5

Resonance Fluorescence Lidar for Measurements of the Middle and Upper Atmosphere

XINZHAO CHU and GEORGE C. PAPEN

Department of Electrical & Computer Engineering,
University of Illinois at Urbana-Champaign, Urbana, IL, USA

5.1. Introduction 182
5.1.1. Lidar Study of the Middle and Upper Atmosphere 182
5.1.2. Lidar Concepts and Classifications 183
5.1.3. Initial Developments of Resonance Fluorescence Lidar 185
5.1.4. New Developments of Resonance Fluorescence Lidar 187
5.1.5. Arrangement of this Chapter 190
5.2. Advanced Technology of Resonance Fluorescence Lidar 190

5.2.1. Lidar Equations for Resonance Fluorescence Lidar 191
5.2.1.1. General Form of the Lidar Equation 191
5.2.1.2. Scattering Form of the Lidar Equation 194
5.2.1.3. Fluorescence Form of the Lidar Equation 196
5.2.1.4. Solutions for the Resonance Fluorescence Lidar Equation..... 206
5.2.2. Na Wind/Temperature Lidar..... 208
5.2.2.1. Introduction..... 208
5.2.2.2. Measurement Principle (Doppler Technique)..... 210
5.2.2.3. Na Doppler Lidar Instrumentation 228
5.2.2.4. Daytime Measurements..... 247
5.2.2.5. Lidar Data and Error Analysis..... 255
5.2.3. Fe Boltzmann Temperature Lidar 268
5.2.3.1. Introduction..... 268
5.2.3.2. Measurement Principle (Boltzmann Technique) 269
5.2.3.3. Fe Boltzmann Lidar Instrumentation 276
5.2.3.4. Temperature and Error Analysis .. 280
5.2.3.5. Rayleigh Temperature Retrieval ... 287
5.2.3.6. Temperature Measurement over the North and South Poles 290
5.2.3.7. Polar Mesospheric Cloud Detection over Both Poles..... 293
5.2.4. Potassium Doppler Lidar 296
5.2.4.1. Introduction..... 296
5.2.4.2. Measurement Principle..... 297
5.2.4.3. K Doppler Lidar Instrumentation . 301
5.2.4.4. Daytime Measurements..... 303
5.2.4.5. Temperature and Error Analysis .. 307
5.2.5. Solid-State Na Doppler Lidar 311
5.2.5.1. Introduction..... 311
5.2.5.2. All-Solid-State Na Temperature Lidar 311

5.2.5.3. Solid-State Na Wind/ Temperature Lidar 316
5.2.6. Comparison of Na, Fe, K, and Rayleigh Lidar Techniques 317
5.2.6.1. Narrowband Na Doppler Lidar Technique..... 318
5.2.6.2. Fe Boltzmann Temperature Lidar Technique 319
5.2.6.3. Narrowband K Doppler Lidar Technique..... 320
5.2.6.4. The Rayleigh Lidar Technique 321
5.3 Key Results of Lidar Measurements in the Middle and Upper Atmosphere 321
5.3.1 Thermal Structure of the Middle and Upper Atmosphere..... 322
5.3.1.1 Thermal Structure at Polar Latitudes..... 323
5.3.1.2 Thermal Structure at Mid-Latitudes 334
5.3.1.3 Thermal Structure at Low Latitudes..... 344
5.3.2 Dynamics of the Middle and Upper Atmosphere 350
5.3.2.1 Heat, Momentum, and Na Flux 350
5.3.2.2 Atmospheric Instability and Gravity Wave Directions..... 358
5.3.2.3 Tidal Study by Full-Diurnal-Cycle Lidar 361
5.3.3 Atmospheric Metallic Layers and Meteor Detection 364
5.3.3.1 Metal Layers and Mesospheric Chemistry 364
5.3.3.2 Meteor Trail Detection by Lidar..... 380
5.3.4 Polar Mesospheric Clouds (Noctilucent Clouds) 386
5.3.4.1 Historical Perspective 386
5.3.4.2 PMC Characteristics Measured by Lidar 389
5.4 Conclusions and Future Outlook 406

Acknowledgments..... 412
References..... 412

5.1. INTRODUCTION

5.1.1. Lidar Study of the Middle and Upper Atmosphere

The middle and upper atmosphere is a complex and important region. It contains a wealth of important geophysical phenomena, for example, the Earth's coldest environment — the mesopause; polar mesospheric clouds (PMCs) and noctilucent clouds (NLCs); the meteoric metal layers of Na, K, Ca, Li, and Fe; the airglow layers of OH, O, and O_2; and planetary, tidal, and gravity wave activities that play vital roles in overall global atmosphere circulation. The growing interest in the middle atmosphere is also tied to this region's importance for detecting and understanding global climate change. Global climate change is not confined to the lower atmosphere, but extends into the middle and upper atmosphere as well. While greenhouse gases such as CO_2 and CH_4 help warm the lower atmosphere by absorbing infrared radiation, they are also efficient radiators of heat and result in cooling in the middle and upper atmosphere (Roble and Dickinson, 1989). Doubling the CO_2 concentration is predicted to cool the stratopause ($\sim$50 km) by 10 to 12 K and the mesopause region ($\sim$80 to 100 km) by 6 to 12 K (Portman et al., 1995). The observed increases in the geographic extent and brightness of PMCs and NLCs in the Polar Regions during the past decade are also believed to be evidence of cooling temperatures and increasing water-vapor concentration in the mesopause region caused by increases of greenhouse gases (Gadsden, 1990; Thomas, 1996). Therefore, the middle and upper atmosphere thermal structure and polar mesospheric clouds are recognized to be important and sensitive indicators of global climate change. Modeling the middle atmosphere is hampered by the lack of observational data and the uncertainties in the parameterization of gravity waves. Observational studies of the middle and upper atmosphere provide crucial tests of

atmospheric general circulation models, which underlie our understanding of the global climate change, thereby leading to improvements in these models and our knowledge of the global atmosphere.

It is clear that data from the middle and upper atmosphere are vital for testing the atmospheric circulation models and studying complex atmospheric dynamics. However, it is challenging to collect data on temperature, wind, PMCs, and gravity waves from this region because the middle and upper atmosphere is usually inaccessible to balloon, aircraft, and most satellites. Optical and radio remote sensing instruments and *in situ* rocket measurements are the most common sources of observational data. Among these remote sensing technologies, lidar techniques are unique in their capabilities of providing high spatial and temporal resolution data on the temperature, wind, and constituent structure of the middle atmosphere region. Because of their daytime measurement capabilities and relatively low cost compared to rockets, lidars can have full seasonal and diurnal coverage of atmosphere measurements. These features make lidars important tools for the middle and upper atmosphere study.

5.1.2 Lidar Concepts and Classifications

Lidar stands for *light detection and ranging*. It is commonly referred to as *laser radar* because it is the optical equivalent of radar. Lidar studies began with searchlights for studying stratospheric aerosols and molecular density (Elterman, 1954). However, the use of a continuous wave (cw) light source prevented the measurement of the height variation of the density. Although it was possible to measure the height variation by separating the cw searchlight and receiving system in a bistatic configuration, modern lidars utilize pulsed laser sources in a monostatic configuration, which provide much better vertical resolution and much better beam collimation. Usually, a lidar consists of a transmitter, a receiver, and a data acquisition and control system. The lidar transmitter consists of a pulsed laser system equipped with spectral analyzing instruments and beam collimating optics to provide laser

pulses with precise frequency, narrow spectral width, high energy, and small divergence. The lidar receiver includes an optical telescope to collect backscattered photons, filters to eliminate background light, and a photon detector to count the return photons. The data acquisition and control system is used to record the photon counts and to provide the timing and system control for the whole system.

For an atmospheric object to be detected by lidar methods, the object must interact with the light in such a manner that a sufficient number of photons are scattered back to the telescope, which are above the system noise level. When laser photons propagate in the atmosphere, they are scattered by different constituents and different mechanisms. The backscattered photons are collected by a telescope, filtered by optical filters, and detected by photomultiplier tubes or photodiodes. The range from the lidar system to the scatterer can be determined precisely by measuring the flight time of photons. This range is given by $R = c\Delta t/2$, where c is the speed of the light and Δt is the flight time. The factor of 1/2 accounts for the two-way propagation. Therefore, the photon count profiles recorded by the lidar receiver are range resolved. By combining knowledge of atmospheric dynamics with principles of light scattering of particles, the range-resolved atmosphere parameters (such as temperature, wind, density, backscatter coefficient, etc.) can be derived.

According to the physical mechanism of the particle-radiation interaction in the atmosphere, lidars can be classified into Mie scattering lidar, Raman scattering lidar, Rayleigh scattering lidar, resonance fluorescence lidar, and differential absorption lidar (DIAL). Lidars using Mie scattering (elastic scattering from small particles and aerosols), Raman scattering (inelastic scattering from atoms and molecules), and differential absorption of atoms and molecules are typically used in the lower atmosphere below 30 km where the backscatter signals from aerosols or molecules and the absorption from molecules are strong enough to derive useful information of

atmosphere. Lidars using Rayleigh scattering (elastic scattering from atoms and molecules) are generally used in the altitude range of 30 to 90 km, where the atmosphere is aerosol-free but still contains enough gas molecules to produce a useful Rayleigh signal. Above approximately 85 km, because of the low atmospheric density, Rayleigh signal is typically too weak to derive any reliable information. Lidars using resonance fluorescence can be used in the range from 75 to 115 km where suitable constituents with large scattering cross-section exist to produce strong resonance fluorescence signals. The backscatter cross-section of resonance fluorescence is approximately 10^{14} times higher than Rayleigh backscatter cross-section. Although the density of tracer atoms or molecules is about 10^{10} times less than the atmosphere molecular density at the same altitude, the overall resonance fluorescence signal, which is proportional to the product of backscatter cross-section and tracer atom density, can be several orders of magnitude higher than Rayleigh backscatter signal for the same laser power and receiving telescope aperture. Thus, resonance fluorescence lidars can generate useful photon signals to derive atmospheric parameters in the altitude range of 75 to 115 km. This region is usually referred to as the MLT region (i.e., the mesosphere and lower thermosphere), and is also called the *mesopause region*.

5.1.3 Initial Developments of Resonance Fluorescence Lidar

The naturally existing metallic atoms in the mesopause region, such as Na, K, Li, Ca, and Fe, are believed to originate from the meteor ablation. Some electronic transitions of these atoms fall in the wavelength range of available laser sources. When a laser is tuned to a resonant line, an atom will absorb a laser photon and then re-emit through spontaneous emission. Since quenching (excited atoms returning to ground state without photon emission, usually caused by collisions with surrounding molecules) is not a problem in the middle and upper

atmosphere because of the low density, fluorescence from these atoms and molecules can be detected. This resonance fluorescence signal can be used to determine the density of the metal species as well as measure the wind and temperature.

Since the first lidar observation of mesospheric Na atoms by Bowman et al. (1969), the Na density and column abundance have been measured by several groups using broadband resonance fluorescence lidars where the spectral width of the laser pulse was much larger than the absorption width of the resonance lines (Sandford and Gibson, 1970; Gibson and Sandford 1971, 1972; Blamont et al., 1972; Hake et al., 1972; Kirchhoff and Clemesha, 1973; Megie and Blamont, 1977). The variations of Na layers at various temporal and spatial scales were observed, including seasonal and diurnal variations, sporadic Na layers, and the variations induced by dynamical processes such as tides and gravity waves. Other metal species were also measured by the broadband resonance fluorescence lidars, for example, K (Felix et al., 1973), Li (Jegou et al., 1980), Ca and Ca^+ (Granier et al., 1985, 1989a; Qian and Gardner, 1995), and Fe (Granier et al., 1989b; Alpers et al., 1990; Bills and Gardner, 1990; Kane and Gardner, 1993a). Simultaneous observations of sporadic E, Na, Fe, and Ca^+ layers were also made by Gardner et al. (1993). By reducing the receiver bandwidth, daytime observations of the atmospheric Na layers were first reported by Gibson and Sandford (1972), and later by three different groups (Clemesha et al., 1982; Granier and Megie, 1982; Kwon et al., 1987).

Studies using broadband resonance fluorescence lidars provided key insights into mesospheric constituents. They also enabled the study of gravity and tidal waves by studying the tracer density variations (Batista et al., 1985; Gardner and Voelz, 1985; Senft and Gardner, 1991). However, broadband lidars cannot directly measure the temperature and wind, which are the key parameters for characterizing the middle and upper atmospheric dynamics. Scientific curiosity called for further developments to make the resonance fluorescence lidars capable of temperature and wind measurements in the mesopause region.

5.1.4 New Developments of Resonance Fluorescence Lidar

The first observation of Na ground-state hyperfine structure in the middle and upper atmosphere was reported by Gibson et al. (1979), establishing the possibility of mesopause temperature measurements with a resonance fluorescence lidar using a narrowband tunable laser. The atmosphere temperature can be determined by measuring the Doppler broadening and the wind can be determined by measuring the Doppler shift of laser-induced fluorescence from atmospheric metal atoms. In this study, a pulsed laser with linewidth of 0.1 GHz was scanned through sodium D_2 line with Doppler broadening. Although the absolute calibration of the laser frequency was not available, a least-square fit using theoretical model of Na spectroscopy gave the best estimate of the atmospheric temperature with an error of ± 15 K (Gibson et al., 1979; Thomas and Bhattacharyya, 1980; Thomas, 1995). The first Na temperature lidar that made routine observations was developed at the University of Bonn with an excimer-pumped pulsed dye-laser system (Fricke and von Zahn, 1985). Fricke and von Zahn (1985) were able to measure the temperature profiles through the mesopause region by tuning and monitoring the frequency of each laser pulse with the help of a wavelength meter and a Na vapor cell. An uncertainty of ± 5 K at the Na layer peak was obtained with 1-km vertical resolution and 10-min integration period. Their measurements at Andoya, Norway with this Na lidar led to an understanding of the temperature structure of a polar mesopause (Lübken and von Zahn, 1991).

A more advanced narrowband Na lidar system using a two-frequency technique was developed through a collaboration between the Colorado State University and the University of Illinois (She et al., 1990; Bills et al., 1991a). The lidar transmitter consisted of a cw frequency-stabilized ring dye laser pumped by an argon ion laser. The cw 589-nm laser was then amplified by a pulsed dye amplifier, which itself was pumped by an injection-seeded frequency-doubled Nd:YAG laser. Such laser system provided excellent frequency

tuning accuracy, pulse-to-pulse stability, and high peak power. A temperature-sensitive ratio of the Na fluorescence signals at two frequencies was used to derive temperature from the Doppler-broadening linewidth. An accuracy of better than ± 3 K at the Na layer peak was achieved with 1-km vertical resolution and 5-min integration period. Since a cw ring dye laser was used, the Doppler-free saturation–absorption features of a Na vapor cell could be used to fine-tune and lock laser frequency onto sodium D_2 line, which provided an absolute calibration of laser frequency. This technique not only improved the accuracy of temperature measurement, but also enabled the wind measurement by determining the Doppler shift of sodium lines (Bills et al., 1991a,b; She et al., 1992; She and Yu, 1994; Gardner and Papen, 1995). Initial horizontal wind measurements were made by tuning the lidar to four frequencies with a stabilized Fabry–Perot etalon (Bills et al., 1991b) and to three frequencies with an acousto-optic modulator (She and Yu, 1994). In currently implemented systems, two- and three-frequency techniques are most widely employed since they can provide higher temporal resolution and the fast switching between the two or three frequencies eliminates errors caused by Na density fluctuations during data collection. After the initial development, temperature and wind measurements were collected on a routine basis as well as during campaigns. The lidar beam was also coupled with large telescope through coude optics and was able to provide flexible beam pointing. This, in combination with the large receiver, enabled the first lidar studies of heat flux and momentum flux (Tao and Gardner, 1995; Gardner and Yang, 1998).

New lidars are continuously being proposed and developed. After a decade of development, the first report on mesopause temperature measurements using a narrowband alexandrite laser probing atmospheric potassium atoms was published by von Zahn and Höffner (1996). A laser-diode-injection-seeded pulsed ring alexandrite laser was used in the transmitter to achieve a narrowband system. The principle of deriving temperatures from a potassium lidar signal is very similar to the Na temperature lidar except

the laser frequency was scanned through potassium D_1 line (instead of D_2) since the potassium D_2 line overlaps with a strong O_2 absorption line. Although potassium abundance is two orders of magnitude lower than sodium, this reduction can be offset by using an alexandrite laser with higher average power and much longer pulse length.

Both the Na and K temperature lidars utilize the Doppler broadening of metal atoms, so narrowband tunable lasers are required in the lidar transmitters to resolve the temperature-dependent Doppler-broadened atomic lineshape. A new concept for temperature measurement was proposed and developed over the past few years. This is the Boltzmann temperature technique. As proposed by Gelbwachs (1994), the temperature dependence of the population ratio on different energy levels of atomic iron (Fe) can be used to derive atmosphere temperatures. Motivated by this initial proposal and the need for polar atmosphere measurement, the University of Illinois at Urbana-Champaign (UIUC) developed a robust Fe Boltzmann temperature lidar using solid-state alexandrite laser technology (Chu et al., 2002a). The temperatures are determined by taking the ratio of the fluorescence photon signals from two independent Fe lines at 372 and 374 nm. This ratio is a very sensitive function of atmosphere temperature, and it obeys the Maxwell–Boltzmann distribution law under thermodynamic equilibrium. Since the ratio is from two transition lines instead of within one line, the lasers can be broadband with a comparable linewidth to Fe Doppler broadening ($\sim$1 GHz at 200 K). The lidar was then deployed on campaigns ranging from the geographic North Pole to the geographic South Pole to observe the polar atmosphere during both summer and winter. Significant scientific findings have been achieved through these measurements (Chu et al., 2001a,b, 2003a; Gardner et al., 2001; Pan et al., 2002).

Other developments in the resonance fluorescence lidars include creating an all-solid-state version of the Na lidar. Until now, the Na wind/temperature lidar is still the most accurate lidar in measuring temperatures. However, the ring dye laser used in the system is so sensitive that it requires a very stable operating environment. This precludes

measurements in remote areas such as the Antarctica and the Arctic or deployment on aircraft. Two types of solid-state Na lidars have been developed. One is to mix the frequencies of two pulsed Nd:YAG lasers operating at 1.06 and 1.3 μm to produce 589-nm laser pulses (Kawahara et al., 2002). Another is to mix the frequencies of two cw single-mode Nd:YAG lasers to produce a cw 589-nm laser (Vance et al., 1998). This solid-state cw 589-nm laser, which replaces the ring dye laser, is then followed by a pulsed dye amplifier to produce 589-nm laser pulses (She et al., 2002b). Both types of lidar have been used to measure the mesopause temperatures in the Polar Regions.

5.1.5 Arrangement of this Chapter

In Section 5.2, we will concentrate on the technologies of the resonance fluorescence lidar developed in the past 20 years. These technologies include the developments of the Na wind/temperature lidar, the Fe Boltzmann temperature lidar, the K Doppler temperature lidar, and the solid-state Na Doppler lidar. A comparison of these lidar technologies is given at the end of Section 5.2. Section 5.3 provides a summary of the key scientific results obtained using resonance fluorescence lidars. These results include the characterization of the MLT region thermal structure, MLT region dynamics, mesospheric metallic layers and meteor trails, and the polar mesospheric clouds (noctilucent clouds). The chapter concludes in Section 5.4 with an outlook on the future of resonance lidar techniques.

5.2. ADVANCED TECHNOLOGY OF RESONANCE FLUORESCENCE LIDAR

Since the development of broadband resonance fluorescence lidar in the late 1960s for atmosphere constituent density measurements, the most important developments in the lidar field are the new resonance fluorescence lidar techniques for the temperature and wind measurements in the middle and upper atmosphere. Wind measurements rely on the effect

of Doppler frequency shift while there are mainly two concepts on how to measure temperatures in the mesopause region. One is to utilize the temperature dependence of Doppler broadening of electronic-dipole transition lines of the tracer gases, such as Na and K. We refer to this as the *Doppler technique*. Another is to utilize the temperature dependence of the ratio of populations on different energy levels of the tracer gases, such as Fe. We refer to this as the *Boltzmann technique*. These concepts and measurement principles will be described in detail in the Na wind/temperature lidar and Fe Boltzmann temperature lidar sections. First, we will introduce the lidar equations for the resonance fluorescence lidars.

5.2.1. Lidar Equations for Resonance Fluorescence Lidar

5.2.1.1. General Form of the Lidar Equation

The lidar equation is the basic equation in the field of laser remote sensing, which relates the received photon count obtained from a scattering region to the emitted laser photon numbers, the concentration of the scatterer, the interaction between the laser radiation and the scatterer, and the lidar system efficiency. The lidar equation is developed under two assumptions: the scattering processes are independent, and only single scattering occurs. For atmospheric scattering, due to the low atmosphere density, air particles are separated adequately and undergo random motion so that the contributions to the total scattered energy by many particles have no phase relations. This allows the total intensity to be calculated by simply adding the intensity from each scattering particle. This process is known as independent scattering. Single scattering implies that a photon is scattered only once. If a photon is scattered two or more times, then multiple scattering occurs. This will be excluded in our considerations since the probability for such multiple scattering to occur is negligibly small in the middle and upper atmosphere.

In general, the lidar interaction between the light photons and the particles is a scattering process. The expected received photon count is equal to the product of the number of

transmitted photons, the probability that a transmitted photon is scattered, the probability that a scattered photon is received, and the system efficiency. The general lidar equation can be written as

$$N_S(\lambda, z) = \left(\frac{P_L(\lambda_L)\Delta t}{hc/\lambda_L}\right)\left(\beta(\lambda, \lambda_L, z)\Delta z\right)\left(\frac{A}{z^2}\right) \times \left(\eta(\lambda, \lambda_L)T(\lambda_L, z)T(\lambda, z)G(z)\right) + N_B\Delta t \quad (5.1)$$

where

$N_S(\lambda, z)$ = expected photon counts detected at wavelength λ in the range interval $(z - \Delta z/2, z + \Delta z/2)$;
λ = detected photon wavelength (m);
$P_L(\lambda_L)$ = laser output power at laser wavelength λ_L (W);
λ_L = laser radiation wavelength (m);
Δt = integration time (s);
h = Planck's constant (6.626×10^{-34} J/s);
c = speed of light (2.99792458×10^8 m/s);
$\beta(\lambda, \lambda_L, z)$ = volume backscatter coefficient at range z under laser radiation at wavelength λ_L for the scattered photons falling into the wavelength λ ($m^{-1}\,sr^{-1}$);
Δz = thickness of the range bin or range interval (m), which is related to the sampling time τ by $\Delta z = c\tau/2$;
A = receiving telescope aperture area (m^2);
z = range from the scatter to the lidar receiver;
$T(\lambda_L, z)$ = one-way transmittance of the atmosphere for wavelength λ_L from the lidar transmitter to the range z;
$T(\lambda, z)$ = one-way transmittance of the atmosphere for wavelength λ from the range z to the lidar receiver;
$\eta(\lambda, \lambda_L)$ = lidar optical efficiency for transmitted wavelength at λ_L and received wavelength at λ;
$G(z)$ = geometrical probability of radiation at range z reaching the detector, based on the geometrical considerations; and
N_B = expected photon counts per range bin per unit time due to background noise and detector dark counts.

The first parenthesis term in Equation (5.1) is the number of transmitted laser photons, in which the factor $P_L \Delta t$ is the total laser energy and hc/λ_L is the energy for a single photon.

The second parenthesis term is the probability that a transmitted photon is backscattered by the scatters into a unit solid angle. This angular scattering probability is equal to the product of the volume backscatter coefficient $\beta(\theta = \pi)$ and the scattering layer thickness (Δz). The volume backscatter coefficient is the volume angular scattering coefficient $\beta(\theta)$ at the scattering angle $\theta = \pi$. The volume angular scattering coefficient $\beta(\theta)$ (m^{-1} sr^{-1}) is defined as the probability per unit distance traveled that a photon is scattered into wavelength λ in unit solid angle at the scattering angle θ. The volume total scattering coefficient β_T (m^{-1}) is the probability per unit distance traveled that a photon is scattered into wavelength λ in all directions. This can be obtained by integrating $\beta(\theta)$ over the 4π steradians. The volume backscatter coefficient can be expressed as

$$\beta(\lambda, \lambda_L, z) = \sum_i \left[\frac{d\sigma_i(\lambda_L)}{d\Omega} n_i(z) p_i(\lambda) \right] \tag{5.2}$$

where

$\frac{d\sigma_i(\lambda_L)}{d\Omega}$ = the differential backscatter cross-section of an individual particle, that is, the angular scattering cross-section at the scattering angle $\theta = \pi$ under laser radiation at wavelength λ_L;
$n_i(z)$ = the number density of scatter species i; and
$p_i(\lambda)$ = the probability of the scattered photons falling into the wavelength λ.

The differential scattering cross-section $\frac{d\sigma_i(\lambda_L)}{d\Omega}$ is defined as the ratio of the rate at which energy is scattered into a unit solid angle at an observation angle θ by an individual particle to the rate at which energy in the incident photon beam crosses a unit area perpendicular to its propagation direction. It has the dimension of the area and can be understood as the area of impinging flux that is scattered into a unit solid angle

at an observation angle θ. It is also called angular scattering cross-section $\sigma(\theta)$. The total scattering cross-section is another quantity used in scattering theory and is defined as the total area of flux scattered in all directions by a scatterer. This can be found by integrating $\sigma(\theta)$ over 4π steradians.

The third parenthesis term in Equation (5.1) is the probability that a scattered photon is collected by the receiving telescope. This receiving probability is equal to the solid angle subtended by the receiver aperture to the scatterer divided by 4π steradians. Since the factor 4π has been incorporated into the angular scattering probability, the receiving probability is just equal to the solid angle A/z^2.

The fourth parenthesis term is the overall system efficiency. The lidar optical efficiency η includes the optical efficiencies of all the optics in the lidar transmitter and receiver (e.g., the beam splitters, mirrors, lenses, optical filters, etc.) as well as the quantum efficiency of the photon detector (such as a photomultiplier tube [PMT] or a photodiode). It can be expressed as the product of the transmitter efficiency $\eta_T(\lambda_L)$ and the receiver efficiency $\eta_R(\lambda)$:

$$\eta(\lambda, \lambda_L) = \eta_T(\lambda_L)\eta_R(\lambda) \tag{5.3}$$

This optical efficiency is usually dominated by the optical filter transmittance and the detector quantum efficiency. The atmospheric transmittance $T(\lambda_L, z)$ and $T(\lambda, z)$ depend on the wavelength and the traveling range of the photons. The geometrical form factor $G(z)$ mainly depends upon the overlap of the area of laser irradiation with the field of view of the receiver optics, but also depends on the details of the receiver optics.

5.2.1.2. Scattering Form of the Lidar Equation

There are a few different forms for the lidar equation, which are determined by the nature of the interaction between the laser radiation and the scatterer. Scattering form (for Rayleigh, Mie, and Raman scatterings) and fluorescence form (for resonance fluorescence) are two representative forms of the lidar equation. Both forms will be used in the data analysis for

the resonance fluorescence lidars. The main differences between these forms lie in the volume backscatter coefficient and the overall system efficiency in the lidar equation.

The Rayleigh, Mie, and Raman scattering processes are instantaneous scattering processes; that is, there are no finite relaxation effects involved. For Rayleigh and Mie scattering, there is no frequency shift when the atmosphere particles are at rest, so $\lambda = \lambda_L$, $p(\lambda) = 1$, and the volume backscatter coefficient can be written as

$$\beta(\lambda, z) = \sum_i \left\lfloor \frac{d\sigma_i(\lambda)}{d\Omega} n_i(z) \right\rfloor \tag{5.4}$$

Therefore, the scattering form of the lidar equation for Rayleigh and Mie scatterings is

$$N_s(\lambda, z) = \left(\frac{P_L(\lambda)\Delta t}{hc/\lambda}\right)(\beta(\lambda, z)\Delta z)\left(\frac{A}{z^2}\right) \times \left(\eta(\lambda) T^2(\lambda, z) G(z)\right) + N_B \Delta t \tag{5.5}$$

For Mie scattering, the summation in Equation (5.4) is mainly for different particle sizes and species, which vary in a large range. For Rayleigh scattering, the differential backscatter cross-section can be written as a common Rayleigh backscatter cross-section $\frac{d\sigma_i(\lambda)}{d\Omega} = \sigma_R(\lambda)$ and the air total number density can be written as $n_R(z) = \sum_i n_i(z)$ for the well-mixed atmosphere below 100 km. Thus, the Rayleigh volume backscatter coefficient can be simplified as

$$\beta(\lambda, z) = \sigma_R(\pi, \lambda) n_R(z) \tag{5.6}$$

The Rayleigh scattering is anisotropic. The relation between the volume angular scattering coefficient and the volume total scattering coefficient is determined by a phase function, which is defined as the ratio of energy scattered into a direction per unit solid angle to the average energy scattered in all directions per unit solid angle. The phase function for Rayleigh scattering is given by (Goody and Yung, 1989)

$$P(\theta) = 0.7629 \times (1 + 0.9324 \cos^2 \theta) \tag{5.7}$$

where θ is the scattering angle, and backscattering corresponds to $\theta = \pi$. The volume angular scattering coefficient can be derived from the volume total scattering coefficient β_T as

$$\beta(\theta) = \frac{\beta_T}{4\pi} P(\theta) = \frac{\beta_T}{4\pi} \times 0.7629 \times (1 + 0.9324 \cos^2 \theta) \tag{5.8}$$

The Rayleigh backscatter cross-section can be computed from quantum mechanics (Loudon, 1983; Measures, 1984; Goody and Yung, 1989). The volume total scattering coefficient can also be expressed through the atmosphere temperature and pressure (Cerny and Sechrist, 1980)

$$\beta_T(z) = 9.807 \times 10^{-23} \left(\frac{273}{T(z)}\right) \left(\frac{P(z)}{1013}\right) \frac{1}{\lambda^{4.0117}} \tag{5.9}$$

where T is the temperature in Kelvin, P is the pressure in millibars, and λ is the wavelength in centimeters. Substituting Equation (5.9) into (5.8) and letting $\theta = \pi$, we get the Rayleigh volume backscatter coefficient as (the wavelength has been changed to the unit of meters)

$$\begin{aligned} \beta(\pi, \lambda, z) &= \sigma_R(\pi, \lambda) n_R(z) \\ &= 2.938 \times 10^{-32} \left(\frac{P(z)}{T(z)}\right) \frac{1}{\lambda^{4.0117}} \end{aligned} \tag{5.10}$$

Therefore, the Rayleigh scattering lidar equation is given by

$$\begin{aligned} N_S(\lambda, z) = &\left(\frac{P_L(\lambda)\Delta t}{hc/\lambda}\right) (\sigma_R(\pi, \lambda) n_R(z) \Delta z) \left(\frac{A}{z^2}\right) \\ &\times \left(\eta(\lambda) T^2(\lambda, z) G(z)\right) + N_B \Delta t \end{aligned} \tag{5.11}$$

where the product of the backscatter cross-section and atmosphere number density is given by Equation (5.10).

5.2.1.3. Fluorescence Form of the Lidar Equation

The resonance fluorescence process involves absorption of a photon and spontaneous emission of a photon by each scattering atom. If the atom is collisionally deactivated (quenched), it will not contribute to the fluorescence intensity. Also, if the decay rate of the atom is too slow (i.e., the lifetime of the

atomic excited state is too long) relative to the laser pulse duration, the saturation effect will be involved. (Details about the saturation effect will be discussed later.) Compared with the instantaneous scattering processes, such as Rayleigh and Mie scatterings that have an infinitely short duration, a few extra factors must be taken into account for the resonance fluorescence in the middle and upper atmosphere: (1) the laser spectral shape and linewidth; (2) the laser and signal polarization; (3) the laser and signal extinction due to the absorption of the resonance atom layers; (4) the finite lifetime of the excited state of the atoms; and (5) the laser pulse duration time and temporal shape. In the following treatment, the first two factors will be considered in the effective backscatter cross-section; the third factor will be considered as the extinction coefficient, which is part of the atmospheric transmission; and the last two factors will be considered as the saturation effect of the atomic layers in the middle and upper atmosphere. Since the collision rate is very low in the mesopause region due to the low atmosphere density, quenching is not a problem and will not be considered here for the measurements in the mesopause region.

5.2.1.3.1. Considerations for the effective backscatter cross-section

In the field of resonance fluorescence lidar, it is common to use the total effective scattering cross-section σ_{eff} instead of the differential backscatter cross-section. The total effective scattering cross-section σ_{eff} is defined as the ratio of the average photon number scattered by an individual atom (in all directions) to the total incident photon number per unit area. Since the resonance scattering is isotropic, the differential backscatter cross-section in Equation (5.2) can be replaced by the total effective scattering cross-section σ_{eff} divided by 4π. The σ_{eff} is determined by the convolution of the atomic absorption cross-section σ_{abs} and the laser spectral lineshape $g_{\text{L}}(\nu)$. The absorption cross-section σ_{abs} is the ratio of the average absorbed single-frequency photons per atom to the total incident photons per unit area, and is proportional

to the probability of a single-frequency photon being absorbed by an atom:

$$\sigma_{\text{abs}}(\nu, \nu_0) = A_{\text{ki}} \frac{\lambda^2}{8\pi n^2} \frac{g_{\text{k}}}{g_{\text{i}}} g_{\text{A}}(\nu, \nu_0) \tag{5.12}$$

where i is the lower energy level; k is the higher energy level; A_{ki} is the transition probability for the transition from level k to level i, that is, the Einstein A coefficient (s^{-1}); λ is the laser wavelength; n is the refraction index; g_{i} and g_{k} are the degeneracy factors for the lower and upper energy levels, respectively; and g_{A} is the atomic transition lineshape. The absorption cross-section σ_{abs} is strongly dependent on the laser frequency ν and the atomic resonance frequency ν_0; that is, σ_{abs} depends on the lineshape g_{A}. There is a certain bandwidth around the resonance frequency. When the laser frequency falls in this bandwidth, resonance occurs, and the absorption cross-section becomes much larger. This bandwidth is mainly determined by the natural linewidth (due to the radiative lifetime of the atomic excited-state), the collisional broadening, and the Doppler broadening. For the atoms found in the mesopause region that we are interested in (such as Na, K, and Fe), their excited-state lifetime τ is longer than 16 ns; therefore, their natural linewidth is less than 10 MHz ($\sim\frac{1}{2\pi\tau}$). The natural linewidth and collisional broadening are both homogenous broadening and can be described by a Lorentzian lineshape

$$g_{\text{H}}(\nu, \nu_0) = \frac{\Delta\nu_{\text{H}}}{2\pi\left[(\nu - \nu_0)^2 + (\Delta\nu_{\text{H}}/2)^2\right]} \tag{5.13}$$

where $\Delta\,\nu_{\text{H}}$ is the homogenous broadened linewidth. Due to the low atmosphere density in the mesopause region, the broadening caused by the collisions between the atoms and the air molecules is negligible, so $\Delta\nu_{\text{H}}$ is mainly the natural linewidth. However, the atom thermal velocity v will cause Doppler frequency shift, and the velocity distribution of the atoms obeys the Maxwell's distribution law: $\text{d}N(v) = Ng_{\text{D}}(v)\text{d}v$, where

$$g_{\text{D}}(v) = \sqrt{\frac{M}{2\pi k_{\text{B}} T}} \exp\left(-\frac{Mv^2}{2k_{\text{B}}T}\right) \tag{5.14}$$

Where M is the mass of the atom, K_B is the Boltzmann Constant, and T is the temperature.

Equation (5.14) is for the ideal gas in the case when the average velocity is zero. Since the observed photon absorption or scattering is the overall statistical results over all the atoms with different velocities, the overall absorption cross-section would be the single atom natural lineshape integrating through the velocity distribution, that is, a statistical average absorption cross-section. This statistical average will cause the Doppler broadening of the transition line. The velocity distribution, essentially the Doppler broadening, is temperature dependent. For the typical temperature of 200 K in the mesopause region, the Doppler broadening linewidth is around 1 GHz. It is much wider than the natural linewidth so that the atomic absorption line is essentially Doppler broadened. The statistically averaged absorption cross-section for each atomic transition line is then given by

$$\sigma_{\text{abs}}(\nu, \nu_0) = \sigma_0 \exp\left(-\frac{(\nu - \nu_0)^2}{2\sigma_{\text{D}}^2}\right) \tag{5.15}$$

where

ν_0 = the resonance frequency of each atomic transition line;
$\sigma_{\text{D}} = \nu_0\sqrt{\frac{k_{\text{B}}T}{Mc^2}}$ is the rms linewidth of Doppler broadening;
$\sigma_0 = \frac{1}{\sqrt{2\pi}\sigma_{\text{D}}}\frac{e^2}{4\epsilon_0 m_{\text{e}} c} f_{\text{ik}}$ is the peak absorption cross-section at resonance;
$f_{\text{ik}} = \frac{\epsilon_0 m_{\text{e}} c\lambda^2}{2\pi e^2}\frac{g_{\text{k}}}{g_{\text{i}}} A_{\text{ki}}$ is the absorption oscillator strength; and
$A_{\text{ki}} = \frac{\text{BF}_{\text{ki}}}{\tau_{\text{k}}}$ is the transition probability from level k to level i, where τ_{k} is the radiative lifetime of the upper level k, and BF_{ki} is the emission branching fraction for this transition.

When the laser has finite spectral linewidth, the total effective scattering cross-section is equal to the convolution of the atomic absorption cross-section and the laser lineshape. The laser lineshape, that is, its frequency distribution function $g_{\text{L}}(\nu)$ is normalized so that

$$\int_0^\infty g_{\text{L}}(\nu, \nu_{\text{L}})\text{d}\nu = 1 \tag{5.16}$$

where ν_L is the central frequency of the laser radiation. It is usually referred as the laser frequency. The effective scattering cross-section is then given by

$$\sigma_{\text{eff}}(\nu_L, \nu_0) = \int_{-\infty}^{+\infty} \sigma_{\text{abs}}(\nu, \nu_0) g_L(\nu, \nu_L) d\nu \tag{5.17}$$

We approximate the laser spectral lineshape by a Gaussian lineshape with a rms width of σ_L:

$$g_L(\nu, \nu_L) = \frac{1}{\sqrt{2\pi}\sigma_L} \exp\left(-\frac{(\nu - \nu_L)^2}{2\sigma_L^2}\right) \tag{5.18}$$

Then the effective cross-section can be written as

$$\sigma_{\text{eff}}(\nu_L, \nu_0) = \frac{\sigma_D \sigma_0}{\sqrt{\sigma_D^2 + \sigma_L^2}} \exp\left(-\frac{(\nu_0 - \nu_L)^2}{2(\sigma_D^2 + \sigma_L^2)}\right) \tag{5.19}$$

The above derivation is intended for the resonance line, which displays no hyperfine structure. In general, if we have a resonance line with m hyperfine structures, we can express a similar relationship by summing up the weighted effect of each hyperfine line as below:

$$\sigma_{\text{eff}}(\nu_L) = \sum_{i=1}^{m} w_i \frac{\sigma_{Di} \sigma_0}{\sqrt{\sigma_{Di}^2 + \sigma_L^2}} \exp\left(-\frac{(\nu_{0i} - \nu_L)^2}{2(\sigma_{Di}^2 + \sigma_L^2)}\right) \tag{5.20}$$

where $w_i = I_i / \sum_{j=1}^{m} I_j$ is the weighting factor for the ith line, and I_i is the relative intensity of the ith line.

The polarization of resonance radiation is a direct consequence of the hyperfine structure of the transitions and of the Zeeman effect when a magnetic field is present. The Earth's magnetic field interacts with occupied atomic states to change their states of polarization. This results in a change in the emitted-radiation pattern from its zero-magnetic-field value. Thus, the relative intensities of the hyperfine components can be modified due to the Hanle effect, and they are dependent on the laser polarization. In the case of hyperfine structures for resonance fluorescence in the geomagnetic field, we count the

polarization effect into the relative intensity I_i of the hyperfine components, which would affect the weighting factors w_i in Equation (5.20). Therefore, the total effective scattering cross-section can still be represented correctly by Equation (5.20) as long as proper weighting factors are used.

For the fluorescence form of the lidar equation, the volume backscatter coefficient given by Equation (5.2) will be modified as follows. The differential backscatter cross section in Equation (5.2) will be replaced by the total effective scattering cross-section given by Equation (5.19) or (5.20) over 4π for the resonance line without and with hyperfine structures, respectively. The $n_i(z)$ in Equation (5.2) will be replaced by the number density of the resonance constituent atoms $n_c(z)$. The $p_i(\lambda)$ in Equation (5.2) will be replaced by the branching ratio R_B, which is defined as the percentage of the fluorescence photons that fall in the absorption resonance wavelength λ, occcuping the entire fluorescence photons emitted by excited-atoms, when the excited atoms may return to several sub-ground-states or go to other excited states to emit photons at different wavelengths rather than the absorption resonance wavelength. Therefore, when the saturation effect of the atomic layers is not considered, the volume backscatter coefficient for the resonance fluorescence can be written as

$$\beta(\lambda, \lambda_L, z) = \frac{\sigma_{\text{eff}}(\lambda_L)}{4\pi} n_c(z) R_B(\lambda) \tag{5.21}$$

5.2.1.3.2 Extinction coefficient

The extinction due to the absorption of the atomic layers can be regarded as part of the atmospheric transmission T in the lidar equation (5.1). In the scattering form of the lidar equation, this transmission T only includes the loss due to the absorption and scattering by the lower atmosphere below the atomic layer in the mesopause region. In the resonance fluorescence case, the total transmission T can be written as the product of the lower atmosphere transmission T_a and the extinction coefficient E due to the absorption of the atomic layers:

$$T(\lambda, z) = T_a(\lambda) E(\lambda, z) \tag{5.22}$$

where the extinction coefficient is defined as the ratio of the transmitted laser power to the incident laser power at the atomic layers. It can be calculated from the effective cross-section and the atomic density:

$$E(z) = \exp\left(-\int_{z_{\text{bottom}}}^{z} \sigma_{\text{eff}}(\lambda, z) n_{\text{c}}(z) \mathrm{d}z\right) \tag{5.23}$$

where z_{bottom} is the bottom altitude of the atomic layers.

Therefore, when the saturation effect of the atomic layers is not considered, the resonance fluorescence lidar equation is given by

$$\begin{aligned} N_{\text{S}}(\lambda, z) = &\left(\frac{P_{\text{L}}(\lambda)\Delta t}{hc/\lambda}\right)(\sigma_{\text{eff}}(\lambda) n_{\text{c}}(z) R_{\text{B}}(\lambda)\Delta z)\left(\frac{A}{4\pi z^2}\right) \\ &\times \left(\eta(\lambda) T_{\text{a}}^2(\lambda) E^2(\lambda, z) G(z)\right) + N_{\text{B}}\Delta t \end{aligned} \tag{5.24}$$

Here, the first parenthesis factor is the transmitted laser photon numbers during integration time; the second factor is the probability that a photon is scattered into the wavelength λ in all directions; the third factor is the probability that a scattered photon can be received by the receiving telescope, that is, the telescope area over the spherical surface area; the fourth factor is the overall system efficiency.

5.2.1.3.3 Saturation and optical pumping effects

For the case that laser pulse duration is in the nanosecond range, Equation (5.24) does not always hold the truth as the radiative lifetime of the atomic excited state is generally on the order of tens of nanoseconds. The finite lifetime of the atomic excited state will cause a saturation effect of the atomic layers. According to Equation (5.24), the number of backscattered photons should be proportional to the flux of the transmitted laser photons. However, for large laser intensities, the returned fluorescence photons are no longer proportional to the laser power, but less than what is expected from Equation (5.24). Saturation is defined as the deviation from this proportionality. The deviation can be caused by

the following reasons. (1) Due to the finite lifetime, atoms at the excited states must relax before they can absorb another photon. In other words, during the lifetime of their excited state, excited atoms cannot absorb another photon. (2) Stimulated emission occurs with the photon emitted in the forward direction; that is, this photon is unobservable. Reasons (1) and (2) contribute about the same order of magnitude to these deviations (von der Gathen, 1991). (3) Optical pumping effect: when a narrowband laser is used to excite the populations from a certain ground state to excited states, the excited atoms can relax back to other ground states. This will cause a partial depopulation of the particular ground state in favor of the other substates. The optical pumping effect contributes only slightly to the deviations for short laser pulses, but becomes important only if the laser pulses are long compared to the radiative lifetime of the excited states.

Here we follow the analysis made by Megie et al. (1978) on the saturation effect. During the laser pulse duration Δt_L, the excited-state population (i.e., the total number density over all velocity distribution) n_e is not in the steady state, but is determined by the rate equation as below:

$$\frac{dn_e}{dt} = -\frac{n_e}{\tau_R} + (n - n_e)\frac{\sigma_{eff} N_L(t) T}{z^2 \Omega} - n_e \frac{\sigma_{eff} N_L(t) T}{z^2 \Omega} \tag{5.25}$$

where τ_R is the radiative lifetime of the excited-state, n is the total number density of the atoms in the layer, i.e., the sum of the ground-state population ($n - n_e$) and the excited-state population n_e, $N_L(t)$ is the total number of laser emitted photons per unit time, T is the one-way atmosphere transmission, and Ω is the solid angle illuminated by the laser beam. The Ω is related to the full divergence θ_L by

$$\Omega = \frac{\pi}{4}\theta_L^2 \tag{5.26}$$

In order to derive an analytical solution of Equation (5.25), we assume that the temporal shape of the laser pulse can be represented by a rectangular shape with a width of Δt_L:

$$N_L(t) = \begin{cases} \frac{N_L}{\Delta t_L}, & 0 < t < \Delta t_L \\ 0, & t > \Delta t_L \end{cases} \tag{5.27}$$

where N_L is the total number of emitted photons related to each pulse energy. By solving Equation (5.25), the excited-state population is given by

$$n_e(t) = \begin{cases} \dfrac{n\tau'}{2t_S}(1 - e^{-t/\tau'}) & t < \Delta t_L \\ \dfrac{n\tau'}{2t_S}(1 - e^{-\Delta t_L/\tau'})e^{-(t-\Delta t_L)\tau_R} & t > \Delta t_L \end{cases} \tag{5.28}$$

where t_S and τ' are defined as

$$t_S = \frac{z^2 \Omega \Delta t_L}{2\sigma_{eff} N_L T} \tag{5.29}$$

$$\frac{1}{\tau'} = \frac{1}{\tau_R} + \frac{1}{t_S} \tag{5.30}$$

Since the decay rate of the excited-state population through spontaneous emission is $n_e(t)/\tau_R$, the number of reemitted photons per unit time in the volume considered will be $(z^2\Omega\Delta z)[n_e(t)/\tau_R]$. Its integration over the time is the number of reemitted photons, and can be used to replace the product of the first and the second factors in Equation (5.24). Therefore, the number of phonons received by the lidar during the integration time will be

$$\begin{aligned} N_S^{Sat}(\lambda, z) = &\left(\frac{A}{4\pi z^2}\right)(\eta(\lambda)T(\lambda, z)G(z))\left(z^2\Omega\Delta z\right) \\ &\times \int_0^{\Delta z/c} \frac{n_e(t)}{\tau_R}\,dt + N_B \Delta t \end{aligned} \tag{5.31}$$

If we consider the case that the integration time is much longer than the radiative lifetime of the atomic excited state and the laser pulse duration time, then the above integration can be extended to infinity so that

$$\begin{aligned} N_S^{Sat}(\lambda, z) = N_S(\lambda, z)\frac{1}{1 + (\tau_R/t_S)}\Bigg\{1 - \frac{\tau_R}{\Delta t_L}\frac{(\tau_R/t_S)}{1 + (\tau_R/t_S)} \\ \times \left[\exp\left(-\frac{\Delta t_L}{\tau}\left(1 + \frac{\tau_R}{t_S}\right)\right) - 1\right]\Bigg\} + N_B \Delta t \end{aligned} \tag{5.32}$$

where $N_S(\lambda, z)$ is the number of photons received when no saturation effect exists, and given by

$$N_S(\lambda,z) = N_L \sigma_{eff}(\lambda) n_c(z) \Delta z \left(\frac{A}{4\pi z^2}\right) \times \left(\eta(\lambda) T^2(\lambda,z) G(z)\right) + N_B \Delta t \tag{5.33}$$

Equation (5.32) shows that the saturation effect is independent of the Na number density and essentially determined by the ratio of the excited-state radiative lifetime to the saturation time (τ_R/t_S). In general, the smaller the ratio, the less the saturation. If $\tau_R = 0$ (such as the Rayleigh or Mie scattering), then $N_S^{Sat} = N_S$; that is, there is no saturation effect. From the definition of saturation time by Equation (5.29), we may find that the longer laser pulse duration Δt_L, lower laser pulse energy N_L, and smaller effective cross-section σ_{eff} will make the saturation time t_S longer linearly, while the beam divergence θ_L will affect the saturation time quadratically. If the beam divergence decreases by factor of 2, then the saturation time will decrease by factor of 4, which can cause much severe saturation effect.

Considering the Na lidar as an example, we have the Na radiative lifetime $\tau_R = 16.23$ nsec. Assuming the laser pulse energy of 50 mJ with a duration time of $\Delta t_L = 6$ nsec, the number of photons emitted per pulse is then given by $N_L = 1.48 \times 10^{17}$. For the beam divergence $\theta_L = 1$ mrad, atmospheric transmission $T = 0.7$, altitude around $z = 90$ km, and the effective cross-section $\sigma_{eff} = 10 \times 10^{-16}$ m^2, we get $t_S = 184$ nsec. Substituting these numbers into Equation (5.32), we obtain the ratio $N_S^{Sat}/N_S = 98.5\%$. If the beam divergence is compressed to $\theta_L = 0.4$ mrad, then $t_S = 29.4$ nsec and $N_S^{Sat}/N_S = 91.5\%$. When using the small beam divergence angle for daytime measurements, if we do not count in the saturation effect, the photon count will be approximately 9% less than expected. This saturation effect should not, therefore, be neglected.

In order to derive an analytical solution of the saturation effect, we made several approximations such as the rectangular pulse shape. When using different pulse shapes, for example, the Gaussian shape, the saturation results will be

slightly different. However, Equation (5.32) still represents a good estimation of the saturation effect for actual situation if the Δt_L is carefully chosen. We would like to point out that the absolute saturation level would directly affect the constituent density measurement. However, the influence on the temperature and wind measurements is related to the relative saturation effects at different frequencies, not the absolute saturation. When a narrowband Na lidar is used, the considerations for the Na saturation computation will be quite different from above treatments that are more suitable for broadband lidars. More detailed analysis and numerical simulation results for the Na saturation can be found in Welsh and Gardner (1989), von der Gathen (1991), and Milonni et al. (1998, 1999).

5.2.1.4. Solutions for the Resonance Fluorescence Lidar Equation

If we include the saturation correction factor into the total effective scattering cross-section σ_{eff}, then the resonance fluorescence lidar equation can be generally written as Equation (5.24). By solving this equation, we can derive the number density of the constituent in the mesopause region as

$$n_c(z) = \frac{N_S(\lambda, z) - N_B \Delta t}{\left(\frac{P_L(\lambda)\Delta t}{hc/\lambda}\right)(\sigma_{\text{eff}}(\lambda) R_B(\lambda) \Delta z)\left(\frac{A}{4\pi z^2}\right)\left(\eta(\lambda) T_a^2(\lambda) E^2(\lambda, z) G(z)\right)} \tag{5.34}$$

If the atmosphere transmission and lidar system parameters are known, the absolute constituent number density can be obtained from Equation (5.34). However, the uncertainty in laser power, atmosphere transmission, and optical efficiency prevents the calculation of absolute number density of the constituent. The most common method for solving the lidar equation is to normalize the resonance fluorescence signals to the Rayleigh scattering signals at a common lower altitude (e.g., 35 km for the Na lidar signal) to obtain a relative number density profile. The relative number density is then modified by the absolute atmosphere number density at this lower

altitude, which is either measured by other instruments or adopted from atmospheric models, to obtain the absolute number density of the constituent in the mesopause region. We rewrite the resonance fluorescence lidar equation and Rayleigh scattering lidar equation as below:

$$N_S(\lambda,z) = \left(\frac{P_L(\lambda)\Delta t}{hc/\lambda}\right)(\sigma_{eff}(\lambda)n_c(z)R_B(\lambda)\Delta z)\left(\frac{A}{4\pi z^2}\right) \times\left(\eta(\lambda)T_a^2(\lambda)E^2(\lambda,z)G(z)\right) + N_B\Delta t \tag{5.35}$$

$$N_R(\lambda,z_R) = \left(\frac{P_L(\lambda)\Delta t}{hc/\lambda}\right)(\sigma_R(\pi,\lambda)n_R(z_R)\Delta z)\left(\frac{A}{z_R^2}\right) \times\left(\eta(\lambda)T_a^2(\lambda,z_R)G(z_R)\right) + N_B\Delta t \tag{5.36}$$

The Rayleigh scattering lidar equation is for the Rayleigh scattering signal at the normalization altitude z_R, and the one-way atmosphere transmission is $T_a(z_R)$. By normalizing Equation (5.35) with Equation (5.36), we obtain the relative number density as

$$\frac{n_c(z)}{n_R(z_R)} = \frac{N_S(\lambda,z)-N_B\Delta t}{N_R(\lambda,z_R)-N_B\Delta t}\frac{z^2}{z_R^2}\frac{4\pi\sigma_R(\pi,\lambda)}{\sigma_{eff}(\lambda)R_B(\lambda)}\frac{T_a^2(\lambda,z_R)G(z_R)}{T_a^2(\lambda,z)E^2(\lambda,z)G(z)} \tag{5.37}$$

The normalization altitude z_R is usually chosen in the region free of aerosol particles to avoid any contributions from Mie scattering, typically above 30 km. The Rayleigh scattering signal at z_R should be of the same order of magnitude of the resonance fluorescence signals in the mesopause region in order to avoid any effects due to the signal dynamic range. We further assume $T_a(\lambda,z) = T_a(\lambda,z_R)$ (i.e., the extinction between z_R and z is negligible) and $G(z) = G(z_R) = 1$ (i.e., the receiving telescope can see the full laser beam at both z and z_R ranges). Thus, the number density of constituent in the mesopause region can be derived as

$$n_c(z) = n_R(z_R)\frac{N_S(\lambda,z)-N_B\Delta t}{N_R(\lambda,z_R)-N_B\Delta t}\frac{z^2}{z_R^2}\frac{4\pi\sigma_R(\pi,\lambda)}{\sigma_{eff}(\lambda)R_B(\lambda)}\frac{1}{E^2(\lambda,z)} \tag{5.38}$$

The resonance fluorescence and Rayleigh scattering photon counts, the ranges, the total effective scattering cross-section, and the extinction coefficient are measured from the lidar observations while the Rayleigh backscatter cross-section can be calculated from quantum mechanics. The atmosphere molecular number density $n_R(z_R)$ at the normalization altitude z_R is usually taken from atmosphere models. Therefore, the absolute number density of the constituent in the mesopause region can be computed from Equation (5.38).

An alternative way is to relate the product of the Rayleigh backscatter cross-section and the atmosphere molecular number density to the atmosphere temperature and pressure at the same altitude using the relationship of Equation (5.10). Then, by taking the temperature and pressure values from atmosphere models or from other instrument measurements, we can also compute the absolute number density of the constituent by the following solution:

$$n_c(z) = \frac{N_S(\lambda,z) - N_B\Delta t}{N_R(\lambda,z_R) - N_B\Delta t}\frac{z^2}{z_R^2}\frac{4\pi}{\sigma_{eff}(\lambda)R_B(\lambda)E^2(\lambda,z)} \times 2.938\times 10^{-32}\left(\frac{P(z_R)}{T(z_R)}\right)\frac{1}{\lambda^{4.0117}} \tag{5.39}$$

where λ is the laser wavelength in meters, T is the atmosphere temperature in Kelvin, and P is the atmosphere pressure in millibars at the normalization range z_R.

5.2.2. Na Wind/Temperature Lidar

5.2.2.1. Introduction

Range-resolved high-resolution temperature profiles of the mesopause region can be obtained by active probing of the Doppler broadening of the hyperfine structure of the sodium D_2 line by a narrowband lidar. This idea was first proposed and demonstrated by Gibson et al. (1979) and Thomas et al. (1980), who were able to derive the temperature near the peak of Na layer with error of ± 15 K. The first practical narrowband Na temperature lidar was developed

at the University of Bonn with an excimer-laser pumped dye-laser system (Fricke and von Zahn, 1985). The lidar was deployed at Andoya, Norway, and routinely obtained Na temperature profiles above Andoya. The system performance was limited by the excimer-dye laser technology. Because the frequency of their pulsed laser was neither reproducible nor predictable, the relative frequency of each individual laser pulse was measured with a wavelength meter, which was calibrated to a sodium vapor cell, and the laser frequency was scanned over a relatively wide frequency range to adequately cover the Doppler-broadened Na spectrum. A least-square fit was then applied to the spectra using a theoretical model of sodium spectroscopy to derive temperature at each altitude. Using this technique, Fricke and von Zahn (1985) reported an uncertainty of ± 5 K at the layer peak with an integration period of 10 min and a vertical resolution of 1 km.

A more sophisticated narrowband Na lidar system using a two-frequency technique was developed through a collaboration between the Colorado State University (CSU) (She's group) and the University of Illinois at Urbana-Champaign (UIUC) (Gardner's group) (She et al., 1990; Bills et al., 1991a). The system employed a pulsed-dye-amplified, frequency-stabilized, single-mode, ring dye laser and the Doppler-free saturation–absorption spectroscopy of a Na vapor cell (She et al., 1990). The new Na lidar system and measurement techniques dramatically improved the resolution and precision of Na temperature measurements. Absolute temperature accuracies at the Na layer peak of better than ± 3 K with a vertical resolution of 1 km and an integration period of approximately 5 min were achieved (She et al., 1990). This measurement accuracy was actually limited by the theoretical and experimental uncertainties in the determination of the Na saturation spectrum. She and Yu (1995) did a detailed theoretical study and experimental measurement to determine the Doppler-free feature within ± 0.1 MHz. These studies improved the measurement accuracy to ± 1 K.

Since the cw ring laser frequency is locked to the Na D_2 line using the Doppler-free spectroscopy, the laser frequency can be calibrated to an absolute frequency reference. This also

enabled the wind measurements in the mesopause region using four-frequency technique as demonstrated in Bills et al. (1991b) and using a three-frequency technique as demonstrated in She and Yu (1994). Error budgets for Na lidar measurements were presented in these early publications. Papen et al. (1995a) gave a detailed systematic error analysis for the two- and three-frequency techniques. The two- and three-frequency lidar techniques dominate the Na temperature and wind measurements today because they provide higher temporal resolutions, and the fast switching between two or three frequencies can help eliminate the bias caused by Na density fluctuations over the measurement time.

Chen et al. (1996) demonstrated the daytime temperature measurements of Na lidar using a Na Faraday anomalous dispersion optical filter (FADOF) to reject solar scattering noise in the CSU lidar receiver. A single etalon was employed in the UIUC lidar receiver to reject sky background noise enabling the full diurnal temperature measurements (Yu et al., 1997). The CSU group further improved the Na lidar receiver part and now is able to measure both temperature and wind through a complete 24-h diurnal cycle with a moderate receiving telescope (She et al., 2003). Concurrently, the UIUC group coupled the Na Doppler lidar transmitter with a large telescope of 3.5 m diameter through coude optics to make a steerable lidar with large power aperture product. This lidar enables more sophisticated atmosphere measurements, such as tracking the persistent meteor trails (Chu et al., 2000b), heat and momentum flux of gravity waves, and their dynamical influences on the atmosphere mean circulation (Tao and Gardner, 1995; Gardner and Yang, 1998; Zhao et al., 2003).

