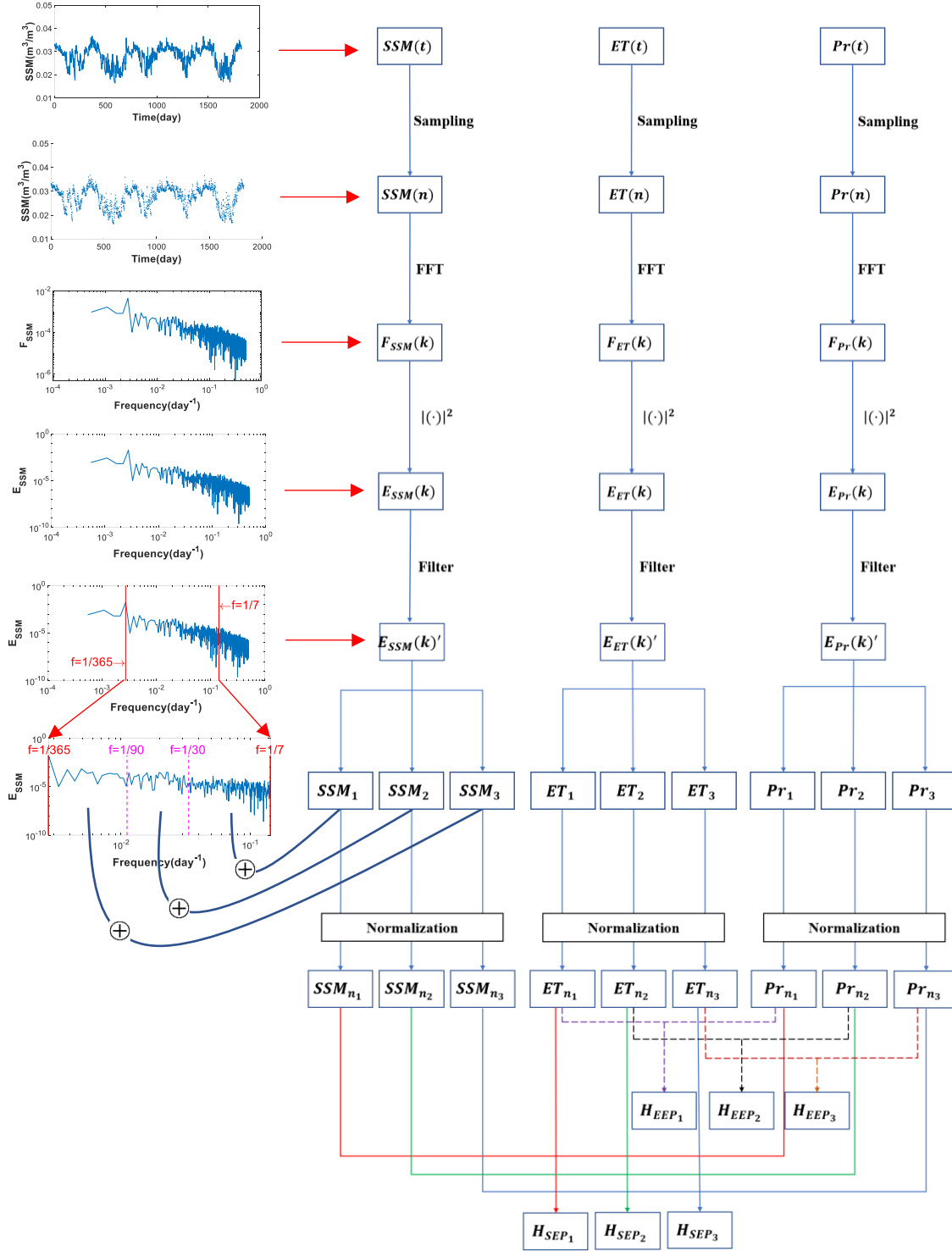
Figure S1. Processes to get the normalized variability of SSM (SSM_{n_1}, SSM_{n_2}, and SSM_{n_3}), ET (ET_{n_1}, ET_{n_2}, and ET_{n_3}), and Pr (Pr_{n_1}, Pr_{n_2}, and Pr_{n_3}), and further the two ratios to analyze the effects of ET and Pr on SSM (i.e., H_{SEP}, H_{EEP}) from the original time series of SSM, ET, and Pr (i.e., $SSM(t)$, $ET(t)$, $Pr(t)$). The left column shows six plots obtained by each

