New Techniques to Eliminate Photon-Noise Biases from Lidar and Radar Data for Tracing Gravity Wave Energy from the Stratosphere to the MLT

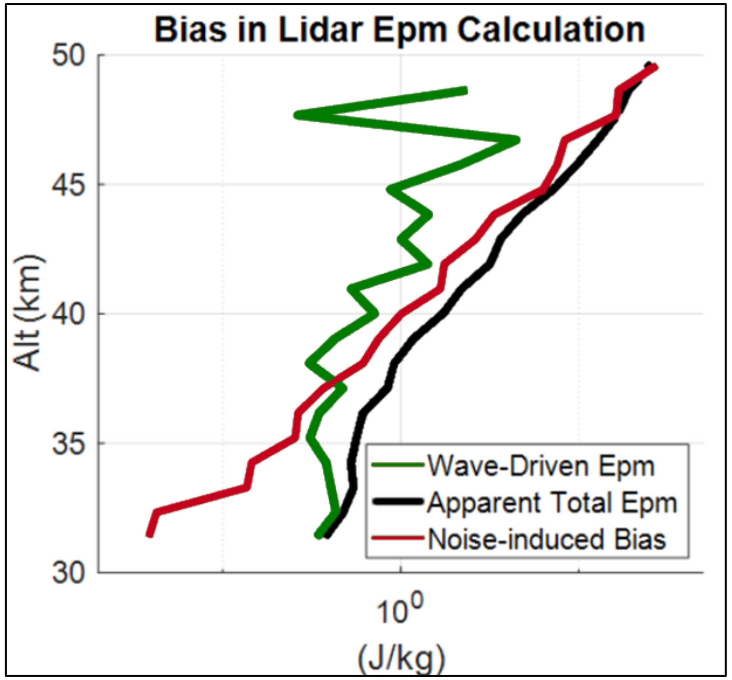
Jackson Jandreau[1,2]
Xinzhao Chu[1,2]

1. University of Colorado Boulder Smead Aerospace
2. CIRES Institute

Motivation

Lidar is a powerful way to study atmospheric waves by analyzing second order statistics of the lidar data. However, **noise in lidar measurements biases second order parameters** like variance and flux, and this bias is often strong enough to entirely prevent reliable calculation of the wave behavior. **Eq. 1** shows the calculation of variance, how the perturbations are comprised of two components, and how this calculation results in a bias of $\overline{(\Delta r)^2}$. The figure on the right demonstrates the strength of this bias under low-SNR conditions, emphasizing the necessity of applying some form of correction method.

This study utilizes potential energy density (Epm), to quantify wave energy, as it directly scales with atmospheric variance and makes a good demonstration of these bias correction methods.



$$\textbf{1.}\ Var[r'_{total}(z)] = \overline{[r'_{total}(z)]^2} = \overline{(r'+\Delta r)^2} = \underbrace{\overline{(r')^2}}_{\text{Signal of interest}} + \underbrace{\overline{(\Delta r)^2}}_{\text{Bias}} + \underbrace{\overline{(2r'\Delta r)}}_{\text{Uncorrelated, trends to 0}}$$

Introduction of variables used:

Atmospheric parameter: $\boldsymbol{r}$ **(density, temperature)**

Perturbation: $\boldsymbol{r' = r - r_{background}}$

Observed perturbation: $\boldsymbol{r'_{total}}$

Wave-induced perturbation: $\boldsymbol{r'}$

Noise induced perturbation: $\boldsymbol{\Delta r}$

$$\textbf{2.}\ E_{pm}(z) = \frac{1}{2}\left[\frac{g(z)}{N(z)}\right]^2 \cdot \frac{Var[r'(z)]}{r^2_{Bkg}(z)}$$

Gravitational Acceleration: $\boldsymbol{g}$

Buoyancy Frequency: $\boldsymbol{N^2}$

Variance Subtraction (VS)

Noise-Variance Subtraction (VS) is the traditional solution to this bias. This method utilizes parameter error to estimate the strength of the noise-induced variance and subtracts it from the total variance. **Eq. 3** and **4** are from Whiteway & Carswell (1995). It is easy to calculate and easy to apply with very low computational expense.

