Lecture 30. Lidar Data Inversion (1)

- Introduction of data inversion
- Resonance-Doppler lidar and basic ideas (clues) for lidar data inversion
- Data retrieval procedure
  - Procedure overview
  - Preprocess
  - Profile process (next lecture)
  - Main process (next lecture)
  - Error analysis (next lecture)
- Summary

From Raw Data to Physical Parameters

Use resonance Doppler lidar as an example

Raw Data

N_S (R)

T (R)

V_R (R)

n_c (R)

β (R)

α (R)

δ (R)

… …

PMC, PSC, Aerosols
Clouds
Constituent Density
Sporadic Layers
Meteors
Climatology
GWs, Tides, PWs
Momentum Flux
Heat Flux
Constituent Flux
Instability
… … …

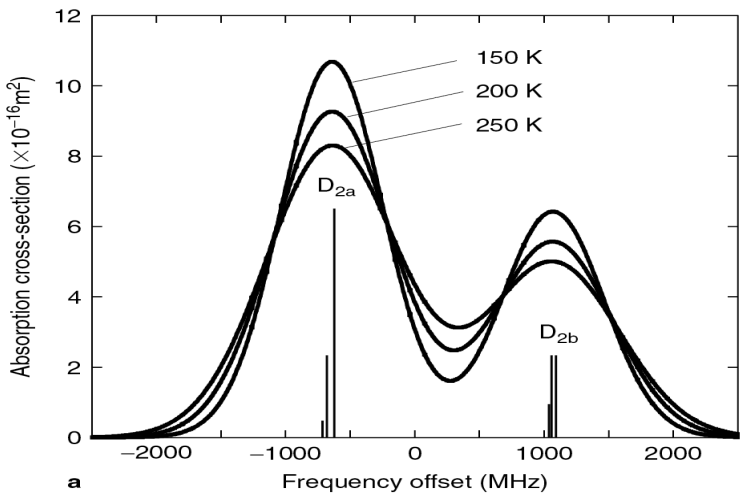
Data Inversion
& Error Analysis
Data Retrieval

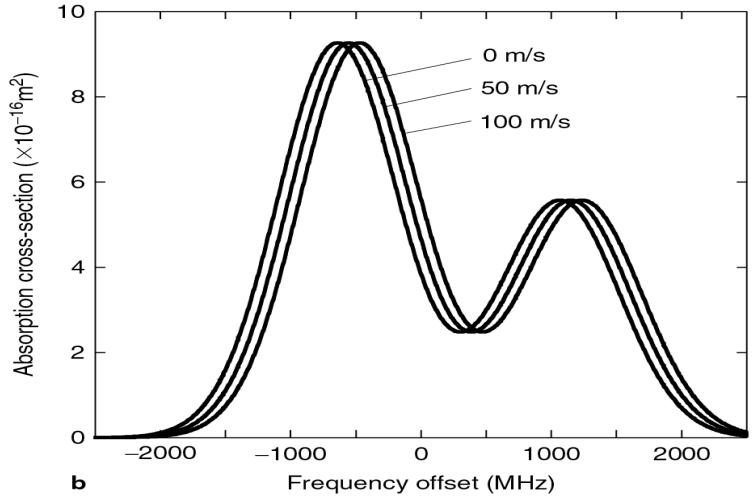
Data Analysis
& Interpretation
Science Study

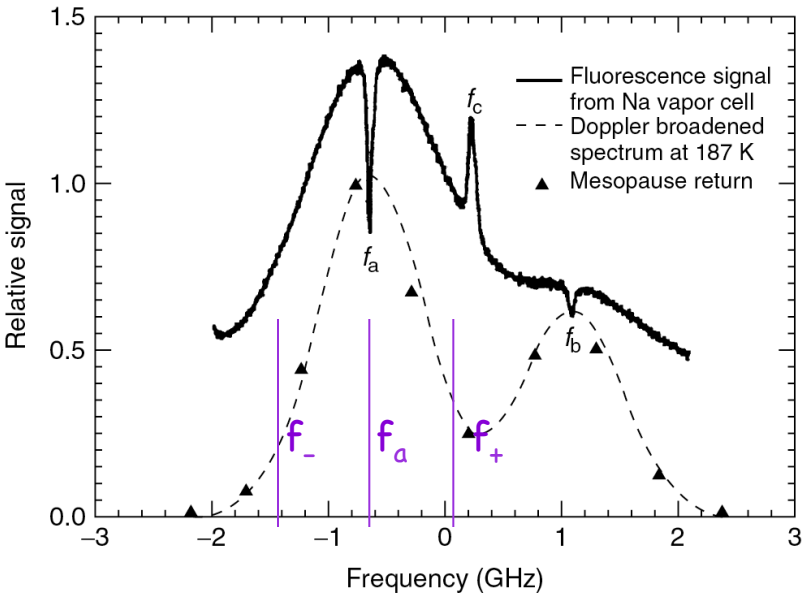
Introduction: Lidar Data Inversion

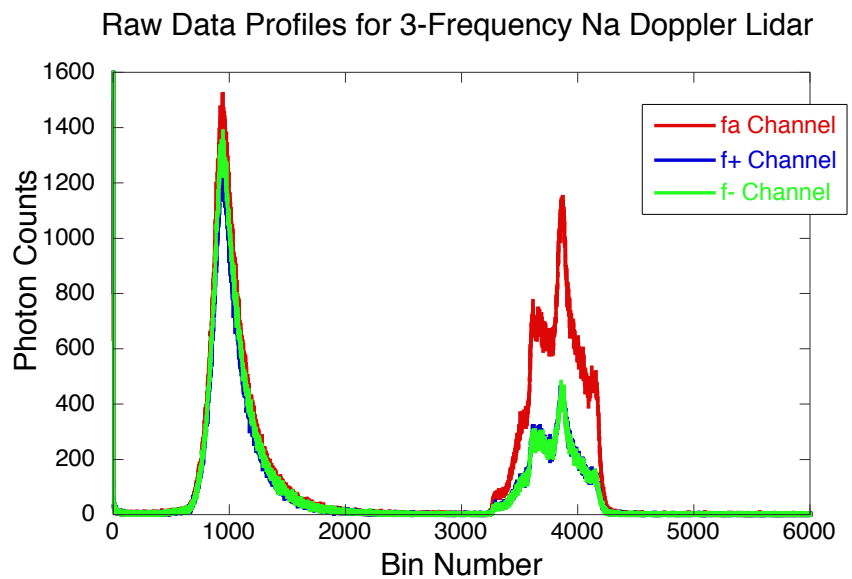
- Lidar data inversion deals with the problems of how to derive meaningful physical parameters from raw data.
- Raw data are usually a column or a row of photon counts, where the positions of photon counts in the column or row mark their time bins, thus the ranges or heights.
- Data inversion is basically a reverse procedure to the development of lidar equation.
- It is necessary to understand the detailed physical procedure from light transmitting, to light propagation, to light interaction with objects, and to light detection, in order to conduct data inversion correctly.
- In this lecture we discuss the data inversion for Na Doppler lidar (K and Fe lidar would be similar) as an example.

Resonance Fluorescence Doppler Lidar









Example Lidar Raw Signals

Doppler-Limited Na Spectroscopy

- Doppler-broadened Na absorption cross-section is approximated as a Gaussian with rms width σ_D

$$\sigma_{abs}(\nu) = \frac{1}{\sqrt{2\pi}\sigma_D}\frac{e^2 f}{4\varepsilon_0 m_e c}\sum_{n=1}^{6} A_n \exp\left(-\frac{[\nu_n - \nu(1 - V_R/c)]^2}{2\sigma_D^2}\right)$$

- Assume the laser lineshape is a Gaussian with rms width σ_L
- The effective cross-section is the convolution of the atomic absorption cross-section and the laser lineshape

$$\sigma_{eff}(\nu) = \frac{1}{\sqrt{2\pi}\sigma_e}\frac{e^2 f}{4\varepsilon_0 m_e c}\sum_{n=1}^{6} A_n \exp\left(-\frac{[\nu_n - \nu(1 - V_R/c)]^2}{2\sigma_e^2}\right)$$

where $\sigma_e = \sqrt{\sigma_D^2 + \sigma_L^2}$ and $\sigma_D = \sqrt{\frac{k_B T}{M\lambda_0^2}}$

The effective cross-section depends on both T and V_R.
How to infer both variables from the lidar-measured σ_{eff}?

