LAWRENCE
LIVERMORE
NATIONAL
LABORATORY

LLNL-TR-654743

Standard Diagnostics for the Diurnal Cycle

C. Covey, P. Gleckler

May 21, 2014

This work performed under the auspices of the U.S. Department of Energy by Lawrence Livermore National Laboratory under Contract DE-AC52-07NA27344.

Standard Diagnostics for the Diurnal Cycle

Curt Covey and Peter Gleckler
PCMDI / LLNL*
<covey1@llnl.gov>
20 May 2014

* Work performed under auspices of the Office of Science, US Department of Energy by Lawrence Livermore National Laboratory under Contract DE-AC52-07NA27344.

Introduction

The purpose of this document is to suggest a strategy for constructing standard diagnostics of the diurnal cycle from both climate model and observational data products. The guiding philosophy is "keep it simple," in the hope that a diagnostic software package can be readily constructed and widely used. Of course this means that the output of the package will form only the beginning of necessary examination of the diurnal cycle in models and in the real world.

The latest version of the Coupled Model Intercomparison Project, CMIP5, provides several fields at 3-hourly time resolution near the surface: air and surface temperatures, pressure, humidity, soil moisture, horizontal wind, energy flux components, overhead cloudiness, evaporation, precipitation, convective precipitation, and snowfall. (See the "3hr" tab in the Standard Output spreadsheet at http://cmip-pcmdi.llnl.gov/cmip5/data_description.html.) Meanwhile satellite observations provide at least equally fine time resolution and global coverage for some of these fields. This data makes possible an extensive study of the diurnal cycle near the surface.

Covey et al. (2011, 2014) have published analyses of CMIP 3-hourly surface pressure fields and shown them to be consistent with the conventional picture of "atmospheric tides." Like their oceanic relatives, atmospheric tides are periodic in time and spatially simple at large scales. Thus the surface-pressure tides provide an easy starting point for study of the diurnal cycle. At the opposite end of the complexity scale, precipitation is very irregular in space and time. This document will use precipitation as an example of the challenges to construction of standard diagnostics of the diurnal cycle.

Analysis of the diurnal cycle of a time series $x(t)$ typically begins by forming a "composite" or average diurnal cycle spanning $0 < t < 24$ hours. Then further processing such as Fourier analysis is applied to the composite. Two questions arise: (1) What further processing is most appropriate for simple standard diagnostics that can produce a few key "metric" numbers when climate models are compared with each other or with observations? (2) Is it necessary to form a composite in the first place? Below we argue that (1) straighforward Fourier analysis is the most appropriate procedure and (2) forming a diurnal cycle composite is not necessary because Fourier analysis of $x(t)$ over its entire domain gives the same diurnal (24h), semidiurnal (12h) and higher harmonics of the diurnal cycle. Nevertheless, a composite may be useful because it can be inspected before succeeding steps are taken, without affecting the final result.

(1) Methods of Analysis

Dai et al. (2007) made a composite diurnal cycle of precipitation at each latitude / longitude grid point for December-January-February and June-July-August over several years. Then they least-squares fit diurnal and semidiurnal cycles to each grid point's composite. The adjustable parameters in each fit are the amplitude and phase of the cycle. Finally they mapped the amplitude (as a percentage of daily mean precip) and the phase (as time of maximum) for both the diurnal and semidiurnal fits at both seasons: 8 maps in all. At this point the results can be used to produce standard metrics such as the numbers plotted in a Taylor diagram. An important caveat is that for time-of-maximum comparisons, one must use modular arithmetic: 24h = 0h for the diurnal harmonic, 12h = 0h for the semidiurnal harmonic, and so on.

