applied optics

Eliminating photon noise biases in the computation of second-order statistics of lidar temperature, wind, and species measurements

Chester S. Gardner[1,3] and Xinzhao Chu[2,*]

[1]Department of Electrical and Computer Engineering, University of Illinois, Urbana, Illinois 61801, USA
[2]Cooperative Institute of Research in Environmental Sciences and Department of Aerospace Engineering Sciences, University of Colorado Boulder, 216 UCB, CIRES, Boulder, Colorado 80309, USA
[3]e-mail: cgardner@illinois.edu
*Corresponding author: xinzhao.chu@colorado.edu

Received 17 June 2020; revised 6 August 2020; accepted 11 August 2020; posted 11 August 2020 (Doc. ID 400375); published 14 September 2020

The precision of lidar measurements is limited by noise associated with the optical detection process. Photon noise also introduces biases in the second-order statistics of the data, such as the variances and fluxes of the measured temperature, wind, and species variations, and establishes noise floors in the computed fluctuation spectra. When the signal-to-noise ratio is low, these biases and noise floors can completely obscure the atmospheric processes being observed. We describe a novel data processing technique for eliminating the biases and noise floors. The technique involves acquiring two statistically independent datasets, covering the same altitude range and time period, from which the various second-order statistics are computed. The efficacy of the technique is demonstrated using Na Doppler lidar observations of temperature in the upper mesosphere and lower thermosphere acquired recently at McMurdo Station, Antarctica. The results show that this new technique enables observations of key atmospheric parameters in regions where the signal-to-noise ratio is far too low to apply conventional processing approaches.

https://doi.org/10.1364/AO.400375

1. INTRODUCTION

Lidars are now widely employed to measure key atmospheric parameters, such as temperature and winds, a wide variety of atmospheric constituents and aerosols, and dynamical features, like tides, gravity waves, and turbulence, from the troposphere into the lower thermosphere [1]. Rayleigh and resonance fluorescence lidars are providing considerable insight on the thermal and wind structure of the mesosphere and lower thermosphere (MLT), the morphology of the meteoric metal layers, and even the influx of cosmic dust into the Earth's upper atmosphere [2–9]. Many of the key observations involve the derivation of second-order statistics from the lidar measurements, including the variances, fluxes (i.e., covariances), and spectra of the wave- and turbulence-induced fluctuations in temperature, winds, and minor species. These statistics are all affected by noise associated with the optical detection process, which can be quite significant under low signal-to-noise ratio (SNR) conditions. Photon noise biases the calculated fluctuation variances and fluxes and introduces noise floors into the computed spectra. Over the years, numerous researchers have employed a variety of techniques to compensate the data for the photon noise [10–12], but these approaches are less effective when the signal levels are low and the species concentrations are small.

In this paper, we demonstrate a novel technique for eliminating the photon noise biases from the derived second-order statistics of the lidar observations. The signal processing technique exploits the fact that the white photon noise fluctuations contaminating the data samples from adjacent range and time bins are statistically independent, while the red wave-driven and turbulence-induced fluctuations are highly correlated. The technique was originally proposed to measure turbulence in the upper mesosphere with Na and Fe lidars [13], and its efficacy was subsequently demonstrated using Na Doppler lidar data acquired at Cerro Pachón, Chile [14]. Turbulence observations in the mesopause region represent an extreme case of low SNR because the measurement range is large (80–100 km), and the required resolutions are small (several seconds and several tens of meters). Thus, the photon counts in each measurement interval are very small, resulting in a very low SNR. Here, we build upon this previous work by fully describing how to apply the technique, analyzing its performance theoretically, and demonstrating how it is adapted for measuring variances of the temperature and lapse rate fluctuations caused by gravity

waves and for measuring the vertical heat flux and temperature fluctuation spectra.

We focus on wave-induced temperature variations because recent work has shown that the variances of the temperature and lapse rate fluctuations are directly related to wave transport of important minor species in the upper atmosphere such as atomic oxygen, NO_x, and meteoric metals [15]. Furthermore, the growth of the temperature variance with increasing altitude, which is related to wave dissipation mechanisms, provides important insight on vertical coupling from the lower to the upper atmosphere [12] and on wave amplitudes in the thermosphere that affect space weather [16]. Additional observations of these key parameters, especially at higher altitudes where signal levels are typically low, are needed from around the globe to fully understand how gravity waves generated in the lower atmosphere impact constituent concentrations in the Earth's upper atmosphere and drive the variability of space weather. In addition to showing how temperature data can be processed to minimize the impact of photon noise, we also show how the results can be adapted in a straightforward manner to measurements of the second-order statistics of wind, wind shear, species concentration, and turbulence fluctuations.

2. ESTIMATING THE TEMPERATURE VARIANCE

We assume the raw photon count data are processed to derive samples of the temperature (T) and wind ($w =$ vertical, $u =$ zonal, $v =$ meridional) at a vertical resolution Δz and a temporal resolution Δt. The fluctuations induced by waves (T', w', u', v') are derived by subtracting the temporal means or linear trends in time from the data obtained during each observation period. The resulting data samples ($T' + \Delta T, w' + \Delta w, u' + \Delta u, v' + \Delta v$) include both the wave-driven fluctuations plus noise associated with the photon counting process ($\Delta T, \Delta w, \Delta u, \Delta v$).

T', w', u', and v' are approximately Gaussian distributed, zero-mean random variables [17]. ΔT, Δw, Δu, and Δv are Poisson distributed, zero-mean random variables, but, in most cases, when the signal levels are large enough to derive scientifically useful temperature and wind measurements, the Poison distribution can be approximated by a Gaussian distribution. The wave fluctuations are statistically independent of the photon noise. Depending on the laser frequencies used to make the temperature and wind observations and the strength of the lidar signals, ΔT, Δw, Δu, and Δv may be weakly correlated [17,18].

For simplicity, we focus on the statistics of the temperature fluctuations, as the final results are easy to adapt to the wind fluctuations. Normally, the temperature variance at each altitude is estimated by computing the sample variance (square of the fluctuations) averaged over the observation period, which is given by

$$s^2_{T'+\Delta T}(z) = \frac{1}{L}\sum_{l=1}^{L}\left[T'(z, t_0 + l\Delta t) + \Delta T(z, t_0 + l\Delta t)\right]^2. \tag{1}$$

The altitude z and time t are integer multiples of Δz and Δt, respectively, and $L = \tau_{\text{obs}}/\Delta t$ is the number of samples, where τ_{obs} is the observation period. τ_{obs} varies from a few hours to about 12 h for lidars restricted to just nighttime observations. However, the observation time can be considerably longer during polar winters. To simplify the subsequent notation, we adopt the convention that time is always measured relative to the beginning of the observation period so that $t_0 = 0$. Because the wave-driven temperature fluctuations and the photon noise uncertainties are statistically independent, the expected value of the sample variance is just the sum of the variances of T' and ΔT. Specifically, if we neglect the small uncertainties in the sample mean or linear trend used to compute the data samples ($T' + \Delta T$), then it is easy to show that

$$\left\langle s^2_{T'+\Delta T}(z)\right\rangle = \frac{1}{L}\sum_{l=1}^{L}\left\langle\left[T'(z, l\Delta t) + \Delta T(z, l\Delta t)\right]^2\right\rangle$$

$$= \text{Var}\left[T'(z)\right] + \text{Var}\left[\Delta T(z)\right], \tag{2}$$

where the angle brackets denote expectation with respect to the probability ensemble

$$\text{Var}\left[T'(z)\right] = \left\langle\left[T'(z)\right]^2\right\rangle, \tag{3}$$

and

$$\text{Var}\left[\Delta T(z)\right] = \left\langle\left[\Delta T(z)\right]^2\right\rangle. \tag{4}$$

The right-hand side of Eq. (2) is obtained because we assume that the wave-driven fluctuations are stationary, at least over the observation period. We also assume that the fluctuations arise from an ergodic random process, where the sample averages can be used to approximate the corresponding ensemble averages. Therefore, in the limit as the number of samples (L) goes to infinity, the sample averages approach the ensemble averages. For a finite number of samples, the uncertainty of the sample variance, i.e., the mean-square difference between the sample and ensemble variances, is related to observation period $L\Delta t$ and the T' correlation time [see Eq. (15) below].

In many cases, when the lidar SNR is low, the photon noise variance is a non-negligible fraction of the wave-induced fluctuation variance, especially for metal lidars at the edges of the MLT layers, where the metal densities are low, and, for Rayleigh lidars in the upper stratosphere and mesosphere, where the atmospheric density is low. In principle, estimates of the photon noise variance can be calculated using the measured photon counts [2,18] and then subtracted from the sample variance to obtain an unbiased estimate of the wave variance:

$$\text{Var}\left[T'(z)\right] \simeq s^2_{T'+\Delta T}(z) - \text{Var}\left[\Delta T(z)\right]. \tag{5}$$

Although Eq. (5) is valid provided T' and ΔT are statistically independent, the challenge is deriving an accurate estimate of $\text{Var}(\Delta T)$. For example, if the precise value of ΔT could be determined for each measurement, which of course is not possible, the relative uncertainty of the calculated sample variance of ΔT is $\sqrt{2\Delta t/\tau_{\text{obs}}}$. For $\Delta t = 2.5$ min and $\tau_{\text{obs}} = 8$ h, which are typical values for resonance fluorescence lidars, the relative uncertainty is still about 10%. In practice, ΔT is estimated from

the measured signal count by linearizing the nonlinear system equations. Under low SNR conditions, these nonlinearities can significantly increase the uncertainties in the estimated values of ΔT and, hence, the estimated value of Var(ΔT). In addition, variations in signal level caused by variations in atmospheric transmittance and densities also introduce considerable uncertainty into the photon noise calculations, especially when the temperature error (ΔT) is comparable to the wave-induced perturbations (T'). Consequently, in practice, it is difficult to accurately compensate for the photon noise bias when the T' variance is estimated according to Eq. (5) [10,12,18].

