Supporting Information for "Long Temporal Autocorrelations in Tropical Precipitation Data and Spike Train Prototypes"

Tristan H. Abbott,[1] Samuel N. Stechmann,[2,3] J. David Neelin[4]

Contents of this file
Text S1-S8
Figures S1-S4

Introduction

Text S1, S2, S3, and S4 contain details about simulation and data analysis methods that are summarized in the main text. Text S5 provides details about the model parameters used to produce Figure 2 in the main text. Text S6, S7, and S8 give additional details about the SN14 point process idealization and the derivation of a power law with exponent -0.5 in the autocorrelation of that idealization.

Figure S1 shows autocorrelations and autocorrelation decompositions calculated for the full Manus and Nauru time series (without seasonal segmentation). Figure S2 illustrates the diurnal cycles discussed in Text S2, and Figure S3 shows autocorrelations in observational precipitation data with and without diurnal cycles removed. Figure S4 supports the discussion in Text S8.

Text S1: Observational Data Analysis

The Manus and Nauru time series used in the main text are subsets of the full ARM time series, which contain 14+ years of data for both Manus and Nauru. The first approximately 50 percent of both full ARM time series were discarded because they contain about 50 percent missing data, and the end of the full Nauru ARM time series was discarded because it contains a 100 hour event with average precipitation rates over 10 mm/h that we feel is most likely an instrumental anomaly. The remaining data contain a small number of missing values (about 1% of all values) and produce near-identical autocorrelations regardless of whether the missing data are replaced by rates of 0 mm/h or by the mean rate of the time series (which effectively excludes missing values when computing autocorrelations).

Observed precipitation time series were coarse grained from 1 minute to 5 minutes and precipitation rates of less than 0.3 mm/h were set to zero to isolate deep convective precipitation. Observational autocorrelations and autocorrelation decompositions are robust with respect to coarse-graining and thresholding for coarsenings between 1 and 15 minutes (1, 5, 10, and 15 minutes) and cutoffs between 0 mm/h and 1 mm/h (0, 0.1, 0.3, 0.5, and 1 mm/h). We chose to use 5 minute coarse-grained data for the figures because the post-coarsening resolution is fine enough to resolve the decay of the intra-event autocorrelation and chose to use a cutoff of 0.3 mm/h as a compromise that is reasonably low while removing small values likely to be affected by instrument error or by drizzle that would have a different physical source than the main convective precipitation. Seasonal autocorrelations for observational time series were calculated from lags of 0 to 10^4 minutes, and raw autocorrelations were calculated from lags of 0 to 5×10^5 minutes. Autocorrelations were collected and averaged in ten logarithmically spaced bins per minute-decade for plotting. Short-lag exponential fits to autocorrelations from observational data were done by taking the logarithm of the autocorrelation function and performing a least-squares linear regression over lags 0 out to 10 minutes. Because of the coarse graining, this regression only uses three data points; however, the results did not change significantly in sensitivity tests where fits were done over longer ranges of lags.

Text S2: Effect of Cyclical Processes on Observational Autocorrelations

Diurnal cycles for precipitation time series from Manus and Nauru were calculated as follows. Let d be the number of time steps in a day, and let p_i be the i^{th} time step in a precipitation time series after coarse graining to 5 minute averages. The diurnal cycle of the time series is c, a function that maps time steps at a particular point within a day onto precipitation rates:

$$c(i) = \text{Exp}[\, p_j : j \equiv i \mod d]. \tag{1}$$

A smoothed diurnal cycle c_k is constructed by taking expectations over ranges of time steps:

$$c_k(i) = \text{Exp}\,[\, p_j : |(j \mod d) - i| \leq k]\,. \tag{2}$$

Note that $c_1(i) = c(i)$.

Figure S2 shows unsmoothed and smoothed diurnal cycles calculated from 5 minute coarse grained time precipitation time series from Manus and Nauru.

The autocorrelation of the full precipitation time series, the autocorrelation of the time series with an unsmoothed diurnal cycle removed, and the autocorrelation of the time series after removing a diurnal cycle smoothed over ± 5 time steps, or ± 25 minutes, are very similar (Figure S3). In both the smoothed and unsmoothed cases, removing the diurnal cycle achieves at best a very small reduction in the size of the error bars of the autocorrelation function and does not change the binned autocorrelation function significantly.

[1]Department of Computer Sciences, University of Wisconsin-Madison, Madison, Wisconsin, USA.
[2]Department of Mathematics, University of Wisconsin-Madison, Madison, Wisconsin, USA.
[3]Department of Atmospheric and Oceanic Sciences, University of Wisconsin-Madison, Madison, Wisconsin, USA.
[4]Department of Atmospheric and Oceanic Sciences, University of California-Los Angeles, Los Angeles, California, USA.