This procedure is the most common one used for diurnal cycle analysis, but there are others. Kikuchi and Wang (2008) simply took a climatological mean of the difference between daily-maximum and daily-minimum precipitation as a measure of diurnal cycle amplitude. For the phase, they produced Empirical Orthogonal Eigenfunctions from a composite diurnal cycle. With TRMM data they found that the first two EOFs represented the diurnal harmonic while the next two represented the semidiurnal harmonic. Finally they plotted time series of the corresponding EOF amplitudes (principal components) for DJF and JJA seasons and for the annual mean. These comprise only 6 line plots of precipitation rate spanning $0 < t < 24$ hours (as opposed to the initial composites in this procedure, and Dai et al.'s, which amount to a line plot for each grid point). Here again the results can be used to produce standard metrics. Wang et al. (2011) propose diagnostic metrics based entirely on such EOFs. By definition, EOFs provide the most compact representation of the principal variations of space-time fields. They have been popular diagnostics since their introduction to meteorology and climatology by Lorenz (1956) and Kutzbach (1967). Compared with Fourier analysis, however, EOFs are not as simple conceptually and not as accessible computationally. Thus they seem inconsistent with a "keep it simple" philosophy for standard diagnostics.

It should be noted that none of the above procedures address the "frequency versus intensity" issue of the diurnal cycle of precipitation. For example, the upper left panel of Figure 4 in Dai et al. shows a composite diurnal cycle for the Southeastern USA. This time series has a smooth once-a-day maximum at a well defined time. But the series is an average over many different days. Does the steady increase and then decrease of precipitation rate exhibited by the composite arise from corresponding steady increases and decreases during most days? Or does it arise from different days' precipitation coming at different times (or not at all) but always at the same rate? Dai et al. argue that the latter explanation is closer to the truth. To make their case, they must reprocess the high-time-frequency raw data in ways that avoid a composite diurnal cycle. This too seems inconsistent with a "keep it simple" philosophy for standard diagnostics.

(2) Are Composites Necessary?

To analyze 3-hourly surface pressure output from climate models, Covey et al. did not form a diurnal cycle composite. They simply applied a Fast Fourier Transform at each grid point to a 32-day detrended time series of anomalies, i.e. values obtained during each day by subtracting that day's mean value. The Fourier analysis produces harmonic components with periods of 32 days, 16 days, and so on down to the Nyquist limit of $2\,\Delta t = 6$ hours, but only the 24h and 12h harmonics were studied.

Procedural details probably make little difference to studying pheonomena as regular as the tides. Indeed, Covey et al. showed that their model analysis agrees well with observational analysis of the tides by Dai and coworkers using the composite technique. But one might worry that the details make a great deal of difference for the diurnal cycle of precipitation.

At first sight a straight-out Fourier analysis seems rather different from the procedure of Dai et al. described above. But the differences are more apparent than real. Although Dai et al. used a least-squares fit of the diurnal and semidiurnal cycles to their composite, it is well known that when a periodic function is approximated by a trigonometric series via least squares, the resulting coefficients are identical to those obtained by Fourier analysis (e.g. Elmore and Heald 1969). So the question “Are the two procedures equivalent?” becomes “Does it matter whether or not a composite is formed before doing the Fourier analysis?” In fact it does not matter, because time-averaging (which forms the composite) and Fourier-transforming are basically linear integral operations that commute. This is shown below, first assuming for simplicity that time is a continuous variable, then for the more pertinent case in which time is measured in discrete steps.

Continuous Time

Taking the units of local solar time t in days, and considering the time period $0 < t < N$ days, the Fourier series for a function over this domain is

$$x(t) = \sum_{m=-\infty}^{\infty} a_m \, e^{2\pi i m t/N} \; ; \qquad a_m = a_{-m}^{*} \text{ for real } x. \tag{1}$$

Note that Eq. (1) always forces $x(N) = x(0)$ so that outside the domain $0 < t < N$, $x(t)$ repeats with time period N. Thus a trend in x over its nominal domain implies a repeating saw-tooth pattern that can be problematic in Fourier analysis. For this reason it is customary to first de-trend the data.