5.2.2.2. Measurement Principle (Doppler Technique)

The energy levels and spectrum of the alkali Na atoms have been studied in detail. The accuracy of the structure of the Na D_2 electric-dipole transition has been determined to a fraction of a megahertz (Arimondo et al., 1977). The energy levels of the ground state ($3s\ ^2S_{1/2}$) and the first excited state ($3p\ ^2P_{1/2}$

and 3p $^2P_{3/2}$) of sodium atoms are shown in Figure 5.1. The D_2 and D_1 lines are the transitions from 3p $^2P_{3/2}$ to 3s $^2S_{1/2}$ and from 3p $^2P_{1/2}$ to 3s $^2S_{1/2}$, respectively. The transition probabilities (Einstein A_{ki} coefficient) for the Na D_2 and D_1 lines are 0.616×10^8 and 0.614×10^8 s^{-1}, respectively. They correspond to the radiative lifetimes of $^2P_{3/2}$ and $^2P_{1/2}$ of 16.23 ± 0.03 and 16.29 ± 0.03 nsec (Jones et al., 1996; Oates et al., 1996; Volz et al., 1996). The Na D_2 line has a natural linewidth of 9.802 ± 0.022 MHz (Oates et al., 1996) and an oscillator strength f_0 of 0.6411. The Na nuclear angular momentum $I = 3/2$. Thus, the ground state $^2S_{1/2}$ splits into two hyperfine levels with total angular momentum $F = 1$ and $F = 2$. The excited states $^2P_{1/2}$ and $^2P_{3/2}$ split to two and four hyperfine levels with F values ranging from 1 to 2 and from 0 to 3, respectively. The D_2 line consists of two groups of lines, D_{2a} and D_{2b}, which are formed by the hyperfine splits of the Na ground state. The frequency separation between these two groups is 1771.6261 MHz (Kasevich et al., 1989). Each group contains three lines resulting from the hyperfine splits of the Na excite state of 3p $^2P_{3/2}$. The related numerical information of Na D_1 and D_2 lines is listed in Table 5.1. Here the zero-frequency reference is the weighted center of the D_2 six

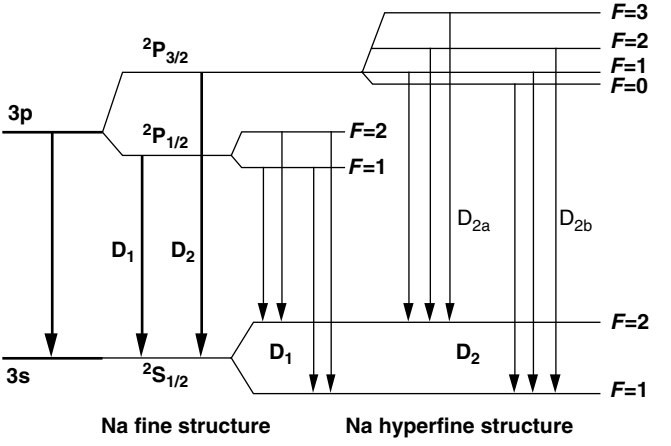


Figure 5.1 Energy level diagram of atomic Na. (Adapted from She, C.Y., and Yu, J.R., *Appl. Opt.*, 34, 1063–1075, 1995. With permission.)

Table 5.1 Parameters of the Na D_1 and D_2 Transition Lines

Transition Line	Central Wavelength (nm)	Transition Probability ($10^8\,s^{-1}$)	Radiative Lifetime (nsec)	Oscillator Strength f_{ik}
D_1 ($^2P_{1/2} \rightarrow {}^2S_{1/2}$)	589.7558	0.614	16.29	0.320
D_2 ($^2P_{3/2} \rightarrow {}^2S_{1/2}$)	589.1583	0.616	16.23	0.641

Group	$^2S_{1/2}$	$^2P_{3/2}$	Offset (GHz)	Relative Line Strength[a]
D_{2b}	$F=1$	$F=2$	1.0911	5/32
		$F=1$	1.0566	5/32
		$F=0$	1.0408	2/32
D_{2a}	$F=2$	$F=3$	−0.6216	14/32
		$F=2$	−0.6806	5/32
		$F=1$	−0.7150	1/32

Doppler-Free Saturation–Absorption Features of the Na D_2 Line

f_a (MHz)	f_c (MHz)	f_b (MHz)	f_+ (MHz)	f_- (MHz)
−651.4	187.8	1067.8	−21.4	−1281.4

[a]Relative line strengths are in the absence of a magnetic field or the spatial average. When Hanle effect is considered in the atmosphere, the relative line strengths will be modified depending on the geomagnetic field and the laser polarization.

transitions, whose wavelength $\lambda_0 = 589.15826\,\text{nm}$ *in vacuo* (Martin and Zalubas, 1981). The relative oscillator strengths listed in the table are the values in the absence of a magnetic field.

For single Na atom, its absorption line has a Lorentzian shape given by Equation (5.13) with natural linewidth of about 10 MHz. Because of the Doppler effect, the absorption line will be broadened. To an atom moving away from the laser with velocity v_R, the laser frequency perceived by the atom appears to be shifted:

$$\nu' = \nu\left(1 - \frac{v_R}{c}\right) \tag{5.40}$$

where ν is the laser central frequency seen by an atom at rest, ν' is the Doppler-shifted laser frequency, v_R is the radial velocity of the atom ($v_R > 0$ if the atom moving away from the laser radiation source), and c is the speed of light. Since the

velocity distribution obeys the Maxwell–Boltzmann distribution (given by Equation 5.14) under thermodynamic equilibrium, the detected average absorption cross-section of the Na (D_2) line will be the single atom absorption cross-section integrating over the velocity distribution. Thus, the absorption line is Doppler broadened. Since the collisional broadening at the mesopause region is negligible due to the low atmosphere density, the D_2 line has almost pure Doppler-broadening lineshape. Plotted in Figure 5.2(a) is the computed Na absorption cross-section for three different atmospheric temperatures at $v_R = 0$ m/s, and Figure 5.2(b) for three different radial wind velocities at $T = 200$ K. The Doppler-broadened line has a Gaussian shape given by

$$\sigma_{abs}(\nu) = \frac{1}{\sqrt{2\pi}\sigma_D} \frac{e^2 f}{4\epsilon_0 m_e c} \sum_{n=1}^{6} A_n \exp\left(- \frac{\left[\nu_n - \nu\left(1 - \frac{v_R}{c}\right)\right]^2}{2\sigma_D^2} \right) \tag{5.41}$$

where f is the oscillator strength of Na D_2 line, ν_n and A_n are the center frequencies and the line strengths of the six electric-dipole-allowed D_2 transitions, respectively, ν is the laser frequency, v_R is the radial velocity of the Na atom moving away from the laser source, and c is the light speed. The rms width of the Doppler-broadened line is σ_D given by

$$\sigma_D = \sqrt{\frac{k_B T}{M\lambda_0^2}} \tag{5.42}$$

where M is the mass of a single Na atom, k_B is the Boltzmann constant, λ_0 is the mean Na D_2 transition wavelength, and T is the temperature.

The Doppler-broadened linewidth is a sensitive function of temperature as shown in Equation (5.42) and Figure 5.2(a). If we measure this Doppler-broadened linewidth, then we can derive the temperature of the Na atoms. Since the Na atoms are in equilibrium with the surrounding atmosphere, the Na temperature would be equal to the atmosphere temperature in the mesopause region. Examining Equation (5.41) and Figure 5.2(b), the radial velocity v_R of Na atoms, and thus of

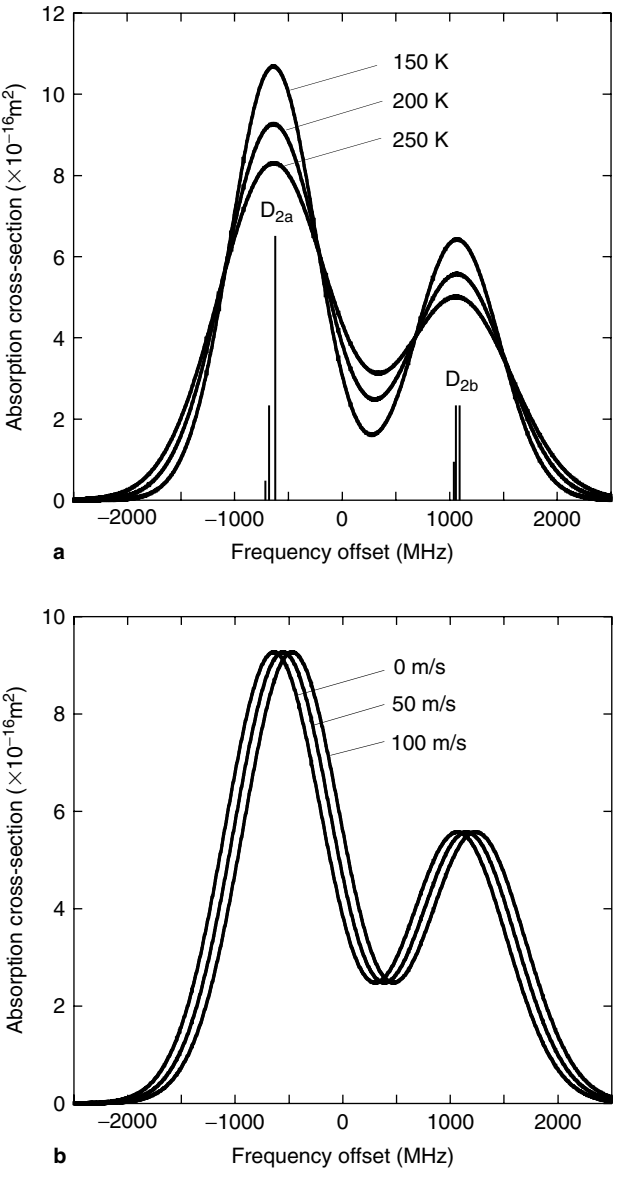


Figure 5.2 (a) Na absorption cross section for three temperatures at $v_R = 0$ m/s. (b) Na absorption cross section for three radial wind velocities at $T = 200$ K.

the atmosphere, can be derived by measuring the Doppler frequency shift of the central frequency of the Na D_2 line. Due to the finite linewidth of the laser beam, the total scattering cross-section σ_{eff} is the convolution of the absorption cross-section and the laser lineshape. The actual σ_{eff} depends on the laser lineshape. Under the assumption of Gaussian lineshape of the laser, σ_{eff} can be written as

$$\sigma_{\text{eff}}(\nu) = \frac{1}{\sqrt{2\pi}\sigma_{\text{e}}} \frac{e^2 f}{4\epsilon_0 m_{\text{e}} c} \sum_{n=1}^{6} A_n \exp\left(- \frac{\left[\nu_n - \nu\left(1 - \frac{v_{\text{R}}}{c}\right)\right]^2}{2\sigma_{\text{e}}^2} \right) \tag{5.43}$$

where $\sigma_{\text{e}} = \sqrt{\sigma_{\text{D}}^2 + \sigma_{\text{L}}^2}$ is the rms width of the total lineshape.

A sodium vapor cell with sufficient saturation vapor pressure can be maintained at reasonable temperature ($\sim 60^\circ$C). Under the simultaneous illumination of two counter-propagating beams at a saturated laser intensity, the Na fluorescence from the vapor cell shows the Doppler-free features at D_{2a} and D_{2b} peaks (f_{a} and f_{b}) as well as at the crossover resonance frequency f_{c}. These features are similar to those observed with the well-known Doppler-free saturated-absorption spectroscopy (Hansch et al., 1971). A measured Doppler-free spectroscopy in a Na vapor cell is shown in Figure 5.3 (Bills et al., 1991a). She and Yu (1995) studied the Doppler-free saturation fluorescence spectroscopy of Na atoms in great detail both theoretically and experimentally. They determined the frequencies of the dominant Doppler-free features to within ± 0.1 MHz from a simulated spectrum. The frequencies for D_{2a}, crossover, and D_{2b} resonances are -651.4, 187.8, and 1067.8 MHz, respectively, relative to the weighted center of the six D_2 hyperfine transition lines. These features are essential for the operation of the narrowband Na resonance fluorescence lidar because these Doppler-free dips or peaks can be used to lock the laser frequency precisely and provide an absolute frequency calibration for the lidar.

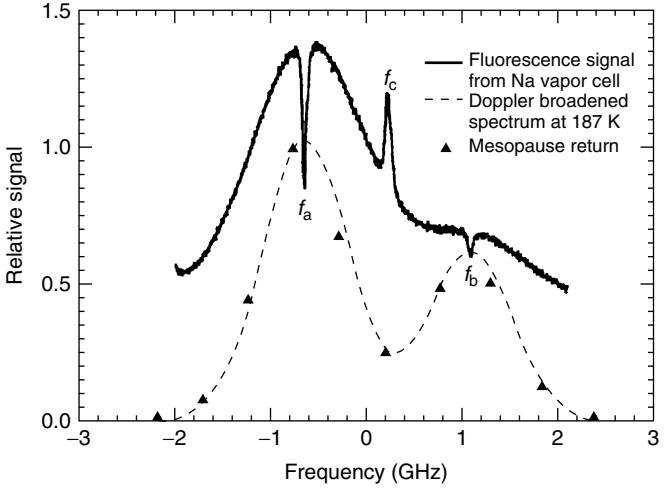


Figure 5.3 Doppler-free saturation–absorption spectroscopy measured from a Na vapor cell. (From Bills, R.E., Gardner, C.S., and She, C.Y., *Opt. Eng.*, 30, 13–21, 1991. With permission.)

5.2.2.2.1. Frequency-scanning method

Starting from the lidar equations (5.35) and (5.36), the measured Na photon counts and Rayleigh photon counts are given as

$$N_{\mathrm{Na}}(\lambda,z) = \left(\frac{P_{\mathrm{L}}(\lambda)\Delta t}{hc/\lambda}\right)(\sigma_{\mathrm{eff}}(\lambda)n_{\mathrm{Na}}(z)\Delta z)\left(\frac{A}{4\pi z^2}\right) \times(\eta(\lambda)T_{\mathrm{a}}^2(\lambda)E^2(\lambda,z)G(z)) \tag{5.44}$$

$$N_{\mathrm{R}}(\lambda,z_{\mathrm{R}}) = \left(\frac{P_{\mathrm{L}}(\lambda)\Delta t}{hc/\lambda}\right)(\sigma_{\mathrm{R}}(\pi,\lambda)n_{\mathrm{R}}(z_{\mathrm{R}})\Delta z)\left(\frac{A}{z_{\mathrm{R}}^2}\right) \times(\eta(\lambda)T_{\mathrm{a}}^2(\lambda,z_{\mathrm{R}})G(z_{\mathrm{R}})) \tag{5.45}$$

Equations (5.44) and (5.45) represent, respectively, the pure Na and Rayleigh photon counts as the background has been subtracted from the raw photon counts. From these two lidar equations, we obtain the measured total scattering cross-section as

$$\sigma_{\text{eff}}(\lambda, z) = \frac{C(z)}{E^2(\lambda, z)} \frac{N_{\text{Na}}(\lambda, z)}{N_{\text{R}}(\lambda, z_{\text{R}})} \tag{5.46}$$

where

$$C(z) = \frac{\sigma_{\text{R}}(\pi, \lambda) n_{\text{R}}(z_{\text{R}})}{n_{\text{Na}}(z)} \frac{4\pi z^2}{z_{\text{R}}^2} \tag{5.47}$$

Here we have assumed $T_{\text{a}}(\lambda,\ z) = T_{\text{a}}(\lambda, z_{\text{R}})$ and $G(z) = G(z_{\text{R}}) = 1$. It is apparent from Equation (5.46) that the wavelength-dependent lineshape of the total scattering cross-section σ_{eff} at each altitude z can be obtained by scanning laser frequency through the Na(D_2) resonance line. If we assume that the Na atom number density $n_{\text{Na}}(z)$ remains unchanged during the period of the laser frequency scan, then $C(z)$ is a constant for each altitude z. By fitting the theoretical function of the total effective scattering cross-section (Equation 5.43) to the measured fluorescence signal scanning over a selected range of frequencies (Equation 5.46), the Na temperature $T(z)$ at each altitude z can be derived from the fitting parameters. Thus, the range-resolved temperature profiles in the mesopause region can be achieved. This scanning method was first demonstrated by Fricke and von Zahn (1985). Shown in Figure 5.4 is an example taken from Fricke and von Zahn (1985). The laser frequency was scanned through a relative large range around the D_{2a} peak. The processed signals for the altitude interval 90.5 to 91.5 km were plotted as the relative intensities versus the relative wavelength in Figure 5.4. A best fit of the Doppler broadened line shape was applied to the data and plotted as a solid line in the figure. The convolution with the instrument profile was included. The fitting parameters give an atmospheric temperature of 195 K at the altitude of 90.5 to 91.5 km. The range-resolved temperature profile was then obtained by applying this method to the altitude range where the Na signals are large enough, as shown in Figure 5.5 (Fricke and von Zahn, 1985).

If the laser frequencies are calibrated against an absolute frequency reference, then the fitting method to the scanned

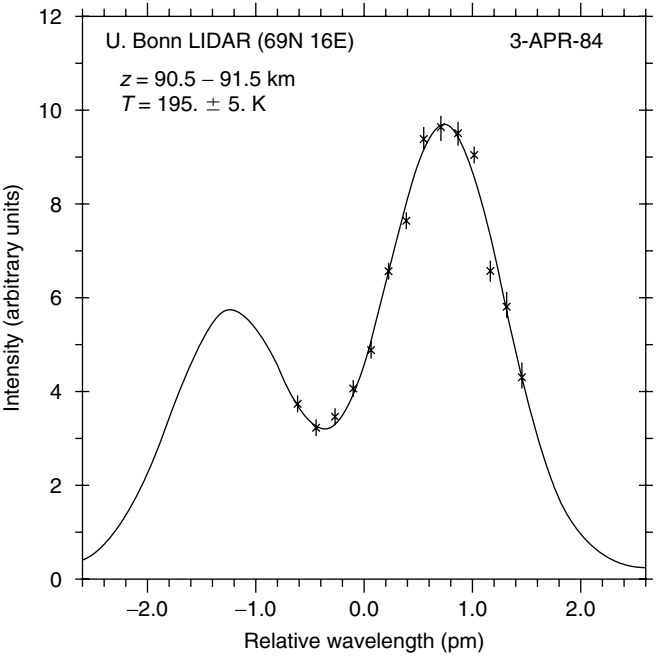


Figure 5.4 Relative intensities from the D_{2a} component of the Na hyperfine structure in the terrestrial atmosphere in the 91 ± 0.5 km range observed on April 3, 1984 at Bleik (69°N, 16°E) on Andoya Island. The solid line is the best fit of the Doppler widened line shape to the data. The convolution with the instrument profile function is included. It corresponds to an atmospheric temperature of 195 K. (From Fricke, K.H., and von Zahn, U., *J. Atmos. Terr. Phys.*, 47, 499–512, 1985. With permission.)

fluorescence signals can also derive the radial wind v_R as indicated by Equation (5.43).

5.2.2.2.2. Ratio techniques: two-frequency

Scanning through a large frequency range reduces the temporal resolution and causes uncertainties because of inherent Na density fluctuations during the measurement time. Examining Figure 5.2(a), the signal levels at the D_{2a} peak and at the intermediate minimum between the D_{2a} and D_{2b} peaks are particularly sensitive to temperature. Since the responses

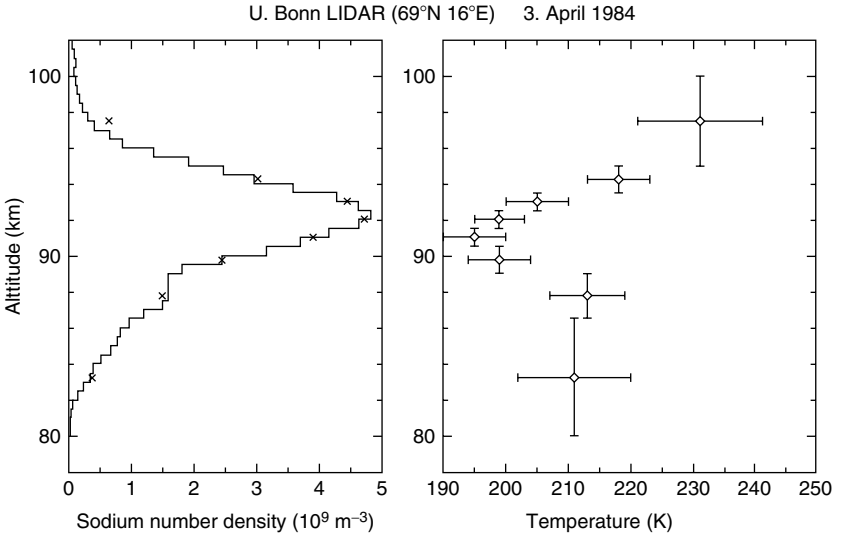


Figure 5.5 Na density and temperature profiles obtained on April 3, 1984 at Andoya (69°N, 16°E) by University of Bonn narrowband Na Doppler lidar. The Na layer was subdivided into eight slices of about equal signal strength to obtain comparable signal/noise ratios throughout the layer. Note that the absolute value of the temperature scale is uncertain by about ± 10 K because of the remaining uncertainty in the lineshape of the lidar transmitter. Error bars in the figure show statistical errors from the fitting procedure only. (From Fricke, K.H., and von Zahn, U., *J. Atmos. Terr. Phys*., 47, 499–512, 1985. With permission.)

of the signal levels at these two frequencies to temperature changes are opposite to each other, the ratio of the fluorescence signals at the intermediate minimum and at the D_{2a} peak is a very sensitive function of temperature. In practice, the laser frequency is tuned to the Doppler-free feature at the D_{2a} peak f_a and at the crossover frequency f_c using a Na vapor cell. We define the normalized Na photon count as

$$N_{\text{norm}}(f,z,t) = \frac{N_{\text{Na}}(f,z,t)}{N_{\text{R}}(f,z,t)E^2(f,z)} \tag{5.48}$$

From the lidar equations (5.44) and (5.45), we have

$$N_{\text{norm}}(f,z,t) = \frac{\sigma_{\text{eff}}(f)n_{\text{Na}}(z)}{\sigma_{\text{R}}(\pi,f)n_{\text{R}}(z_{\text{R}})}\frac{z_{\text{R}}^2}{4\pi z^2} \tag{5.49}$$

The ratio of the normalized Na photon counts at the crossover frequency f_c to the D_{2a} peak frequency f_a is then defined as

$$R_{\text{T}}(z) = \frac{N_{\text{norm}}(f_c,z,t_1)}{N_{\text{norm}}(f_a,z,t_2)} = \frac{\sigma_{\text{eff}}(f_c,z)n_{\text{Na}}(z,t_1)}{\sigma_{\text{eff}}(f_a,z)n_{\text{Na}}(z,t_2)} \approx \frac{\sigma_{\text{eff}}(f_c,z)}{\sigma_{\text{eff}}(f_a,z)} \tag{5.50}$$

To reach the last equality in Equation (5.50), two assumptions have been made: (1) the frequency difference between f_c and f_a is so small that their Rayleigh scattering cross sections at these two frequencies are the same; (2) the Na densities are the same at time t_1 and t_2. The σ_{eff} in Equation (5.50) is given by Equation (5.43). Plotted in Figure 5.6 is the ratio R_{T} versus the temperature T from 100 to 300 K for $v_{\text{R}} = 0$ m/s. Figure 5.6 illustrates that R_{T} is a sensitive function of temperature and can be used to accurately derive the Na temperature.

There are several reasons for choosing f_a and f_c as the two operational frequencies. (1) The ratio of the fluorescence signals at these two frequencies is a very sensitive function of

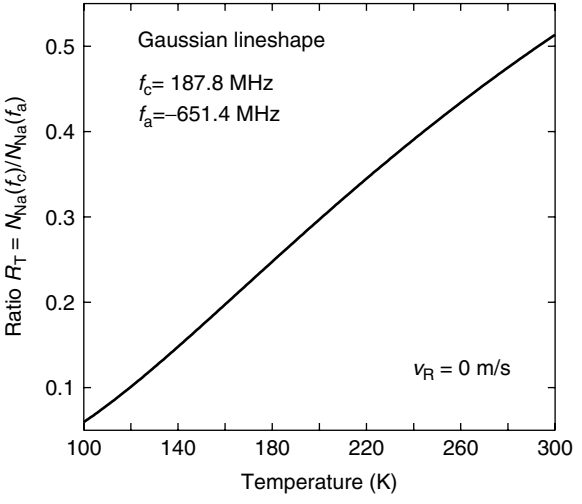


Figure 5.6 Calibration curve (R_{T} versus T) for two-frequency ratio technique of Na wind/temperature lidar.

temperature. (2) Since f_a is at the D_{2a} peak and f_c is near the minimum between the D_{2a} and D_{2b} peaks, where the slope of scattering cross-section versus frequency is close to zero, the fluorescence signal levels from the Na layer at these two frequencies are relatively insensitive to small frequency tuning errors of the laser. (3) Unlike the frequency (f_{min}) at the minimum between the D_{2a} and D_{2b} peaks, which depends on temperature, both f_a and f_c are independent of temperature and can be determined from the Doppler-free features to a very high precision (within ± 0.1 MHz) (She and Yu, 1995). These characteristics simplify the calculation of temperature from the measured ratio of the fluorescence signals and provide a good frequency calibration. Data are collected by tuning the laser to the D_{2a} peak and accumulating the photon returns from certain laser shots, and then tuning to the crossover resonance and accumulating the returns. This process was repeated continuously throughout the observation period. By taking the ratio of the photon counts for two frequencies collected at each altitude, one can derive a range-resolved temperature structure throughout the Na layer region. This is called the *two-frequency ratio technique*, which was first demonstrated by She et al. (1990) and Bills et al. (1991a). The temperature accuracy that is obtained with either the scanning method or the two-frequency ratio technique should be comparable as long as other conditions remain equal. The ratio techniques require precise and repeatable laser tuning as well as pulse-to-pulse stability. Its instrumentation will be described later. Since the laser switches only between two wavelengths, the temporal resolution of the two-frequency ratio technique can be much higher than the scanning method.

5.2.2.2.3. *Ratio techniques: three-frequency*

The Na atom velocity, that is, the wind velocity, has been involved in the effective cross-section given by Equation (5.43). Since the two-frequency measurements cannot derive the Na temperature, number density and the radial wind at the same time, the radial wind has to be assumed as zero when deriving the temperature from the ratio R_T. Typically, the vertical wind is less than 3 m/s, so the temperature error

caused by this assumption would be less than 0.5 K (She et al., 1990). On the other hand, if we measure the fluorescence signals at three or more different frequencies, then we can derive the Na temperature, radial wind, and number density simultaneously. This is called the *multiple-frequency ratio technique* or *three-frequency ratio technique*, first proposed by Bills et al. (1991a) and She et al. (1992), and then demonstrated by Bills et al. (1991b) and She and Yu (1994). Bills et al. (1991b) tuned the laser frequency to each of four frequencies at f_a, $f_a + \Delta f$, f_c, and $f_a - \Delta f$. They derived the temperature using photon count profiles collected at f_a and f_c, and derived the radial wind using photon count profiles collected at $f_a + \Delta f$ and $f_a - \Delta f$. She and Yu (1994) tuned the laser frequency to each of three frequencies at f_a, f_c, and $f_a - \Delta f$, and then derived the temperature and radial wind at the same time.

The modern three-frequency ratio technique is to tune the laser frequency to each of the three frequencies at f_a, $f_a + \Delta f$, and $f_a - \Delta f$, where $\Delta f = 630$ MHz. The temperature is derived from the following ratio:

$$R_T(z) = \frac{N_{norm}(f_+, z, t_1) + N_{norm}(f_-, z, t_2)}{N_{norm}(f_a, z, t_3)} \approx \frac{\sigma_{eff}(f_+, z) + \sigma_{eff}(f_-, z)}{\sigma_{eff}(f_a, z)} \tag{5.51}$$

and the radial wind will be derived from the following ratio:

$$R_W(z) = \frac{N_{norm}(f_-, z, t_2)}{N_{norm}(f_+, z, t_1)} \approx \frac{\sigma_{eff}(f_-, z)}{\sigma_{eff}(f_+, z)} \tag{5.52}$$

Plotted in Figure 5.7(a) and (b) are the ratio R_T versus temperature for the radial wind $v_R = 0$ m/s, and the ratio R_W versus radial wind for the temperature $T = 200$ K. Here, the radial wind $v_R > 0$ if the Na atoms move along the same direction as the laser photon propagation, that is, move away from the laser source. Figure 5.7(b) illustrates that the ratio R_W is a sensitive function of the radial wind. The ratio R_T in the three-frequency technique is not as sensitive as the ratio R_T in the two-frequency technique, but it is still sensitive enough for a good temperature measurement. Indeed, the

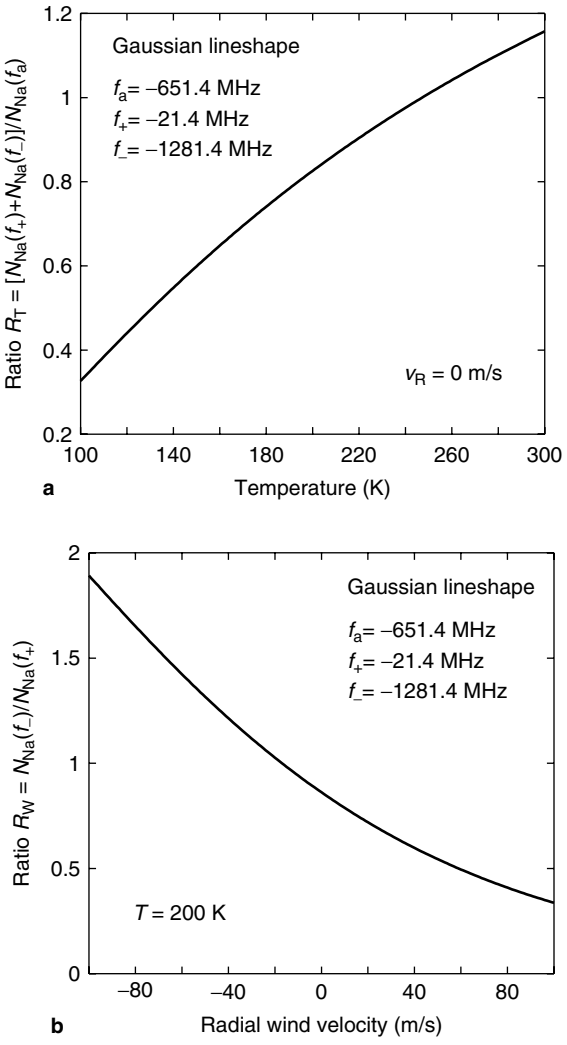


Figure 5.7 (a) Na wind/temperature lidar three-frequency ratio technique: R_T versus T. (b) Na wind/temperature lidar three-frequency ratio technique: R_W versus wind.

temperature and the radial wind are involved with each other in the ratios R_T and R_W, given by Equations (5.51) and (5.52), respectively. Therefore, it is necessary to solve Equations (5.51) and (5.52) simultaneously using iteration.

Both the two- and three-frequency ratio techniques require that the laser pulses have repeatable frequency and lineshape. The radial wind measurement also requires an accurate absolute frequency reference in order to derive the Doppler frequency shift. At a given temperature, the spectrum of the laser-induced fluorescence is the convolution of the laser lineshape with the thermal Doppler-broadened Na D_2 fluorescence spectrum. Therefore, one must use a narrowband laser system with a known lineshape function to experimentally retrieve the spectral information, and thereby the wind and temperature information, from the induced fluorescence of atmospheric Na atoms. The strength of the detected fluorescence signal strongly depends on the excitation frequency. The success of the ratio techniques relies on having a stable, tunable, and narrowband laser transmitter whose center frequency as well as lineshape can be characterized accurately against a known absolute frequency reference. Such a frequency marker should be conveniently available during the data acquisition phase of the lidar operation. Fortunately, it is convenient to utilize a Na vapor cell and the Doppler-free saturation–absorption spectroscopy to achieve this absolute frequency reference, as described earlier. The laser frequency can be locked to these Doppler-free features to have the absolute frequency calibration and to ensure the precise derivation of atmospheric temperature and radial wind from the convoluted laser-induced Na fluorescence.

5.2.2.2.4. Lidar wind measurement versus radar wind measurement

Although both the resonance fluorescence lidar and the radar utilize the effect of Doppler frequency shift to measure wind, the concepts of how to measure the Doppler frequency shift are quite different between the lidar and radar. Briefly, the resonance fluorescence lidar employs incoherent detection method; that is, it converts the frequency shift to the optical intensity information and then measures the frequency shift from the intensity ratios. The radar measures the frequency shift between the returned and transmitted beams directly using coherent detection method.

The radar scattering from charged particles is similar to the Rayleigh scattering of photons from the atmosphere molecules in the sense that they are both elastic scattering, and do not include any real resonance absorption. Suppose a particle has velocity $\vec{v}_1$ before the scattering and has new velocity $\vec{v}_2$ after the scattering. An incident photon with frequency ω_1 has a momentum of $\hbar\vec{k}_1$, where $k_1 = \omega_1/c$. The scattered photon with frequency ω_2 has a new momentum of $\hbar\vec{k}_2$, where $k_2 = \omega_2/c$. Considering the particle and the photon as one system, there is no external force interacting with this system. Therefore, we have the momentum conservation and energy conservation for this particle and photon system before and after the scattering, as shown in Equations (5.53) and (5.54), respectively:

$$m\vec{v}_1 + \hbar\vec{k}_1 = m\vec{v}_2 + \hbar\vec{k}_2 \tag{5.53}$$

$$\frac{1}{2}mv_1^2 + \hbar\omega_1 = \frac{1}{2}mv_2^2 + \hbar\omega_2 \tag{5.54}$$

Elimination of v_2 from Equation (5.54) with the help of Equation (5.53) gives

$$\omega_1 = \omega_2 + \vec{k}_1 \cdot \vec{v}_1 - \vec{k}_2 \cdot \vec{v}_2 + \frac{\hbar k_1^2}{2m} - \frac{\hbar k_2^2}{2m} \tag{5.55}$$

The last two terms on the right-hand side of Equation (5.55) are second-order terms, whose magnitudes are much smaller than the second and the third terms. Thus, the last two terms can be neglected altogether. For backscattered radar or backscattered Rayleigh signals, the returned photons propagate in the opposite direction of the transmitted signals, and the frequency shift is very small compared with the carrier frequency, so $\vec{k}_2 \approx -\vec{k}_1$. A single photon momentum $\hbar\vec{k}$ is much smaller than a particle's momentum, so the particle's velocity will only slightly change after scattering; thus, $\vec{v}_2 \approx \vec{v}_1$. Under these considerations, Equation (5.55) can be simplified, and we obtain the frequency shift of the returned signal relative to the transmitted beam as

$$\Delta\omega_{\text{radar}} = \omega_2 - \omega_1 = -2\vec{k}_1 \cdot \vec{v}_1 \tag{5.56}$$

Thus, in the radar scattering and the Rayleigh scattering cases, the frequency shift between the returned and transmitted signals is two times the Doppler frequency shift. The radar wind measurement relies on coherent detection; that is, taking the beat signal between the transmitted and the returned beams to obtain the frequency shift corresponding to the wind velocity. This is a direct measurement of the Doppler frequency shift.

The resonance fluorescence actually consists of two steps: first, the resonance absorption of an incident photon by an atom; and second, the spontaneous emission of a photon from the excited atom. Here, let us consider a simplified two-energy-level atom. Suppose the atom is in the ground state with energy E_1 and has velocity $\vec{v}_1$ before the absorption takes place. Absorption of a photon of energy $\hbar\omega_1$ promotes the atom into its excited energy level E_2. The photon has a momentum of $\hbar\vec{k}_1$, where $k_1 = \omega_1/c$. The absorption of the photon causes a recoil of the atom to a new velocity $\vec{v}_2$. The total momentum conservation and the energy conservation before and after the absorption give

$$m\vec{v}_1 + \hbar\vec{k}_1 = m\vec{v}_2 \tag{5.57}$$

$$E_1 + 1/2mv_1^2 + \hbar\omega_1 = E_2 + 1/2mv_2^2 \tag{5.58}$$

The resonance absorption frequency of an atom with zero velocity is given by

$$\omega_0 = (E_2 - E_1)/\hbar \tag{5.59}$$

Solving Equations (5.57) and (5.58) gives

$$\omega_1 = \omega_0 + \vec{k}_1 \cdot \vec{v}_1 + \frac{\hbar k_1^2}{2m} \tag{5.60}$$

Neglecting the last term on the right-hand side of Equation (5.60), we obtain the frequency shift of the perceived laser frequency relative to the transmitted laser frequency during the absorption as

$$\Delta\omega_{\text{abs}} = \omega_0 - \omega_1 = -\vec{k}_1 \cdot \vec{v}_1 \tag{5.61}$$

The excited atom then spontaneously emits a photon and decays back to its ground state. For backscattered photons received by a detector of the lidar, the emitted photon travels towards the lidar. Suppose the atom has a new velocity $\vec{v}_3$ and the returned photon has a momentum of $\hbar\vec{k}_2$, where $k_2 = \omega_2/c$. The momentum and energy conservations yield

$$m\vec{v}_2 = m\vec{v}_3 + \hbar\vec{k}_2 \tag{5.62}$$

$$E_2 + \frac{1}{2}mv_2^2 = E_1 + \frac{1}{2}mv_3^2 + \hbar\omega_2 \tag{5.63}$$

Solving Equations (5.62) and (5.63), we obtain

$$\omega_2 = \omega_0 + \vec{k}_2 \cdot \vec{v}_3 + \frac{\hbar k_2^2}{2m} \tag{5.64}$$

Recalling the similar discussion above, we have $\vec{k}_2 \approx -\vec{k}_1$ and $\vec{v}_3 \approx \vec{v}_1$. Thus, the frequency shift of the returned photon relative to the resonance frequency during spontaneous emission is given by

$$\Delta\omega_{\text{sp}} = \omega_2 - \omega_0 = -\vec{k}_1 \cdot \vec{v}_1 \tag{5.65}$$

We can obtain the frequency shift of the returned fluorescence photon frequency relative to the transmitted photon frequency from Equations (5.61) and (5.65) as follows:

$$\Delta\omega_{\text{overall}} = \omega_2 - \omega_1 = -2\vec{k}_1 \cdot \vec{v}_1 \tag{5.66}$$

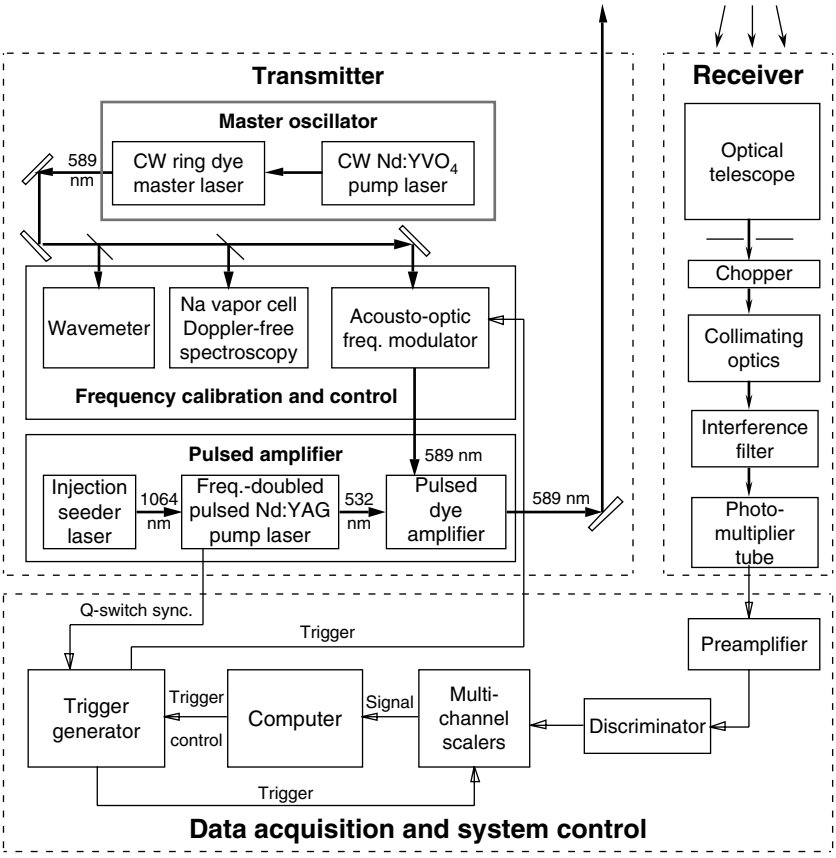
For the two-energy-level atom, the returned fluorescence photon has two times the Doppler frequency shift relative to the transmitted photon frequency, which is similar to the radar scattering and Rayleigh scattering. However, the fluorescence photons have random phases, so the frequency shift cannot be detected by the coherent method used by the radar. As described in the three-frequency ratio technique above, the resonance fluorescence lidar actually converts the frequency information to the intensity signals of returned fluorescence photons and then derives the frequency shift from the intensity ratios. In other words, for each operating frequency of the laser, the lidar receiver only counts the total number of

returned fluorescence photons no matter what frequencies they have. This is especially true when using the broadband optical filters in the lidar receiver. The filter function is flat in the frequency range we are interested in, so all returned fluorescence photons will be counted in the same way. Then the ratio of the photon counts collected at each operating frequency of the laser (e.g., $f_a \pm \Delta f$) is taken to give the frequency-shift information. The returned fluorescence photon number is proportional to the absorption cross-section of the atom given by Equation (5.41). Since only the number (intensity), not the frequency, of the fluorescence photons is considered in the wind measurements, the effective scattering cross-section will be given by Equation (5.43), which only includes the Doppler frequency shift associated with the resonance absorption process, but excludes the shift associated with the spontaneous emission. Therefore, for the resonance fluorescence lidar, the Doppler frequency shift derived from the photon intensity signals will be determined by Equation (5.61), rather than Equation (5.66).

The frequency shift of the fluorescence photons associated with the spontaneous emission needs to be considered when using narrowband optical filter (such as a Na FADOF, see Section 5.2.2.4) in the lidar receiver. In this case, the filter function is no longer flat at the frequencies of interest. So the filter function needs to be taken into account when computing the photon intensity ratio, and the frequency shift of the fluorescence photons will influence the transmission through the filter, and thus, the receiving efficiency. Nevertheless, the Doppler frequency shift derived from the photon intensity signals will still be determined by Equation (5.61), that is, only one-time Doppler frequency shift.

5.2.2.3. Na Doppler Lidar Instrumentation

As discussed earlier, a narrowband Na lidar is required in order to resolve the temperature and wind information from the laser-induced fluorescence signals. To illustrate the instrumentation of the modern Na wind/temperature lidar, a schematic diagram of the UIUC Na wind/temperature lidar



Na Wind/Temperature Lidar System

Figure 5.8 Schematic diagram of the University of Illinois Na wind/temperature lidar system. The lidar consists of three subsystems: a transmitter, a receiver, and data acquisition and system control. The thick solid lines with solid arrows denote the optical path, while the thin solid lines with open arrows indicate the electronic signals. The major components of the lidar transmitter are the master oscillator, frequency calibration and control, and pulsed amplifier.

is plotted in Figure 5.8. The lidar system consists of three subsystems: a transmitter, a receiver, and a data acquisition and control system. The lidar transmitter is to provide laser pulses with high precision frequency, narrowband linewidth,

high power, and small divergence. The lidar receiver is to collect backscatter photons, reject the solar background noise, and count the returned photon numbers. The data acquisition and control system interacts with both the transmitter and the receiver to record the photon counts and to provide the timing and system control for the whole lidar system. A list of the system parameters is given in Table 5.2.

Table 5.2 System Parameters for the University of Illinois Na Wind/Temperature Lidar

Lidar Transmitter		
Characteristics	Ring Dye Laser Master Oscillator	Pulsed Dye Amplifier
Model	Coherent 899-21	Spectra-Physics Quanta-Ray PDA-1
Optical pump laser	Coherent Nd:YVO_4 laser Verdi V-5 CW at 532 nm	Spectra-Physics Nd:YAG Laser Quanta-Ray GCR-3 pulsed at 532 nm
Pump power/energy per pulse	4 W	300 mJ
Gain medium	Dye Rhodamine 6G	RH640 + KR620
Pulse repetition rate	CW	30 Hz
Pulse length	–	6 nsec
Linewidth	500 kHz (rms)	60 MHz (rms)
Average output power	500–600 mW (350–400 mW at PDA input)	1.2–1.5 W (40–50 mJ per pulse)
Beam divergence	–	1 mrad FW at e^{-2}
Wavelength	589.1583 nm	589.1583 nm

Lidar Receiver	
Characteristics	
Telescope	SOR reflecting telescope
Aperture diameter	3.5 m
Aperture area	9.6 m^2
Telescope focal length	315 m
Field of view	1 mrad
Interference filter bandwidth	3 nm (FWHM)
PMT quantum efficiency	11%
Range resolution	24 m

The Na lidar transmitter can be further divided into three sections for ease of explanation: (1) master oscillator, (2) frequency calibration and control, and (3) pulsed amplifier. The master oscillator is a high-performance single-frequency cw ring dye laser, optically pumped by a frequency-doubled cw Nd:YVO_4 laser. (Originally, an argon ion laser was used to pump the ring dye laser.) The frequency calibration and control monitors the ring dye laser wavelength with a wavemeter, locks the laser frequency to the D_{2a} Lamb dip or crossover peak of the Na D_2 line using the Doppler-free saturation–absorption spectroscopy of a Na vapor cell, and creates the two wing-operating frequencies for the three-frequency ratio technique by using an acousto-optic modulator. The pulsed amplifier provides pulsed amplification to the cw master laser to generate laser pulses needed for the lidar application. The pulsed amplification is achieved by the interaction of the cw master laser beam with the pump beam from a powerful frequency-doubled pulsed Nd:YAG laser. The laser pulses are then transmitted to the atmosphere. Return photons are collected by a receiving optical telescope, counted by a PMT, binned by multichannel scalers, and saved by a computer. Interference filters are used to reject most of the background skylight. Such a lidar system achieves the excellent laser pulse-to-pulse stability and provides a high degree of tuning accuracy, which are difficult to achieve with pulsed laser oscillators, but are essential for the temperature and wind measurements. In the following text, we discuss several key components for the system.

5.2.2.3.1. Master oscillator

The ring dye master laser (Model: Coherent 899-21) is a well-designed unidirectional-lasing ring cavity. The operation of unidirectional lasing prevents the linewidth broadening caused by the spatial hole burning. The ring cavity contains several frequency selection components. There are three passive frequency filters inside the ring dye cavity: a birefringent filer, a thin etalon, and a thick etalon. These filters altogether select a single longitudinal mode of the ring cavity. In theory, the single longitudinal mode should have bandwidth on the

order of 1 Hz or less. However, due to environmental factors (such as the cavity length fluctuation due to vibrations, the change of air refraction index, or the change of dye jet thickness) that modulate the effective cavity optical length, the actual oscillation bandwidth is on the order of tens of megahertz if only the passive optical filters are used in the cavity. To achieve a better performance, an internal active frequency-stabilization servo loop is used to stabilize the effective cavity length. The servo loop locks the laser cavity frequency to an external temperature-stabilized confocal Fabry–Perot reference cavity. An error signal is produced if the laser frequency drifts from the lock point of the reference cavity. The fast varying error signal is separated from the slower signal and coupled to a piezoelectric-transducer-driven mirror (tweeter), while the slower varying signal is directed to a scanning Brewster plate. By rotating the Brewster plate and adjusting the tweeter, the effective cavity length can be stabilized and the cavity frequency can be locked to the reference-cavity. The resulting cavity is a single-frequency laser with TEM_{00} mode and rms bandwidth of 500 kHz.

The frequency of the ring cavity can be scanned in a range of several gigahertz by continuously varying the reference-cavity optical length with a rotating galvanometer-driven Brewster plate. The ring cavity frequency will then track the reference-cavity frequency due to the feedback of the internal frequency-stabilization servo loop. For the two-frequency ratio technique, the ring dye master laser must be locked respectively to the Na D_{2a} resonance frequency f_a and the crossover frequency f_c of the Doppler-free saturation–absorption spectroscopy of a Na vapor cell. For the three-frequency technique, the ring dye laser is locked to frequency f_a at all times, and the other two wing frequencies are generated by an acousto-optic modulator. Locking to a frequency is accomplished by first tuning the reference cavity to either f_a or f_c. Then the internal frequency servo-loop drives the ring cavity frequency to track the reference cavity. Thus, the ring dye master-laser frequency can be tuned to the Na D_{2a} line frequencies f_a and f_c with very high accuracies. This is an

essential feature for the accurate temperature and wind measurements.

The dye jet used in the ring laser is highly sensitive to vibrations and the ring cavity is sensitive to the change of room temperatures, so the ring dye laser requires a well-controlled laboratory environment for operation. This is the main drawback for the applications of the Na wind/temperature lidar.

5.2.2.3.2. Doppler-free spectroscopy for precise frequency calibration and control

The experimental setup for the Doppler-free spectroscopy is illustrated in Figure 5.9. A small amount of ring dye laser beam is split into two beams by a cubic beam splitter. With the help of three mirrors, the two beams become two counter-propagating beams when going through a Na vapor cell. The Na cell is heated to and stabilized at 60°C in order to provide

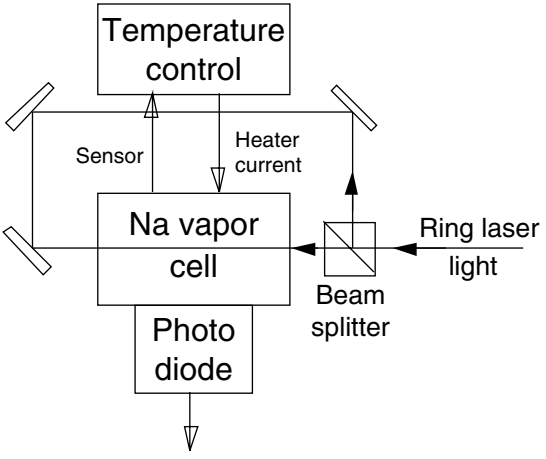


Figure 5.9 Schematic diagram of Doppler-free saturation–absorption spectroscopy experiment with a Na vapor cell. The ring laser beam enters from the right, and is split into two beams by a cubic beam splitter. With the help of three mirrors, the two beams pass through the Na vapor cell along the same line but in opposite directions. A photodiode is placed to the side of the cell to detect laser-induced fluorescence.

sufficient atomic sodium vapor density. A photodiode is placed adjacent to the Na cell to detect the laser-induced fluorescence. Under the simultaneous illumination of the two counter-propagating beams with intensities exceeding a saturation intensity ($\sim$0.4 mW/mm^2), the Na fluorescence spectrum exhibits two Doppler-free dips at the peak frequencies f_a and f_b of the Na D_{2a} and D_{2b} lines, and a Doppler-free peak at the crossover frequency f_c between D_{2a} and D_{2b}, as shown in Figure 5.3.

These Doppler-free features can be explained through a simplified three-level atomic energy system. Suppose the Na atom ground state is split into two levels ($|\,a\,\rangle$ and $|\,b\,\rangle$), and the Na atoms in these two states can be excited to state $|\,e\,\rangle$ by the photons with frequencies ν_a and ν_b, respectively. If a laser with frequency ν is incident on the Na vapor cell, the laser-induced fluorescence will only occur for those Na atoms with the correct velocity for ν to be Doppler-shifted to ν_0, where $\nu_0 = \nu_a$ or ν_b. When the laser intensity increases, the fluorescence intensity also increases but not in proportion to the intensity. The deviation from the linearity is called saturation. The saturation arises because (1) the population of the corresponding ground state decreases since the excited Na atoms can decay to another ground state (i.e., the optical pumping effect) and the population on the excited state is no longer zero, and (2) when the excited state population is not zero, the stimulated emission occurs, and it does not contribute to the fluorescence signal in the direction perpendicular to the laser beam because the stimulated emission photons are along the laser beam direction. Despite saturation, the laser-induced fluorescence spectrum follows the velocity distribution as ν is scanned. If the laser frequency $\nu < \nu_0$, the beam incident on the cell from the left interacts with atoms traveling towards the left with a radial speed v_R giving the shifted frequency $\nu' = \nu_0$, and the beam incident on the cell from the right interacts with atoms traveling towards the right with a same radial speed. As long as ν is not equal to ν_0, the two counter-propagating beams interact with two separate velocity groups of Na atoms, and the laser-induced fluorescence at a given laser intensity is simply double what it was with only one

beam. This is also true for $\nu > \nu_0$. However, for $\nu = \nu_0$, the two counter-propagating beams interact with the same group atoms with zero radial velocity ($v_R = 0$). The population on the proper ground state will be reduced dramatically since the effective laser intensity is doubled. Therefore, the laser-induced fluorescence will be significantly reduced, resulting in a Lamb dip exactly at $\nu = \nu_0$, where $\nu_0 = \nu_a$ or ν_b. The dip exhibits a nearly natural linewidth and therefore is Doppler-free. As described above, the optical pumping effect reduces the population on the corresponding ground state so the reduction in the fluorescence intensity occurs at all frequencies across both D_{2a} and D_{2b} Gaussian spectra (with Lamb dips centered at ν_a and ν_b) except at the crossover frequency ν_c. At $\nu = \nu_c$, the atoms on the $|\ a\ \rangle$ state with a correct velocity to interact with the laser beam incident from the right will be excited to the $|\ e\ \rangle$ state. They can either decay back to $|\ a\ \rangle$ to continue interacting with the same laser beam, or decay to the $|\ b\ \rangle$ state to automatically be the correct velocity group to interact with the laser beam incident from the left. Thus, the optical pumping effect does not decrease the fluorescence intensity, and a sharp (relative) enhancement peak occurs in the laser-induced fluorescence spectrum at exactly $\nu = \nu_c$. This peak is also Doppler-free with nearly a natural linewidth.

The Lamb dip and crossover peak features in sodium saturation–absorption spectroscopy were first described by Hansch et al. (1971), and that paper still remains an illuminating reference. The actual Na atoms consist of four sublevels in the excited state, separated by 15 to 60 MHz, and six allowed transitions to the two ground sublevels. For each pair of allowed transitions, there exist two Lamb dips and a crossover peak. The natural linewidth of each transition is about 10 MHz, that is, the same order as the separation of the excited sublevels. Therefore, the transitions within the D_{2a}, D_{2b}, and crossover are not resolved. She and Yu (1995) did detailed experimental and theoretical studies of the Na Doppler-free saturation–absorption spectroscopy and determined the central frequencies of these Doppler-free features to ± 0.1 MHz as illustrated in Table 5.1. These Doppler-free features provide feedback for locking the ring dye master

laser. The central frequencies of these Doppler-free features are independent of temperature, so they are an absolute frequency reference required for the wind and temperature measurements.

5.2.2.3.3. *Acousto-optic modulator for precise frequency shift*

The acousto-optic modulator (AOM) is used to generate the two wing frequencies $f_a \pm \Delta f$ required by the three-frequency ratio technique. This is a dual acousto-optic frequency shifter system. One is to shift the frequency output from the ring dye master laser (which is locked to the Lamb dip f_a) up by 630 MHz, and another is to shift it down by 630 MHz. The AOM is controlled by a computer through a trigger generator so that the operating laser frequency cycles between f_a, $f_+ = f_a + \Delta f$, and $f_- = f_a - \Delta f$. The idea of the AOM was first demonstrated by She and Yu (1994) and then employed in both the CSU and the UIUC Na lidar systems. A comprehensive review paper for the CSU Na wind/temperature lidar can be found in Arnold and She (2003). In the following, we provide an analysis of the operating principle of the acousto-optic modulator following Yariv (1997).

A piezoelectric transducer attached to the side of an acousto-optic crystal is driven by an electromagnetic field to convert the radio-frequency (RF) energy into acoustic energy. The vibration of this ultrasonic transducer produces a traveling acoustic wave with wave vector $\vec{k}_s$ to propagate across the crystal (Figure 5.10). The acoustic wave consists of a sinusoidal perturbation of the density of the AO crystal that travels at the sound velocity v_s. The variation in the density causes a change in its index of refraction because the index of refraction is higher in the compressed portions of the sound wave and lower in the rarefied region. A change in index of refraction results in partial reflection, so the acoustic wave can be represented by a series of partially reflecting plane mirrors, separated by the sound wavelength λ_s, that are moving with a wave vector $\vec{k}_s$ ($k_s = \omega_s/v_s$, where ω_s and v_s are the angular frequency and velocity of the acoustic wave, respectively). A laser beam with wave vector $\vec{k}_i$ ($k_i = \omega_i/c$, where ω_i and c are

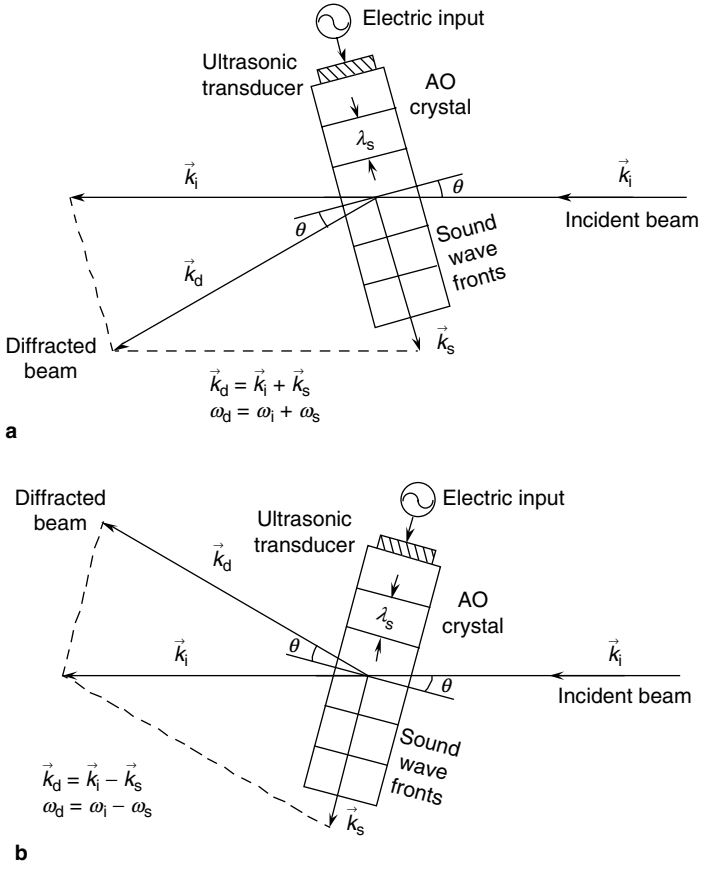


Figure 5.10 (a) The Bragg diffraction of the acousto-optic modulator (frequency up-shifting): the Bragg vector diagram and corresponding physical configuration for the diffraction of light from oncoming sound wave. (b) The Bragg diffraction of the acousto-optic modulator (frequency down-shifting): the Bragg vector diagram and corresponding physical configuration for the diffraction of light from retreating sound wave.

the laser angular frequency and speed of light, respectively) is incident on the AO crystal with a small angle θ relative to the sound wave front. According to the first-order Bragg diffraction condition (Yariv, 1988), if the incident laser beam, the acoustic wave and the angle θ satisfy the following equations:

$$k_s = 2k_i \sin\theta \tag{5.67}$$

or

$$2\lambda_s \sin\theta = \frac{\lambda_i}{n} \tag{5.68}$$

where n is the index of refraction for the laser beam, then part of the incident laser beam will be diffracted by the acoustic wave and exit the AO crystal with the same angle θ on the other side of the crystal as illustrated in Figure 5.10. The diffracted laser beam with a wave vector $\vec{k}_d$ ($k_d = \omega_d/c$, where ω_d is the laser angular frequency) will experience a Doppler frequency shift due to the moving of the acoustic wave at velocity v_s. The Doppler frequency shift between the diffracted beam and the incident beam is equal to

$$\Delta\omega = \omega_d - \omega_i = -(\vec{k}_i \cdot \vec{v}_s - \vec{k}_d \cdot \vec{v}_s) \tag{5.69}$$

Since $\vec{k}_d \cdot \vec{v}_s \approx -\vec{k}_i \cdot \vec{v}_s$, the above Doppler frequency shift can be written as

$$\Delta\omega = -2\vec{k}_i \cdot \vec{v}_s = \pm 2k_i v_s \sin\theta \tag{5.70}$$

where the "+" is for the oncoming sound wave shown in Figure 5.10(a), and the "−" is for the retreating sound wave shown in Figure 5.10(b). From Equation (5.67), it is easy to obtain that

$$\omega_s = 2k_i v_s \sin\theta \tag{5.71}$$

Therefore, the frequency shift is given by

$$\Delta\omega = \pm\omega_s \tag{5.72}$$

and the frequency of the diffracted laser beam is given by

$$\omega_d = \omega_i \pm \omega_s \tag{5.73}$$

that is, the laser frequency is shifted up or down by ω_s for the oncoming and retreating sound waves, respectively.