The photon noise bias can be eliminated if we compute the variance using two statistically independent measurements of the $T' + \Delta T$ time series, such as employing two lidars to probe the same volume of the atmosphere at the same time. Obviously, this is not a cost-effective solution. Alternatively, the T' variance can also be estimated by recognizing that the wave-induced fluctuations at adjacent time (and altitude) measurement intervals are highly correlated because their temporal frequency (ω) and vertical wavenumber (m) spectra are red (i.e., larger-scale waves dominate the energy spectra), while the white photon noise is uncorrelated. The simplest approach for acquiring two independent $T' + \Delta T$ time series is to use the odd numbered temperature measurements as one time series and the even numbered measurements as the second. These two time series are nearly identical but displaced in time by the temporal resolution Δt, where the odd series leads the even series. Because the photon noise contaminating the odd samples is statistically independent of T' and also independent of the noise contaminating the even samples, the average value of the product of the two time series is approximately equal to the temperature covariance function, $\mathrm{Cov}_{T'}(\Delta t)$. For sufficiently small Δt, the covariance is approximately equal to Var(T'). Similarly, independent measurements of $T' + \Delta T$ can also be acquired by using the odd range bins for one set of time series and the even bins for the second set. The average value of the product of the time series from two adjacent altitudes is approximately equal to $\mathrm{Cov}_{T'}(\Delta z)$.

However, there is a more elegant way to derive independent estimates of the temperature versus time using measurements made by a single lidar. Modern atmospheric lidars typically acquire the raw photon count data at temporal (δt) and vertical (δz) resolutions that are much smaller than the required resolutions of the processed scientific data (Δt and Δz). For example, the University of Colorado lidar group acquired raw photon count data with an Fe Boltzmann lidar at fundamental resolutions of $\delta t = 1$ min and $\delta z = 48$ m and with a Na Doppler lidar at $\delta t = 0.5$ min and $\delta z = 96$ m, but they derived temperature and Fe/Na density profiles at much lower resolutions of $\Delta t = 15$ min and $\Delta z = 960$ m [19]. Specifically, for the Na lidar, these authors combined 30 consecutive photon count samples, each acquired by integrating the Na signals over 0.5 min, to compute estimates of T every 15 min. If, instead, suppose two estimates of T are derived, one using the 15 odd photon count samples (T_{odd}) and a second using the 15 even samples (T_{even}), these two times series are statistically independent of each other because they are derived from different but interleaved photon count samples. The resolution of the two temperature time series ($\Delta t - \delta t$) is still approximately $\Delta t = 15$ min, but the temperatures are offset by $\delta t = 0.5$ min, where T_{odd} leads T_{even} by δt. Of course, the integrated signal counts drop by half so the temperature uncertainties for both T_{odd} and T_{even} increase by $\sqrt{2}$. More importantly, the temperature error caused by photon noise computed with the odd samples is uncorrelated with the temperature error computed using the even samples. This interleaved data processing approach is also equivalent to having two independent lidars, each probing the same volume of the atmosphere but at different, multiple, interleaved time periods, offset by δt. Thus, the atmospheric temperatures measured by the "two lidars" are highly correlated, but their uncertainties are statistically independent. A temperature profile, with the required measurement precision and resolution Δt, can still be derived by using the standard processing method or by computing $T = (T_{\mathrm{odd}} + T_{\mathrm{even}})/2$. However, the T' variance is now estimated by computing the covariance between T'_{odd} and T'_{even}. In this case the sample T' covariance and its expected value are

$$\rho_{T'+\Delta T}(\delta t) = \frac{1}{L}\sum_{l=1}^{L}\left[T'(z, l\Delta t) + \Delta T(z, l\Delta t)\right]_{\mathrm{odd}} \times \left[T'(z, l\Delta t) + \Delta T(z, l\Delta t)\right]_{\mathrm{even}}, \tag{6}$$

and

$$\langle \rho_{T'+\Delta T}(\delta t)\rangle = \frac{1}{L}\sum_{l=1}^{L}\left\langle T'(z, l\Delta t)\, T'(z, l\Delta t + \delta t)\right\rangle = \mathrm{Cov}_{T'}(\delta \mathrm{t}), \tag{7}$$

where the temperature covariance function is related to its power spectrum $F_{T'}(\omega)$:

$$\mathrm{Cov}_{T'}(\delta t) = \frac{1}{2\pi}\int F_{T'}(\omega)\exp(i\omega\delta t)\, \mathrm{d}\omega. \tag{8}$$

As before, Eq. (7) follows because we assume T' is a stationary, ergodic random process and because ΔT and T' are uncorrelated. Therefore, in the limit as the number of samples (L) goes to infinity, the sample temperature covariance given by Eq. (6) approaches the ensemble temperature covariance given by the right-hand side of Eq. (7).

Since the temperature ω spectrum is roughly proportional to ω^{-p} with $p \sim 5/3 - 2$ (see [20] and references therein), and δt is typically small compared to the correlation time of the wave fluctuations [see Eq. (16) below], $\mathrm{Cov}_{T'}(\delta t)$ is approximately equal to the variance. This can be quantified by expanding the covariance function given by Eq. (8) in a power series in terms of δt whenever $|\omega\delta t| < 1$ for the important values of ω and then retaining just the first three terms in the series:

$$\mathrm{Cov}_{T'}(\delta t) \simeq \frac{1}{2\pi}\int F_{T'}(\omega)\left[1 + i\omega\delta t - \frac{(\omega\delta t)^2}{2}\right]\mathrm{d}\omega = \mathrm{Var}\left[T'(z)\right] - \frac{\delta t^2}{2}\mathrm{Var}\left[\frac{\partial T'(z)}{\partial t}\right], \tag{9}$$

where

$$\mathrm{Var}\left[T'(z)\right] = \frac{1}{2\pi}\int F_{T'}(\omega)\mathrm{d}\omega = \mathrm{Cov}_{T'}(0), \tag{10}$$

and

$$\mathrm{Var}\left[\frac{\partial T'(z)}{\partial t}\right] = \frac{1}{2\pi}\int \omega^2 F_{T'}(\omega)\mathrm{d}\omega. \tag{11}$$

In evaluating the integral in Eq. (9), we note that the power spectrum is an even function of ω so that the odd term in ω integrates to zero. However, gravity wave spectra are typically modeled as single-sided spectra, where the integrals are carried out over just the positive wavenumbers and frequencies (see Appendix A). By combining Eqs. (6), (7), and (9), we obtain

$$\mathrm{Var}\left[T'(z)\right] \simeq \rho_{T'+\Delta T}(\delta t) + \frac{\delta t^2}{2}\mathrm{Var}\left[\frac{\partial T'(z)}{\partial t}\right]. \tag{12}$$

Notice that according to Eq. (12), the photon noise bias has been eliminated from the T' variance when it is computed using the sample covariance, but it is replaced by another term, which is related to the variance of $\partial T'/\partial t$. This term, which accounts for the slightly smaller value of the sample covariance compared to the variance, can be easily estimated from the data set by calculating the sample variance of $\partial T'/\partial t$ (see Section 4), or it can be evaluated by using a model for the temperature spectrum (see Appendix A). If the temperature spectrum is proportional to ω^{-2}, then the correction term is given by [see Eq. (A15) in Appendix A]

$$\frac{\delta t^2}{2}\mathrm{Var}\left[\frac{\partial T'(z)}{\partial t}\right] = \frac{\delta t^2}{2}\omega_{\max}\omega_{\min}\mathrm{Var}\left[T'(z)\right], \tag{13}$$

where $\omega_{\max} = \min(N, \pi/\Delta t)$ is the highest observed gravity wave frequency, and $\mathrm{N} = 2\pi/\tau_B$ ($\tau_B \approx 5$ min) is the buoyancy frequency. $\omega_{\min} = \max(f, 2\pi/\tau_{obs})$ is the lowest observed gravity wave frequency, and $f = 2\pi/\tau_i$ ($\tau_i = 12\,\mathrm{h}/\sin\theta$, $\theta =$ latitude) is the inertial frequency. By substituting Eq. (13) intoEq. (12) and rearranging terms, we obtain this final expression for the temperature variance:

$$\begin{aligned}\mathrm{Var}\left[T'(z)\right] &\simeq \rho_{T'+\Delta T}(\delta t) + \frac{\delta t^2}{2}\omega_{\max}\omega_{\min}\mathrm{Var}\left[T'(z)\right] \\ &= \frac{\rho_{T'+\Delta T}(\delta t)}{\left[1 - \delta t^2\omega_{\max}\omega_{\min}/2\right]}.\end{aligned} \tag{14}$$

The correction term can be made arbitrarily small by acquiring and processing the raw photon count data with a very small fundamental resolution δt so that the sample covariance will be nearly equal to the temperature variance. For example, the largest gravity wave frequency is N, so if we choose $\Delta t = \tau_B/2 = 2.5$ min, acquire and process the data so that $\delta t = 15$ s, and if $\tau_{\mathrm{obs}} = 8$ h, the correction term given by Eq. (13) is only 0.05% $\mathrm{Var}(T')$ and can be neglected. This example illustrates the major advantage of interleaving the raw photon counts. If we had simply used the odd and even temperature measurements to compute the temperature covariance, the correction term would be $(\Delta t/\delta t)^2 = 100$ times larger and could not be neglected. As will be shown in Section 4, this method for acquiring the raw data and computing the temperature covariance is especially useful for estimating the T' variance in high photon noise scenarios where the lidar SNR is small.