Basic Clue (1): Lidar Equation & Solution

- From lidar equation and its solution to derive preprocess procedure of lidar data inversion

$$N_S(\lambda,z) = \left(\frac{P_L(\lambda)\Delta t}{hc/\lambda}\right)\left[\sigma_{eff}(\lambda,z)n_c(z)R_B(\lambda) + 4\pi\sigma_R(\pi,\lambda)n_R(z)\right]\Delta z\left(\frac{A}{4\pi z^2}\right) \times \left(T_a^2(\lambda)T_c^2(\lambda,z)\right)\left(\eta(\lambda)G(z)\right) + N_B$$

+

$$N_S(\lambda,z_R) = \left(\frac{P_L(\lambda)\Delta t}{hc/\lambda}\right)\left[\sigma_R(\pi,\lambda)n_R(z_R)\right]\Delta z\left(\frac{A}{z_R^{\,2}}\right)T_a^{\,2}(\lambda,z_R)\left(\eta(\lambda)G(z_R)\right) + N_B$$

⇩

$$N_{Norm}(\lambda,z) = \frac{N_{Na}(\lambda,z)}{N_R(\lambda,z_R)T_c^2(\lambda,z)}\frac{z^2}{z_R^{\,2}} = \frac{\sigma_{eff}(\lambda,z)n_c(z)}{\sigma_R(\pi,\lambda)n_R(z_R)}\frac{1}{4\pi}$$

$$= \frac{N_S(\lambda,z) - N_B}{N_S(\lambda,z_R) - N_B}\frac{z^2}{z_R^{\,2}}\frac{1}{T_c^2(\lambda,z)} - \frac{n_R(z)}{n_R(z_R)}$$

Physics

Photon Counts

Principle of Doppler Ratio Technique

- Lidar equation for resonance fluorescence (Na, K, or Fe)

$$N_S(\lambda,z) = \left(\frac{P_L(\lambda)\Delta t}{hc/\lambda}\right)\left[\sigma_{eff}(\lambda,z)n_c(z)R_B(\lambda) + \sigma_R(\pi,\lambda)n_R(z)\right]\Delta z\left(\frac{A}{4\pi z^2}\right) \times \left(T_a^2(\lambda)T_c^2(\lambda,z)\right)\left(\eta(\lambda)G(z)\right) + N_B$$

R_B = 1 for current Na Doppler lidar since return photons at all wavelengths are received by the broadband receiver, so no fluorescence is filtered off.

- Pure Na signal and pure Rayleigh signal in Na region are

$$N_{Na}(\lambda,z) = \left(\frac{P_L(\lambda)\Delta t}{hc/\lambda}\right)\left[\sigma_{eff}(\lambda,z)n_c(z)\right]\Delta z\left(\frac{A}{4\pi z^2}\right)\left(T_a^2(\lambda)T_c^2(\lambda,z)\right)\left(\eta(\lambda)G(z)\right)$$

$$N_R(\lambda,z) = \left(\frac{P_L(\lambda)\Delta t}{hc/\lambda}\right)\left[\sigma_R(\pi,\lambda)n_R(z)\right]\Delta z\left(\frac{A}{z^2}\right)\left(T_a^2(\lambda)T_c^2(\lambda,z)\right)\left(\eta(\lambda)G(z)\right)$$

- So we have

$$N_S(\lambda,z) = N_{Na}(\lambda,z) + N_R(\lambda,z) + N_B$$

Principle of Doppler Ratio Technique

- Lidar equation at pure molecular scattering region (35-55km)

$$N_S(\lambda,z_R)=\left(\frac{P_L(\lambda)\Delta t}{hc/\lambda}\right)\left[\sigma_R(\pi,\lambda)n_R(z_R)\right]\Delta z\left(\frac{A}{z_R^{\,2}}\right)T_a^{\,2}(\lambda,z_R)\left(\eta(\lambda)G(z_R)\right)+N_B$$

- Pure Rayleigh signal in molecular scattering region is

$$N_R(\lambda,z_R)=\left(\frac{P_L(\lambda)\Delta t}{hc/\lambda}\right)\left[\sigma_R(\pi,\lambda)n_R(z_R)\right]\Delta z\left(\frac{A}{z_R^{\,2}}\right)T_a^{\,2}(\lambda,z_R)\left(\eta(\lambda)G(z_R)\right)$$

- So we have

$$N_S(\lambda,z_R)=N_R(\lambda,z_R)+N_B$$

- The ratio between Rayleigh signals at z and z_R is given by

$$\frac{N_R(\lambda,z)}{N_R(\lambda,z_R)}=\frac{\left[\sigma_R(\pi,\lambda)n_R(z)\right]T_a^{\,2}(\lambda,z)T_c^2(\lambda,z)G(z)}{\left[\sigma_R(\pi,\lambda)n_R(z_R)\right]T_a^{\,2}(\lambda,z_R)G(z_R)}\frac{z_R^{\,2}}{z^2}=\frac{n_R(z)}{n_R(z_R)}\frac{z_R^{\,2}}{z^2}T_c^2(\lambda,z)$$

Where n_R is the (total) atmospheric number density, usually obtained from atmospheric models like MSIS00.

Principle of Doppler Ratio Technique

- From above equations, the pure Na and Rayleigh signals are

$$N_{Na}(\lambda,z) = N_S(\lambda,z) - N_B - N_R(\lambda,z)$$

$$N_R(\lambda,z_R) = N_S(\lambda,z_R) - N_B$$

- Normalized Na photon count is defined as

$$N_{Norm}(\lambda,z) = \frac{N_{Na}(\lambda,z)}{N_R(\lambda,z_R)T_c^2(\lambda,z)}\frac{z^2}{z_R^2}$$

- From physics point of view, the normalized Na count is

$$N_{Norm}(\lambda,z) = \frac{N_{Na}(\lambda,z)}{N_R(\lambda,z_R)T_c^2(\lambda,z)}\frac{z^2}{z_R^2} = \frac{\sigma_{eff}(\lambda,z)n_c(z)}{\sigma_R(\pi,\lambda)n_R(z_R)}\frac{1}{4\pi}$$

- From actual photon counts, the normalized Na count is

$$N_{Norm}(\lambda,z) = \frac{N_{Na}(\lambda,z)}{N_R(\lambda,z_R)T_c^2(\lambda,z)}\frac{z^2}{z_R^2} = \frac{N_S(\lambda,z) - N_B - N_R(\lambda,z)}{N_R(\lambda,z_R)T_c^2(\lambda,z)}\frac{z^2}{z_R^2}$$

$$= \frac{N_S(\lambda,z) - N_B}{N_S(\lambda,z_R) - N_B}\frac{z^2}{z_R^2}\frac{1}{T_c^2(\lambda,z)} - \frac{n_R(z)}{n_R(z_R)}$$

Basic Clue (2): Ratio Computation

- From physics, we calculate the ratios of R_T and R_W as

$$R_T = \frac{\sigma_{eff}(f_+,z) + \sigma_{eff}(f_-,z)}{\sigma_{eff}(f_a,z)}$$

$$R_W = \frac{\sigma_{eff}(f_+,z) - \sigma_{eff}(f_-,z)}{\sigma_{eff}(f_a,z)}$$

- From actual photon counts, we calculate the ratios as

$$R_T = \frac{N_{Norm}(f_+,z) + N_{Norm}(f_-,z)}{N_{Norm}(f_a,z)}$$

$$R_W = \frac{N_{Norm}(f_+,z) - N_{Norm}(f_-,z)}{N_{Norm}(f_a,z)}$$

Note: There are other formats of temperature and wind ratios R_T and R_W that are much more complicated than these two simple ratios. The purpose is to reduce the cross-talk between temperature and LOS wind.

Principle of Doppler Ratio Technique

- From physics, the ratios of R_T and R_W are then given by

$$R_T = \frac{N_{Norm}(f_+,z) + N_{Norm}(f_-,z)}{N_{Norm}(f_a,z)} = \frac{\dfrac{\sigma_{eff}(f_+,z)n_c(z)}{\sigma_R(\pi,f_+)n_R(z_R)} + \dfrac{\sigma_{eff}(f_-,z)n_c(z)}{\sigma_R(\pi,f_-)n_R(z_R)}}{\dfrac{\sigma_{eff}(f_a,z)n_c(z)}{\sigma_R(\pi,f_a)n_R(z_R)}} = \frac{\sigma_{eff}(f_+,z) + \sigma_{eff}(f_-,z)}{\sigma_{eff}(f_a,z)}$$

$$R_W = \frac{N_{Norm}(f_+,z) - N_{Norm}(f_-,z)}{N_{Norm}(f_a,z)} = \frac{\dfrac{\sigma_{eff}(f_+,z)n_c(z)}{\sigma_R(\pi,f_+)n_R(z_R)} - \dfrac{\sigma_{eff}(f_-,z)n_c(z)}{\sigma_R(\pi,f_-)n_R(z_R)}}{\dfrac{\sigma_{eff}(f_a,z)n_c(z)}{\sigma_R(\pi,f_a)n_R(z_R)}} = \frac{\sigma_{eff}(f_+,z) - \sigma_{eff}(f_-,z)}{\sigma_{eff}(f_a,z)}$$

Here, Rayleigh backscatter cross-section is regarded as the same for three frequencies, since the frequency difference is so small. Na number density is also the same for three frequency channels, and so is the atmosphere number density at Rayleigh normalization altitude.