The AOM can also be interpreted through a particle picture of Bragg diffraction of light by sound. The acoustic wave can be represented by a stream of phonons with momentum $\hbar\vec{k}_s$ and energy of $\hbar\omega_s$, and the incident laser can be

represented by a stream of photons with momentum $\hbar\vec{k}_i$ and energy $\hbar\omega_i$. In the case of Figure 5.10(a), an incident photon and a phonon are annihilated and simultaneously create a diffracted photon with a momentum $\hbar\vec{k}_d = \hbar\vec{k}_i + \hbar\vec{k}_s$ and energy $\hbar\omega_d = \hbar\omega_i + \hbar\omega_s$. These are required by the conservation of momentum and energy, respectively. Thus, the laser frequency is up-shifted to $\omega_d = \omega_i + \omega_s$. In the case of Figure 5.10(b), an incident photon is annihilated and simultaneously creates a new phonon and a diffracted photon with the momentum and energy of $\hbar\vec{k}_d = \hbar\vec{k}_i - \hbar\vec{k}_s$ and $\hbar\omega_d = \hbar\omega_i - \hbar\omega_s$, again required by the conservation of momentum and energy. Thus, the laser frequency is down-shifted to $\omega_d = \omega_i - \omega_s$. The up- or down-shifting is determined by the relative direction between the acoustic wave and the incident laser beam. When the acoustic wave travels towards or away from the incident laser beam, the laser frequency is up- or down-shifted, respectively.

The experimental setup for the AOM used in the UIUC Na wind/temperature lidar is illustrated in Figure 5.11. Two AO shifters are driven by an RF field with frequency f_s of 315 MHz. The laser beams at three frequencies f_a, f_+, and f_- need to be injected into the pulsed dye amplifier (PDA) with the same propagation direction and polarization direction independent of the final frequency. To achieve this goal, a

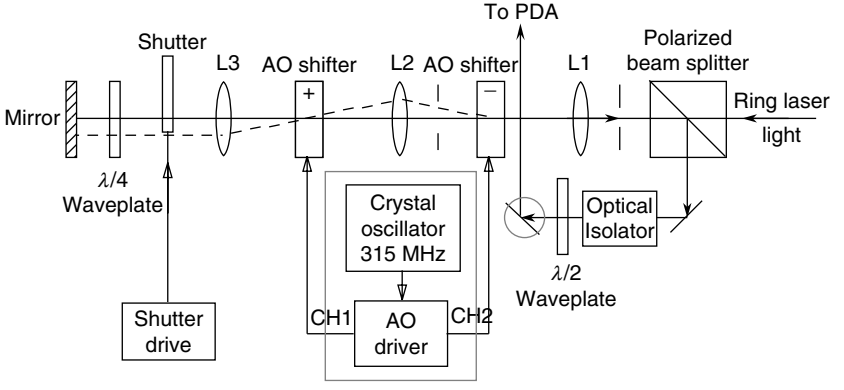


Figure 5.11 Schematic diagram of the acousto-optic modulator unit used in the UIUC Na wind/temperature lidar.

double-pass arrangement with two AO crystals separated by lenses is used. Lenses L_1, L_2, and L_3 are placed in such a way that the focal points of L_1 and L_3 are coincident with the two-time focal length points of lens L_2 on each side. Two AO crystals are placed at these two overlapping points. The incident ring laser beam from the right-hand side is horizontally polarized and passes the cubic polarizing beam splitter. It is focused by lens L_1 onto the AO shifter "−", then imaged by lens L_2 onto the AO shifter "+". Lens L_3 is used to collimate the divergent light. The mirror is used to reflect the beam back exactly along its original path. When the piezoelectric transducer is turned off, each AO crystal behaves as a dielectric medium, and the laser beam follows the solid line from the beam splitter to the reflecting mirror and reflects back through the same path from the mirror to the beam splitter. Since the beam passes through the $\lambda/4$ wave plate twice, the polarization of the reflected beam is rotated by 90°. Thus, the return beam has perpendicular polarization and will be reflected by the beam splitter for going to the PDA.

When only AO shifter "−" is activated, the acoustic wave will diffract part of the incident beam from the right by the phonon creation process as shown in Figure 5.11(b). The resulting diffracted beam will exit the AO crystal to the left following the dashed line. Along the dashed line, it is imaged by lens L_2, passes the unactivated AO crystal "+", and collimated by lens L_3. It is then reflected by the mirror back to AO shifter "−". It is diffracted by the acoustic wave again and exits from the right following the solid line to the beam splitter. Since the laser beam is diffracted by the acoustic wave twice, the return beam has a net frequency shift of $\Delta f = -2\, f_s = -630\,\text{MHz}$, that is, down-shifted. When only AO shifter "+" is activated, the incident laser beam is diffracted twice by the phonon annihilation process as shown in Figure 5.11(a), leading to a net frequency shift of $\Delta f = 2\, f_s = 630\,\text{MHz}$, that is, up-shifted. In this case, the incident beam follows the solid line except between the AO crystal "+" and the mirror where it takes the dashed line path. With sufficient RF power to drive the piezoelectric transducer, the single-pass first-order diffraction efficiency is usually 85%. Thus, the

double-pass diffraction efficiency is $(85\%)^2$, which is more than 70%. When either AO shifter "+" or AO shifter "−" is activated, the undiffracted beam (∼15%) from the first pass has unshifted frequency. We do not want it to be reflected back to the beam splitter, so a shutter is placed in the undiffracted beam path to block the unshifted beam. The shutter and the AO drives are controlled by the computer through the trigger generator box to synchronize with the pulse timing: the shutter is open to let the master laser beam pass when both AO shifters are not activated, and then the shutter is closed to block the unshifted beam when one of the AO shifters is activated. This double-pass geometry ensures the optical beam exits from the beam splitter in the same direction (perpendicular to the incident beam) independent of final frequency, and allows the use of a lower acoustic modulation frequency by a factor of 2.

5.2.2.3.4. Pulsed amplification and frequency chirp measurement

After the laser beam exits the polarized beam splitter in Figure 5.11, it passes through an optical isolator, a $\lambda/2$ wave plate, and a periscope before entering the PDA. The isolator is to prevent optical feedback from PDA in order not to disturb the ring dye master laser. The half wave plate is to rotate the laser polarization from the perpendicular to the horizontal direction. The periscope then changes it again to a perpendicular polarization, which is required by the PDA dye cells to match the Brewster angle in order to have minimum loss. The periscope is used to match the PDA beam height and to provide the adjustment for the beam to enter the PDA with precise angle and position.

Pulsed amplification of the single-mode cw ring dye laser produces laser radiation combining high peak power with a bandwidth approaching the Fourier-transform limit. This is ideal for lidar application. As shown in Figure 5.12, the PDA is a three-stage amplifier, consisting of two transversely pumped dye cells as the first two stages and one longitudinally pumped dye cell as the last stage. The cw ring laser beam enters three dye cells at the Brewster angle. Without the

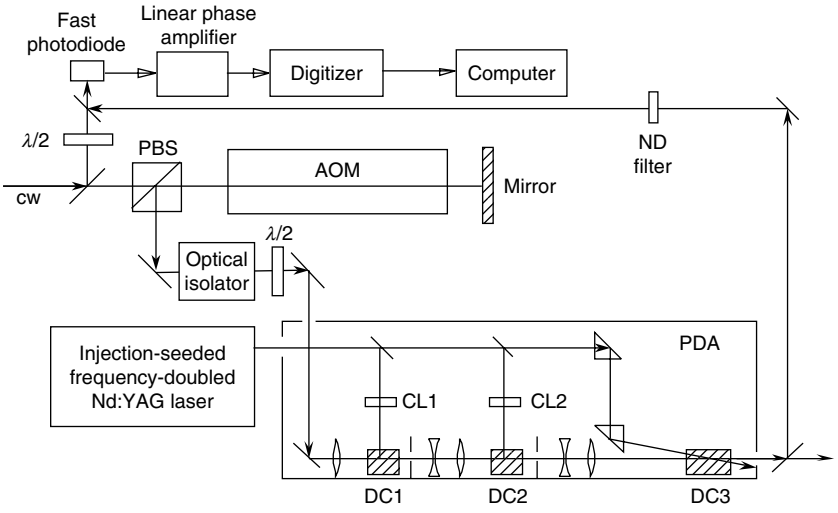


Figure 5.12 Proposed optical heterodyne measurement of the frequency shift and chirp of the pulsed dye amplifier (PDA). AOM is the acousto-optic modulator described in Figure 5.11, which is used to shift the cw beam up or down by 630 MHz. PBS is a polarized beam splitter. Two half-wave plates are used to rotate the polarizations of the cw beam and the amplified pulse to keep them match. DC1, DC2, and DC3 are three-stage dye cells of the PDA. Optical isolator is used to prevent the optical feedback from the PDA to the ring dye master laser. ND filter is a set of neutral density filters used to decrease the pulse peak amplitude in order to match the cw beam amplitude.

pump beam, the dye cells absorb most of the cw ring laser. However, a weak residual of the cw beam can still pass through the dye cells and exit the PDA. When a pump pulse from a frequency-doubled Nd:YAG laser (532 nm) is injected into the dye cells, the PDA will amplify the cw ring beam and produce a 589-nm laser pulse. Thus, the output of PDA is a quasi-pulsed laser beam with a high pulse peak and a very weak cw beam during idle period between pump laser pulses. The duration time (around 6 nsec) of PDA output pulse is determined by the pump laser. The spectrum of the PDA output is strongly influenced by the pulse shape of the pump

laser and also is limited by the short pulse duration time (Fourier transform). When the pump laser is unseeded, the pulses acquire numerous large side bands stemming from mode beating in the Nd:YAG pump laser that is partially preserved in the amplification process. Thus, the PDA output pulses have significant shot-to-shot variations. In order to generate highly reproducible pulses, the pump laser is injection-seeded by a cw seed laser working at the fundamental wavelength (1064 nm). The injection-seeding reduces the possible number of modes that can lase and makes the Nd:YAG pump laser pulses to have nearly pure Gaussian shape with stable width and height. The resulting PDA output spectrum is nearly Fourier transform limited and highly reproducible. The rms bandwidth of the PDA output pulse is approximately 60 MHz, which corresponds to a full width at half maximum (FWHM) of 140 MHz.

For an ideal PDA, the output pulse would be a Fourier-transform-limited Gaussian lineshape centered precisely on the frequency of the cw input beam. However, the actual spectrum of the PDA output not only has a broadened linewidth (larger than the transform-limited) but also has a shift of the central frequency. These effects are mainly caused by three factors: the amplified spontaneous emission (ASE), unseeded Nd:YAG laser pulses, and nonlinear effects associated with the pulsed amplification process. The ASE is present even in the complete absence of the cw input beam. It is the result of spontaneous emission when the dye is excited by the pump laser. If it is amplified by successive pumped dye cells, it will add a broad spectral pedestal under the normal pulse lineshape, increasing the overall baseline. Fortunately, with careful alignment of the PDA to eliminate reflections along the beam path, ASE is typically reduced to a few percent of the output power. Thus, its effect is negligible. Unseeded or improperly seeded Nd:YAG pulses would produce a strong modulation of the PDA output with a much broader linewidth. Modern Nd:YAG laser seeding technique is well developed, and the seeding status can be monitored through Q-Switch build-up time or by Fabry–Perot etalon fringes.

Thus, well-seeded performance of the Nd:YAG laser can be assured.

The nonlinear effects associated with the pulsed amplification process are not completely identified, but the main factors are the optical phase perturbations during pulsed amplification, which can both broaden and shift the PDA output frequency spectrum (Eyler et al., 1997). The possible mechanisms that can cause the optical phase distortion include heating of the dye solvent, the intensity dependence of the refractive index, and the time dependence of the gain. The experimental and theoretical studies made by Gangopadhyay et al. (1994) and Melikechi et al. (1994) suggested that the principal source of the phase shifts is the time-varying susceptibility of the dye solution, which is a consequence of the changing excited-state population, that is, the time-varying gain. The phase distortion causes the instantaneous frequency of PDA output to vary with time (which is called the *frequency chirp*), resulting in broadened linewidth and shifted central frequency of the laser pulses. If the frequency shift cannot be measured precisely, it will cause a systematic error in wind and temperature measurements, especially in the wind measurement that is more sensitive to an absolute frequency shift.

There are two approaches to measure the frequency shift and linewidth broadening. One is to utilize the optical heterodyne technique to measure the shift and broadening directly. Another is to use an iodine vapor cell to convert the frequency information to intensity ratio signals. The optical heterodyne technique was developed by Fee et al. (1992), and applied by Gangopadhyay et al. (1994), Melikechi et al. (1994), and Eyler et al. (1997) to measure the frequency chirp and shift in the pulsed amplification of a cw single-mode laser. The basic idea is to use an AOM to shift the frequency of part of the cw beam over a range larger than a Fourier-transform-limited linewidth. This shifted beam provides a reference for heterodyne mixing. The rest of the cw beam is pulse-amplified, and then mixed with the frequency-shifted sample of the cw laser on a fast photodiode. The beat signal is then processed in the Fourier domain to reveal the phase and the instantaneous

frequency of the pulse laser. Since an AOM has been used in the Na wind/temperature lidar, the following schematic (Figure 5.12) can make the optical heterodyne measurement of the PDA frequency chirp and shift. Before entering the AOM, a small amount of the cw beam is split to the fast photodiode as a reference. The rest of the cw beam is frequency-shifted by the AOM and pulse-amplified by the PDA. A small amount of PDA output pulses is attenuated and sent collinear with the reference beam onto a broadband fast photodiode detector. The output of the detector amplifier is digitized by a fast digital oscilloscope (with high sampling rate and wide bandwidth) and the resulting digitized signal is transferred to a computer for processing. The main frequency of the beat signal is equal to the sum of the AOM frequency shift (630 MHz) and the PDA frequency shift (on the order of 10 MHz). Both the PDA frequency shift and the pulse bandwidth can be derived from the beat signals since the bandwidth of the cw reference beam is very narrow compared with the pulse bandwidth.

Another method to measure the PDA frequency shift and chirp was developed by CSU group using an iodine vapor cell with edge technique (White, 1999). The molecular iodine has strong absorption bands in the vicinity of the Na D_{2a} line: one is about 1 GHz higher and another is about 1.5 GHz lower than the Na D_{2a} Doppler-free peak in frequency. Thus, these two absorption lines form an iodine transmission peak in between them. Plotted in Figure 5.13 is an example of this iodine transmission peak with respect to the three operating frequencies (f_a, f_+, and f_-) of the Na Doppler lidar when the iodine cell is heated to 80°C. At this temperature, the FWHM of this iodine transmission peak is about 1.1 GHz. It is fortunate that the Na f_a and f_- operating frequencies locate on the two edges of the iodine transmission peak, so the ratio of the transmission signals at these two frequencies is sensitive to the frequency shift in the transmitter. For example, if the transmitter has a positive frequency shift, the transmission at f_- should increase while the transmission at f_a should decrease. Since the two wing frequencies are generated by the AOM, the separation between the operating frequency is

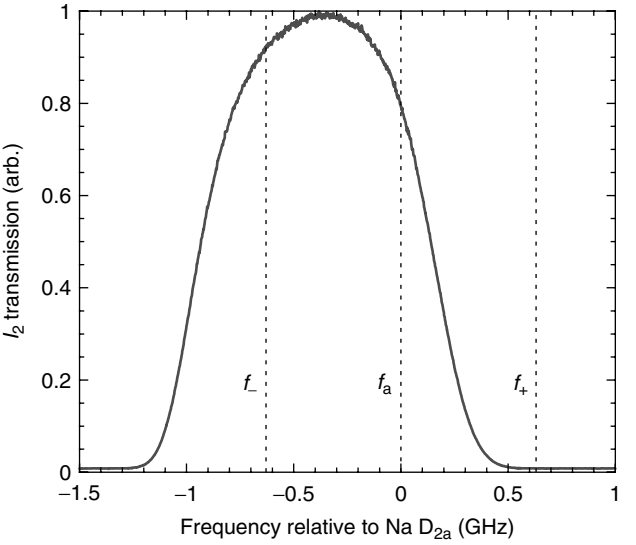


Figure 5.13 An iodine transmission line with respect to the three operating frequencies of the Na Doppler lidar. The spectrum was acquired using a 4-in. iodine cell at 80°C with cw laser light. The FWHM is about 1.1 GHz. (Courtesy of T. Yuan and C.-Y. She, Colorado State University, Fort Collins, Colorado. With permission.)

assumed to be constant regardless of any frequency shift imparted by the pulsed dye amplifier. Thus, the frequency shift can be derived from the intensity ratio, whose principle is similar to the wind measurement using the Na Doppler lidar described earlier. Since the location of the iodine peak is known, given a stable iodine temperature and other parameters, the frequency shift is measured in absolute frequency. Such technique has been employed successfully by the CSU group in the Na wind/temperature lidar to monitor the PDA frequency shift and to calibrate the wind measurements.

5.2.2.3.5. Receiving and photon counting

As illustrated in Figure 5.8, the Na lidar receiver consists of a receiving optical telescope, a field-stop, a chopper, collimating lenses, narrowband filters (interference filter), and a PMT working in photon-counting mode. The field-stop is used

to limit the field of view in order to reject most of the sky background. The mechanical chopper is used to block the extremely strong scattering from the low altitudes (~0–15 km) to prevent the PMT being saturated. An alternative to the mechanical chopper is an electronic gain switch to vary the voltage between the PMT cathode and the first dynode to control the PMT gain in order to avoid the PMT saturation. The interference filter is used to reject the sky background while transmitting the signal photons, which is enough for the nighttime measurement. During daytime operation, an ultra-narrowband filter (e.g., a Faraday optical filter or a Fabry–Perot etalon) must be used to reject more sky background. The collected and filtered photons are then transmitted to detectors. The CSU system employs a fiber for transmitting the photons to the detector. This makes the receiver more flexible. Since the return signals are extremely weak, the highly sensitive PMT is used to count the photon number. The data acquisition and control system is a computer-controlled electronic system to time-gate the signals, record the photon counts, and save frequency calibration data.

5.2.2.4. Daytime Measurements

It is highly desirable to make the temperature and wind measurements in the mesopause region through the complete 24-h diurnal cycle since a large number of geophysical phenomena and long period changes of the Earth's atmosphere can be explored only if the instrument allows daylight operation with high quality data. Narrowband optical filters are necessary in daytime lidar measurements to reject the broadband background photons while passing the narrowband signals. Conventional interference filters cannot provide extremely narrow bandwidth (0.01 to 1 Å) with a highly stable transmission peak and a reasonably large field of view. Daytime operation of a broadband Na lidar was realized by using a Fabry–Perot etalon as a narrowband filter in the receiver (Granier and Megie, 1982; Kwon et al., 1987). However, only Na density was measured with such a system. In 1995, the CSU group introduced a FADOF into the narrowband Na lidar

receiver to reject the sky background, and realized the daytime measurements of Na temperature in the mesopause region (Chen et al., 1996). Later, the daytime measurements of wind were also achieved with such lidar system (She et al., 2003). The UIUC group employed an ultra-narrowband Fabry–Perot etalon to help reject the skylight and also achieved the daytime observations of the Na temperature (Yu et al., 1997). Both the CSU and the UIUC groups reported a large number of new scientific results from these 24-h full diurnal cycle measurements, which will be summarized in Section 5.3. Here we describe the instrumentation aspects.

5.2.2.4.1. Faraday anomalous dispersion optical filter (FADOF)

FADOF is an atomic resonance filter based on the resonant Faraday effect. A Na vapor cell is placed within a permanent magnetic field between two crossed polarizers (Figure 5.14). The magnetic field is oriented parallel to the optical axis. As explained in the following text, for light on Na resonance that was transmitted through the first polarizer, the Na vapor rotates the plane of light polarization by 90° so that it can pass through the second polarizer without loss. The refraction index n of a dilute Na atomic vapor at frequencies close to its resonance absorption lines can be expressed as a complex refraction index:

$$n = \sqrt{1+\chi} \cong 1 + \frac{1}{2}\chi = 1 + \frac{1}{2}\chi' - \mathrm{i}\frac{1}{2}\chi'' \tag{5.74}$$

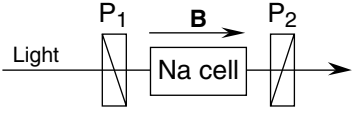


Figure 5.14 Schematic of the Faraday anomalous dispersion optical filter (FADOF). P_1 and P_2 are two crossed polarizers. A Na vapor cell between P_1 and P_2 is housed in permanent magnets such that a strong axial magnetic field $\vec{B}$ (along the direction of the light propagation) permeates the Na cell.

where χ is the electric susceptibility of the Na vapor, $\chi = N\alpha$, N is the number density of the Na vapor, and α is the atomic polarizability. In the vicinity of an atomic resonance line, the electric susceptibility is given by $\chi = \chi' - i\chi''$, where

$$\chi' = \frac{Ne^2 f}{2m\omega\epsilon_0} \frac{\omega_0 - \omega}{(\omega_0 - \omega)^2 + (\gamma/2)^2} \tag{5.75}$$

$$\chi'' = \frac{Ne^2 f}{2m\omega\epsilon_0} \frac{\gamma/2}{(\omega_0 - \omega)^2 + (\gamma/2)^2} \tag{5.76}$$

Here, e is the electron charge, m is the electron mass, f is a dimensionless oscillator strength, ε_0 is the dielectric permittivity of vacuum, ω_0 is the Na atomic resonance frequency, ω is the light frequency, and γ is the transition probability of Na excited states, that is, $\gamma = A_{21} = 1/\tau_R$ (τ_R is the spontaneous lifetime of the Na excited states). Here χ' represents the dispersion near the resonance absorption lines, whereas χ'' represents the resonance absorption. When a linear polarized light with frequency ω close to the Na resonance frequency ω_0 propagates along the magnetic field $\vec{B}$ through the Na vapor, the linear polarization may be decomposed into a superposition of right- and left-circular polarized light. The static magnetic field $\vec{B}$ across the atomic vapor cell causes the Zeeman splitting of Na energy levels, which results in different Na resonance frequencies for right- and left-circular polarized light: $\omega_0^{\pm} = \omega_0 \pm \Delta\omega$, where $\Delta\omega$ is the Zeeman splitting. Therefore, according to Equation (5.74), the right- and left-circular polarized light will experience different indices of refraction at frequencies near ω_0 (the Na resonance frequency in absence of the magnetic field) as illustrated in Figure 5.15. The difference in the index of refraction will result in different propagating velocities and introduce phase shift between the two circular polarizations. The phase shift is given by

$$\Delta\varphi = 2\pi \frac{l\Delta n}{\lambda} \tag{5.77}$$

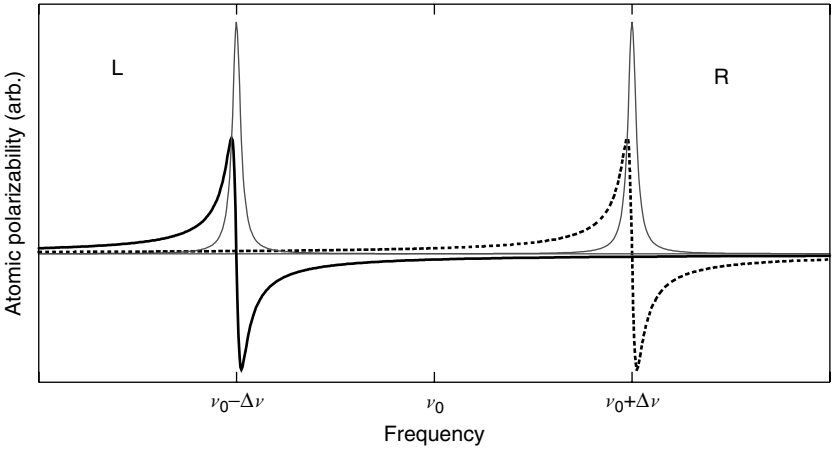


Figure 5.15 Theoretical atomic polarizability (thicker lines) and Lorentzian absorption (thin lines) curves of Na vapor in the presence of an axial magnetic field. Zeeman effect causes the right circular light has a higher resonance frequency while the left circular light has a lower resonance frequency. At ν_0 (the resonance frequency in absence of the magnetic field), there is no absorption for either component of the light, but there is a difference in polarizability (thus index of refraction). In the actual Faraday Na cell, all these features are thermally broadened. The magnetic field is set so that the Zeeman splitting is larger than the thermally broadened linewidth.

where l is the length of the Na cell, λ is the light wavelength, and Δn is the difference of the refraction index between the two circular polarizations. The resulting light will have a linear polarization rotated by certain angle from its initial linear polarization. This dispersive polarization rotation phenomenon is called the *resonant Faraday effect*. The degree of polarization rotation depends on the light frequency, the magnetic field strength, the vapor cell temperature (i.e., the Na density), and the vapor cell length. For a given frequency and cell length, the magnetic field and the temperature of the Na cell are adjusted so that the phase shift $\Delta\varphi = \pi$, corresponding to a rotation of polarization of $\pi/2$. The resulting light will pass the second polarizer without attenuation and result in

nearly 100% transmission at the Na resonance frequency. Meanwhile, the polarization rotation for frequencies a few linewidth from the FADOF transmission peaks is vanishingly small. This results in the out-of-band rejection of the filter on the order of 10^{-5} determined by the extinction of polarizers. More detailed theory behind the FADOF can be found in Yeh (1982), Dick and Shay (1991), and Yin and Shay (1991).

The original Na FADOF consisted of a Na vapor cell of 0.76 cm long at 189°C in a 1750 G magnetic field (Chen et al., 1993). However, the Na vapor cell was easily degraded by the high temperature. Chen et al. (1996) improved the Na FADOF by using a 2.54-cm long Na vapor cell in a magnetic field of 1800 G in an oven operated at a temperature of 168°C. Plotted in Figure 5.16 is a measured transmission function of a Na FADOF operated at 168°C with 1850 G of magnetic field (Arnold and She, 2003). The peak transmission at the Na line center is 86%, and the FWHM of the peak is 1.9 GHz. One of the advantages for the Na FADOF is that the FADOF

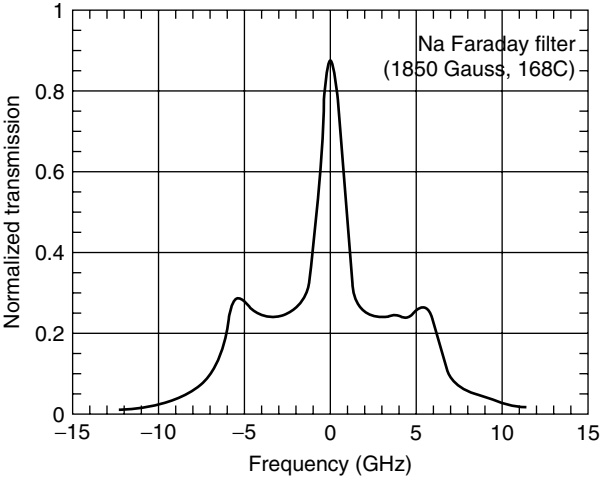


Figure 5.16 Normalized transmission curve of a CSU Na lidar Faraday filter. (From Arnold, K.S., and She, C.Y., *Contemp. Phys.*, 44, 35–49, 2003. With permission.)

does not need highly collimated light, and therefore, has large field of view. The combination of the ultra-narrow Na FADOF and the reduced beam divergence and receiver field of view makes it possible to reduce the detected daytime sky background to a level comparable with the typical values encountered at night, leading to a peak Na signal-to-background ratio of 15 at high noon (Chen et al., 1996). This signal-to-noise ratio enables the daytime temperature and wind measurements in the mesopause region.

Because of the hyperfine structure and the magnetic Zeeman splitting of the Na energy levels, an excited Na atom can decay via different energy paths and thus reradiate at a different frequency than that of the absorbed laser photon. With two ground states, four upper states and their magnetic substates, the selection rules from quantum mechanics allow many absorption/reemission possibilities, each having characteristic emission frequencies. All these transitions will be excited with different probabilities depending on the laser frequency, so the returned fluorescence is spectrally dispersed. This dispersion in returned fluorescence frequency is not important for wide bandwidth receivers. For nighttime configuration, when a broadband filter is used in the receiver, the filter transmission function is constant, and the receiving efficiency is the same over the spectral range of the Na D_2 Doppler lines. The nighttime filters with a bandwidth of hundreds of gigahertz pass all the dispersed light with the same transmission value. Therefore, the filter transmission function is canceled when taking the ratios (R_T and R_W) between fluorescence intensity signals at different frequencies. According to Equations (5.51) and (5.52), the ratios R_T and R_W are thus determined by the ratio of effective cross sections. However, for daytime configuration, the bandwidth of the Na Faraday filter is so narrow that the spectrally dispersed fluorescence will have different transmissions through the Faraday filter. For different laser exciting frequencies, the returned fluorescence will contain different portions of the 10 emission components, resulting in different receiving efficiencies. Therefore, the filter transmission function must be taken into account when calculating the ratios R_T and R_W. As discussed in Section

5.2.2.2, the Doppler frequency shift of the fluorescence should also be considered when computing the filter transmission function. By using the frequency-dependent scattering cross sections and taking into account the Hanle effect as well as the transmission function of the Faraday filter, one can deduce the calibration curve between the intensity ratio and temperature for the FADOF daytime measurements. Shown in Figure 5.17 are the calibration curves with and without FADOF for CSU Na wind/temperature lidar system taken from Chen et al. (1996). Notice that the intensity ratio for the FADOF channel at the same temperature is considerably smaller than that for the nighttime configuration. This is expected since FADOF transmission at the frequency f_c or $f_\pm$ is much lower than at the peak frequency f_a. The temperature of the mesopause region can be derived from the measured intensity ratio by using these calibration curves. The difference between the temperature determined from the nighttime configuration and daytime

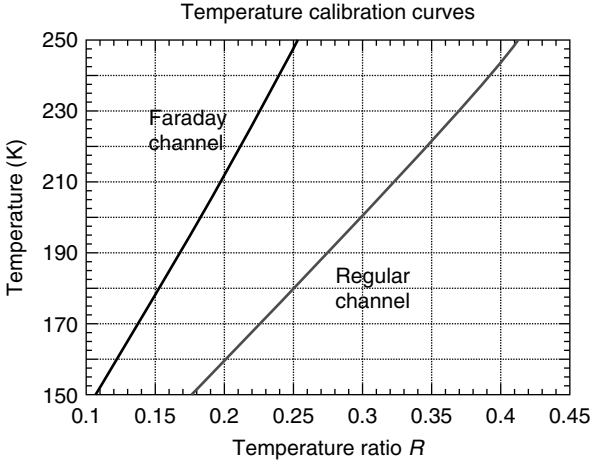


Figure 5.17 Calibration curves for the conversion of measured intensity ratio to temperature in the mesopause region for the regular channel (without Faraday filter) and the Faraday channel (with Faraday filter). (From Chen, H., White, M.A., Krueger, D.A., and She, C.Y., *Opt. Lett.*, 21, 1093–1095, 1996. With permission.)

configuration is within the experimental error (Chen et al., 1996).

5.2.2.4.2. Fabry–Perot etalon for daytime operation

The UIUC Na wind/temperature lidar can also make daytime observations (Yu et al., 1997). To reduce background noise from the bright daytime sky, the laser beam divergence and telescope field of view were reduced to 0.5 mrad. Further smaller divergence angles would cause serious saturation effects, which result in intolerable measurement distortion. A temperature stabilized, narrowband interference filter and a temperature-stabilized pressure-tuned Fabry–Perot etalon were employed in the daytime receiving telescope to reduce the optical bandwidth to 10 GHz (FWHM). These narrowband filters help reject the sky background. To optimize performance, different operating configurations and scenarios were used for day and night measurements. During the daytime, a 0.35-m diameter telescope was employed and only Na density and temperature profiles were measured. The major error source was signal photon noise and background noise. When the daytime data are averaged over a vertical range of 1 km and temporally for 30 min, the rms temperature errors between 85 and 100 km vary between 1.7 and 4.6 K. The smallest errors are at the peak of the Na layer near 92.5 km where the signal is strongest. Like the FADOF, the transmission function of the etalon is not constant over the spectral range. In order to derive accurate temperature, it is also necessary to derive the calibration curve after considering the transmission function. This requires precise knowledge of the etalon transmission function, including the central frequency, bandwidth, and the lineshape. However, the center of the etalon transmission band is not known accurately. There is no way to analytically correct for this transmission function and optical pumping effects. This can be solved only by an examination of the calculated temperatures near the night/day transition. Any temperature jump between the two configurations is associated with the combined effects of optical pumping and the etalon transmission band curve. From the

observed bias, the etalon central frequency is calculated and the data are reprocessed considering this central frequency.

5.2.2.5. Lidar Data and Error Analysis

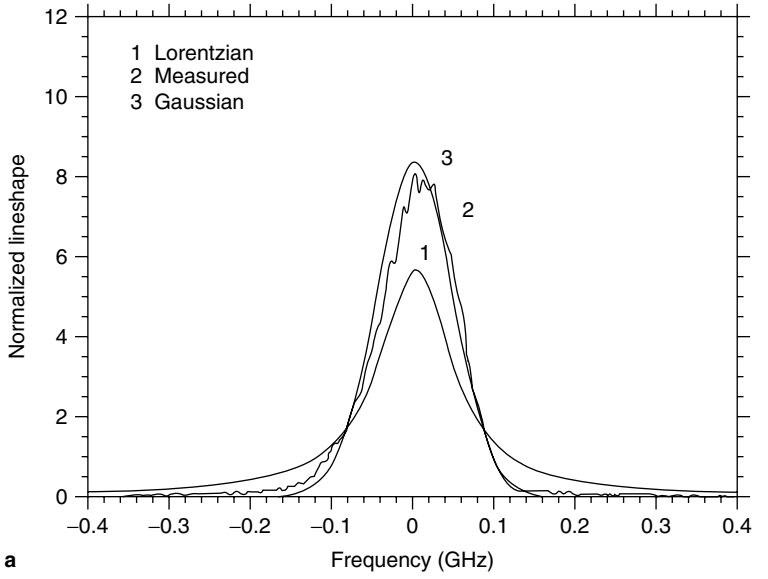
5.2.2.5.1. Lidar data analysis

Here we briefly describe the lidar data analysis on the aspects of how to derive the atmosphere temperature, wind, and density information from the lidar photon count profiles. A standard preprocess of the raw photon counts includes the correction of the PMT and discriminator saturation effects, the correction of the chopper function or PMT gain effect, and the subtraction of the background photon counts. After the preprocess, the lidar equations for the Na resonance fluorescence and Rayleigh signals can be expressed as Equations (5.44) and (5.45), respectively. Notice that the background term no longer presents in these two equations as compared with the lidar equation (5.1). The Na photon counts are then normalized to the Rayleigh photon count at 35 km, which is obtained by averaging the Rayleigh signals between 30 and 40 km. This is defined as the normalized Na photon count given by Equation (5.48). With the Rayleigh term in the denominator, the time-dependent laser power and atmospheric transmittance can be canceled because the Rayleigh and Na photons resulted from the same laser pulses at nearly the same time. Further assumptions are made that the receiver has the same geometric factors for the Na signals and Rayleigh signal at Z_R. The receiver also has broadband filters so that the Na fluorescence and Rayleigh scattering photons experience the same filter transmission even if their frequencies are slightly different. Thus, these two factors can be canceled out and the normalized Na photon counts are then given by Equation (5.49).

To derive the atmosphere temperature and wind, the ratio of the normalized Na photon counts is calculated. For the two-frequency technique, the ratio for temperature R_T is calculated using the crossover frequency f_c and the D_{2a} peak frequency f_a, as shown in Equation (5.50). The frequency difference between f_c and f_a is very small, so the Rayleigh

scattering cross sections at these two frequencies can be regarded as equal. Thus, the Rayleigh cross-section and atmosphere number density at Z_R can also be eliminated and the second equality in Equation (5.50) can be achieved. However, the Na densities at different times t_1 and t_2 are still involved in this equality. For the two-frequency technique, the laser is tuned between f_a and f_c. Approximately 1 min is required for tuning and 1 min for data acquisition at each frequency. At this interval length (1 min), the Na density fluctuations can bias the temperature measurement. To minimize this bias, the photon count profiles are weighted and averaged before taking the ratio R_T using the method described in She et al. (1990): four (three) consecutive photon count profiles obtained with the laser tuned to f_c (f_a) are added using weighting factors of 1/15, 1, 1, 1/15 (3/10, 1, 3/10) to form a single f_c (f_a) profile which corresponds to an equivalent data integration period of ~5 min. The weighting factors were chosen to minimize the effects of Na and atmospheric density variations. By using these weighting factors, only sixth- and higher-order temporal variations of the density influence the final temperature profiles. Detailed discussion about this average can be found in Appendix C of Papen et al. (1995a). Therefore, canceling the Na number density in Equation (5.50) is a good approximation. Now the ratio R_T is simply proportional to the ratio of the effective scattering cross sections at these two frequencies, shown as the third equality in Equation (5.50).

Since the lineshape of effective cross-section is a convolution of the atomic Doppler broadened lineshape and the laser lineshape, the ratio R_T will be influenced by the lineshape of the laser. Shown in Figure 5.18(a) is a comparison of three laser lineshape functions: a measured lineshape compared to the Lorentzian and Gaussian lineshapes. Using different laser lineshapes, the calculated theoretical ratios of R_T at the same temperatures are different, as plotted in Figure 5.18(b). The actual laser lineshape is very close to the Gaussian shape, and it is highly reproducible. The curve for measured laser lineshape in Figure 5.18(b) is served as a calibration curve, that is, when R_T is calculated from the



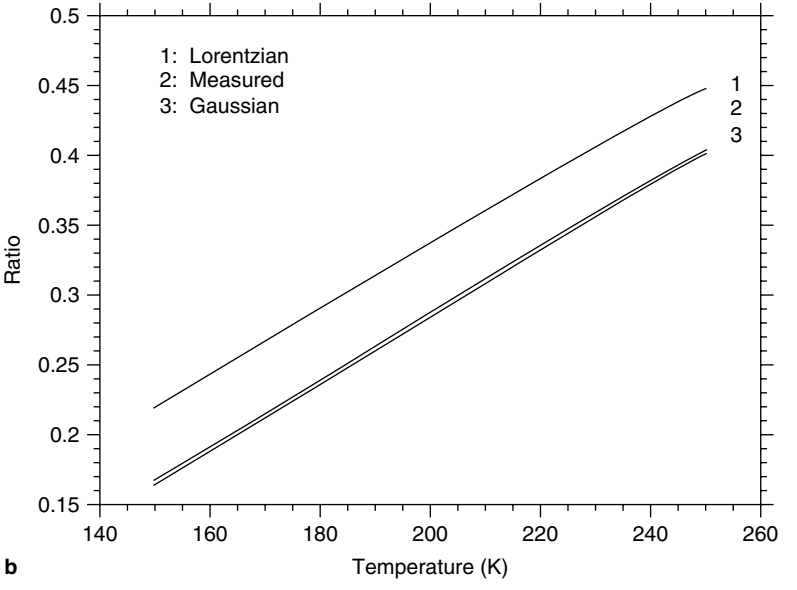


Figure 5.18 (a) Three laser lineshape functions for the Na wind/temperature lidar: a measured lineshape along with two hypothetical line shapes with equal FWHM = 112 MHz. (b) The temperature calibration lines of these lineshapes. (From She, C.Y., Yu, J.R., Latifi, H., and Bills, R.E., *Appl. Opt.*, 31, 2095–2106, 1992. With permission.)

observational data of photon counts, the temperature can be then inferred from the curve accordingly. For the two-frequency technique, there is no wind information available, so the radial wind v_R is set to 0 m/s during the determination of temperatures. When the lidar beam points to zenith, the radial wind is usually less than 3 m/s so that the resulting temperature errors are negligible (<0.5 K).

For the three-frequency technique, the ring dye laser is locked to the Na Doppler-free feature at the D_{2a} peak f_a throughout the observations, and an AOM is used to shift the laser frequency by ± 630 MHz. The experiment scenario is that the outgoing lidar beam switches its frequency from the peak frequency f_a to the plus wing frequency f_+ then to the minus wing frequency f_-, and then repeats this sequence again. The data are accumulated at three frequencies accordingly. The frequency switching is very fast, either every single pulse (as in the CSU system) or every 50 pulses (as in the UIUC system), so the data at three frequencies are collected with very short intervals, such as 1/50 or 1 s. This interleaving of observations eliminates any biases due to Na density fluctuations. Therefore, when taking the ratio for temperature R_T and the ratio for wind R_W, the Na density terms can be safely canceled so these two ratios are given by Equations (5.51) and (5.52), respectively.

The radial wind velocities are determined by the measurement of the Doppler shift of the Na atoms along the laser beam path. A positive wind component moving along the laser beam, v_R, shifts the absorption spectrum of the atmospheric Na to higher frequency by v_R/λ, where λ is the laser wavelength. The Doppler frequency shift can be reflected from the change of the return signal ratio $R_W = N(f_-)/N(f_+)$. As the velocity increases ($v_R > 0$), the ratio of returns R_W decreases as shown in Figure 5.2(b). The maximum variation in R_W occurs for measurements taken at f_+ and f_- on each side of the D_{2a} resonance. Ignoring the contribution of the D_{2b} feature and using the analytic expression for the absorption cross-section, we find that the optimal frequencies for highest sensitivity of R_W to the wind variations are $f_\pm = f_a \pm \sqrt{2}\sigma_D$ (She et al., 1992; Papen et al., 1995a). For the atmospheric Na

atoms, we choose $f_{\pm} = f_a \pm 630\,\text{MHz}$. The actual ratio R_W used for the UIUC Na wind/temperature lidar is more complicated than the simple ratio given by Equation (5.52). The ratio used is

$$R_{W2} = \frac{\ln\left(\frac{N_{norm}(f_-,t_2)}{N_{norm}(f_+,t_1)}\right)}{\ln\left(\frac{N_{norm}(f_-,t_2)N_{norm}(f_+,t_1)}{N^2_{norm}(f_a,t_3)}\right)} = \frac{\ln\left(\frac{\sigma_{effNa}(f_-,T,v_R)}{\sigma_{effNa}(f_+,T,v_R)}\right)}{\ln\left(\frac{\sigma_{effNa}(f_-,T,v_R)\sigma_{effNa}(f_+,T,v_R)}{\sigma^2_{effNa}(f_a,T,v_R)}\right)} \tag{5.78}$$

The fast switching between frequencies by the AOM allows the cancelation of the densities. This ratio is chosen to minimize the uncertainty in the measurement.

Sodium densities can also be calculated from the observations. Only one frequency at the D_{2a} peak frequency f_a is necessary, but a weighted average of all three frequencies is used to obtain a better estimate. The weighting is chosen to minimize the sensitivity of the density measurement to changes in the temperature and wind. The weighted cross-section is defined as

$$\sigma_{eff\_wgt} = \sigma_a + \alpha\sigma_+ + \beta\sigma_- \tag{5.79}$$

where σ_f is the effective cross-section at frequency f. The weighting factors, α and β, are chosen so that

$$\frac{\partial\sigma_{eff\_wgt}}{\partial T} = 0; \quad \frac{\partial\sigma_{eff\_wgt}}{\partial v_R} = 0 \tag{5.80}$$

After the weights are determined, the Na density is calculated by

$$n_{Na}(z) = 4\pi n_R(z_R)\sigma_R \frac{z^2}{z_R^2} \times \frac{N_{norm}(f_a,z) + \alpha N_{norm}(f_+,z) + \beta N_{norm}(f_-,z)}{\sigma_a + \alpha\sigma_+ + \beta\sigma_-} \tag{5.81}$$

Once the ratios R_T and R_W are calculated from the observed photon counts, the temperature and radial wind can be inferred from the calibration curves similar to the ones shown in Figure 5.7(a) and (b), respectively. Since R_T and R_W depend on both temperature and wind, the solution for temperature

and wind is iterative: by assuming 0 m/s wind, we calculate the temperature using R_T, and then calculate the wind using the calculated temperature and the ratio R_W. We then recalculate the temperature again using the computed wind. This is repeated until the values converge. After the temperature and wind are determined, the Na density can then be calculated using Equations (5.79)–(5.81).

5.2.2.5.2. Lidar error analysis

There are several sources of measurement errors. They can be classified into two categories: systematic errors and random errors. Systematic errors are mainly caused by our imprecise information of the Na atom absorption cross-section $\sigma_{abs}(\nu)$, the laser absolute frequency calibration, the laser lineshape $g_L(\nu)$, the receiver filter function $\eta_R(\nu)$, and the geometric factor $G(z)$. Random errors are mainly caused by the photon noise, the detector shot noise, Na density fluctuations, and the random laser and electronic jitter. These system characterization errors and photon-counting errors will propagate and contribute to the errors of temperature, wind, and Na density measurements. Systematic errors determine the measurement accuracy (bias), and random errors determine the measurement precision (uncertainty).

A. Systematic errors

a. Determination of $\sigma_{abs}(\nu)$: The accurate determination of $\sigma_{abs}(\nu)$ is mainly obstructed by the Hanle effect, the Na layer saturation, and the optical pumping effect. The $\sigma_{abs}(\nu)$ given in Equation (5.41) and the values of relative line strength A_n given in Table 5.1 are true for the case when Zeeman splitting of hyperfine structures (hfs) of Na atoms, saturation, and optical pumping effects are not present. However, in the geomagnetic field in the mesopause region, the Na hyperfine energy levels exhibit Zeeman splitting. The resulted frequency shift of the Na transition lines does not exceed 1.4 MHz, which is less than 0.3% of the hfs splitting and may be ignored in the temperature conversion procedure. Nevertheless, the existence of the geomagnetic field will cause the Hanle effect through the Zeeman splitting and the Hanle

effect will modify the relative line strength A_n. Here, A_n is not the strength for a single transition line, but defined as a sum over the corresponding Zeeman transitions. This new relative line strength of Na atoms caused by the Hanle effect can be calculated using the equations derived by Fricke and von Zahn (1985) and Papen et al. (1995a) once the laser polarization, the strength, declination and inclination angles of the geomagnetic field are known. A_n shown in Table 5.1 are the spatially integrated values (5, 5, 2, 14, 5, 1). For the Urbana Atmospheric Observatory (UAO), the Hanle-effect-modified A_n values are (5, 5.48, 2, 15.64, 5, 0.98). The relative line strengths must be recomputed for each lidar site, beam direction, and beam polarization because of the large variations in the magnetic-field direction around the globe. If the Hanle effect is ignored, the temperature derived by using spatially averaged line strength will result in a bias of 1.4 K (Papen et al., 1995a).

For large laser pulse intensities, the intensity dependence of $\sigma_{\text{abs}}(\nu)$ becomes important because single atoms no longer respond in the linear fashion that is assumed for the calculation of the transition strengths. Saturation and optical pumping effects must be included to determine A_n for use in Equation (5.41). As described in Section 5.2.1.3, the saturation of Na density measurements results from the finite lifetime of the excited state. Once in the excited state, the Na atoms cannot absorb another photon before relaxing back to the ground state, and may be involved in stimulated emission further reducing the backscattered signal. If the laser linewidth is narrow enough that the laser excites Na atoms only from one individual ground state, for example, the $F=2$ state, optical pumping occurs if the excited atoms decay to the $F=1$ ground state. Optical pumping reduces the available Na density on $F=2$ state for absorbing the incoming laser photons, and effectively reduces the absorption cross-section. Laser pulse lengths that are shorter than the lifetimes of the excited states do not suffer from optical pumping effects because no significant shift in the ground-state population can occur. As the laser pulse length increases, the error induced by optical pumping becomes more significant. For modern Na wind/temperature lidar, due to the short Nd:YAG laser pulse

duration, the lidar outgoing pulse has a length around 6 to 7 nsec. It is much shorter than the Na 16.23 nsec lifetime. So, the optical pumping effect is not significant and will be excluded in the analysis.

The absolute saturation level directly affects the accuracy of the Na density measurement. However, the temperature and wind measurements are influenced by the relative saturation effect. In other words, if the Na layer saturation levels at different frequencies are the same, the saturation will not affect the temperature and wind measurements since the ratio of effective cross sections remains the same as nonsaturated. Unfortunately, the saturation at frequency near the absorption peak is higher than at wing frequencies or the crossover frequency. When the laser intensity increases, the ratio R_T for both two- and three-frequency techniques will be different from the nonsaturated ratio. The saturated R_T will always be higher than nonsaturated R_T, and thus the saturation will result in higher measured temperature. The effect of saturation on temperature measurement is not as sensitive as the effect on Na density measurement. Saturation effects have been considered by several groups. Megie et al. (1978) made the first quantitative analysis using a rate-equation model. Their method and results were very comprehensive and founded the basis for the Na saturation study, although they are more suitable for the broadband laser excitation instead of the narrowband laser excitation. Welsh and Gardner (1989), von der Gathen (1991), and Papen et al. (1995a) adapted Megie et al.'s (1978) rate-equation methods and applied them to the narrowband laser cases. Milonni et al. (1998) made further detailed analysis using density matrix simulations for laser pulse excitations with long, short, and intermediate duration time. Here, we give the results from von der Gathen (1991) in Figure 5.19: the saturation at maximum, minimum, and temperature correction in relation to the energy area density. The curves were computed under the conditions of pulse length 10 nsec, laser linewidth 130 MHz, atmospheric transmission 0.7, and Na temperatures 150, 200, and 250 K. Normal operation conditions of Na wind/temperature lidar have the laser pulse energy of 40 mJ, the laser

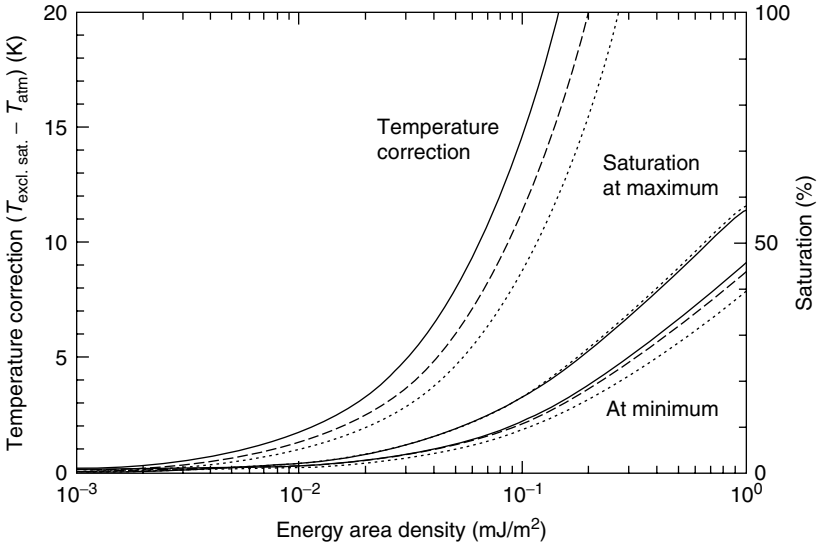


Figure 5.19 Na saturation at maximum, minimum, and temperature correction in relation to the energy area density. Basic assumptions are atmospheric temperature 200 K (dashed lines), laser linewidth 130 MHz, and pulse length 10 nsec. Further atmospheric temperatures are 150 K (dotted lines) and 250 K (solid lines). (From von der Gathen, P., *J. Geophys. Res.*, 96 (A3), 3679–3690, 1991. With permission.)

divergence of 1 mrad, and the laser linewidth of about 130 MHz. This gives the energy area density of 4.4×10^{-3} mJ/m^2 at 90 km. According to Figure 5.19, the estimated saturation for Na density is about 1% and the temperature correction is less than 0.5 K. This temperature bias caused by the saturation effect is less than the random errors due to photon noise.

b. Absolute laser frequency calibration and laser lineshape: The absolute laser frequency calibration and the laser lineshape influence the derived temperature and radial wind velocity because the effective backscatter cross-section $\sigma_{\text{eff}}(\nu)$ is the convolution of the Na absorption cross section $\sigma_{\text{abs}}(\nu)$ and the laser lineshape $g_{\text{L}}(\nu,\ \nu_{\text{L}})$. The factor $g_{\text{L}}(\nu,\ \nu_{\text{L}})$ contains the information of laser central

frequency, linewidth, and lineshape. In the Na wind/temperature lidar, the Doppler-free saturation–absorption features of a Na vapor cell serve as the absolute frequency reference to lock the CW ring dye laser. The PDA is seeded by the dye laser, so the frequency accuracy of the Na vapor cell response is supposed to transfer to PDA pulses. However, as described in Section 5.2.2.3, the central frequency of the PDA output pulse is shifted by the nonlinear effects associated with the pulse amplification. This frequency shift is on the order of 10 MHz. While this shift does not have a strong effect on temperature measurements, it can bias wind measurements. This frequency shift can be measured through carefully designed experiments as described in Section 5.2.2.3. Discrepancies between the lineshape (e.g., Gaussian shape) used in data reduction and the actual lineshape also introduce temperature and wind errors. The laser lineshape needs to be measured by using a narrowband etalon in order to derive accurate temperature and wind. As shown in Figure 5.18, since the actual laser lineshape is very close to be a Gaussian shape, the temperature error caused by using a Gaussian assumption is less than 0.5 K.

c. Receiver filter function and geometric factor: Accurate knowledge of the effective backscatter cross-section $\sigma_{\text{eff}}(\nu)$ is essential to both temperature and wind measurements. All the factors we considered above influence $\sigma_{\text{eff}}(\nu)$ by affecting either the determination of $\sigma_{\text{abs}}(\nu)$ or the laser frequency and lineshape. Our imprecise information of the effective backscatter cross-section $\sigma_{\text{eff}}(\nu)$ will result in systematic errors. In addition, our imprecise knowledge of the receiver filter function and geometric factor will also result in systematic errors. Through careful design of the laser and optical system, good arrangement of experimental setup (e.g., beam divergence), and precise calibration of laser frequency, lineshape, and filter transmission, these systematic errors can be eliminated or reduced to be within the limit of random errors caused by photon noise.

B. Random errors

Random errors for the lidar can be caused by the random uncertainty associated with laser and electronic jitter, random Na density fluctuations, and the shot noise associated with the photon-counting system. The laser and electronic jitter varies from system to system and also depends on the system alignment and environment stability. With proper laser and optics alignment in controlled lab environment, the jitter is usually quite small. The influence of Na density fluctuations can be minimized by proper weighting in two-frequency technique or eliminated by fast switching between three frequencies when taking data.

Shot noise ultimately limits the measurement precision of any photon-counting system because of the statistical nature of the photon-detection processes. The physics of the photon-counting process are well known. For the two-frequency ratio technique, the relative error of ratio R_{T} is given by

$$\frac{\Delta R_{\mathrm{T}}}{R_{\mathrm{T}}} = \frac{\left(1+\frac{1}{R_{\mathrm{T}}}\right)^{1/2}}{\left(N_{f_{\mathrm{a}}}\right)^{1/2}} \left[1+\frac{B}{N_{f_{\mathrm{a}}}} \frac{\left(1+\frac{1}{R_{\mathrm{T}}^2}\right)}{\left(1+\frac{1}{R_{\mathrm{T}}}\right)}\right]^{1/2} \tag{5.82}$$

where B is the background photon count, $N_{f_{\mathrm{a}}}$ is the Na photon count at peak frequency, and $R_{\mathrm{T}} = N_{f_c}/N_{f_{\mathrm{a}}}$. For the three-frequency ratio technique, the relative error of ratio R_{T} is given by

$$\frac{\Delta R_{\mathrm{T}}}{R_{\mathrm{T}}} = \frac{\left(1+\frac{1}{R_{\mathrm{T}}}\right)^{1/2}}{\left(N_{f_{\mathrm{a}}}\right)^{1/2}} \left[1+\frac{B}{N_{f_{\mathrm{a}}}} \frac{\left(1+\frac{2}{R_{\mathrm{T}}^2}\right)}{\left(1+\frac{1}{R_{\mathrm{T}}}\right)}\right]^{1/2} \tag{5.83}$$

while the relative error of ratio R_{W} is given by

$$\frac{\Delta R_{\mathrm{W}}}{R_{\mathrm{W}}} = \frac{\left(1+\frac{1}{R_{\mathrm{W}}}\right)^{1/2}}{\left(N_{f_{+}}\right)^{1/2}} \left[1+\frac{B}{N_{f_{+}}} \frac{\left(1+\frac{1}{R_{\mathrm{W}}^2}\right)}{\left(1+\frac{1}{R_{\mathrm{W}}}\right)}\right]^{1/2} \tag{5.84}$$

where N_{f_+} is the Na photon count at the plus wing frequency and $R_W = N_{f_-}/N_{f_+}$.

C. Calculation of errors

Systematic errors and random errors can be converted into the errors of derived temperature and radial wind velocity by the use of differentials of the corresponding ratio. Here we take the two-frequency temperature ratio as an example to explain the calculation method. A similar method can be used to analyze the error propagation from the uncertainties of lidar system parameters and photon noise to the uncertainties of the temperature and wind measurements by the three-frequency ratio technique.

For the two-frequency technique, the ratio R_T used to derive temperature error is

$$R_T(f_a, f_c, T, v_R, \sigma_L) = \frac{\sigma_{eff}(f_c, T, v_R, \sigma_L)}{\sigma_{eff}(f_a, T, v_R, \sigma_L)} \tag{5.85}$$

The temperature determined from Equation (5.85) is a function of R_T, f_a, f_c, σ_L, and v_R in implicit form. Therefore, the temperature errors are given by

$$\Delta T = \frac{\partial T}{\partial R_T}\Delta R_T + \frac{\partial T}{\partial f_a}\Delta f_a + \frac{\partial T}{\partial f_c}\Delta f_c + \frac{\partial T}{\partial \sigma_L}\Delta \sigma_L + \frac{\partial T}{\partial v_R}\Delta v_R \tag{5.86}$$

Implicitly differentiating Equation (5.86), we have

$$\Delta T = \Delta R_T\left(\frac{\partial R_T/\partial R_T}{\partial R_T/\partial T}\right) + \Delta f_a\left(\frac{\partial R_T/\partial f_a}{\partial R_T/\partial T}\right) + \Delta f_c\left(\frac{\partial R_T/\partial f_c}{\partial R_T/\partial T}\right) + \Delta \sigma_L\left(\frac{\partial R_T/\partial \sigma_L}{\partial R_T/\partial T}\right) + \Delta v_R\left(\frac{\partial R_T/\partial v_R}{\partial R_T/\partial T}\right) \tag{5.87}$$

The derivatives of the ratio R_T with respect to each of the system parameters are

$$\frac{\partial R_T}{\partial x} = R_T \left[\frac{\partial \sigma_{eff}(f_c)/\partial x}{\sigma_{eff}(f_c)} - \frac{\partial \sigma_{eff}(f_a)/\partial x}{\sigma_{eff}(f_a)} \right] \tag{5.88}$$

where x is one of the parameters (f_a, f_c, σ_L, v_R, T). If the laser lineshape is assumed to be Gaussian, the Na effective scattering cross-section σ_{eff} can be represented by Equation (5.43). By computing the partial differentials in Equation (5.87) with the help of Equation (5.43), the temperature errors introduced by photon noise and uncertainties in system parameters can be calculated. For example, the temperature error associated with the uncertainty in ratio R_T caused by photon noise is given by

$$\Delta T = \frac{\partial T}{\partial R_T} \Delta R_T = \frac{\Delta R_T}{R_T} \left[\frac{\partial \sigma_{eff}(f_c)/\partial T}{\sigma_{eff}(f_c)} - \frac{\partial \sigma_{eff}(f_a)/\partial T}{\sigma_{eff}(f_a)} \right]^{-1} \tag{5.89}$$

Similar expression can be derived for other system parameters and measurement techniques.

The derivatives and their magnitudes were computed for an operating point Q_T ($T = 200\,K$, $v_R = 0\,m/s$, $f_a = -651.5\,MHz$, $f_c = 187.8\,MHz$, $\sigma_L = 60\,MHz$) by Papen et al. (1995a). According to the computation results, the major source of temperature error is the uncertainty of laser linewidth, while the major source of wind error is the uncertainty of absolute laser frequency. An uncertainty of 4 MHz in laser linewidth will result in 0.46 K error in temperature. The temperature error is also sensitive to the laser frequency at f_c ($\sim$0.26 K per 2.45 MHz), but insensitive to the peak frequency f_a. The wind error is sensitive to the uncertainties of the wing frequencies f_+ and f_- ($\sim$0.7 m/s per 2.4 MHz) while insensitive to the uncertainty of laser linewidth. Because of the way the frequencies are chosen, the measurement errors are not sensitive to an error in the other measurement. For example, 1 m/s wind error results in temperature error by 0.017 K, while 1 K temperature error results in wind error by 0.087 m/s (Papen et al., 1995a).