The uncertainty (precision) of the estimated temperature variance is dominated by the uncertainty of the sample covariance not the uncertainty of the small correction term. The uncertainty of the sample covariance can be computed by following the approach described in the appendix of Ref. [17] and summarized here in Appendix B. The final result for the uncertainty associated with estimating the temperature variance from the sample covariance function is

$$\begin{aligned}&\Delta\rho_{T'+\Delta T}(\delta t) \\ &\simeq \sqrt{\frac{2\tau_{T'}}{\tau_{\mathrm{obs}}}\mathrm{Var}^2(T') + \frac{\Delta t}{\tau_{\mathrm{obs}}}\left[2\mathrm{Var}\,(T')\,\mathrm{Var}\,(\Delta T) + \mathrm{Var}^2\,(\Delta T)\right]},\end{aligned} \tag{15}$$

where (see Appendix A)

$$\tau_{T'} = \frac{\frac{1}{4\pi}\int_{\omega_{\min}}^{\omega_{\max}} F_{T'}^2(\omega)\mathrm{d}\omega}{\mathrm{Var}^2\,(T')} \simeq \frac{\pi}{3\omega_{\min}}. \tag{16}$$

The right-hand side of Eq. (16) was evaluated assuming $\mathrm{F}_{T'}(\omega) \propto \omega^{-2}$ and $\omega_{\min} \ll \omega_{\max}$. $\tau_{T'}$ is the correlation time of the measured temperature fluctuations, which varies between about 1–2 h, depending on the gravity wave activity, resolution, and length of the observation period, which determines what portion of the gravity wave spectrum is observed. $\tau_{T'}$ could also be derived from the dataset using Eq. (16), the measured $\mathrm{Var}(T')$, and the measured ω spectrum of T'. Because $\Delta t \ll \tau_{T'}$, the uncertainty in the T' variance estimate is usually dominated by the statistical noise associated with the calculation of the sample covariance not photon noise. In this case, the uncertainty is simply $\sqrt{2/k}\mathrm{Var}(T')$, where $k = \tau_{\mathrm{obs}}/\tau_{T'}$ is the number of statistically independent observations of T' used to estimate the covariance. For typical values of $\mathrm{Var}(\Delta T)$ and Δt, the effects of photon are negligible. However, when observations are made in regions where the SNR is very low so that $\mathrm{Var}(T') \ll \mathrm{Var}(\Delta T)$, the photon noise will increase the uncertainty of the estimated temperature covariance (see Section 4), even though the large noise bias has been eliminated.

For comparison, the uncertainty of the sample variance given by Eq. (1) is (see Appendix B)

$$\begin{aligned}&\Delta s^2_{T'+\Delta T} \\ &\simeq \sqrt{\frac{2\tau_{T'}}{\tau_{\mathrm{obs}}}\mathrm{Var}^2\,(T') + \frac{2\Delta t}{\tau_{\mathrm{obs}}}\left[2\mathrm{Var}\,(T')\,\mathrm{Var}\,(\Delta T) + \mathrm{Var}^2\,(\Delta T)\right]}.\end{aligned} \tag{17}$$

It is important to note that $\mathrm{Var}(\Delta T)$ in Eq. (17) is half the value of that in Eq. (15) because, for the conventional method, T is estimated from the full dataset while, for the interleaved method, T_{odd} and T_{even} are each computed from half the data. Notice also that the uncertainties in the estimated temperature covariance and variance given by Eqs. (15) and (17), computed during a single observation period of length $\tau_{\mathrm{obs}} \sim 8-12$ h, are quite large. This is because the T' correlation time ($\tau_{T'}$) is fairly long so that for a single observation period only a small number of statistically independent samples of T' are obtained. To derive estimates of the temperature variance with uncertainties $\sim 20\%$ $\mathrm{Var}(T')$ or less with either data processing approach, one must

average the covariance or variance calculated using $k = 50$ statistically independent samples of T', which, for a correlation time of ~2 h, requires ~100 h of lidar measurements made over 8–12 different observation periods (see Section 5).

3. ESTIMATING THE LAPSE RATE VARIANCE

The fluctuations of the lapse rate ($\partial T'/\partial z$) and vertical shear of horizontal winds ($\partial u'/\partial z$ and $\partial v'/\partial z$) are very important parameters because they are directly related to the stability of the atmosphere [21–23]. They are also among the more difficult parameters to measure, because applying the derivative, which is typically computed as a finite difference, enhances the white photon noise relative to the red temperature and wind fluctuations. We assume the lapse rate is calculated from temperature measurements derived at a vertical resolution of Δz. By taking into consideration the additive photon noise, the estimated lapse rate fluctuation is

$$\frac{\partial\left[T'(z,t)+\Delta T(z,t)\right]}{\partial z} \simeq \frac{T'(z+\Delta z,t)+\Delta T(z+\Delta z)-T'(z,t)-\Delta T(z,t)}{\Delta z}. \tag{18}$$

As before, the conventional approach for computing the lapse rate variance is to simply square the calculated lapse rate fluctuations and average the squared samples over the observation period, which yields

$$s^2_{\partial(T'+\Delta T)/\partial z}(z) = \frac{1}{L}\sum_{l=1}^{L}\left[\frac{T'(z+\Delta z,l\Delta t)+\Delta T(z+\Delta z,l\Delta t)-T'(z,l\Delta t)-\Delta T(z,l\Delta t)}{\Delta z}\right]^2 \simeq \left\langle s^2_{\partial(T'+\Delta T)/\partial z}(z)\right\rangle, \tag{19}$$

where

$$\left\langle s^2_{\partial(T'+\Delta T)/\partial z}(z)\right\rangle = \mathrm{Var}\left[\frac{T'(z+\Delta z)-T'(z)}{\Delta z}\right] + \frac{2Var(\Delta T)}{\Delta z^2} \simeq \mathrm{Var}\left[\frac{\partial T'(z)}{\partial}\right] + \frac{2\mathrm{Var}(\Delta T)}{\Delta z^2}, \tag{20}$$

so that

$$\mathrm{Var}\left[\frac{\partial T'(z)}{\partial z}\right] \simeq s^2_{\partial(T'+\Delta T)/\partial z}(z) - \frac{2\mathrm{Var}(\Delta T)}{\Delta z^2}. \tag{21}$$

The photon noise bias in Eq. (21) can be quite large. For example, for a typical temperature measurement uncertainty of $\Delta T = 2-3\ K$ and vertical resolution $\Delta z = 0.5$ km, the photon noise bias ($32-72\ \mathrm{K^2/km^2}$) is comparable to the actual lapse rate variance ($\sim 30\ \mathrm{K^2/km^2}$) in the mesopause region [24].

Alternatively, to eliminate the photon noise bias, we first compute two, statistically independent, lapse rate estimates using the derived temperatures, $T_{\mathrm{odd}}(z,t)$ and $T_{\mathrm{even}}(z,t)$, described in the previous section. The lapse rate covariance is computed by averaging the product of $\partial T'_{\mathrm{odd}}(z,t)/\partial z$ and $\partial T'_{\mathrm{even}}(z,t)/\partial z$ over the observation period,

$$\rho_{\partial(T'+\Delta T)/\partial z}(\delta t) = \frac{1}{L}\sum_{l=1}^{L}\left\{\frac{\partial\left[T'(z,l\Delta t)+\Delta T(z,l\Delta t)\right]}{\partial z}\right\}_{\mathrm{odd}} \times \left\{\frac{\partial\left[T'(z,l\Delta t)+\Delta T(z,l\Delta t)\right]}{\partial z}\right\}_{\mathrm{even}}, \tag{22}$$

where

$$\left\langle\rho_{\partial(T'+\Delta T)/\partial z}(\delta t)\right\rangle = \frac{1}{L}\sum_{l=1}^{L}\left\langle\frac{\partial T'(z,l\Delta t)}{\partial z}\frac{\partial T'(z,l\Delta t+\delta t)}{\partial z}\right\rangle = \mathrm{Cov}_{\partial T'/\partial z}(\delta t). \tag{23}$$

The estimated lapse rate variance is

$$\mathrm{Var}\left[\frac{\partial T'(z)}{\partial z}\right] \simeq \rho_{\partial(T'+\Delta T)/\partial z}(\delta t) + \frac{\delta t^2}{2}\mathrm{Var}\left[\frac{\partial^2 T'(z)}{\partial t\partial z}\right] = \frac{\rho_{\partial(T'+\Delta T)/\partial z}(\delta t)}{\left[1-(\delta t\omega_{\max}/2)^2/\ln(\omega_{\max}/\omega_{\min})\right]}, \tag{24}$$

where the right-hand side of Eq. (24) is obtained when the lapse rate spectrum is proportional to ω^{-1} (see Appendix A). As before, the expected lapse rate variance does not include contributions from the photon noise, but it does include a term related to higher-order derivatives of the temperature fluctuations. This term accounts for the slightly smaller value of the lapse rate covariance compared to the variance. It can be derived directly from the dataset or evaluated using a model for the temperature spectrum [see Eq. (A10) in Appendix A and Section 4]. However, like the estimated temperature variance, for sufficiently small δt, the correction is negligible regardless of the required resolution. For example, when $\tau_{\mathrm{obs}} = 8$ h, $\Delta t = 2.5$ min, and $\delta t = 15$ s, the correction term is ~0.5%Var($\partial T'/\partial z$) and can be neglected. The uncertainty in the lapse rate fluctuation variance is given by Eq. (15) with T' replaced by $\partial T'/\partial z$ and ΔT replaced by $\partial \Delta T/\partial z$. The lapse rate correlation time $\tau_{\partial T'/\partial z} \sim 10-15$ min is given in Appendix A by Eq. (A11). It is shorter than $\tau_{T'}$ because the lapse rate fluctuations have more energy at the higher frequencies than T'. Specifically, the lapse rate spectrum is proportional to

ω^{-p+1}, where $p \sim 2$, which is shallower than the T' spectrum. which is proportional to ω^{-p}.

4. EXAMPLES

The conventional and interleaved data processing methods are illustrated using Na Doppler lidar measurements from McMurdo (77.84°S, 166.67°E), Antarctica. The data were acquired on 26–27 and 29–30 May 2020 by the University of Colorado lidar group. The instrument employs the well-known three-frequency technique to infer simultaneously the temperature (T), vertical wind (w), and atomic Na density ([Na]) in the MLT region [2]. The laser transmitter runs at a 50 Hz pulse repetition rate with a single pulse energy of ~10−20 mJ. The telescope aperture area is 0.5 m^2, so the power-aperture product (PA) is between 0.25 and 0.5 Wm^2, depending on the laser power. This is comparable to the PA of many Na lidars currently in operation around the globe. The quality of the data used here, to demonstrate the new processing technique, is similar to that being acquired by other lidar groups, who should be able to obtain comparable results. However, as it will be shown, the technique is effective even under low SNR conditions and can also be applied to situations where the signal levels are quite low.

