Rockland Technical Note 061

— Goodman coherent noise removal— spectral bias

2021-11-18

Rockland Scientific
520 Dupplin Rd
Victoria, BC, CANADA, V8Z 1C1
www.rocklandscientific.com

Copyright Notice

info@rocklandscientific.com
tel: +1 250 370 1688

Version History

Date	Description
2021-11-18	Initial version
2022-06-02	Update intro to include ODAS release

1 Introduction

The data collected from shear probes is very sensitive to contamination from various source of vibration and movements. To mitigate this effect, the spectra of turbulence shear are cleaned by removing their vibration-coherent components using the technique of Goodman et al. (2006), which relies on estimating the coherency between shear-probe and vibration signals. However, estimates of coherency are always larger than zero when using a finite length of data, even when the signals are statistically independent and completely incoherent. The unavoidably finite coherency biases low the cleaned spectra and leads to an underestimation of the rate of dissipation of kinetic energy from the variance of shear. The bias decreases with increasing number of fft-segments that are used to estimate a spectrum of shear and its coherency with vibrations. The bias increases with the number of vibration (or other contamination) signals that are used to clean the spectrum of the measured shear. The bias is wavenumber independent, and does not depend on the variance of either the vibrations or the shear.

The paper presented in this Technical Note provides a theoretical derivation of the bias and an equation that can be used to correct the bias of the spectra. This correction has now been implemented in the ODAS Matlab Library, released under version 4.5.

The bias in coherent-noise removal

Rolf G. Lueck * and Duncan MacIntyre † and Justine McMillan ‡

Rockland Scientific, Victoria, Canada

ABSTRACT

A simple formula is provided for correcting the bias of vibration-coherent noise removal. Spectra of turbulence shear are often cleaned by removing their vibration-coherent components using the technique of Goodman et al. (2006), which relies on estimating the coherency between shear-probe and vibration signals. However, estimates of coherency are always larger than zero when using a finite length of data, even when the signals are statistically independent and completely incoherent. The unavoidably finite coherency biases low the cleaned spectra and leads to an underestimation of the rate of dissipation of kinetic energy from the variance of shear. The bias decreases with increasing number of fft-segments that are used to estimate a spectrum of shear and its coherency with vibrations. The bias increases with the number of vibration (or other contamination) signals that are used to clean the spectrum of the measured shear. The bias is wavenumber independent, and does not depend on the variance of either the vibrations or the shear.
2021-03-24

1. Introduction

Measuring the rate of dissipation of turbulent kinetic energy, ϵ, is important for understanding ocean mixing because it provides a means to estimate the diapycnal eddy diffusivity (Osborn 1980). Such measurements are usually done with velocity shear probes mounted on profilers which, like nearly all fluid velocity sensors, measure velocity relative to the platform to which they are mounted (Siddon and Ribner 1965; Siddon 1971; Osborn 1974). Vibrations of the platform induce a contamination that must be removed from the measured shear signal because they can raise considerably the spectrum of shear at the frequencies of the vibrations, for example Figure 4 in Levine and Lueck (1999). The wavenumber-dependent bias induced by vibrations causes an overestimation of the rate of dissipation of kinetic energy, because it is proportional to the variance of shear, which is usually derived by integrating the spectrum of shear using

$$\epsilon = \frac{15}{2}\nu\overline{\left(\frac{\partial v}{\partial x}\right)^2} = \frac{15}{2}\nu\int_0^\infty \Psi(k)\,\mathrm{d}k \approx \frac{15}{2}\nu\int_0^{k_c} \Psi(k)\,\mathrm{d}k \tag{1}$$

where ν is the kinematic viscosity, v is any velocity component orthogonal to the direction of profiling, x is any direction of profiling, $\Psi(k)$ is the spectrum of shear, k is the wavenumber in the x-direction and $k_c < \infty$ is an upper wavenumber imposed by practical considerations (Taylor 1935; Pope 2009).

A popular method for removing vibration-induced contamination of a shear-probe measurements is that of Goodman et al. (2006) which relies on estimating the multivariate squared-coherency between the signals produced by shear-probes and vibration sensors to identify the variance in the shear-probe signals that is induced by platform vibrations. However, the squared-coherency (hereafter, coherency) is a positive definite quantity that is always bigger than zero if it is estimated from a finite length of data, even when the the signals are completely incoherent. Consequently, the technique of Goodman et al. (2006) will remove some spectral variance even when its application is not required.