corresponding step on the right (take SSM as an example). This example is based on the data located at (51.57°N, 1.25°E) of the "GFDL-ESM2M" model within CMIP5 from January 1, 2001, to December 31, 2005.
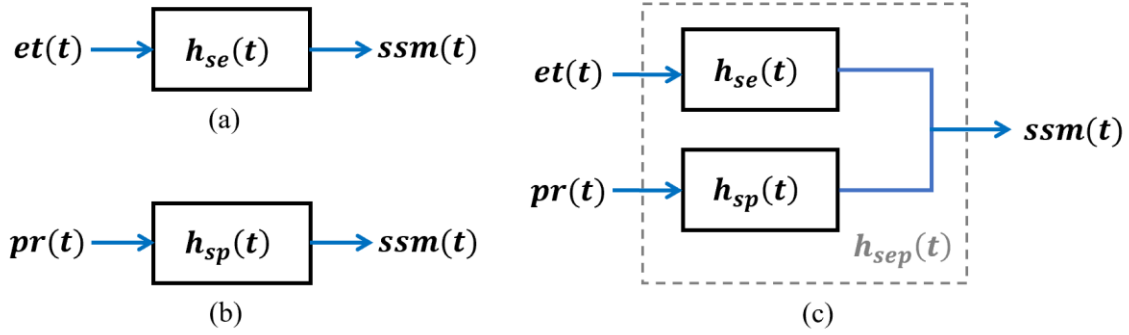

Figure S2. Conceptual diagrams of the assumed "ET-SSM" LTI system (a) and "Pr-SSM" LTI system (b) and a combination of them (c). The excitations (i.e., input) of the system (a) and (b) are $et(t)$ and $pr(t)$, respectively. The responses (i.e., output) of the two systems are all $ssm(t)$. The transfer functions of the system (a) and (b) are $h_{se}(t)$ and $h_{sp}(t)$, respectively. For figure(c), the inputs are $et(t)$ and $pr(t)$ together, and the output is $ssm(t)$. The grey dashed box includes the two transfer functions of system (a) and (b) and represented by an identical transfer function $h_{sep}(t)$.
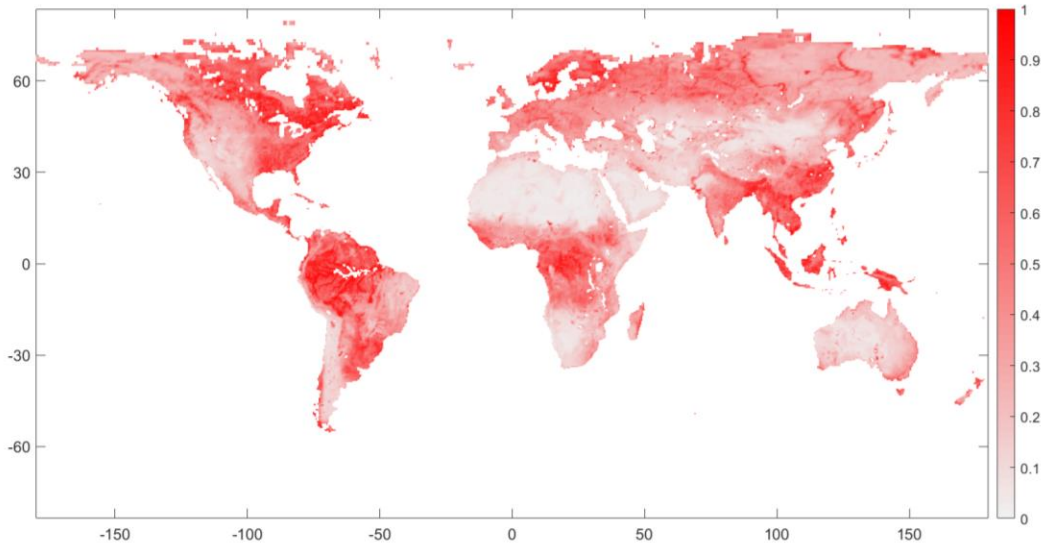