0094-8276/16/$5.00

Text S3: SN14 Model Simulations and Autocorrelation Computations

SN14 precipitation time series were produced using numerical simulations. The analytical results in SN14 provide stationary CWV distributions and precipitation probabilities conditioned on CWV, which allowed stationary simulations to be initialized with rejection sampling rather than a spin up to a stationary state. Model simulations were done using time steps of 0.01 hours, and the model's stochastic differential equations were advanced using the Euler-Maruyama method [see e.g. *Higham*, 2001]. After each time step, jumps between states were done if CWV passed the relevant threshold during that time step. Simulations were run for 4×10^7 time steps; this length provides reasonable convergence of autocorrelation statistics at lags of up to 10^5 time steps. Autocorrelations of precipitation time series from SN14 model runs were calculated for lags of 0 out to 10^5 time steps (0 to 60,000 minutes, or just over 40 days), and were collected and averaged in 10 logarithmically spaced bins per minute-decade for plotting. Short-lag exponential fits to autocorrelations from model data were done by taking the logarithm of the autocorrelation function and performing a least-squares linear regression over lags 0 out to 10 minutes. Because of the short model time step, this range includes 17 data points. Again, fit results did not change significantly when done over larger numbers of lags.

Text S4: Autocorrelation Decomposition Computations

Plots of autocorrelation decompositions were prepared for observed time series using the methods described for in Section 2.1 in the main text, and pairs of points were assigned to intra- and inter-event components just before autocorrelations were calculated and after coarse graining and applying the cutoff. For model output time series, plots of autocorrelation decompositions were prepared using methods analogous to those described for full autocorrelations in Section 2.2 in the main text, and pairs of points were again assigned to intra- and inter-event components just before calculating autocorrelations. Autocorrelation decompositions were computed from lags of 0 out to 5×10^5 minutes for raw Manus and Nauru autocorrelation decompositions, from lags of 0 out to 10^4 minutes for seasonal Manus and Nauru autocorrelation decompositions, and from lags of 0 out to 10^5 minutes for SN14 model autocorrelation decompositions.

A simple conjecture for the form of the intra-event cutoff is $R_a \sim \tau^{-\alpha} \exp(-\tau/\tau_a)$, where one would expect the exponent α to be approximately that of the apparent exponent of the full correlation in that range, and the cutoff scale τ_a to characterize the scale at which the rapid drop of intra-event correlation occurs. Fits of this form with fixed $\alpha = -0.8$ (not shown) over the full range for which data points are available for the intra-event correlation yielded good correspondence for the Nauru observations. However, noticeable departures from this form occurred for both the Manus observations and SN14 model, suggesting a potentially more complex form of the cutoff. The resulting cutoff scale values wuth $\alpha = -0.8$ were approximately 70 min for Manus, 180 min for Nauru, and 110 min for the SN14 model, which can be useful as a rough estimate of the scale at which intra-event correlation drops.

Text S5: Parameter Settings and Estimated Exponents for Figure 2

The exponents shown in Figure 2 in the main text were estimated using a log-space least squares regression on the full autocorrelations between t_{Sd} and t_{Ld}. The parameter settings for each plot are given below:

Subfigure	t_{Sd} (min)	t_{Ld} (min)	t_{Sw} (min)	t_{Lw} (min)
a	33.75	10667	4.22	76.80
b	18.36	1076	4.22	32.83
c	79.86	237700	4.22	255.5
d	33.75	1.07×10^6	4.22	10.54

Text S6: Precipitation Time Series from the SN14 Point Process Idealization

Stechmann and Neelin [2014] give the mean $\langle s \rangle$ and variance $\langle s^2 \rangle - \langle s \rangle^2$ of SN14 event sizes:

$$\langle s \rangle = b \tag{3}$$

$$\langle s^2 \rangle - \langle s \rangle^2 = b \frac{D_1^2}{P_*}. \tag{4}$$

To reach the SN14 point process idealization, we fix b and take $D_1^2, P_* \to \infty$ while $D_1^2/P_* \to 0$. This brings both wet spell cutoffs to 0 and also brings the event size variance to 0 while leaving the mean unaffected. We represent instantaneous events at time t_i with size b by $b\delta(t_i)$, and so precipitation time series in the SN14 point process idealization are

$$P(t) = b \sum_{t_i} \delta(t - t_i) \tag{5}$$

Note the difference from traditional slow-fast dynamics scenarios in meteorology: here, the fast component turns on intermittently rather than being always active.