5.2.3. Fe Boltzmann Temperature Lidar

5.2.3.1. Introduction

Knowledge of the temperature and constituent density in the mesopause region, especially in the polar mesopause region, is important for understanding a wide range of geophysical phenomena, such as the global climate change, the gravity wave effects on atmosphere dynamics, and polar mesospheric clouds. Currently, Na wind/temperature lidars are one of the most accurate remote sensing instruments used to probe the mesopause region. Unfortunately, the demanding environmental requirements of narrowband Na systems generally preclude their use at many important remote sites such as the North and South Poles. For global scale observations, the development of a robust temperature lidar is clearly attractive. To help address this measurement need, the University of Illinois in collaboration with the Aerospace Corporation and the National Center for Atmospheric Research (NCAR) developed a new lidar system — the Fe Boltzmann temperature lidar — for measuring temperature profiles from the middle stratosphere to the lower thermosphere, which can be deployed on research aircraft or operated at remote sites during both day and night (Chu et al., 2002a). This Fe Boltzmann temperature lidar was deployed to the North Pole in 1999 aboard the National Science Foundation/National Center for Atmospheric Research (NSF/NCAR) Electra aircraft and to the South Pole from 1999 to 2001 at a ground-based station. During these campaigns the instrument made the first lidar measurements of middle atmosphere temperature, Fe density, and PMC over both the North Pole and the South Pole during midsummer.

It is a completely different concept to use the Boltzmann technique to measure temperature. Unlike the Doppler technique that relies on the temperature dependence of the Doppler broadening of atomic resonance absorption lines, the Boltzmann technique relies on the temperature dependence of the Maxwell–Boltzmann distribution of atomic populations on different atomic energy levels under the assumption of thermodynamic equilibrium. The higher the temperature,

the greater is the population in upper energy levels. The population ratio between the upper and lower energy levels is determined by the temperature T and the energy difference ΔE between these two levels; that is, the population ratio is proportional to $\exp(-\Delta E/k_B T)$, where k_B is the Boltzmann constant. This well-known concept has been applied to imagers for deriving atmosphere temperatures from airglow. Applying the Boltzmann concept to the lidar field is a revolution for resonance fluorescence lidar, compared to conventional Doppler lidars. It lowers the requirement on laser linewidth and allows the lidar to be constructed in a robust way to enable difficult measurements at remote sites and in aircrafts. One of the key points for a Boltzmann lidar is to find a constituent (atom or molecule) that satisfies the following conditions: (1) the energy difference between ground-state sublevels is large enough to be detected, and also small enough to ensure adequate population on the upper ground state; (2) the transition wavelengths from its ground states to excited states fall in the wavelength range of available laser sources. Gelbwachs (1994) first pointed out that the mesospheric Fe atoms have the energy level distribution suitable for this type of system. Based upon his suggestion, a solid-state Fe Boltzmann temperature lidar using twin solid-state alexandrite lasers was designed, developed, and deployed from the North Pole to the South Pole by the UIUC for both daytime and nighttime measurements (Chu et al., 2002a). Following the success of the first Fe Boltzmann temperature lidar, the Arecibo group constructed an Fe Boltzmann lidar using a single dye laser and obtained temperature profiles during nighttime operation (Raizada and Tepley, 2002). In the following, we will describe the principle, technology, and error analysis of Fe Boltzmann temperature lidar based upon the UIUC system.

5.2.3.2. Measurement Principle (Boltzmann Technique)

There are four natural isotopes of Fe atoms: ^{54}Fe, ^{56}Fe, ^{57}Fe, and ^{58}Fe. Among them, ^{56}Fe is the most abundant isotope

Table 5.3 Isotopic Data of Fe Atoms

	^{54}Fe	^{56}Fe	^{57}Fe	^{58}Fe
Z	26	26	26	26
A	54	56	57	58
Nuclear spin	0	0	1/2	0
Natural abundance	5.845%	91.754%	2.119%	0.282%
Standard Atomic weight		55.845 g/mol		

with natural abundance of 91.754%. The isotopic data of Fe atoms are summarized in Table 5.3. The isotope shifts of ^{54}Fe, ^{57}Fe, and ^{58}Fe with respect to ^{56}Fe are around 1 GHz for the two resonance lines 372 and 374 nm that we are interested. Since the energy difference between the two lowest ground states is about 416 $cm^{-1} = 1.248 \times 10^4$ GHz, the contribution of these isotope shifts to the Boltzmann factor $\exp(-\Delta E/k_B T)$ is negligible (less than 1×10^{-4}). In addition, as shown below, the temperature determination is directly related to the effective cross-section ratio between 372 and 374 nm. Since the two lines have very similar isotope shifts, the effective cross-section ratio when considering all Fe isotopes is nearly identical to the ratio when considering ^{56}Fe only, which results in nearly identical Boltzmann temperature. Therefore, the influence of Fe isotopes to Boltzmann temperature is negligible. Here we only consider the ^{56}Fe atoms for the simplicity of description.

Since the nuclear spin of ^{56}Fe is zero, the Fe atoms have no hyperfine structures. The Fe Boltzmann temperature lidar relies on the unique energy level distribution diagram of the ground-state manifold of Fe atoms (see Figure 5.20). In thermodynamic equilibrium, the ratio of the populations in the $J = 3$ and $J = 4$ sublevels in the ground-state manifold is given by Maxwell–Boltzmann distribution law:

$$\frac{P_{374}(J=3)}{P_{372}(J=4)} = \frac{\rho_{\mathrm{Fe}}(374)}{\rho_{\mathrm{Fe}}(372)} = \frac{g_2}{g_1} \exp[-\Delta E/k_B T] \tag{5.90}$$

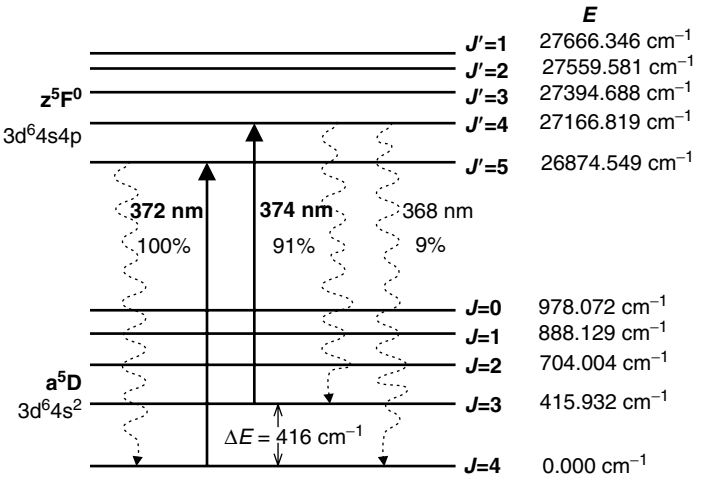


Figure 5.20 Energy level diagram of atomic Fe used for the Boltzmann technique. (From Chu, X., Pan, W., Papen, G.C., Gardner, C.S., and Gelbwachs, J.A., *Appl. Opt.*, 41, 4400–4410, 2002. With permission.)

where P_{372} and P_{374} are the populations in the two ground states with degeneracy factors $g_1 = 9$ and $g_2 = 7$, ρ_{Fe} (372) and ρ_{Fe} (374) are the associated Fe number densities, ΔE is the energy difference between the two levels ($\sim 416\,\mathrm{cm}^{-1}$), k_{B} is the Boltzmann constant, and T is the atmospheric temperature. The subscripts 372 and 374 refer to the excitation wavelengths of two transitions in nanometers (actually, $\lambda_{372} = 372.0993$ nm and $\lambda_{374} = 373.8194$ nm in vacuum). The line parameters for these two transitions are summarized in Table 5.4. At 200 K, the ratio of the populations is about 26. The temperature is then given by

$$T = \frac{\Delta E / k_{\mathrm{B}}}{\ln\left(\dfrac{g_2}{g_1}\dfrac{\rho_{\mathrm{Fe}}(372)}{\rho_{\mathrm{Fe}}(374)}\right)} \tag{5.91}$$

The Fe densities in these two states can be measured using resonance fluorescence lidar techniques. The densities are proportional to the number of backscattered photon counts from Fe atoms (N_{Fe}) detected for each wavelength probed by the lidar. The detected Fe photon count is given by the lidar equation

Table 5.4 Fe Resonance Line Parameters

Transition wavelength λ	372.0993 nm	373.8194 nm
Degeneracy for ground state	$g_1 = 9$	$g_2 = 7$
Degeneracy for excited state	$g_1' = 11$	$g_2' = 9$
Radiative lifetime of excited state (ns)	61.0	63.6
Einstein coefficient A_{ki} (10^8 s^{-1})	0.163	0.142
Oscillator strength f_{ik}	0.0413	0.0382
Branching ratio R_B	0.9959	0.9079
σ_0 (10^{-17} m^2)	9.4	8.7

$$N_{Fe}(\lambda, z) = \left(\frac{P_L \Delta t T_a}{hc/\lambda} E(\lambda, z)\right) \sigma_{eff}(\lambda, T, \sigma_L) R_{B\lambda} \rho_{Fe}(\lambda, z) \times \Delta z \left(E(\lambda, z) T_a \frac{A_R}{4\pi z^2} \eta\right) \tag{5.92}$$

where z is the altitude, P_L is the laser power, Δt is the temporal resolution, Δz is the vertical resolution, T_a is the one-way transmittance of the lower atmosphere, λ is the wavelength, c is the speed of light in vacuum, h is Planck's constant, $E(\lambda, z)$ is the extinction coefficient of the signal associated with Fe absorption, $\sigma_{eff}(\lambda, T, \sigma_L)$ is the effective scattering cross-section of the Fe transition (which is a function of temperature T, laser wavelength λ, and laser linewidth σ_L), $R_{B\lambda}$ is the branching ratio ($R_{B374} = 0.9079$, $R_{B372} = 0.9959$), A_R is the area of the receiving telescope, and η is the overall optical efficiency of the system. The effect of the radial velocity on the absorption cross-section has not been included, but it is negligible for zenith observations because the Doppler shift caused by the vertical wind velocity is typically less than 30 MHz, which is very small compared to the Doppler broadening and the laser linewidth ($\sim$1 GHz). The total detected signal also includes the background noise count $N_B(\lambda, z)$, which must be added to the Fe signal count given by Equation (5.92).

The current system uses two separate lasers and receivers because the two Fe lines are too far apart in wavelength to probe easily by tuning a single laser, yet they are too close to be separated in the receiving telescope using a dichroic beam splitter. To compensate for signal variations in each of these

systems, the photon counts from the Fe layer at each wavelength are normalized by the Rayleigh counts (N_R) over a common lower altitude range, typically near 50 km:

$$N_R(\lambda, z_R) = \left(\frac{P_L \Delta t T_a}{hc/\lambda}\right) \sigma_R(\lambda) \rho_{atmos}(z_R) \Delta z \left(T_a \frac{A_R}{4\pi z_R^2} \eta\right) \tag{5.93}$$

where z_R is the normalization altitude, $\rho_{atmos}(z_R)$ is the atmospheric number density at normalization altitude, and $\sigma_R(\lambda)$ is the effective atmospheric Rayleigh backscatter cross-section that is 4π times the Rayleigh angular backscatter cross-section $\sigma_R(\theta = \pi)$; that is, $\sigma_R(\lambda) = 4\pi\sigma_R(\theta = \pi)$, where θ is the scattering angle. The product of atmospheric number density and effective Rayleigh backscatter cross-section can be expressed as

$$\sigma_R(\lambda)\rho_{atmos}(z_R) = 1.370 \times 10^{-30} \times \frac{273}{T(z_R)} \times \frac{P(z_R)}{1013} \times \frac{1}{\lambda^{4.0117}} \tag{5.94}$$

where $T(z_R)$ is the atmospheric temperature in Kelvin, $P(z_R)$ is the atmospheric pressure in millibars, and λ is the laser wavelength in meters. The Rayleigh-normalized background-corrected Fe signal count is defined as

$$\begin{aligned} N_{norm}(\lambda, z) &= \frac{N_{Fe}(\lambda, z) + N_B(\lambda, z) - \hat{N}_B(\lambda)}{N_R(\lambda, z_R) + N_B(\lambda, z_R) - \hat{N}_B(\lambda)} \\ &= \frac{z_R^2 E^2(\lambda, z) R_{B\lambda} \sigma_{eff}(\lambda, T, \sigma_L) \rho_{Fe}(\lambda, z)}{z^2 \sigma_R(\lambda) \rho_{atmos}(z_R)} \end{aligned} \tag{5.95}$$

where $\hat{N}_B(\lambda)$ is the estimated background noise count. The Fe densities at each wavelength can be derived from Equation (5.95):

$$\rho_{Fe}(\lambda, z) = \frac{z^2}{z_R^2} \frac{N_{norm}(\lambda, z) \sigma_R(\lambda) \rho_{atmos}(z_R)}{E^2(\lambda, z) R_{B\lambda} \sigma_{eff}(\lambda, T, \sigma_L)} \tag{5.96}$$

By using Equation (5.90), the ratio of the normalized Fe signal counts at the two wavelengths (R_T) can be related to temperature

$$
\begin{aligned}
R_{\mathrm{T}}(z) &= \frac{N_{\mathrm{norm}}(\lambda_{374}, z)}{N_{\mathrm{norm}}(\lambda_{372}, z)} \\
&= \frac{g_2}{g_1}\frac{R_{\mathrm{B374}}}{R_{\mathrm{B372}}}\left(\frac{\lambda_{374}}{\lambda_{372}}\right)^{4.0117}\frac{E^2(\lambda_{374}, z)}{E^2(\lambda_{372}, z)} \\
&\quad \times \frac{\sigma_{\mathrm{eff}}(\lambda_{374}, T, \sigma_{\mathrm{L374}})}{\sigma_{\mathrm{eff}}(\lambda_{372}, T, \sigma_{\mathrm{L372}})}\exp(-\Delta E/k_{\mathrm{B}}T) \\
&= 0.7221 R_E^2(z) R_\sigma \exp(-598.44/T)
\end{aligned}
\tag{5.97}
$$

where

$$
R_E(z) = \frac{E(\lambda_{374}, z)}{E(\lambda_{372}, z)}, \quad R_\sigma = \frac{\sigma_{\mathrm{eff}}(\lambda_{374}, T, \sigma_{\mathrm{L374}})}{\sigma_{\mathrm{eff}}(\lambda_{372}, T, \sigma_{\mathrm{L372}})} \tag{5.98}
$$

At 200 K, the 372 nm Fe signal is approximately 30 times stronger than the 374 nm signal so that $R_{\mathrm{T}} \sim 1/30$. By solving Equation (5.97) for temperature, we obtain the final result

$$
\begin{aligned}
T(z) &= \frac{\Delta E/k_{\mathrm{B}}}{\ln\left[\dfrac{g_2}{g_1}\dfrac{R_{\mathrm{B374}}}{R_{\mathrm{B372}}}\left(\dfrac{\lambda_{374}}{\lambda_{372}}\right)^{4.0117}\dfrac{R_E^2(z)R_\sigma}{R_{\mathrm{T}}(z)}\right]} \\
&= \frac{598.44K}{\ln\left[\dfrac{0.7221 R_E^2(z) R_\sigma}{R_{\mathrm{T}}(z)}\right]}
\end{aligned}
\tag{5.99}
$$

Thus, by measuring the Fe signal level at the two wavelengths and computing $R_T(z)$ using Equations (5.95) and (5.97), the temperature can be derived from Equation (5.99) provided the ratios of the extinction and the effective backscatter cross sections are known.

As described earlier in the text, the two main concepts on how to measure temperature in the mesopause region by resonance fluorescence lidars are the Doppler and the Boltzmann techniques. No matter which method is used, the laser-induced fluorescence is the convolution of the laser lineshape with the thermal Doppler-broadened atomic absorption spectrum. Since the Doppler technique retrieves temperature from the shape of the Doppler-broadened atomic absorption spectrum, a narrow bandwidth laser ($\sim$0.1 GHz) is required in

order to precisely resolve the atomic absorption spectrum from the convoluted laser-induced fluorescence lineshape. On the other hand, the Boltzmann technique does not depend on an individual Doppler-broadened atomic absorption spectrum, but relies on the population ratio between two energy levels, that is, photon count ratio between two Doppler-broadened atomic absorption lines. Therefore, the laser bandwidth can be much broader and comparable to the atomic Doppler broadening linewidth in the mesopause region ($\sim$1 GHz) as long as sufficient number of fluorescence photons can be excited from the mesopause region. This feature allows the relative broadband lasers (such as solid-state lasers) to be deployed for the temperature measurements, and also simplifies the calibration for daytime measurements using a Fabry–Perot etalon as a narrowband filter. These features enable the development of a robust temperature lidar and the measurements of temperature in the Polar Regions.

To compare the performance of the Fe Boltzmann temperature lidar with the Na wind/temperature lidar, we define the sensitivity as the normalized change in the ratio R_T per degree of temperature change:

$$S_T = \frac{\partial R_T / \partial T}{R_T} \tag{5.100}$$

The sensitivity is a measure of the relative change of the ratio R_T with temperature variation. For accurate temperature measurements, high sensitivity is required. Illustrated in Figure 5.21 is a comparison of the sensitivity of the ratio R_T for the Fe Boltzmann technique and the Na Doppler two-frequency technique, adapted from Papen and Treyer (1998). Over typical mesospheric temperature range (100 to 250 K) the sensitivity of the Na Doppler technique is approximately half of that of the Fe Boltzmann technique. For both techniques, as the temperature increases, the sensitivity decreases. A significant difference between the two techniques is their sensitivity to radial winds and to frequency tuning errors. As described in Section 5.2.2, the Na Doppler technique is sensitive to the radial winds and frequency tuning

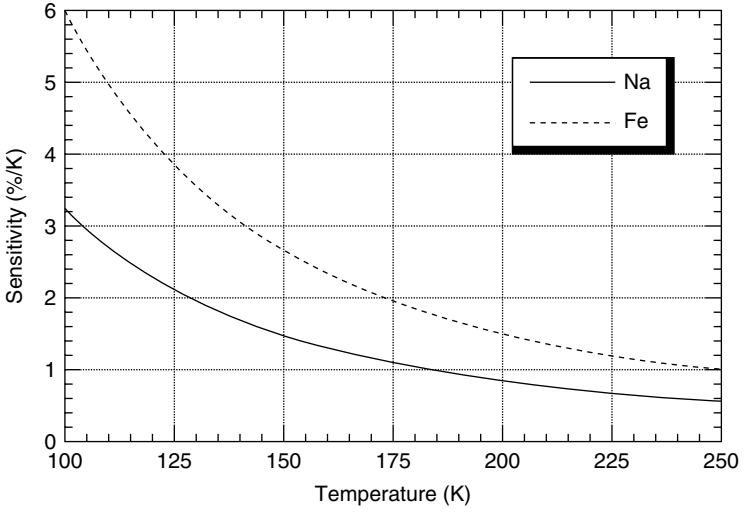


Figure 5.21 Comparison of sensitivity of the ratio R_T for the Fe Boltzmann technique and the Na Doppler technique (two-frequency). (From Papen, G.C., and Treyer, D., *Appl. Opt.*, 37, 8477–8481, 1998. With permission.)

errors. The nominal operating points for the broadband Fe system are at the optimal frequencies (peak frequencies), so the Fe Boltzmann lidar is insensitive to radial winds and frequency tuning errors. However, this robust insensitivity to radial winds also implies that the broadband Fe Boltzmann lidar system cannot be used for wind measurements.

5.2.3.3. Fe Boltzmann Lidar Instrumentation

A schematic of the system is shown in Figure 5.22. The system contains two laser transmitters and two optical receivers. Each laser transmitter consists of an injection-seeded, frequency-doubled, flash-lamp-pumped, pulsed alexandrite laser (Light Age, Inc., Model: PAL 101). The injection seed laser is a commercial tunable external cavity diode laser (Formerly EOSI, Inc., Model: 2010). The wavelength control of the seed laser is accomplished via a Burleigh wavemeter (Model: WA-1500) that has a frequency accuracy of ± 40 MHz at the fundamental wavelengths of 744 and 748 nm. This corresponds to a fre-

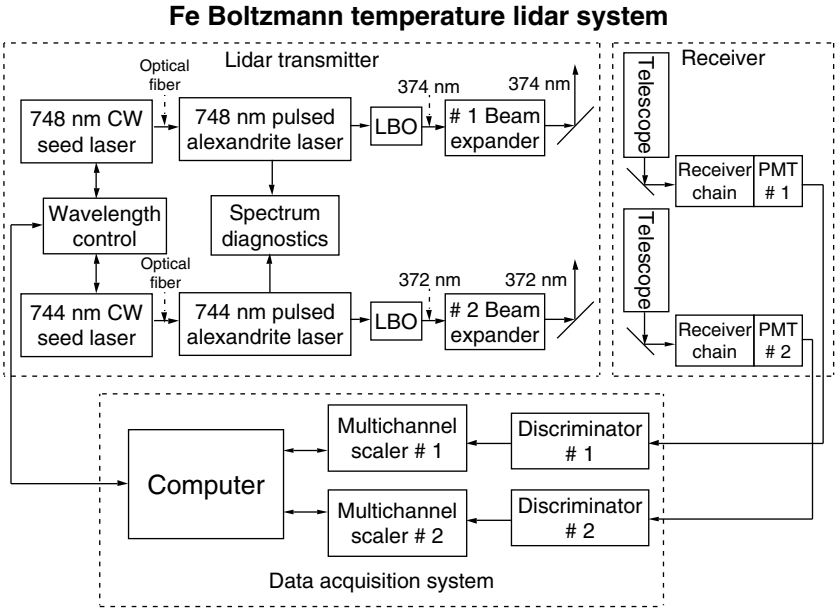


Figure 5.22 Diagram of the University of Illinois Fe Boltzmann temperature lidar system. (From Chu, X., Pan, W., Papen, G.C., Gardner, C.S., and Gelbwachs, J.A., *Appl. Opt.*, 41, 4400–4410, 2002. With permission.)

quency accuracy of ± 80 MHz when the lasers are frequency doubled to probe the Fe fluorescence lines. Each seed laser wavelength is sequentially monitored using a computer-controlled flip mirror. The wavelength is adjusted via a software control loop. The elliptical output of the external cavity laser is corrected using an anamorphic prism pair and then the circularized output is fiber coupled for injection into the main laser cavity. The fiber coupling allows flexibility in locating the seed lasers with respect to the alexandrite lasers.

The frequency-doubled output of the alexandrite laser is expanded to decrease the divergence to about 350 μrad to enable daytime operation. The pulsed output of the alexandrite laser is monitored using a commercial pulsed laser spectrum analyzer (Burleigh Instruments, Inc., Model: RFP-3600). The pulsed output is not single frequency, but consists of three to four modes within a frequency range of approximately

800 MHz. The effective cross-section of each Fe transition is determined by scanning the wavelength of each laser through the Fe transition using the atmospheric returns. This will be discussed in detail in the next section.

Each optical receiver includes a commercial Meade 0.4-m LX200 Schmidt–Cassegrain telescope with custom coatings to enhance ultraviolet (UV) transmission. The field of view of the telescope is controlled via a pinhole at the focal plane and is set to match the outgoing divergence of the laser. A mechanical chopper is placed close to the focal plane to enable blanking of the low altitude laser returns. The return signal is collimated, passed through an interference filter and a pressure-tuned, temperature-stabilized Fabry–Perot etalon before being detected by a photon-counting PMT. The output signal from the PMT is then thresholded using a discriminator, and the photon counts are integrated using a multichannel scaler. A list of the system parameters is given in Table 5.5. Compared to narrowband Na systems, the Fe density at 372 nm is, on average, about a factor of 2 higher than Na and the effective cross-section is about an order of magnitude lower than Na. Thus, the average power required to produce an Fe signal count at 372 nm that is comparable to the Na signal count is roughly a factor of 5 higher for Fe assuming equal efficiencies for the rest of the system.

In the Section 5.2.2.4, we mentioned that a Fabry–Perot etalon used for the narrowband Na temperature lidar to reject the sky background had a calibration issue. Due to the hyperfine structures of the Na atoms and the optical pumping effect when narrowband laser excitation is used, the returned fluorescence photons have different frequencies than the transmitted laser photons. The spread of the returned photon frequency can be over 2 GHz. This requires a precise knowledge of the etalon transmission function in order to derive the temperature accurately. However, such issue does not exist in the Fe Boltzmann temperature lidar. This is because the Fe atoms have no hyperfine structure due to the zero nuclear spin. Thus, the returned fluorescence photons have the same frequency as the transmitted laser photons when using resonant frequency of laser excitation. The Rayleigh scattering

Table 5.5 System Parameters for the University of Illinois Fe Boltzmann Temperature Lidar

Lidar Transmitter		
Characteristics	Channel 1 (374 nm)	Channel 2 (372 nm)
Alexandrite laser	Light Age, Inc., Model: PAL 101 Frequency doubled pulsed alexandrite laser	
Linewidth (FWHM) (MHz)	864	876
Pulse energy (mJ)	100	100
Pulse length (nsec)	55–65	55–65
Repetition rate (Hz)	33.2	34.1
Average power (W)	3	3
Beam divergence (mrad)	0.35	0.35
Injection-seeder laser	Formerly EOSI, Inc., Model: 2010 External cavity diode laser	
Wavelength (nm)	747.6390	744.1990
Output power (mW)	3	3
Linewidth (MHz) (50 msec)	<100 kHz	<100 kHz

Lidar Receiver		
Characteristics	Nighttime	Daytime
Telescope	Meade, Model: LX200 (f/10) 0.4064 m diameter Schmidt–Cassegrain design	
Aperture area (m^2)	0.130	
Telescope focal length (mm)	4064	
Field-of-view (mrad)	1	0.5
Interference filter bandwidth (nm)	4.2	0.3
Interference filter peak transmission (%)	77	46
Fabry–Perot etalon bandwidth (GHz)	–	30
Etalon free-spectra-range (GHz)	–	650
Etalon finesse	–	22
Etalon spacing (μm)	–	230
Etalon peak transmission (%)	–	50
PMT quantum efficiency (%)	28	
Receiver optical efficiency (%)	58.7	18.0
Total receiver efficiency (%)	16.4	5.0

photons also have the same frequency as the transmitted laser photons since the Doppler shift due to wind velocity in zenith direction is negligible. Thus, the returned Fe fluorescence and

Rayleigh scattering photons have the same frequency and undergo through the same etalon transmission function. Therefore, the etalon filter function can be canceled out when normalizing the Fe signals to the Rayleigh signals as given by Equation (5.95).

5.2.3.4. Temperature and Error Analysis

The approach to derive the atmosphere temperature and Fe density information from the lidar photon count profile is similar to what has been used for the Na wind/temperature lidar. The calibration curve (R_T versus T) used for the Fe Boltzmann temperature lidar is plotted in Figure 5.23. Here, R_T is given by Equation (5.97). Once the ratio R_T is computed from the observed photon counts of the 374 and 372 nm channels, the temperature can be inferred from this calibration curve. Then the Fe density can be calculated from Equation (5.96) using derived temperature.

There are two major sources of error, viz. fluctuations in the signal levels associated with photon noise and fluctuations in the effective backscatter cross sections associated with laser tuning errors and laser linewidth fluctuations. The

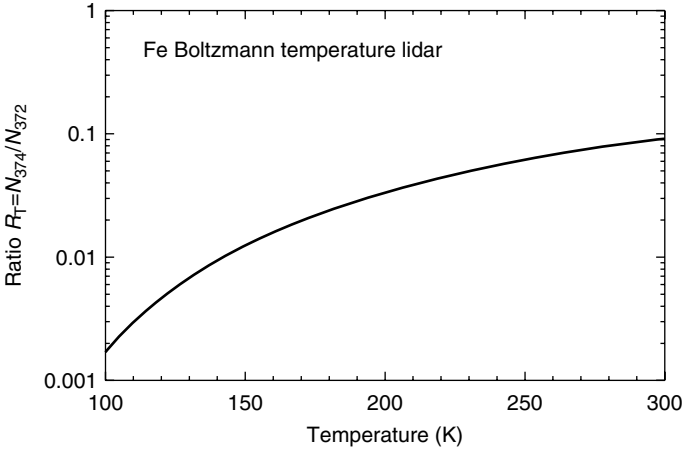


Figure 5.23 Calibration curve for the Fe Boltzmann temperature lidar: ratio R_T versus temperature T.

extinction is calculated from the measured Fe densities and temperatures and also includes errors. However, the two-way extinction correction is typically no more than about 1 to 2% at the top of the 372 nm Fe layer and much less than a percent for the 374 nm Fe layer. For this reason, extinction errors are small and can be neglected. Because the laser tuning errors and linewidth fluctuations are uncorrelated with signal photon noise, by using $\Delta T = (\partial T/\partial R_{\mathrm{T}})\Delta R_{\mathrm{T}} + (\partial T/\partial R_{\sigma})\Delta R_{\sigma}$, the total RMS temperature error can be derived from Equation (5.99) as

$$\Delta T_{\mathrm{rms}} = \frac{T^2}{\Delta E/k_{\mathrm{B}}} \sqrt{\left\langle \left(\frac{\Delta R_{\mathrm{T}}}{R_{\mathrm{T}}}\right)^2 \right\rangle + \left\langle \left(\frac{\Delta R_{\sigma}}{R_{\sigma}}\right)^2 \right\rangle} \tag{5.101}$$

Since $\Delta E/k_{\mathrm{B}} \sim 600\,\mathrm{K}$, and T is in the range of 150 to 250 K, the total relative error for both R_{T} and R_{σ} in Equation (5.101) must be no larger than 1.5% for a $\pm 1\,\mathrm{K}$ accurate temperature measurement at $T = 200\,\mathrm{K}$. We now consider each of these two error sources separately.

The total effective backscatter cross-section is proportional to the effective absorption cross-section

$$\sigma_{\mathrm{eff}}(\lambda, T, \sigma_{\mathrm{L}}) = \sigma_{\mathrm{effabs}}(\lambda, T, \sigma_{\mathrm{L}}) \tag{5.102}$$

We have accounted for the decay of the excited state to multiple ground states by including the appropriate branching ratios in the lidar equation. The effective absorption cross-section is the convolution of the atomic absorption cross section and laser lineshape

$$\sigma_{\mathrm{eff}}(\lambda, T, \Delta\nu, \sigma_{\mathrm{L}}) = \int_{-\infty}^{+\infty} \sigma_{\mathrm{abs}}(\nu', \lambda_0, T) g_{\mathrm{L}}(\nu', \lambda, \sigma_{\mathrm{L}}) \mathrm{d}\nu' \tag{5.103}$$

where $\sigma_{\mathrm{abs}}(\nu', \lambda_0, T)$ is the intrinsic atomic absorption cross-section centered at wavelength λ_0, and $g_{\mathrm{L}}(\nu', \lambda, \sigma_{\mathrm{L}})$ is the normalized laser lineshape centered at wavelength λ, ν' is the frequency, and $\Delta\nu$ is the frequency difference between the Fe line center λ_0 and the laser line center λ. Assuming a Gaussian functional form for both the laser lineshape and the absorption cross-section, the effective cross-section is

$$\sigma_{\text{eff}}(\lambda, T, \Delta\nu, \sigma_{\text{L}}) = \frac{1}{\sqrt{2\pi}\sigma_\lambda} \frac{e^2}{4\epsilon_0 m_{\text{e}} c} f_\lambda \exp\left(-\frac{\Delta\nu^2}{2\sigma_\lambda^2}\right) \quad (5.104)$$

where f_λ is the oscillator strength of the transition ($f_{372} = 0.0413$ and $f_{374} = 0.0382$), ε_0 is the permittivity of free space, e is the charge of an electron, m_{e} is the mass of the electron, c is the velocity of light, and

$$\sigma_\lambda = \sqrt{\sigma_{\text{D}}^2 + \sigma_{\text{L}}^2} \quad (5.105)$$

where σ_{D} is the rms Doppler width of the intrinsic atomic absorption cross-section,

$$\sigma_{\text{D}} = \frac{1}{\lambda}\sqrt{\frac{k_{\text{B}}T}{m_{\text{Fe}}}} \quad (5.106)$$

and m_{Fe} is the mass of an Fe atom.

In practice, we tune the lasers to the peaks of the Fe lines so that uncertainties in the assumed values for the cross-section arise because of tuning errors ($\Delta\nu$) and uncertainties in the assumed values for the laser linewidths ($\Delta\sigma_{\text{L}}$). The ratio of the cross sections is given by

$$R_\sigma = 0.9249 \frac{\sigma_{372}}{\sigma_{374}} \exp\left(\frac{\Delta\nu_{372}^2}{2\sigma_{372}^2} - \frac{\Delta\nu_{374}^2}{2\sigma_{374}^2}\right) \quad (5.107)$$

While the temperature is required to evaluate the Doppler linewidths, the cross-section ratio is insensitive to temperature so that errors in R_σ associated with errors in assumed-temperature are negligible. Because the lasers are identical and the wavelength difference is small, we assume the laser linewidths are similar so that σ_{372} is approximately equal to σ_{374}. In this case, the cross-section ratio can be simplified

$$R_\sigma \cong \begin{cases} 0.9292\sqrt{1 + \frac{\left(\sigma_{\text{L372}}^2 - \sigma_{\text{L374}}^2\right)}{\sigma_\lambda^2}} \exp\left[\frac{\left(\Delta\nu_{372}^2 - \Delta\nu_{374}^2\right)}{2\sigma_\lambda^2}\right] \sigma_{\text{L}} \ll \sigma_{\text{D}} \\ 0.9249\sqrt{1 + \frac{\left(\sigma_{\text{L372}}^2 - \sigma_{\text{L374}}^2\right)}{\sigma_\lambda^2}} \exp\left[\frac{\left(\Delta\nu_{372}^2 - \Delta\nu_{374}^2\right)}{2\sigma_\lambda^2}\right] \sigma_{\text{D}} \ll \sigma_{\text{L}} \end{cases} \quad (5.108)$$

If both the laser linewidths and the squares of the tuning errors are identical, then $0.9249 < R_\sigma < 0.9292$. For identical lasers, the cross-section ratio varies less than 0.5% as the laser linewidth varies from zero to infinity. Because the cross-section ratio is insensitive to the absolute value of the laser linewidths when they are approximately equal, it is more important to know the linewidth difference rather than the absolute values of the individual linewidths. We can express the relative error in the cross-section ratio in terms of the laser tuning errors and uncertainties in the assumed values for the laser linewidths, or in terms of the linewidth difference.

$$\frac{\Delta R_\sigma}{R_\sigma} \cong \begin{cases} \frac{|\Delta\nu_{372}^2 - \Delta\nu_{374}^2|}{2\sigma_\lambda^2} \\ \sqrt{2}\left(\frac{\sigma_L}{\sigma_\lambda}\right)^2 \frac{\Delta\sigma_L}{\sigma_L} \\ \left(\frac{\sigma_L}{\sigma_\lambda}\right)^2 \frac{|\sigma_{L372} - \sigma_{L374}|}{\sigma_L} \end{cases} \tag{5.109}$$

By combining Equation (5.109) with Equation (5.101), we obtain

$$\Delta T_{\text{rms}}(\text{K}) \cong \begin{cases} \pm \frac{T^2}{2(\Delta E/k_B)} \frac{|\Delta\nu_{372}^2 - \Delta\nu_{374}^2|}{\sigma_\lambda^2} \\ \pm \frac{\sqrt{2}T^2}{(\Delta E/k_B)} \left(\frac{\sigma_L}{\sigma_\lambda}\right)^2 \frac{\Delta\sigma_L}{\sigma_L} \\ \pm \frac{T^2}{(\Delta E/k_B)} \left(\frac{\sigma_L}{\sigma_\lambda}\right)^2 \frac{|\sigma_{L372} - \sigma_{L374}|}{\sigma_L} \end{cases} \tag{5.110}$$

To compute the effective cross sections from Equation (5.104) or their ratio from Equation (5.107), we need to determine σ_λ for each line. This is done by scanning the lasers through the Fe fluorescence lines and recording the atmospheric returns as a function of frequency. A Gaussian function of frequency is fitted to the normalized signal counts to determine the total rms linewidths. The laser linewidths are computed from Equations (5.105) and (5.106). Plotted in Figure 5.24 is an example fit to 372 nm data. The results for the cross-section and laser parameters and the statistical errors of the fits are listed in Table 5.6.

There are two ways to determine the cross-section ratio and the errors. The cross sections can be computed from the

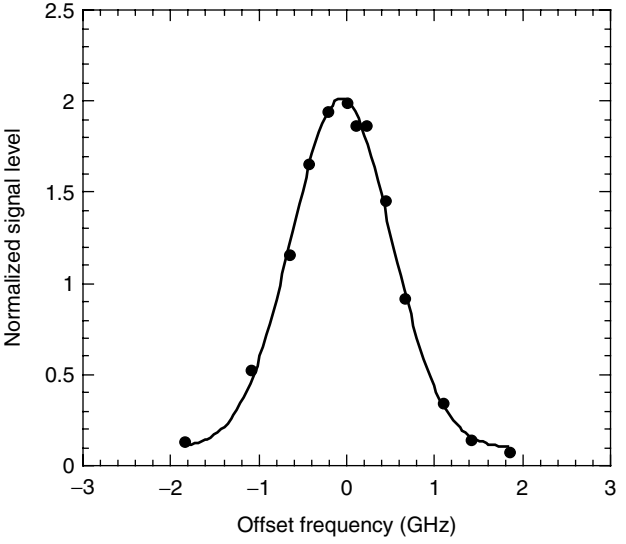


Figure 5.24 Plot of the effective backscatter cross section spectrum for the Fe 372-nm line obtained by the scanning of the lidar frequency and observation of the returns from the mesospheric Fe layer. (From Chu, X., Pan, W., Papen, G.C., Gardner, C.S., and Gelbwachs, J.A., *Appl. Opt.*, 41, 4400–4410, 2002. With permission.)

laser scans using Equation (5.104). In this case the relative error in the cross-section ratio is given by

$$\frac{\Delta R_\sigma}{R_\sigma} = \sqrt{\left(\frac{\Delta\sigma_{\text{eff}}(372)}{\sigma_{\text{eff}}(372)}\right)^2 + \left(\frac{\Delta\sigma_{\text{eff}}(374)}{\sigma_{\text{eff}}(374)}\right)^2} \tag{5.111}$$

Table 5.6 Measured Cross Sections and Laser Parameters

Parameter	372 nm system	374 nm system
$\sigma_\lambda \pm \Delta\sigma_\lambda$ (MHz)	566.8 ± 16.7	562.1 ± 20.0
σ_D (MHz) at $T = 170$ K	427.7	425.7
$\Delta\nu \pm \delta(\Delta\nu)$ (MHz)	59.7 ± 10.8	82.2 ± 13.8
$\sigma_L \pm \Delta\sigma_L$ (MHz)	371.9 ± 25.4	367.0 ± 30.7
$\sigma_{\text{eff}} \pm \Delta\sigma_{\text{eff}}$ ($\times 10^{-12}$ cm^2)	0.770 ± 0.023	0.718 ± 0.026
$\Delta\sigma_\lambda/\sigma_\lambda$ (%)	2.95	3.56
$\Delta\sigma_L/\sigma_L$ (%)	6.83	8.37
$\Delta\sigma_{\text{eff}}/\sigma_{\text{eff}}$ (%)	2.99	3.62

By using the data in Table 5.6, the cross-section ratio is 0.9348 and the relative error is 4.7%, which according to Equation (5.101) yields an error of about ± 3.1 K at $T = 200$ K. Alternatively, we can assume the laser linewidths are identical and that the lasers are both tuned to the peaks of the fluorescence lines. In this case, the cross-section ratio is assumed to be 0.9292 when $\sigma_L \ll \sigma_D$ or 0.9249 when $\sigma_D \ll \sigma_L$ (Equation 5.108). In our case, where the laser linewidths are both approximately 370 MHz and do not satisfy either extreme, the assumed value of the cross-section ratio would be 0.9274 (Equation 5.107). According to the linewidth scan data in Table 5.6, tuning errors are comparable to the ± 80 MHz accuracy of the Burleigh wavemeter at the frequency-doubled UV lines. Tuning errors would introduce a relative error of less than 1% in the cross-section ratio, which corresponds to a temperature error of ± 0.7 K. Uncertainties in the laser linewidths (25 and 31 MHz) introduce a temperature error of ± 3.1 K. Alternatively, if the laser linewidths are similar, we can use the linewidth difference to compute the temperature error, rather than the uncertainties of the measured values. The laser linewidths differ by 4.9 ± 40 MHz. If we assume the laser linewidth difference is no larger than the uncertainty in the measured difference (40 MHz), then the relative cross-section error is less than 4.6%, which corresponds to a temperature error of ± 3.1 K. If the actual linewidth difference were assumed to be the measured value of 4.9 MHz, then the temperature error would be approximately ± 0.4 K.

The impact of photon noise is calculated using the same analysis for the Na lidar technique where winds and temperatures are determined by computing the ratio of normalized photon counts obtained at several frequencies within the Na D_2 fluorescence line. Because the Fe signals are normalized by the strong Rayleigh signal at lower altitudes near 50 km, the photon noise contributed by the Rayleigh signal is negligible. The dominant error source is photon noise contributed by the weaker 374 nm Fe channel. If we assume the Rayleigh normalizing signals for the 372 and 374 nm channels are comparable, the temperature ratio R_T is given approximately by

$$R_{\mathrm{T}} \cong \frac{N_{\mathrm{Fe}}(374,z) + N_{\mathrm{B}}(374,z) - \hat{N}_{\mathrm{B}}(374)}{N_{\mathrm{Fe}}(372,z) + N_{\mathrm{B}}(372,z) - \hat{N}_{\mathrm{B}}(372)} \tag{5.112}$$

The rms relative error in the temperature ratio associated with photon noise is easily calculated from Equation (5.112) by recognizing that the Fe signal and background counts are uncorrelated Poisson distributed random numbers. If we assume that the mean background counts on the 372 and 374 nm channels are approximately equal and $T = 200\,\mathrm{K}$ so that $1/R_{\mathrm{T}} \sim 30$, the rms relative error is given by

$$\begin{aligned}\sqrt{\left\langle \left(\frac{\Delta R_{\mathrm{T}}}{R_{\mathrm{T}}}\right)^2 \right\rangle} &= \frac{\sqrt{1 + (1/R_{\mathrm{T}})}}{\sqrt{N_{\mathrm{Fe}}(372,z)}} \sqrt{1 + \frac{[1 + (1/R_{\mathrm{T}}^2)]}{[1 + (1/R_{\mathrm{T}})]} \frac{1}{\mathrm{SBR}_{372}}} \\ &\approx \frac{5.6}{\sqrt{N_{\mathrm{Fe}}(372,z)}} \sqrt{1 + \frac{29}{\mathrm{SBR}_{372}}}\end{aligned} \tag{5.113}$$

where the signal-to-background ratio (SBR) is

$$\mathrm{SBR}_{372} = \frac{N_{\mathrm{Fe}}(372,z)}{\hat{N}_{\mathrm{B}}(372)} \tag{5.114}$$

By substituting Equation (5.113) into Equation (5.101), the rms temperature error associated with photon noise at $T = 200\,\mathrm{K}$ is

$$\Delta T_{\mathrm{rms}}(\mathrm{K}) = \frac{\pm 372}{\sqrt{N_{\mathrm{Fe}}(372,z)}} \sqrt{1 + \frac{29}{\mathrm{SBR}_{372}}} \tag{5.115}$$

At night when SBR_{372} is very large, approximately 140×10^3 Fe signal counts are required in each resolution cell of the stronger 372 nm channel to achieve a temperature accuracy of $\pm\,1\,\mathrm{K}$. This requirement drops to about 15×10^3 counts for an accuracy of $\pm 3\,\mathrm{K}$. During daytime when SBR_{372} can be much less than 1, the signal requirements increase substantially. For example if $\mathrm{SBR}_{372} = 1$, the required signal level is 4×10^6 counts for $\pm 1\,\mathrm{K}$ accuracy. Long averaging times, typically several hours, are required to obtain accurate temperature

measurements during the daytime. The error budget for the Fe Boltzmann lidar is summarized in Table 5.7.

5.2.3.5. Rayleigh Temperature Retrieval

Above the stratospheric aerosol layers (~30 km) and below the Fe layer, the received signal results from pure molecular scattering. Molecular scattering is proportional to the atmospheric density and the temperature profile can be derived from the relative atmospheric density profile using the Rayleigh lidar technique. The Fe lidar Rayleigh signal between 30 and 75 km can be used to study the atmospheric temperature and density profile. The use of high power lasers with the Rayleigh technique was pioneered by Hauchecorne and Chanin (1980). The data analysis approach is very similar to that employed by Elterman in the early 1950s to measure stratospheric temperatures with search light technique (Elterman, 1951, 1953, 1954). This involves integrating the relative density profile downward using a starting temperature at the highest altitude in combination with the hydrostatic equation and the ideal gas law. The starting temperature may be chosen from a model because when the equation has been integrated downward by about one and half scale heights (atmospheric scale height is the altitude range in which density decreases by the factor $1/e$), the calculated

Table 5.7 Fe Boltzmann Temperature Lidar Error Budget for $T = 200$ K and $\sigma_L = 370$ MHz

Parameter	Temperature Error (K)	Value for $\Delta T_{rms} = \pm 1$ K
Laser tuning errors $\Delta\nu$	$\pm 33 \frac{\lvert\Delta\nu_{372}^2 - \Delta\nu_{374}^2\rvert}{\sigma_\lambda^2}$	100 MHz
Laser linewidth errors $\Delta\sigma_L$	$\pm 95 \left(\frac{\sigma_L}{\sigma_\lambda}\right)^2 \frac{\Delta\sigma_L}{\sigma_L}$	9 MHz
Laser linewidth difference $\lvert\sigma_{L372} - \sigma_{L374}\rvert$	$\pm 67 \left(\frac{\sigma_L}{\sigma_\lambda}\right)^2 \frac{\lvert\sigma_{L372} - \sigma_{L374}\rvert}{\sigma_L}$	13 MHz
Signal level $N_{Fe}(372, z)$	$\pm \frac{372}{\sqrt{N_{Fe}(372,z)}} \sqrt{1 + \frac{29}{SBR_{372}}}$	1.8×10^5 at $SBR_{372} = 100$ 4.2×10^6 at $SBR_{372} = 1$

temperature is relatively insensitive to the starting estimation. The hydrostatic equation

$$dP(z) = -\rho(z)g(z)dz \tag{5.116}$$

can be combined with the ideal gas law

$$P(z) = \frac{\rho(z)RT(z)}{M(z)} \tag{5.117}$$

and integrated to yield

$$T(z) = T(z_0)\frac{\rho(z_0)}{\rho(z)}\frac{M(z)}{M(z_0)} + \frac{M(z)}{R}\int_z^{z_0}\frac{\rho(z')g(z')}{\rho(z)}dz' \tag{5.118}$$

where

$T(z)$ = atmospheric temperature profile (K)
$P(z)$ = atmospheric pressure profile (mbar)
$\rho(z)$ = atmospheric mass density profile (kg/m^3)
$g(z)$ = gravitational acceleration (m/s^2)
$M(z)$ = mean molecular weight of the atmosphere
R = universal gas constant (8.31432 J/mol/K)
z_0 = altitude of the upper level starting temperature (m)

The atmospheric mass density $\rho(z)$ and number density $n(z)$ have the following relationship:

$$\rho(z) = n(z)M(z)/N_A \tag{5.119}$$

where N_A is the Avogadro constant. Substituting Equation (5.119) into Equation (5.118), we obtain

$$T(z) = T(z_0)\frac{n(z_0)}{n(z)} + \frac{M(z)}{R}\int_z^{z_0}\frac{n(z')M(z')g(z')}{n(z)M(z)}dz' \tag{5.120}$$

Below 100 km for the well-mixed atmosphere, the mean molecular weight of the atmosphere $M(z)$ is a constant. Thus, the $M(z')$ and $M(z)$ are canceled out in Equation (5.120), and we obtain the temperature derived from Rayleigh signals by

$$T(z) = T(z_0)\frac{n(z_0)}{n(z)} + \frac{M(z)}{R}\int_z^{z_0}\frac{n(z')g(z')}{n(z)}dz' \tag{5.121}$$

Notice that the atmosphere number density $n(z)$ appears as a ratio of densities, that is, the relative number density. Thus, to determine temperature, it is only necessary to measure the relative density, which can be derived from the Rayleigh photon count $N_R(z)$ at different altitudes given by Equation (5.93). The accuracy of this technique depends upon the Rayleigh photon count $N_R(z)$ and the accuracy of the upper level temperature estimate $T(z_0)$. The variance of derived temperature is given as below (Gardner, 1989):

$$\text{var}[T(z)] \approx \frac{T^2(z)}{N_R(z)} + \left\{ \text{var}[T(z_0)] + \frac{T^2(z_0)}{N_R(z_0)} \right\} \exp[-2(z_0 - z)/H] \tag{5.122}$$

where H is the atmospheric scale height ($\sim$7 km for the upper stratosphere and the lower mesosphere).

Most middle atmosphere Rayleigh lidars employ frequency-doubled pulsed Nd:YAG lasers operating at 532 nm in the green region of the visible spectrum. Typical systems employ telescopes with diameters near 1 m and lasers with average powers levels in the 10 to 15 W range. These systems have power aperture products of about 10 W m^2. The Fe Boltzmann lidar operates in the near-UV where Rayleigh scattering is more than four times stronger than at 532 nm (Equation 5.94). By combining the molecular (Rayleigh) scattered signal from the 372 and 374 nm channels, temperatures can be derived from about 30 to about 75 km using the Rayleigh technique. The Fe Boltzmann temperature at the bottom of the Fe layer near 80 km, rather than a model estimate, can be used as the starting temperature for the retrieval. For our system, which employs two 3-W lasers and two 0.4-m diameter telescopes, the equivalent power aperture product at 532 nm is about 3 W m^2. Thus, the current Fe Boltzmann lidar system can derive temperature profiles with an accuracy and vertical resolution comparable to many existing Rayleigh lidars if the signals are integrated about three times longer. If the same 1-m diameter telescopes are used for the Fe system, the equivalent power aperture product at 532 nm will be about 18 W m^2, which is actually better than the 532-nm Rayleigh system.

5.2.3.6. Temperature Measurement over the North and South Poles

As described previously, the Fe Boltzmann temperature lidar can measure the temperature in the upper mesosphere and lower thermosphere by the Boltzmann technique, and in the upper stratosphere and lower mesosphere by the Rayleigh technique. The combination of these two capabilities makes the Fe Boltzmann temperature lidar able to measure the temperature from 30 to 110 km. If further combining with the balloon-sonde data from the ground to 30 km, the temperature measurements will be able to cover the whole altitude range from the ground to 110 km. These measurement capabilities were clearly demonstrated in the lidar observations from the North Pole to the South Pole made by the University of Illinois (Chu et al., 2002a; Gardner et al., 2001; Pan et al., 2002).

The Fe Boltzmann temperature lidar was first deployed on the NSF/NCAR Electra aircraft to make observations of Leonid meteor shower over Okinawa in November 1998 (Chu et al., 2000a). In June and July 1999, the lidar system was again deployed on the NSF/NCAR Electra to make temperature, Fe density, and PMC observations over the north polar cap during the Arctic Mesopause Temperature Study (AMTS). The Electra aircraft flew at varying altitudes from 6 to 8.5 km with speed about 150 m/s. There were totally 10 flights during AMTS campaign. AMTS began on June 16, 1999 with a ferry flight from Broomfield, Colorado (40°N, 105°W) to Winnipeg, Canada, and then to Resolute Bay, Canada (75°N, 95°W) where the campaign was based. One roundtrip flight was made to Sondrestromfjord, Greenland (67°N, 50°W) on June 19, 1999. The first flight to the North Pole was made on June 21, 1999, and then another three flights were made to the geographic North Pole on July 1, July 2, and July 4, 1999. On the last three flights to the geographic Pole, the return flight path passed directly over the magnetic North Pole (79°N, 105°W). On July 5, 1999, the system was flown to Anchorage, Alaska (61°N, 150°W) and one additional flight was made on July 8, 1999 to probe NLC over the Gulf of

Alaska. Finally, the lidar system was flown back to Broomfield, Colorado passing through Fort Collins on July 9, 1999. A total of 52 h of airborne lidar observations were made during AMTS. In November 1999, the Fe lidar system was installed in the Atmospheric Research Observatory on the ground, which is 488 m north of the geographic South Pole at the Amundsen–Scott South Pole Station. Observations began on December 2, 1999 and continued through the austral summer and winter from 1999 to 2001, until November 2001. More than 1700 h of temperature, Fe density, and polar mesospheric clouds observations were made at the South Pole during this 2-year deployment.

The mesospheric Fe chemistry is temperature dependent. The primary sink reaction $FeO + O_2 \rightarrow FeO_3$ on the layer bottom side proceeds most rapidly at low temperatures (Rollason and Plane, 2000). This reaction drives the seasonal variations in the Fe column abundance. Because mesopause temperatures are extremely low over the polar caps at midsummer, the Fe densities are also expected to be quite low. Thus, summertime observations over the polar caps, which are sunlit 24 h per day, represent the greatest challenge to making accurate temperature measurements with the Fe Boltzmann lidar. A composite profile of temperatures measured during the summer solstice periods over the North and South Poles is plotted in Figure 5.25(a). The corresponding Fe density profiles are plotted in Figure 5.25(b). During midsummer over the polar caps, the Fe layer is thin and the peak densities are typically quite low. The June 21, 1999 data over the North Pole are a notable exception. On this flight a sporadic Fe layer (Fe_s) formed near 106 km in lower thermosphere with peak densities exceeding $2 \times 10^5\,cm^{-3}$. This prominent feature was observed on the three subsequent flights to the North Pole. However, the peak densities and abundances of the Fe_s on these later flights were much smaller. This dense sporadic layer provided an exceptionally strong backscatter signal so that excellent temperature data were obtained between 102 and 109 km. The thin layers observed at the South Pole near 91 km on December 24, 2000 and 98 km on December 29, 2000 are also sporadic layers, but the lower

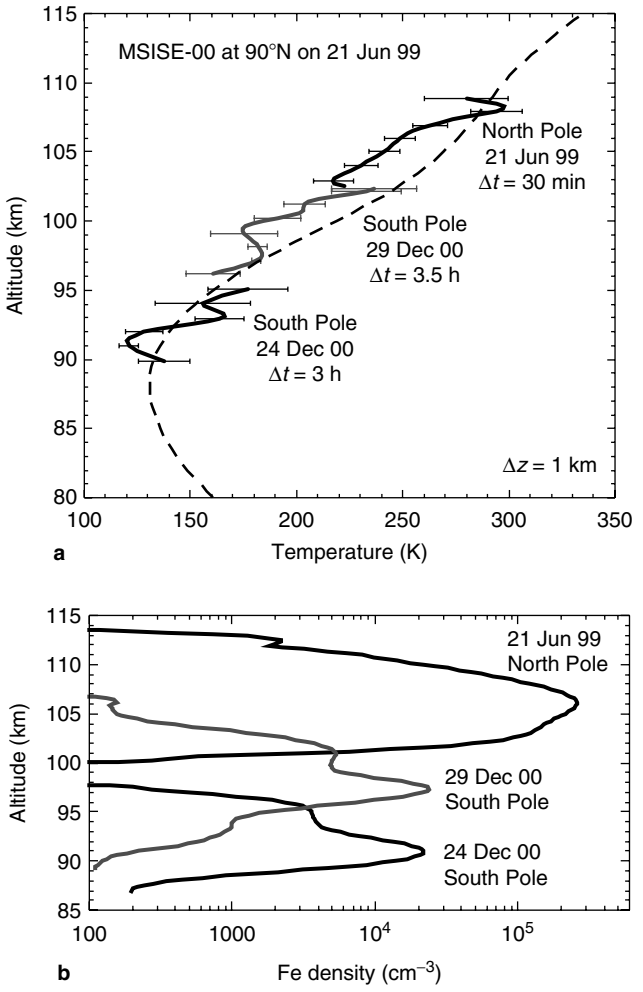


Figure 5.25 (a) Temperature profiles measured by the University of Illinois Fe Boltzmann temperature lidar over the North Pole on June 21, 1999 and over the South Pole on December 24, 2000 and December 29, 2000 along with MSISE-00 model data for June 21, 1999 at the North Pole. (b) The 372-nm ground-state Fe density profiles corresponding to the temperature profiles plotted in (a). (From Chu, X., Pan, W., Papen, G.C., Gardner, C.S., and Gelbwachs, J.A., *Appl. Opt.*, 41, 4400–4410, 2002. With permission.)

densities just above them are more typical. The combination of low densities and high background noise makes temperature observations difficult at midsummer over either pole. Even so, these airborne and ground-based profiles demonstrate that the lidar is capable of making useful temperature measurements, even in the daytime, whenever Fe densities (at 372 nm) are at least a few thousand atoms per cm^3.

The Polar Regions are more sensitive than elsewhere to global change effects associated with greenhouse gas warming. Profiles of atmospheric parameters and constituents at the geographic poles can provide a convenient means of validating and calibrating global circulation models. However, measurements of key parameters, such as temperature profiles, have only been conducted in the troposphere and lower stratosphere at the geographic poles with balloon borne sensors to altitudes less than 30 km. A major goal of the South Pole observations is to characterize the atmospheric temperature profile from the surface to the edge of space throughout the year. Plotted in Figure 5.26 is a composite temperature profile determined using balloon and Fe/Rayleigh lidar data collected on May 8, 2000, which illustrates this capability. The lower mesosphere is about 20 K warmer, and the upper mesosphere is about 20 K colder than predicted by the MSISE-00 model. By combining the Fe and Rayleigh lidar temperatures with the radiosonde balloon data, the atmospheric temperatures from the surface to 110 km were characterized for the first time at the South Pole throughout the year. The comparisons of the measured monthly mean temperature with the TIME-GCM model and MSIS00 model have revealed several significant differences. More detailed information about these results can be found in Section 5.3.

5.2.3.7. Polar Mesospheric Cloud Detection over Both Poles

Besides the capability of temperature measurements from 30 to 110 km, the Fe Boltzmann temperature lidar is highly capable of detecting PMCs in the polar region. PMCs and their visual counterparts NLCs are the thin water–ice layers

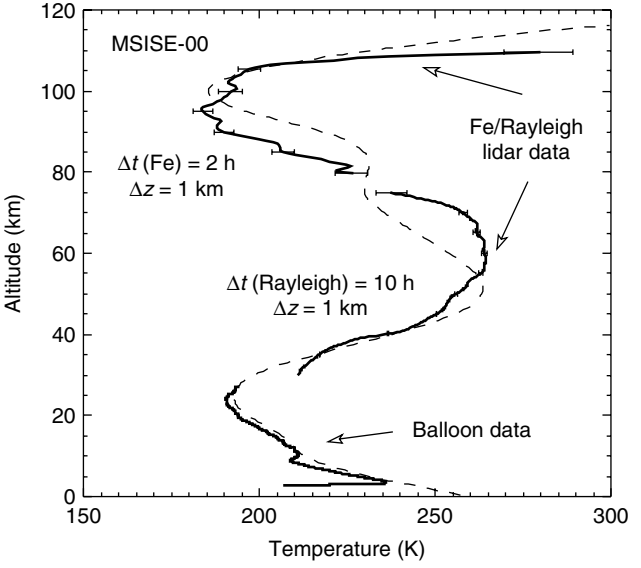


Figure 5.26 Composite temperature profile measured over the South Pole on May 8, 2000 by use of balloon and Fe-Rayleigh lidar data along with the MSISE00 model for May 8, 2000 at the South Pole. (From Chu, X., Pan, W., Papen, G.C., Gardner, C.S., and Gelbwachs, J.A., *Appl. Opt.*, 41, 4400–4410, 2002. With permission.)

occurring in the mesopause region (∼85 km) at high latitudes (∼poleward of 50°) in both hemispheres mainly during the 3 months surrounding summer solstice when the temperatures fall below the frost point (∼150 K). PMCs/NLCs are sensitive tracers of middle and upper atmospheric water vapor and temperature. The observed increases in the geographic extent and brightness of PMCs and NLCs in the Polar Regions during the past few decades are believed to be the evidence of cooling temperature and increasing water-vapor concentration in the mesopause region caused by the increases of greenhouse gases. The possible link of PMCs to the global climate change has intensified the research on PMCs. However, the observations of PMCs are not easy, especially in the Polar Regions due to the continuous 24-h sunlight in the summer. The ground-based visual and photographic observations are limited to twilight conditions, so usually confined within 55°

to 65° latitudes. The space satellites can observe PMCs in the polar summer. Unfortunately, they have limited vertical resolution of about 2 to 3 km. Although *in situ* rocket experiments can measure PMC vertical distribution precisely, they are so sporadic that they cannot be used to characterize the seasonal and diurnal variations in PMC properties.