Principle of Doppler Ratio Technique

- From actual photon counts, the ratios R_T and R_W are

$$R_T = \frac{N_{Norm}(f_+,z)+N_{Norm}(f_-,z)}{N_{Norm}(f_a,z)}$$

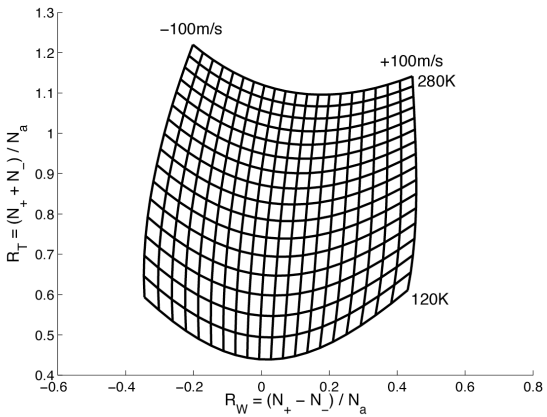
$$= \frac{\left(\frac{N_S(f_+,z)-N_B}{N_S(f_+,z_R)-N_B}\frac{z^2}{z_R{}^2}\frac{1}{T_c^2(f_+,z)}-\frac{n_R(z)}{n_R(z_R)}\right)+\left(\frac{N_S(f_-,z)-N_B}{N_S(f_-,z_R)-N_B}\frac{z^2}{z_R{}^2}\frac{1}{T_c^2(f_-,z)}-\frac{n_R(z)}{n_R(z_R)}\right)}{\frac{N_S(f_a,z)-N_B}{N_S(f_a,z_R)-N_B}\frac{z^2}{z_R{}^2}\frac{1}{T_c^2(f_a,z)}-\frac{n_R(z)}{n_R(z_R)}}$$

$$R_W = \frac{N_{Norm}(f_+,z)-N_{Norm}(f_-,z)}{N_{Norm}(f_a,z)}$$

$$= \frac{\left(\frac{N_S(f_+,z)-N_B}{N_S(f_+,z_R)-N_B}\frac{z^2}{z_R{}^2}\frac{1}{T_c^2(f_+,z)}-\frac{n_R(z)}{n_R(z_R)}\right)-\left(\frac{N_S(f_-,z)-N_B}{N_S(f_-,z_R)-N_B}\frac{z^2}{z_R{}^2}\frac{1}{T_c^2(f_-,z)}-\frac{n_R(z)}{n_R(z_R)}\right)}{\frac{N_S(f_a,z)-N_B}{N_S(f_a,z_R)-N_B}\frac{z^2}{z_R{}^2}\frac{1}{T_c^2(f_a,z)}-\frac{n_R(z)}{n_R(z_R)}}$$

Basic Clue (3): T, V_R, n_C, or β Derivation

- Compute actual ratios R_T and R_W from photon counts, and then look up these two ratios on the calibration curves to infer the corresponding Temperature (T) and Wind (V_R) from isoline/isogran



- If there is only R_T ratio, then infer the temperature from the calibration curve, like the Fe Boltzmann lidar case.

- Constituent (e.g., Na) density can be inferred from the peak freq signal

$$n_{Na}(z) = \frac{N_{norm}(f_a,z)}{\sigma_a} 4\pi n_R(z_R)\sigma_R = \frac{N_{norm}(f_a,z)}{\sigma_a} 4\pi \times 2.938 \times 10^{-32} \frac{P(z_R)}{T(z_R)} \cdot \frac{1}{\lambda^{4.0117}}$$

- Volume backscatter coefficient can be derived as

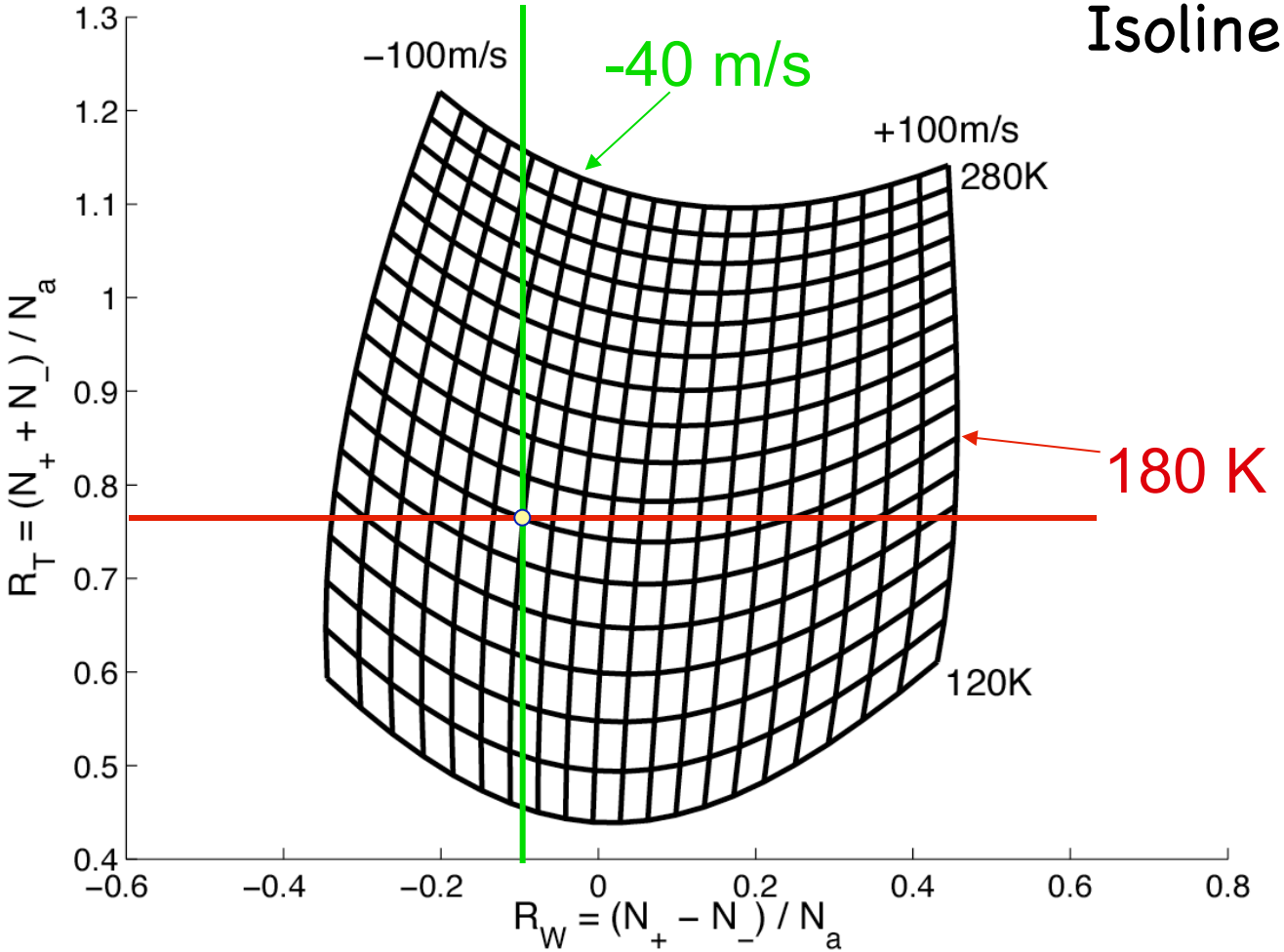
$$\beta_{PMC}(z) = \left[\frac{[N_S(z) - N_B] \cdot z^2}{[N_S(z_{RN}) - N_B] \cdot z_{RN}^2} - \frac{n_R(z)}{n_R(z_{RN})} \right] \cdot \beta_R(z_{RN})$$

where $$\beta_R(z_{RN},\pi) = \frac{\beta}{4\pi} P(\pi) = 2.938 \times 10^{-32} \frac{P(z_{RN})}{T(z_{RN})} \cdot \frac{1}{\lambda^{4.0117}}$$

How Does Ratio Technique Work?