Text S7: Derivation of the Analytic -0.5 Power Law

An analytical technique for calculating the autocovariance of a temporal point process

$$S(t) = \sum_{t_i} \delta(t - t_i) \tag{6}$$

is presented in *Gerstner and Kistler* [2002], where they derive a formula for the Fourier transform $\hat{R}_*(\omega)$ of the autocovariance:

$$\hat{R}_*(\omega) = \nu \Re \left\{ \frac{1 + \hat{p}(\omega)}{1 - \hat{p}(\omega)} \right\} + 2\pi\nu^2 \delta(\omega). \tag{7}$$

Here, $\hat{p}(\omega)$ denotes the Fourier transform of the interspike interval distribution, $\nu = \langle S(t) \rangle$ is the mean firing rate of the spike train, and $\Re$ denotes the real part of a complex variable. We use their technique is used along with our own asymptotics to derive a true power law with exponent -0.5 in the autocovariance of the SN14 spike train idealization.

Spike trains from the SN14 point process idealization are of the form

$$P(t) = b \sum_{t_i} \delta(t - t_i), \tag{8}$$

which have autocovariances with Fourier transform

$$\hat{R}_*(\omega) = \frac{b^2}{\sqrt{t_{Sd}t_{Ld}}}\Re\left\{\frac{1+\hat{p}(\omega)}{1-\hat{p}(\omega)}\right\} + 2\pi\frac{b^2}{t_{Sd}t_{Ld}}\delta(\omega). \quad (9)$$

Intervals between spikes are governed by the dry-spell duration PDF $p_{t0}(t)$ given in SN14:

$$p_{t0}(t) = \sqrt{\frac{t_{Sd}}{\pi}}\exp\left(\sqrt{\frac{4t_{Sd}}{t_{Ld}}}\right)\exp\left(-\frac{t_{Sd}}{t}\right)\exp\left(-\frac{t}{t_{Ld}}\right)t^{-3/2}. \quad (10)$$

$p_{t0}(t)$ is an inverse gaussian distribution with Fourier transform

$$\hat{p}(\omega) = \exp\left[2\sqrt{\frac{t_{Sd}}{t_{Ld}}}\left(1-\sqrt{1-i\omega t_{Ld}}\right)\right] \quad (11)$$

[see e.g. *Johnson et al.*, 1994].

Although Equation 9 does not have an obvious analytical reverse transform, an asymptotics-based argument can be used derive the -0.5 exponent. For the sake of simplicity, the argument will use a modified version of Equation 9:

$$\hat{R}_+(\omega) = \frac{b^2}{\sqrt{t_{Sd}t_{Ld}}}\Re\left\{\frac{1+\phi(\omega)}{1-\phi(\omega)}\right\}, \ \ \omega > 0. \quad (12)$$

Because $\hat{R}_*(\omega)$ is even and real, the inverse Fourier transform of $\hat{R}_*(\omega)$ is

$$\int_{-\infty}^{\infty}\hat{R}_*(\omega)\exp\left(-i\omega t\right) = 2\left(\int_0^{\infty}\hat{R}_+(\omega)\cos\left(\omega t\right)\right) + \frac{b^2}{t_{Sd}t_{Ld}}. \quad (13)$$

The $\delta(\omega)$ term in $\hat{R}_*(\omega)$ becomes a constant when the inverse transform is applied, so any power laws in the inverse Fourier transforms of Equations 9 and 12 will have the same exponent.

The argument proceeds by assuming t_{Ld} and t_{Sd} are well separated and examining the regime where ω^{-1} is large relative to t_{Sd} and small relative to t_{Ld}. First, assuming $\omega t_{Ld} >> 1$ and expanding the second square root in Equation 11,

$$\sqrt{1-i\omega t_{Ld}} = \left(-i\omega t_{Ld}\right)^{1/2}\left(1-\frac{1}{i\omega t_{Ld}}\right)^{1/2} \quad (14)$$

$$= \left(-i\omega t_{Ld}\right)^{1/2}\left(1-\frac{1}{2i\omega t_{Ld}}+...\right), \quad (15)$$

and truncating the expansion at zero order,

$$\sqrt{1-i\omega t_{Ld}} \approx \left(-i\omega t_{Ld}\right)^{1/2} \quad (16)$$

Substituting this approximation in Equation 11,

$$\hat{p}(\omega) \approx \exp\left[2\sqrt{\frac{t_{Sd}}{t_{Ld}}}\left(1-\frac{i-1}{\sqrt{2}}\sqrt{\omega t_{Ld}}\right)\right] \quad (17)$$

$$= \exp\left[2\sqrt{\frac{t_{Sd}}{t_{Ld}}}\right]\exp\left[(1-i)\sqrt{2\omega t_{Sd}}\right]. \quad (18)$$