Lidar technology is unique in detecting PMCs because of its high vertical resolution, daytime measurement capability, and full seasonal and diurnal coverage. As demonstrated at the South Pole, the Fe Boltzmann temperature lidar is a powerful tool for observing PMCs. This is attributed to its daytime measurement capability, its short operating wavelength, and its high transmitted power. PMCs consist of water ice particles. Without PMC particles in the mesopause region, the Rayleigh scattering from pure atmosphere molecules decreases with height exponentially, and becomes negligibly small when reaching 80 to 85 km. When PMC particles are present in the mesopause region around 85 km, the Mie scattering from the PMC water–ice particles, which are much larger than atmosphere molecules, is significantly stronger than the Rayleigh scattering at the same altitude. Thus, a sharp, narrow peak will stand out in the lidar photon count profile. Plotted in Figure 5.27 is an example of a PMC profile obtained by the Fe Boltzmann temperature lidar at the South Pole on January 18, 2000. The photon counts have been converted to the volume backscatter coefficient. The challenge of detecting PMCs in the polar region comes from the high solar scattering background. Only when the PMC peak is higher than the background, it is possible to distinguish PMCs from background noise. The narrowband filters used in the Fe lidar receivers effectively reject the sky background and make it capable of PMC measurements in the daylight. The volume backscatter coefficient of PMCs is proportional to the sixth power of the particle radius while it is inversely proportional to the fourth power of light wavelength. Thus, the shorter wavelength of the Fe lidar at 374 nm will produce much stronger PMC backscattering, compared to normal Rayleigh lidar at 532 nm and the Na lidar at 589 nm. The background level and the PMC signal level together determine the lidar

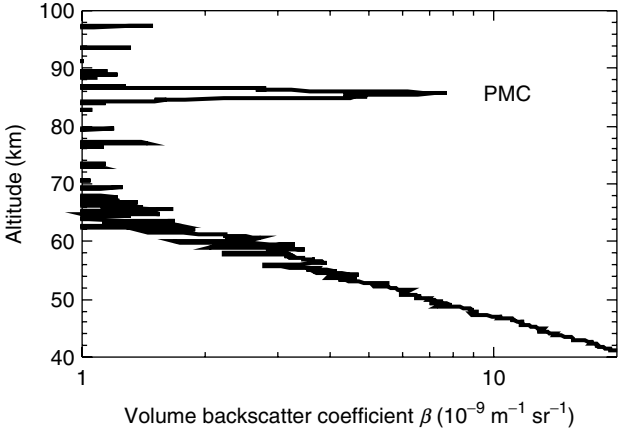


Figure 5.27 Volume backscatter coefficient profile of PMC and Rayleigh scattering of atmosphere molecules versus altitude obtained by the UIUC Fe Boltzmann temperature lidar between 22:38 and 23:00 UT on January 18, 2000 at the South Pole. (From Chu, X., Gardner, C.S., and Papen, G., *Geophys. Res. Lett.*, 28, 1203–1206, 2001a. With permission.)

detection sensitivity. The combination of its daylight measurement capability, short operation wavelength, and high output power ensures that the Fe Boltzmann temperature lidar has very high sensitivity for PMC detection in the polar summer. Extensive data on PMCs have been collected at the South and North Poles with the Fe Boltzmann temperature lidar. Several important scientific results and findings have merged from these data, such as the inter-hemispheric difference in PMC altitude (Chu et al., 2003a). More detailed description can be found in Section 5.3.

5.2.4. Potassium Doppler Lidar

5.2.4.1. Introduction

The geographic and seasonal dependence of the complex thermal structure of the mesopause region is a strong motivating factor for the development of transportable Doppler temperature lidars. The application of ring dye laser that is extremely sensitive to environmental conditions in the Na Doppler tem-

perature lidar precludes its use for transportable systems. Fortunately, there are other neutral metal species besides Na, such as K, Ca, and Fe in the mesopause region. Among these species, the energy level diagram of the K atoms is very similar to that of the Na since both are alkali-metal atoms, and has been well studied (Arimondo et al., 1977). The potassium D_1 and D_2 line wavelengths (769.9 and 766.5 nm) fall in the fundamental gain range of several solid-state lasers including alexandrite, Ti:Sapphire, and Cr:LiSrAlF$_6$. Therefore, it is possible to develop robust Doppler temperature lidars based upon solid-state laser technology using potassium as a tracer. The idea was explored by Papen et al. (1995b) and Höffner and von Zahn (1995) simultaneously. Papen et al. (1995b) did a parametric study on a proposed K lidar system and pointed out that a potassium temperature and wind lidar had the potential for high-resolution atmosphere applications. Their work was mainly based on the K(D_2) line but the analysis for the K (D_1) line was added in their revised manuscript. At the same time, the University of Bonn/Leibniz Institute for Atmospheric Physics (IAP) shifted their alexandrite ring laser wavelength from the Na to the K, and started to construct a narrow-band K Doppler lidar (Höffner and von Zahn, 1995). The first useful K temperature profiles were obtained in 1995 (von Zahn and Höffner, 1996). Subsequently, the IAP K lidar was deployed to Arctic and other locations to study the global atmosphere temperatures. A similar narrowband K lidar based on a ring alexandrite laser was then built at Arecibo Observatory (18.3°N, 66.75°W) and began to make routine observations in the mesopause region (Friedman et al., 2002). In this section, we will discuss the measurement principles, instrumentation, and error analysis based upon the IAP K lidar. Other K lidars will only be mentioned briefly to compare their instrumentation with the IAP K lidar.

5.2.4.2. Measurement Principle

The partial energy levels of the potassium are shown in Figure 5.28. The transition from 4 $^2P_{1/2}$ to 4 $^2S_{1/2}$ is the D_1 line, while the transition from 4 $^2P_{3/2}$ to 4 $^2S_{1/2}$ is the D_2 line.

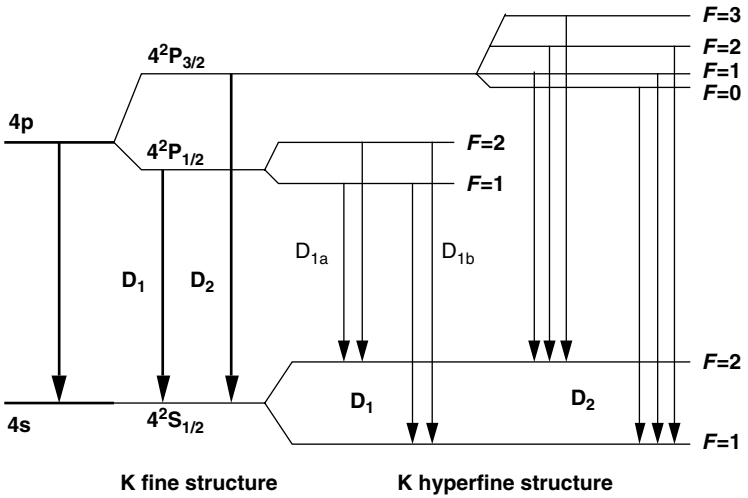


Figure 5.28 Energy level diagram of atomic K. The nuclear spin of K is 3/2.

The K (D_2) line at 766.491 nm cannot be used for ground-based lidar soundings as it is imbedded in a strong O_2 absorption line. Only the K (D_1) line at 769.898 nm can be used for ground-based lidar observations. Although it seems unfortunate that the D_1 resonance cross-section is only half that of the D_2 line, the Hanle effect is zero for D_1 resonance and need not be taken into account in either the experimental arrangement or the analysis.

The potassium D_1 line consists of four hyperfine structure lines in two groups D_{1a} and D_{1b}. However, the D_{1a} and D_{1b} peaks are significantly closer than those of the Na (D_2) line, and they are not resolved at typical mesospheric temperatures. This results in a nearly Gaussian shape for the K (D_1) backscatter cross-section for temperatures above 175 K. At lower temperatures, the K (D_1) lineshape becomes noticeably non-Gaussian. There are three natural isotopes of potassium ^{39}K, ^{40}K and ^{41}K. The abundance of ^{40}K is only 0.0117%, so it is ignored in the K lidar data analysis. The most abundant potassium isotope is ^{39}K with an isotopic abundance of 93.2581%. The presence of the minor isotope ^{41}K with an abundance of 6.7302% slightly modifies the shape of the K

(D_1) fine structure line and must be taken into account during data analysis. With respect to the ^{39}K line, the center of the ^{41}K line is shifted towards shorter wavelengths and the hyperfine structure transitions are somewhat closer together. The calculated backscatter cross sections are shown in Figure 5.29. The quantum number, frequency offsets, and relative line strength are listed in Table 5.8 for both isotopes.

The Doppler broadening of the K (D_1) resonance is temperature dependent. For a natural mixture of ^{39}K and ^{41}K, the variation of the FWHM of the K (D_1) line versus temperature is shown in Figure 5.30. At 200 K, the linewidth is 936 MHz. The mesopause temperature can be derived from the measured linewidth of the Doppler broadened D_1 resonance. To determine the FWHM of the received signal, the resonance

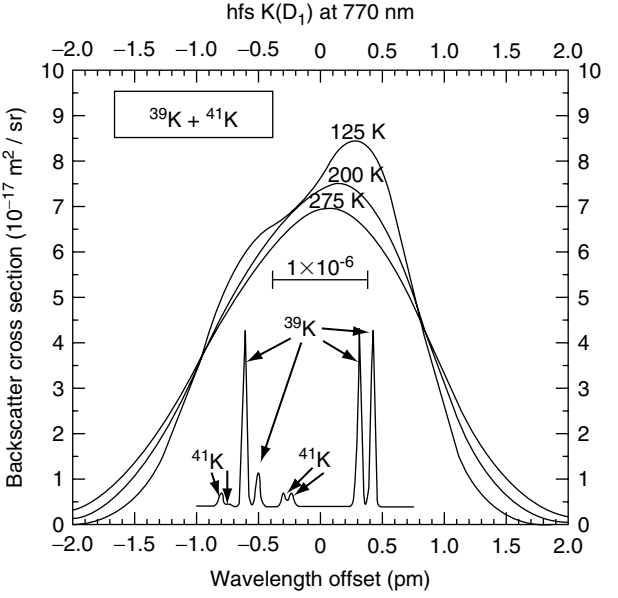


Figure 5.29 Variation of the backscatter cross section of the K (D_1) transition with temperature for $^{39}K + ^{41}K$. The positions and relative strengths of the hyperfine transitions are indicated. The FWHM of these transitions is plotted with the estimated maximum bandwidth of the used laser (20 MHz). (From von Zahn, U., and Höffner, J., *Geophys. Res. Lett.*, 23, 141–144, 1996. With permission.)

Table 5.8 Quantum Numbers, Frequency Offsets, and Relative Line Strength for K (D_1) Hyperfine Structure Lines

$^2S_{1/2}$	$^2P_{1/2}$	^{39}K (MHz)	^{41}K (MHz)	Relative Line Strength
$F=1$	$F=2$	310	405	5/16
	$F=1$	254	375	1/16
$F=2$	$F=2$	−152	151	5/16
	$F=1$	−208	121	5/16

Source: From von Zahn, U., and Höffner, J., *Geophys. Res. Lett.*, 23, 141–144, 1996.

backscatter cross-section, convoluted with the laser bandwidth and assumed Gaussian lineshape, is least-square-fitted to the lidar-induced fluorescence lineshape. The temperature is then derived from the parameters of the fit. By scanning the laser frequency through the K (D_1) line and fitting the theoretical model to the fluorescence signals at each altitude, the range-resolved atmospheric temperature profiles can be derived from the temperature-dependent line shape of the resonance fluorescence signals (von Zahn and Höffner, 1996).

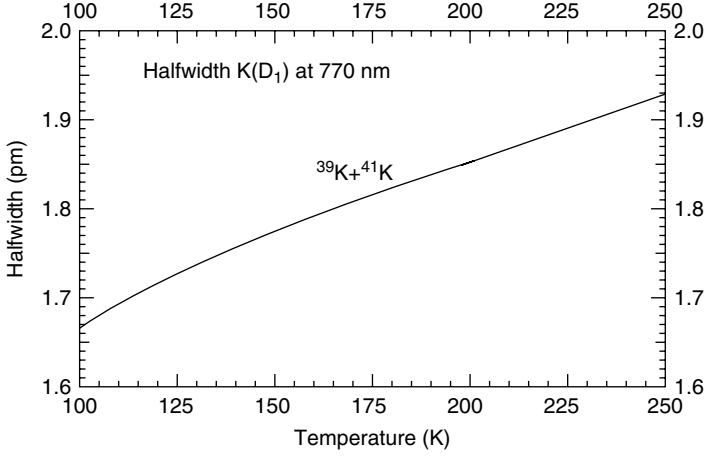


Figure 5.30 Variation of the FWHM of the backscatter cross section of K (D_1) with temperature for $^{39}K + {}^{41}K$. The gradient is about 0.0017 pm/K. (From von Zahn, U., and Höffner, J., *Geophys. Res. Lett.*, 23, 141–144, 1996. With permission.)

This is similar to the original Na temperature system operated by the same group.

5.2.4.3. K Doppler Lidar Instrumentation

To achieve high spectral resolution for the scan, a narrowband laser source is needed. The IAP K lidar consists basically of (1) an alexandrite ring laser, which is injection-seeded by an external cavity diode laser; (2) a high-resolution Fabry–Perot etalon, which is used as both a wavelength meter and a spectrum analyzer to precisely measure the wavelength and the bandwidth of the each transmitted laser pulse; (3) a receiving telescope; and (4) a single photon counting device with computer. The whole system fits in a 20-ft standard container.

The pulsed alexandrite laser is configured in a unidirectional ring cavity that avoids the spatial-hole-burning problem suffered by linear cavities. With the help of injection-seeding from a single-frequency external cavity diode laser, the alexandrite laser can run in a nearly perfect single-longitudinal and single transverse mode (TEM_{00}) with a linewidth around 20 MHz (FWHM). Because of the low gain of the alexandrite crystal, the alexandrite laser pulse has much longer duration time (>200 nsec), compared to the high gain Nd:YAG laser ($\sim$ 7 nsec). This long pulse duration time allows for an ultra-narrow spectral width of the alexandrite laser and also helps avoid the saturation effect of potassium layer in the mesopause region. As a typical operation condition, the IAP K lidar has average output energy of 150 mJ per pulse with duration time of 250 nsec at a repetition rate of 33 Hz. The laser is tunable from pulse to pulse over a range greater than 50 pm. The alexandrite laser frequency is tuned by tuning the wavelength of the injection seed laser.

The seeder (diode laser) wavelength is controlled by a piezo-driven grating. The long-time stabilization of the seed laser is achieved by a lock-in technique with reference to a potassium vapor cell. However, in the IAP K lidar system, the diode laser is not locked to a Doppler-free spectroscopy. In fact, the seed laser is constantly scanned from one edge to the other of the K D_1 line. The alexandrite laser pulses then follow the

seed laser and scan through the potassium D_1 line. The wavelength and bandwidth of each alexandrite laser pulse are measured in real time with a high-resolution etalon. The high-resolution etalon is a Fabry–Perot interferometer with 5.2 cm plate spacing. Its interference fringes are imaged onto a linear array of 1024 photodiodes. The data from these diodes are used to calculate the wavelength of each pulse in real time. The wavelength can be determined with 1 MHz resolution, and the bandwidth can be determined with a resolution of 20 MHz. About 99% of alexandrite laser pulses are in a TEM_{00} single mode and have a spectral width lower than the resolution of the spectrum analyzer (20 MHz).

The receiving telescope for the IAP K lidar uses a primary mirror of 80-cm diameter with the focus being fiber coupled to an optical bench. The single-photon counting system has an altitude resolution of 200 m. Depending on the measured wavelength of the output laser pulse, individual returned photon counts are stored to one of 18 different wavelength channels for later data processing, each being 0.16-pm wide (von Zahn and Höffner, 1996). Typically, the seed and the alexandrite lasers can scan through the entire potassium D_1 line in a few seconds. During a typical integration time of a few minutes, the returned data fill in all 18 channels with more shots towards the edge and less at the center of the D_1 line to compensate the low photon counts at the off-resonance wavelengths.

A narrowband K lidar developed by the Arecibo Observatory also uses an alexandrite ring laser, but the laser frequency is not scanned. Instead, the injection seed laser in the Arecibo system is locked to a Doppler-free feature near the center of the K (D_1) spectrum (Friedman et al., 2002). An AOM is then used to shift the frequency to achieve two wing-frequencies. During observations, the K lidar frequency is shifted between the peak frequency f_0 and the two wing-frequencies $f_\pm$ quickly. By taking the ratio between wing-frequency signals to the peak frequency signals, the atmospheric temperature is then derived from this ratio using the similar three-frequency ratio-technique initially developed for the Na Doppler lidar (She and Yu, 1994; Papen et al., 1995a).

Compared to the IAP K lidar system, the Arecibo K lidar using ratio-technique has the advantage of switching quickly only among three frequencies, so potentially has higher temporal resolution if other conditions are the same. However, the Arecibo system relies on the stability of the seed laser frequency, and the pulsed laser system tends to have a frequency chirp caused by the changing length of laser cavity due to variations of the refraction index of the laser rod during pulse formation. This results in a blue shift of the laser output frequency that can be tens of megahertz. This frequency shift results in a systematic offset in the temperature measurement of approximately 0.3 K/MHz (Friedman et al., 2003b). For typically 40 MHz frequency shift in the Arecibo K lidar, the laser chirp could introduce as large as 10 K uncertainties to the temperature measurements. Currently, the average laser chirp can be measured and removed from the temperature measurements at Arecibo (Friedman et al., 2003b). The IAP K lidar measures the wavelength and spectral width of each laser pulse in real time, so their system does not rely on the seed laser frequency and is immune to chirp influences. By scanning laser frequency through the K (D_1) line, they can derive the temperatures from the fitting parameters. The shortcoming is that the scan takes a longer time than the ratio-technique.

5.2.4.4. Daytime Measurements

To achieve daytime operation, the IAP K lidar employs a twin Faraday anomalous dispersion optical filter (FADOF) and a high-performance photon-counting avalanche photodiode (APD) (instead of a photomultiplier tube) along with reduced field of view (Fricke-Begemann et al., 2002). This improves the signal-to-noise ratio during daytime operations and allows ground-based lidar measurements of upper-atmospheric temperatures in full daylight.

In order to measure temperature accurately, the spectral transmission shape of the daylight filter must be uniform over the $\sim$3 pm range required to scan the K (D_1) line, or its transmission must be measured with high accuracy for correction. With a suitable combination of the magnetic field strength

and vapor cell temperature of the FADOF, the spectral transmission curve of the IAP FADOF is uniform over the K lidar scan range (D_1 line) as shown in Figure 5.31. The full theor-

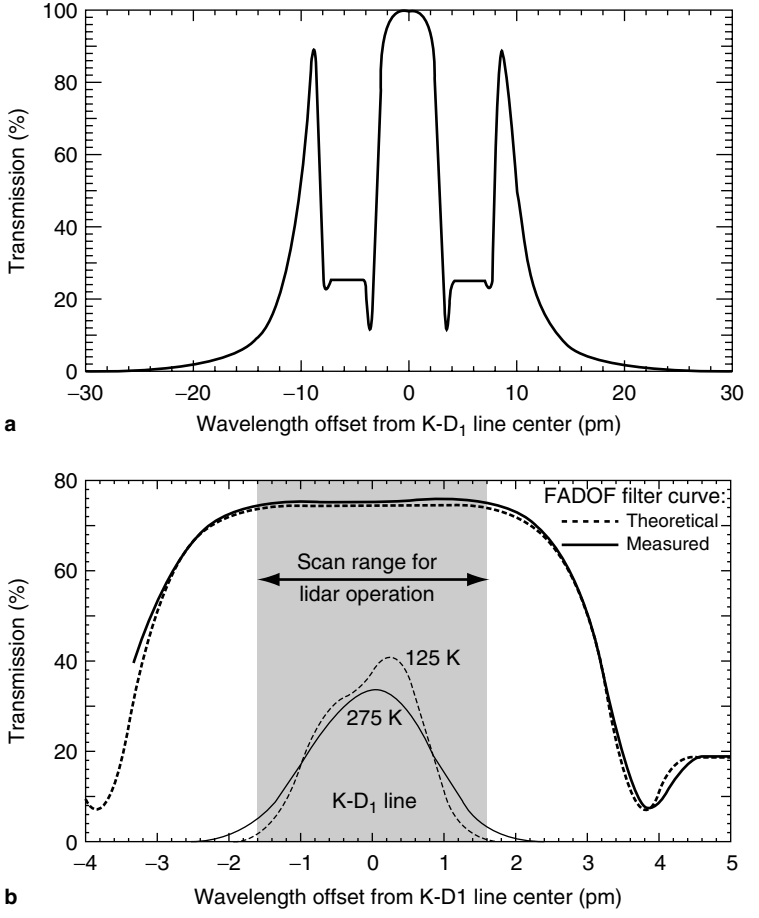


Figure 5.31 Spectral transmission curve of the K FADOF at the K D_1 line (770 nm). (a) Full theoretical shape. (b) Comparison of the theoretical shape of the center peak with the measured transmission shape for the experimental configuration of the IAP K lidar (thick solid curve) and spectral Doppler shape of the K D_1 fine structure line for two different atmospheric temperatures (in arbitrary units). (From Fricke-Begemann, C., Höffner, J., and von Zahn, U., *Geophys. Res. Lett.*, 29 (22), doi: 10.1029/2002GL015578, 2002. With permission.)

etical transmission curve (Figure 5.31a) has a central peak of 6.7 pm (FWHM) and two side peaks of about 4 pm width at 10 pm distance. The central peak has a nearly uniform transmission of more than 99% within ± 1.8 pm around its center. This uniform transmission function avoids the influence of the spectral filter shape on the temperature measurements. The FADOF filter shape is equivalent to that of a Fabry–Perot etalon with 10 pm FWHM, finesse of 50, and peak transmission of 85%. The measured transmission shape for the IAP potassium FADOF is compared with theoretical curve in Figure 5.31(b). The lower absolute value of the measured transmission ($\sim$75%) is due to losses in the optical components. The filter transmission modulation of 1% over the whole scan range requires a small correction of 0.5 K to the temperature. Therefore, any uncertainties in filter-shape determination have only negligible influence on the temperature results. In comparison to Na systems, the two peaks of the K (D_1) line are much closer than corresponding peaks D_{2a} and D_{2b} of the Na (D_2) line. Therefore, much smaller magnetic fields and lower vapor cell temperatures are required to achieve uniform transmission curve for K-FADOF. In fact, uniform transmission is not practical for Na systems. Thus, the nonuniform transmission curve of the Na-FADOF must be measured to high precision for calibration.

The operation of a FADOF requires linearly polarized light. However, because of the nature of resonance fluorescence and the depolarization in the fiber cable between the receiving telescope and the detection bench, the return light of the resonance lidar is unpolarized. Therefore, with a single FADOF, only half of the total backscatter signal can be detected. This can be avoided by using two parallel FADOFs with orthogonal polarizations as shown in Figure 5.32. The unpolarized signal from the receiving telescope is split by a polarizer into two beams with perpendicular polarizations. Each beam passes through one FADAF. After the analyzing polarizer of each FADOF, the two beams are recombined before detection.

With respect to the IAP K lidar configuration used for nighttime measurements, the background must be reduced at least three orders of magnitude to achieve a signal-to-noise

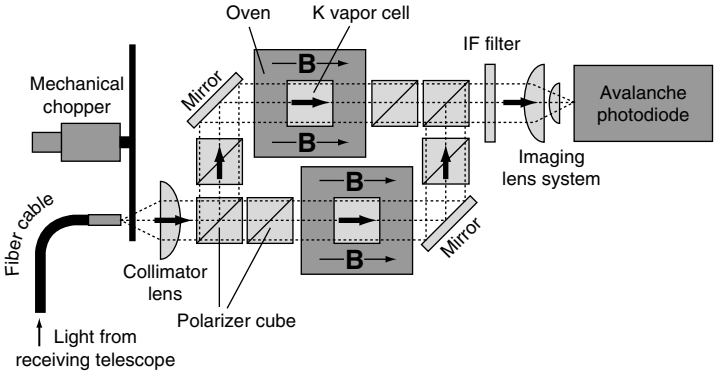


Figure 5.32 IAP K resonance fluorescence lidar detection bench with a twin FADOF system. B represents magnetic field. (From Fricke-Begemann, C., Höffner, J., and von Zahn, U., *Geophys. Res. Lett*., 29 (22), doi: 10.1029/2002GL015578, 2002. With permission.)

ratio (SNR) greater than 1 for daytime operation. The FADOF reduces the solar background by two orders of magnitude, compared with an interference filter with 1 nm spectral width. A further reduction is achieved by reducing the field of view. The field of view of the IAP K lidar can be changed from ~790 to 192 μrad. This reduces the solar background by a factor of 17. The use of a photon-counting APD instead of a PMT also significantly enhances the quantum efficiency (QE) of the receiver. At 770 nm, the APD has QE of 68% while the PMT reaches only 15%. The signal enhancement with the APD by a factor of 4 reduces the statistical error of the lidar retrieval and improves the signal-to-noise ratio by a factor of 2. However, this type of low-noise APD (Perkin-Elmer SPCM-AQ) has a very small active area of only about 170 μm in diameter, requiring careful alignment of the imaging optics in front of the APD for proper operation.

The mobile IAP K Doppler lidar with a twin FADOF, in combination with a reduced field of view and a high-QE APD, allows ground-based lidar measurements of mesopause region temperatures in full daylight. Initial 24-h temperature data have been obtained on the islands of Tenerife (28°N) and Spitsbergen (78°N).

5.2.4.5. Temperature and Error Analysis

Temperatures are derived from the accurately measured shape of the K (D_1) fine-structure line. An example of measured lineshapes is shown in Figure 5.33 for the night of May 4 to 5, 1995 for different altitudes and an integration period of about 3 h. The least-square fits are calculated over altitude range of 2 km, but with 1-km separation from one another. Important parameters of the fits are listed on both sides of

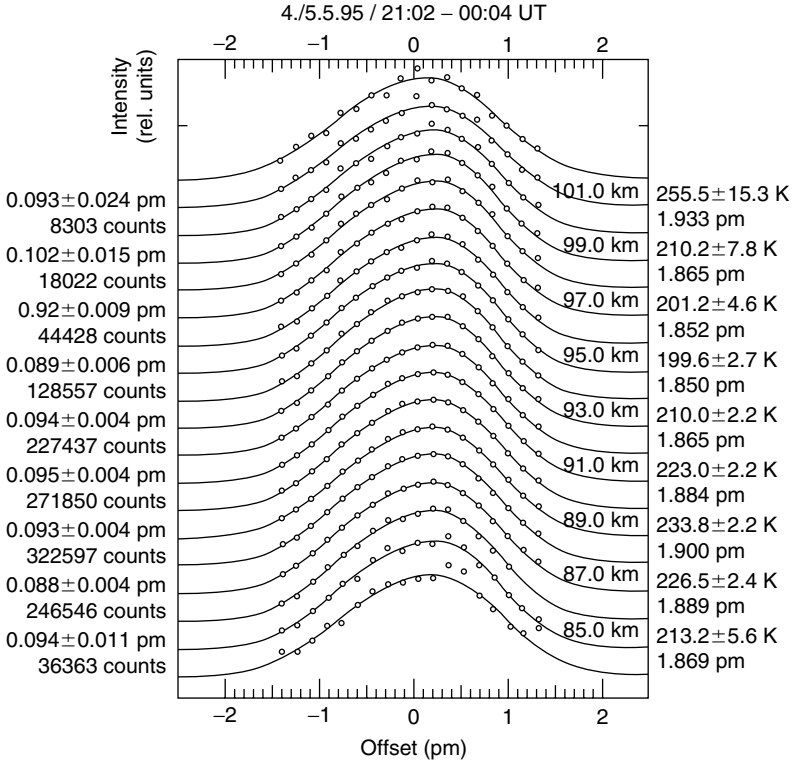


Figure 5.33 Measured number of photon counts (circles) for the 16 wavelength channels and computer fits (solid lines) for the K (D_1) fine-structure line at different altitudes. Information on the numbers on either side of the curves is given in the text. (From von Zahn, U., and Höffner, J., *Geophys. Res. Lett.*, 23, 141–144, 1996. With permission.)

each profile. On the left side, the upper values are the offsets of the line maxima from an arbitrary wavelength "zero." The lower ones are the total photon counts over the entire line. On the right side, the upper value gives the calculated temperature and its error. Errors are one-sigma errors of the least-square-fit excluding systematic errors. At the peak of the K layer, the error is about 2.2 K. The largest potential source for systematic errors is the accuracy of the wavemeter calibration used in the IAP K lidar (estimated contribution about ± 5 K). Due to the ultra-narrow laser bandwidth, the systematic error stemming from the convolution of laser bandwidth with the potassium backscatter cross-section is small, being less than 0.5 K. As von Zahn and Höffner (1996) pointed out, the Hanle effect of the K (D_1) resonance does not occur for linearly polarized laser light. Thus, the backscattered intensity is independent of the geomagnetic field.

Plotted in Figure 5.34 is an example of a measured temperature profile in the mesopause region. Both K isotopes have been taken into account for this temperature calculation. The data were acquired late in the night of May 4 to 5, 1995 and integrated over 3 h. The mesopause has a temperature of 200 K at an altitude of 96 km. In these observations, corrections of the calculated temperature due to saturation effects in the K layer are negligible. This is demonstrated in Figure 5.35 that shows the degree of saturation depending on the divergence of a 100 mJ laser beam. For the parameters of the IAP K lidar (beam divergence of 0.4 mrad and pulse duration time of 275 ns), the saturation is less than 0.2% for one of the hyperfine lines. Computer simulations show that even a saturation of 10% would change the temperature only about 0.2 K. The reason for such a small sensitivity of the calculated temperature to the saturation effect is because that the all three strongest hyperfine structure lines have the same line strength and therefore always exhibit the same degree of saturation. Only one hyperfine structure transition has a much smaller line strength, but since its spectral position is between the strong lines and because of its small line strength, it has only a marginal influence on the overall line shape and hence temperature determination (von Zahn and Höffner, 1996). For

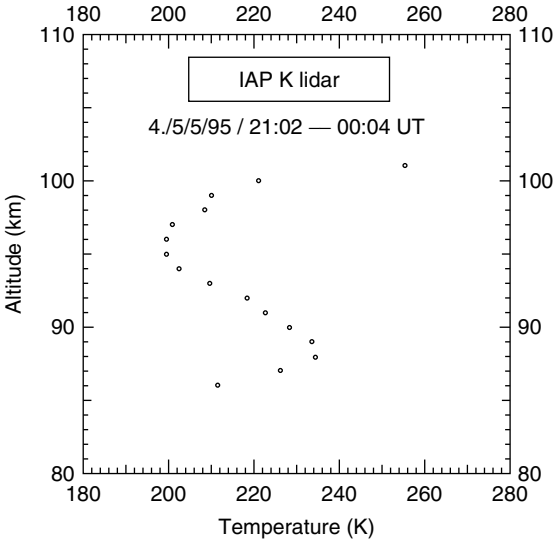


Figure 5.34 Typical temperature profile, integrated over 3 hours on May 4, 1995 for $^{39}K + {}^{41}K$. The statistical 1σ-error at the maximum of the potassium layer is ± 2 K. (From von Zahn, U., and Höffner, J., *Geophys. Res. Lett.*, 23, 141–144, 1996. With permission.)

daytime observations, the lidar beam with a reduced divergence could cause saturation in K layer. Because saturation is more pronounced at the center of the K spectrum, it distorts the K lineshape. By scanning over a broad spectral range of the K (D_1) line, the IAP K lidar can actually detect this distortion effect on the lineshape and measure the saturation effect. In practice, the IAP K lidar runs at a relative lower power during daylight in order to avoid the saturation (Höffner, personal communication, 2003). For the daytime observations by the IAP K lidar, the systematic errors are slightly larger than in the nighttime. The statistical error per hour of integration can range from 1 to 20 K depending on the highly variable potassium densities at different locations and times (Höffner, personal communication, 2003).

When Papen et al. (1995b) analyzed the narrowband K lidar, the K abundance used in the calculation was taken from the measurement results of Megie et al. (1978), in which the Na to K abundance ratio was about 15. Assuming an 8-W

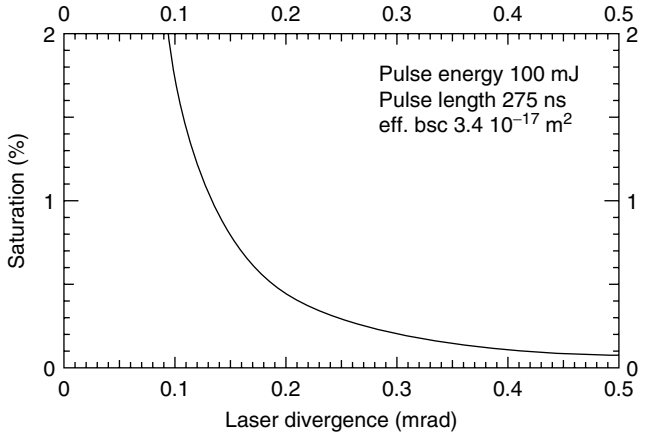


Figure 5.35 Dependence of the degree of saturation (in %) on the laser beam divergence for one of the strong hyperfine structure lines of ^{39}K, calculated for a laser pulse of 100 mJ energy and 275 nsec duration, and K layer height of 90 km. (From von Zahn, U., and Höffner, J., *Geophys. Res. Lett.*, 23, 141–144, 1996. With permission.)

766-nm K lidar, a 1-W 589-nm Na lidar and a ratio of Na to K density being 15, Papen et al. (1995b) concluded that the temperature performance of a K (D_1) lidar would be 37% of the Na lidar, while the K return photon counts would be about 55% of the Na counts. However, the IAP K lidar measurements show that the K abundance is about 80 times less than the Na abundance, at least five times lower than the number reported by Megie et al. (1978). This low K abundance has been confirmed by other K lidar measurements (Friedman et al., 2002) and twilight airglow measurements (Sullivan and Hunten, 1964). This unexpected low K density reduces the temporal and spatial resolution of the K lidar, and dramatically increases the integration time required for obtaining accurate temperature measurements. The higher pulse energy of the K lidar compared to the Na lidar can partially compensate the negative effects of the small K density. Although it is possible to measure the wind in the mesopause region using the K lidar as analyzed by Papen et al. (1995b), this has not been achieved to date because of the extremely low K density in the mesosphere.

5.2.5. Solid-State Na Doppler Lidar

5.2.5.1. Introduction

Three kinds of resonance fluorescence lidars have been developed so far for measuring the mesopause region temperatures: the Na Doppler technique, the K Doppler technique, and the Fe Boltzmann technique. Among them, the narrowband Na lidar currently provides the highest resolution and most accurate temperature measurements. This is because the combination of relatively high Na densities and large backscatter cross-section yields strong signals even with a relatively modest telescope of a few tens of centimeters in diameter. Unfortunately, the conventional sophisticated Na Doppler lidars require a well-controlled laboratory environment to maintain proper operation conditions, which precludes their deployment in aircrafts or at remote sites. It would be attractive to develop a narrowband Na lidar with all-solid-state lasers to take the advantages of both the Na lidar and the solid-state laser technology. Such a lidar would be robust and easy to operate for deploying to different geographical locations to study the global atmosphere. Three models of solid-state Na lidars have been tested so far: the sum frequency generation of two pulsed Nd:YAG lasers (Jeys et al., 1989; Chiu et al., 1994; Kawahara et al., 2002), the Raman-shifted and frequency-doubled alexandrite ring laser (Schmitz et al., 1995), and the sum-frequency generation of two cw Nd:YAG lasers followed by pulse amplification (Vance et al., 1998; She et al., 2002b). The two models with Nd:YAG lasers have been developed into practical lidar systems and performed temperature measurements in polar regions, while the alexandrite laser has been converted to a K lidar. Here we will briefly introduce the ideas of these solid-state Na lidar systems.

5.2.5.2. All-Solid-State Na Temperature Lidar

An interesting coincidence of nature is that the sum frequency of two appropriately tuned Nd:YAG lasers operating near 1064.591 and 1319.250 nm may be made resonant with the

sodium D_2 transition wavelength at 589.1583 nm (Jeys et al., 1989). Shown in Figure 5.36 is a diagram of energy levels used in Nd:YAG lasers. The strong emission of 1064 nm is the lasing transition from the upper level $^4F_{3/2}$ to the lower level $^4I_{11/2}$ level. The weaker transition of 1319 nm is from the upper level $^4F_{3/2}$ to the lower level $^4I_{13/2}$ level. Jeys et al. (1989) were the first to obtain the sodium resonance radiation by sum-frequency mixing of these two Nd:YAG laser lines in $LiNbO_3$ crystal, although the 589 nm radiation was broadband with spectral width of about 2 GHz and contained multimodes. They made the lidar observations of the Na layer in the mesopause region using this 589 nm radiation and obtained sodium density profiles.

However, to make temperature measurements, a narrowband radiation at the sodium D_2 resonance is needed. The first all-solid-state single-frequency sodium resonance radiation was obtained by Chiu et al. (1994) using the sum-frequency mixing of the output from two injection-seeded pulsed Nd:YAG lasers. The overall layout of the Chiu et al. laser system is shown in Figure 5.37. The 1.319 μm Nd:YAG laser was configured as an injection-seeded, Q-switched mas-

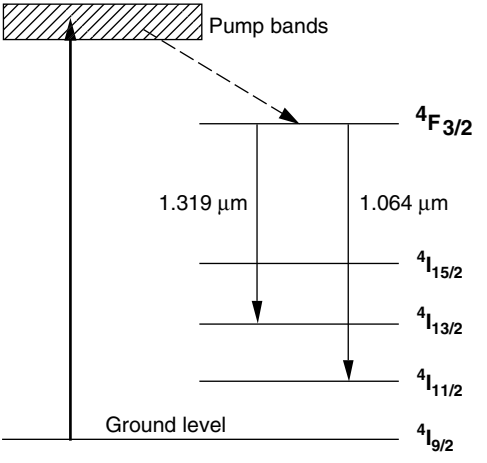


Figure 5.36 Energy level diagram for Nd:YAG laser.

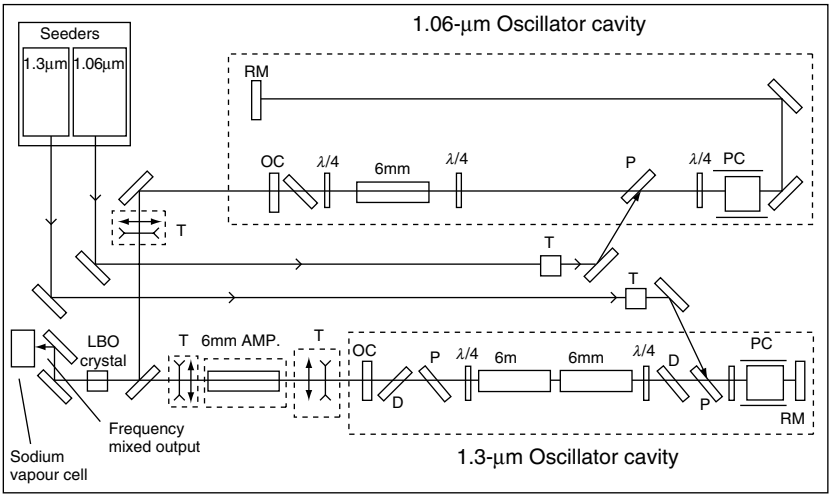


Figure 5.37 Layout of the laser system for all-solid-state single-mode sum-frequency generation of Na resonance radiation. RMs, rear mirrors; OCs, output couplers; PCs, Pockels cells; Ts, telescopes; Ps, polarizers; Ds, dichroic mirrors. (From Chiu, P.H., Magana, A., and Davis, J., *Opt. Lett.*, 19, 2116–2118, 1994. With permission.)

ter oscillator followed by a power amplifier. The 1.064 μm Q-switched master oscillator has a typical folded, long-pulse cavity design. The 1.064 and the 1.319 μm beams were combined on a common beam path by a dichroic mirror and then passed through a 5 mm × 5 mm × 15 mm noncritically phase-matched, temperature-tuned LBO type I crystal to generate the 589 nm radiation. Both the 1.064 and the 1.319 μm laser oscillators are well synchronized with temporal overlap adjustable to better than half a nanosecond. With input energies of 85 mJ at 1319 nm and 180 mJ at 1064 nm, the 589-nm radiation was generated as 65 mJ per pulse at a repetition rate of 20 Hz. The 589-nm radiation has a pulse duration time of 18 to 20 nsec. The measured linewidth of the 589-nm radiation was less than 100 MHz (FWHM) and the frequency stability was better than 100 MHz/h (Chiu et al., 1994). The injection seeders used in this laser system were the diode-pumped single-frequency cw Nd:YAG lasers with a continuous

tuning range of more than 10 GHz. The seeder frequency can be tuned by adjusting the laser-crystal temperature with a Peltier thermoelectric cooler or by stressing the laser-crystal with a piezoelectric transducer bonded to the crystal. The 589-nm radiation frequency can be tuned over a 4 to 5 GHz range with a near-Gaussian beam profile in the near and far fields (Chiu et al., 1994).

No lidar observations of the mesopause region were ever reported with the Chiu et al. laser system. The Shishu University of Japan adopted the Chiu et al. laser techniques and developed a Na temperature/density lidar transmitter using the sum-frequency generation of two injection-seeded single-frequency pulsed Nd:YAG lasers (Kawahara et al., 2002). A different nonlinear crystal (BBO) was used in the Shishu University system, compared to the LBO crystal used in the Chiu et al. system. With input energies of 100 mJ at 1319 nm and 150 mJ at 1064 nm, Kawahara et al. generated the 589-nm laser pulses with energy of $\sim$40 mJ per pulse at a repetition rate of 10 Hz. The pulse duration time (28 nsec) is longer compared with the conventional Na lidar (about 7 nsec). This helps avoid the sodium layer saturation problem. The 589-nm laser bandwidth is estimated to be about 40 MHz. By tuning the injection-seeding laser wavelengths through controlling the temperature of seed laser crystal, the sum laser frequency can be tuned through sodium D_2 resonance line. The tuning error is within 100 MHz. During the operation, the 1319-nm seed laser wavelength was held at 1319.2012 nm while the 1064-nm seeder wavelength was tuned between 1064.6221 and 1064.6187 nm by adjusting crystal temperatures. The corresponding wavelength shift of mixing laser pulses was between 589.1589 nm (near sodium D_{2a} peak) and 589.1578 nm (near the minimum between D_{2a} and D_{2b}). The photon count ratio at these two wavelengths was then converted to temperatures using the two-frequency technique as described by She et al. (1990). The measured-temperature error due to the locking wavelength uncertainty is estimated to be ± 1 K.

Besides the 589-nm lidar beam, half of the output energy of the 1064-nm laser passed through a KDP crystal to produce

laser pulses at 532 nm for Rayleigh temperature observations. This all-solid-state narrowband Na/Rayleigh temperature lidar was deployed to the Japanese Syowa Station (69°S, 39°E), Antarctica in the austral winters of 2000 to 2002. The lidar beams were pointed to zenith direction and the backscatter photons were collected by a 0.5-m diameter Dall–Kirkham Cassegrain telescope. The collected 589 and 532-nm signals were optically separated and then detected by two different PMTs. The lidar observations were limited to the nighttime periods from March to October. Even so, the Na/Rayleigh lidar has collected important temperature data in Antarctica and made significant contributions in characterizing the thermal structure of polar mesopause region and understanding the atmosphere dynamics (Kawahara et al., 2004).

A few nonlinear crystals have been used for this frequency mixing: $LiNbO_3$ (Jeys et al., 1989), LBO (Chiu et al., 1994), and BBO (Kawahara et al., 2002). Compared with $LiNbO_3$ crystal, the LBO and BBO crystals have much higher damage thresholds and larger phase-matching temperature range: about 4 to 5°C near 50°C for LBO and near 40°C for BBO compared to the 0.1°C range in the vicinity of 224°C for $LiNbO_3$. However, the mixing technique has a critical requirement on the spatial overlap of two Nd:YAG laser beams since the spatial overlap has a dramatic influence on the mixing conversion efficiency. The changes of seed laser alignment usually change the spatial overlap. In addition, the pulsed 589-nm laser is difficult to be locked to sodium D_2 Doppler-free spectrum, which results in the lack of absolute frequency reference for wind measurements. Therefore, it is unlikely that this all-solid-state narrowband Na lidar can be applied to mesospheric wind measurements.

Another idea to make all-solid-state Na lidar was tested on ring alexandrite lasers (Schmitz et al., 1995). The output of the alexandrite laser working at 791 nm can be Raman-shifted to 1178 nm (first Stokes line) using a H_2 Raman cell, and then frequency doubled to 589 nm. An alternate method is to use the first anti-Stokes line emerging from the H_2 Raman cell to convert the alexandrite laser to the 589 nm directly. However,

a few serious shortcomings, such as the low energy conversion rate, low spectra purity, sensitive to environmental pressure and temperature, and a broadening of the laser spectral width, prevent this technique from being practical for the Na lidar applications. This idea has been abandoned, and the ring alexandrite laser has been converted to 770 nm for the K lidar applications as described in the Section 5.2.4.

5.2.5.3. Solid-State Na Wind/Temperature Lidar

Another idea to make solid-state Na lidar is to utilize the sum-frequency mixing of two cw single-mode Nd:YAG lasers operating at 1064 and 1319 nm to generate the 589-nm radiation in order to replace the ring dye master oscillator in the conventional Na wind/temperature lidar (Vance et al., 1998). The sum-frequency mixing of two cw single-mode Nd:YAG lasers in a doubly resonant congruent lithium niobate resonator generated TEM_{00} beam of single-frequency 589-nm radiation with power up to 400 mW as demonstrated in Vance et al. (1998). The two Nd:YAG lasers are based on a monolithic diode-pumped Nd:YAG ring cavity and incorporate both a relatively fast piezo-tuning mechanism (a 30-kHz response bandwidth) with a 30-MHz tuning range and a slow (1 GHz/s) thermal tuning mechanism with a tuning range of 60 GHz. Their narrow linewidth (5 kHz over 1 msec), TEM_{00} spatial mode, and low amplitude noise (0.05% rms between 10 Hz and 100 MHz) combined with their tenability makes them well suited for the generation of narrowband tunable radiation near 589 nm. The cw 589-nm beam still needs to be amplified by a PDA, which is pumped by an injection-seeded frequency-doubled pulsed Nd:YAG laser, to produce high peak power laser pulses for the lidar applications. Since the sum-frequency generated cw 589-nm radiation can be easily locked to the Doppler-free features of the sodium D_2 line using a Na vapor cell, the absolute frequency reference is available for the wind measurements. Therefore, it is still a Doppler wind and temperature lidar, and can make both temperature and wind measurements in the mesopause region.

Such a solid-state Na wind/temperature lidar has been developed by the Colorado State University in collaboration with the Colorado Research Associates (She et al., 2002b). This lidar is called the Weber lidar. Its transmitter consists of a sum-frequency generator (SFG) producing 100 to 200 mW cw tunable single-frequency light at 589 nm, an AOM that up- or down-shifts the SFG output by 630 MHz on command, and a PDA that is pumped by a 20 W frequency-doubled Nd:YAG laser working at a repetition rate of 50 Hz. The output 589-nm pulses have a bandwidth of 120 MHz and a power about 1 to 1.5 W. The solid-state master laser is much more robust than the ring dye laser, and is easy to operate once the alignment of frequency mixing is done. The Weber Na wind/temperature lidar has been successfully deployed to ALOMAR at Andoya, Norway (69°N, 16°E) (She et al., 2002b). With the large steerable twin telescopes at ALOMAR and demonstrated line-of-sight wind and dual-beam momentum flux measurement capability, our understanding of gravity wave dynamics in the polar region will be much enhanced by the Weber lidar.

To obtain an all-solid-state Na wind/temperature lidar, an option is to replace the PDA by an optical parametric amplifier (OPA) (She, private communication, 2003). This all-solid-state Na lidar can make both wind and temperature measurements with comparable accuracy and resolution to the conventional Na wind/temperature lidar. If the OPA could slow down the pulse (i.e., increase the pulse duration time), it would also help avoid the Na layer saturation problem.

5.2.6. Comparison of Na, Fe, K, and Rayleigh Lidar Techniques

As of today, four lidar techniques have been developed to measure temperature profiles in the middle and upper atmosphere: the narrowband Na Doppler lidar technique, the broadband Fe Boltzmann lidar technique, the narrowband K Doppler lidar technique, and the Rayleigh lidar technique. It will be worthwhile to compare their measurement capabilities and technical challenges for the applications in the atmosphere observations, as shown in Table 5.9.

Table 5.9 Comparison of Na, Fe, and K Lidar Performances

	Na	Fe	K
Wavelength in vacuum(nm)	589.158	372.099; 373.819	770.109
Peak backscatter cross section (m^2) at 200 K	15.0×10^{-16} Na (D_2)	0.94×10^{-16} (372 nm) 0.87×10^{-16} (374 nm)	13.4×10^{-16} K (D_1)
Abundance (cm^{-2})	5×10^9	10×10^9	6×10^7
Laser power (W)	1	3 (×2)	3
Laser bandwidth FWHM (MHz)	140	800–1000	20
Pulse duration time (nsec)	6–7	60–100	250–275
Rayleigh range (km)	Up to 50	Up to 80	Up to 50
Daytime filter	Na-FADOF Fabry–Perot etalon	Fabry–Perot Etalon	K-FADOF

5.2.6.1. Narrowband Na Doppler Lidar Technique

The Na Doppler technique currently provides the highest resolution and most accurate temperature measurements of any system in the mesopause region. The combination of relatively high Na densities and large backscatter cross-section yields strong signals even with a relatively modest telescope of a few tens of centimeters in diameter. Besides temperature, the Na Doppler lidar is also capable of wind measurements in the mesopause region, as demonstrated by the UIUC and the CSU group. Because of the application of Na FADOF, which effectively rejects the sky scattering background and has absolute frequency calibration, the Na Doppler lidar is capable of both temperature and wind measurements in the mesopause region in both day and night, as achieved by the CSU group. However, to avoid the Na layer saturation effect, the transmitted laser power is limited to low power (1 to 2 W). Combined with its relative longer wavelength at 589 nm, the molecular (Rayleigh) scattering signal from Na lidars can only

be used to derive Rayleigh temperatures to altitudes of 50 to 60 km with poor temporal resolution.

For the conventional Na wind/temperature lidar, the sophisticated laser technology requires a well-controlled laboratory environment to maintain proper operating conditions. Na systems cannot be easily deployed at remote sites and cannot be used for airborne observations. The full solid-state narrowband Na Doppler lidar has proved to be rugged enough for deployment in the Antarctica. However, it is unlikely to make wind measurements due to the lack of absolute frequency reference. The solidified Na wind/temperature lidar developed by CSU has made reliable temperature and wind measurements in Arctic. Once the frequency mixing is made more robust and reliable, this would be a good system for exploring the middle and upper atmosphere. However, this is still a complicated and sensitive system. It would be difficult to be deployed in the aircraft.

5.2.6.2. Fe Boltzmann Temperature Lidar Technique

The Fe Boltzmann temperature lidar does not require the narrowband lasers, but is comfortable with the broadband lasers with linewidth of about 1 GHz. The Fe Boltzmann lidar based upon the solid-state alexandrite lasers is rugged enough to be deployed at remote sites and in the aircrafts. The Fe Boltzmann temperature technique has higher sensitivity compared to the Na Doppler technique, as shown in Figure 5.21. Although limited by the lower Fe density on the 374 nm channel, the Fe Boltzmann lidar can still provide relatively accurate temperature measurements in the mesopause region. Because of its shorter wavelength and higher output power, the Fe lidar has much stronger molecular scattering and so the Rayleigh temperature measurement can be extended to the range 75 to 80 km. The combination of the Boltzmann technique and Rayleigh technique for the Fe Boltzmann lidar provides exceptionally wide altitude coverage for temperature measurements from 30 to 110 km. It is also capable of daytime temperature measurements. In

addition, the Fe Boltzmann lidar has been proved to be a great tool for detecting the polar mesospheric clouds in the polar summer under 24-h sunlight. The combination of its daylight measurement capability, short operation wavelength, and high output power ensures that the Fe Boltzmann temperature lidar has very high sensitivity for PMC detection. Extensive data on PMC have been collected at the South and North Poles with the Fe Boltzmann temperature lidar. Several important scientific results and findings have merged from these data.

Due to the use of the broadband laser, the Fe Boltzmann lidar is not capable of wind measurements in the mesopause region. The biggest drawback to the Fe Boltzmann technique is the fact that the system is actually two complete lidar systems operating at 372 and 374 nm. The signal levels on the weak 374 nm channel limit the performance of the system. Typical Fe densities in this ground state at the peak of the layer vary from about 50 to 300 cm^{-3} so that daytime observations are doable but challenging and require long integration times. The observations conducted near summer solstice over the North and South Poles when the Fe densities are minimum and the background noise from the daytime sky is maximum, provided the toughest environmental tests of the instrument. In these first field campaigns, the impressive measurement capabilities of the Fe lidar have been demonstrated under the most extreme operating conditions.

5.2.6.3. Narrowband K Doppler Lidar Technique

The narrowband K Doppler lidars based upon solid-state alexandrite lasers operating in a pulsed ring configuration have been shown to be rugged enough to be deployed on ships and at remote sites to make temperature measurements. The low natural K abundance is partially compensated by the high detection efficiency of the avalanche photodiode and the high output power of the alexandrite laser. The development of twin K FADOFs enables the daytime temperature measurements. Therefore, as demonstrated by the IAP group, the K Doppler

lidar has provided high quality temperature measurements in the mesopause region. In principle, the narrowband K Doppler lidar could make wind measurements. However, no wind measurement have been reported so far. Because the K resonance line is in the near IR region of the spectrum, the molecular scattering is weak and so the Rayleigh temperature capabilities of K lidars are limited to about 50 km.

5.2.6.4. The Rayleigh Lidar Technique

The Rayleigh technique, which depends on the relatively weak molecular scattering, requires a power aperture product of several hundred W m^2 to make useful temperature measurements at altitudes near 100 km. This requires 4-m class telescopes and 10 to 20 W lasers. While achievable, it is unlikely these systems will be deployed to remote locations like the South Pole anytime soon, and they certainly cannot be deployed on aircraft. However, its simplicity makes the Rayleigh technique easy to be built and used for atmosphere temperature and gravity wave study below 80 km, even with moderate laser power and telescope size.

5.3 KEY RESULTS OF LIDAR MEASUREMENTS IN THE MIDDLE AND UPPER ATMOSPHERE

The advanced resonance fluorescence lidar technologies have dramatically enhanced human's capabilities in the measurements of the middle and upper atmosphere, such as the temperature, wind, meteor, and polar mesospheric clouds in the mesopause region. These lidar measurements enable the systematic study of the atmospheric thermal structure, dynamics, and composition, as well as their relations to the global climate change. Significant progress has been made in these areas since the 1980s. It is impossible to cover all these results in this brief summary section. We will highlight several key results obtained by the resonance fluorescence lidars that are unique and important to the study of the middle and upper atmosphere.

5.3.1 Thermal Structure of the Middle and Upper Atmosphere

Temperature is one of the fundamentally important parameters for understanding the global atmosphere and long-term climate change. The range-resolved temperature profiles in the middle and upper atmosphere are one of the key results obtained by the advanced resonance fluorescence lidars. These lidar measurements provide, for the first time, the insight of the global thermal structure from the North Pole to the South Pole.

Many radiative, chemical and dynamic processes play important roles in establishing the thermal structure of the middle and upper atmosphere (Garcia, 1989; Roble, 1995). The mean background temperature is largely determined by the radiative forcing, which mainly includes the solar heating associated with absorption of solar UV radiation by O_3 and O_2 (Mlynczak and Solomon, 1993), and radiative cooling associated with infrared (IR) emissions of CO_2 (Andrews et al., 1987; Rodgers et al., 1992). The chemical heating from exothermic reactions contributes energy to the middle atmosphere (Mlynczak and Solomon, 1991; Reise et al., 1994) while energy can be lost from the atmosphere by airglow from excited photolysis products or by chemiluminescent emission from product species of exothermic chemical reactions. The middle atmosphere thermal structure is also strongly influenced by dynamic forcing. Lindzen (1981) pointed out that the momentum deposited by breaking gravity waves in the mesopause region slows down or reverses the zonal winds, resulting in a strong mean meridional circulation. Associated with this single-cell meridional circulation is strong upwelling at the summer pole, strong downwelling at the winter pole, and a meridional flow from summer pole to winter pole elsewhere. These vertical motions produce significant departures from radiative equilibrium through adiabatic heating and cooling (Walterscheid, 1981; Garcia and Solomon, 1985; Fritts and van Zandt, 1993; Fritts and Luo, 1995). The measurements of background thermal structure and dynamic behavior of the middle and upper atmosphere are essential for developing

accurate atmosphere general circulation models, and help understand more about the energy budget in this region.

Since the first range-resolved temperature profile through the mesopause region was made by the first practical narrowband Na Doppler lidar (Fricke and von Zahn, 1985), the thermal structures of the middle and upper atmosphere have been characterized by the lidars in the polar regions (Neuber et al., 1988; Lübken and von Zahn, 1991; von Zahn et al., 1996; Gardner et al., 2001; Chu et al., 2002a; Pan et al., 2002; Kawahara et al., 2002; Pan and Gardner, 2003), at the mid-latitudes (She et al., 1993, 2000, 2002a; Senft et al., 1994; Leblanc et al., 1998; Chen et al., 2000; States and Gardner, 2000a,b), and in the low-latitudes (Fricke-Begemann et al., 2002; Friedman, 2003a). The comparisons of the measurement results with the theoretical models improved our understanding of the middle and upper atmosphere, their relations to the global climate change, and their relations to the solar cycle effect (She et al., 2002c; She and Krueger, 2004).

5.3.1.1. Thermal Structure at Polar Latitudes

The middle and upper atmosphere at polar latitudes is a very interesting region of the global atmosphere. At the polar latitudes, the mesopause is coldest in summer and warmest in winter. The lowest temperatures of the entire terrestrial atmosphere are found at the polar summer mesopause. It also contains different layered phenomena, such as the PMCs, sporadic metal layers, and airglow layers. It is also a region of enhanced NO production through auroral processes. In particular, during winter, the NO can reach the stratosphere and act on the upper ozone layer. All these layers may or may not be related to the thermal structure of the atmosphere. In order to reach a deeper understanding of the processes acting in this interesting region, it is important to improve our knowledge of the thermal structure of the polar middle and upper atmosphere.