Performance: Regardless of the amount of data, the VS method can yield negative values for variance/Epm if the error is high enough (i.e., bad data conditions, high altitude, etc.) as seen in the below Figs. a-d. More data improves precision but does not drive result closer to model.

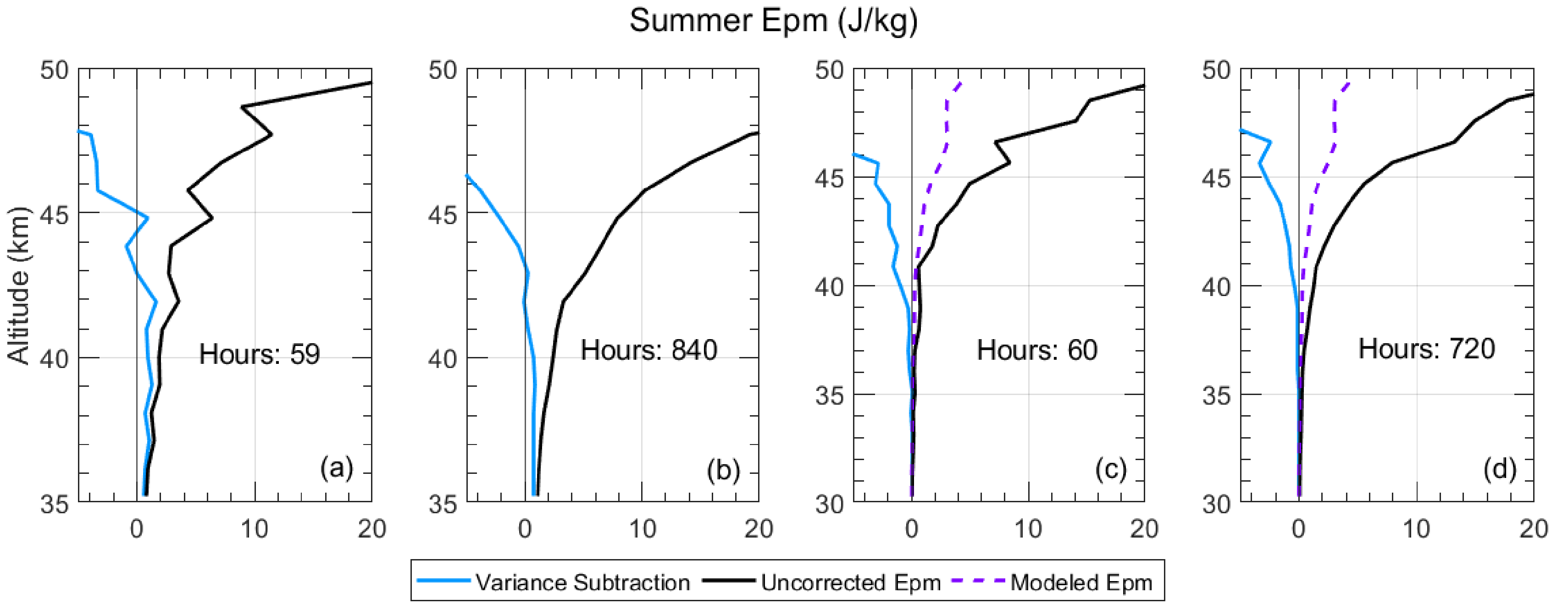
Atmospheric parameter: r (density, temperature)

Parameter error: $\boldsymbol{\delta r}$

$$\textbf{3.}\ Var_{VS}(\Delta r) = \frac{1}{k}\sum_{i=1}^{k}[\delta r(z,t_i)]^2$$

$$\textbf{4.}\ E_{pm}(z) = \frac{1}{2}\left[\frac{g(z)}{N(z)}\right]^2 \cdot \frac{Var(r'_{Total}) - Var_{VS}(\Delta r)}{r_{Bkg}^2(z)}$$

Figure Note: *These methods are demonstrated on short and long sets of McMurdo lidar data (a and b) and forward models (c and d)*