Second, assuming $\omega t_{Sd} << 1$ (in addition to $\omega t_{Ld} >> 1$) makes the first exponential in Equation 18 approximately 1 and allows it to be discarded. Expanding the resulting approximation for $\hat{p}(\omega)$ as a Maclaurin series and retaining up to first order terms gives, to leading order,

$$1+\hat{p}(\omega) \approx 2+(1-i)\sqrt{2\omega t_{Sd}} \approx 2 \quad (19)$$

and

$$1-\hat{p}(\omega) \approx (i-1)\sqrt{2\omega t_{Sd}}. \quad (20)$$

Substituting Equations 19 and 20 into Equation 9 and performing some algebraic simplifications,

$$\hat{R}_+(\omega) \approx \frac{b^2}{\sqrt{2t_{Sd}^2t_{Ld}}}\omega^{-1/2} = \tilde{\hat{R}}_+(\omega) \quad (21)$$

We now consider the implications of this approximation in the inverse Fourier transform Equation 13. Evaluating the first term using $\tilde{\hat{R}}_+(\omega)$ from 0 to ∞ yields

$$\int_0^{\infty}\tilde{\hat{R}}_+(\omega)cos(\omega\tau)d\omega = \frac{b^2}{\sqrt{2t_{Sd}^2t_{Ld}}}\tau^{-1/2}. \quad (22)$$

The assumptions $t_{Ld}^{-1} << \omega << t_{Sd}^{-1}$ do not hold for all of $0 \le \omega < \infty$, so before we can conclude that Equations 22 and 13 together yield an autocovariance with a $\tau^{1/2}$ power law, the contributions from the low and high end of the range of integration must be addressed. An analytical argument can show that, for $t_{Sd} < \tau < t_{Ld}$, the parts of the integral in Equation 22 where these assumptions do not hold contribute a constant that depends only on model parameters to the value of the integral. Similarly, numerical experiments show that for $t_{Sd} < \tau < t_{Ld}$, the parts of the integral in the LHS of Equation 13 where the assumptions do not hold contribute a constant amount to the value of that integral. These arguments are dicussed in more detail in the following section of the SI. It follows that the difference between the true autocovariance (the RHS of Equation 13) and the approximation given in Equation 22 is a constant. That difference can be combined with the constant from the $\delta(\omega)$ term in Equation 13 in a constant C. Then, for sufficiently separated t_{Sd} and t_{Ld}, the spike train idealization has an autocovariance that has a $\tau^{-1/2}$ power law range:

$$R_*(\tau) \approx b^2\sqrt{\frac{2}{t_{Sd}^2t_{Ld}}}\tau^{-1/2} + C. \quad (23)$$

Text S8: Fourier Integral Contributions for Large and Small ω

The approximation presented in the previous section applies in the regime where $t_{Sd}^{-1} >> \omega >> t_{Ld}^{-1}$. However, the Fourier integral used to transform $\hat{R}(\omega)$ into $R(\tau)$ is over the range $-\infty < \omega < \infty$. This section argues that the value of the Fourier integral is set by the behavior of the integral in the range $t_{Sd}^{-1} >> \omega >> t_{Ld}^{-1}$.

Because $\hat{R}(\omega)$ is even and real, its Fourier transform will also be even and real, and

$$\int_{-\infty}^{\infty}\hat{R}(\omega)\exp\left[i\omega t\right]d\omega = \int_0^{\infty}\hat{R}(\omega)\cos\left(\omega t\right)d\omega. \quad (24)$$

Let $\tilde{\hat{R}}(\omega)$ be the asymptotic approximation of $\hat{R}$ derived in Text S2. Then

$$\begin{aligned}\int_0^\infty \hat{R}(\omega)\cos(\omega t)\,d\omega &= \int_0^{t_{Ld}^{-1}} \hat{R}(\omega)\cos(\omega t)\,d\omega \\ &+ \int_{t_{Ld}^{-1}}^{t_{Sd}^{-1}} \tilde{\hat{R}}(\omega)\cos(\omega t)\,d\omega \\ &+ \int_{t_{Sd}^{-1}}^{\infty} \hat{R}(\omega)\cos(\omega t)\,d\omega \\ &+ \int_{t_{Ld}^{-1}}^{t_{Sd}^{-1}} \left[\hat{R}(\omega) - \tilde{\hat{R}}(\omega)\right]\cos(\omega t)\,d\omega. \end{aligned} \tag{25}$$