The University of Bonn Na Doppler lidar (Fricke and von Zahn, 1985) made nighttime temperature measurements at

Andoya, Norway (69°N, 16°E) in the winter conditions (Neuber et al., 1988) and late summer conditions (Kurzawa and von Zahn, 1990). Combining Na lidar technique with three other techniques: the passive falling sphere technique, the rocketborne mass spectrometer, and the rocketborne ionization gauge, Lübken and von Zahn (1991) presented the first standard atmosphere thermal structure of the mesopause region at 69°N based upon the 10-year data measured on 180 days. Most of the profiles were measured during winter and summer, few in early spring and during autumn. No lidar temperature measurements were available from May through July due to the permanent daylight during this period at Andoya, Norway. The altitude range of the temperature profiles was from 50 to 120 km. In total, nine monthly mean temperature profiles were obtained, as shown in Figure 5.38. The seasonal variations of the mesopause temperatures and mesopause altitudes are shown in Figure 5.39 and Figure 5.40. The polar mesopause is much colder and is at a lower altitude in summer than in winter, a fact known at least qualitatively since the pioneering work by Stroud et al. (1959). Lübken and von Zahn (1991) found the mesopause at approximately 98 km from October to March and at 88 km in the summer months. Both the mesopause temperature and mesopause altitude essentially had two values only: a "winter" value and a "summer" value. For the temperature, there was an extended period from September until April where the mesopause temperature was around 190 K and a shorter one during summer (end of May until middle of August) where the mesopause was much colder (typically 130 K). Lübken and von Zahn suggested that in the mesosphere the transition between winter and summer conditions (and vice versa) is faster than commonly anticipated. Lübken and von Zahn also compared the measured thermal structure with an empirical reference atmosphere, the COSPAR International Reference Atmosphere (CIRA) (Fleming et al., 1990), and found significant deviations. The measured temperature in summer is significantly lower than the CIRA temperature during the same period, but in winter the measured temperature in lower mesosphere is up to 20 K warmer than that of the CIRA. The

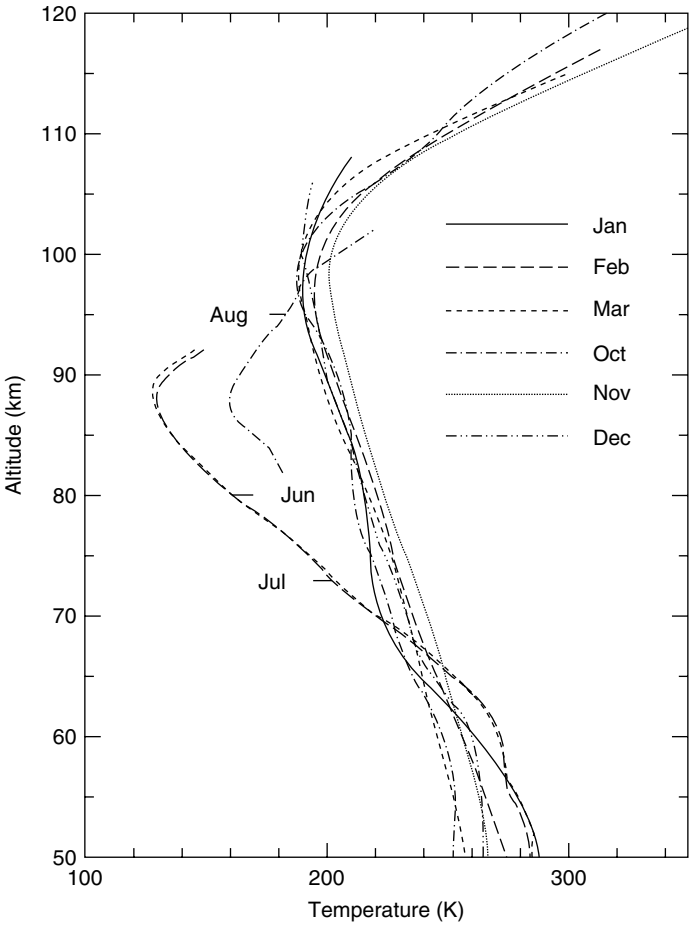


Figure 5.38 Smoothed monthly mean temperature profiles obtained by ground-based lidar and *in situ* density measurements at Andoya, Norway (69°N). (From Lübken, F.J., and von Zahn, U., *J. Geophys. Res*., 96 (D11), 20841–20857, 1991. With permission.)

mean static stability was found to be 3.5 and 7.9 K/km in summer and winter, respectively. Lübken and von Zahn (1991) also found statistically significant (95% confidence level) positive correlations between solar flux and Na temperatures between 82 and 100 km.

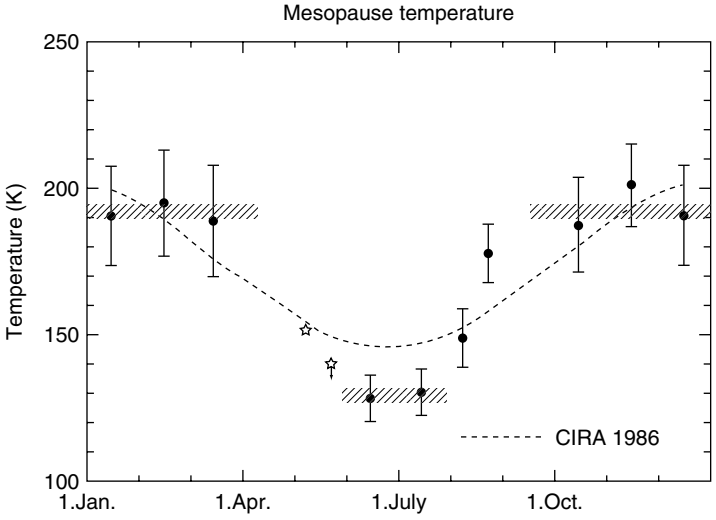


Figure 5.39 Mesopause temperature as a function of season at Andoya, Norway (69°N). The vertical bars give the variability (not error bars) of the temperature in that particular month. (From Lübken, F.J., and von Zahn, U., *J. Geophys. Res.*, 96 (D11), 20841–20857, 1991. With permission.)

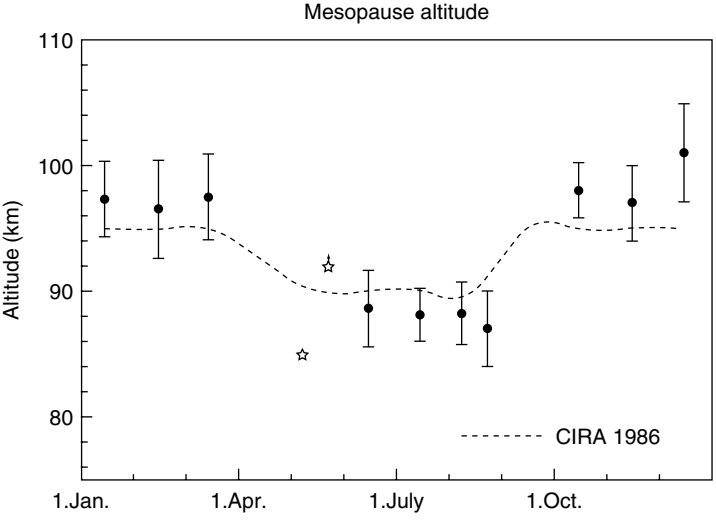


Figure 5.40 Altitude of the mesopause as a function of season at Andoya, Norway (69°N). Also shown are the error bars obtained from a polynomial fit. (From Lübken, F.J., and von Zahn, U., *J. Geophys. Res.*, 96 (D11), 20841–20857, 1991. With permission.)

The IAP narrowband K Doppler lidar (von Zahn and Höffner, 1996) was on board a ship (Polarstern) and made the mesopause temperature measurements from 80 to 105 km between 71°S and 54°N from late April to early July 1996 (von Zahn et al., 1996). Throughout the observations, the mesopause altitude was located at altitudes of either 100 $\pm$ 3 or 86 $\pm$ 3 km. The higher level was in general connected with winter conditions, while the lower with summer conditions. The high "winter" level extended from 71°S to 23°N, the low "summer" level extended from 24°N until the end of the field observations at 54°N. Shown in Figure 5.41 are the measured mesopause altitudes and temperatures versus latitudes during this shipborne campaign. From the latitudinally distributed observations, von Zahn et al. (1996) proposed that the mesopause altitude has a worldwide bimodal character. However, the mesopause temperature does not exhibit a bimodal character, but changes from the winter to summer state without noticeable jump. The mesopause temperatures measured in the southern hemisphere winter (von Zahn et al., 1996) are generally 20 K lower than previously measured in the northern hemisphere (Lübken and von Zahn, 1991). With the additional datasets from Fort Collins, USA (41°N, 105°W) by the CSU Na Doppler lidar and from Kühlungsborn, Germany (54°N, 12°E) by the IAP K Doppler lidar, She and von Zahn (1998) further investigated the concept of a two-level mesopause altitude. They concluded that the mesopause exists globally at only two altitudes: a winter state at an altitude near 100 km and a summer state near 88 km. But in contrast with the finding by von Zahn et al. (1996) that the southern mesopause in winter was colder than the northern sites, She and von Zahn (1998) concluded that in the winter state, the mesopause is at a nearly constant temperature of about 180 K, independent of latitude. Based upon these observational results, Berger and von Zahn (1999) used a three-dimensional nonlinear model of the middle-upper atmosphere (0 to 150 km) to explore the processes that lead to this two-level structure of the mesopause. In particularly, they studied the influences of photochemical heating processes, gravity wave momentum deposition, and vibrational excitation of CO_2 molecules by

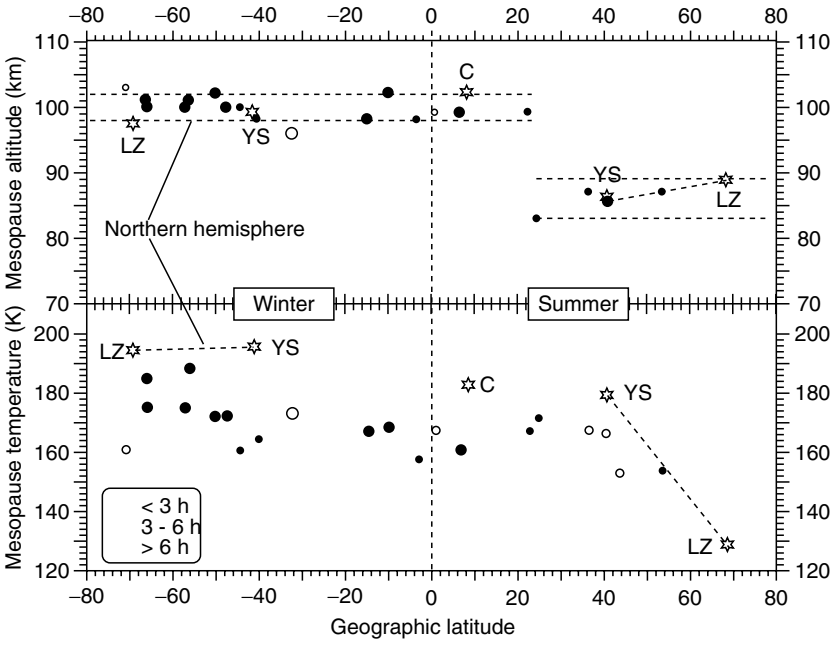


Figure 5.41 Mesopause altitude (upper panel) and temperature (lower panel) as function of latitude obtained by von Zahn et al. during their "Polarstern" cruise. The size of the dots increases with increasing length of the integration period. Dots are drawn hollow if during that night strong wave activity or strongly variable potassium densities raise some doubts as to the representativeness of the data as a climatological mean value. "LZ" refers to Lübken and von Zahn (1991), "YS" refers to Yu and She (1995), and "C" refers to Cole et al. (1979). (From von Zahn, U., Höffner, J., Eska, V., and Alpers, M., *Geophys. Res. Lett.*, 23, 3231–3234, 1996. With permission.)

collisions with O atoms on the global mesopause structure. They also calculated the effects of solar-induced tides locally on the diurnal variations of the mesopause altitude. The results of their numerical simulations suggest that two distinct mesopause altitude levels indeed exist worldwide even in terms of zonal mean states; the two-level feature is mostly due to the photochemical heating processes in combination with the gravity wave momentum deposition; and the two-

level mesopause structure is not a local phenomenon of tidal wave activity.

The Shishu University all-solid-state narrowband Na Doppler lidar (Kawahara et al., 2002) made the measurements of the mesopause temperatures from 80 to 110 km between March and September during 2000 and 2001 at the Syowa Station (69°S, 39°E), Antarctica. Figure 5.42 illustrates the monthly mean temperature profiles along with MSIS90 model at the Syowa location and the lidar observation at the northern site Andoya (69°N, 16°E). The error bars shown in the figure are the day-to-day standard deviation since the temperature measurement uncertainty is considerably less than daily variability. No apparent monthly change of the

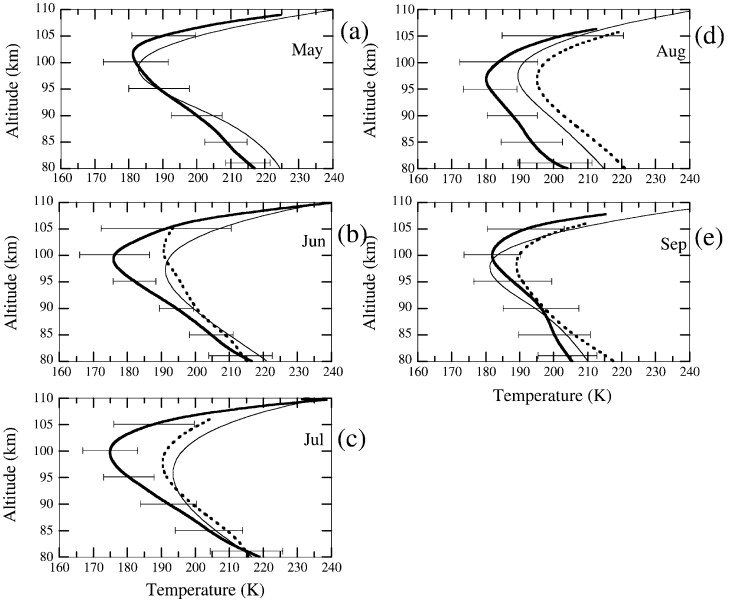


Figure 5.42 Monthly average Syowa (69°S) temperature (thick solid line) over 2 years (2000 and 2001) from May to September together with Andoya (69°N) data (dotted line) (Lübken and von Zahn, 1991) and MSIS model (thin solid line). (From Kawahara, T.D., Kitahara, T., Kobayashi, F., Saito, Y., Nomura, A., She, C.Y., Krueger, D.A., Tsutsumi, M., *Geophys. Res. Lett.*, 29, doi: 10.1029/2002GL015244, 2002. With permission.)

structures is seen in these 5 months. The mesopause altitudes are between 96 and 101 km, which is the same as the northern hemisphere winter level. The measured temperatures in the mesopause region (80 to 105 km) are, however, much lower than those at Andoya, and are also much lower than the MSIS model temperatures except in May and September. The spring (September) temperatures measured at Syowa shows closer agreement with the Andoya data and MSIS model. The measured winter mesopause altitudes at Syowa are 99 km within the day-to-day variation of ± 3 km, which are in good agreement with the observations at Andoya in wintertime. However, the averaged mesopause temperature in winter at Syowa is 175 K, about 20 K lower than those observed at Andoya in winter months (June, July, and August). The lower winter mesopause temperatures are consistent with southern hemisphere mesopause temperatures measured by a shipborne lidar (von Zahn et al., 1996). The Syowa results suggest the existence of a hemispheric difference. Though the cause of the observed asymmetry in mesopause temperature is not completely clear, Kawahara et al. (2002) suggested that it is partly due to the difference in gravity wave activities between the Arctic and the Antarctica.

The UIUC Fe Boltzmann temperature lidar (Chu et al., 2002a) was deployed from the geographic North Pole to the geographic South Pole from 1999 to 2001 to measure the temperature profiles between 30 and 110 km (Gardner et al., 2001; Chu et al., 2002a; Pan et al., 2002). With the daytime capability of this lidar system, it is possible to measure the mesopause region temperature around summer solstice at both poles under continuously sunlit (Chu et al., 2002a). An example has been shown in Figure 5.25 in the previous section. The measurement results from this Fe Boltzmann temperature lidar were combined with the balloon sonde observations at the geographic South Pole to produce the monthly mean winter temperature profiles from the surface to 110 km (Pan et al., 2002). Plotted in Figure 5.43 are the measured monthly mean temperatures along with the predictions of TIME-GCM (thermosphere–ionosphere–mesosphere-electrodynamics general circulation model) and

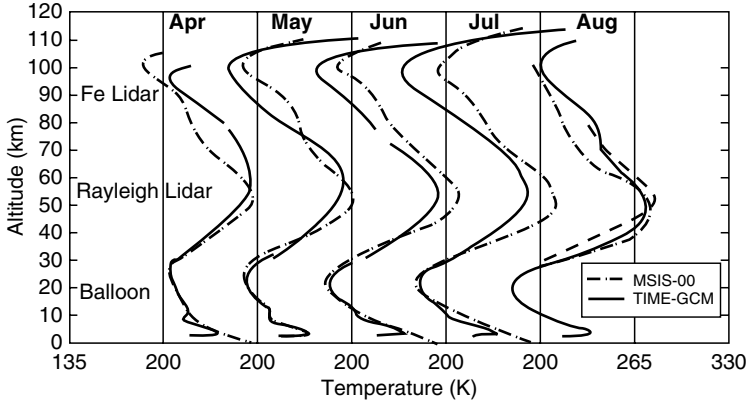


Figure 5.43 South Pole monthly mean temperature profiles for April through August (separated by 65 K for successive months). Lidar/balloon measurements are in solid lines, MSIS-00 model in dotted lines, and TIME-GCM model in dashed lines. (From Pan, W., Gardner, C.S., and Roble, R.G., *Geophys. Res. Lett.*, 29, doi: 10.1029/2002GL015288, 2002. With permission.)

the predictions of MSIS-00 (mass spectrometer incoherent scatter extended model). It is obvious from Figure 5.43 that the measured temperatures during midwinter (from May to July) in both stratopause and mesopause regions are 20 to 30 K colder than the current model predictions. Observations made at Syowa, Antarctica (69°S, 39°E) by the Shishu University Na temperature lidar also exhibit much colder mesopause temperatures in the winter (Kawahara et al., 2002). As plotted in Figure 5.44, the measured Syowa temperature in June by the Na lidar is similar to the measured South Pole temperature by the Fe lidar in the same month. Both observations are 20 to 30 K colder than TIME-GCM and MSIS-00 model predictions. Solar heating is absent during polar winter and the stable polar vortex in midwinter prevents the transport of warmer air from mid-latitudes into the south polar cap. After sunset in March, the stratosphere and mesosphere rapidly cool through thermal radiative emissions in the 15-μm band of CO_2. The wintertime radiative cooling is accompanied by downwelling over the south polar cap, which is associated

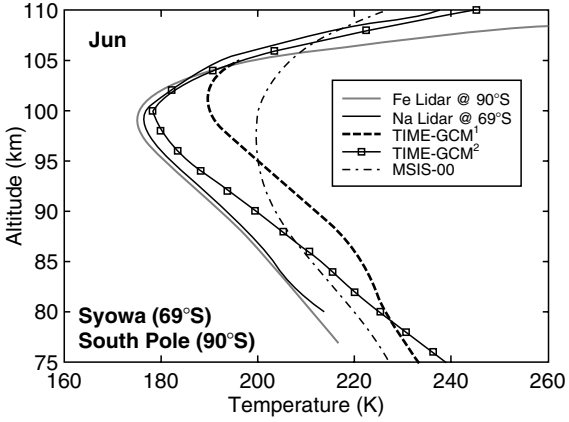


Figure 5.44 Comparison of the observed temperature at the South Pole and Syowa in June with the original TIME-GCM[1] predictions, and the TIME-GCM[2] predictions with weaker gravity wave forcing under June solstice conditions. (From Pan, W., Gardner, C.S., and Roble, R.G., *Geophys. Res. Lett*., 29, doi: 10.1029/2002GL015288, 2002 and Kawahara, T.D., Kitahara, T., Kobayashi, F., Saito, Y., Nomura, A., She, C.Y., Krueger, D.A., Tsutsumi, M., *Geophys. Res. Lett*., 29, doi: 10.1029/2002GL015244, 2002. With permission.)

with the mean meridional circulations driven primarily by the gravity wave forcing in the middle atmosphere. The downwelling heats the atmosphere through adiabatic compression and partially offsets the effects of radiative cooling, resulting in relatively warm winter temperatures in south polar cap (Garcia and Solomon, 1985). Observed colder winter temperatures in Antarctica suggest a weaker meridional circulation with subsequent weaker adiabatic warming than what is employed in models. When less gravity wave forcing was incorporated into TIME-GCM (TIME-GCM[2] in Figure 5.44), the weakened meridional circulation and down welling over the south polar cap produced colder mesopause temperatures as observed.

By combining the Fe and Rayleigh temperature measured by the UIUC Fe Boltzmann temperature lidar with the radiosonde balloon data, the atmospheric temperatures from the surface to 110 km were characterized for the first time at

the South Pole throughout the whole year. Shown in Figure 5.45 is the University of Illinois South Pole temperature model (UISP-02) derived from the 2-year measurements at the South Pole (Pan and Gardner, 2003). Although high altitude lidar coverage above 70 km was limited during summer when the Fe densities were low and the background noise from solar scattering was high, the data were sufficient to characterize the dominant 12- and 6-month oscillations in the temperature profiles. The data show that the stratosphere and lower mesosphere between 10 and 60 km altitude are dominated by a 12-month oscillation, with the coldest temperatures in winter when solar heating is absent. The mesopause region between 70 and 100 km is dominated by 12- and 6-month temperature oscillations with maximum

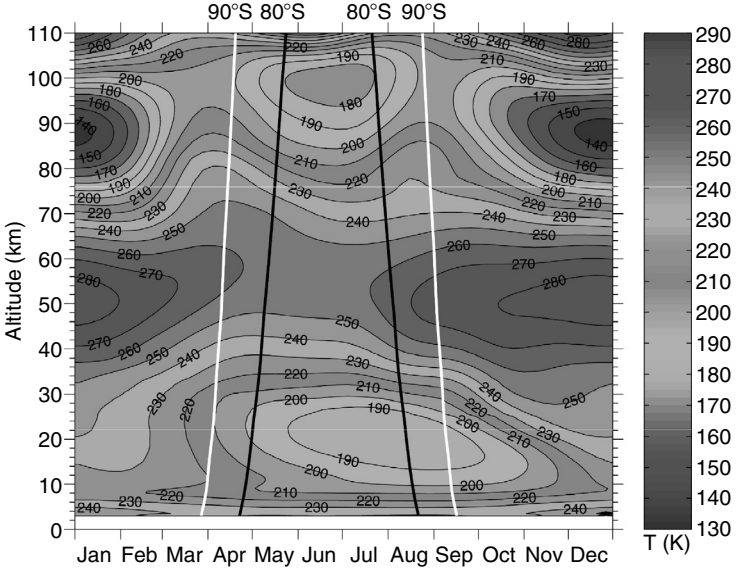


Figure 5.45 (Color figure follows page 398). The observed weekly mean temperature structure of the atmosphere above South Pole (UISP-02) plotted from 3 to 110 km. Polar nights (24 h darkness) occur between the white curve at 90°S and between the black curves at 80°S. The vertical resolution is 500 m. (From Pan, W., and Gardner, C.S., *J. Geophys. Res.*, 108 (D18), 4564, doi: 10.1029/2002JD003217, 2003. With permission.)

temperatures during the spring and fall equinoxes. The rapid response of this region to temperature changes as the sun rises and sets suggests that solar heating is stronger and upwelling is weaker than current model predictions.

5.3.1.2. Thermal Structure at Mid-Latitudes

Since the development of a new generation of the narrowband Na wind/temperature lidar through a collaboration between the Colorado State University and the University of Illinois (She et al., 1990; Bills et al., 1991a), observations of the thermal structure in the mesopause region at mid-latitudes were conducted routinely at Fort Collins, Colorado (41°N, 105°W) and at Urbana, Illinois (40°N, 88°W) for many years. The extensive observations and active researches by the CSU group and the UIUC group produced many scientific results. These results enhanced our knowledge of the background thermal structure and dynamic behavior in the mesosphere and lower thermosphere, which is essential for developing accurate global circulation and chemical models of the middle and upper atmosphere.

She et al. (1993) first reported two prevailing temperature minima observed at altitudes 86.3 $\pm$ 2.5 and 99.0 $\pm$ 2.9 km from their nighttime temperature profiles in the mesopause region, as shown in Figure 5.46. These two temperature minima appear to be a persistent phenomenon in the nighttime thermal structures at mid-latitudes, although the double minima are obvious during the spring and fall while one minimum is much weaker in the summer and winter. Similar to the bistable mesopause altitudes in the polar region reported by Lübken and von Zahn (1991), the mid-latitude mesopause also appears to be bistable with winter altitudes near 100 km and summer altitudes near 86 to 88 km (She et al., 1993; Bills and Gardner, 1993).

Senft et al. (1994) combined the nighttime observational results obtained at both Fort Collins and Urbana to provide the first comprehensive study of mesopause region temperatures at mid-latitude. Illustrated in Figure 5.47 are the contour plots of the seasonal variations of temperature structure in

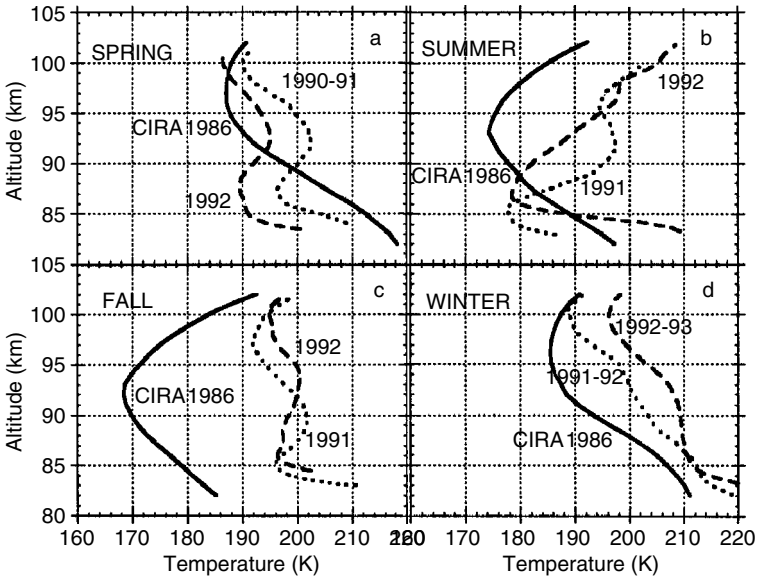


Figure 5.46 Seasonally averaged temperature profiles of the mesopause region obtained by the Colorado State University Na wind/temperature lidar over Fort Collins, Colorado, along with CIRA 1986 temperature profiles. (From She, C.Y., Yu, J.R., and Chen, H., *Geophys. Res. Lett.,* 20, 567–570, 1993. With permission.)

the mesopause region at Urbana, Illinois and Fort Collins, Colorado. The seasonal variations of the mesopause temperature and altitude are plotted in Figure 5.48. These data were obtained from the nightly mean temperature profiles at each site. The solid curves are annual plus semiannual harmonic fits to the data while the dashed curves are one standard deviation above and below the harmonic fits. The seasonal temperature variations and annual mean mesopause temperatures at these two sites are similar. The harmonic fits reach maximum values in winter of approximately 190 K and minimum values in summer of approximately 175 K. This annual variation (∼15 K) is much smaller than the mesopause temperature variation (∼60 K) at Andoya, Norway (69°N, 16°E) ranging from 190 K in winter to 130 K in summer (Lübken and von Zahn, 1991). The seasonal variations of the mesopause

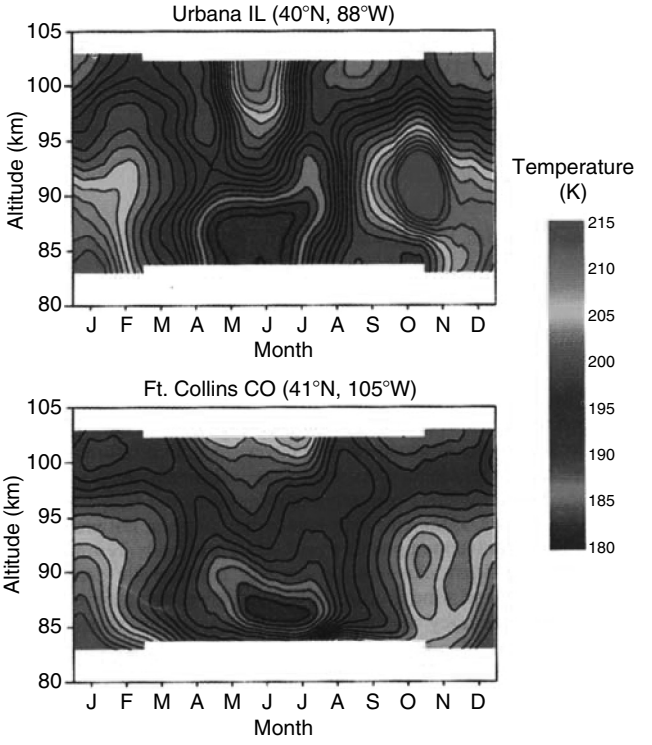


Figure 5.47 (Color figure follows page 398). Contour plots of the seasonal variations in the mesopause region temperature structure at Urbana, Illinois and Fort Collins, Colorado. The contour interval is 2 K. The plots were derived from the nightly mean temperature profiles by smoothing the data vertically and seasonally using Hamming windows with full widths of 2 km and 45 days, respectively. (From Senft, D.C., Papen, G.C., Gardner, C.S., Yu, J.R., Krueger, D.A., and She, C.Y., *Geophys. Res. Lett.*, 21, 821–824, 1994. With permission.)

heights at these two sites are significant. Both harmonic fits have strong annual components with minima near 86 km in June and maxima near 100 km from November to February. This mesopause altitude variations are quite similar to that observed by Lübken and von Zahn (1991) at Andoya, Norway. The aforementioned general features of the thermal structure in the mesopause region have been confirmed by more extensive nighttime datasets obtained at Fort Collins (She et al.,

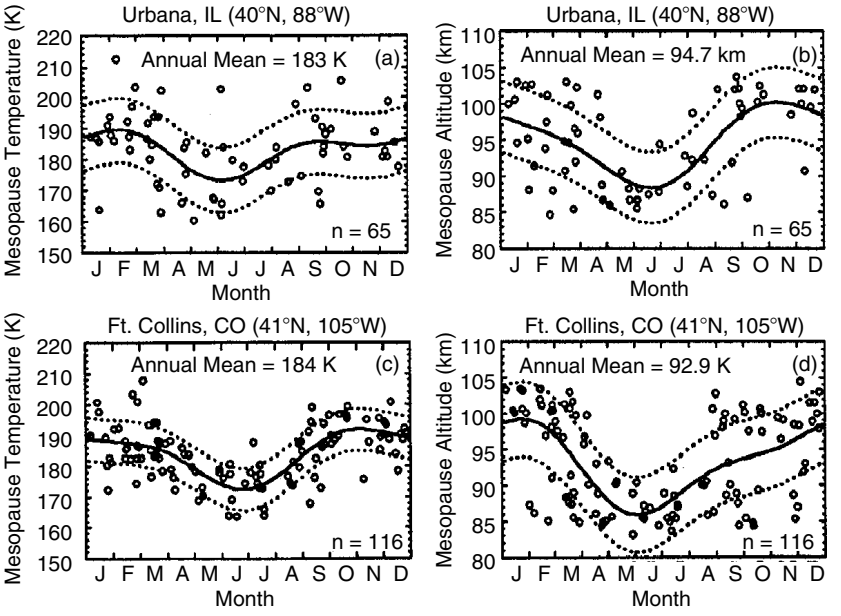


Figure 5.48 Seasonal variations of the mesopause temperatures and altitudes at Urbana (40°N, 88°W), Illinois and Fort Collins (41°N, 105°W), Colorado. The open circles are the values inferred from the nightly mean temperature profiles. (From Senft, D.C., Papen, G.C., Gardner, C.S., Yu, J.R., Krueger, D.A., and She, C.Y., *Geophys. Res. Lett.*, 21, 821–824, 1994. With permission.)

1995, 2000; Yu and She, 1995), although some details are slightly different. The mid-latitude climatology supports the concept of a two-level thermal structure with mesopause occurring at higher altitudes in winter and lower altitudes in summer, and the sharp transitions between them taking place in May and August (She et al., 2000).

The two prevailing temperature minima reported by She et al. (1993) appear to be associated with the mesosphere temperature inversion layers (MILs). The MILs are observed in low- to mid-latitude regions in two different altitude ranges at ~70 and ~95 km, which at midnight tend to be separated by the equivalent of one diurnal tidal wavelength (~25 km) (Meriwether and Gardner, 2000). Each warm air layer typically

displays a thickness of ~10 km. The lower MIL (~70 km) is a particularly characteristic and persistent feature of the winter mesosphere thermal structure with a typical temperature enhancement of about 10 to 25 K (Hauchecorne et al., 1987). The upper MIL is typically seen with amplitude ranging from 10 to 35 K and is also a persistent feature of the mesopause region (She et al., 1993; Yu and She, 1995). The lower MIL was first found by a rocket measurement (Schmidlin, 1976) then characterized mostly by Rayleigh lidars, while the upper MIL was characterized by the narrowband Na Doppler lidars.

The aforesaid measurements have been restricted to nighttime, including the thermal structure climatology and the temperature inversion layers. After the Na wind/temperature lidar was upgraded to daytime measurement capability by Chen et al. (1996) and Yu et al. (1997), it was highly desirable to investigate all these phenomena over diurnal coverage. When temperature profile was averaged over a complete diurnal cycle, the mesopause temperature profile showed little evidence of the prominent inversion layers. The temperature inversion layers frequently reported in nighttime measurements appear to be artifacts associated with incomplete sampling of the diurnal tide (States and Gardner, 1998).

The diurnal-mean climatology of the thermal structure in the mesopause region at the mid-latitude was then reported by States and Gardner (2000a) for Urbana, Illinois and by Chen et al. (2000) for Fort Collins, Colorado. Using more than 1000 h of lidar observational data obtained from February 1996 through January 1998 at Urbana, Illinois, States and Gardner (2000a) derived the seasonal variations of the thermal structure from 80 to 105 km, as shown in Figure 5.49, with (a) data covering the complete diurnal cycle and (b) data covering only the nighttime period. The nighttime mean climatology Figure 5.49(b) is very similar to the earlier datasets published by Senft et al. (1994), She et al. (1995), and Leblanc et al. (1998) using nighttime observational data. The main difference between the diurnal mean and nighttime mean thermal structures are the increased annual variability in

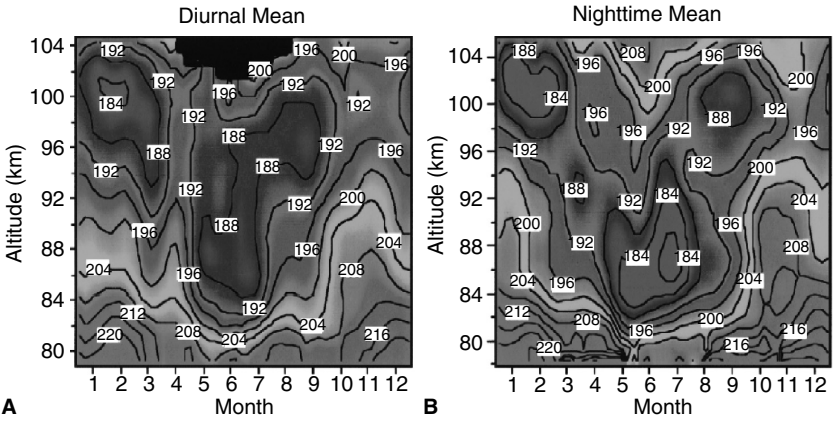


Figure 5.49 (Color figure follows page 398). Contour plot of the mesopause region annual temperature structure at Urbana, Illinois using (A) data covering the complete diurnal cycle and (B) data covering just the nighttime period. (From States, R.J., and Gardner, C.S., *J. Atmos. Sci.*, 57, 66–77, 2000. With permission.)

the nighttime data below 96 km and increased semiannual variability above 96 km, and inversion layers appearing only in the nighttime data. Plotted in Figure 5.50 are the comparison of mesopause altitude, width, and temperature between the diurnal mean and the nighttime mean. Here, the mesopause width is defined as the altitude range where the temperature is within 5 K of the temperature minimum. The quasi-two-level mesopause altitude is clearly presented here. According to the diurnal mean (solid line in Figure 5.50), the mesopause is near the high winter state (98 to 101 km) until about May 7 when it abruptly switches to the lower summer state (86 to 88 km). The mesopause remains in the summer state for about 70 days until about July 15 when it abruptly moves to about 96 km and then slowly increases over the next several months back to the winter state (~101 km). Both transitions last less than a week. For the nighttime data (dotted line), the mesopause transits to the summer state earlier in the spring and remains there longer. From the observational data shown in Figure 5.50(a),

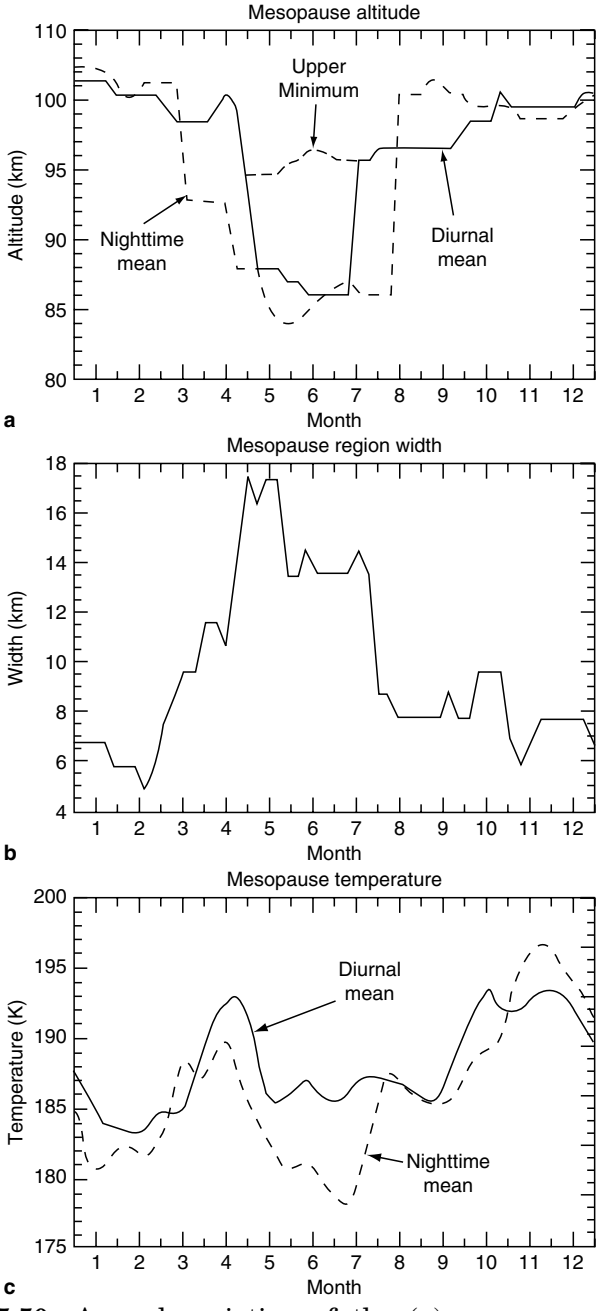


Figure 5.50 Annual variation of the (a) mesopause altitude, (b) mesopause width, and (c) mesopause temperature at Urbana, Illinois. (From States, R.J., and Gardner, C.S., *J. Atmos. Sci.*, 57, 66–77, 2000. With permission.)

States and Gardner (2000a) suggested that thermospheric heating forces the mesopause to a lower altitude during summer. This trend is interrupted when the adiabatic cooling in the lower mesopause region assumes control of the mesopause altitude in late spring. Without the strong summer adiabatic cooling below 90 km, the mesopause altitude would oscillate between a winter maximum of about 101 km and a summer minimum of about 96 km. The mesopause temperature derived from the diurnal mean is warmer than the nighttime data because of daytime solar heating as shown in Figure 5.50(c). The nighttime mesopause temperature averages 2.0 K colder than the diurnal mean, with biggest deviation occurring during summer time as expected. States and Gardner (2000a) also presented a brief summary of the energy budget for the thermal structure in the mesopause region. Plotted in Figure 5.51 is their annual mean temperature profile over diurnal-mean with their suggestions for the

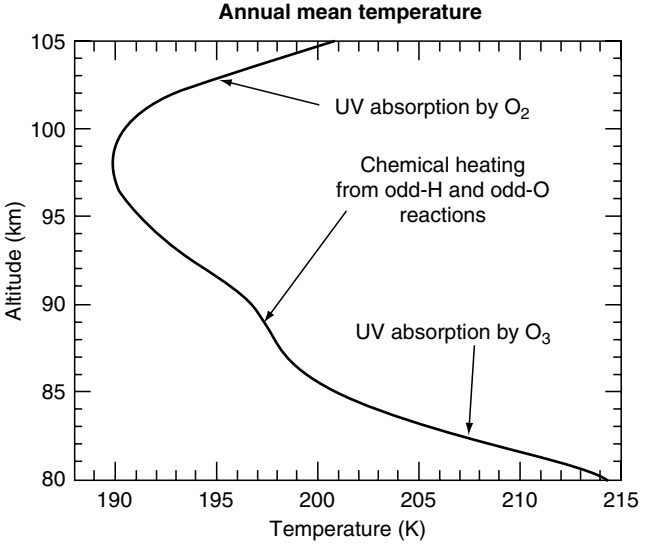


Figure 5.51 Mean background temperature averaged over both the diurnal and annual cycles at Urbana, Illinois (40°N, 88°W). (From States, R.J., and Gardner, C.S., *J. Atmos. Sci.*, 57, 66–77, 2000. With permission.)

main energy sources. This observed diurnal mean annual temperature structure is largely consistent with the assumption of radiative equilibrium between direct solar UV heating and radiative cooling by IR emission (States and Gardner, 2000a).

Chen et al. (2000) used 18 datasets of continuous temperature measurements covering a complete diurnal cycle over Fort Collins, Colorado from February 1997 to January 1999 to derive the seasonal variations of diurnal mean thermal structure in the mesopause region. Illustrated in Figure 5.52 are the temperature profiles for four seasons (spring, summer, fall, and winter) over the diurnal mean, nighttime mean, and daytime mean. The difference between annual diurnal, nighttime, and daytime means may be ascribed to a temperature wave propagating with a downward phase speed of ~0.8 km/h. The nighttime annual mean is colder than the

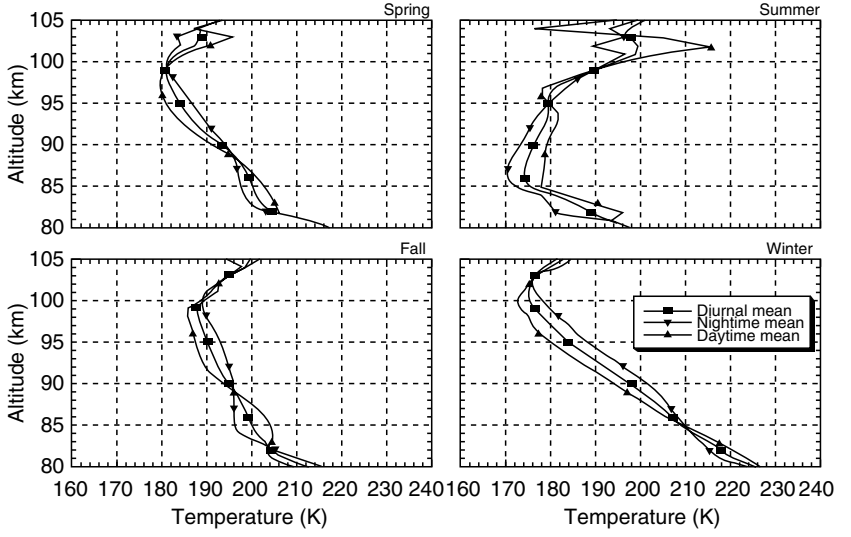


Figure 5.52 Profiles of seasonal averages of temperature means over 24 hours of data, means over nighttime (18:00 to 06:00 LST), and means over daytime (06:00 to 18:00 LST) from 18-day data obtained at Fort Collins, Colorado. (From Chen, S., Hu, Z., White, M.A., Chen, H., Krueger, D.A., and She, C.Y., *J. Geophys. Res.*, 105 (D10), 12371–12379, 2000. With permission.)

diurnal annual mean by no more than 2 K below 88 km and warmer by no more than 3 K above 88 k where the three means are nearly the same. Shown in Figure 5.53 are the mesopause altitude and temperature for the 18 datasets. In addition, the altitudes of the secondary minima above 91 km for the three summer data are marked by × in Figure 5.53(a). The winter-to-summer transitions, based on diurnal means, are abrupt and clear, further strengthening the concept of

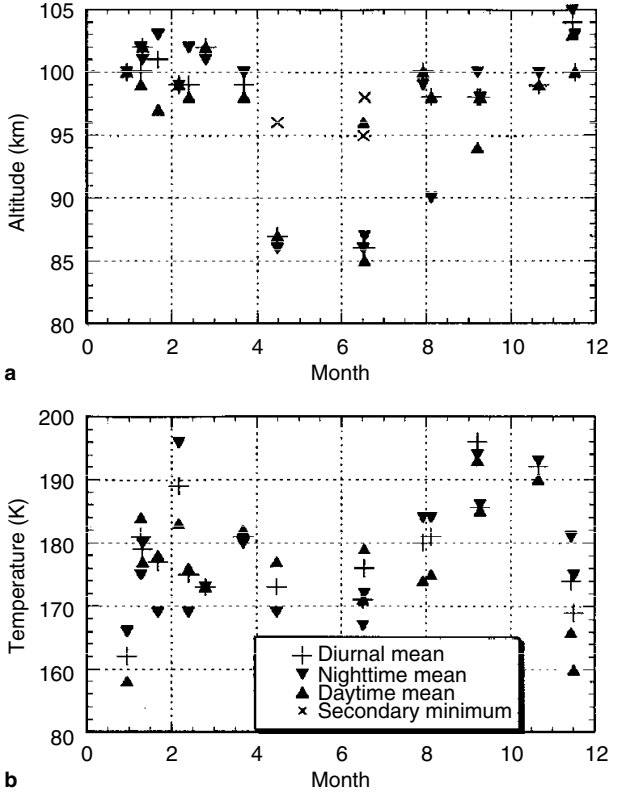


Figure 5.53 Mesopause altitudes (a) and temperatures (b) for the 18-day data taken at Ft. Collins, Colorado. The altitudes of the secondary minima for the three summer days are marked by × in (a). (From Chen, S., Hu, Z., White, M.A., Chen, H., Krueger, D.A., and She, C.Y., *J. Geophys. Res.*, 105 (D10), 12371–12379, 2000. With permission.)

two-level mesopause. Similar to the Urbana data, Chen et al. (2000) also observed the altitude of a secondary temperature minimum in the summer to be ~96 km, which, without strong dynamic cooling in the summer, would be the summer mesopause altitude.

The full diurnal coverage of the Na Doppler lidar measurements of the thermal structure in the mesopause region enables the studies of thermal tides at the mid-latitudes. States and Gardner (2000b) used the same Na lidar dataset as in States and Gardner (2000a) to characterize the seasonal behavior of solar thermal tides in the mesopause region between 80 and 105 km. She et al. (2002) used the same 18 datasets of 24-h continuous lidar temperature observations as in Chen et al. (2000) to analyze the mean amplitude and phase of oscillations with 24, 12, 8, and 6 h periods for four seasons. Both observational results are compared to the recent global-scale wave model (GSWM) (Hagan et al., 1995). Good general agreements between the observed and model tidal phases were found for both sites, implicating the prevalence of migrating diurnal and semidiurnal tides. However, the GSWM consistently underestimates the diurnal and semidiurnal amplitude (States and Gardner, 2000b). These observations demonstrate the high potential of full-diurnal-coverage lidars for tidal studies.

5.3.1.3. Thermal Structure at Low Latitudes

The transportable IAP K Doppler lidar, operating at scanning K (D_1) line mode (von Zahn and Höffner, 1996), was located at the Observatorio del Teide (28°N, 17.5°W) on the island of Tenerife. Four campaigns were conducted in 1999. Using the data obtained in 55 nights, Fricke-Begemann et al. (2002) characterized the nocturnal mean temperatures and their seasonal variations. Shown in Figure 5.54 are the mean temperature profiles for the four campaigns. They provide the new evidence for the global two-level structure of the mesopause. Around midsummer, the mesopause is close to 86 km with about 183 K for a period of 4 weeks from the end of May to the end of June. The transition periods between the regular

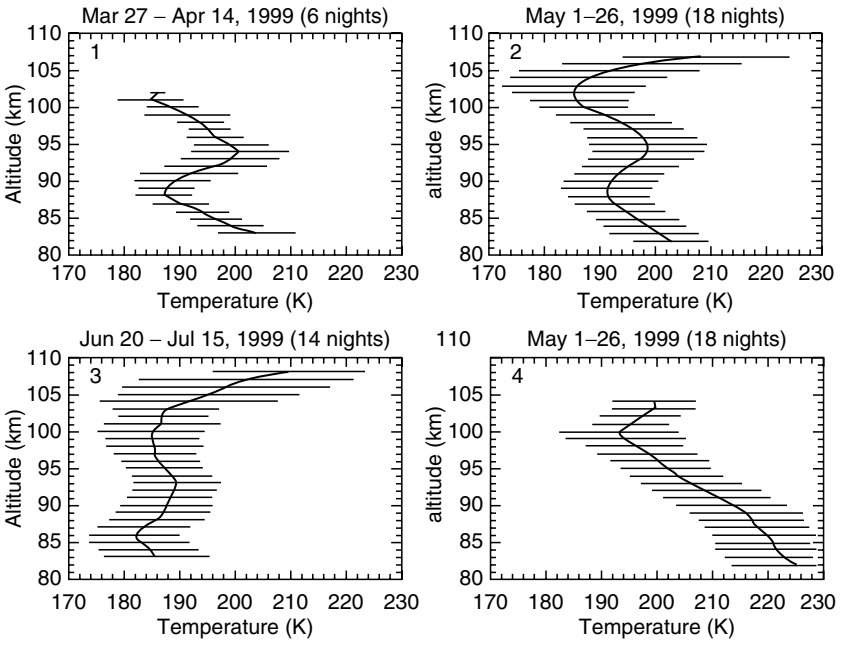


Figure 5.54 The mean temperature profiles obtained by the IAP K Doppler lidar at Tenerife (28°N) during 55 nights in 1999. The bars represent the rms variability of the nightly means. Values are given only for altitudes where more than 50% of the nights have reliable data. (From Fricke-Begemann, C., Höffner, J., and von Zahn, U., *Geophys. Res. Lett.*, 29 (22), doi: 10.1029/2002GL015578, 2002. With permission.)

(high) and the low altitude state of the mesopause are 2 weeks each. In November the observed regular mesopause is at 100 km with 192 K temperature. The thermal structure observed at this 28°N site is similar to what was observed at 41°N and 40°N.

The Arecibo K Doppler lidar, operating at three separate frequencies (Friedman et al., 2002), was operated routinely at Arecibo Observatory (18.35°N, 66.76°W). Based upon 74 nights temperature profiles of observing periods greater than 5 h, Friedman (2003a) derived the climatology of tropical thermal structure in the mesopause region. Plotted in Figure 5.55 is the annual mean thermal structure in

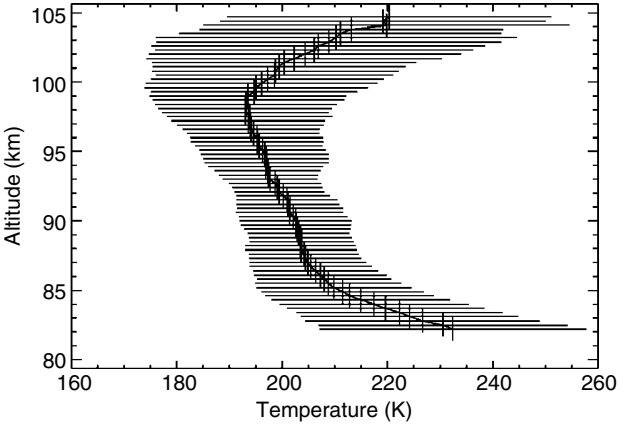


Figure 5.55 Annual mean mesopause temperature at Arecibo Observatory (18.35°, 66.75°W) based on the average of the K lidar data from 74 nights. The error bars represent the rms variability of the layer at each height. (From Friedman, J.S., *Geophys. Res. Lett.*, 30 (12), doi: 10.1029/2003GL016966, 2003. With permission.)

the mesopause region based on the average of the data from the 74 nights. The mean mesopause is at 98 km with a temperature of 193 K. This annual mean temperature profile is similar to what was seen at mid-latitudes. Shown in Figure 5.56 is the contour plot of the seasonal variation of thermal structure from 80 to 110 km at Arecibo. It presents a temperature minimum of 176 K at 100 km during mid-summer. The winter period minimum of about 190 K is found mostly in a broad flat distribution between 90 and 95 km, and less frequently near 100 km (Friedman, 2003a). These features are quite different from the 28°N observations made by Fricke-Begemann et al. (2002). Figure 5.57 shows the mesopause temperature and altitude observed at Arecibo over the 22 months. The high altitude of summer mesopause is clearly shown at the center of the plot, with only 3 of 29 nights from May through August have mesopause altitudes below 95 km. The harmonic fit to the mesopause temperature exhibits that the annual mean mesopause temperature is 179 K, and the mesopause temperature is 21 K warmer in winter than in summer. The temperature inversion layers

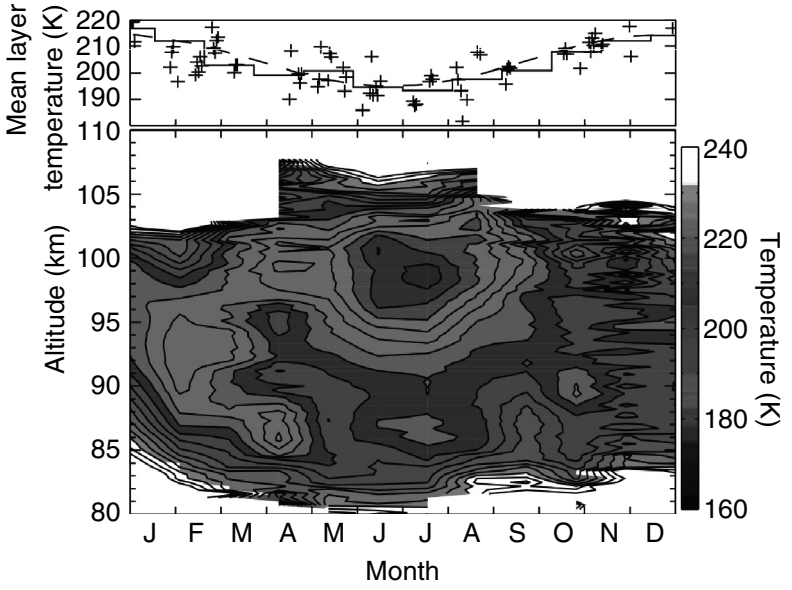


Figure 5.56 (Color figure follows page 398). Plotted in the upper panel are the nightly (+) and monthly (—) average temperatures along with a harmonic fit (- - -). Shown in the lower panel is the seasonal variation of mesopause region temperature above Arecibo. (From Friedman, J.S., *Geophys. Res. Lett.*, 30 (12), doi: 10.1029/2003GL016966, 2003. With permission.)

around 90 km are usually observed during the summer period with amplitudes often greater than 30 K. Inversions are not frequently observed in winter. Therefore, a dual minimum is observed in summer, with the upper level virtually always experiencing the lower temperature. The seasonal variation of the mesopause region temperature structure observed at Arecibo provides a challenge for both theoretical models and empirical studies. Both the summer mesopause altitude and the summer mean temperature near 176 K are consistent with the earlier observations by She and von Zahn (1998) and the model prediction by Berger and von Zahn (1999). But the broad flat temperature profile observed in February was not predicted by the model, and deserves further study.

Clemesha et al. (1997) proposed and demonstrated a new low-cost technique, using a comb laser, to determine the

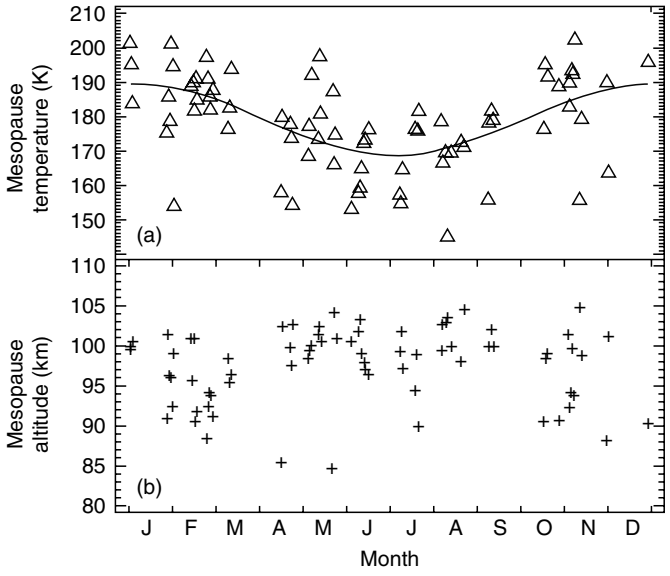


Figure 5.57 Shown, binned by date, are (a) the temperature and (b) the altitude of the mesopause for 67 nightly observations made by the Arecibo K Doppler lidar at Arecibo. (From Friedman, J.S., *Geophys. Res. Lett.*, 30 (12), doi: 10.1029/2003GL016966, 2003. With permission.)

Doppler temperature by a fairly simple modification to the dye laser transmitter of a broadband Na lidar. By replacing the front reflector with a low reflectivity Fabry–Perot interferometer, a broadband dye laser is modified to a comb laser with narrowband multi-line output where the spacing of the multi-line is tuned exactly equal to the separation of the D_{2a} and D_{2b} transition groups of Na (D_2) line. The lidar return from Na atom scattering is maximum when the laser lines coincide with the D_{2a} and D_{2b} hyperfine groups, and minimum when one of the lines lies approximately halfway between D_{2a} and D_{2b}. Therefore, the depth of modulation of the lidar response is a sensitive function of the Na atom temperature. By sweeping the Fabry–Perot interferometer through several free spectral ranges, the lidar response from atmosphere Na layers is recorded, and the temperature at each altitude is determined by least mean square fits to the lidar response versus

wavelength. With this simple, low-cost lidar system, Clemesha et al. (1997) were able to measure temperature profiles in the mesopause region and claimed a potential precision of ± 5 K with height resolution of 1 km and temporal resolution of 6 min. The actual results give an estimate of ± 7 K error in the absolute temperatures, and ± 2 K error in the relative temperatures when the data are averaged over 10 min. Clemesha et al. (1999) reported the mesopause temperature profiles measured at southern latitude Sao Jose dos Campos (23°S, 46°W), Brazil. As shown in Figure 5.58, the average temperature profile based on the measurements of 15 nights from July to October 1998 reveals a mesopause temperature of 190 K at a height of 102.5 km. Between 83 and 100 km, the measured temperatures are typically 20 K warmer than the temperature predicted by the CIRA86 standard atmosphere model for the corresponding time period at 23°S location. This southern mean temperature profile is very similar to the December average at Fort Collins, supporting the suggestion of She and

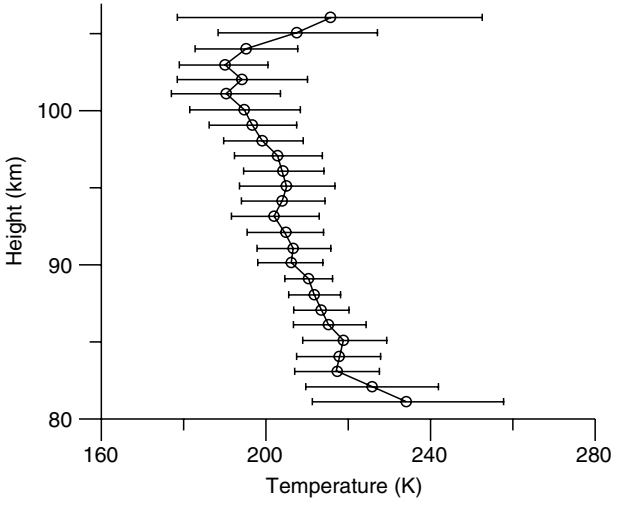


Figure 5.58 Average temperature profile for measurements made by the INPE Na lidar at Sao Jose dos Campos (23°S, 46°W), Brazil between July and October, 1998. (From Clemesha, B.R., Veselovskii, I., Batista, P.P., Jorge, M.P.P.M., and Simonich, D.M., *Geophys. Res. Lett.*, 26, 1681–1684, 1999. With permission.)

von Zahn (1998) that between 23°S and 23°N, the mesopause is probably always located at the high "winter" level.