Figure S3. The observational mean SSM (surface soil moisture) after spatiotemporal normalization ($\overline{SSM_n}$). More than five years' data from the SMAP Level-3 product, spanning 1

April 2015 - 31 December 2020, are used. We first use original data to get the daily average SSM ($\overline{SSM}$) for each pixel and then normalize them between zero and one based on the min-max normalization as: $\overline{SSM}_n = (\overline{SSM} - \overline{SSM}_{min})/(\overline{SSM}_{max} - \overline{SSM}_{min})$.
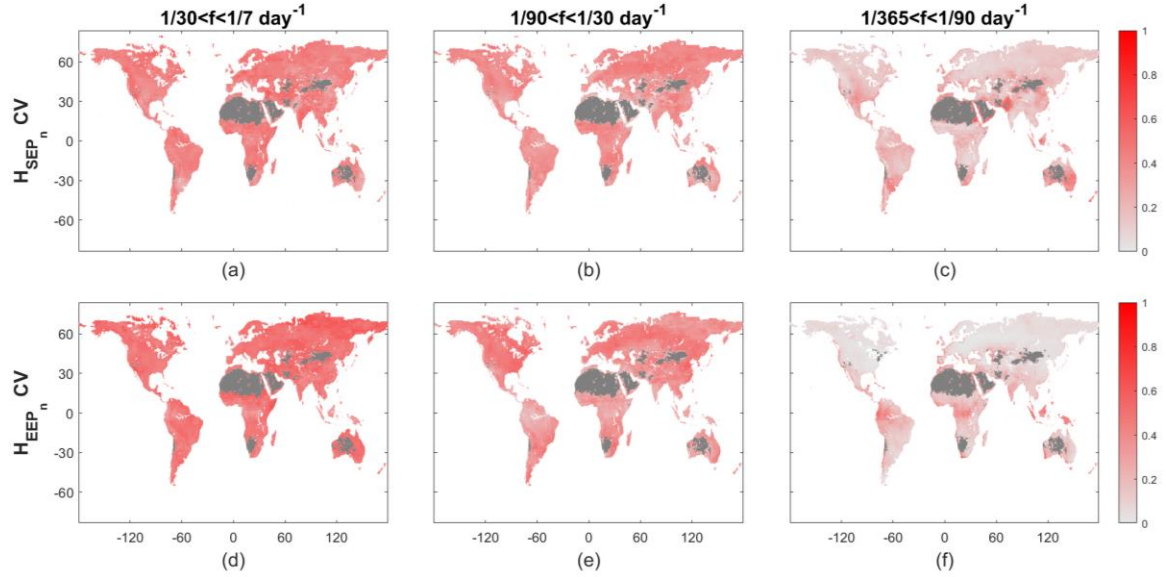