In order to show that the $[0,\infty)$ integral of $\hat{R}(\omega)$ is a $t^{-1/2}$ power law, we show that the $[0, t_{Ld}^{-1}]$ integral of $\hat{R}(\omega)$ and the $[t_{Sd}^{-1}, \infty)$ integral of $\hat{R}(\omega)$ both contribute only a constant to the value of the $[0,\infty)$ integral, show that the $[t_{Ld}^{-1}, t_{Sd}^{-1}]$ integral of $\tilde{\hat{R}}(\omega)$ is a $t^{-1/2}$ power law, and assume that, because $\tilde{\hat{R}}$ is an approximation for $\hat{R}$ for $t_{Ld}^{-1} << \omega << t_{Sd}^{-1}$, the $[t_{Ld}^{-1}, t_{Sd}^{-1}]$ integral of $\hat{R}(\omega) - \tilde{\hat{R}}(\omega)$ can be neglected. Rather than showing this for all t, we show this for $t_{Sd} << t << t_{Ld}$, the range of t where we expect to see scale-free behavior.

For the large-ω and small-ω ranges of the $\hat{R}(\omega)$ integral, we argue the desired property using numerical experiments. Specifically, we explore the behavior of

$$I(\omega', t) = \int_0^{\omega'} \hat{R}(\omega)\cos(\omega t)\mathrm{d}\omega \tag{26}$$

using an adaptive quadrature (`gsl_integration_qawf` from the GNU Scientific Library). Because $\hat{R}(\omega)$ is singular at $\omega = 0$, the quadrature routine fails when asked to approximate an integral over $[0,\infty)$. Instead, we used the quadrature to approximate integrals over $[\epsilon,\infty)$. Our results were the same for all $\epsilon \in [10^{-9}, 10^{-7}]$ h^{-1}, and we therefore believe that integrals over $[\epsilon,\infty)$ with sufficiently small ϵ behave the same as integrals over $[0,\infty)$.

Figure S3 displays two plots of $I(\omega', t)$ that differ in the value of t_{Ld}. For values of t in the range $[t_{Sd}, t_{Ld}]$, the value of $I(\omega', t)$ grows with ω' like a power law with exponent 1 up to $\omega' \approx t_{Ld}^{-1}$. For $\omega' > t_{Sd}^{-1}$, $I(\omega', t)$ remains approximately constant. This suggests that the size of the contribution from the small-ω range of the RHS of Equation 27 depends linearly on t_{Ld}^{-1} but is independent of t, and that the size of the contribution of the large-ω range is negligible.

For

$$\int_{t_{Ld}^{-1}}^{t_{Sd}^{-1}} \tilde{\hat{R}}(\omega)\cos(\omega t)\,d\omega, \tag{27}$$

we use an analytic argument. Because the Fourier transform of $\delta(\omega)$ is a constant, the $\delta(\omega)$ terms of both sides of Equation 27 contribute only a constant to the values of the integrals and can be ignored. Omitting the $\delta(\omega)$ term and the factor $b^2/\sqrt{2t_{Sd}t_{Ld}}$, the integral for Equation 27 is

$$\int_{t_{Ld}^{-1}}^{t_{Sd}^{-1}} \omega^{-1/2}\cos(\omega t)\,d\omega = \frac{2\pi C\left(\sqrt{\frac{2t}{\pi t_{Sd}}}\right)}{\sqrt{t}} - \frac{2\pi C\left(\sqrt{\frac{2t}{\pi t_{Ld}}}\right)}{\sqrt{t}} \tag{28}$$

where C is the Fresnel C function as defined in *Olver et al.* [2010]. For $s >> t$, $C(s) \approx 1/2$, and for $s << 1$, $C(s) \approx 0$ [*Olver et al.*, 2010], so the value of Equation 28 is, for $t_{Sd} << t << t_{Ld}$, approximately $\pi t^{-1/2}$.

References

Gerstner, W., and W. M. Kistler (2002), *Spiking Neuron Models: Single Neurons, Populations, Plasticity*, Cambridge University Press.

Higham, D. J. (2001), An Algorithmic Introduction to the Numerical Simulation of Stochastic Differential Equations, *SIAM Review*, *43*(3), 525–546.

Johnson, N. L., S. Kotz, and N. Balakrishnan (1994), *Continuous Univariate Distributions*, Wiley.

Olver, F. W., et al. (Eds.) (2010), *NIST Handbook of Mathematical Function*, Cambridge University Press.

Stechmann, S., and J. D. Neelin (2014), First-Passage-Time Prototypes for Precipitation Statistics, *Journal of the Atmospheric Sciences*, *71*, 3269–3291.