5.3.2 Dynamics of the Middle and Upper Atmosphere

Studies of atmospheric dynamics depend on the availability of temperature and wind data. Na wind/temperature lidars can provide these data in the mesopause region and thus have a significant impact on the progress in dynamic studies. There are currently two types of experimental configurations for these studies: the coupled lidar beam with large aperture telescope through coude optics, which can be steered together, as the system developed by the UIUC group (Tao and Gardner, 1995; Zhao et al., 2003), and the dual-beam lidar system with two fixed beams and telescopes, as the system developed by the CSU group (She et al., 2003).

5.3.2.1. Heat, Momentum, and Na Flux

During the Airborne Lidar and Observations of Hawaiian Airglow (ALOHA-93) campaign (Gardner, 1995), the University of Illinois Na wind/temperature lidar was operated from the summit of Haleakala Mountain (20.7°N, 156°W) at Maui, Hawaii. As the system measured all three components of the wind and temperature, Tao and Gardner (1995) used the data obtained on October 21, 1993 to estimate the heat flux and heating rate profiles between 85 and 100 km. This is known as the first study of heat flux with a lidar. The momentum flux and mean flow acceleration in the troposphere, stratosphere, and mesosphere have been studied by numerous groups using the dual-beam radar technique pioneered by Vincent and Reid (1983). However, there was no report of heat flux measurements until Tao and Gardner (1995) because these radar observations did not include temperature measurements.

The Na lidar beam was coupled to a 0.8 m diameter Beam Director Telescope that provided full scanning capabilities for the lidar. The lidar was pointed at zenith (Z) and 15° off zenith to the north (N) and east (E) in the sequence of ZNZEZNZE. The laser was tuned to four frequencies: the Na D_{2a} peak

frequency f_a, crossover frequency f_c, and 600 MHz above and below D_{2a} peak frequency ($f_{\pm} = f_a \pm 600$ MHz) in the sequence of f_-, f_a, f_+, and f_c. At each frequency the signal was integrated for 900 laser pulses over a period of 30 s at a range resolution of 48 m. Temperature, vertical wind, and Na density profiles were obtained at the zenith position with a temporal resolution of ~7.5 min. Radial wind, temperature, and Na density profiles were obtained at the off-zenith positions with a temporal resolution of ~15 min. With these data, Tao and Gardner (1995) calculated the Na density, temperature, vertical wind, zonal wind, meridional wind, and vertical shear of meridional wind shown as the contour plots in Figure 5.59. The temperature and wind perturbation profiles (T', u', v', and w') were computed by subtracting the corresponding mean profiles for the night. To improve the accuracy of the flux measurements, the perturbation profiles were smoothed using rectangular windows with full widths of 60 min and 4 km. Thus, the computed perturbations include the effects of waves with periods between about 2 and 10 h and vertical wavelengths between about 8 and 30 km. The heat fluxes were computed by calculating the covariances of temperature and wind perturbations. The results are plotted in Figure 5.60. Tao and Gardner (1995) did an analysis on the instrument errors and statistical variability in the measurement of heat fluxes. The error limits plotted in Figure 5.60 include the combined effects of instrument errors and the statistical uncertainties of the flux estimates. The zonal heat flux (Figure 5.60b) is positive everywhere, indicating strong eastward heat transport. Since the standard deviation of meridional flux measurement exceeds the mean (Figure 5.60c), we can only conclude that the meridional heat transport was weak. From the vertical heat flux, Tao and Gardner (1995) computed the heating rate associated with wave transport of thermal energy as illustrated in Figure 5.61 along with the rms error limit imposed by instrument errors and the statistical variability of the heat flux estimate. Because of the strong convergence of the vertical flux, the heating rate is positive between 87 and 97 km and negative at the extreme edges of the observation range.

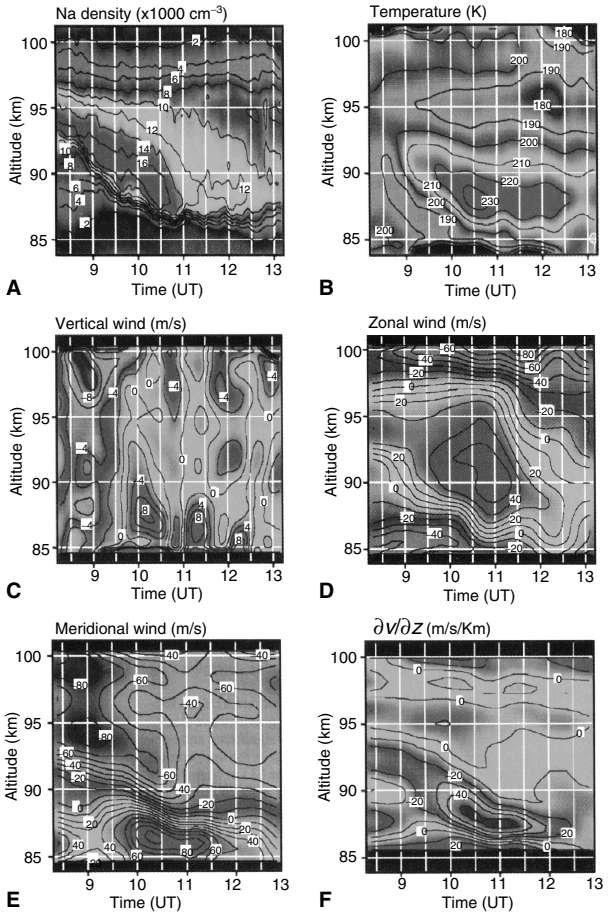


Figure 5.59 (Color figure follows page 398). Contour plots of (A) Na density, (B) temperature, (C) vertical wind, (D) zonal wind, (E) meridional wind, and (F) vertical shear of meridional wind as functions of time and altitude on October 21, 1993 at Haleakala, Maui by the UIUC Na wind/temperature lidar. To generate these plots the data were smoothed to resolution of 30 min and 1 km. (From Tao, X., and Gardner, C.S., *Geophys. Res. Lett.*, 22, 2829–2832, 1995. With permission.)

The vertical heat flux and heating rate reported by Tao and Gardner (1995) are substantially different from the values predicted by gravity wave theory (Schoeberl et al., 1983; Gardner, 1994). However, the measured values are statistic-

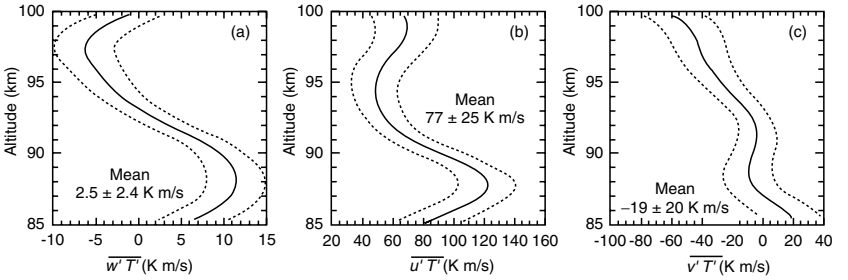


Figure 5.60 Vertical profiles of (a) vertical heat flux, (b) zonal heat flux, and (c) meridional heat flux obtained by the UIUC Na Doppler lidar at Haleakala. The dashed lines are 1 standard deviation error limits imposed by instrument errors and the statistical variability of the flux estimates. (From Tao, X., and Gardner, C.S., *Geophys. Res. Lett.*, 22, 2829–2832, 1995. With permission.)

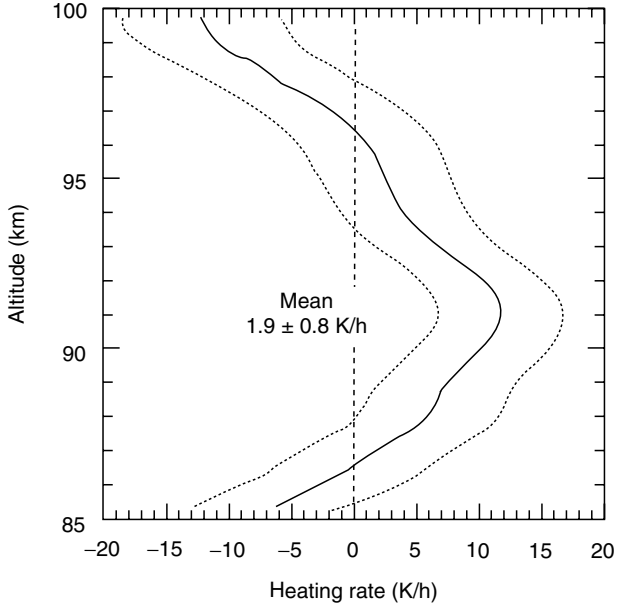


Figure 5.61 Vertical profile of heating rate (solid line) inferred from the UIUC Na wind/temperature lidar measurements at Haleakala in 1993. The dashed lines are the 1 standard deviation error limits. (From Tao, X., and Gardner, C.S., *Geophys. Res. Lett.*, 22, 2829–2832, 1995. With permission.)

ally significant and consistent with the observed short-term variations in the temperature profile (Figure 5.59b). The disagreement between the observation and theoretical predictions may be due to the short observational period (~5 h), which may bias the mean state of the mesopause region. In addition, the wind and temperature data were heavily averaged to reduce instrument errors to tolerable levels. Thus, the heat flux estimates include only the effects of waves with periods between 2 and 10 h and vertical wavelengths between 8 and 30 km. Fritts and Vincent (1987) have shown that in the upper mesosphere, waves with periods less than 2 h are responsible for the momentum deposition. If this is also true for the heat flux, then we need higher resolution data to compute the heat fluxes and heating rates. Because the dominant error source is photon noise, higher resolution heat flux measurements can be obtained only by substantially increasing the power aperture product of the lidar. Using larger telescopes is one of the best ways to improve the accuracy and resolution of the heat flux measurements.

Gardner and Yang (1998) used more extensive, higher-resolution observations made at the Starfire Optical Range (SOR) (35°N, 107°W) near Albuquerque, New Mexico to compute vertical heat flux and cooling rate profiles in the mesopause region. The University of Illinois Na wind/temperature lidar was interfaced with the SOR steerable telescope (3.5 m in diameter) to make nighttime measurements of temperature and all three wind-components between 82 and 102 km. The power-aperture product of the lidar was approximately 10 W m^2. A total of 65 h of measurement was conducted on eight different nights in November 1994, February 1995, and April 1995. The lidar and telescope were pointed at zenith (Z) and 15° off-zenith to the north (N), south (S), east (E), and west (W) in the sequence of NZEZSZWZ. The laser was tuned to the D_{2a} peak (f_a) and the wing frequencies $f_a \pm 600$ MHz. The laser was locked to f_a using the Doppler-free feature of a Na vapor cell, and to $f_a \pm 600$ MHz using a temperature-stabilized etalon (Bills et al., 1991b). Tuning was cycled rapidly through the three frequencies several times during the 2.5 min data accumulation period to minimize errors associated with Na

density fluctuations. Due to the high signal levels, the vertical wind and temperature profiles computed from SOR data were obtained with temporal resolution of 2.5 min and vertical bin width of 480 m. Then the wind and temperature perturbations were computed by removing a linear trend in time at each altitude (Gardner and Yang, 1998). The mean observing period for each night was about 8 h. Thus, the computed heat fluxes include the effects of waves with periods between 5 min and 16 h, and vertical wavelengths between 1 and 24 km. Gardner and Yang (1998) also computed the dynamical cooling rate from the heat flux estimates. They did a detailed error analysis in the estimate of the heat flux and cooling rate. The seven-night mean heat flux profile, dynamical cooling rate, and their error limits are plotted in Figure 5.62. When averaged within 92.5 ± 6 km, the mean vertical heat flux was -2.29 ± 0.43 K m/s, the mean cross-correlation coefficient for the wind and temperature fluctuations was -0.12 ± 0.02, and the mean cooling rate was -30.9 ± 5.2 K/day. Despite the large geophysical variability of the nightly measurements,

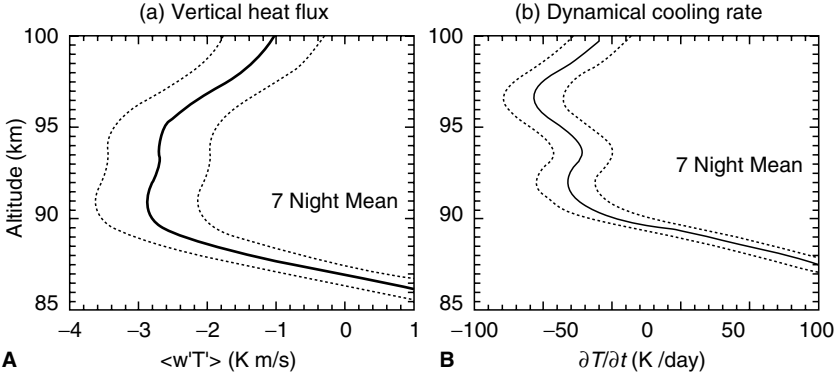


Figure 5.62 (A) Mean vertical heat flux and (B) mean dynamical cooling rate profiles obtained by the UIUC Na wind/temperature lidar at Starfire Optical Range, Albuquerque, New Mexico during 1994 to 1995 campaigns. The mean profiles were smoothed vertically using a 4 km FWHM Hamming window and exclude the February 3, 1995 data. (From Gardner, C.S., and Yang, W., *J. Geophys. Res.*, 103 (D14), 16909–16927, 1998. With permission.)

Gardner and Yang (1998) believed that these seven-night means are representative of the dynamical heat transport and cooling associated with dissipating gravity waves in the mesopause region. The measured cooling rate (−31 K/day) is consistent with theoretical predications (−36 K/day). The vertical wind and temperature data were acquired with exceptionally high spatial (480 m) and temporal (2.5 min) resolution, so the computed heat fluxes and cooling rates include the effects of the important short-period short-vertical-scale waves, which are expected to make the most significant contributions to momentum deposition and heat transport.

The vertical heat flux is defined as the expected value of the product of the vertical wind perturbation (w') and temperature perturbation (T'), that is, $\langle w'T' \rangle$. The measured perturbations from a lidar are the sum of the gravity wave perturbation (w' and T') and the instrumental errors (Δw and ΔT). Thus, the heat flux estimate is given by (Gardner and Yang, 1998):

$$\overline{\mathrm{HF}} = \frac{1}{\tau L} \int_{z_0 - L/2}^{z_0 + L/2} \mathrm{d}z \int_{t_0 - \tau/2}^{t_0 + \tau/2} \mathrm{d}t (w' + \Delta w)(T' + \Delta T) \qquad (5.123)$$

where τ is the averaging period and L is the height range for sample average. Equation (5.123) is an unbiased estimator of $\langle w'T' \rangle$, and its variance is

$$\mathrm{Var}(\overline{\mathrm{HF}}) \approx \frac{\Delta z_{\mathrm{HF}}}{L} \frac{\Delta t_{\mathrm{HF}}}{\tau} [\langle (w')^2 \rangle + \langle (\Delta w)^2 \rangle][\langle (T')^2 \rangle + \langle (\Delta T)^2 \rangle] \qquad (5.124)$$

where $\Delta t_{\mathrm{HF}} \approx 10$ min and $\Delta z_{\mathrm{HF}} \approx 1$ km are the correlation time and vertical correlation length of the heat flux estimate. The vertical heat flux is a measure of the correlation between the vertical wind and the temperature perturbation. Because the temperature perturbations are approximately proportional to the vertical displacement, and w' is nearly orthogonal to T', the vertical heat flux is expected to be small. Typical values are a few Kelvin meter per second. To achieve statistically significant heat flux estimate, the standard deviation associated with Equation (5.124) must be significantly smaller

than the mean heat flux value. The standard deviation is mainly caused by two types of errors: one is the instrumental errors, whose dominant source is photon noise; another is the statistical fluctuation due to the geophysical variability of gravity waves. The photon noise (instrumental) errors can be reduced by longer integration period and large spatial average. However, longer integration time will exclude many high frequency waves that make the most significant contributions to the momentum and heat fluxes. The most sufficient way to reduce photon noise but still preserve the high spatial and temporal resolutions is to use much large telescopes to improve the photon signal levels, as demonstrated in Gardner and Yang (1998). Nevertheless, even in the absence of the instrumental errors (mainly photon noise), that is, Δw and ΔT are zeros, the standard deviation due to the statistical fluctuations caused by the geophysical variability of gravity wave heat flux is still significant if the flux estimate is only averaged over a short period like $\tau = 8$ h. Several thousand independent measurements must be averaged in time and altitude to obtain statistically significant heat flux estimation. The same difficulties are also applied to the vertical flux of horizontal momentum of gravity waves, as recognized by Kudeki and Franke (1998) and Gardner and Yang (1998). These challenges require the heat flux and momentum flux measurements to be conducted with Na lidar coupled with large telescopes, such as 3.5 m or bigger diameter, and many independent measurements must be made to reduce the statistical fluctuations. An example of momentum flux lidar measurements can be found in Gardner et al. (1999).

Eddy transport is believed to play an important role in establishing the constituent structure of the middle atmosphere below the turbopause. Gravity waves contribute to this transport process by generating turbulence when they break. Dynamical transport occurs when dissipating (nonbreaking) gravity waves impart a net vertical displacement in the constituent as they propagate through a region. It can occur even in the absence of a concentration gradient. Unlike eddy transport, the effects of dynamical transport are usually neglected in chemical models of the middle and upper

atmosphere. Over 400 h of high-resolution Na lidar data obtained at SOR were used to study the vertical dynamical transport of Na in the mesopause region between 85 and 100 km (Zhao, 2000; Liu and Gardner, 2004). The vertical Na flux was computed from Na and vertical wind data and it was related to the vertical heat flux in a simple way. The observed Na flux profile is consistent with theoretical predictions in the region between 85 and 93 km where the wave dissipation is strong. The direct measurements of the vertical dynamical flux of Na show that the dynamical transport is at least as important as eddy transport in the mesopause region and should be included in chemical models.

5.3.2.2. Atmospheric Instability and Gravity Wave Directions

Static (convective) and dynamic (shear) instabilities are believed to play major roles in gravity wave dissipation. The static (convective) stability of the atmosphere is characterized by the square of the buoyancy frequency N defined as

$$N^2 = \frac{g}{T}\left(\frac{\partial T}{\partial z} + \frac{g}{C_p}\right) \tag{5.125}$$

where g is the gravitational acceleration (9.5 m/s^2 in the mesopause region), T is the atmospheric temperature, and $C_p = 1004\,\mathrm{J\,K^{-1}\,kg^{-1}}$ is the specific heat at constant pressure. When N^2 is negative (i.e., the atmospheric lapse rate is larger than the adiabatic lapse rate), the atmosphere is unstable. Dynamic (shear) stability is characterized by Richardson number Ri:

$$Ri = \frac{N^2}{(\partial u/\partial z)^2 + (\partial v/\partial z)^2} = \frac{N^2}{S^2} \tag{5.126}$$

where u and v are the zonal and meridional wind profiles, respectively, and $S = [(\partial u/\partial z)^2 + (\partial v/\partial z)^2]^{1/2}$ is the total vertical shear of horizontal wind. Dynamic instability is induced by large vertical shears of the horizontal wind in combination with low static stability. The atmosphere is considered to be dynamically unstable when $0 < Ri < 1/4$.

Although the static and dynamic instabilities are important wave saturation mechanism in the mesopause region, there is still no detailed understanding of the instability due to the difficulties in obtaining temperature and wind data with sufficient spatial and temporal resolution and accuracy. Most studies of the instability properties in the mesopause region have been limited to theoretical work and lab experiments. The buoyancy period is about 5 min at mesopause heights and the viscous dissipation limit is about 1 km. Thus, to capture most gravity wave effects, at least 500-m vertical resolution and 2.5-min temporal resolution are needed. If Δz is the vertical resolution, the measurement rms errors in the temperature and horizontal wind must satisfy the following criteria according to Zhao et al. (2003):

$$\Delta T_{\mathrm{rms}} \ll \Delta z \cdot \mathrm{std}(\partial T'/\partial z)/\sqrt{2} \approx 3\mathrm{K} \tag{5.127}$$

$$\Delta u_{\mathrm{rms}} \ll \Delta z \sqrt{\langle S^2 \rangle}/2 \approx 6\,\mathrm{m/s} \tag{5.128}$$

where "std" is the standard deviation, Δz is 500 m, the mean observed values for the rms temperature lapse rate are 7.7 K/km, and total wind shear is 23 m/s/km. These accuracy and resolution criteria present considerable observational challenges. Only by employing large telescope with high power Na wind/temperature lidar, the above demanding requirements can be satisfied, such as the observations made by the UIUC Na wind/temperature lidar at SOR (with a 3.5-m telescope in diameter), at Maui (with a 3.75-m telescope) and at UAO (with a 1-m telescope). The mean diurnal and annual temperature profiles at UAO demonstrate that in the absence of gravity wave and tidal perturbations, the background atmosphere is statically stable throughout the day and year (Gardner et al., 2002). Thin layers of instability can be generated only when the combined perturbations associated with tides and gravity waves induce large vertical shears in the horizontal wind and temperature profiles. Figure 5.63 shows the atmospheric instability during the night of February 18, 1999 at SOR. The structure and seasonal variations of static and dynamic instabilities in the mesopause region are

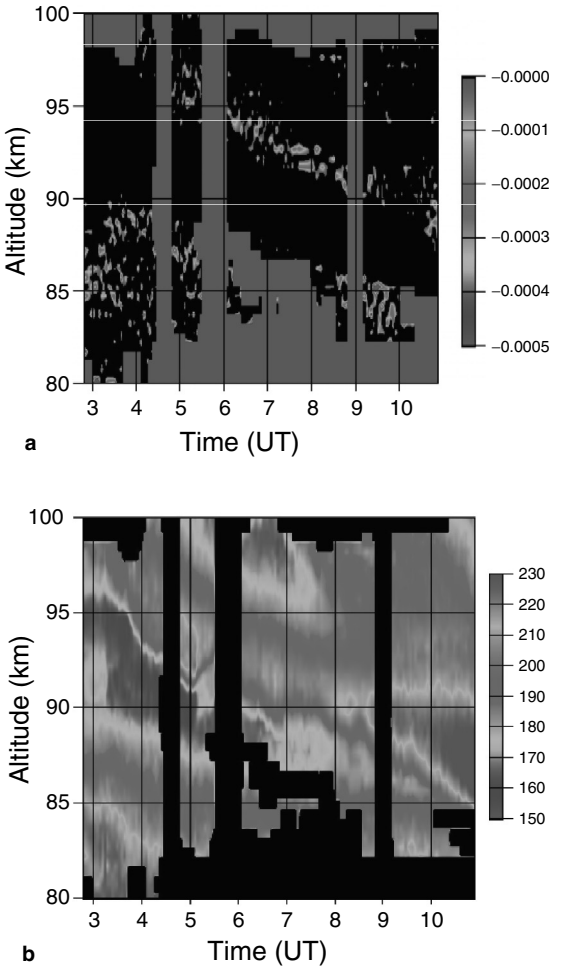


Figure 5.63 (Color figure follows page 398). (a) Convectively unstable regions for February 18, 1999, determined by N^2 calculated from 500-m 1.5-min resolution data taken by the UIUC Na wind/temperature lidar at Starfire Optical Range, New Mexico. The black region is where $N^2 > 0$; colored region is where $N^2 < 0$. No data are available in the purple region. (b) Temperature for the same night. (From Zhao, Y., Liu, A.Z., and Gardner, C.S., *J. Atmos. Sol. Terr. Phys.*, 65, 219–232, 2003. With permission.)

characterized at SOR (Zhao et al., 2003). The annual mean values of the buoyancy frequency N and the total vertical shear S in the horizontal winds are $0.021\,s^{-1}$ and 23 m/s/km, respectively. The probabilities of static and dynamic instabilities are maximum in midwinter averaging about 10 and 12%, and are minimum in summer with the average being around 7 and 5%. The instability probabilities vary considerably from night to night and the structure of the unstable regions is significantly influenced by atmospheric tides. Tides alone are usually not strong enough to induce instability but they can establish the environment for instabilities to develop. As the tidal temperature perturbations propagate downward, they reduce the stability on the topside of the positive temperature perturbation. Instabilities are then induced as gravity waves propagate through this layer with reduced static stability. In a region of reduced stability below the mesopause between 80 and 90 km, the temperature lapse rate is large, the buoyancy parameter N^2 is low, and the vertical heat flux is maximum. It suggests that this is a region of significant wave dissipation (Gardner et al., 2002).

With horizontal wind and temperature profiles simultaneously measured at SOR by the UIUC Na wind/temperature lidar, vertical wavelengths, intrinsic periods and propagation directions of the dominant monochromatic gravity waves in the mesopause region were extracted using the hodograph method (Hu et al., 2002). The results are illustrated in Figure 5.64. A total of 700 monochromatic gravity waves were determined from about 300 h of observations. Among them, 84.4% of the waves were propagating upwards. The mean vertical wavelength was 12.6 and 9.9 km for upward and downward propagating waves, respectively. The intrinsic period is about 10 h.

5.3.2.3. Tidal Study by Full-Diurnal-Cycle Lidar

The CSU Na wind/temperature lidar has been upgraded to a two-beam system capable of simultaneous measurements of temperature and wind in the mesopause region over full diurnal cycles, weather permitting (She et al., 2003; She, 2004). Although this lidar is a modest system with a power-aperture

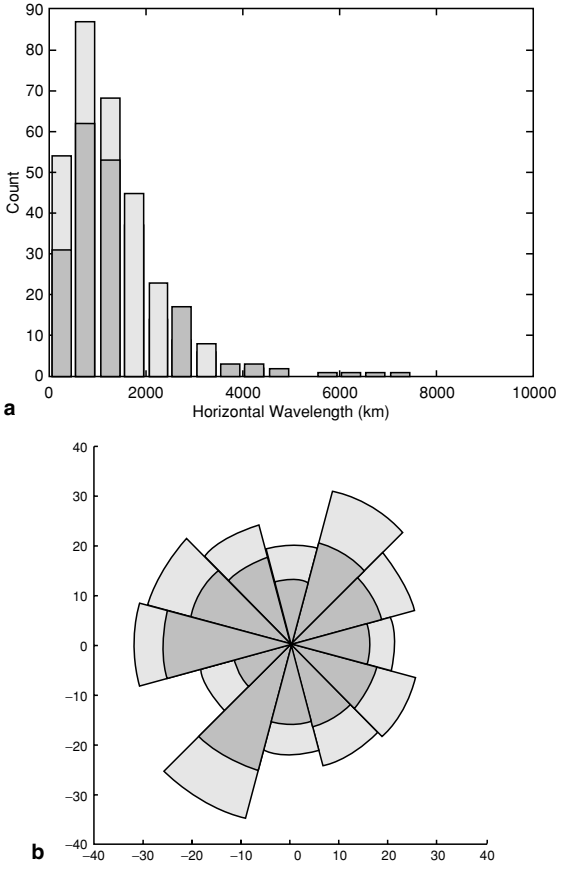


Figure 5.64 Histogram of (a) horizontal wavelengths and (b) propagation direction derived from the UIUC Na wind/temperature lidar data taken at Starfire Optical Range, New Mexico from 1998 to 2000. In plot (b), north is up and east is right. Darker (lighter) shading indicates upward (downward) propagating waves. (From Hu, X., Liu, A.Z., Gardner, C.S., and Swenson, G.R., *Geophys. Res. Lett.*, 29 (24), 2169, doi: 10.1029/2002GL014975, 2002. With permission.)

product of only 0.6 W m^2, good data quality is demonstrated by means of contour plots (Figure 5.65) depicting an 80-h continuous observation between August 9 and 12 showing the existence of atmospheric waves with different periods along with their coherence and interactions. The salient feature of

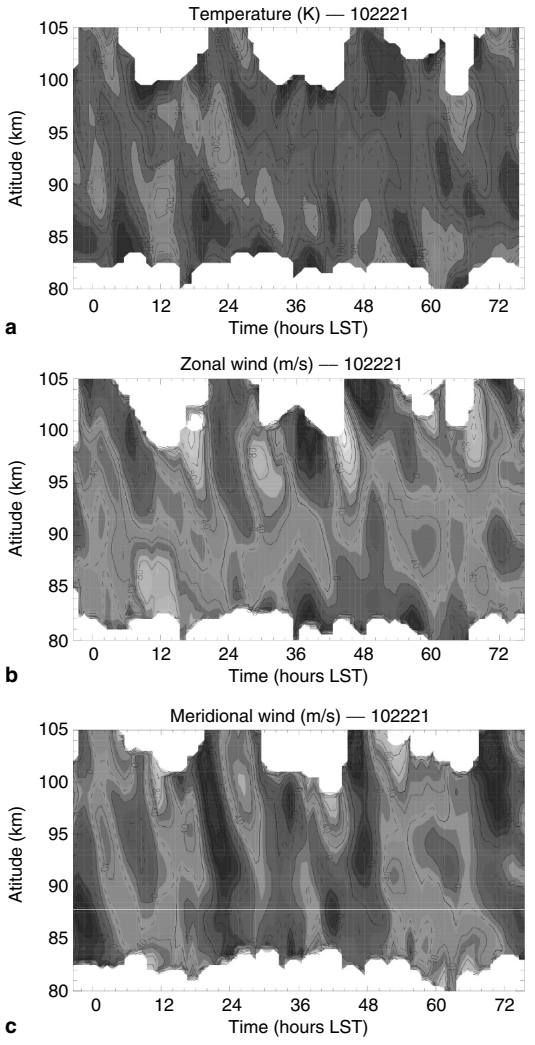


Figure 5.65 (Color figure follows page 398). (a) Mesopause region temperature, (b) zonal wind, and (c) meridional wind obtained by the CSU Na wind/temperature lidar between August 9 and 12, 2002 over Fort Collins, Colorado (41°N, 105°W). The origin of x-axis is 0:00 hours on August 9, 2002, LST. The ranges contours shown are approximately between (a) (160 K, 220 K), (b) (−90 m/s, 120 m/s), and (c) (−90 m/s, 90 m/s). (From She, C.Y., *J. Atmos. Sol. Terr. Phys.,* 66, 663–674, 2004. With permission.)

data with full-diurnal-cycle coverage lies in its ability to describe the vertical profiles of dynamical fields (temperature, zonal, and meridional winds) as a unique linear superposition of diurnal-mean and oscillations with different tidal periods, plus a residual term. In this manner, She (2004) investigated the variability of diurnal-means and of diurnal tidal amplitudes and phases. Using six datasets obtained between July 17 and August 12, each covering a full-diurnal-cycle, considerable day-to-day variability was found, as much as 20 K, 35 m/s, and 75 m/s for diurnal-mean temperature, zonal wind, and meridional wind, respectively, and 15 K and 40 m/s for the diurnal tidal amplitude of temperature and winds. The coherence of tidal excitation and of mean solar forcing is found to prevail over the variability, leading to the convergence to the climatological mean in a multiday observation. While a minimum of three full diurnal cycles appears to be required, the 6-day composite yields means and diurnal tides that resemble model predictions very well. Since the resulting amplitudes and phases of the observed diurnal oscillations agree well with the GSWM00 predictions, She (2003) concluded that the migrating diurnal tide contributes significantly to the observed oscillations with diurnal period over Fort Collins, Colorado (41°N, 105°W). These initial results demonstrate the high potential of the full-diurnal-cycle lidar systems for the atmospheric tides and dynamic studies.

5.3.3 Atmospheric Metallic Layers and Meteor Detection

5.3.3.1. Metal Layers and Mesospheric Chemistry

Meteoric ablation is believed to be the major source of the atomic metal layers that occur globally in the upper mesosphere (Plane, 1991). The resonance fluorescence lidar technique has been used to study the structure and chemistry of the mesospheric metal layers since Bowman et al. (1969) made the first lidar observations of Na layer in 1969. These studies are motivated by the increasing use of metal atoms (such as Na, K, and Fe) as tracers of atmospheric dynamics, chemistry, and temperature. Significant progress has also

been made with laboratory investigations of the chemistry that controls these metal layers. The combined laboratory work and lidar observations have led to the development of models that are able to reproduce the characteristic features of the metal layers.

5.3.3.1.1. Na

The atomic Na is one of the most well-studied species in the middle and upper atmosphere. The broadband resonance Na lidars made remarkable measurements on the Na layers, as described in Section 5.1.3 (Bowman et al., 1969; Sandford and Gibson, 1970; Gibson and Sandford, 1971, 1972; Blamont et al., 1972; Hake et al., 1972; Kirchhoff and Clemesha, 1973; Megie and Blamont, 1977). Since the development of narrowband Na Doppler lidars, the Na layer has been studied more comprehensively combining with chemistry modeling (von Zahn et al., 1987; von Zahn and Hansen, 1988; Hansen and von Zahn, 1990; Plane et al., 1998, 1999; States and Gardner, 1999; She et al., 2000; Clemesha et al., 2002; Gong et al., 2002; Yi et al., 2002).

The IAP group measured the Na layer at Andoya (69°N, 16°E), Norway with their narrowband Na Doppler lidar and discovered the sporadic metal layers as a distinct phenomenon (von Zahn et al., 1987; von Zahn, 1988; Hansen and von Zahn, 1990). The rapid growth of narrow Na layers appears between 90 and 110 km altitude. The Na density increases (sometimes by large factors) within a few minutes in a narrow layer ($\sim$1 km of FWHM). Their appearance shows a strong, positive correlation with that of *f*-type E sporadic layers, but not showing a strong correlation with either riometer absorption or meteor showers (Hansen and von Zahn, 1990).

The University of Illinois operated a broadband Na density lidar (consisting of an excimer laser pumped dye laser) on board an aircraft and measured the summertime Na density profiles during both day and night flights to the Arctic Ocean during ANLC-93 campaign. The University of Illinois also operated another broadband Na lidar (consisting of an Nd:YAG laser pumped dye laser) at the South Pole station and measured the Na density in January 1995 and

1996 (Plane et al., 1998). Measurements at the South Pole reveal a Na layer that has less column abundance and are significantly higher and thinner than at mid-latitudes, as shown in Figure 5.66. Plane et al. (1998) modeled satisfactorily the characteristic features of the South Pole Na layer by choosing a proper rate coefficient for the reaction between sodium bicarbonate and atomic hydrogen. In particular, the model can reproduce the small-scale height of about 2 km observed on the underside of the Na layer.

Using the narrowband Na wind/temperature lidar, the CSU and the UIUC lidar groups made comprehensive measurements and studies of the Na layers at mid-latitudes in Fort Collins, Colorado (41°N) and Urbana, Illinois (40°N) (Kane and Gardner, 1993a; Plane et al., 1999; States and Gardner, 1999; She et al., 2000). Combining the Na data collected at both sites from 1991 through mid-1994, Plane et al. (1999) obtained the seasonal variations of the nighttime Na layers as shown in Figure 5.67. The Na density exhibits a strong annual variation at all heights between 81 and 107 km, with the

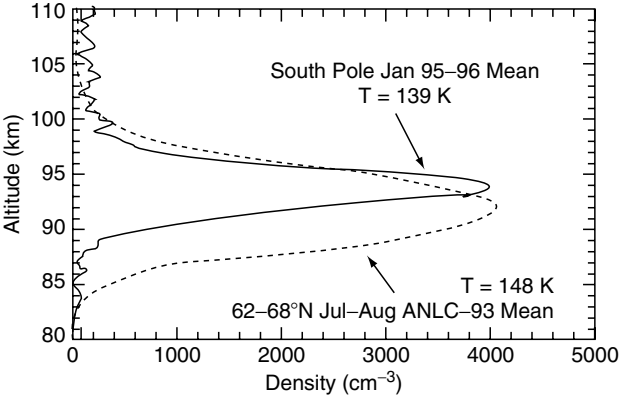


Figure 5.66 Mean Na density profiles computed by averaging all data collected at latitudes between 62° and 68°N during the ANLC-93 flights and mean Na density measured at the South Pole in January 1995 to 1996. (From Plane, J.M.C., Cox, R.M., Qian, J., Pfenninger, W.M., Papen, G.C., Gardner, C.S., Espy, P.J., *J. Geophys. Res.*, 103 (D6), 6381–6389, 1998. With permission.)

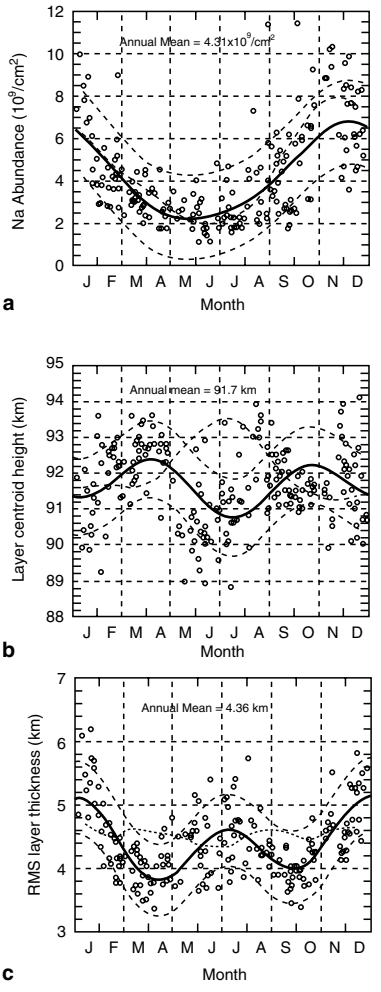


Figure 5.67 The Na layer (a) column abundance, (b) centroid height, and (c) rms width versus data obtained at Urbana, Illinois (40°N) and at Fort Collins, Colorado (41°N). The data were computed from the nightly mean Na density profiles. The thick solid curves are the least-square fits to the data including annual plus semiannual variations. The dashed curves are one standard deviation above and below the fits. The model predictions of these layer properties are shown as thick dotted curves. (From Plane, J.M.C., Gardner, C.S., Yu, J.R., She, C.Y., Garcia, R.R., and Pumphrey, H.C., *J. Geophys. Res.*, 104 (D3), 3773–3788, 1999. With permission.)

column abundance of the layer peaking in early winter and then decreasing by nearly a factor of 4 to a midsummer minimum (Figure 5.67a). The Na layer centroid altitude exhibits strong semiannual variations, with the altitude peaking at spring and fall equinox and minimizing at midsummer (Figure 5.67b). The annual mean profiles of the measured nighttime Na density and temperature are plotted in Figure 5.68. The annual mean Na density profile has a peak density around 3000 cm^{-3} near 92 km. The fundamental features of the nighttime Na layers were confirmed by the 8-year (1991 to 1999) nocturnal Na density measurements at Fort Collins (She et al., 2000) and further measurements (1996 to 1998) at Urbana, Illinois (States and Gardner, 1999). Plane et al. (1999) constructed a seasonal model of the Na layer incorporating the laboratory studies of the pertinent neutral and ionic reactions of the metal Na. The modeled annual mean Na density profile is also plotted in Figure 5.68 for comparison. The model was able to reproduce many observed features of the Na layer remarkably well, including the monthly variation in column abundance and layer shape. Both the observations and the model show that the Na density and temperature are highly correlated below 96 km (with correlation coefficient equal to 0.8 to 0.95), mostly as a result of the influence of odd oxygen/hydrogen chemistry on the partitioning of Na between atomic Na and its principal reservoir species $NaHCO_3$.

With full diurnal coverage, States and Gardner (1999) further studied the seasonal and diurnal variations of the Na layer at 40°N. They found that the seasonal variations in Na density and abundance are influenced primarily by changes in mesopause region temperatures, which are coldest during midsummer. Below 95 km, reactions leading to the bicarbonate sink and the liberation of Na from this reservoir dominate Na chemistry. These reactions and their temperature dependencies are responsible for the large annual variations in Na abundance from 2.3×10^9 to 5.3×10^9 cm^{-2}. The annual mean Na density profiles for both the diurnal mean and nighttime mean are plotted in Figure 5.69. The peak Na density for the annual mean is 3040 cm^{-3} at 91.8 km. The

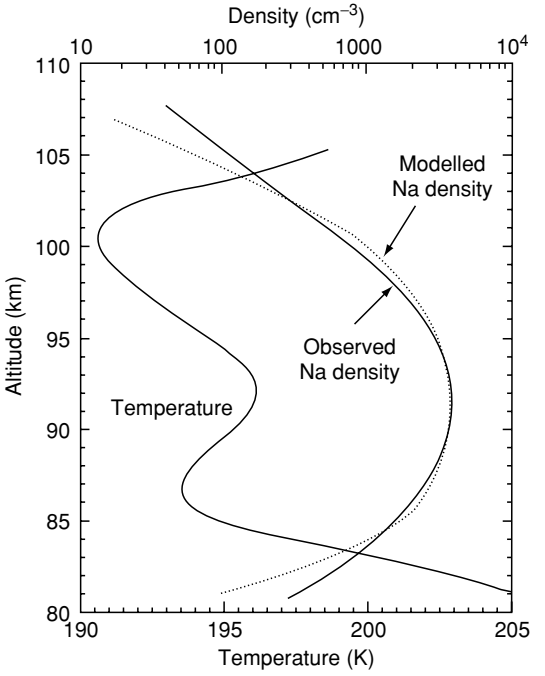


Figure 5.68 Annual mean profiles of the measured Na density (thin solid curve) and temperature (thick solid curve) at 40°N and 41°N. The modeled mean Na density profile (thin dotted curve) is shown for comparison. (From Plane, J.M.C., Gardner, C.S., Yu, J.R., She, C.Y., Garcia, R.R., and Pumphrey, H.C., *J. Geophys. Res.*, 104 (D3), 3773–3788, 1999. With permission.)

largest difference between the diurnal and nighttime means is on the underside of the layer, between 76 and 85 km, where the density is larger for the diurnal mean. This difference is associated with photochemistry, which affects the daytime Na densities. The measured diurnal variations of Na layers are plotted in Figure 5.70 (States and Gardner, 1999). Although there is little evidence of direct tidal perturbations in Na density, 24-h oscillations dominate the diurnal variations. The Na abundance is maximum at sunrise, and the peak-to-peak diurnal variation averages more than 30%. The combined effects of photoionization above 90 km and photochemistry below 85 km induce a strong 24-h oscillation in the layer

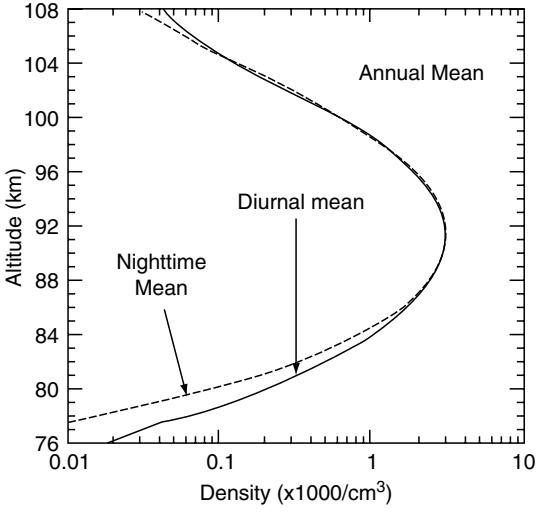


Figure 5.69 Annual mean Na density profiles for observations covering the complete diurnal cycle (solid curve) and observations covering just the nighttime period (dashed curve) at 40°N. (From States, R.J., and Gardner, C.S., *J. Geophys. Res*., 104 (D9), 11783–11798, 1999. With permission.)

centroid altitude. The peak-to-peak variation exceeds 1 km. The lowest centroid height is at local noon when Na densities below 85 km are maximum and photoionization above 90 km is the strongest.

Clemesha et al. (2002) presented the simultaneous measurements of meteor winds and Na density in the 80 to 110 km region at Sao Jose dos Campos (23°S, 46°W) and studied the variations in Na density profiles. They found that the observed vertical oscillations in the Na isopleths are closely correlated with the meridional winds, which confirms their earlier conclusion (Batista et al., 1985) that the Na density oscillations are driven by the solar diurnal and semidiurnal tides. Clemesha et al. (2002) claimed that their conclusion is in direct opposition to those of States and Gardner (1999) that the observed diurnal variations of Na density were the result of solar zenith angle driven changes in photo- and ion-chemistry, but not related to tides. The difference of tidal influence on Na variations may be related to the different amplitude of

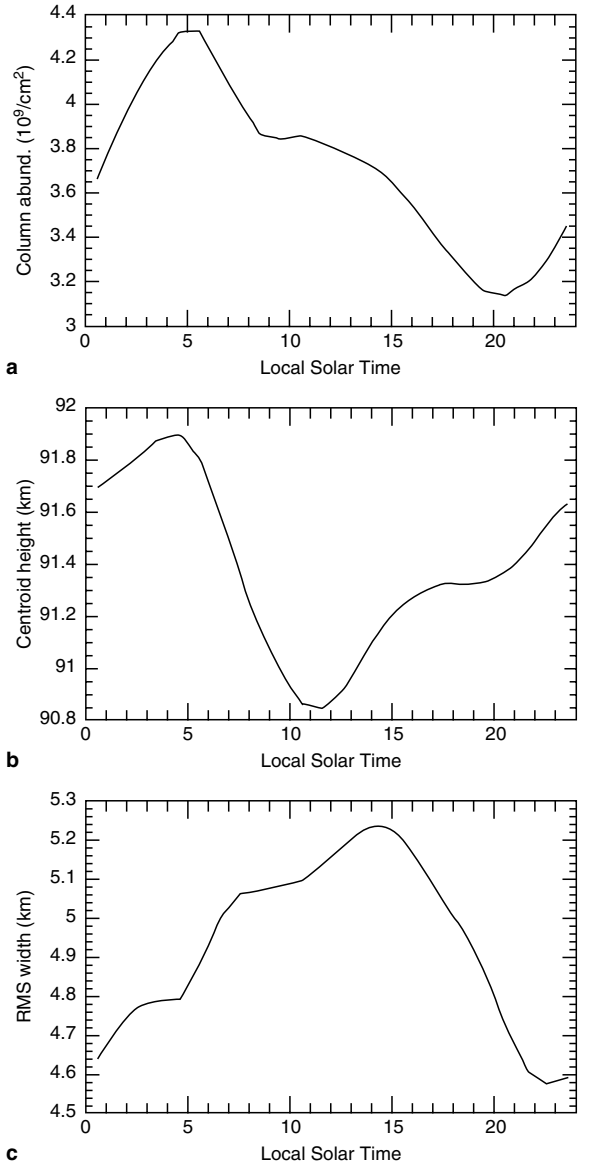


Figure 5.70 Diurnal variation of (a) Na abundance, (b) layer centroid height, and (c) rms layer width at 40°N. These parameters were derived from the annual mean day data. (From States, R.J., and Gardner, C.S., *J. Geophys. Res.*, 104 (D9), 11783–11798, 1999. With permission.)

tides at the locations. Further investigations in both observations and modeling efforts are needed.

Recently, systematic studies of sporadic Na layers (Na_S) were made by Gong et al. (2002) and Yi et al. (2002) with broadband Na lidar systems at Wuhan (31°N, 114°E), China. The sporadic Na layers at Wuhan show some different features than the low and high latitudes. The Na_S layers at Wuhan often exhibit broader layer widths and a longer formation time than those at low and high latitudes (Yi et al., 2002). Gong et al. (2002) found a fairly good correlation between the sporadic Na layer and the sporadic E layers. Through analyzing the observed perturbations during most of the Na_S developments, Gong et al. (2002) suggested that the role of dynamic processes of atmosphere in the Na_S formation should not be ignored. Further investigations are needed to understand the formation mechanism of sporadic Na layers.

5.3.3.1.2. Fe

The first lidar observations of the atomic Fe layers were performed by Granier et al. (1989b) at the Observatoire de Haute-Provence (44°N, 6°E), France. It was followed by the lidar observations of normal and sporadic Fe layers at 40°N (Bills and Gardner, 1990) and at 69°N (Alpers et al., 1990). The Fe layers exhibited considerably different properties in these three measurements, especially on the Fe column abundance values and sporadic Fe occurrence. Densities at the normal Fe layer peak vary from 3000 cm^{-3} (Granier et al., 1989b) to 7000 cm^{-3} (Alpers et al., 1990) to approximately 15,000 cm^{-3} (Bills and Gardner, 1990), while the Fe column abundance is 3×10^9, 9×10^9, and 15×10^9 cm^{-2} for the three measurements, respectively. Sporadic Fe layers were observed by Bills and Gardner (1990) and Alpers et al. (1990), but not shown in Granier et al.'s (1989b) measurements.

Kane and Gardner (1993a) presented the first climatology of the seasonal variations of mesospheric Fe layer at mid-latitude 40°N. The nighttime averages of the Fe layer column abundance, centroid altitude, and rms width, summarized from 75 nights of Fe lidar data from 1989 to 1992, are shown in Figure 5.71 (Kane and Gardner, 1993a). The

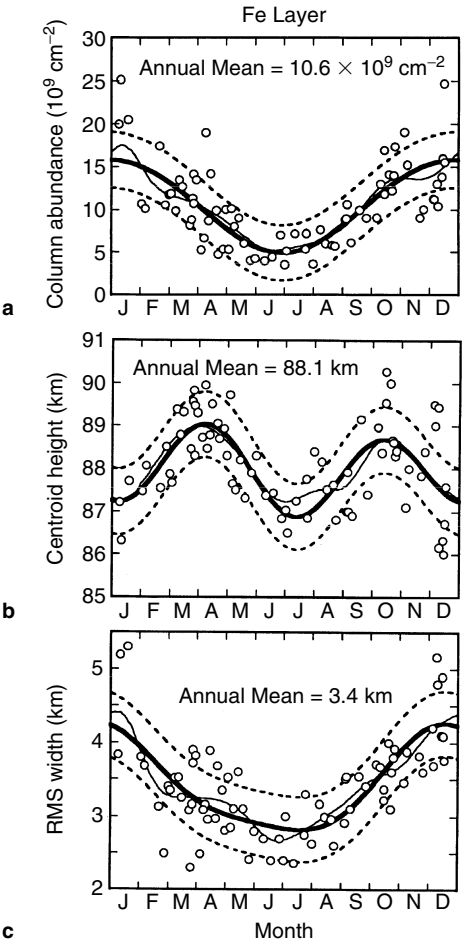


Figure 5.71 Nightly averages of the Fe layer (a) column abundance, (b) centroid height, and (c) RMS width plotted versus date at Urbana, Illinois (40°N). The annual mean Fe layer abundance, centroid height, and width are $10.6 \times 10^9\,\mathrm{cm}^{-2}$, 88.1 m, and 3.4 m, respectively. The thin curves are running means of the data computed using a Hamming window with a full width of 90 days. The dark curves are annual plus semiannual fits to the data, with the dashed curves lying 1 standard deviation above and below the harmonic fit. (From Kane, T.J., and Gardner, C.S., *J. Geophys. Res.*, 98, 16875–16886, 1993. With permission.)

annual mean Fe column abundance is $10.6 \times 10^9\,cm^{-2}$ with the mean centroid altitude at 88.1 km and rms width of 3.4 km. Kane and Gardner (1993a) also provided a climatology of the Na layer seasonal variations: the annual mean Na column abundance is $5.35 \times 10^9\,cm^{-2}$ with the mean centroid altitude at 92.1 km and rms width of 4.4 km. These data show that the Fe layer column abundance is about twice the Na layer abundance, while the Fe peak altitude is about 4 km lower than the Na peak, and the Fe layer RMS width is about 24% narrower than the Na layer width. The Fe layer abundance shows much stronger annual variations than the Na layer, but similar annual variations in the layer centroid altitudes. Both layers reach maximum centroid height around spring and fall equinox.

Sporadic Fe layers occurred in about 27% of the total observation time (Kane and Gardner, 1993a). Sporadic Fe and Na layers were further studied by Alpers et al. (1994) with Bonn University's narrowband Fe lidar and Na lidar at polar, middle, and low latitudes. The comparison between Fe_S and Na_S shows that the overall occurrence rate for Fe_S is much higher than for Na_S; both the normal and sporadic Fe layers are considerably more dynamic than the Na layers; Fe_S layer shape is typically broader, slower growing, and longer lasting than Na_S layer; the correlation with E_S appears to be weaker for Fe_S than for Na_S (Alpers et al., 1994).

The mesospheric Fe layers were measured at both the North Pole and the South Pole by the University of Illinois Fe Boltzmann temperature lidar (Gardner et al., 2001; Chu et al., 2002a). A huge sporadic Fe layer was observed at the North Pole on the summer solstice (June 21, 1999) with peak density of $2.3 \times 10^5\,cm^{-3}$ at 106 km (Figure 5.25). The peak sporadic Fe density is comparable to the meteor trail density observed during Leonid meteor shower (Chu et al., 2000a), but its width ($\sim$10 km) is much wider than a meteor trail ($\sim$250 m). Although this huge sporadic Fe layer is likely linked to sporadic E layer, its abnormal high peak density is not yet completely understood. Further modeling and observation efforts are needed. The annual variation of Fe density at the South

Pole was characterized by the UIUC Fe lidar as shown in Figure 5.72. The South Pole Fe layer has a significant higher peak altitude (~92 km) in summer compared to its winter peak altitude around 88 km.

Raizada and Tepley (2003) present the annual variation of the mesospheric Fe layer for a low latitude at Arecibo Observatory (18.35°N, 66.75°W). The comparison between Arecibo results and the mid-latitude results (Kane and Gardner, 1993a) shows that the annual variation of the height and width of the Fe layers at Arecibo are essentially anticorrelated (opposite) with the mid-latitude observations, while the Fe column abundance lags its mid-latitude counterpart by roughly 3 months. These features are illustrated in Figure 5.73. Raizada and Tepley (2003) suggested that the latitudinal differences in Fe layer properties result from

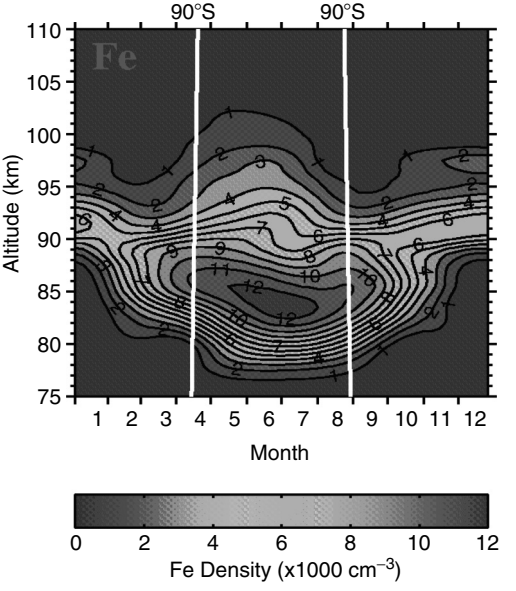


Figure 5.72 (Color figure follows page 398). Seasonal variations of Fe density observed by the UIUC Fe Boltzmann temperature lidar at the South Pole (90°S).

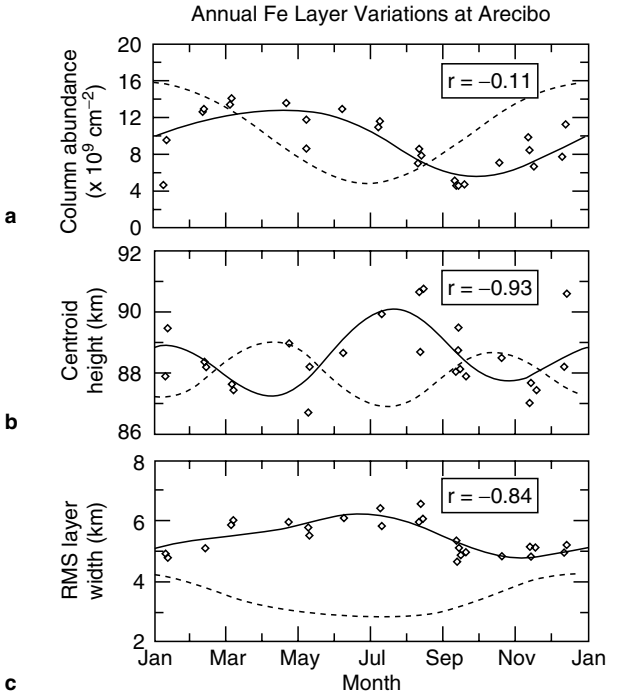


Figure 5.73 Annual variation of (a) column abundance, (b) centroid height, and (c) rms width for Fe measured at Arecibo (18.35°N). A fit (solid line) to the data points is shown together with a comparison to similar observations at 40°N latitude (dashed line). The correlation coefficients, *r*, are listed in each panel. (From Raizada, S., and Tepley, C.A., *Geophys. Res. Lett.*, 30 (2), doi: 10.1029/2002GL016537, 2003. With permission.)

variations in the distribution of mesospheric O_3 and O_2, which in turn affect the Fe chemistry at the two different latitudes. A one-dimensional steady-state model of the mesospheric Fe layer has been developed by Helmer et al. (1998) for the mid-latitude 40°N. It considered three processes thought to be most influential in affecting the atomic Fe layer: the Fe deposition via meteor ablation, the vertical distribution of the ablated material, and the partitioning of the metal along 10

constituent species as a result of neutral and ion gas phase reactions in dynamic steady state. Most observed features of the Fe layer can be successfully reproduced by the model (Helmer et al., 1998).

5.3.3.1.3. K

The K layer was first measured by Felix et al. (1973). Fifteen nights of lidar observations made by Megie et al. (1978) did not show significant seasonal variations of the K column and peak densities. After the development of narrowband K Doppler lidar, comprehensive measurements of the K layer and its seasonal variations were made by the IAP group (Eska and Höffner, 1998; Eska et al., 1998, 1999; Fricke-Begemann et al., 2002) and the Arecibo group (Friedman et al., 2002). Eska et al. (1998) present the seasonal variations of the mesospheric K layers over Kühlungsborn, Germany (54°N, 12°E) based on the 110 nights data obtained between June 1996 and June 1997. The measured nightly mean K layer column abundance, centroid height, and rms width are plotted in Figure 5.74. The averaged K column abundance is 4.4×10^7 cm^{-2}, with peak altitude at 90.5 km and rms width of 4.0 km. The nightly mean column and peak densities exhibit dominantly semiannual variations with maxima in summer and winter. The centroid height varies semiannually too, attaining highest altitudes during the equinoxes. The rms layer width shows a strong annual variation with a maximum width in winter. The seasonal variations of the Na and Fe layers (Kane and Gardner, 1993a) are compared with the K layer in Figure 5.74. For all three metal layers, their centroid heights are dominated by semiannual variations with quite similar phases. The column abundances are dominated by a semiannual variation in the case of the K layer but are dominated by an annual variation for the Na and Fe layers. The layer rms widths are dominated by annual variations in the case of the K and Fe layers, while the Na layer exhibits a semiannual variation (Eska et al., 1998).

Observations of the nighttime K layer were performed on the German research vessel *Polarstern* from March to June

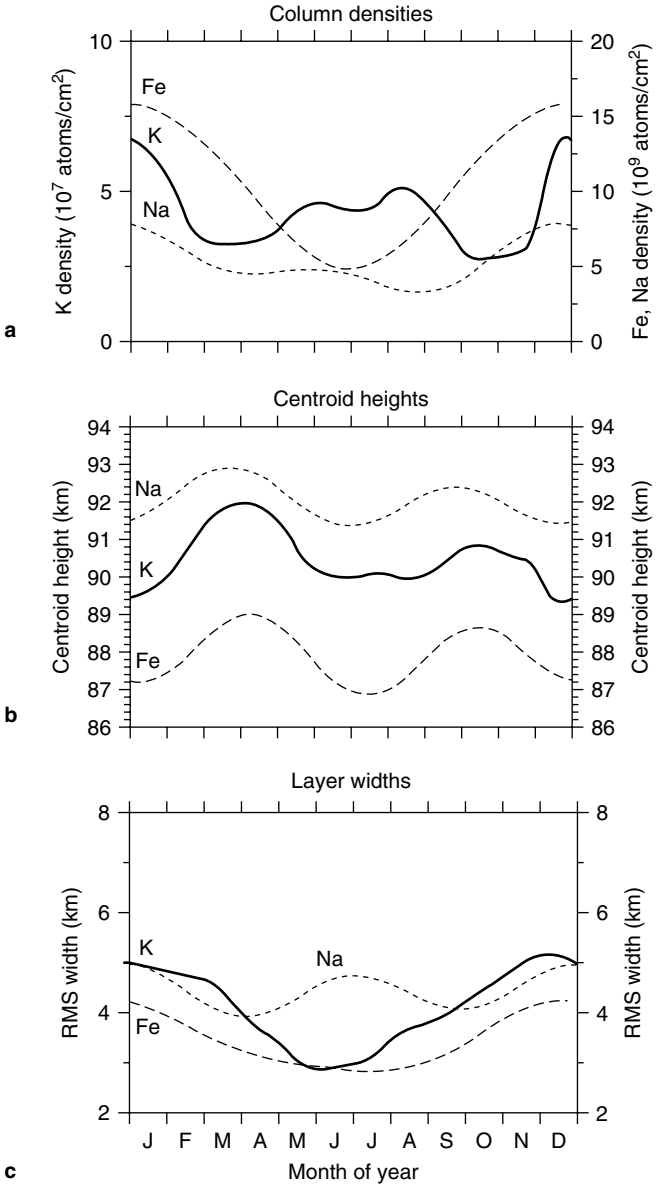


Figure 5.74 (a) Column density, (b) centroid height, and (c) layer RMS width for K (solid curve), Na (dotted curve), and Fe (dashed curve) plotted versus month of the year. (From Eska, V., Höffner, J., and von Zahn, U., *J. Geophys. Res.*, 103 (A12), 29207–29214, 1998. With permission.)