In section 2 we examine the principal nature of coherent noise removal. In section 3 we demonstrate the bias by (*i*) using statistically independent sequences for shear and vibrations, for which the shear signals do not require noise removal, (*ii*) adding several sinusoidal coherent components to these sequences that do require removal, and (*iii*) using real shear-probe and vibration signals collected in a tidal channel where the shear is strong and the vibrations are weak. The results are discussed and concluded in section 4.

2. Background

The multivariate method of coherent noise removal of Goodman et al. (2006) is most easily understood in its univariate form. Let a shear-probe measurement, $\hat{s}(x)$, be given by

$$\hat{s}(x) = s(x) + h(x) * a(x) + e(x) \tag{2}$$

**Corresponding author*: Rolf Lueck, rolf@rocklandscientific.com

†University of British Columbia, Vancouver, BC, Canada.

‡Luna Sea Solutions, Freeport, Nova Scotia, Canada.

where s is the oceanic shear, a is the vibration of the platform in the direction of sensitivity of the shear probe, e is some added noise in the shear-probe measurement, h is the impulse response that relates the measured shear-probe signal to the vibrations, and the symbol $*$ demotes a convolution. It is seldom possible to derive h from an analysis of the elastic properties of the platform that carries the shear probes. Instead, the impulse response is usually estimated from the simultaneous measurements of $\hat{s}$ and a.

The Fourier transform of (2) is

$$\hat{S}(k) = S(k) + H(k)A(k) + E(k) \tag{3}$$

where the uppercase symbols are the transforms of their lowercase versions and k is the wavenumber. Multiplying both sides of (3) by their respective complex conjugates gives

$$\begin{aligned}\hat{S}\hat{S}^* &= SS^* + HH^*AA* + EE^* + \\ &2\Re(SH^*A^*) + 2\Re(SE^*) + \Re(HAE^*)\end{aligned} \tag{4}$$

where the operator $\Re$ extracts the real part. Averaging (4) reduces the last three terms zero because there should be no relationship between oceanic shear and platform vibrations (SH^*A^*), and between the added noise and either the oceanic shear (SE^*) or the vibrations (HAE^*). Thus, the spectrum of the measured shear is

$$C_{\hat{s}\hat{s}} = C_{ss} + |H|^2 C_{aa} + C_{ee} \tag{5}$$

where each term is implicitly a function of wavenumber, k. An estimate of the transfer function H is provided by multiplying both sides of (3) by A^* and invoking the expectation of the various terms, which yields

$$C_{\hat{s}a} = HC_{aa} \tag{6}$$

where $C_{\hat{s}a}$ is the cross-spectrum of the measured shear and vibration. Substituting into (5) gives

$$C_{\hat{s}\hat{s}} = C_{ss} + \frac{|C_{\hat{s}a}|^2}{C_{aa}} + C_{ee} = C_{ss} + C_{\hat{s}\hat{s}}\,\Gamma_{\hat{s}a} + C_{ee} \tag{7}$$

where

$$\Gamma^2_{\hat{s}a} = \frac{|C_{\hat{s}a}|^2}{C_{aa}C_{\hat{s}\hat{s}}} \tag{8}$$

is the coherency between the vibration and shear-probe signals. Thus, the best estimate of the oceanic shear spectrum is

$$C_{ss}(k) = \left(1 - \Gamma^2_{\hat{s}a}(k)\right) C_{\hat{s}\hat{s}}(k) - C_{ee}(k) \tag{9}$$

where the spectrum of the measurement noise, $C_{ee}(k)$ is estimated by other means or ignored because it is usually small. The multivariate method of Goodman et al. (2006) adjusts downward the spectrum of all shear-probe signals using all vibration sensors simultaneously, and takes into account the partial coherency between vibration signals. But, it does rely on an estimate of the coherency, which is always positive ($0 < \Gamma < 1$) for any estimate that uses a finite length of data, and will reduce the spectrum of shear even at wavenumbers at which there are no vibrations.