Corresponding author: Samuel N. Stechmann, Department of Mathematics, University of Wisconsin-Madison, 480 Lincoln Dr, Madison, WI 53706, USA. (stechmann@wisc.edu)

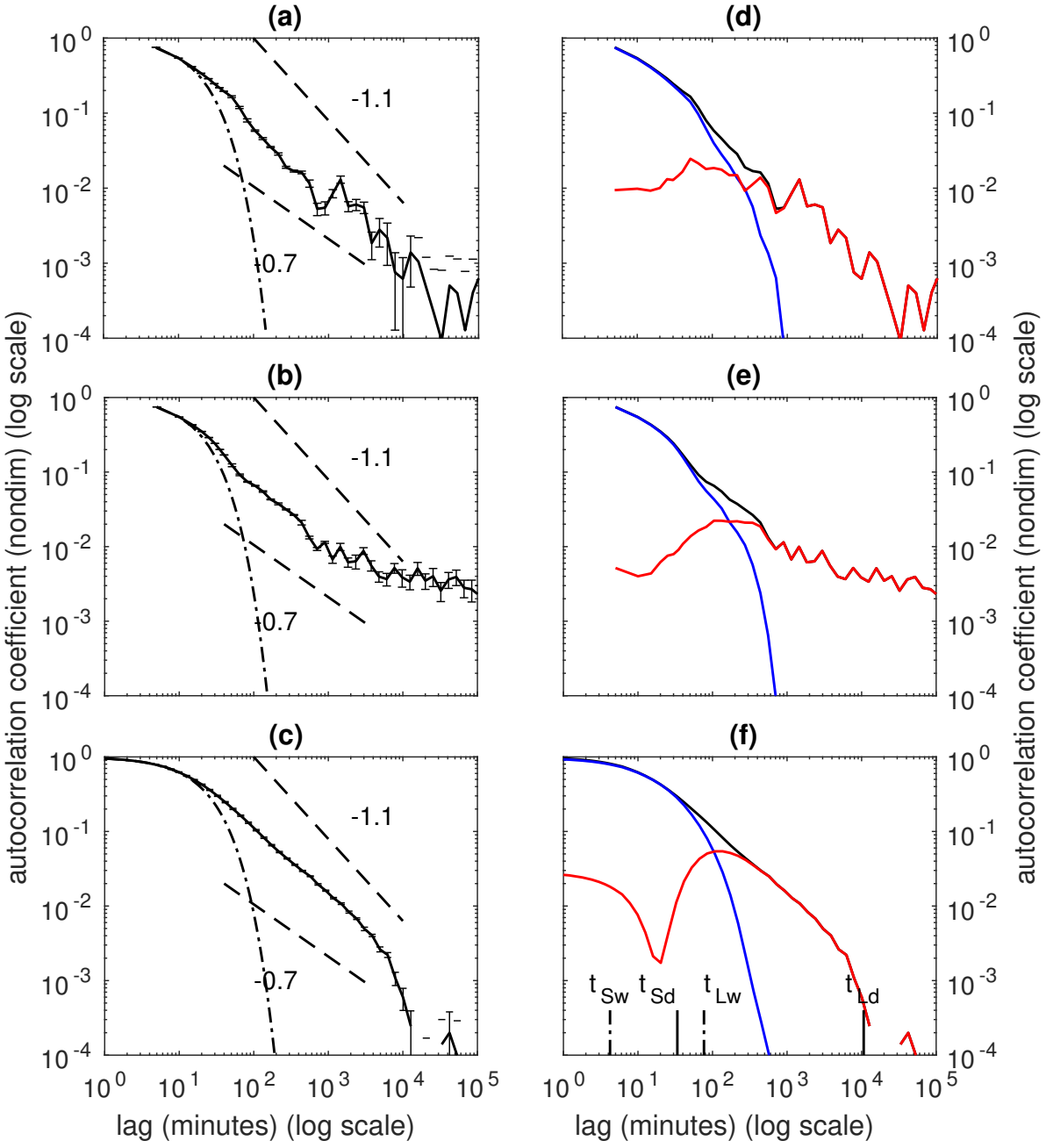

Figure S1. Temporal autocorrelations and autocorrelation decompositions for observed precipitation time series from Manus (a,d) and Nauru (b,e) and precipitation time series from an SN14 two state model simulation (c,f). In contrast with Figure 1 in the main text, the autocorrelations and autocorrelation decompositions shown here for Manus and Nauru are from the full time series, not from averages of seasonal segments. On the left, solid black lines indicate autocorrelations after binning, error bars show the standard error of the mean for each bin, dashed black lines show the power laws $t^{-0.7}$ and $t^{-1.1}$, and dot-dashed black lines show an exponential fit from lag 0 to lag 10 minutes. On the right, black lines show the full autocorrelation of the time series, blue lines show the intra-event component, and red lines show the inter-event component. On the lower right, dashed (resp. solid) black ticks indicate the wet-spell (resp. dry-spell) duration cutoffs of the two state model.