- Compute Doppler calibration curves from physics
- Look up these two ratios on the calibration curves to infer the corresponding Temperature and Wind from isoline/isogram.

Isoline / Isogram



Considerations in Data Inversion

- How to obtain related information like date, time, location, base altitude, operation conditions?

-- from data header and other info sources

- How to obtain range or altitude information?

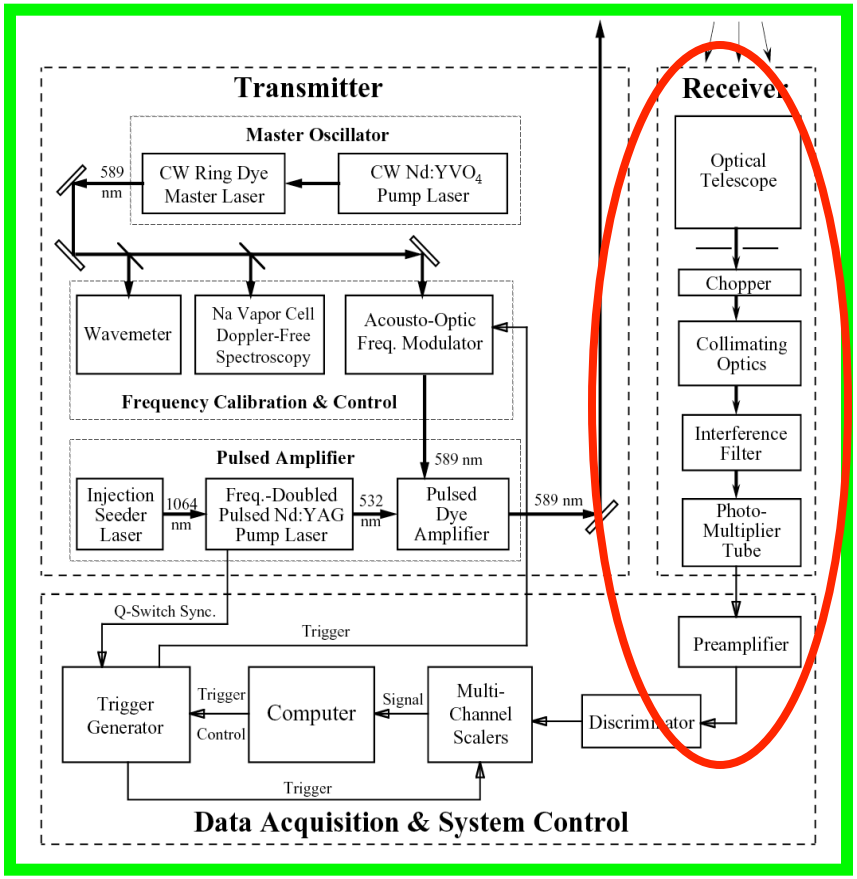
-- from bin number, data header and other source

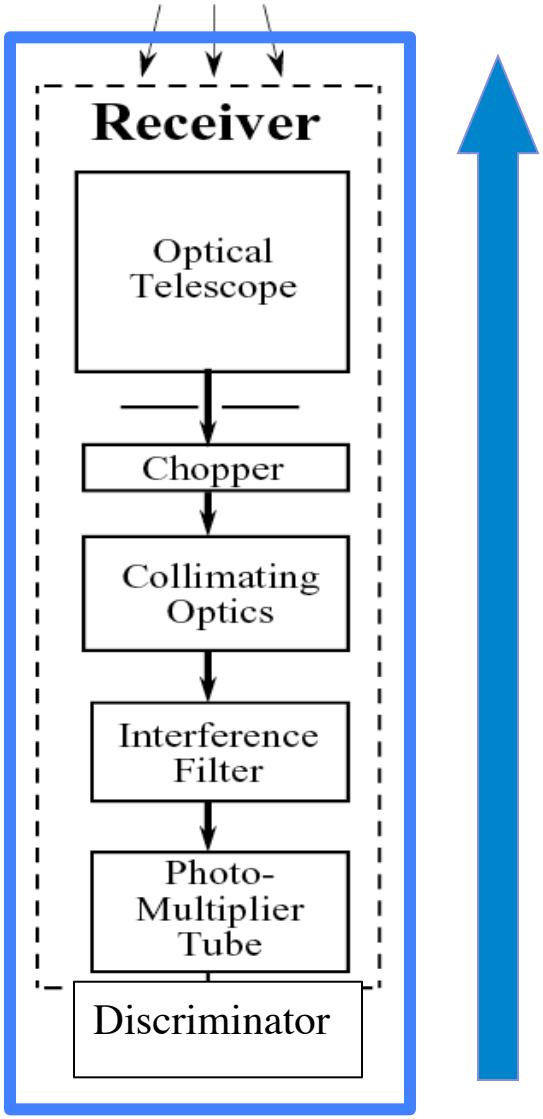
$$R = n_{bin} \cdot t_{bin} \cdot c/2$$

$$z = R \cdot \cos\theta + z_{base}$$

R is range, n_{bin} is bin number, t_{bin} is bin width in time, c is light speed, z is absolute altitude, θ is off-zenith angle, and z_{base} is the base altitude relative to sea-level.

Data Retrieval Procedure





Preprocess Procedure

Data inversion is a reverse procedure to lidar equation development.

Preprocess Procedure and Profile-Process Procedure for Na/Fe/K Doppler Lidar

Read Data File
↓
PMT/Discriminator Saturation Correction
↓
Chopper/Filter Correction
↓
Subtract Background
↓
Remove Range Dependence (x R^2)
↓
Add Base Altitude
↓
Estimate Rayleigh Signal @ z_R km
↓
Normalize Profile By Rayleigh Signal @ z_R km

- Read data: for each set, and calculate T, W, and n for each set
- PMT/Discriminator saturation correction
- Chopper/Filter correction

Integration

- Background estimate and subtraction
- Range-dependence removal (xR^2, not z^2)
- Base altitude adjustment
- Take Rayleigh signal @ z_R (Rayleigh fit or Rayleigh mean)
- Rayleigh normalization $N_N(\lambda,z) = \frac{N_S(\lambda,z) - N_B}{N_S(\lambda,z_R) - N_B} \frac{z^2}{z_R^2}$

Main Process

- Subtract Rayleigh signals from Na/Fe/K region after counting in the factor of T_C

Main Process

Load Atmosphere n_R, T, P Profiles from MSIS00

Start from Na layer bottom
T_C ($z=z_b$) = 1
Calculate Nnorm ($z=z_b$) from photon counts and MSIS number density for each freq

$$N_{Norm}(\lambda,z) = \frac{N_S(\lambda,z) - N_B}{N_S(\lambda,z_R) - N_B} \frac{z^2}{z_R{}^2} \frac{1}{T_c^2(\lambda,z)} - \frac{n_R(z)}{n_R(z_R)}$$

Calculate R_T and R_W from N_{Norm}

Are ratios reasonable?