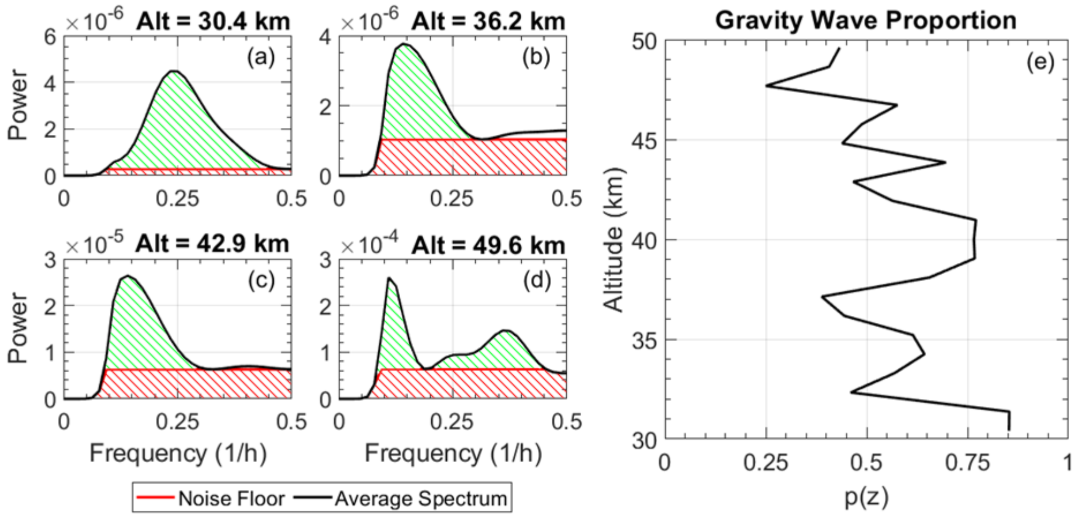


New Techniques to Eliminate Photon-Noise Biases from Lidar and Radar Data for Tracing Gravity Wave Energy from the Stratosphere to the MLT

Spectral Proportion (SP)

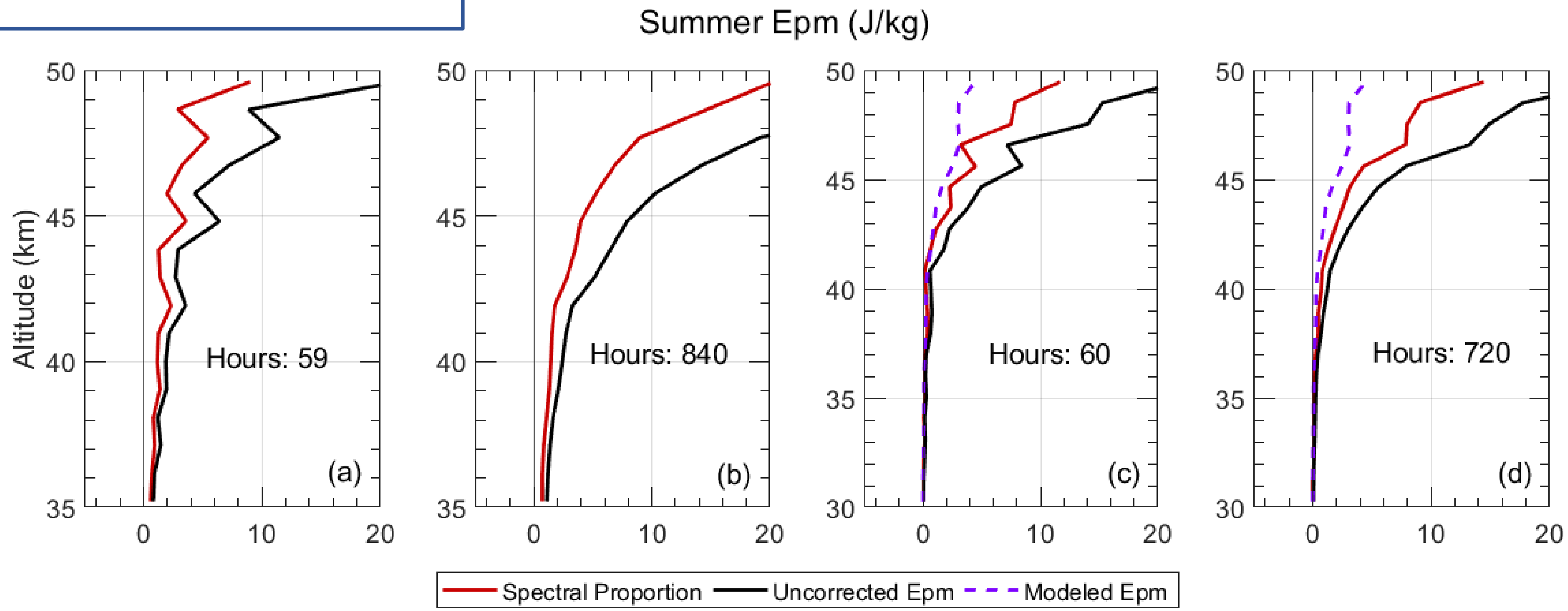
The **Spectral Proportion Method** was developed by Chu et al., 2018. This method involves a Monte Carlo simulation where 1000 replicas of the existing observation are generated, random noise is applied onto each (scaled by parameter error value), and a 1D-FFT of each noisy-replica is averaged at every altitude. Then, we find noise floor of resulting average and calculate $p(z)$ via **Eq. 5**. Results using SP are always positive due to scaling as opposed to subtraction