Figure S4. The coefficient of variation (CV) of H_{SEP_n} (Figure a-c) and H_{EEP_n} (Figure d-f) across all models in the three frequency bands. Similar to the CV of SSM_n (Figure 5d – 5f), for each model, the CV of H_{SEP_n} and H_{EEP_n} are calculated as their standard deviation divided by their mean values for each frequency band, and we then normalize CV values between zero and one across the three frequency bands. The dark grey parts are regions with $\overline{SSM}_n$ less than 0.1.
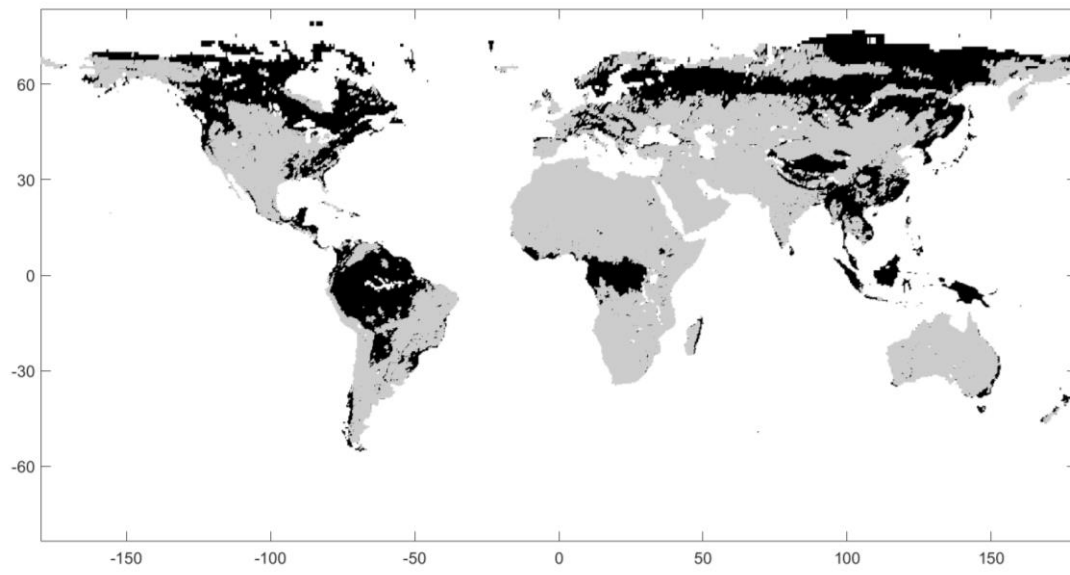
Figure S5. Display of the regions where have potential errors in the "uncertainty analysis". Grey parts are land surface coverage analyzed in this study. Black parts are regions where being masked due to potential uncertainties.

Model version	Center	Forcing	Spatial Resolution
BCC-CSM1.1	Beijing Climate Center, China Meteorological Administration	Nat Ant GHG SD Oz Sl Vl SS Ds BC OC	128*64
BNU-ESM	College of Global Change and Earth System Science, Beijing Normal University	Nat, Ant	128*64
CanESM2	Canadian Centre for Climate Modeling and Analysis	GHG, Oz, SA, BC, OC, LU, Sl,Vl (GHG includes CO_2, CH_4, N_2O, CFC11, effective CFC12)	128*64
CNRM-CM5	Centre National de Recherches Meteorologiques / Centre Europeen de Recherche et Formation Avancees en Calcul Scientifique (CNRM/CERFACS)	GHG, SA, Sl, Vl, BC, OC	256*128
CSIRO-Mk3.6	Commonwealth Scientific and Industrial Research Organization/Queensland Climate Change Centre of Excellence (CSIRO-QCCCE)	Ant, Nat (all forcings)	192*96
GFDL-CM3	Geophysical Fluid Dynamics Laboratory	GHG, SA, Oz, LU, Sl, Vl, SS, BC, MD, OC (GHG includes CO_2, CH_4, N_2O, CFC11, CFC12, HCFC22, CFC113)	144*90
GFDL-ESM2G	Geophysical Fluid Dynamics Laboratory	GHG, SD, Oz, LU, Sl, Vl, SS, BC, MD, OC (GHG includes CO_2, CH_4, N_2O, CFC11, CFC12, HCFC22, CFC113)	144*90
GFDL-ESM2M	Geophysical Fluid Dynamics Laboratory	GHG, SD, Oz, LU, Sl, Vl, SS, BC, MD, OC (GHG includes CO_2, CH_4, N_2O, CFC11, CFC12, HCFC22, CFC113)	144*90
MIROC5	Atmosphere and Ocean Research Institute (The University of Tokyo), National Institute for Environmental Studies, and Japan Agency for Marine-Earth Science and Technology	GHG, SA, Oz, LU, Sl, Vl, SS, Ds, BC, MD, OC (GHG includes CO_2, N_2O, methane, and fluorocarbons; Oz includes OH and H_2O_2; LU excludes change in lake fraction)	256*128