Raw photon counts were acquired at McMurdo with resolutions of $\delta t = 4.5$ s and $\delta z = 24$ m. The data acquisition system saves the raw data in a convenient format designed for ease of data retrieval, which has also proven useful for interleaving the photon count samples during post-processing. Each 4.5 s set of three-frequency photon count data (1.5 s per frequency) is stored in pairs. The first and second sets from each pair are then integrated separately to yield two interleaved, statistically independent, times series of three-frequency photon counts with the desired resolution ($\Delta t = 2.5$ min and $\Delta z = 0.96$ km, in this case). By storing the high-resolution raw data in pairs, we ensure that the two lower-resolution, interleaved times series are offset by exactly 4.5 s regardless of the data gaps. Two statistically independent sets of temperature times series are then derived from the interleaved photon counts, $T_{\text{odd}}(z, t)$ and $T_{\text{even}}(z, t)$, along with corresponding sets of w and [Na] measurements. Next, two sets of temperature perturbations, T'_{odd} and T'_{even}, are derived by subtracting the individual temporal means for each of the observation periods. The sample covariances of the temperature and lapse rate fluctuations are then computed using these two interleaved datasets according to methods described in Sections 2 and 3. These calculations provide estimates of the wave-driven fluctuation variances [$\text{Var}(T')$ and $\text{Var}(\partial T'/\partial z)$]. The total variances of the measured temperature [$\text{Var}(T') + \text{Var}(\Delta T)$] and lapse rate [$\text{Var}(\partial T'/\partial z) + \text{Var}(\partial \Delta T/\partial z)$] fluctuations are calculated using the conventional processing method. First, the photon count pairs are added to form a single set of three-frequency counts (to increase the signal levels by two). These datasets are then summed to achieve the proper resolution and

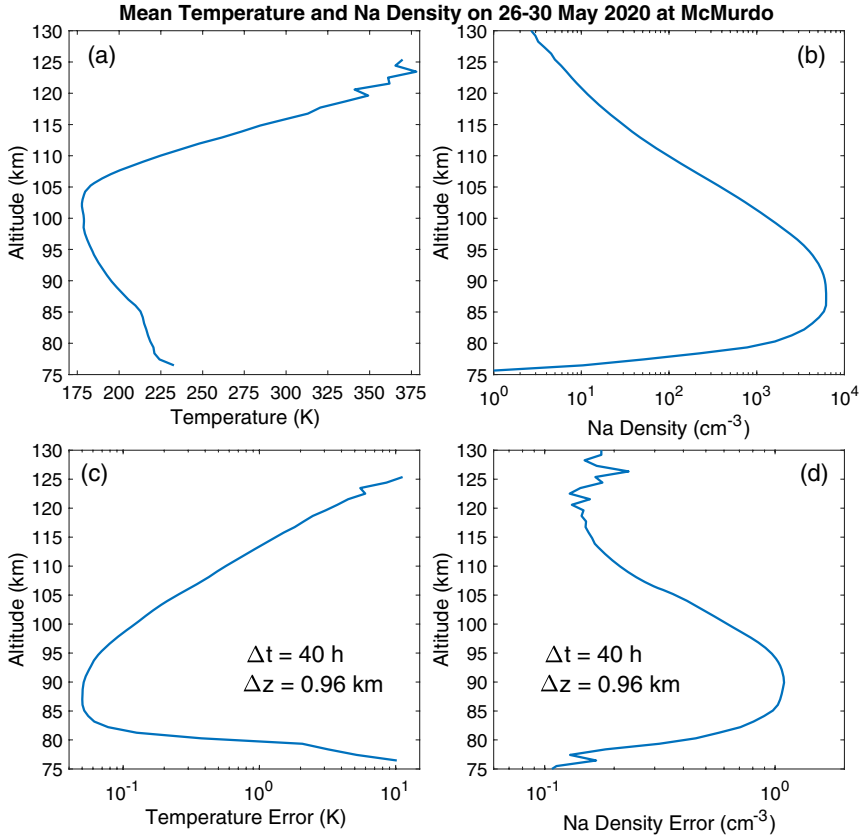

Fig. 1. (a) Mean temperature profile averaged over 26–27 and 29–30 May 2020 datasets taken at McMurdo, Antarctica with the University of Colorado Na Doppler lidar. (b) Mean Na density profile observed over the same period. (c) Measurement uncertainty (precision) of the mean temperature profile. (d) Measurement uncertainty (precision) of the mean Na density. Note that temperatures and Na densities were first retrieved at resolutions of 30 min and 0.96 km and then averaged over the 40 h of observations.

used to derive estimates of T, w, and [Na] and their fluctuations. Finally, the sample variances of the measured temperature and lapse rate fluctuations are computed using the conventional method, as described in Sections 2 and 3.

The two lidar datasets, acquired on 26–27 and 29–30 May 2020, include, respectively, ~22 and 18 h of continuous measurements. Both observation periods exceed the inertial period (~12.24 h) at McMurdo. The mean temperature and Na density profiles, averaged over the full 40 h of observations at a vertical resolution of 0.96 km, are shown in Figs. 1(a) and 1(b), and the associated measurement uncertainties (precision) are shown in Figs. 1(c) and 1(d). The mesopause temperature is ~178 K at 102.5 km. Temperatures below the mesopause gradually increase to ~220 K at 79 km, while temperatures above the mesopause quickly increase to ~368 K at 125 km in the thermosphere. The precision of the mean temperature profile varies from less than 0.1 K near the Na layer peak to 8–10 K at 77 and 125 km. The Na densities range from ~6200 cm^{-3} at 88 km to ~2–3 cm^{-3} at 76 and 130 km. The precision of the mean Na profile varies from 1 cm^{-3} to less than 0.2 cm^{-3}, but the relative errors increase substantially near the layer top edge.

In Fig. 2(a), the wave-driven temperature variance (blue curve), calculated using the interleaved data method at $\Delta t = 2.5$ min and $\Delta z = 0.96$ km, is compared with the total temperature variance (red curve), calculated using the conventional method. The 2.5 min temporal resolution was chosen to achieve sufficient SNR while also enabling the detection of the shortest period gravity waves ($\tau_B \sim 5$ min). The difference between these two curves is the noise-induced temperature variance (yellow curve). Between 90 and 95 km, this photon noise bias for the conventional processing method is small, ~4 K^2. But, near the bottom and top edges of the Na layer, the noise bias increases considerably. In fact, it exceeds the wave-driven variance (blue curve) below 79 km and above 112 km. At 115 km, the noise bias (~2100 K^2) is triple the wave-driven temperature variance (~690 K^2). It is obvious from Fig. 2(a) that the total variance, calculated using the conventional method, is dominated by the photon noise bias at the extremes of the Na layer where the SNR is very low. The precision (rms error) of the high-resolution ($\Delta t = 2.5$ min, $\Delta z = 1$ km) temperature measurements is plotted on a linear scale in Fig. 3(a). This curve is simply the square root of the photon noise variance plotted in Fig. 2(a). The rms error is about ±2 K near the Na layer peak between 90 and 95 km, which is typical for well-designed Na lidars operating at night with clear skies. However, at the highest altitudes on the topside of the Na layer near 115 km, where the mean Na density is only about 25 cm^{-3}, the temperature error increases to about ±46 K. In comparison, the rms value of the wave-driven temperature fluctuations is much smaller ($\sqrt{690 K^2} = 26.3$ K). In spite of this large measurement error, it is still possible to derive scientifically useful estimates of the wave-driven temperature variance by using the interleaved data processing technique. As shown by Eq. (12), the calculated temperature covariance is smaller than the variance by the term $\frac{\delta t^2}{2}\mathrm{Var}[\frac{\partial T'(z)}{\partial t}]$, which can be calculated from the interleaved temperature data. For the McMurdo data, $\mathrm{Var}(\partial T'/\partial t)$ averages ~4300 K^2 h^{-2} between 80 and 100 km. The covariance correction term for $\delta t = 4.5$ s averages to ~0.003 K^2, which is considerably smaller than the estimated values of $\mathrm{Var}(T')$ over this altitude range and can be neglected.

The uncertainties of the variance estimates, computed using both the interleaved and conventional methods, are plotted in Fig. 2(b) [see Eqs. (15) and (17)]. The related parameters and equations are listed in Table 1. The uncertainties of both estimates are near 32% over the majority of the height range, but they increase by a few percent at the top and bottom edges of the Na layer where the photon noise makes small, but noticeable, contributions to the measurement uncertainties. Nevertheless, the accuracy of the estimated wave-driven temperature variance, calculated using the interleaved data method, has been improved significantly, because the bias induced by photon noise has been eliminated. The uncertainties for both methods are largely determined by statistical noise, the length of the

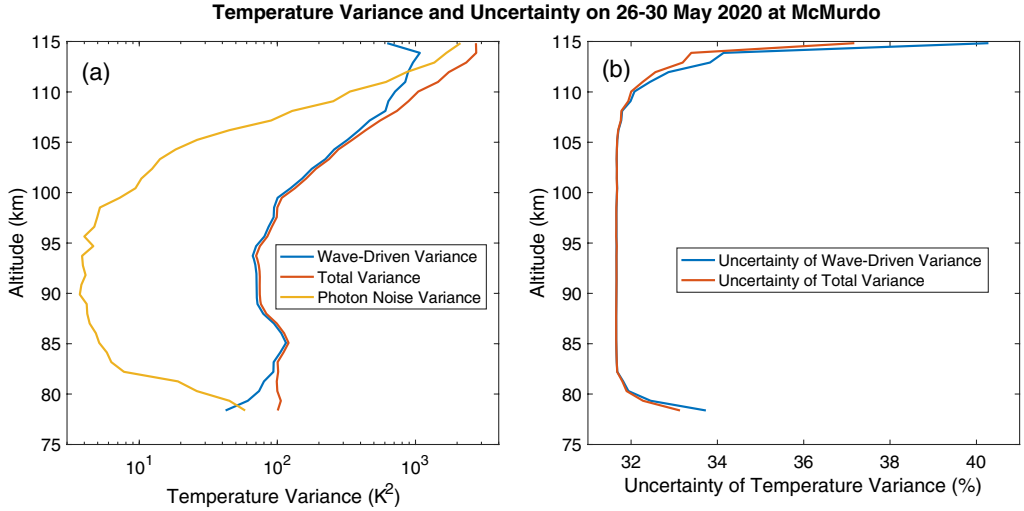

Fig. 2. (a) Wave-driven and total temperature variances computed from the interleaved and conventional methods, respectively, along with the difference between them (photon noise variance) using the two datasets on 26–27 and 29–30 May 2020 over McMurdo, Antarctica. (b) Uncertainties (precision) of wave-driven and total temperature variances in percentage. Note that temperatures were retrieved at resolutions of 2.5 min and 0.96 km, and the T' variances derived using the two methods were averaged over the 40 h of observations.