Performance: Yields results much closer to modeled Epm (Figs. c and d), though it overestimates under low-SNR. Precision is increased greatly by the addition of more samples, yet accuracy (proximity to model result) remains similar regardless of sample size.

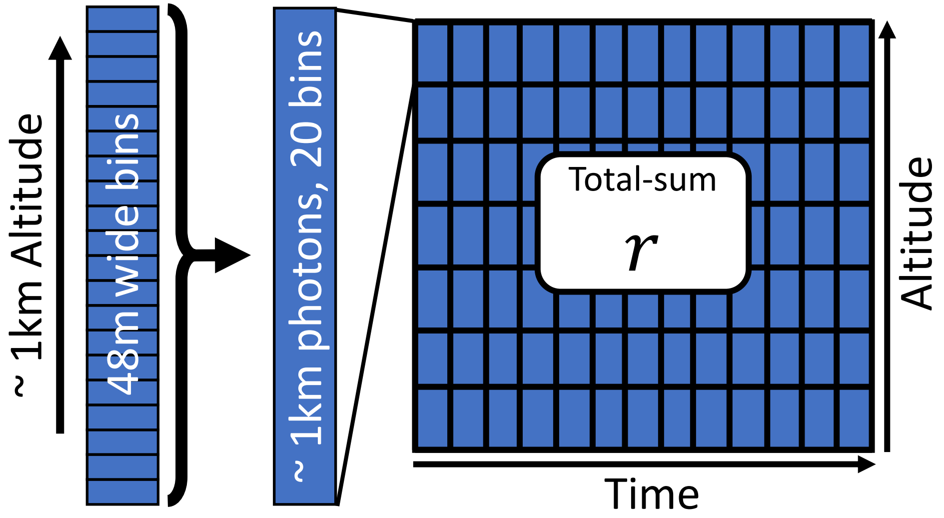


5. $p(z) = \dfrac{\textit{Area under curve} - \textit{Area under floor}}{\textit{Area under curve}}$

6. $E_{pm,SP}(z) = p(z) \cdot E_{pm,Total}(z)$

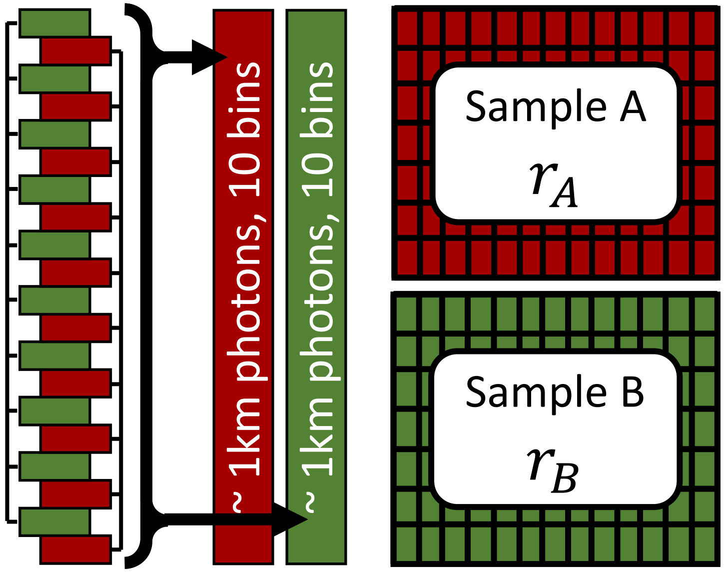


Interleaved Method (INT)