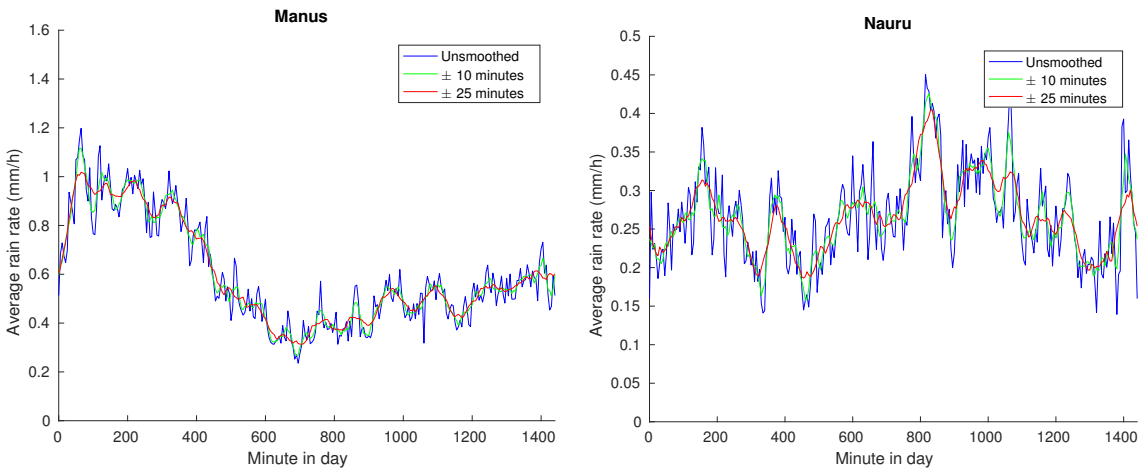

Figure S2. Reconstructions of the diurnal cycle for Manus (left) and Nauru (right) based on precipitation data coarsegrained to 5 minute intervals. Blue lines represent the unsmoothed reconstruction, green lines represent a reconstruction smoothed over ranges of ±2 time steps (±10 minutes), and red lines represent the reconstruction smoothed over ranges of ±5 time steps (±25 minutes).