Table 1. RMS Uncertainties of the Estimated T' and $\partial T'/\partial z$ Variances for McMurdo Observations[a]

Wave-Driven Temperature Variance: $\tau_{T'} \approx 2$h, $\Delta t = 2.5$ min, $\tau_{obs} \approx 40$ h
Interleaved Method: $\Delta\rho_{T'+\Delta T}(\delta t) \simeq \sqrt{\frac{1}{10}\mathrm{Var}^2(T') + \frac{1}{960}[2\mathrm{Var}(T')\mathrm{Var}(\Delta T) + \mathrm{Var}^2(\Delta T)]}$
Conventional Method: $\Delta s^2_{T'+\Delta T} \simeq \sqrt{\frac{1}{10}\mathrm{Var}^2(T') + \frac{1}{480}[2\mathrm{Var}(T')\mathrm{ar}(\Delta T) + \mathrm{Var}^2(\Delta T)]}$
Wave-Driven Lapse Rate Variance: $\tau_{\partial T'/\partial z} \approx 15$ min, Δt = 2.5 min, $\tau_{obs} \approx 40$ h
Interleaved Method: $\Delta\rho_{\partial(T'+\Delta T)/\partial z}(\delta t) \simeq \sqrt{\frac{1}{80}\mathrm{Var}^2(\partial T'/\partial z) + \frac{1}{960}[2\mathrm{Var}(\partial T'/\partial z)\mathrm{Var}(\partial\Delta T/\partial z) + \mathrm{Var}^2(\partial\Delta T/\partial z)]}$
Conventional Method: $\Delta s^2_{\partial(T'+\Delta T)/\partial z} \simeq \sqrt{\frac{1}{80}\mathrm{Var}^2(\partial T'/\partial z) + \frac{1}{480}[2\mathrm{Var}(\partial T'/\partial z)\mathrm{Var}(\partial\Delta T/\partial z) + \mathrm{Var}^2(\partial\Delta T/\partial z)]}$

[a]Var(T') is obtained from the interleaved method, while Var(ΔT) for conventional processing is obtained from the difference between the total and wave-driven variances. Var(ΔT) for interleaved processing is two times the value for conventional processing. Same for the lapse rate.

total observation period, and the temperature correlation time, which is about 2 h for the McMurdo dataset. If the variances were recomputed using 160 h of lidar observations instead of 40 h, the uncertainties would be reduced by a factor of 2 to about 16–18%.

Similar profiles for the lapse rate fluctuations are plotted in Figs. 4(a), 4(b), and 3(b), and the related parameters and equations for calculating the uncertainties of the estimated variances are listed in Table 1. The major differences are that the photon noise bias is larger, relative to the lapse rate variance, and the uncertainty of the lapse rate variance, computed using the interleaved method, is smaller, except at the edges of the Na layer where the SNR is low. Between 85 and 95 km, the uncertainty of the lapse rate variance is about 11%. It increases substantially above 105 km rising to about 25% at 110 km. In fact, at 110 km the photon noise variance (∼750 K^2/km^2) is about three times the true lapse rate variance (∼225 K^2/km^2). These differences between the temperature and lapse rate variances arise primarily because the lapse rate correlation time in the MLT at McMurdo is ∼15 min, eight times smaller than the temperature correlation time (∼2 h). Consequently, statistical noise makes a much smaller contribution to the lapse rate uncertainty than it does for the temperature variance uncertainty. In regions of low SNR, the uncertainty of the estimated lapse rate variance is dominated by photon noise, which is ±27.4 K/km at 110 km, as shown in Fig. 3(b), compared to ∼15 K/km for the rms value of the wave-driven lapse rate fluctuations. Nevertheless, by employing the interleaved data processing method, it is possible to derive scientifically useful estimates of the lapse rate variance well into the thermosphere to at least 110 km where the photon noise is significantly larger, as this example demonstrates.

In summary, the interleaved data processing method provides two statistically independent measurements of the temperature perturbations and lapse rate fluctuations, thereby eliminating the photon-noise-induced biases in the computation of the temperature and lapse rate variances. The measurement accuracies are significantly improved by eliminating the biases, allowing the measurements to be extended into the lower thermosphere where the lidar SNR is low. The uncertainties (precision) associated with the variance measurements can be improved by increasing the total observation time to provide more independent samples of the variances or by averaging in altitude. Increasing the observation time by four times decreases the uncertainty by two times. However, there is a limit to the

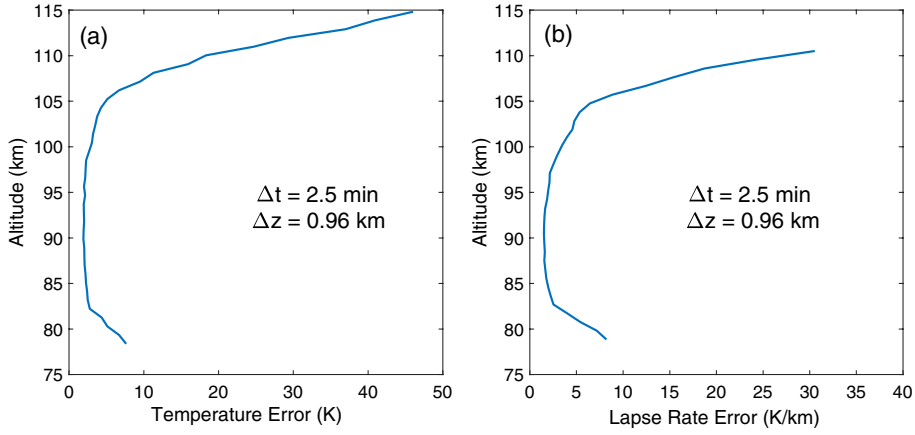

Fig. 3. (a) RMS temperature error at resolutions of 2.5 min and 0.96 km during 26–27 and 29–30 May 2020 lidar observations at McMurdo. (b) RMS lapse rate error at the same resolutions. These errors are calculated by taking the square root of the photon noise variances in Figs. 2(a) and 4(a).

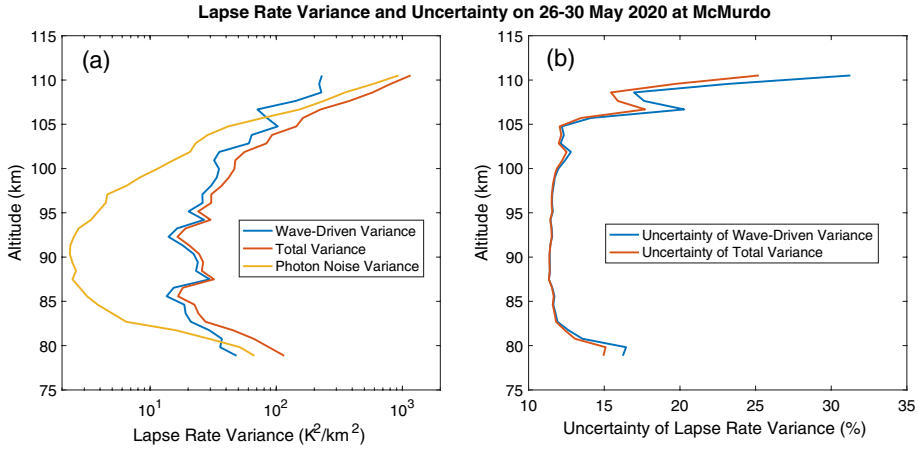

Fig. 4. (a) Wave-driven and total lapse rate variances computed from the interleaved and conventional methods, respectively, along with the difference between them (photon noise variance) using the two datasets on 26–27 and 29–30 May 2020 over McMurdo, Antarctica. (b) Uncertainties (precision) of wave-driven and total lapse rate variances in percentage. Note that temperatures were retrieved at resolutions of 2.5 min and 0.96 km, and the lapse rate variances derived were averaged over the 40 h of observations.

amount of averaging since we are interested in observing the geophysical variations of these parameters both in altitude and time. The McMurdo data shown here have a temporal resolution of 2.5 min, and the observational window lengths exceed the inertial period. As a result, the temperature and lapse rate variances include contributions from the entire gravity wave spectrum that ranges from the inertial frequency (f) up to the buoyancy frequency (N). Furthermore, because the dataset was acquired with a fundamental resolution of 4.5 s, the variance corrections associated with the interleaved data processing method are negligible [$\sim$0.003%Var(T') and $\sim$0.045%Var($\partial T'/\partial z$)]. In contrast, if we had employed the 2.5 min resolution data, multiplied each T' sample by the next sample and then computed the average to estimate $\mathrm{Cov}_{T'}(\Delta t)$, instead of estimating $\mathrm{Cov}_{T'}(\delta t)$, the covariance corrections would be $(\Delta t/\delta t)^2 = (150s/4.5s)^2 = 1111$ times larger, and, at least for the lapse rate, the correction would be substantial ($\sim$50%). This simple example demonstrates the significant advantage of acquiring data at a high fundamental resolution and then interleaving the samples to compute the second-order statistics at lower resolutions that are consistent with the processes being observed.