The **Interleaved Method** was developed by Gardner & Chu 2020. This method is more involved than the previous two and starts earlier in the data processing procedure. This technique begins with the photon bins at the earliest level. When summing the photons into larger-sized bins, as is typical of lidar processing, we instead take alternating bins and create two samples, as demonstrated in the figure on the left. It is essential to interleave in such a way as finely as possible so that samples A and B represent the most similar parcel of atmosphere as possible, therefore the interleaving is done on raw photon bins before any other integration is done. The results is like having two independent, adjacent lidar systems, yet only requires one.

Below is demonstrated the reason for this splitting. We substitute the variance used in prior Epm calculations for the covariance of the two samples as shown in **Eq. 8**. We then have a variance without a bias term, which is only dependent on wave perturbations as the noise dependent terms have dropped out due to non-correlation.



Three calculated perturbations

7.
$$\begin{cases} \boldsymbol{r'} = r' + \Delta r \\ \boldsymbol{r_A}' = r_A' + \Delta r_A \\ \boldsymbol{r_B}' = r_B' + \Delta r_B \end{cases}$$

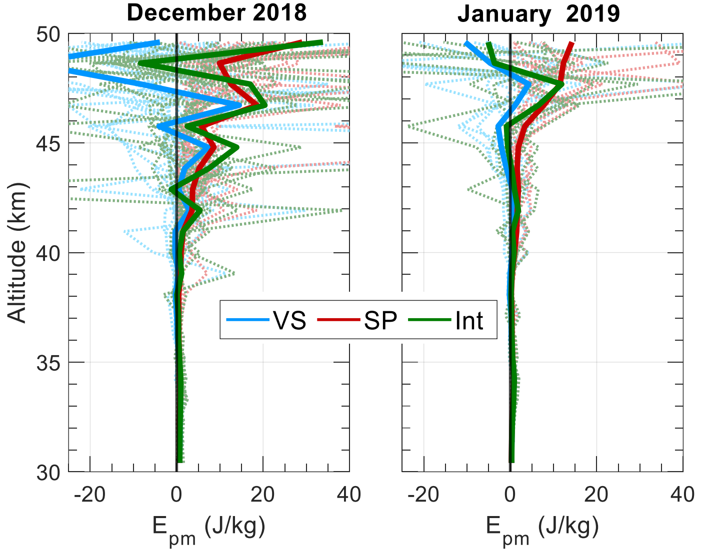
Calculating second order parameters

8.
$$\begin{cases} Variance = \overline{(\boldsymbol{r'})^2} = \overline{(r')^2} + \overline{(\Delta r)^2} + \cancel{\overline{2r'\Delta r}} \\ \quad = Var_{Waves} + Var_{Noise} \\ Covariance = \overline{\boldsymbol{r'}_A \boldsymbol{r'}_B} = \\ \overline{r'_A r'_B + \cancel{r'_A \Delta r_B} + \cancel{r'_B \Delta r_A} + \cancel{\Delta r_A \Delta r_B}} \\ \quad = Cov_{Wave} \end{cases}$$

Crossed terms drop due to non-correlation!