MIROC-ESM	Atmosphere and Ocean Research Institute (The University of Tokyo), National Institute for Environmental Studies, and Japan Agency for Marine-Earth Science and Technology	GHG, SA, Oz, LU, Sl, Vl, MD, BC, OC	128*64
MIROC-ESM-CHEM	Atmosphere and Ocean Research Institute (The University of Tokyo), National Institute for Environmental Studies, and Japan Agency for Marine-Earth Science and Technology	GHG, SA, Oz, LU, Sl, Vl, MD, BC, OC (Ozone is predicted)	128*64
MRI-CGCM3	Meteorological Research Institute	GHG, SA, Oz, LU, Sl, Vl, BC, OC (GHG includes CO2, CH4, N2O, CFC-11, CFC-12, and HCFC-22)	320*160
MRI-ESM1	Meteorological Research Institute	GHG, SA, Oz, LU, Sl, Vl, BC, OC (GHG includes CO2, CH4, N2O, CFC-11, CFC-12, and HCFC-22)	320*160
NorESM1-M	Norwegian Climate Centre (NorClim)	GHG, SA, Oz, Sl, Vl, BC, OC	144*96

Table S1. Fourteen CMIP5 models used in this research and some of their specific information. The model simulations have the same temporal coverage from 01/01/1950 to 12/31/2005.

Form of Fourier Transform	Time Domain	Frequency Domain
Fourier Transform (FT)	aperiodic, continuous	aperiodic, continuous
Fourier Series (FS)	periodic, continuous	aperiodic, discrete
Discrete Time Fourier Transform (DTFT)	aperiodic, discrete	periodic, continuous
Discrete Fourier Transform (DFT)	periodic, discrete	periodic, discrete

Table S2. Four different forms of Fourier transform.

Algorithm	Complex multiplication (#)	Complex addition (#)
DFT	$\dfrac{N}{2}\log_2 N$	$N\log_2 N$
FFT	N^2	$N(N+1)$

Table S3. Computation complexity comparison between DFT and FFT.

Frequency band (day^{-1})	1/7 ~ 1/30	1/30 ~ 1/90	1/90 ~ 1/365
H_{SEP_n}	0.4127 (0.3733)	0.8708 (0.7883)	0.5129 (0.5662)
H_{EEP_n}	0.2064 (0.1966)	0.3393 (0.3601)	0.7586 (0.7622)

Table S4. Observational-based H_{SEP_n} and H_{EEP_n} in the three frequency bands. H_{SEP_n} and H_{EEP_n} here are original values without normalization across the three frequency bands. The numbers in brackets are corresponding values masked by regions with potential uncertainties (see main text).

Significance	100% significance test			75% significance test		
Frequency band (day^{-1})	1/7 ~ 1/30	1/30 ~ 1/90	1/90 ~ 1/365	1/7 ~ 1/30	1/30 ~ 1/90	1/90 ~ 1/365
H_{SEP_n}	-0.3365 (-0.2816)	-0.6792 (-0.5898)	0.4797 (0.4168)	-0.2871 (-0.2366)	-0.4492 (-0.3797)	0.4011 (0.3402)
H_{EEP_n}	0.1259 (0.1471)	0.0677 (0.0532)	-0.0872 (-0.0919)	0.0770 (0.0899)	0.0515 (0.0449)	-0.0597 (-0.0628)

Table S5. Multimodel average differences of H_{SEP_n} and H_{EEP_n} within CMIP5. The numbers in brackets are corresponding values masked by regions with potential uncertainties (see main text).

Significance	100% significance test			75% significance test		
Frequency band (day^{-1})	1/7 ~ 1/30	1/30 ~ 1/90	1/90 ~ 1/365	1/7 ~ 1/30	1/30 ~ 1/90	1/90 ~ 1/365
H_{SEP_n}	0.5120 (0.4648)	0.4409 (0.4029)	0.2325 (0.2174)	0.5148 (0.4690)	0.4489 (0.4032)	0.2370 (0.2261)
H_{EEP_n}	0.3352 (0.3335)	0.2287 (0.2153)	0.0894 (0.0877)	0.3474 (0.3507)	0.2558 (0.2466)	0.0966 (0.0922)

Table S6. Coefficient of variation (CV) of H_{SEP_n} and H_{EEP_n} across the 14 CMIP5 models. Values here are original values without normalization across the three frequency bands. The numbers in brackets are corresponding values masked by regions with potential uncertainties (see main text).