5. PRACTICAL CONSIDERATIONS AND OTHER APPLICATIONS

When making observations of wave effects in the atmosphere, the vertical resolution, altitude range of observations, temporal resolution, and length of the observation period should be chosen to enable the most energetic waves to be observed. The smallest-scale waves have vertical wavelengths of $\sim$1 km with periods comparable to the buoyancy period $\tau_B \sim 5$ min (smaller scales are turbulence). The most energetic large-scale waves have vertical wavelengths comparable to λ_z^*, which is 15–20 km at mesopause heights, and periods comparable to the inertial period $\tau_i = 12\,\mathrm{h}/\sin\theta$, where θ is the latitude, thus, to capture the most important waves in the spectrum requires $h \sim 10-20$ km, $\tau_{\mathrm{obs}} \sim 6-12$ h, $\Delta z \sim 0.25-1$ km, and $\Delta t \sim 2.5-10$ min. Furthermore, when estimating the T' and $\partial T'/\partial z$ variances, we prefer the correction terms to be small so that uncertainties in these corrections make only small contributions to the overall uncertainties in the measured parameters. Thus, the correction terms should be derived from the dataset with δt chosen to ensure that they are small compared to the measured parameters. For example, if $\tau_{\mathrm{obs}} = 8$ h, $\Delta t = 2.5$ min $\approx \tau_B/2$, and $\delta t = 30$ s, the correction term for the T' variance is $\sim$0.2%Var(T') [see Eqs. (13) and (A15)] and for the $\partial T'/\partial z$ variance is $\sim$2%Var($\partial T'/\partial z$) [see Eqs. (24) and (A10)], which meet the criteria. These correction terms are even smaller if $\Delta t > \tau_B/2$.

In practice, observations of the second-order statistics of the wave-driven fluctuations of temperature and winds are conducted to characterize the impact of wave sources in the lower atmosphere and to assess the seasonal variations of wave activity and its influence on the wind, thermal, and constituent structure throughout the atmosphere. Because the temperature and lapse rate correlation times are relatively long ($\sim$1–2 h and 10–15 min, respectively), the uncertainties in the estimated variances of the temperature and lapse rate fluctuations, that are derived from a single observation period are large. Measurements from multiple observation periods (which are statistically independent) are averaged to reduce the uncertainties and thereby obtain scientifically useful estimates of these parameters. Decades of measurements made throughout the atmosphere have shown that the ω and m spectra of the temperature fluctuations are approximately proportional to ω^{-2} and m^{-3} [20,25–27], in which case the rms uncertainties in the measured T' and $\partial T'/\partial z$ variances are [see Eqs. (15), (A5), and (A11)]

$$\Delta\mathrm{Var}\left(T'\right) \simeq \sqrt{\frac{2\pi\,\mathrm{Var}^2\left(T'\right)}{3n\tau_{\mathrm{obs}}\omega_{\mathrm{min}}} + \frac{\Delta t}{n\tau_{\mathrm{obs}}}\left[2\mathrm{Var}\left(T'\right)\mathrm{Var}\left(\Delta T\right) + \mathrm{Var}^2\left(\Delta T\right)\right]}, \tag{25}$$

and

$$\Delta\mathrm{Var}\left(\frac{\partial T'}{\partial z}\right) \simeq \sqrt{\frac{2\pi\,\mathrm{Var}^2\,(\partial T'/\partial z)}{n\tau_{\mathrm{obs}}\omega_{\min}\ln^2\,(\omega_{\max}/\omega_{\min})} + \frac{\Delta t}{n\tau_{\mathrm{obs}}}\left[2\mathrm{Var}\left(\frac{\partial T'}{\partial z}\right)\mathrm{Var}\left(\frac{\partial \Delta T}{\partial z}\right) + \mathrm{Var}^2\left(\frac{\partial \Delta T}{\partial z}\right)\right]}, \tag{26}$$

where n is the number of observation periods employed, τ_{obs} is their average duration, and $n\tau_{\mathrm{obs}}$ is the total observation period. For $n = 12$, $\tau_{\mathrm{obs}} = 8$ h, and $n\tau_{\mathrm{obs}} = 96$ h, $\Delta\mathrm{Var}(T') \approx 20\%\mathrm{Var}(T')$ and $\Delta\mathrm{Var}(\partial T'/\partial z) \approx 7\%\mathrm{Var}(\partial T'/\partial z)$, which should be adequate to characterize the monthly variations in wave activity.

The vertical fluxes of heat, horizontal momentum, and constituents are also important second-order statistics that can be biased by photon noise if the measurement errors are correlated. For example, the vertical heat flux is defined as the covariance between the vertical wind and temperature fluctuations. The photon noise bias $\langle \Delta w \Delta T \rangle$ can be eliminated from the measured heat flux by deriving w from one set of measured photon counts and T from another statistically independent set. This can be accomplished by using the same technique of interleaving the measured photon counts that we applied to temperature measurements. The heat flux is then estimated as follows:

$$\langle w' T' \rangle \simeq \frac{1}{2}\left(\overline{w'_{\mathrm{odd}} T'_{\mathrm{even}}} + \overline{w'_{\mathrm{even}} T'_{\mathrm{odd}}}\right), \tag{27}$$

where the overbar denotes the sample average. Similar approaches can be adapted for estimating the momentum and constituent fluxes that also eliminate the photon noise biases.

When conventional spectral analysis techniques are used to estimate the power spectra of the temperature, wind, and constituent fluctuations, the white photon noise establishes a noise floor equal to $2\Delta t\mathrm{Var}(\Delta T)$ for the temporal frequency spectrum and $2\Delta z\mathrm{Var}(\Delta T)$ for the vertical wavenumber spectrum. These noise floors can be eliminated by computing the spectra, again using two statistically independent datasets, such as T'_{odd} and T'_{even}. The approach involves deriving the spectra from the discrete Fourier transforms (DFT). For example, the estimated temporal frequency spectrum is

$$F_{T'}(\omega) \simeq \left|\overline{\mathrm{DFT}_\omega\left(T'_{\mathrm{odd}}\right)\mathrm{DFT}^*_\omega\left(T'_{\mathrm{even}}\right)}\right|. \tag{28}$$

Of course, it is important to apply the best practices for computing power spectra, which might include optimizing the data window function, zero padding the data series, and implementing pre-whitening/post-coloring processing.

Although we focused on temporally interleaving the raw photon counts to construct two statistically independent datasets, it is obvious that independent datasets can also be constructed by interleaving range bins. For some instruments and/or for some observations, it may be advantageous to interleave the range bins or employ a combination of interleaved range and time bins. It is also possible to weight the interleaved samples to effectively eliminate the offsets in time and altitude between the two datasets. This approach is useful when the fundamental resolutions of the acquired data (δt and δz) are rather large. The key is to construct the two datasets so that the computed variances, fluxes, and spectra closely approximate the underlying statistics of the observed variable and do not require significant correction. As we have shown here, temporal interleaving with δt on the order of a few tens of seconds is certainly adequate for estimating the variances of the temperature and lapse rate fluctuations induced by gravity waves in the upper atmosphere, even in regions where the lidar SNR is so low that conventional processing techniques cannot be employed.

6. CONCLUSIONS

We have shown how lidar observations of temperature can be processed in a straightforward manner to completely eliminate the photon noise biases and noise floors that normally occur in the calculated second-order statistics of the measured temperature fluctuations. To be most effective, the raw photon count data should be acquired with temporal and vertical resolutions (δt and δz) that are smaller than the required resolutions of the derived temperature data (Δt and Δz). This allows statistically independent measurements of the temperature time series and profiles to be derived by interleaving the odd and even numbered time or range bin photon count data. In such cases, it is possible to measure the fluctuation variances and spectra, even in regions where the SNR is quite low, as was confirmed by the Na Doppler lidar examples shown in Section 4.

Although the results were demonstrated specifically, by analyzing wave-driven temperature fluctuations, the results are directly applicable to wave-driven horizontal wind fluctuations, because the horizontal wind spectra are approximately proportional to the temperature spectra [20]. The results can also be applied to vertical winds, but, in this case, it would be necessary to rederive the statistical formulas in Appendix A by using the appropriate spectral models for the vertical wind fluctuations. Similarly, the results can be adapted to turbulence measurements. Because turbulence occurs at much smaller scales than waves, the raw photon count data must be acquired at a temporal resolution of a few seconds and a vertical resolution of a few tens of meters. Furthermore, the derivation of the temperature and wind fluctuations from the high-resolution, low SNR photon count data is a more involved process [13]. Of course, the key statistical formulas in Appendix A would also need to be reevaluated by employing the spectral models for the turbulence fluctuations.

Because the photon noise biases and noise floors are eliminated from the variance, flux, and spectrum estimates, these second-order statistics can even be reliably estimated from noisy high-resolution data to provide better characterizations of the smaller-scale waves. The SNR of the raw data need only be large enough to permit the derivation of the temperature and wind fluctuations, which are typically nonlinear functions of the measured photon counts. However, the system equations can be readily linearized to avoid the data processing problems associated with nonlinear inversions in low SNR environments [13].