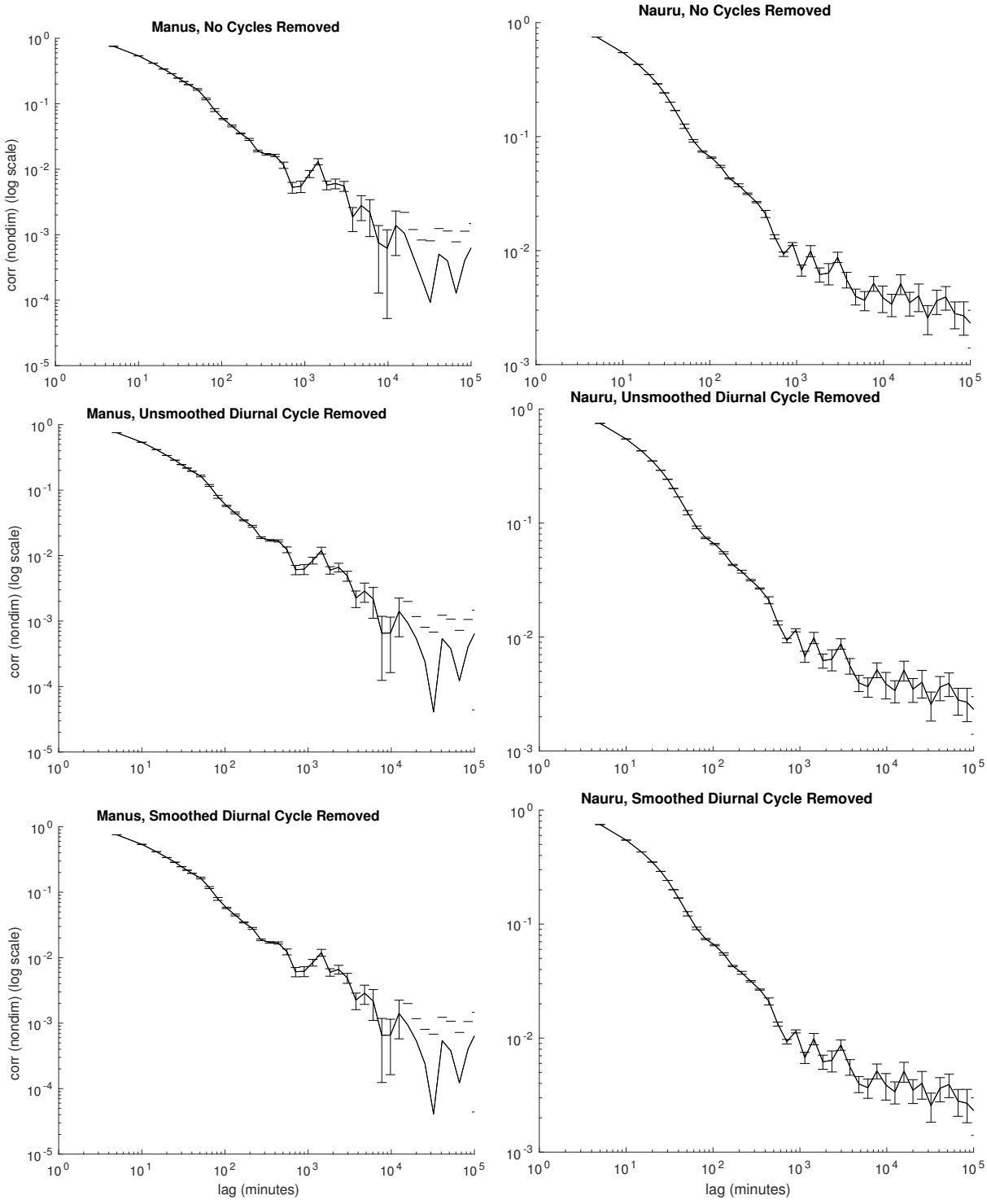

Figure S3. Precipitation autocorrelations from Manus (left) and Nauru (right). Plots in the top row are autocorrelations calculated without removing diurnal cycles, plots in the middle row are autocorrelations calculated after removing unsmoothed diurnal cycles, and plots on the bottom row are autocorrelations calculated after removing diurnal cycles smoothed over ranges of ±5 time steps (±25 minutes).

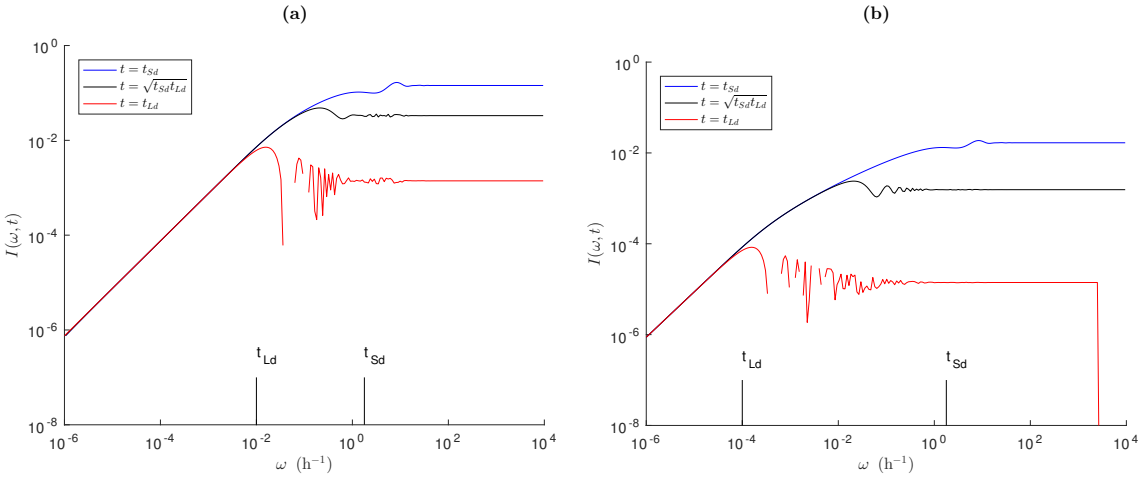

Figure S4. $I(\omega, t)$ from Equation 26 against ω. Each plot displays the value of $I(\omega, t)$ plotted against ω for $t = t_{Sd}$ (blue), $t = \sqrt{t_{Sd} t_{Ld}}$ (black), and $t = t_{Ld}$ (red). $\omega = t_{Sd}^{-1}$ and $\omega = t_{Ld}^{-1}$ are indicated by black ticks at the bottom of the plots. For plot (a), $t_{Ld} = 100$ h, and for plot (b), $t_{Ld} = 6400$ h.