From the orthogonality of trigonometric functions it follows that

$$a_m = N^{-1} \int_0^N x(t) \, e^{-2\pi i m t/N} \, dt. \tag{2}$$

As with any continuous-time Fourier series, a_0 gives the (constant) average of x over the domain, the $a_{\pm 1}$ give amplitude and phase for the longest allowed period (N days), the $a_{\pm 2}$ give amplitude and phase for half this period ($N/2$ days), the $a_{\pm 3}$ give amplitude and phase for a third this period ($N/3$ days), and so on ad infinitum. But for the diurnal cycle, the only relevant coefficients are $a_{\pm N}$ for the diurnal harmonic, $a_{\pm 2N}$ for the semidiurnal harmonic, $a_{\pm 3N}$ for the terdiurnal harmonic, and so on. For comparison with the composite diurnal cycle, it is convenient to write these coefficients as

$$a_{nN} = N^{-1} \int_0^N x(t) \, e^{-2\pi i n t} \, dt = N^{-1} \sum_{k=1}^{N} \int_{k-1}^{k} x(t) \, e^{-2\pi i n t} \, dt. \tag{3}$$

The composite itself is formed by averaging over the time of each day:

$$\overline{x}(t) = N^{-1} \sum_{k=1}^{N} x(t + k - 1) \; ; \qquad 0 < t < 1, \tag{4}$$

e.g. the composite 0800h LST value is $\overline{x}\left(\frac{1}{3}\right) = \left[x\left(\frac{1}{3}\right) + x\left(1 + \frac{1}{3}\right) + x\left(2 + \frac{1}{3}\right) + \ldots + x\left(N - \frac{2}{3}\right)\right] \Big/ N$ in our units.

A Fourier series for $\overline{x}(t)$ is just a special case of Eqs. (1)–(2) with $N = 1$:

$$\overline{x}(t) = \sum_{m=-\infty}^{\infty} \overline{a}_m \, e^{2\pi i m t} \; ; \qquad \overline{a}_m = \overline{a}_{-m}^{*} \text{ for real } \overline{x} \; ; \tag{5}$$

$$\overline{a}_m = \int_0^1 \overline{x}(t) \, e^{-2\pi i m t} \, dt = N^{-1} \int_0^1 \sum_{k=1}^{N} x(t + k - 1) \, e^{-2\pi i m t} \, dt. \tag{6}$$

In this case all of the coefficients that emerge from the procedure are relevant to the diurnal cycle. The $\overline{a}_{\pm 1}$ give amplitude and phase of the diurnal harmonic, the $\overline{a}_{\pm 2}$ give amplitude and phase of the semidiurnal harmonic, the $\overline{a}_{\pm 3}$ give amplitude and phase of the terdiurnal harmonic, and so on ad infinitum.

The relationship between the coefficients in Eqs. (3) and (6) is revealed by interchanging the order of summation and integration in (6) and then substituting $t' \equiv t + k - 1$ in each of the integrals. Since the limits $t=0$ and $t=1$ become $t'=k-1$ and $t'=k$, we have

$$\overline{a}_m = N^{-1} \sum_{k=1}^{N} \int_{k-1}^{k} x(t')\, e^{-2\pi i m (t'-k+1)}\, dt'. \qquad (7)$$

After substituting $m \to n$, dropping the prime on the dummy integration variable t', and recognizing that $e^{-2\pi i n (t-k+1)} = e^{-2\pi i n t}$, comparison with Eq. (3) shows that $\overline{a}_n = a_{nN}$, QED.

Discrete Time

Continuing to measure time in days, suppose each day is divided into S time-segments (e.g. $S=8$ in for 3-hourly data) and the continuous function $x(t)$ is replaced by the sequence x_0, x_1, x_2, ..., x_{SN} at the segment boundaries over $0<t<N$. Since x repeats with time period N (q.v.) $x_{SN}=x_0$, so there are exactly $S\times N$ independent time points x_0, x_1, x_2, ..., x_{SN-1}. The correspondence between $x(t)$ and the x_j is given by $x(t)=x(j/S)\equiv x_j$. Therefore integrals in the equations above are replaced by sums according to

$$\int_a^b f(t)\, dt \to S^{-1} \sum_{j=Sa}^{Sb} f(j/S). \qquad (8)$$

Also, because the interval between time points Δt (=S^{-1}days) is finite, a Nyquist limit applies to the highest frequency that can be resolved by Fourier analysis. The equation analogous to (1) is thus

$$x_j = \sum_{m=-SN/2}^{+SN/2} a_m\, e^{2\pi i m j/SN}; \qquad j=0, 1, 2, \ldots, SN-1. \qquad (1')$$

At the Nyquist frequency limits $m=\pm SN/2$ the period is twice the interval between time points ($2\,\Delta t = 6$ hours for 3-hourly data: a quadradiurnal harmonic). "2 Δt waves" are the highest frequency that discrete Fourier analysis can resolve.