APPENDIX A

To evaluate the correction factors, we employ the single-sided two-dimensional (2D) gravity wave temperature fluctuation spectrum model that was derived from diffusive filtering theory, which is not separable in m and ω [28]:

$$F_{T'}(m,\omega)=\frac{(2\pi)^2\overline{(T')^2}}{\left[1-(f/N)^{p-1}\right]}\frac{(s+1)}{m_*}\left(\frac{m}{m_*}\right)^s \times\frac{(p-1)}{f}\left(\frac{f}{\omega}\right)^{p+\frac{(s+1)}{2}}, \tag{A1}$$

where $0\le m\le m_*\sqrt{\frac{\omega}{f}}$ and $f\le\omega\le N$. $f=2\pi/\tau_i$ is the inertial frequency, $\tau_i=12\,\mathrm{h}/\sin\theta$ is the inertial period, θ is the latitude, $\mathrm{N}=2\pi/\tau_B$ is the buoyancy frequency, $\tau_B\approx 5$ min is the buoyancy period, and $\overline{(T')^2}$ represents the temperature fluctuation variances associated with the full spectrum of waves. The parameter s is related to the gravity wave source characteristics. It is difficult to measure, but is not critical, and so most researchers assume $s=1$. The parameter p is also related to the source charateristics. For this model, the one-dimensional (1D) frequency spectrum is proportional to ω^{-p}, and the 1D vertical wavenumber spectrum is proportional to $m^{-(2p-1)}$ for $m_*<m$. A wide variety of observational studies have shown that 1D spectra are approximately proportional to ω^{-2} and m^{-3} ([20] and references therein), so $p=2$ is a reasonable choice. For $s=1$ and $p=2$, the characteristic vertical wavenumber is given by

$$m_*^2=\left(\frac{2\pi}{\lambda_z^*}\right)^2\simeq\frac{2}{\ln(N/f)}\frac{\overline{(\partial T'/\partial z)^2}}{\overline{(T')^2}}. \tag{A2}$$

λ_z^* is the characteristic vertical wavelength. For the diffusive filtering model, none of the waves, for which $\lambda_z^*\le\lambda_z$, are damped by diffusion. λ_z^* is 15–20 km at mesopause altitudes, which is comparable to h, the typical height range of observations.

By taking into account the region of the gravity wave temperature spectrum observed by the lidar, the 1D ω spectrum is

$$F_{T'}(\omega)\simeq\frac{2\pi\,\mathrm{Var}\left(T'\right)}{\left[1-(\omega_{\min}/\omega_{\max})^{p-1}\right]}\frac{(p-1)}{\omega_{\min}}\left(\frac{\omega_{\min}}{\omega}\right)^p, \quad f\le\omega_{\min}\le\omega\le\omega_{\max}\le N, \tag{A3}$$

where $\omega_{\max}=\min(N,\pi/\Delta t)$ is the highest observed gravity wave frequency, and $\omega_{\min}=\max(f,2\pi/\tau_{\mathrm{obs}})$ is the lowest. $\mathrm{Var}(T')\le\overline{(T')^2}$ is the temperature fluctuation variance corresponding to the region of the gravity wave spectrum observed by the lidar. The 1D m spectrum is

$$F_{T'}(m)\simeq\frac{2\pi\,\mathrm{Var}\left(T'\right)}{\left[1-(m_{\min}/m_{\max})^{2(p-1)}\right]}\frac{2(p-1)}{m_{\min}}\left(\frac{m_{\min}}{m}\right)^{2p-1}, \times m_*\le m_{\min}\le m\le m_{\max}, \tag{A4}$$

where $m_{\max}\approx\pi/\Delta z$ is the highest observed wavenumber, and $m_{\min}\approx 2\pi/h$ is the lowest. The temporal correlation time of the temperature fluctuations is given by [17]

$$\tau_{T'}=\frac{\frac{1}{4\pi}\int_{\omega_{\min}}^{\omega_{\max}}F_{T'}^2(\omega)\mathrm{d}\omega}{\mathrm{Var}^2\left(T'\right)}=\frac{\left[1-(\omega_{\min}/\omega_{\max})^{2p-1}\right]}{\left[1-(\omega_{\min}/\omega_{\max})^{p-1}\right]} \times\frac{\pi(p-1)^2}{(2p-1)\,\omega_{\min}}\simeq\frac{(p-1)^2}{(2p-1)}\frac{\pi}{\omega_{\min}}. \tag{A5}$$

Because the 2D lapse rate ($\partial T'/\partial z$) fluctuation spectrum is

$$F_{\partial T'/\partial z}(m,\omega)=m^2F_{T'}(m,\omega), \tag{A6}$$

it is easy to show that the 1D ω spectrum is proportional to ω^{-p+1}, and the 1D m spectrum is proportional to m^{-2p+3} for $m_*\le m$. For $p=2$, the 1D spectra are

$$F_{\partial T'/\partial z}(\omega)\simeq\frac{2\pi\,\mathrm{Var}\left(\partial T'/\partial z\right)}{\ln(\omega_{\max}/\omega_{\min})\,\omega}, \quad f\le\omega_{\min}\le\omega\le\omega_{\max}\le N, \tag{A7}$$

and

$$F_{\partial T'/\partial z}(m)\simeq\frac{2\pi\,\mathrm{Var}\left(\partial T'/\partial z\right)}{\ln(m_{\max}/m_{\min})\,m}, \quad m_*\le m_{\min}\le m\le m_{\max}, \tag{A8}$$

where

$$\mathrm{Var}\left(\frac{\partial T'}{\partial z}\right)=\frac{2m_{\min}^2\ln(m_{\max}/m_{\min})}{\left[1-(m_{\min}/m_{\max})^2\right]}\mathrm{Var}\left(T'\right) \simeq 2m_{\min}^2\ln(m_{\max}/m_{\min})\,\mathrm{Var}\left(T'\right), \tag{A9}$$

and

$$\mathrm{Var}\left(\frac{\partial^2T'}{\partial t\partial z}\right)=\frac{\left(\omega_{\max}^2-\omega_{\min}^2\right)}{2\ln(\omega_{\max}/\omega_{\min})}\mathrm{Var}\left(\frac{\partial T'}{\partial z}\right) \simeq\frac{\omega_{\max}^2}{2\ln(\omega_{\max}/\omega_{\min})}\mathrm{Var}\left(\frac{\partial T'}{\partial z}\right). \tag{A10}$$

The temporal correlation time of $\partial T'/\partial z$ is

$$\tau_{\partial T'/\partial z}=\frac{\frac{1}{4\pi}\int_{\omega_{\min}}^{\omega_{\max}}F_{\partial T'/\partial z}^2(\omega)\mathrm{d}\omega}{\mathrm{Var}^2\left(\partial T'/\partial z\right)}=\frac{\pi}{\ln^2(\omega_{\max}/\omega_{\min})} \times\left(\frac{1}{\omega_{\min}}-\frac{1}{\omega_{\max}}\right)\simeq\frac{\pi}{\omega_{\min}\ln^2(\omega_{\max}/\omega_{\min})}. \tag{A11}$$

Similarly, because the 2D spectrum of $\partial T'/\partial t$ is

$$F_{\partial T'/\partial t}(m,\omega)=\omega^2F_{T'}(m,\omega), \tag{A12}$$

the 1D ω spectrum is proportional to ω^{-p+2}, and the 1D m spectrum is proportional to m^{5-2p} for $m_*\le m$. For $p=2$, the 1D spectra are

$$F_{\partial T'/\partial t}(\omega)\simeq\frac{2\pi\,\mathrm{Var}\left(\partial T'/\partial t\right)}{(\omega_{\max}-\omega_{\min})}, \quad f\le\omega_{\mathrm{in}}\le\omega\le\omega_{\max}\le N, \tag{A13}$$

and

$$F_{\partial T'/\partial t}(m) \simeq \frac{4\pi \mathrm{Var}\left(\partial T'/\partial t\right) m}{\left(m_{\max}^2 - m_{\min}^2\right)},\ m_* \le m_{\min} \le m \le m_{\max}, \tag{A14}$$

where

$$\mathrm{Var}\left(\frac{\partial T'}{\partial t}\right) = \omega_{\max}\omega_{\min}\mathrm{Var}\left(T'\right). \tag{A15}$$

APPENDIX B

The derivation of the uncertainties of the estimated sample variance and convariance follows the approach described in the appendix of Ref. [17]. In our case, where we are only averaging over time, we rewrite Eq. (A4) in [17] as

$$\begin{aligned}\mathrm{Var}\left[\overline{(x+\Delta x)(y+\Delta y)}\right] &= \frac{1}{\tau_{\mathrm{obs}}}\int_{-\tau_{\mathrm{obs}}}^{\tau_{\mathrm{obs}}} \mathrm{d}t \left(1-\frac{|t|}{\tau_{\mathrm{obs}}}\right) \\ &\times \left[C^2_{(x+\Delta x)(y+\Delta y)}(t) + C_{x+\Delta x}(t)\, C_{y+\Delta y}(t)\right] \\ &= \frac{1}{\tau_{\mathrm{obs}}}\int_{-\tau_{\mathrm{obs}}}^{\tau_{\mathrm{obs}}} \mathrm{d}t \left(1-\frac{|t|}{\tau_{\mathrm{obs}}}\right)\left\{\left[C_{xy}(t) + C_{\Delta x\Delta y}(t)\right]^2\right. \\ &+ \left.\left[C_x(t) + C_{\Delta x}(t)\right]\left[C_y(t) + C_{\Delta y}\right]\right\}, \end{aligned} \tag{B1}$$

where the overbar denotes time averaging, $C_x(t)$ is the auto-covariance of x, and $C_{xy}(t)$ is the cross covariance between x and y. Recall that x and y are statistically independent of the measurement errors Δx and Δy. Notice that for convenience we have expressed the discrete time averaging by the corresponding continuous time average. For conventional processing, $x(t) = y(t) = T'(t)$ and $\Delta x(t) = \Delta y(t) = \Delta T(t)$ so that $C_{xy} = C_{T'}$ and $C_{\Delta x\Delta y} = C_{\Delta T}$, and Eq. (B1) reduces to

$$\begin{aligned}\mathrm{Var}\left(s^2_{T'+\Delta T}\right) &= \frac{1}{\tau_{\mathrm{obs}}}\int_{-\tau_{\mathrm{obs}}}^{\tau_{\mathrm{obs}}} \mathrm{d}t \left(1-\frac{|t|}{\tau_{\mathrm{obs}}}\right) \\ &\times \left[2C^2_{T'}(t) + 4C_{T'}(t)\, C_{\Delta T}(t) + 2C^2_{\Delta T}(t)\right]. \end{aligned} \tag{B2}$$