Yes

No

Set to nominal values or MSIS
T = 200 K, W = 0 m/s

Create look-up table or calibration curves From physics

$$R_T = \frac{\sigma_{eff}(f_+,z) + \sigma_{eff}(f_-,z)}{\sigma_{eff}(f_a,z)}$$

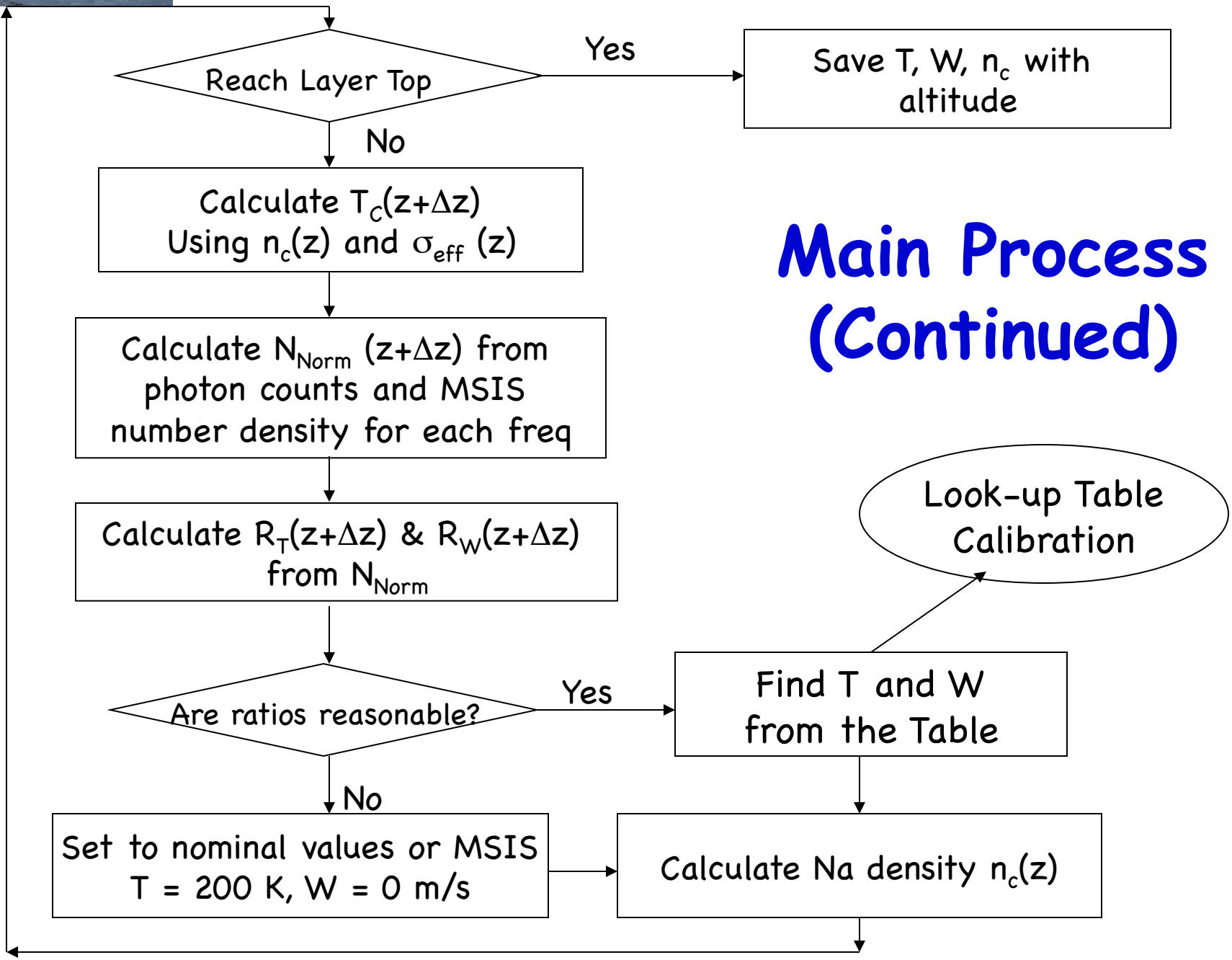
$$R_W = \frac{\sigma_{eff}(f_+,z) - \sigma_{eff}(f_-,z)}{\sigma_{eff}(f_a,z)}$$

Look-up Table Calibration

Find T and W from the Table

Calculate Na density $n_c(z)$

Main Process (Continued)



Considerations behind Profile-Process and Preprocess

- Indicated from the lidar equation and its solution, the profile process for Na Doppler lidar data is
  - Background estimation and subtraction (- N_B)
  - Range-dependence removal (x R^2)
  - Base altitude adjustment (+ z_{Base})
  - Rayleigh normalization [$1/(N_S(z_R)-N_B)$]

- More considerations on lidar hardware and detection - preprocess procedure
  - PMT and discriminator saturation correction
  - Chopper or electronic gain correction
  - Integration in time and/or range bins to obtain sufficient signal-to-noise ratio (SNR)

Step1. Read Raw Data

- Headers + One Column Photon Counts (ASCII or Binary)

Total Bin #	Low Bin #	# of Freq	Set No.	Profile No.	# of Shots
10240	1	3	10	1	1500
4	11	2	7.060833		
0	7	0	1	3.05	20.708	-156.258

Photon Counts:
12
1543
0
0
1574
4694
…
…
…

Month
Day
Time (UT)
Year-2000
Azimuth Angle
Bin Resolution
Off-Zenith Angle
Base Altitude (km)
Lat
Longi
Frequency Number

Another Example of Lidar Raw Data

Total Bin #	Low Bin #	# of Freq	Set No.	Profile No.	# of Shots
10240	1	3	11	2	1500
4	11	2	7.090278		
180	7	30	2	3.05	20.708	-156.258

Photon Counts:
12
1532
0
0
2400
3771
…
…
…

Month

Day

Time (UT)

Year-2000

Azimuth Angle

Bin Resolution

Off-Zenith Angle

Base Altitude (km)

Lat

Longi

Frequency number

Step 2. Nonlinearity of PMT + Discriminator

For small input photon flux, PMT output photon counts are proportional to the input photon counts:

$$\lambda_{oP} = \lambda_S = \lambda_i \eta_{QE}$$

When the input photon flux is considerably large, the output photon counts are no longer linear with input photons. Nonlinearity of PMT occurs:

$$\lambda_{oP} = \lambda_S e^{-\lambda_S \tau_p}$$

A discriminator is used to judge real photon signals and also has a saturation effect, i.e., its output photon counts are smaller than input photon counts when input count rate is large:

$$\lambda_o = \frac{\lambda_{iD}}{1 + \lambda_{iD} \tau_d}$$

Nonlinearity of PMT + Discriminator

Since PMT output is the input of discriminator

$$\lambda_{iD} = \lambda_{oP}$$

we obtain

$$\lambda_o = \frac{\lambda_S e^{-\lambda_S \tau_p}}{1 + \lambda_S \tau_d e^{-\lambda_S \tau_p}} = \frac{\lambda_S e^{-\lambda_S \tau_p}}{1 + \lambda_S \tau_d e^{-\lambda_S \tau_p}}$$

where

$\lambda_S = \lambda_i \eta_{QE}$ η_{QE} is the quantum efficiency of cathode

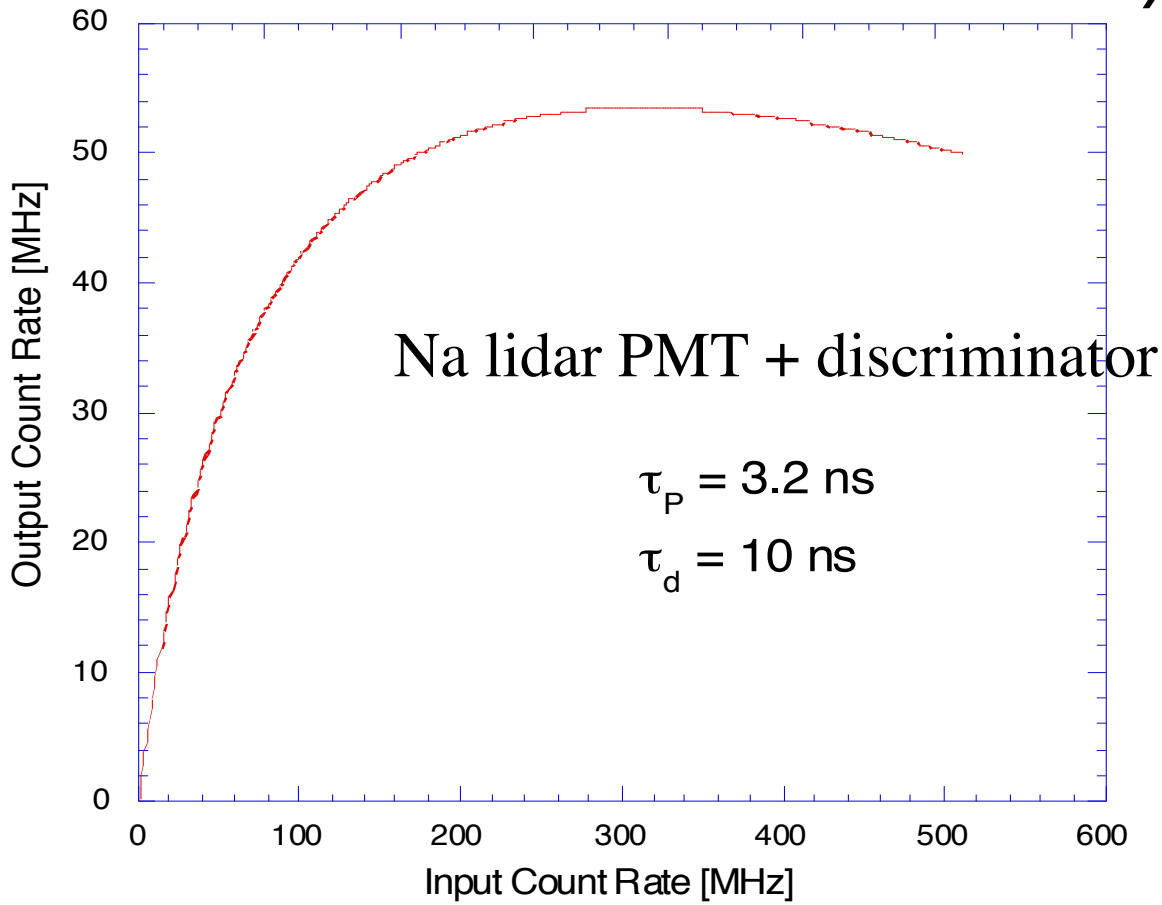
Maximum output count rate is reached when $\lambda_S = 1/\tau_p$

$$\lambda_{o\,\max} = \frac{1}{\tau_p e + \tau_d} \quad \Rightarrow \quad \tau_p = \frac{\frac{1}{\lambda_{o\max}} - \tau_d}{e}$$