Interleaved Method (INT)

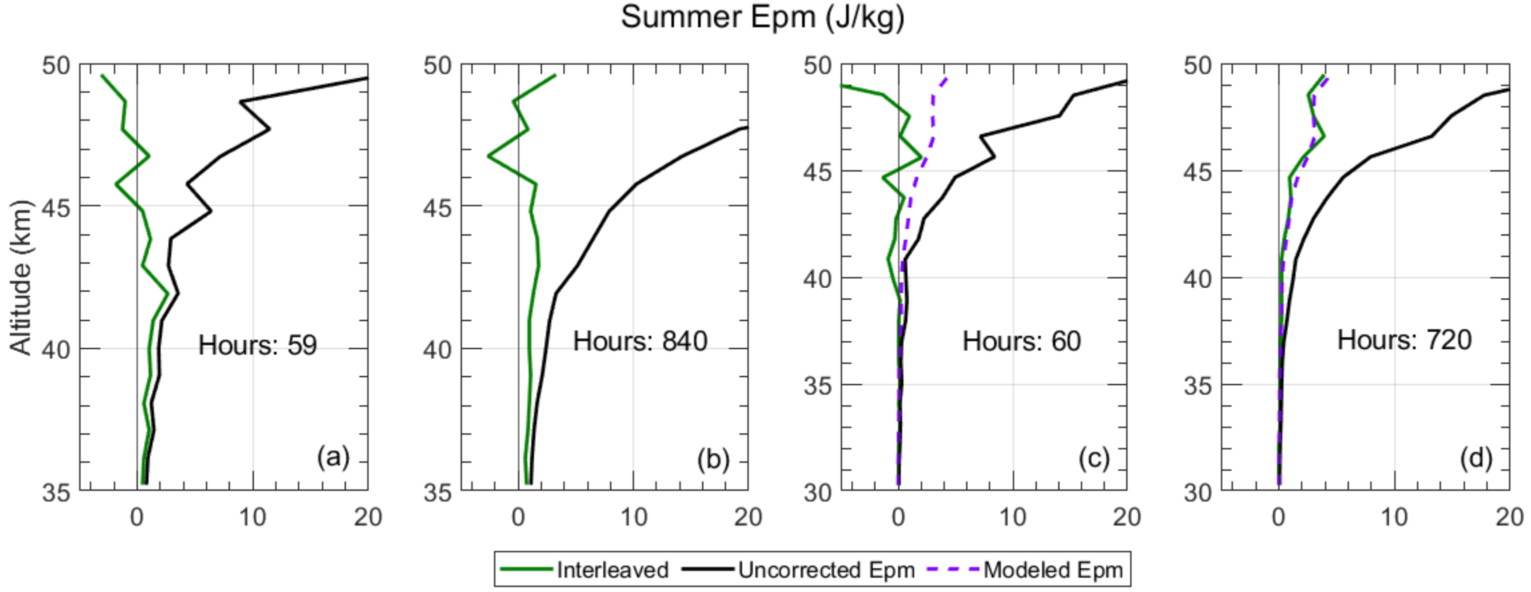
The biggest thing to consider when using the interleaved method is that by splitting the photon counts in half two create the two samples, the SNR has been decreased considerably. This results in some individual runs being even noisier than the noisy VS runs were, and some rune even having negative values as shown in the Figure on the right. The difference between INT and VS, however, is that the INT profiles are not consistently negative, and are just noisy, meaning that under sufficient sample size, the resulting profile will begin to trend back towards all positive values, the real Epm as determined by atmospheric perturbations. This can be seen in the bold profile in the Figure on the right, where the light profiles show individual runs.



From Jandreau and Chu, 2022

The results of this method are generally the strongest of the three. The Figure to the right clearly shows that the Epm generally is positive, and that under sufficient sample size, very well approximates the true atmospheric Epm. This is especially evident in (d), though (b)'s analysis of real data is still noisy in the upper altitudes and has a few negative bins.