For interleaved processing, $x(t) = T'(t)$ and $y(t) = T'(t+\delta t) \approx T'(t)$, while $\Delta x(t) = \Delta T(t)$ and $\Delta y(t) = \Delta T(t)$ are the temperature errors for the respective interleaved temperature estimates. But, Δx and Δy are statistically independent so that $C_{\Delta x\Delta y}(t) = 0$. In this case, Eq. (B1) reduces to

$$\begin{aligned}\mathrm{Var}\left(\rho_{T'+\Delta T}\right) &= \frac{1}{\tau_{\mathrm{obs}}}\int_{-\tau_{\mathrm{obs}}}^{\tau_{\mathrm{obs}}} \mathrm{d}t \left(1-\frac{|t|}{\tau_{\mathrm{obs}}}\right) \\ &\times \left[2C^2_{T'}(t) + 2C_{T'}(t)\, C_{\Delta T}(t) + C^2_{\Delta T}(t)\right]. \end{aligned} \tag{B3}$$

Equations (B2) and (B3) may be simplified by expressing the auto-covariance functions in terms of the corresponding power spectra, as was done in the appendix of [17]. This approach permits an accurate calculation of the correlation times for T' (viz. $\tau_{T'}$) and ΔT (viz. Δt). Alternatively, we may approximate the integrals in Eqs. (B2) and (B3) by first assuming that τ_{obs} is larger than the correlation times of both T' and ΔT. Since the auto-covariance functions are approximately equal to the variances for small $|t|$ less than the respective correlation times and are approximately zero for large $|t|$ greater than the respective correlation times, Eqs (B2) and (B3) can be evaluated approximately as

$$\begin{aligned}\mathrm{Var}\left(s^2_{T'+\Delta T}\right) &\simeq \frac{2\tau_{T'}}{\tau_{\mathrm{obs}}}\mathrm{Var}^2\left(T'\right) + \frac{4\Delta t}{\tau_{\mathrm{obs}}}\mathrm{Vat}\left(T'\right)\mathrm{Var}\left(\Delta T\right) \\ &+ \frac{2\Delta t}{\tau_{\mathrm{obs}}}\mathrm{Var}^2\left(\Delta T\right), \end{aligned} \tag{B4}$$

and

$$\begin{aligned}\mathrm{Var}\left(\rho_{T'+\Delta T}\right) &\simeq \frac{2\tau_{T'}}{\tau_{\mathrm{obs}}}\mathrm{Var}^2\left(T'\right) + \frac{2\Delta t}{\tau_{\mathrm{obs}}}\mathrm{Vat}\left(T'\right)\mathrm{Var}\left(\Delta T\right) \\ &+ \frac{\Delta t}{\tau_{\mathrm{obs}}}\mathrm{Var}^2\left(\Delta T\right). \end{aligned} \tag{B5}$$

Funding. National Science Foundation (AGS-1452351, AGS-2029162, OPP-1246405, OPP-1246431, OPP-1443726).

Acknowledgment. We gratefully acknowledge winter-over scientists Xianxin Li and Ying-Tsen (Cissi) Lin for collecting the lidar data in May 2020 at McMurdo Station in Antarctica. We thank the staff of the United States Antarctic Program, Antarctica New Zealand, McMurdo Station, and Scott Base for their superb support.

Disclosures. The authors declare no conflicts of interest.

REFERENCES

1. T. Fujii and T. Fukuchi, eds., *Laser Remote Sensing* (CRC Press, Taylor & Francis, 2015).
2. X. Chu and G. C. Papen, "Resonance fluorescence lidar for measurements in the middle and upper atmosphere," in *Laser Remote Sensing*, T. Fujii and T. Fukuchi, eds.,(CRC Press, Taylor & Francis, 2005).
3. C.-Y. She, D. A. Krueger, R. Akmaev, H. Schmidt, E. Talaat, and S. Yee, "Long-term variability in mesopause region temperatures over Fort Collins, Colorado (41°N, 105°W) based on lidar observations from 1990 through 2007," J. Atmos. Terr. Phys. **71**, 1558–1564 (2009).
4. X. Chu, Z. Yu, C. S. Gardner, C. Chen, and W. Fong, "Lidar observations of neutral Fe layers and fast gravity waves in the thermosphere (110–155 km) at McMurdo (77.8°S, 166.7°E), Antarctica," Geophys. Res. Lett. **38**, L23807 (2011).
5. J. M. C. Plane, W. Feng, and E. C. M. Dawkins, "The mesosphere and metals: chemistry and changes," Chem. Rev. **115**, 4497–4541 (2015), Special issue: 2015 Chemistry in Climate.
6. W. Huang, X. Chu, C. S. Gardner, J. D. Carrillo-Sánchez, W. Feng, J. M. C. Plane, and D. Nesvorný, "Measurements of the vertical fluxes of atomic Fe and Na at the mesopause: implications for the velocity of cosmic dust entering the atmosphere," Geophys. Res. Lett. **42**, 169–175 (2015).
7. J. D. Carrillo-Sánchez, D. Nesvorný, P. Pokorný, D. Janches, and J. M. C. Plane, "Sources of cosmic dust in the Earth's atmosphere," Geophys. Res. Lett. **43**, 11979–11986 (2016).
8. A. Z. Liu, Y. Guo, F. Vargas, and G. R. Swenson, "First lidar measurements of horizontal wind and temperature in the lower thermosphere (105–140 km) with a Na lidar at Andes lidar observatory," Goephys. Res. Lett. **43**, 2374–2380 (2016).

9. T. Yuan, S. C. Solomon, C.-Y. She, D. A. Krueger, and H.-L. Liu, "The long-term trends of nocturnal mesopause temperature and altitude revealed by Na lidar observations between 1990 and 2018 at midlatitude," J. Geophys. Res. Atmos. **124**, 5970–5980 (2019).
10. T. J. Duck, J. A. Whiteway, and A. I. Carswell, "The gravity-wave arctic stratospheric vortex interaction," J. Atmos. Sci. **58**, 3581–3596 (2001).
11. R. J. Sica and A. Haefele, "Retrieval of temperature from multiple-channel Rayleigh-scatter lidar using an optimal estimation technique," Appl. Opt. **54**, 1872–1889 (2015).
12. X. Chu, J. Zhao, X. Lu, V. L. Harvey, R. M. Jones, E. Becker, C. Chen, W. Fong, Z. Yu, B. R. Roberts, and A. Dörnbrack, "Lidar observations of stratospheric gravity waves from 2011 to 2015 at McMurdo (77.84°S, 166.69°E), Antarctica: 2. Potential energy densities, lognormal distributions, and seasonal variations," J. Geophys. Res. Atmos. **123**, 7910–7934 (2018).
13. C. S. Gardner and A. Z. Liu, "Measuring eddy heat, constituent, and momentum fluxes with high-resolution Na and Fe Doppler lidars," J. Geophys. Res. Atmos. **119**, 10583–10603 (2014).
14. Y. Guo, A. Z. Liu, and C. S. Gardner, "First Na lidar measurements of turbulence heat flux, thermal diffusivity, and energy dissipation rate in the mesopause region," Geophys. Res. Lett. **44**, 5782–5790 (2017).
15. C. S. Gardner, Y. Guo, and A. Z. Liu, "Parameterizing wave-driven vertical constituent transport in the upper atmosphere," Earth Space Sci. **6**, 904–913 (2019).
16. H. L. Liu, "Variability and predictability of the space environment as related to lower atmosphere forcing," Space Weather **14**, 634–658 (2016).
17. C. S. Gardner and W. Yang, "Measurements of the dynamical cooling rate associated with the vertical transport of heat by dissipating gravity waves in the mesopause region at the Starfire optical range, New Mexico," J. Geophys. Res. **103**, 16909–16926 (1998).
18. L. Su, R. L. Collins, D. A. Krueger, and C.-Y. She, "Statistical analysis of sodium Doppler wind-temperature lidar measurements of vertical heat flux," J. Atmos. Ocean. Technol. **25**, 401–415 (2007).
19. W. Huang, X. Chu, C. S. Gardner, Z. Wang, W. Fong, J. A. Smith, and B. R. Roberts, "Simultaneous, common-volume lidar observations and theoretical studies of correlations among Fe/Na layers and temperatures in the mesosphere and lower thermosphere at Boulder table mountain (40°N, 105°W), Colorado," J. Geophys. Res. Atmos. **118**, 8748–8759 (2013).
20. C. S. Gardner, "Testing theories of atmospheric gravity wave saturation and dissipation," J. Atmos. Terr. Phys. **58**, 1575–1589 (1996).
21. P. K. Kundu, *Fluid Dynamics* (Academic, 1990), p. 638.
22. K. A. Emanuel, *Atmospheric Convection* (Oxford University, 1994), p. 580.
23. Y. Zhao, A. Z. Liu, and C. S. Gardner, "Measurements of atmospheric stability in the mesopause region at Starfire optical range," J. Atmos. Sol. Terr. Phys. **65**, 219–232 (2003).
24. C. S. Gardner and A. Z. Liu, "Seasonal variations of the vertical fluxes of heat and horizontal momentum in the mesopause region at Starfire optical range, New Mexico," J. Geophys. Res. **112**, D09113 (2007).
25. T. E. VanZandt, "A universal spectrum of buoyancy waves in the atmosphere," Geophys. Res. Lett. **9**, 575–578 (1982).
26. T. Tsuda, T. Inoue, D. C. Fritts, T. E. VanZandt, S. Kato, T. Sato, and S. Fukao, "MST radar observations of a saturated gravity wave spectrum," J. Atmos. Sci. **46**, 2440–2447 (1989).
27. Y. Muraoka, S. Fukao, T. Sugiyama, M. Yamamoto, T. Nakamura, T. Tsuda, and S. Kato, "Frequency spectra of mesospheric wind fluctuations observed with the MU radar," Geophys. Res. Lett. **17**, 1897–1900 (1990).
28. C. S. Gardner, "Diffusive filtering theory of gravity wave spectra in the atmosphere," J. Geophys. Res. **99**, 20601–20622 (1994).