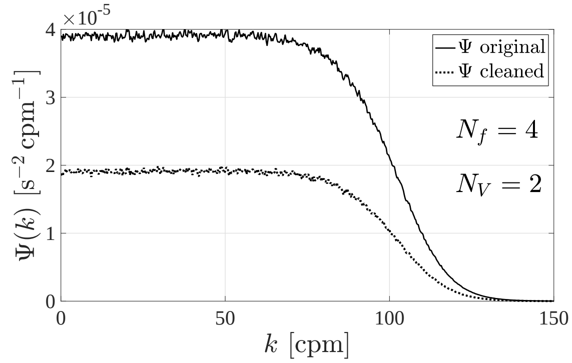

FIG. 1. The average of 2002 spectra of low-pass filtered white Gaussian noise (solid line) and the average of the same spectra after they have been cleaned using $N_V = 2$ incoherent vibration signals (dotted line), using spectra formed from $N_f = 4$ fft-segments.

3. Spectral Bias Examples

In this section we demonstrate the bias induced by vibration-coherent noise removal, how this bias depends on the number of fft-segments that are used to estimate a shear spectrum and on the number of vibration signals that are used to clean a shear spectrum.

The statistical reliability of coherency estimates improves with increasing length of data used to make this estimate. In particular, for a finite length of data the estimate improves with the number of fft-segments that are used. We calculate spectra and cross-spectra using the periodogram method. The span of data used to estimate a spectrum is decimated into N_f segments that overlap by 50 %. Each of these segments is de-trended and then tapered with a cosine window before it is transformed with a fast Fourier transform. The first half (plus one) of the transforms are retained. For auto spectra, the transform magnitudes are squared and added to an accumulator. For cross-spectra, the transform is multiplied by the complex conjugate of the transform of the other signal, and added to an accumulator. After the last transformed segment is added to the accumulator, the values are divided by the number of segments used, N_f, to produce the average periodogram. This periodogram is then scaled to produce a spectrum. This spectrum has the property that its integral over all wavenumbers (approximated by, for example, the trapezoidal method) equals the variance of the original signal, or the co-variance of the original pair of signals.

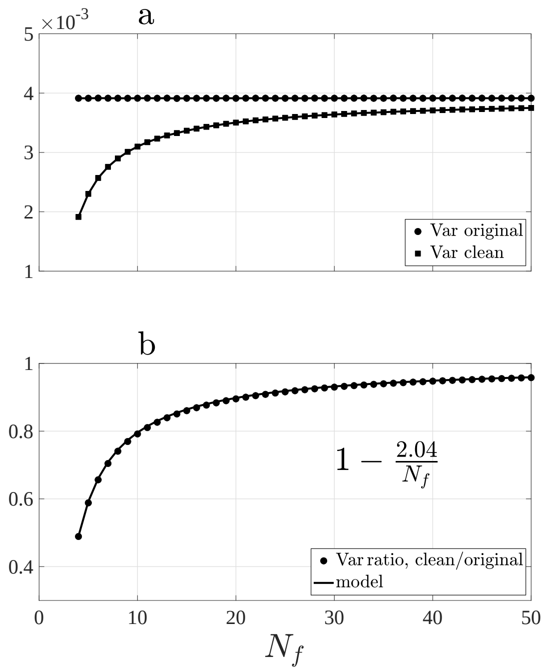

FIG. 2. (a) — The variance of the Gaussian noise signals (disks) (determined by spectral integration) and the variance of the same signal after the removal of vibration-coherent noise (squares), as a function of the number of fft-segments, N_f, used to estimate the spectra. (b) — The ratio of the variance of the cleaned and original signals (disks) and a mathematical model of their ratio. Th number of vibration signals is $N_V = 2$.

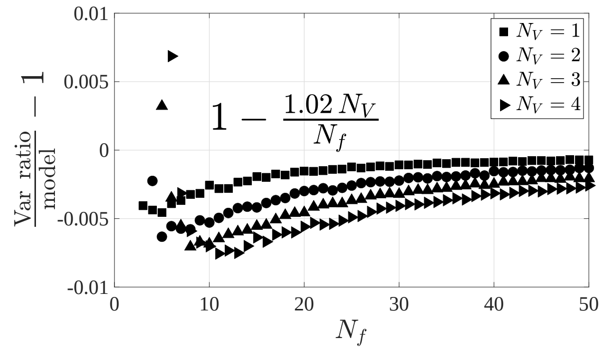

FIG. 3. The ratio of the variance derived from the cleaned and the original spectra of Gaussian white noise signals (Var ratio), divided by the mathematical model of (11) less 1, as a function of the number of fft-segments, N_f, used to estimate the spectra, for $N_V = 1$ to 4 vibration signals.