Text S1. Fourier Transform

1. Overview

Fourier transform is a linear integral transform. The basic idea was first systematically put forward by French mathematician and physicist Joseph Fourier in 1822. The purpose of the Fourier transform is to establish a specific transformation relationship between the signal with time as the independent variable and the frequency spectrum function with frequency as the independent variable, that is, to realize the transformation from the time domain to the frequency domain. Considering various types of signals (periodic, aperiodic, continuous, discrete), there can be four different forms of Fourier transform. Their corresponding periodicity and continuity in the time domain and frequency domain are shown in Table S2.

Generally speaking, the Fourier transform is referred to the first form in Table S2, which can be expressed as:

$$X(f) = \int_{-\infty}^{\infty} x(t) e^{-j2\pi ft} \, dt \tag{1}$$

$$x(t) = \int_{-\infty}^{\infty} X(f) e^{j2\pi ft} \, dt \tag{2}$$

where $x(t)$ is the signal in the time domain, and $X(f)$ is the spectrum function of $x(t)$ in the frequency domain. $x(t)$ and $X(f)$ form a transform pair.

For the first three forms of Fourier transform in Table S2 (i.e., FT, FS, and DTFT), since there is always a variable that is continuous in either time or frequency domain, they are not suitable for the calculation by computer. Compared to the first three forms, DFT can be applied on the computer since its transform pairs are discrete in both time and frequency domains.

2. Discrete Fourier Transform (DFT)

Discrete Fourier Transform (DFT) is a discrete form of continuous Fourier transform in both time and frequency domains. DFT is aimed at a finite-length sequence, and its essence is to discretize the continuous Fourier transform of the sequence and transform the sampling of the signal in the time domain into the sampling of DTFT in the frequency domain. In this way, the discretization of the frequency domain results in a periodic time domain, so the Fourier series is limited to one cycle. The transformation pair in the form of DFT series can be expressed as:

$$X(kf_1) = \sum_{n=0}^{N-1} x(nT_s) e^{-j\frac{2\pi}{N}nk} \tag{3}$$

$$x(nT_s) = \frac{1}{N} \sum_{k=0}^{N-1} X(kf_1) e^{j\frac{2\pi}{N}nk} \tag{4}$$

where $X(kf_1)$ is the periodic discrete time function in the time domain, $x(nT_s)$ is the periodic discrete frequency function in the frequency domain. Here, the time interval T_s of the discrete time function and the repetition period f_s of the frequency function satisfy: $f_s = \frac{1}{T_s}$, and the

interval f_1 of the discrete frequency function and the period T_1 of the time function satisfy: $f_1 = \frac{1}{T_1}$. Besides, there are the following relationships in each cycle of the time domain and the frequency domain:

$$\frac{T_1}{T_s} = N \text{ or } \frac{f_s}{f_1} = N \tag{5}$$

that is there are N sampling points in each cycle.

   The discrete Fourier series is commonly used for periodic sequence analysis. Actually, the periodic sequence only has a finite number of meaningful sequence values, so the finite-length sequence $x(n)$ of length N can be regarded as a period of the periodic sequence of period N, and the DFT of a finite sequence can be calculated by the Fourier series of the periodic sequence. The transform pair of DFT of a finite sequence can be expressed as:

$$X(k) = \sum_{n=0}^{N-1} x(n) W_N^{nk}, 0 \le k \le N - 1 \tag{6}$$

$$x(n) = \frac{1}{N} \sum_{n=0}^{N-1} X(k) W_N^{-nk}, 0 \le k \le N - 1 \tag{7}$$

where $W_N = e^{-j\frac{2\pi}{N}}$.

3.   Fast Fourier Transform (FFT)

   Because DFT calculation is relatively cumbersome, DFT has not been widely used for a long time, until 1965, Curry and Atlas proposed a fast DFT algorithm (Cooley & Tukey, 1965). This method and a series fast of DFT algorithms later are collectively referred to as Fast Fourier Transform (FFT) (Cochran et al., 1967; Gentleman & Sande, 1966). There are two commonly used FFT methods, one is decimation-in-time (DIT), another one is decimation-in-frequency (DIF). FFT is not a new transformation but a fast algorithm to implement DFT.