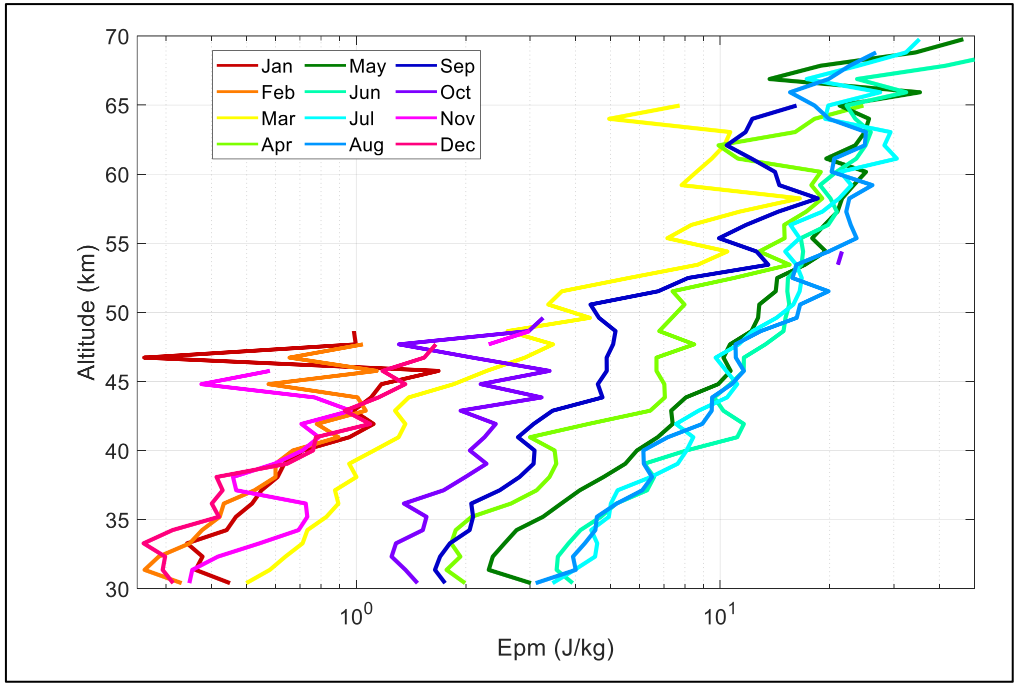
$$9.\ E_{pm,INT} = \frac{1}{2}\left[\frac{g(z)}{N(z)}\right]^2 \cdot \frac{Cov(T_A',T_B')}{T_{Bkg}(z)^2}$$



Conclusions and Results

The goal of this study was to determine when each method should be used. By the results in the last few sections, two main factors drive this decision. The most important is the SNR of the data followed by the amount of data available for the study. Each method responds differently to changes in each of these variables, leading us to determine the following guidelines:

- Variance Subtraction
  - Should be used only under very high SNR such as low-altitude measurements. Its major benefit is ease of application, yet it fails easily.
- Spectral Proportion
  - This method has very strong performance under high-SNR, yielding near perfect results in simulated winter conditions (Jandreau and Chu, 2022) and agreeing with INT at most altitudes. However, under low-SNR, it is difficult to determine the noise floor and the result undercorrects the bias (resulting in overestimated variance/Epm). It should be used when a small set of samples is being analyzed.
- Interleaved Method
  - The interleaved method is the only one of the three which yields can statistically eliminate the bias entirely. However, this method fails heavily when there are not sufficient samples available. It should be utilized whenever a large batch of samples is being processed, as this will counteract the increased uncertainty resulting in a reliable result.



These results showcase the application of the newly developed methods. Previous studies of the dataset were only able to reliably reach ~60 km in the winter, with summer 50 km results not being nearly as reliable. While the summer profiles here still become noisy at their upper levels, the trends are still observable. This study has enabled a detailed look into McMurdo's upper atmosphere wave dynamics, a work which is now in progress.

References:

Chu, X., et al.: Lidar Observations of Stratos. … Potential Energy Densities, Lognormal Distributions, and Seasonal Variations. JGR: Atmos. 123(15), 7910–7934. (2018).

Whiteway, J. A., & Carswell, A. I.: Lidar observations of gravity wave activity in the upper stratosphere over Toronto. Journal of Geophysical Research, 100(D7), 14113. (1995).

Gardner, C. S., & Chu, X.: Eliminating photon noise biases in the computation of second-order … and species measurements. Applied Optics, 59(27), 8259–8271. (2020).

Chu, X. & Yu, Z., et al.: Lidar observations of neutral Fe layers … thermosphere (110-155 km) at McMurdo (77.8°S, 166.7°E), Antarctica. GRL, 38, L23807. (2011).

Jandreau, J., & Chu, X.: Comparison of Three Methodologies for Removal of Random-Noise- … Measurements. Earth and Space Science, 9, e2021EA002073. (2022).