a. White noise

Two shear-probe signals that are completely unrelated to vibration signals are simulated using independent Gaussian random sequences, with a sampling rate of $512\,\mathrm{m}^{-1}$. These signals are low-pass filtered at 98 cpm to simulate a typical data acquisition system, using a cascade of two 4-pole Butterworth filters. All auto- and cross-spectra are estimated using fft-segments of length 1024 samples (i.e., 2 m). The number of fft-segments used to estimate the spectra is varied from $N_f = 3$ to 50. The calculations are repeated 1001 times to determine the average spectrum for a particular N_f. These calculations are repeated using $N_V = 1$ to 4 vibration signals. The vibration-coherent noise is removed from the shear spectra to produce 'cleaned' spectra, even though there is no coherency between the vibration and shear-probe signals because they are independent Gaussian sequences.

The average cleaned shear spectrum is reduced by a factor of approximately 2 when the number of fft-segments is only four and the number of vibration signals is two (Fig. 1). The variance of the original shear signals is determined by integrating the spectrum over all wavenumbers and equals $3.9\times10^{-3}\,\mathrm{s}^{-2}$ for all choices of N_f (Fig. 2a disks). In contrast, the variance determined from an integration of the cleaned spectra is smaller (Fig. 2a squares). The ratio of the variance determined from the cleaned and original spectra (Fig. 2b) is approximately

$$R = 1 - \frac{2.04}{N_f} \ . \tag{10}$$

The ratio of variance (10) is independent of the variance of the shear-probe and the vibration signals, and is also independent of the number of shear-probe signals that are cleaned simultaneously by the vibration-coherent noise removal algorithm. However, the results do depend on the number of vibration signals, N_V, that are used to clean the shear-probe spectra. A model expression that fits the simulations to within better than ± 0.75 %, for $N_V \le 4$, is

$$R = 1 - \frac{1.02 N_V}{N_f} \tag{11}$$

(Fig. 3). That is, (11) provides a model for the amount by which the vibration-coherent noise removal algorithm reduces the spectrum of a shear-probe measurement at all wavenumbers, due to estimating the spectrum with a finite number of fft-segments, N_f.

b. Coherent sinusoidal signals

Real platform vibrations tend to be quasi-sinusoidal. Real vibrations were simulated by adding 4 sinusoidal components to the Gaussian vibration signal. A fraction of these vibrations were also added to the shear-probe signals to produce strongly contaminated signals (Fig. 4b). The vibration-coherent noise removal algorithm readily removes the sinusoidal components from the shear-probe signals even when $N_f = 3$ (Fig. 4a). However, the entire shear-probe spectra are still biased low by the factor given by (11). For example, the wavenumber independent part of the original shear spectra is $4\times10^{-5}\,\mathrm{s}^{-2}\,\mathrm{cpm}^{-1}$ (Fig. 4b), while the same range of the cleaned spectra is $1.3\times10^{-5}\,\mathrm{s}^{-2}\,\mathrm{cpm}^{-1}$ (Fig. 4a), i.e. three time smaller.

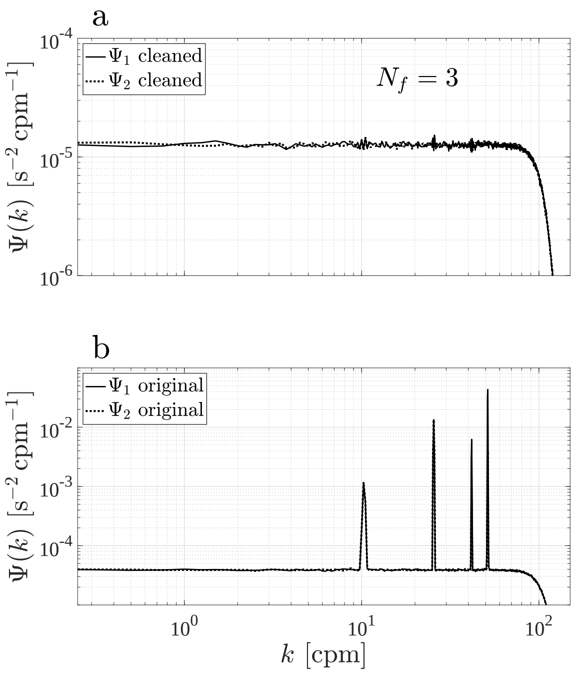

FIG. 4. (a) The average spectra for shear-probes 1 and 2 after cleaning using $N_f = 3$ and $N_V = 2$. (b) The same spectra before cleaning.