   Recall the equation of DFT for N-point sequence, generally, $x(n)$ and W_N^{nk} are plural. Each calculation of an $X(k)$ value requires N complex multiplications and $(N-1)$ complex additions. Therefore, for an N-point sequence, DFT needs to do N^2 complex multiplications and $N(N-1)$ complex additions, which is a very large amount of computation.

   FFT utilizes the periodicity and symmetry of W_N^{nk} to decompose the DFT operation with a length of N points into a shorter sequence of DFT operations. Specifically, the periodicity of W_N^{nk} can be expressed as:

$$W_N^{nk} = W_N^{((nk))_N} \tag{8}$$

where $((nk))_N$ is the value for nk modulo of N, and the symmetry of W_N^{nk} can be expressed as:

$$W_N^{(nk+\frac{N}{2})} = -W_N^{nk} \tag{9}$$

The N-point DFT can be decomposed into two sets of $\frac{N}{2}$-point DFT, and then take the sum of them, which can be expressed as:

$$X(k) = \sum_{r=0}^{\frac{N}{2}-1} x(2r)W_{\frac{N}{2}}^{rk} + W_N^k \sum_{r=0}^{\frac{N}{2}-1} x(2r+1)W_{\frac{N}{2}}^{rk} \tag{10}$$

$$X\left(\frac{N}{2}+k\right) = \sum_{r=0}^{\frac{N}{2}-1} x(2r)W_{\frac{N}{2}}^{rk} - W_N^k \sum_{r=0}^{\frac{N}{2}-1} x(2r+1)W_{\frac{N}{2}}^{rk} \tag{11}$$

where $k = 0,1,\dots,\frac{N}{2}-1$, $2r$ represents even numbers, and $2r+1$ represents odd numbers.

Equation (10) and equation (11) give the value of the first $\frac{N}{2}$ points and the last $\frac{N}{2}$ points of $X(k)$, respectively.

By performing FFT, the computation complexity can be reduced from $O(n*n)$ to $O(n*logn)$. The computation complexity comparison between DFT and FFT is shown in Table S3.

4. Spectrum analysis

Spectrum (including amplitude spectrum and phase spectrum) describes signal characteristics in the frequency domain. Spectrum reflects the distribution of the amplitude and phase of the components contained in the signal with frequency. Spectrum analysis is the process of obtaining the frequency structure of the signal by calculating the amplitude and phase of the signal at each frequency (Kay & Marple, 1981).

For computer applications, an analog signal $x(t)$ is usually converted into a discrete-time signal $x(n)$ through sampling in the time domain, and then use DFT and FFT for spectrum analysis, which can be expressed as:

$$X(k) = DFT[x(n)] = \sum_{n=0}^{N-1} x(n)W_N^{kn} = X_R(k) + jX_l(k) \tag{12}$$

The amplitude and phase corresponding to each frequency value are as follows:

$$|X(k)| = \sqrt{X_R^2(k) + X_l^2(k)} \tag{13}$$

$$\theta(k) = arctan\frac{X_l(k)}{X_R(k)} \tag{14}$$

where $k = 0,1,\dots,N-1$, equation (13) and (14) are the amplitude and phase for k frequency, respectively.

Besides, the power spectrum can also be used to describe the signal, which indicates the energy of the signal varying with frequency in the frequency domain. The energy of the signal is based on its amplitude and can be expressed as:

$$E(k) = |X(k)|^2 = X_R^2(k) + X_l^2(k) \tag{15}$$

Text S2. Color of noise

Noise is a stochastic process. The power spectrum, which describes the variance as a sum of sinusoidal waves of different frequencies, is an important characteristic of environmental noise (Vasseur and Yodzis, 2004). There are many ways to characterize different noise sources. Noise distributed in the whole frequency domain and with the form that variance scales with frequency according to an inverse power law, $1/f^\beta$, can be used to describe noise in nature, and is called power-law noise (Mandelbrot, 1982). For power-law noise, its spectrum can be used to characterize different noise and categorize noise into different "colors". The color of the environmental noise has been investigated for some time. For instance, it was brought to attention in ecology by Steele, who proposed the color of terrestrial and marine noise should be different (Steele, 1985). Based on this, a wide range of studies examined different climatological and hydrological variables based on various colored noise and their influence on population dynamics (Vasseur and Yodzis, 2004). In this paradigm, white noise ($\beta = 0$) is a special case with the same variance at all frequencies. Therefore, the power spectral density of white noise is flat, and its corresponding spectral slope is zero. The spectrum of precipitation sets was assumed as white noise in previous studies (Delworth and Manabe, 1988; Katul et al., 2007; Nakai et al., 2014). Compared to white noise, colored noise refers to noise whose power spectral density function is not flat, which is dominated by frequencies in a certain band.

According to the slope of the power spectral density (i.e., β in inverse power law $1/f^\beta$), the colored noise can be mainly divided into five types: violet noise, blue noise, pink noise, red noise (also known as Brownian noise (Gilman et al., 1963)), and black noise. In a limited frequency band, the spectral density of blue and violet noise increases with the increase of frequency by 3dB and 6dB per octave, and the spectral density of pink and red noise decreases with the increase of frequency by 3dB and 6dB per octave. In other words, the spectral density of blue and violet noise is proportional to the frequency and the square of the frequency, respectively, while the spectral density of pink and red noise is inversely proportional to the frequency and the square of the frequency, respectively. Therefore, the spectral slopes of violet, blue, pink, and red noise are 2, 1, -1, and -2, and the spectral slope of black noise is less than -2 (Nakai et al., 2014).

References

Cochran, W. T., Cooley, J. W., Favin, D. L., Helms, H. D., Kaenel, R. A., Lang, W. W., ... & Welch, P. D. (1967). What is the fast Fourier transform?. *Proceedings of the IEEE*, 55(10), 1664-1674.

Cooley, J. W., & Tukey, J. W. (1965). An algorithm for the machine calculation of complex Fourier series. *Mathematics of computation*, 19(90), 297-301.

Delworth, T. L., & Manabe, S. (1988). The influence of potential evaporation on the variabilities of simulated soil wetness and climate. *Journal of Climate*, *1*(5), 523-547.

Entekhabi, D., Njoku, E. G., O'Neill, P. E., Kellogg, K. H., Crow, W. T., Edelstein, W. N., ... & Kimball, J. (2010). The soil moisture active passive (SMAP) mission. *Proceedings of the IEEE*, 98(5), 704-716.

Gentleman, W. M., & Sande, G. (1966, November). Fast Fourier transforms: for fun and profit. *In Proceedings of the November 7-10, 1966, fall joint computer conference* (pp. 563-578).

Gilman, D. L., Fuglister, F. J., & Mitchell Jr, J. M. (1963). On the power spectrum of "red noise". *Journal of the Atmospheric Sciences*, *20*(2), 182-184.

Katul, G. G., Porporato, A., Daly, E., Oishi, A. C., Kim, H. S., Stoy, P. C., ... & Siqueira, M. B. (2007). On the spectrum of soil moisture from hourly to interannual scales. *Water Resources Research*, *43*(5).

Kay, S. M., & Marple, S. L. (1981). Spectrum analysis—a modern perspective. *Proceedings of the IEEE*, 69(11), 1380-1419.

Mandelbrot, B. B. (1982). *The fractal geometry of nature* (Vol. 2). New York: WH freeman.

Nakai, T., Katul, G. G., Kotani, A., Igarashi, Y., Ohta, T., Suzuki, M., & Kumagai, T. O. (2014). Radiative and precipitation controls on root zone soil moisture spectra. *Geophysical Research Letters*, *41*(21), 7546-7554.

Steele, J. H. (1985). A comparison of terrestrial and marine ecological systems. Nature, 313(6001), 355-358.

Taylor, K. E., Stouffer, R. J., & Meehl, G. A. (2012). An overview of CMIP5 and the experiment design. *Bulletin of the American Meteorological Society*, *93*(4), 485-498.

Vasseur, D. A., & Yodzis, P. (2004). The color of environmental noise. *Ecology*, *85*(4), 1146-1152.