PMT+Discriminator Saturation Correction

$$\lambda_o = \frac{\lambda_S e^{-\lambda_S \tau_p}}{1 + \lambda_S \tau_d e^{-\lambda_S \tau_p}}$$

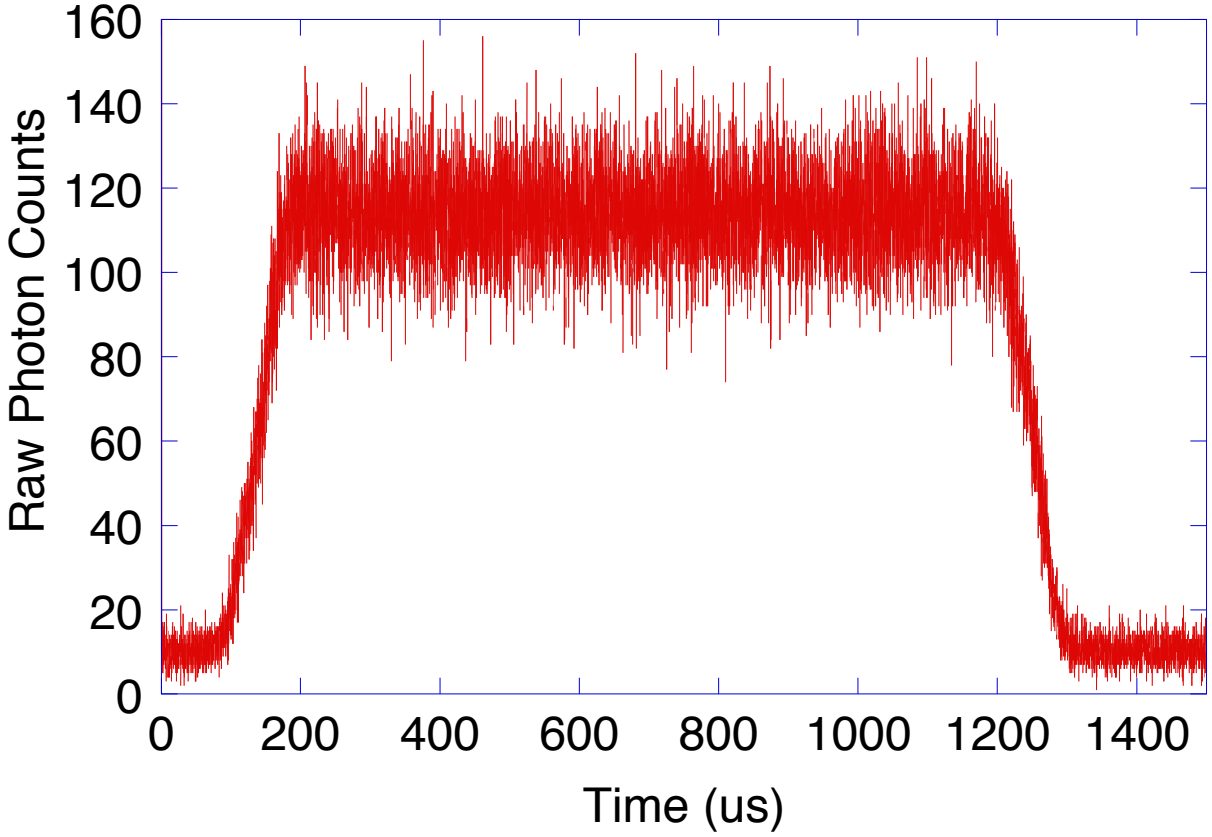
$$\lambda_S = \lambda_i \eta_{QE}$$



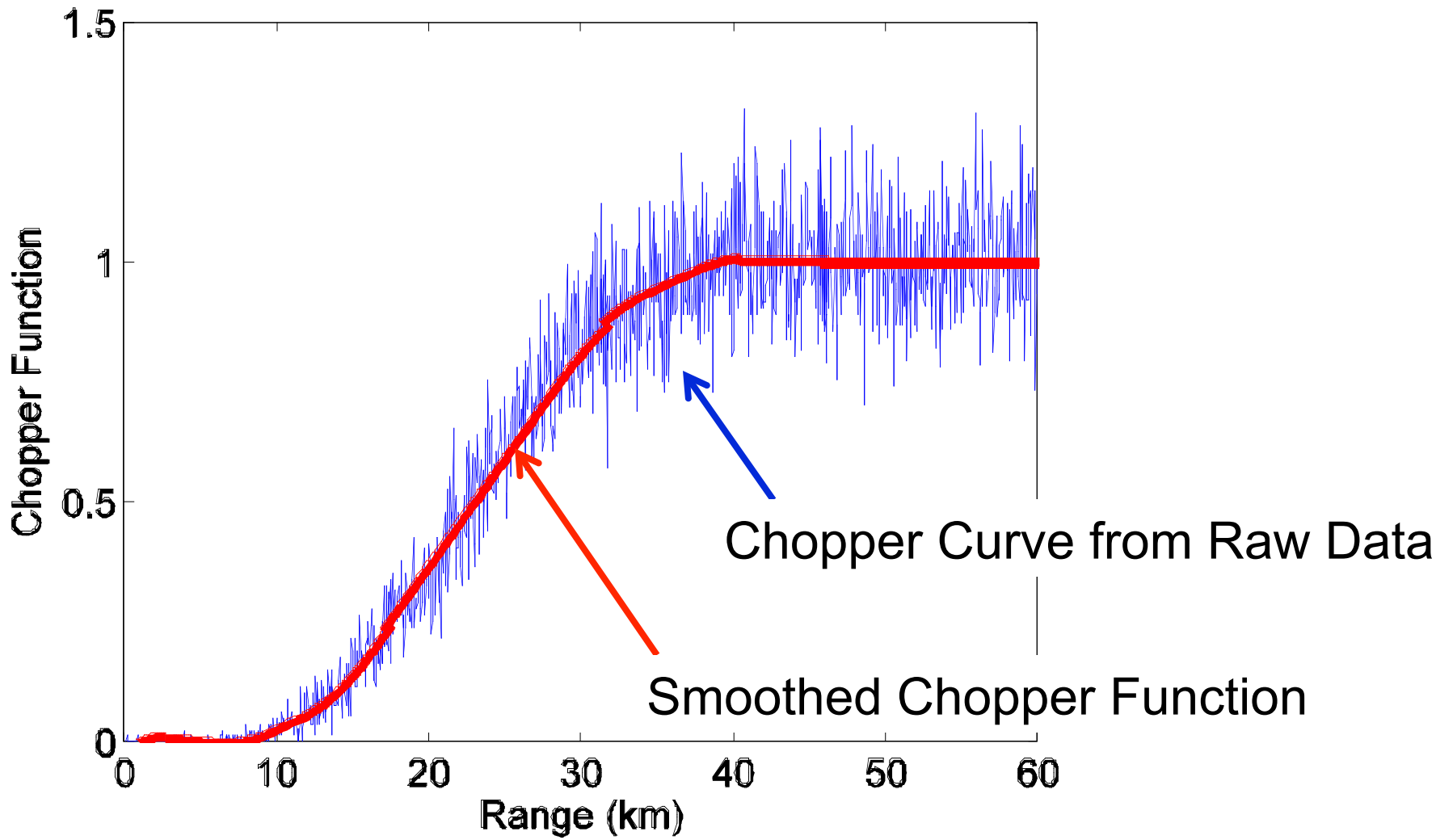
λ_S

Step 3. Chopper Correction

- Chopper function is measured and then used to do chopper correction for lower atmosphere signals



Chopper Correction



Step 4. Subtract Background N_B

$$N_S(\lambda,z)=\left(\frac{P_L(\lambda)\Delta t}{hc/\lambda}\right)\left[\sigma_{eff}(\lambda,z)n_c(z)R_B(\lambda)+\sigma_R(\pi,\lambda)n_R(z)\right]\Delta z\left(\frac{A}{4\pi z^2}\right)\times\left(T_a^2(\lambda)T_c^2(\lambda,z)\right)\left(\eta(\lambda)G(z)\right)+N_B$$