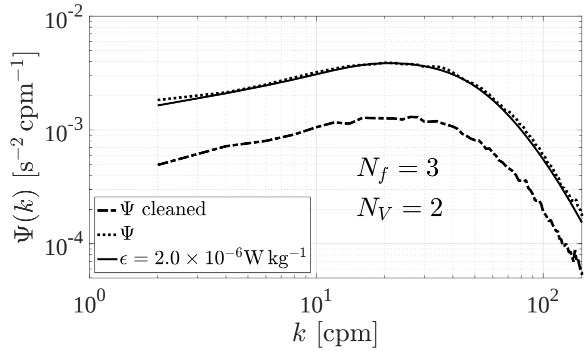

FIG. 5. The average of 915 spectra from each of 4 shear probes in a tidal channel (dotted line) using $N_f = 3$ fft-segments, and the average spectrum after the shear-probe signals were cleaned using $N_V = 2$ vibration signals (dot-dash line), and the spectral model for shear of Lueck (2021b) (solid line).

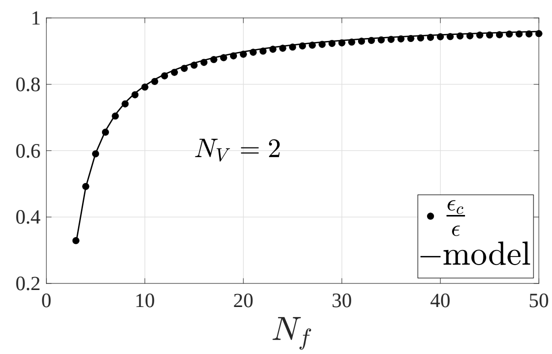

FIG. 6. The ratio of the rate of dissipation of kinetic energy estimated using the clean spectra, ϵ_c, and the original spectra, ϵ, for spectra derived from N_f fft-segments (disks), and the model of (11) (line).

c. Real shear-probe signals

Shear-probe signals are not Gaussian white noise. They have a spectrum that rises with increasing wavenumber, reaches a peak and diminishes rapidly with increasing wavenumber due to viscosity. Does the model of (11) hold for real oceanic turbulence signals? We use 915 m of data collected in a tidal channel that have been processed according to Lueck (2021b) to demonstrate the effects of vibration-coherent noise removal using real shear-probe and vibration data. The length of the fft-segments is 0.5 m. For these data, the shear signals are strong while the platform vibrations are weak. That is, there is very little vibration-induced contamination in the four shear-probe signals. The instrumentation had two vibration sensors mounted close to its four shear probes. When using only $N_f = 3$ fft-segments, the average spectrum of the shear-probe signals is reduced by a factor of approximately three just as was found for white noise signals (Fig. 5). The ratio of the four-probe average rate of dissipation, calculated using (1) and a value $k_c = 150$ cpm, for the cleaned and original spectra follows closely the model of (11) (Fig. 6). The modelled dissipation rate is higher than the observed rate by less than 0.5 % for $N_f > 20$.

4. Discussion and conclusions

The technique of vibration-coherent noise removal is not specific to just shear probes and vibration sensors. It can be used to clean any signal for which there is a simultaneous measure of the source of contamination. The measure of the contamination only needs to be linearly related to the actual contamination of the signal that one wishes to clean.

When the data segments that are used to estimate a spectrum are short (i.e., the number of fft-segments used to form the spectrum, N_f, is small), the bias introduced by the coherent noise removal technique is quite large, but it is wavenumber (or frequency) independent. Cleaned spectra should be corrected (boosted) using the model of (11).

The degrees of freedom of a spectral estimate (derived using the periodogram technique) is approximately $2N_f$ (Nuttall 1971). The degrees of freedom of a cleaned spectrum is smaller by $2N_V$, because that many degrees of freedom are consumed in the estimation of the coherency. Therefore, one should never use fewer than $3 + N_V$ fft-segments, and preferably many more.