Eq. (1′) represents exactly SN equations in the SN unknowns a_0, $a_{\pm 1}$, $a_{\pm 2}$, ..., $a_{|SN/2|}$. (The coefficients $a_{-SN/2}$ and $a_{+SN/2}$ represent a single Fourier term because $m=\pm SN/2$ gives the same exponential factor $e^{\pm i\pi j}=(-1)^j$.) These equations can be solved to give

$$a_m = (SN)^{-1} \sum_{j=0}^{SN-1} x_j\, e^{-2\pi i m j/SN} \qquad (2')$$

by using the identity $\sum_{j=0}^{M-1} e^{2\pi i n j/M} = M\,\delta_{nM}$, the discrete version of orthogonality for trigonometric functions (or one may simply apply the transformation (8) to Eq. (2)).

Continuing with this procedure, the analysis of Eqs. (1)–(7) can be repeated to reach the same conclusion. Instead of interchanging the order of summation and integration in going from Eq. (6) to Eq. (7), one interchanges the order of summation in a double sum. In both cases the operations commute, and the diurnal and higher harmonics are identical whether or not one first forms a composite diurnal cycle.

Practical Implementation

The equations above are perfectly valid but do not represent the best algorithm for Fourier analysis. The Fast Fourier Transform speeds up the procedure by several orders of magnitude (Press et al. 2007). Standard software packages employ FFTs ubiquitously. For example, Covey et al. wrote scripts in a Python-based climate data analysis language (Williams et al. 2013) that in turn invokes the Numerical Python module **`fft`** for discrete Fourier transforms (http://www.numpy.org). This module is consistent with Eqs. (1′) and (2′) above, or equivalently Eqs. (12.1.7) and (12.1.9) in Press et al. 2007.

References

C. Covey, A. Dai, D. Marsh and R. S. Lindzen: “The Surface-Pressure Signature of Atmospheric Tides in Modern Climate Models,” *Journal of the Atmospheric Sciences 68,* pp. 495-514, 2011

C. Covey, A. Dai, R. S. Lindzen and D. Marsh: “Atmospheric Tides in the Latest Generation of Climate Models,” ibid., in press 2014 (early online release at http://journals.ametsoc.org/doi/abs/10.1175/JAS-D-13-0358.1)

A. Dai, X. Lin and K.-L. Hsu: “The Frequency, Intensity, and Diurnal Cycle of Precipitation in Surface and Satellite Observations over Low- and Mid-Latitudes,” *Climate Dynamics 29,* pp. 727-744, 2007

W. C. Elmore and M. A. Heald: *Physics of Waves*, McGraw-Hill, 1969 (Problem 1.7.2)

K. Kikuchi and B. Wang: “Diurnal Precipitation Regimes in the Global Tropics,” *Journal of Climate 21,* 2680-2696, 2008

J. E. Kutzbach: “Empirical Eigenvectors of Sea-Level Pressure, Surface Temperature and Precipitation Complexes over North America,” *Journal of Applied Meteorology 6,* pp. 791-802, 1967; see also M. Hirose and J. E. Kutzbach: “An Alternate Method for Eigenvector Computations,” ibid. *8,* p. 701

E. N. Lorenz: *Empirical Orthogonal Functions and Statistical Weather Prediction,* Statistical Forecasting Project Scientific Report No. 1, MIT Department of Meteorology, 1956 (http://www.o3d.org/abracco/Atlantic/Lorenz1956.pdf)

W. H. Press, S. A. Teukolsky, W. T. Vetterling and B. P. Flannery: *Numerical Recipes: The Art of Scientific Computing,* 3rd Ed., Cambridge University Press, 2007

D. N. Williams and the UV-CDAT Project Team: “Ultrascale Visualization of Climate Data,” *Computer*, September 2013, pp. 68-75