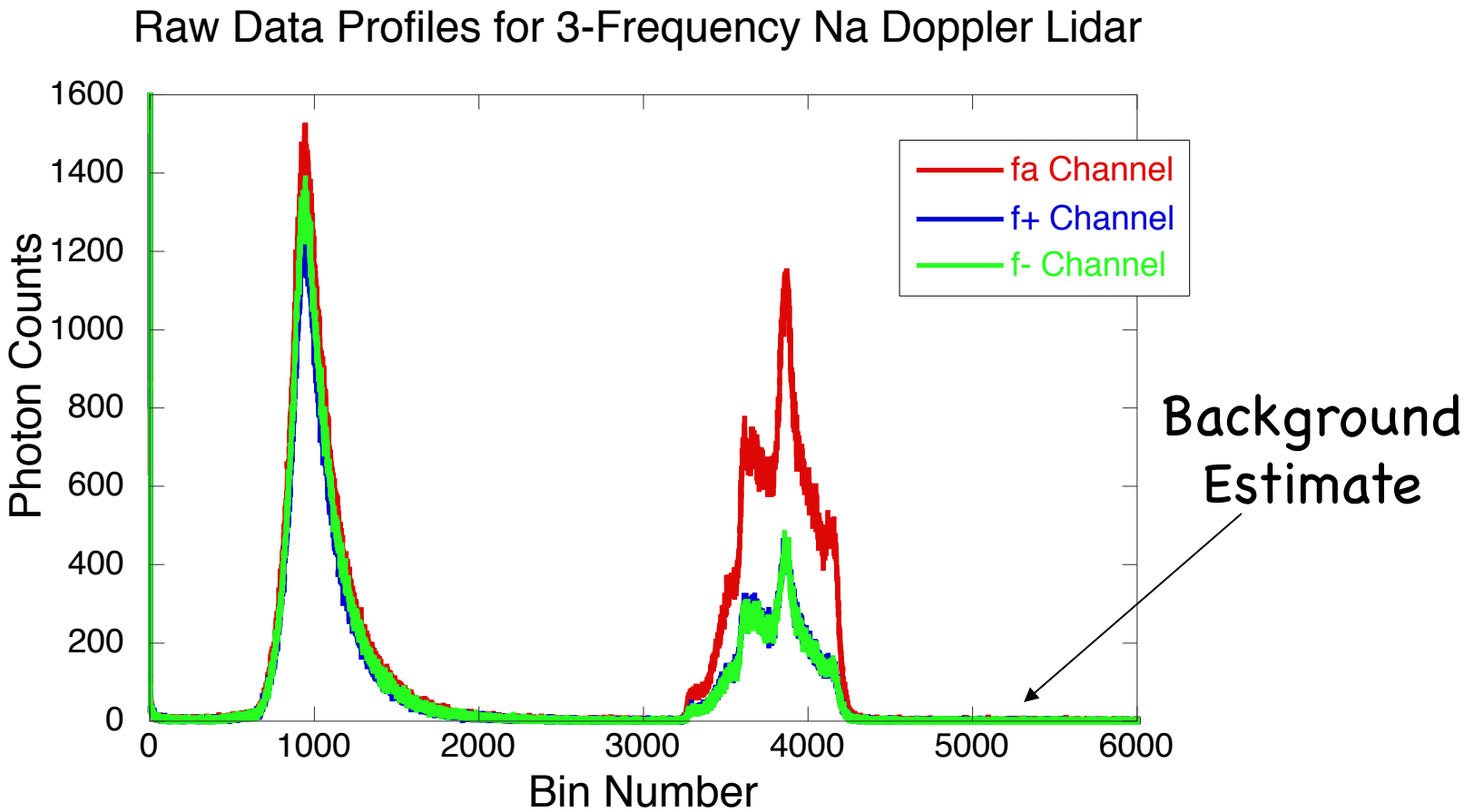
$$N_S(\lambda,z_R)=\left(\frac{P_L(\lambda)\Delta t}{hc/\lambda}\right)\left[\sigma_R(\pi,\lambda)n_R(z_R)\right]\Delta z\left(\frac{A}{z_R^2}\right)T_a^2(\lambda,z_R)\left(\eta(\lambda)G(z_R)\right)+N_B$$

$$\Downarrow$$

$$N_S(\lambda,z_R)-N_B=\left(\frac{P_L(\lambda)\Delta t}{hc/\lambda}\right)\left[\sigma_R(\pi,\lambda)n_R(z_R)\right]\Delta z\left(\frac{A}{z_R^2}\right)T_a^2(\lambda,z_R)\left(\eta(\lambda)G(z_R)\right)$$

Background Estimate

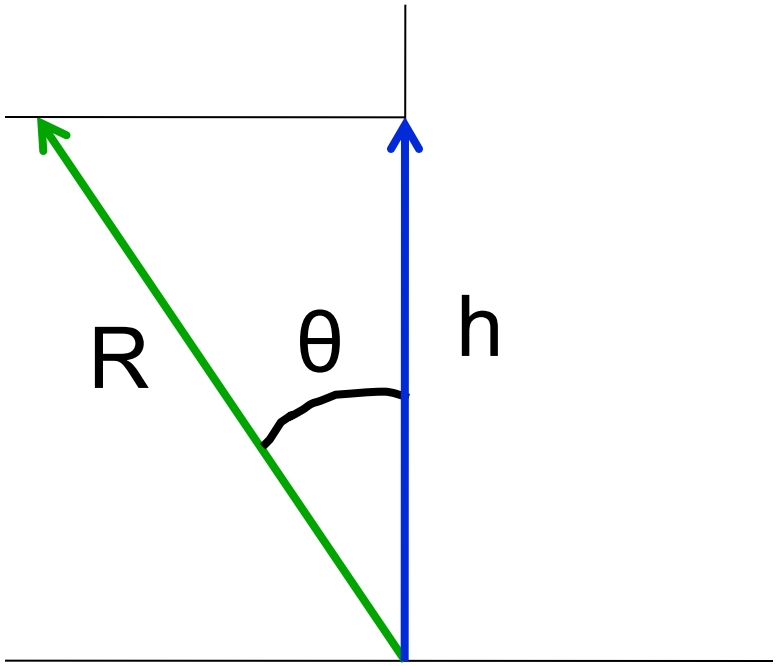
- Background is estimated from high altitude signal



There could be titled background due to PMT saturation

Step 5. Remove Range Dependence

$$\left[N_S(\lambda,z_R)-N_B\right]R^2=\left(\frac{P_L(\lambda)\Delta t}{hc/\lambda}\right)\left[\sigma_R(\pi,\lambda)n_R(z_R)\right]\Delta R(A)T_a^{\,2}(\lambda,z_R)\left(\eta(\lambda)G(z_R)\right)$$