The 95 % confidence interval of a spectral estimate of turbulent shear is

$$\exp\left(\pm 1.96\sigma_{\ln\Psi}\right) \tag{12}$$

where

$$\sigma^2_{\ln \Psi} = \frac{5}{4} N_f^{-7/9} \tag{13}$$

(Lueck 2021a). When estimating the statistical uncertainty of a spectrum, the value of N_f in (13) must be reduced by the number of vibration signals, N_V, that are used to clean the spectrum. That is, (13) should be replaced by

$$\sigma^2_{\ln \Psi} = \frac{5}{4} \left(N_f - N_V\right)^{-7/9} \tag{14}$$

when estimating the confidence limits of a shear spectrum that has been cleaned. Although the bias induced by spectral cleaning is significant, it is always smaller than the 95 % confidence interval of a turbulent shear spectrum.

However, the spectral bias should not be ignored for two reasons. Integrating the spectrum to obtain the shear variance and, hence, an estimate of the rate of dissipation, produces estimates with a much tighter confidence interval than (12). However, integrating the spectrum does not reduce the bias. When trying to obtain high spatial resolution of dissipation estimates it is necessary to estimate spectra using very few fft-segments. The bias will be large (a factor of 3) such seen in Fig. 9e in Wijesekera et al. (2020) where the authors used 1 s of data and fft-segments of 0.5 s to estimate their spectra, that is $N_f = 3$ and $N_V = 2$. However, in many cases the bias has largely gone unnoticed because typical spectral estimates use $N_f \geq 9$ and $N_V = 2$ for which the bias is less than ≈20 %, while the confidence interval range from 0.36 to 2.8.

Acknowledgments. The tidal channel data were supplied by Alex Hay, of Dalhousie University.

Data availability statement. The data used to create Fig. 5 and 6 are available from the corresponding author.

References

Goodman, L., E. R. Levine, and R. G. Lueck, 2006: On measuring the terms of the turbulent kinetic energy budget from an AUV. *J. Atmos. Oceanic Technol.*, **23**, 977–990, doi:10.1175/JTECH1819.1.

Levine, E. R., and R. G. Lueck, 1999: Turbulence measurements from an autonomous underwater vehicle. *J. Atmos. Oceanic Technol.*, **16**, 1533–1544.

Lueck, R. G., 2021a: The statistics of oceanic turbulence measurements: Shear spectra and a new spectral model. *J. Phys. Oceanogr.*, **–**, submitted, doi:--.

Lueck, R. G., 2021b: The statistics of oceanic turbulence measurements: Shear variance and dissipation rates. *J. Phys. Oceanogr.*, **–**, submitted, doi:--.

Nuttall, A., 1971: Spectral estimation by means of overlapped fast fourier transform processing of windowed data. NUSC Tech. Rep. 4169, [Available online at https://apps.dtic.mil/dtic/tr/fulltext/u2/a182402.pdf], Naval Underwater Systems Center.

Osborn, T. R., 1974: Local vertical profiling of velocity microstructure. *J. Phys. Oceanogr.*, **4**, 109–115, doi:10.1175/1520-0485(1974)004.

Osborn, T. R., 1980: Estimates of the local rate of vertical diffusion from dissipation measurements. *J. Phys. Oceanogr.*, **10**, 83–89, doi: 10.1175/1520-0485(1980)010.

Pope, S. B., 2009: *Turbulent Flows*. Cambridge University Press.

Siddon, T. E., 1971: A miniature turbulence gauge for utilizing aerodynamic lift. *Review of Scientific Instruments*, **42**, 653–656, doi: 10.1063/1.1685193.

Siddon, T. E., and H. S. Ribner, 1965: An aerofoil probe for measuring the transverse component of turbulence. *J. American Inst. Aeronautics and Astronautics*, **3**, 747–749, doi:10.2514/3.2963.

Taylor, G., 1935: Statistical theory of turbulence. *Proc. R. Soc. A*, **151**, 421–444, doi:10.1098/rspa.1935.0158.

Wijesekera, H. W., D. W. Wang, and E. Jarosz, 2020: Dynamics of the diurnal warm layer: surface jet, high frequency internal waves, and mixing. *J. Phys. Oceanogr.*, **50**, 2053 – 2070, doi: 10.1175/JPO-D-19-0285.1.

End of document