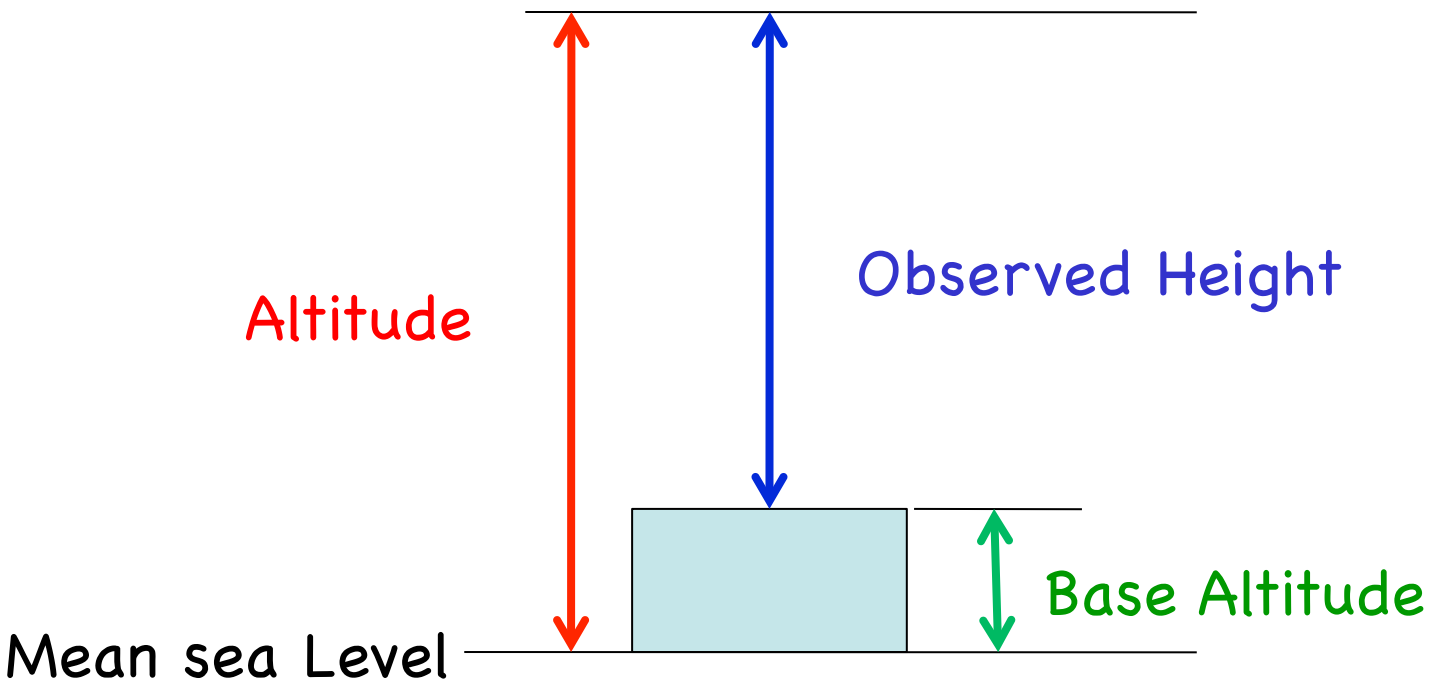


$$h=R\cos\theta$$

Step 6. Add Base Altitude

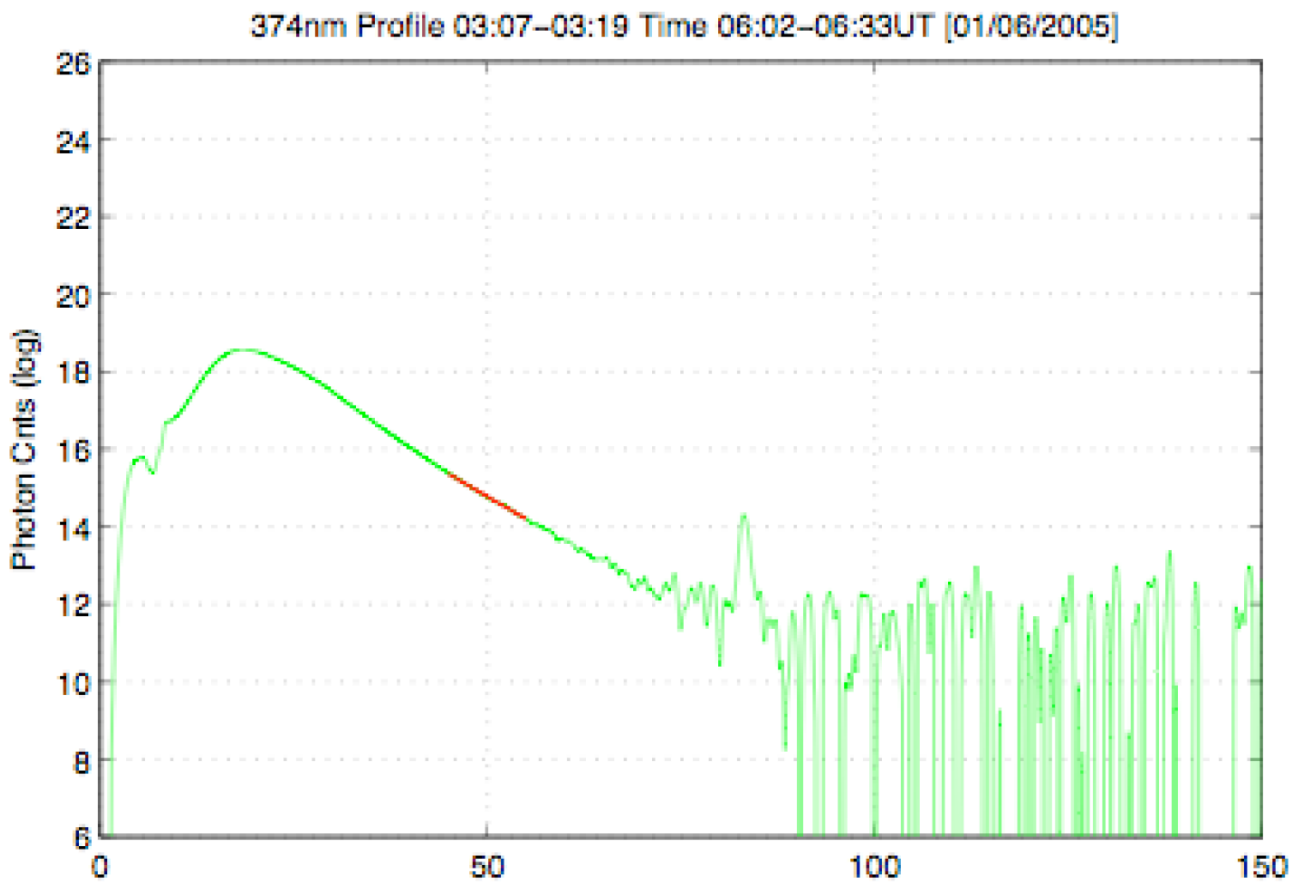
Altitude (relative to mean sea level)
= Observed Height + Base Altitude

$$\boxed{z = h + z_{Base} = R\cos\theta + z_{Base}}$$



Step 7. Rayleigh Normalization

- Estimate of Rayleigh Normalization Signal - Rayleigh Fit or Sum

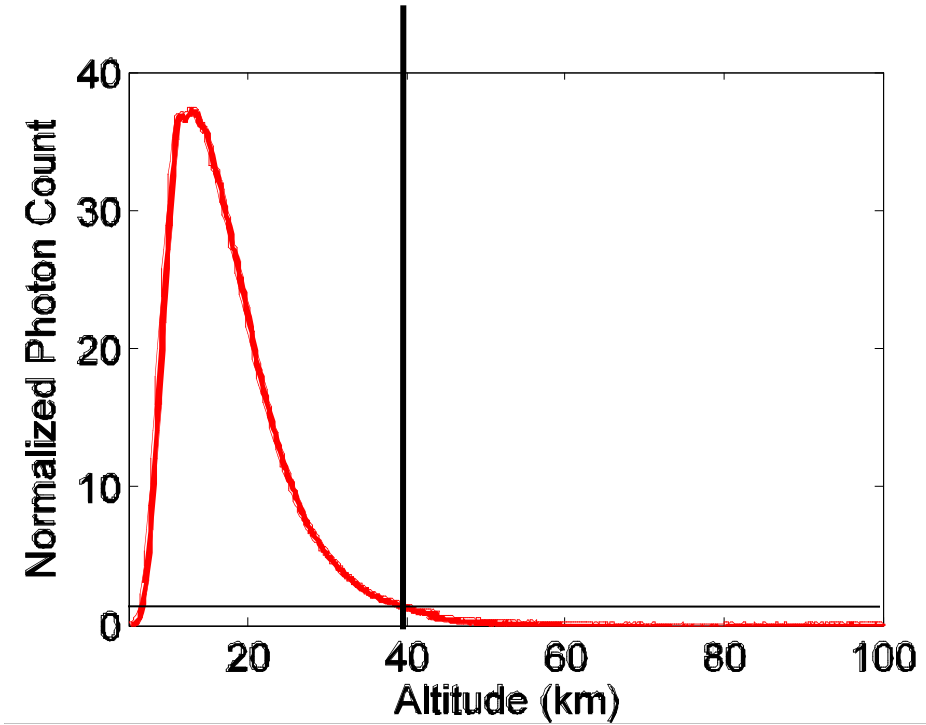


Rayleigh Normalization

$$N_{Normalization}(\lambda, z) = \frac{N_S(\lambda, z) - N_B}{N_S(\lambda, z_R) - N_B} \frac{z^2}{z_R^{\ 2}}$$

Photon Counts at Rayleigh Normalization Altitude (30-55 km)

Normalization Altitude (40 km)



Summary

- Lidar data inversion is to convert raw photon counts to meaningful physical parameters like temperature, wind, number density, and volume backscatter coefficient. It is a key step in the process of using lidar to study science.
- The basic procedure of data inversion originates from solutions of lidar equations, in combination with detailed considerations of hardware properties and limitations as well as detailed considerations of light propagation and interaction processes.
- Output of the preprocess and profile process is Normalized Photon Count, which is a preparation for the main process to derive temperature, wind, density, or backscatter coefficient, etc.