1996. The K densities were obtained between 71°S and 45°N. The nightly mean peak densities varied from 140 cm^{-3} in the equatorial region to 10 cm^{-3} in the Antarctic, and the column abundances decreased from 1.2×10^8 to 1.3×10^7 cm^{-2} for low to high latitudes (Eska et al., 1999). The global mean peak height of the normal K layer was found to be 88.3 km, which is lower than what reported in Eska et al. (1998) for Kühlungsborn (54°N) observations. A one-dimensional model of the K layer was developed, which includes meteoric deposition, vertical transport through eddy diffusion, and a full chemical scheme (Eska et al., 1999). This model was able to reproduce very satisfactorily the seasonal behavior of the K layer at 54°N if the wintertime deposition flux of the metal was reduced by 30% compared to the summer. The mid-latitude ratio of K to Na was about 1%, much less than either the chondritic or cosmic ratios of the two metals (~8 or 6%, respectively). The most likely reason is that potassium vaporizes less efficiently from meteoroids than sodium. The model was generally very successful in reproducing the latitudinal variations in the K layer (Eska et al., 1999).

The observations of the K layer at Arecibo Observatory (18.3°N, 66.75°W) show quite different features of the K layer compared to the Kühlungsborn results (Friedman et al., 2002). At 18°N, there is less variability in the layer width and a generally lower peak height with no winter minimum. More sporadic K layers were observed at this low-latitude site than were seen at higher latitude. However, the semiannual variation of K column abundance was seen at both the low and mid-latitudes. Further studies are needed to understand the differences in the behavior of the K layer at different latitudes.

5.3.3.1.4. Ca and Ca^+

The Ca and Ca^+ layers have been studied by lidars at mid-latitudes 44°N (Granier et al., 1985, 1989a), 40°N (Gardner et al., 1993; Qian and Gardner, 1995), and 54°N (Alpers et al., 1996; Gerding et al., 2000) and at a low-latitude 18°N (Tepley et al., 2003). All the lidar systems used for these measurements consist of either Nd:YAG laser pumped or

excimer laser pumped dye lasers (broadband) working at 423 nm for Ca and at 393 nm for Ca^+. The comprehensive observations of the Ca and Ca^+ layers were obtained by ground-based lidars at Kühlungsborn (54°N, 12°E), Germany between December 1996 and December 1998. Gerding et al. (2000) studied the data on 112 nights of Ca and 58 nights of Ca^+ during this period. The Ca layer has an average column abundance of $2.1 \times 10^7\,cm^{-2}$ with a mean peak density of $22\,cm^{-3}$ at 89.9 km altitude. The Ca^+ layer has a column abundance of $4.9 \times 10^7\,cm^{-2}$ with a mean peak density of $29\,cm^{-3}$ at 91.9 km. The high variability of the layer profiles was found as a basic feature of the Ca layer and Ca^+ layer. The column density of Ca is influenced by both annual and semiannual variations with a maximum in summer and autumn/winter. Ca^+ shows even more obvious short-time variations of sporadic layers. Gerding et al. (2000) also presented a one-dimensional steady-state chemistry model of the nighttime Ca and Ca^+ layers, based on the laboratory studies of CaO reaction kinetics. This model can reproduce satisfactorily the characteristic features of the annual mean layers and can provide a possible explanation for the unusual seasonal variation of the Ca layer, which exhibits a pronounced summertime enhancement around 87 km.

These metal layer observations and chemistry modeling provide us more and deeper understanding of the atmosphere constituents, their variations, and their relations to the temperature, solar activity, and atmosphere dynamics.

5.3.3.2. Meteor Trail Detection by Lidar

The daily influx of extraterrestrial matter into the Earth's upper atmosphere has been deduced to be of the order of 100 tons per day. The ionization and ablation of meteors give rise to a number of research topics including radar meteor echo, sporadic-E formation, and optical meteor investigation. In order to understand all these processes, it is necessary to know the origin, the influx rate, the chemical composition, size distribution, and incoming velocities of this extraterrestrial matter as well as the temporal and spatial variations of

these properties. The lidar observations give us the capability to quantify the abundances of a number of elements in these meteor trails and to allow a characterization of the ablation processes. Meteoric ablation is also believed to be the major source of the neutral metal layers of Na, Fe, K, Ca, and Li, etc. Lidar studies of the meteor trails also seek further support for this belief. If this belief is indeed true, one would expect that the abundances of the metals in this region and their temporal and spatial variations should reflect, in some fashion, the properties of the incoming meteoroid population (von Zahn et al., 2002).

Lidar observations of five short-lived meteor trails were first reported by Beatty et al. (1988) with a broadband Na lidar at Urbana (40°N), Illinois. She et al. (1991) reported about a single meteor trail observed with a narrowband Na Doppler lidar. Kane and Gardner (1993b) reported a comprehensive study of 101 Na meteor trails and five Fe meteor trails observed during 4 years at Urbana, Illinois; Arecibo, Puerto Rico; and near Hawaii. They found that the vertical distribution of these 106 meteor trails is approximately Gaussian shape with a centroid height of 89.0 $\pm$ 0.3 km and an rms width of 3.3 $\pm$ 0.2 km. The 1996 Leonid meteor trails were studied by the IAP narrowband K lidar (Höffner et al., 1999) and a March 1997 meteor shower was studied by simultaneous and common-volume observations of a K and a Ca lidar (Gerding et al., 1999). Grime et al. (1999) reported a first two-laser-beam observation of a meteor trail on March 23, 1998 by a broadband Na lidar at Arecibo. The lidar was operating with two beams probing different volumes of the Na layer separated zonally by 15.7 $\pm$ 0.8 km. A single meteor trail was observed near 89 km altitude in both lidar field of views with a 310 $\pm$ 50 s temporal displacement. The observational separation suggests a westward zonal wind of 50 $\pm$ 10 m/s, while trail dispersion yielded an upper bound for the total diffusion coefficient of 2.6 $\pm$ 0.5 m^2/s, which is consistent with dispersion seen in other meteor trails.

The 1998 Leonid meteor shower was comprehensively studied by a few different lidars at different locations: Chu

et al. (2000a) observed 18 Fe meteor trails over Okinawa (26.3°N, 127.7°E) with an airborne Fe Boltzmann temperature lidar; Chu et al. (2000b) observed seven persistent Na meteor trails at SOR (35°N, 106.5°W) with a steerable Na wind/temperature lidar coupled with a 3.5 m diameter optical telescope; von Zahn and Bremer (1999) observed the 1998 Leonid meteor trails by a cluster of K, Ca, and Fe lidars. Additional studies of the 1998 Leonid meteor showers observed at SOR using the Na lidar data and imagers were reported by Kelley et al. (2000), Grime et al. (2000), and Drummond et al. (2001). A comprehensive summary of the lidar observations of meteor trails was given by von Zahn et al. (2002).

The University of Illinois Fe Boltzmann temperature lidar was installed on the NSF/NCAR Electra aircraft and flew above Okinawa on the night of November 17/18, 1998 as part of the 1998 Leonid Multi-Instrument Aircraft Campaign (Chu et al., 2000a). During the 8-h Leonid flight, 18 meteor trails were observed with this Fe lidar. One of the most dense meteor trails probed by the airborne lidar over Okinawa is plotted in Figure 5.75 with the peak density reaching $2.33 \times 10^5\,\mathrm{cm}^{-3}$. The average altitude of the 18 trails from the high velocity (72 km/s) Leonid meteors, 95.67 ± 0.93 km, is approximately 6.7 km higher than previously observed for slower (~30 km/s) sporadic meteors (Kane and Gardner, 1993b). This height difference is consistent with the assumption that meteor ablates when atmospheric drag reaches a critical threshold. The average abundance of the trails is 10% of the abundance of the background Fe layer. Observations suggested that the 1998 Leonid shower did not have a significant impact on the abundance of the background Fe layer (Chu et al., 2000a).

SOR is located on the Kirtland Air Force Base near Albuquerque, New Mexico. The facility is operated by the Air Force Research Laboratory, Directed Energy Directorate and includes a 3.5-m astronomical telescope. The University of Illinois Na wind/temperature lidar is coupled to this telescope through the coude optics so that the beam can be pointed in any direction above 5° elevation. A 1-m diameter portion of

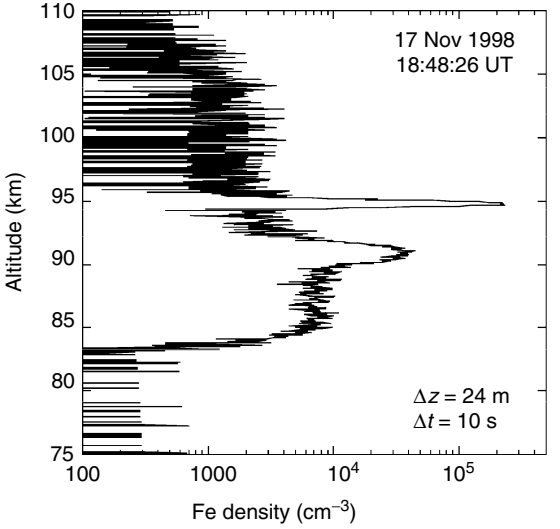


Figure 5.75 Fe meteor trail profile probed by the UIUC airborne lidar on November 17, 1998 over Okinawa. The vertical resolution is 24 m and integration period is 10 s. The dense narrow layer at 94.8 km is a meteor ablation trail, while the broader layer at 91 km is a sporadic Fe layer. (From Chu, X., Pan, W., Papen, G., Gardner, C.S., Swenson, G., and Jenniskens, P., Characteristics of Fe ablation trails observed during the 1998 Leonid meteor shower, *Geophys. Res. Lett.*, 27, 1807–1810, 2000a. With permission.)

the telescope primary mirror is used to project the laser beam while the remainder is used for collecting the backscattered light and focusing it onto the detector. The beam divergence is approximately 1 mrad. At 100 km range, the beam diameter (full width at e^{-2}) is 100 m. The lidar operates at 30 pps and the laser output power varies between 1 and 1.5 W. For this Leonid experiment, the range resolution was 24 m and the integration period was 15 s. The lidar was operated in the normal scanning mode to measure radial wind, temperature, and Na density profiles at zenith and 10° off-zenith to the north, east, south, and west. These data are used to derive profiles of the three wind components of the atmosphere as well as the temperature and Na density profiles. An observer

was positioned outside the telescope dome to look for persistent trails associated with fireballs. When one was spotted, the telescope operator was apprised of the approximate azimuth and elevation angle and the telescope was moved to that position. By using a bore sighted video camera mounted on the telescope and the lidar profile data, the telescope operator positioned the lidar beam on the persistent trail. In this way, seven persistent trails were tracked and probed with the lidar for as long as 30 min. Visible CCD cameras, all-sky airglow imagers, and meteor radar also collected correlative data at SOR during the meteor shower. The persistent trails observed on 17/18 November at SOR were given descriptive names that characterized their visual appearances in the video and CCD images. Figure 5.76 is a false color CCD image of the one of

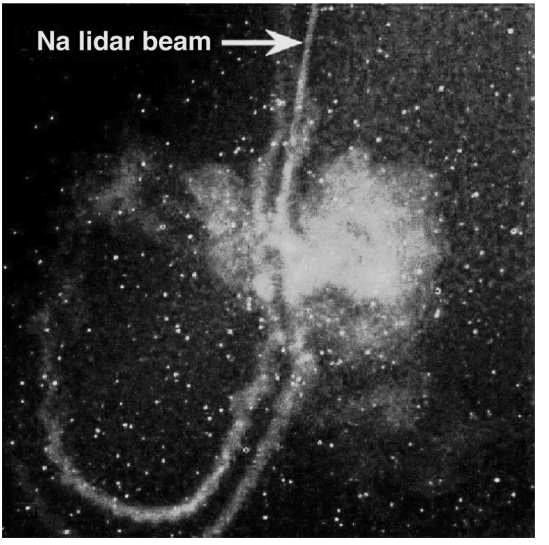


Figure 5.76 (Color figure follows page 398). CCD image of the persistent chemiluminescence trail (Diamond Ring) observed at 09:30 UT November 17, 1998 at the Starfire Optical Range, New Mexico. The straight line near the top center of the image is the UIUC Na lidar beam. (From Chu, X., Liu, A.Z., Papen, G., Gardner, C.S., Kelley, M., Drummond, J., and Fugate, R., *Geophys. Res. Lett.*, 27, 1815–1818, 2000b. With permission.)

these trails (Diamond Ring) observed at 09:30 UT on 17 November. The bright trail is chemiluminescence associated with a variety of emissions including Na and OH airglow (Kelley et al., 2000). The corresponding lidar profile for the diamond-ring trail is shown in Figure 5.77. The peak density of the Na trail reached 3.41×10^5 cm^{-3}. The duration time of the persistent Na meteor trails observed at SOR ranges from 3 min to more than 30 min, with average peak altitude of 94.0 $\pm$ 1.6 km (Chu et al., 2000b). The average Na abundance within the trails was 52% of the background Na layer abundance, which suggests the corresponding masses of the meteors were from 1 g up to 1 kg. Elevated temperatures were observed by the lidar in the bright chemiluminescent meteor

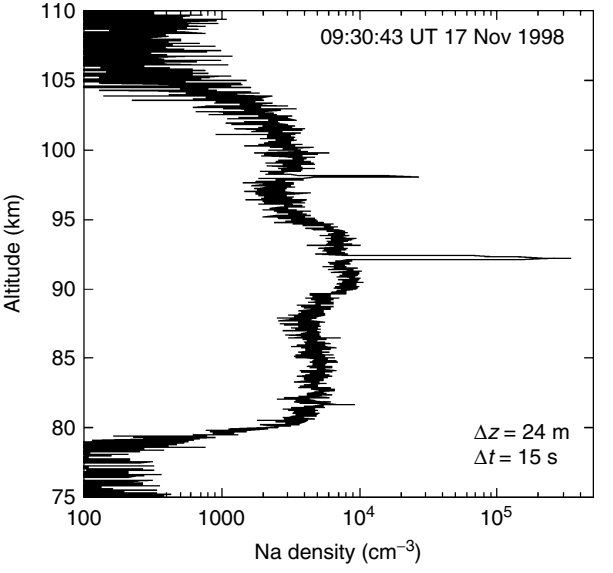


Figure 5.77 Na meteor trail profile of the Diamond Ring shown in Figure 5.76, probed by the UIUC Na wind/temperature lidar at SOR during 1998 Leonid meteor shower. The vertical resolution is 24 m and the integration period is 15 s. (From Chu, X., Liu, A.Z., Papen, G., Gardner, C.S., Kelley, M., Drummond, J., and Fugate, R., *Geophys. Res. Lett.*, 27, 1815–1818, 2000b. With permission.)

trails. Approximately 3 min after ablation, the temperature at the tube walls was about 20 to 50 K warmer than the tube core and background atmosphere (Chu et al., 2000b). Neither chemical nor frictional heating provides a satisfactory explanation for the observations.

5.3.4 Polar Mesospheric Clouds (Noctilucent Clouds)

5.3.4.1. Historical Perspective

Polar mesospheric clouds (PMCs) and their visual counterparts noctilucent clouds (NLCs) are the thin scattering layers (water–ice particles) occurring in the mesopause region ($\sim$85 km) at high latitudes ($\sim$poleward of 50°) in both hemispheres mainly during the three months surrounding summer solstice when the temperatures fall below the frost point ($\sim$150 K) (Chu et al., 2003). NLCs were first identified as a distinct atmospheric phenomenon in 1885 when the summer-night unusually bright clouds came to people's attention throughout Europe, Great Britain, and Russia (Backhouse, 1885). The idea that PMC particles are composed of water ice was first suggested by Humphreys (1933), and then later taken up by Hesstvedt (1961, 1962), Chapman and Kendall (1965), Charlson (1965, 1966), Reid (1975), and Gadsden (1981). More comprehensive models including microphysics of ice formation were developed by Turco et al. (1982), Jensen and Thomas (1988, 1994), Jensen et al. (1989), Thomas (1996), Klostermeyer (1998, 2001), Rapp et al. (2002), Berger and von Zahn (2002), and von Zahn and Berger (2003a). Satellite observations provided the first evidence of the water–ice composition of PMC particles and the enhanced water vapor layer in the PMC region (Hervig et al., 2001; Summers et al., 2001; Stevens et al., 2001).

Due to the sudden appearance of NLCs in 1885 and their close relations to the mesopause conditions on temperature and water vapor (Thomas, 1991), PMCs/NLCs are generally regarded as a sensitive tracer of middle and upper atmospheric water vapor and temperature, and possibly as an

indicator of long-term global climate change (Thomas, 1991; Avaste, 1993). Since the first observation of NLCs (Backhouse, 1885), PMCs/NLCs have been studied from ground by numerous visual, photographic, spectroscopic, photometric, and polarimetric observations for more than a century as reviewed by Gadsden and Schröder (1989). NLCs are generally visible only within the latitude range of 50 to 65° at night when the mesopause region is still sunlit. The silver-blue NLCs are typically cirrus-like clouds rippled by complex waves and billows, and can occasionally provide a spectacular visual display (Gadsden and Schröder, 1989). PMCs/NLCs have also been investigated by *in situ* rocket experiments (Witt, 1969; Gumbel and Witt, 1998) and by space observations from the satellites OGO-6, SME, NIMBUS 7, UARS, SPOT 3, MSX, and METEOSAT 5 (Donahue et al., 1972; Thomas and Olivero, 1989; Thomas et al., 1991; Evans et al., 1995; Debrestian et al., 1997; Carbary et al., 1999; Gadsden, 2000) in both hemispheres. Satellite observations in the 1970s led to the discovery of extensive scattering layers (PMCs) near 85 km over the entire north and south polar caps during local summers (Donahue et al., 1972). NLCs are believed to be the "ragged edges" of these much more persistent and extensive PMC layers that are visible at lower latitudes. The geographic coverage, occurrence frequency, and brightness of PMCs/NLCs have been increasing in the past decades. Recently, NLCs have been observed at mid-latitude such as Utah (41°N) (Taylor et al., 2002; Wickwar et al., 2002). The observed increases in the geographic extent and brightness of PMCs and NLCs in polar regions during the past decades have been related to the decreasing mesopause temperature and increasing mesospheric water vapor concentration possibly caused by the rising level of mesospheric CO_2 and CH_4, respectively (Thomas, 1996). The possible link of PMC to the global climate change has intensified the research on PMCs.

The passive optical observations from ground and satellite observations from space have made great contributions to the study of PMC/NLC. However, these observations either

have limited vertical resolution of about 2 to 3 km or difficulties in covering the 24-h diurnal cycles. Although *in situ* rocket experiments can measure PMC vertical distribution precisely, they are so sporadic that they cannot be used to characterize the seasonal and diurnal variations in PMC properties. The active lidar sounding technology is unique in detecting and studying PMCs because of its high vertical resolution, daytime measurement capability, full seasonal and diurnal coverage, and atmospheric temperature and gravity wave measurement capability. The ability to detect PMC throughout the 24-h diurnal cycle is one of the key advantages that lidars have for PMC/NLC studies, since passive observations are limited to twilight conditions and satellite observations are either limited by their Sun-synchronous orbits or the need for illumination of the NLC region by the Sun.

The first lidar observation of NLCs in the northern hemisphere was made by Hansen et al. (1989) at Andoya (69°N, 16°E), Norway with the University of Bonn narrowband Na Doppler lidar, and the first lidar observations of PMCs in the southern hemisphere was made by Chu et al. (2001a,b) at the South Pole (90°S) with the University of Illinois Fe Boltzmann temperature lidar. Numerous lidar observations have been made in the northern hemisphere at the North Pole (90°N) and the Gulf of Alaska (58°N, ~150°W) (Gardner et al., 2001), Svalbard (78.2°N, 15.4°E) (Höffner et al., 2003), Andoya (69°N, 16°E) (Hansen and von Zahn, 1994; Langer et al., 1995; Nussbaumer et al., 1996; Lübken et al., 1996; von Cossart et al., 1997, 1999; von Zahn et al., 1998; von Zahn and Bremer, 1999; Baumgarten et al., 2002a,b; Fiedler et al., 2003), Sondrestrom (67.0°N, 50.9°W) (Thayer et al., 1995, 2003; Hecht et al., 1997; Gerrard et al., 1998), Poker Flat (65°N, 147°W) (Collins et al., 2003), Juliusruh (54.6°N, 13.4°E) (von Cossart et al., 1996), Kühlungsborn (54°N, 12°E) (Alpers et al., 2000, 2001), Aberystwyth (52.4°N, 4.1°W) (Thomas et al., 1994), and Logan (41.7°N, 111.8°W) (Wickwar et al., 2002). Lidar observations of PMCs have also been made in the southern hemisphere at the South Pole (90°S) (Chu et al., 2001a,b, 2003; Gardner et al., 2001), Rothera (67.5°S,

68°W) (Chu et al., 2004), and Davis (68.6°S, 78°E) (Klekociuk et al., 2003). These lidar observations provide precise information on PMC/NLC altitude, layer vertical structure, and backscatter coefficient as well as valuable information on particle size and shape, seasonal and diurnal variations, and wave signatures.

5.3.4.2. PMC Characteristics Measured by Lidar

The detection of PMCs/NLCs by lidar arises from the Mie scattering of transmitted laser photons by PMC ice particles. Without PMC particles in the mesopause region, the Rayleigh scattering from pure atmosphere molecules decreases exponentially with height due to the decreasing number density of atmosphere molecules. The Rayleigh scattering becomes negligibly small at altitudes above 80 to 85 km. When PMC particles present in the mesopause region around 85 km, the Mie scattering from the PMC water–ice particles that are much larger than atmosphere molecules is significantly stronger than the Rayleigh scattering at the same altitude. Thus, a sharp, narrow peak will stand out in the lidar photon count profile. Figure 5.27 provides an example of PMC profile obtained by the Fe Boltzmann temperature lidar at the South Pole on January 18, 2000. The photon counts have been converted to the volume backscatter coefficient. The narrow peak at 85.5 km is a PMC layer with peak volume backscatter coefficient approximate 8×10^{-9} m^{-1} sr^{-1}.

The lidar detection limit/threshold of PMCs is primarily determined by the volumetric backscatter properties of the PMC particles and the performance capabilities of the lidar system for background noise compression, laser wavelength, and power. The challenge of detecting PMCs in the polar region comes from the high solar background. Only when the PMC backscatter peak is higher than the background, is it possible to distinguish PMC from background noise. Narrowband filters are needed in the lidar receivers to effectively reject the sky background. The volume backscatter coefficient of PMC is approximately inversely proportional to the fourth power of light wavelength. Thus, the shorter wavelength,

such as the Fe lidar operating at 374 nm, will produce much stronger PMC backscattering, compared to normal visible light, such as the Rayleigh lidar operating at 532 nm and the Na lidar at 589 nm. The background level and the PMC signal level together determine the lidar detection sensitivity. The combination of daytime measurement capability, short operation wavelength, and high output power is necessary to make a lidar capable of PMC detection during the polar summer. The volume backscatter coefficient of PMCs is approximately proportional to the sixth power of the particle radius, which means the larger PMC particles will produce stronger backscatter. For the lidar wavelength ranges from near UV to visible light, PMC particles can be detected by lidar only when their radius in excess of about 20 nm (von Cossart et al., 1999). Therefore, lidars observe the larger radii of the water–ice particles in the mesopause region.

Commonly, the parameters used in the lidar field to characterize PMCs are the volume backscatter coefficient, the backscatter ratio, the total backscatter coefficient, the centroid altitude, and the layer rms width. The definition of volume backscatter coefficient $\beta(z)$ is the same as Equation (5.4). The backscatter ratio $R(z)$, the total backscatter coefficient β_{Total}, the centroid altitude Z_{C}, and the layer rms width σ_{rms} are defined as

$$R(z) = \frac{\text{Total signal}(z)}{\text{Molecular signal}(z)} = \frac{\beta_{\text{PMC}}(z) + \beta_{\text{R}}(z)}{\beta_{\text{R}}(z)} \tag{5.129}$$

$$\beta_{\text{Total}} = \int \beta_{\text{PMC}}(z)\mathrm{d}z \tag{5.130}$$

$$Z_{\text{C}} = \frac{\sum_i z_i \beta_{\text{PMC}}(z_i)}{\sum_i \beta_{\text{PMC}}(z_i)} \tag{5.131}$$

$$\sigma_{\text{rms}} = \sqrt{\frac{\sum_i (z_i - Z_{\text{C}})^2 \beta_{\text{PMC}}(z_i)}{\sum_i \beta_{\text{PMC}}(z_i)}} \tag{5.132}$$

The data processing methods and the equations used to compute these parameters from lidar photon count profiles could be found in Chu et al. (2003).

5.3.4.2.1. Mean characteristics

Numerous lidar and satellite observations have shown that the PMC/NLC occurrence varies with latitude and season. Although PMCs/NLCs have been observed by lidar from the North Pole to 41°N and from the South Pole to 67.5°S, the PMC/NLC occurrence in polar latitudes is more frequent, persistent, and stronger while the mid-latitude PMCs/NLCs are less frequent, more sporadic, and weaker. PMCs/NLCs usually occur in the summer, from late May to the end of August in the northern hemisphere and from late November to the end of February in the southern hemisphere. This is because the PMCs contain water–ice particles that only form in the saturation region. The very small amount of water vapor in the mesopause region (typically 3 to 5 ppmv) requires very cold temperatures to reach saturation. Lidar, rocket, and satellite measurements have shown that the polar mesopause temperature is coldest in summer but warmer in winter. This is caused by the atmosphere dynamic interaction — upwelling in summer and downwelling in winter resulting from the global mean meridional circulation. These upwelling and downwelling motions result in adiabatic cooling and heating in summer and winter, respectively. The upwelling is strongest at the poles and becomes weaker towards lower latitudes, so the polar summer mesopause temperature is coldest in the atmosphere, often lower than 150 K in the polar region, which is usually the frost point at mesopause.

Systematic lidar measurements of PMCs can reveal the mean characteristics of PMC layers. Shown in Figure 5.78 are the histograms of the measured hourly mean PMC peak backscatter ratio $R_{\max}$, peak volume backscatter coefficient $\beta_{\max}$, total backscatter coefficient β_{Total}, layer centroid altitude Z_C, and layer rms width σ_{rms} obtained by the Fe Boltzmann temperature lidar during the 1999 to 2000 and 2000 to 2001 austral summer seasons at the South Pole (Chu et al., 2003a). More than 430 h worth of data were collected at the South Pole. The mean peak backscatter ratio, peak volume backscatter coefficient, total backscatter coefficient, centroid altitude, and layer RMS width are 58.4, 3.75×10^{-9} m^{-1} sr^{-1},

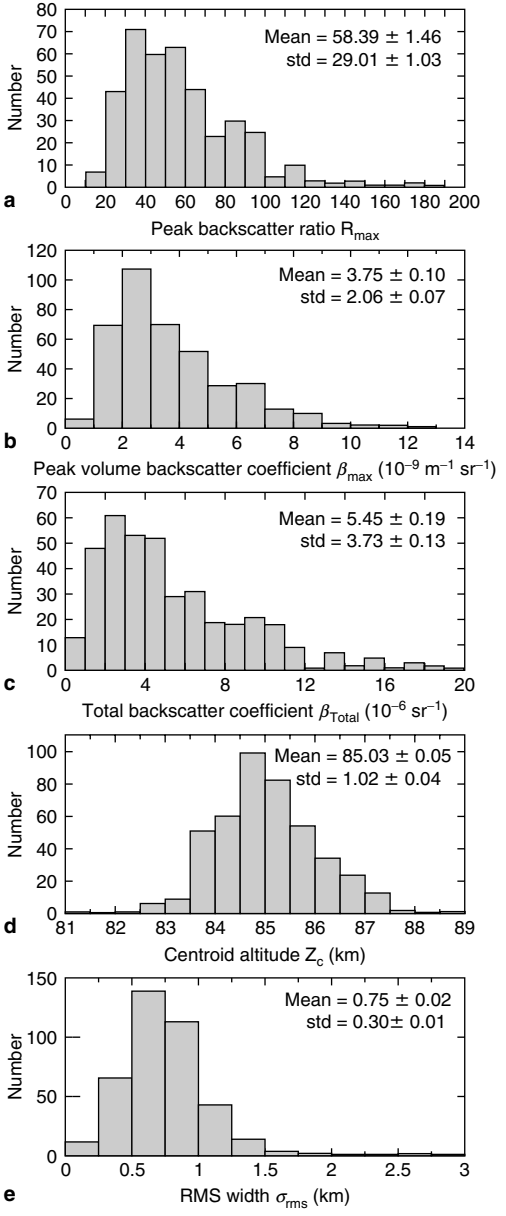


Figure 5.78 Histogram of PMC characteristics obtained by the Fe Boltzmann temperature lidar in the 1999 to 2001 summer seasons at the South Pole. (From Chu, X., Gardner, C.S., and Roble, R.G., *J. Geophys. Res.*, 108 (D8), 8447, doi: 10.1029/2002JD002524, 2003. With permission.)

$5.45 \times 10^{-6}\,\mathrm{sr}^{-1}$, 85.03 km, and 0.75 km, respectively. The overall PMC occurrence frequency at the South Pole from 1999 to 2001 is 67.4%. Besides the South Pole, extensive lidar measurements of PMCs/NLCs have been made at Rothera by the same Fe Boltzmann temperature lidar, at Svalbard by the IAP K Doppler lidar, and at Andoya and Sondrestrom by two Rayleigh lidars. The lidar detection of PMC/NLC made at other locations is much less frequent. A comparison of mean characteristics of PMCs observed by lidars at the South Pole, Rothera, Svalbard, Andoya, and Sondrestrom is given in Table 5.10.

When comparing the PMC brightness (expressed as the peak volume backscatter coefficient), the wavelength dependence of the backscatter coefficient must be taken into account. We assume a lognormal size distribution of spherical ice particles with radius of 40 nm and a width of 1.4 as used in Höffner et al. (2003). Then the color ratios of 374/770 and 374/532 nm are 7.3 and 2.3, respectively. These two ratios are used when converting the 532 nm (for Andoya and Sondrestrom) and 770 nm (for Svalbard) backscatter coefficient to 374 nm as of the South Pole and Rothera. The PMC layer's FWHM was converted to rms width by assuming a Gaussian distribution of the layer vertical structure so that $\mathrm{FWHM} = \sqrt{8 \ln 2}\sigma_{\mathrm{rms}}$. Most of the locations exhibit PMCs around 83 to 85 km. Table 5.10 indicates that the PMCs at higher latitudes (such as the South Pole and Svalbard) usually have higher occurrence frequency, higher brightness, and higher centroid altitude compared to lower latitude PMCs.

Table 5.10 Comparison of Mean Characteristics of PMC Observed by Lidars at the South Pole, Rothera, Svalbard, Andoya, and Sondrestrom

	South Pole (90°S)	Rothera (67.5°S)	Svalbard (78°N)	Andoya (69°N)	Sondrestrom (67.5°N)
$\beta_{\max}$ ($\times 10^{-9}$ m^{-1} sr^{-1})	3.75	0.98	2.85	2.19	2.17
Z_{C} (km)	85.03 ± 0.05	83.7 ± 0.25	83.6	83.3	82.9
σ_{rms} (km)	0.75 ± 0.02	0.83 ± 0.06	0.72	0.51	0.38
Occurrence frequency (%)	67.4	17.3	77	43	20

5.3.4.2.2. Geographic difference in PMC altitudes

A significant hemispheric difference in PMC altitudes was found through our lidar observations over both poles (Chu et al., 2001a, 2003a; Gardner et al., 2001). The mean altitude, computed from 437 h of PMCs observed in two seasons at the South Pole, is 85.03 km, which is approximately 2 km higher than the commonly observed PMC altitude (~83 km) in the northern hemisphere. To further explore the geographic differences in PMC altitudes, we plot in Figure 5.79 the mean PMC altitudes measured by lidars and ground-based triangulation versus latitude for both the southern and northern hemispheres. The lidar data were from the South Pole (90°S) (Chu et al., 2003); Davis (68.6°S) (Klekociuk et al., 2003); Rothera (67.5°S) (Chu et al., 2004); North Pole

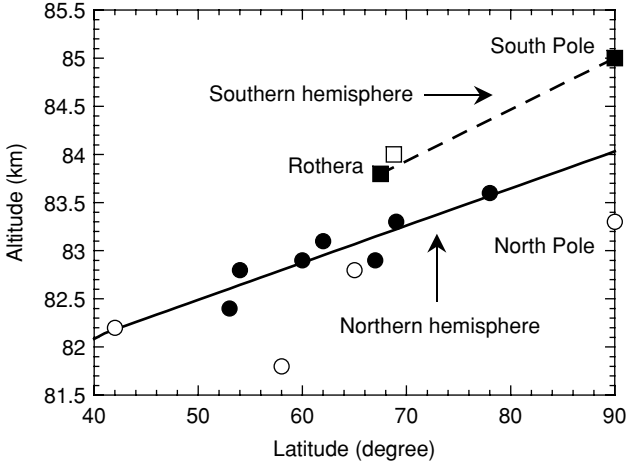


Figure 5.79 PMC altitude versus latitude in both hemispheres. Circles are for the northern hemisphere, and squares are for the southern hemisphere. Filled circles indicate large data sets and are used to obtain the linear fit (solid line). Open circles indicate limited data sets and are excluded in the fit. The open square is for the Davis data. See text for references. (From Chu, X., Nott, G.J., Espy, P.J., Gardner, C.S., Diettrich, J.C., Clilverd, M.A., and Jarvis, M.J., *Geophys. Res. Lett.*, 31, L02114, doi: 10.1029/2003GL018638, 2004. With permission.)

(90°N), and Gulf of Alaska (58°N) (Gardner et al., 2001); Svalbard (78°N) (Höffner et al., 2003); Andoya (69°N) (Fiedler et al., 2003); Sondrestrom (67°N) (Thayer et al., 2003); Poker Flat (65°N) (Collins et al., 2003); Juliusruh and Kühlungsborn (55/54°N) (von Cossart et al., 1996; Alpers et al., 2000, 2001); and Logan (42°N) (Wickwar et al., 2002). Three other northern data points were adapted from von Zahn and Berger (2003b), in which they summarized the triangulation measurements of NLC altitudes from several sources including Grahn and Witt (1971) for 62°N, Gadsden and Schröder (1989) for 60°N, and Jesse (1896) for 53°N. The lidar data at 90°S, 67.5°S, 78°N, 69°N, 67°N, 65°N, and 55/54°N were averaged over many hours for one or more seasons. There were only a few hours of observations at 90°N, 65°N, 58°N, and 42°N. These four points (open circles) were excluded in the linear fit (solid line) to the northern PMC altitudes, but are plotted for reference. The data at 90°N, 65°N, and 58°N fall below this trend line, while the data from 42°N lie on the line. The two southern points (filled squares) were obtained from many days of measurements by the same Fe lidar at the South Pole and Rothera. The single Davis (68.6°S) observation (open square), obtained on January 9, 2002, is slightly above the dashed line linking the Rothera and South Pole means. Both the southern and northern hemisphere trend lines indicate that PMC altitudes are generally higher at higher latitudes. What is more interesting in Figure 5.79 is that the southern line is significantly higher than the northern line. The altitude difference between the two hemispheres is about 1 km for similar latitudes. We conclude from these measurements that there is a significant inter-hemispheric difference in PMC altitudes. The southern hemisphere PMCs are about 1 km higher than their northern counterparts observed at similar latitudes.

The formation and evolution of PMC particles are a complex process that depends on the temperature, water vapor, and vertical wind structures of the mesopause region. PMC particles actually consist of water ice, and the water ice particles are only formed in the supersaturation region where water vapor pressure P_w is higher than saturation water

vapor pressure P_s, or in other words, the atmosphere temperature T is lower than the frost point T_s. Once they fall into the subsaturation region where $P_w < P_s$, the ice particles will be destroyed quickly by sublimation due to higher temperatures at lower altitudes (Thomas, 1991). Therefore, PMC altitudes will be determined primarily by the altitude of the supersaturation region. To study the possible causes for this hemispheric difference in PMC altitudes, TIME-GCM simulation data are used to represent the temperature, water vapor, and vertical wind structures of atmosphere at the poles. The supersaturation regions at the South and the North Poles predicted by TIME-GCM are plotted in Figure 5.80 along with the two seasons of PMC data at the South Pole. The lines denote the boundary where the atmosphere temperature is equal to the frost point. The scattered points in Figure 5.80

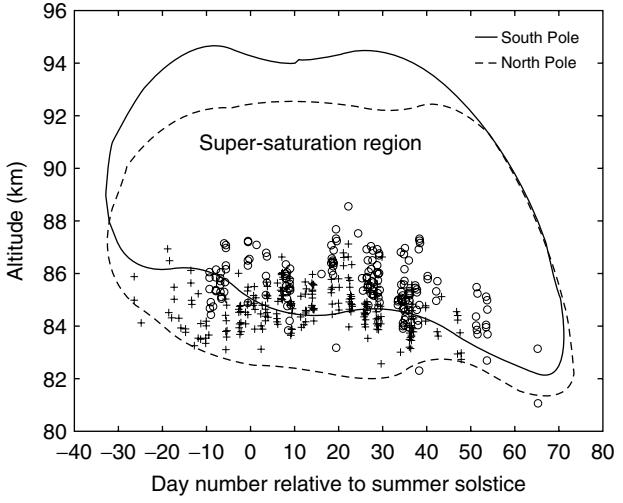


Figure 5.80 Supersaturation region predicted by the TIME-GCM at the South and North Poles. The lines denote the boundary where the atmosphere temperature is equal to the frost point. Scattered data points are South Pole PMC data obtained by the Fe Boltzmann temperature lidar in (○) 1999 to 2000 season and (+) 2000 to 2001 season. (From Chu, X., Gardner, C.S., and Roble, R.G., *J. Geophys. Res.*, 108 (D8), 8447, doi: 10.1029/2002JD002524, 2003. With permission.)

are the lidar data of PMC at the South Pole. They mainly stay within the saturation region clustered along its bottom. Comparing the saturation regions at the South and North Poles, we find that the South Pole saturation region is generally 2 km higher than the North Pole saturation region. This hemispheric difference in super-saturation region altitudes is approximately equal to the hemispheric difference in PMC altitudes found by the lidar measurements over both poles. Once PMC particles form in the saturation region around mesopause, under gravitational pull, the PMC particles will slowly settle down through the supersaturation region while growing in size and mass. The particles also experience a buoyancy force mainly due to the upwelling atmosphere over the summer pole. When the buoyancy force of the upwelling atmosphere balances the gravitational force on the PMC particle, PMC particles will stay around that altitude, which is several kilometers below mesopause, and PMC layers will be observed there. Because atmospheric density is smaller at higher altitude and this buoyancy provided by the upwelling air is proportional to the product of atmospheric density and vertical wind, a higher vertical wind is required to maintain the South Pole PMCs staying at 2 to 3 km higher altitudes than the North Pole if PMC particle mass and drag coefficient are similar at both poles. Checking with the TIME-GCM predictions, the vertical wind on the 13th day after summer solstice is about 3.5 cm/s around 85 km at the South Pole, which is larger than the vertical wind of 2.5 cm/s around 83 km at the North Pole. The South Pole mesopause (∼91 km) is about 2 km higher than the North Pole mesopause (∼89 km), just like the South Pole saturation region is about 2 km higher than that of the North Pole, as shown in Figure 5.80. The PMC layers are higher at the South Pole because they begin forming in the higher saturation region near the higher mesopause. As they grow in size and slowly fall to lower altitudes within the supersaturation region, the larger vertical wind of upwelling atmosphere at the South Pole would provide sufficient buoyancy to maintain the PMC particles at altitudes that are about 2 km higher than at the North Pole. Thus, the South Pole PMCs are expected to be

higher than the North Pole PMC as we observed. Since the upwelling is strongest at the pole and becomes weaker towards lower latitudes, it is not surprising that the PMC appears to be lower at lower latitudes as shown in Figure 5.79.

Checking with the simulation of TIME-GCM, the hemispheric differences in mesopause altitudes and upwelling wind are mainly caused by the solar flux in January being 6% greater than the solar flux in July because of the Earth's orbital eccentricity where the Earth is closest to the Sun on January 3 and farthest from the Sun on July 5. The increased solar heating at the South Pole during austral summer results in a warmer stratopause that causes a dynamic adjustment. The dynamic adjustment results in higher mesopause altitude and stronger upwelling vertical winds in summer at the South Pole than at the North Pole. These differences cause the supersaturation region to be at higher altitudes over the South Pole than over the North Pole (Chu et al., 2003). Besides the hemispheric differences in PMC altitudes, the lidar observations we made in two different years also indicate major inter-annual differences that the PMC in 1999 to 2000 season is about 0.8 km higher than the 2000 to 2001 season. Undoubtedly, the model solar forcing variation alone cannot account for the observed inter-annual variability at the South Pole. It is most likely to be dynamically forced and related to gravity wave and planetary wave variability. Nevertheless, it is believed that the main process responsible for the hemispheric differences is related to the difference in solar heating caused by the Earth's orbital eccentricity and that there is considerable dynamic variability superimposed the measurements.

5.3.4.2.3. Inter-annual and seasonal variations

The backscatter coefficient and centroid altitude of PMC observed at the South Pole exhibit considerable seasonal variations. Figure 5.81 shows the smoothed total backscatter coefficient, centroid altitude, and rms width for the two-season data taken at the South Pole. The smoothing full width is 11 days. It clearly shows that the 2000 to 2001 season PMCs are

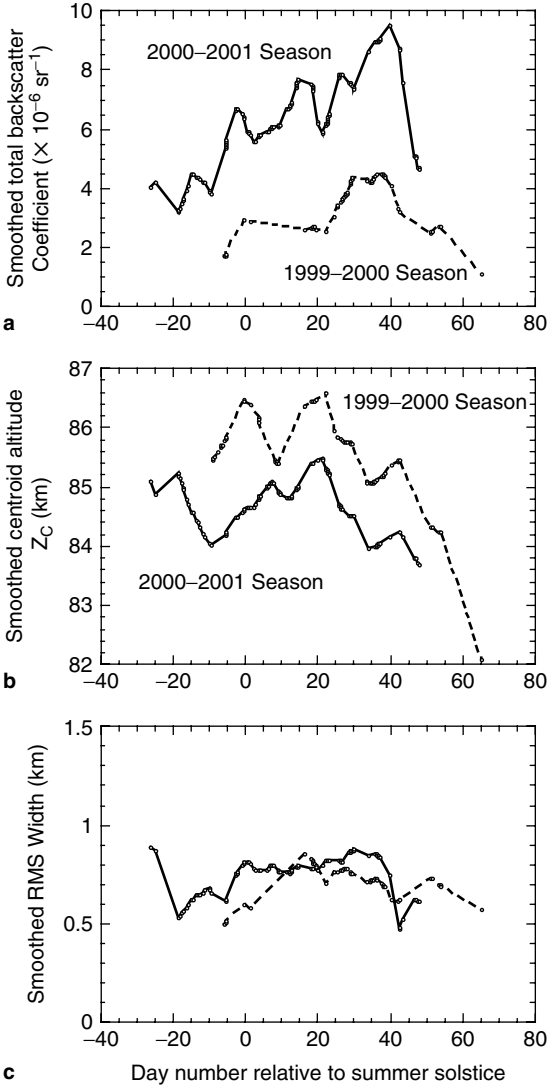


Figure 5.81 Eleven-day smoothed PMC parameters versus day number at the South Pole: (a) smoothed total backscatter coefficient; (b) smoothed centroid altitude; and (c) smoothed RMS width. Circles are the smoothed data points, dashed line is for the 1999 to 2000 season and solid line is for the 2000 to 2001 season. (From Chu, X., Gardner, C.S., and Roble, R.G., *J. Geophys. Res.*, 108 (D8), 8447, doi: 10.1029/2002JD002524, 2003. With permission.)

much brighter than the ones in the 1999 to 2000 season, and both seasons reach maximum brightness in the period of 25 to 40 days after summer solstice, which is also the period when the PMC occurrence probabilities are maximum during both seasons.

Significant seasonal trends emerge from the scattered PMC altitude data as shown in Figure 5.81(b). During both seasons, since 10 days before summer solstice, the altitudes generally increase as the season progresses, reach maximum values around 10 to 20 days after summer solstice, and then decrease throughout the remainder of the season. PMC altitudes exhibit similar descending rate from about 18 to 50 days after solstice in both seasons. PMC altitudes decrease by about 64 m/day from day 18 to day 53 after solstice in the 1999 to 2000 season (Chu et al., 2001a), while the decreasing rate is 65 m/day from day 18 to day 48 for the 2000 to 2001 season data. At the beginning of the 2000 to 2001 season, PMC altitudes start at a relatively high value of 85 km, and then decrease to slightly below 84 km around 10 days before solstice. Due to the lack of data in that period, this feature does not appear in the 1999 to 2000 season data. There is no obvious seasonal trend in the PMC layer rms width as shown in Figures 5.81(c). The rms width is approximately homogeneous through the PMC seasons.

The PMCs at Andoya (69°N) do not show the same seasonal trends as the South Pole (Fiedler et al., 2003). Similar analysis for PMC seasonal variations has not been done for other locations. It is an open question whether the seasonal variations of PMC at lower latitudes are different from the South Pole. Further investigations are needed. If they are indeed different, it would indicate differences in atmosphere dynamics between the pole and lower latitudes.

5.3.4.2.4. Diurnal variations

The full-diurnal-coverage of lidar detection of PMCs allows the study of PMC diurnal variations. Significant diurnal variations have been observed in PMC backscatter coefficients and altitudes at Andoya (69°N) (von Zahn et al., 1998) and at the South Pole (Chu et al., 2001b) as shown in Figure 5.82.

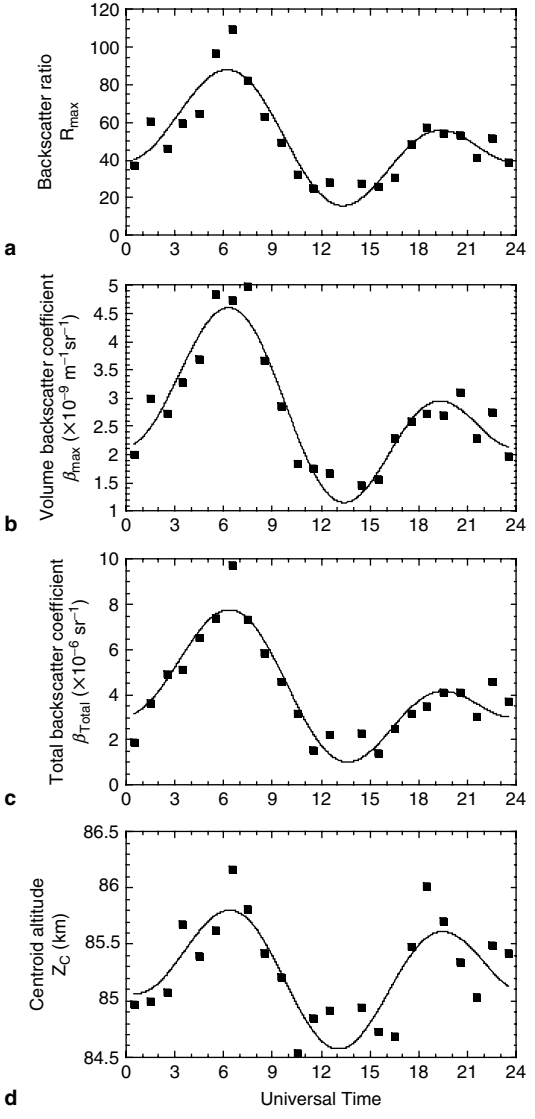


Figure 5.82 Diurnal variations of the South Pole PMC (a) peak backscatter ratio, (b) peak volume backscatter coefficient, (c) total backscatter coefficient, and (d) layer centroid altitude plotted versus UT. The solid curves are the 12 and 24 h harmonic fits. (From Chu, X., Gardner, C.S., and Papen, G., *Geophys. Res. Lett.*, 28, 1937–1940, 2001b. With permission.)

A clear in-phase relationship between backscatter coefficients and altitudes was found at the South Pole (Chu et al., 2001b). In contrast, a clear out-of-phase relationship was observed at Andoya (von Zahn et al., 1998). The observed diurnal variation at Andoya was attributed to a stable semidiurnal tide (von Zahn et al., 1998), while the diurnal variation at the South Pole was attributed to tidal variations in temperature and vertical wind caused by a zonally symmetric tide (Chu et al., 2001b).

More data obtained in the second season at the South Pole allow us to further investigate the diurnal variations. An overall average of all data in both seasons versus UT hour shows that there is an in-phase relation for the 12-h oscillation and an out-of-phase relation for the 24-h oscillation between PMC backscatter coefficient and centroid altitude. The backscatter coefficient variation has similar amplitudes in semidiurnal and diurnal variations, but the altitude variation has a larger amplitude in diurnal variation than in semidiurnal variation. The observed overall diurnal oscillations at the South Pole may be explained as the following: the positive-correlation between backscatter coefficient and the altitude is caused by the in-phase semidiurnal variation, while the anti-correlation between the maxima of backscatter coefficient and altitude is determined by the out-of-phase diurnal variation.

The PMC total backscatter coefficients are plotted versus PMC centroid altitude in Figure 5.83(a) for the South Pole. The circles are for 1999 to 2000 season, and the crosses are for 2000 to 2001 season. The plot is quite different from those observed at Andoya shown in Figure 5.83(b) (Fiedler et al., 2003). At Andoya, the largest backscatter coefficients are observed at the lowest altitudes, which is consistent with the current theory of PMC particles forming at higher altitudes near the mesopause and then falling to lower altitudes as they grow in size and mass. At the South Pole, the distribution is roughly symmetric, and PMCs exhibit the largest backscatter coefficients near the center of the altitude distribution around 85 km. The backscatter coefficients tend to be smaller both above and below the mean altitude. The aforementioned

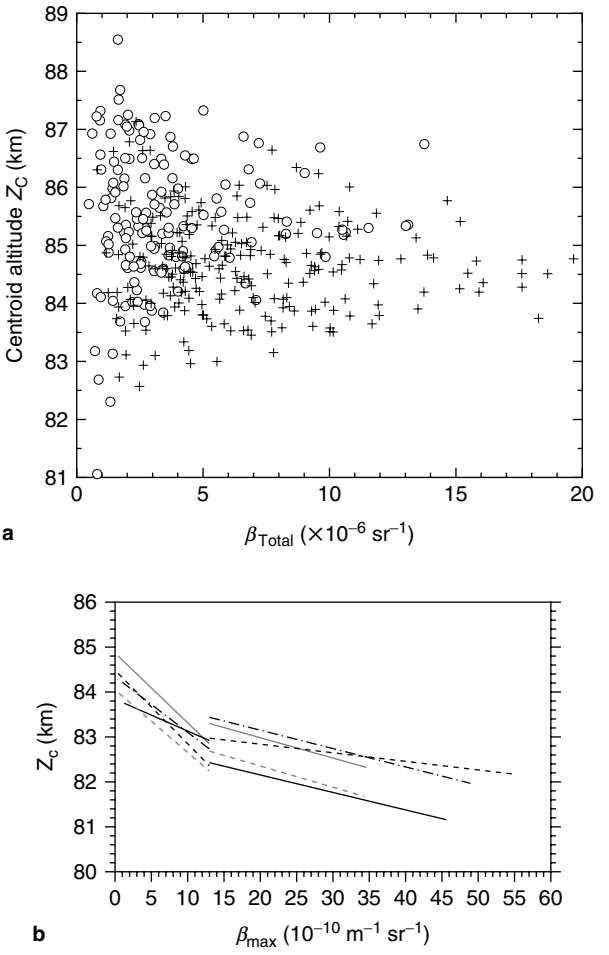


Figure 5.83 (a) PMC total backscatter coefficients versus centroid altitudes for the 1999 to 2000 (○) and 2000 to 2001 (+) summer seasons at the South Pole. (From Chu, X., Gardner, C.S., and Roble, R.G., *J. Geophys. Res.*, 108 (D8), 8447, doi: 10.1029/2002JD002524, 2003.) (b) Relation between altitude and brightness of PMC above ALOMAR at Andoya (69°N, 16°E), Norway. Linear fits for each year are shown: 1997 (black solid line), 1998 (black dashed line), 1999 (black dashed–dotted line), 2000 (gray dashed line), and 2001 (gray solid line). The brightness has been subdivided into two ranges, and the border between them is $\beta_{max} = 13 \times 10^{-10}\,m^{-1}\,sr^{-1}$. (From Fiedler, J., Baumgarten, G., and von Cossart, G., *J. Geophys. Res.*, 108 (D8), 8453, doi: 10.1029/2002JD002419, 2003. With permission.)

differences observed between the South Pole and lower latitudes need modeling investigations to understand.

5.3.4.2.5. Wave signature in PMCs

Gerrard et al. (1998) observed wave structures in NLCs with a period of 2 to 3 h, and wave perturbations in stratosphere at the same time. Their data indicate that the increasing stratospheric wave perturbations may cause the decreasing of NLC backscatter values. They suggested that wave activity has a negative effect on NLCs (either by inducing sublimation on existing NLC particles, reducing their radius, or by limiting the creation of new NLC particles). Hecht et al. (1997) also observed an NLC sublimating due to the passage of an acoustic gravity wave through 85 km. Thayer et al. (2003) reported that gravity-wave signatures were routinely observed in the NLC detection. Upon estimating stratospheric wave activity in the lidar data, they observed stronger cloud backscatter during low gravity-wave activity and weak cloud backscatter during high gravity-wave activity.

5.3.4.2.6. PMC particle size, number density, and shape

Multicolor lidar observations can determine PMC/NLC particle size based upon the assumptions of spherical water ice particles and monomodal lognormal size distribution. Most results show that the diameter of PMC particles is 20 to 51 nm, the width is 1.2 to 1.6, and the number density is 82 to 610 cm^{-3} (von Cossart et al., 1997, 1999; Alpers et al., 2000). Alpers et al. (2001) reported a strange NLC that showed a lack of backscattering at 770 nm while showing strong scattering at three other shorter wavelengths. This phenomenon cannot be explained by backscattering on any size distribution of homogeneous water–ice spheres. Baumgarten et al. (2002a) deduced the shape of PMC particles with a polarization lidar technique. They measured the polarization state of 532 nm laser light backscattered from NLC particles. The observed depolarization, averaged over the altitude range of 84.2 to 85.5 km, was $1.7 \pm 1.0\%$. The layer of enhanced depolarization was centered 1 km above the maximum

of the PMC layer. A comparison of the observed depolarization with that calculated for cylinder-shaped PMC particles shows that the observed depolarization can be explained by the presence of elongated particles with a length-over-diameter ratio larger than 2.5. This experiment indicates that the depolarizing particles occur primarily in a region where PMC particles grow.

5.3.4.2.7. PMC and PMSE

Simultaneous and common volume lidar/radar/rocket observations show that PMCs and PMSEs (polar mesosphere summer echoes) can be either tightly coupled or loosely coupled (Nussbaumer et al., 1996; von Zahn and Bremer, 1999; Stebel et al., 2000). In the tightly coupled cases, PMCs and PMSEs occur at the same time, and PMCs exist along the bottom of PMSEs. Goldberg et al. (2001) showed that for loosely coupled cases, PMCs and PMSEs can be either temporally or spatially separated. These observations suggest that PMCs and PMSEs may have a common origin. However, PMSEs prefer large number density but small subvisible charged aerosols, whereas PMCs/NLCs prefer small number density but relatively large visible particles.

5.3.4.2.8. PMC and Fe depletion in summer at the South Pole

PMCs are a nearly persistent phenomenon at the South Pole in summer. The observed PMCs are commonly located around 85 km, which overlaps in altitude with the mesospheric Fe layers produced by the ablation of meteoroids entering the atmosphere. The Fe Boltzmann temperature lidar can observe both the Fe layer (by resonance fluorescence scattering) and the PMC (by Mie scattering) with high vertical resolution. When PMCs occur, the Fe data would be "contaminated" by the PMCs around 85 km. By taking the difference between 372- and 374-nm channel densities and then scaling back to the 372-nm density, we obtained the pure Fe density shown in Figure 5.84 as an example on January 19, 2000 (Chu et al., 2002b). An obvious Fe bite-out occurs at the PMC peak altitude. Plane et al. (2004) explained this phenomenon by the

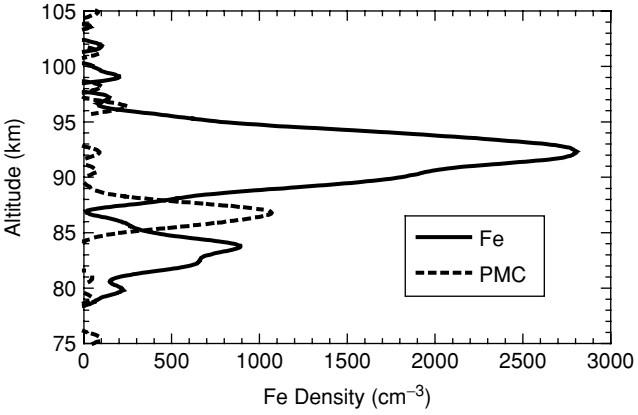


Figure 5.84 Simultaneous observations of the Fe density (black solid line) and PMC (gray dashed line) by the UIUC Fe Boltzmann temperature lidar at the South Pole on January 19, 2000. An Fe bite-out occurs at the altitude of PMC around 85 to 86 km.

efficient uptake of Fe atoms on the surface of the PMC ice particles. Combining the laboratory measurements of Fe uptake rate on the ice surface with the lidar observations of PMC at the South Pole, Plane et al. (2004) developed a one-dimensional mesospheric model to simulate the Fe layer under the influence of PMC ice particles. The model reproduces the Fe bite-out, higher peak height (~92 km), and overall depletion of the layer in the presence of the PMCs in January, and the higher Fe density and lower peak height (below 90 km) of the Fe layer in the absence of the PMCs during February. This is compelling evidence for heterogeneous removal of Fe on the ice particles causing these Fe depletion events.

5.4 CONCLUSIONS AND FUTURE OUTLOOK

Atmospheric lidar research and application is a field that integrates science, technology, and engineering. Lidar has proven to be a unique technology for studying the middle and upper atmosphere. Exciting advances in lidar design and development have been proposed and implemented in

recent years. Advances in resonance fluorescence lidar technology, including the development of the narrowband Doppler technique lidars (Na and K) and the Boltzmann technique lidar (Fe), enabled temperature and wind measurements in the middle and upper atmosphere as well as the atmospheric constituent and cloud particle detections. Over the past decade, these resonance fluorescence lidars along with Rayleigh lidars have made great contributions to the understanding of the global atmosphere thermal structure, dynamics, composition, and their relations to the global climate change. Exciting lidar measurements of the middle and upper atmosphere have been made from the geographic North Pole to the geographic South Pole. Compared with 20 years ago, our knowledge about the MLT thermal structures, the MLT dynamics, the polar mesospheric clouds (noctilucent clouds), the influences of gravity waves on the MLT region, the atmospheric metal layers, and mesospheric metal chemistry has been greatly enhanced by lidar observations in the middle and upper atmosphere.

These sophisticated Doppler lidars (Na and K) and Boltzmann lidars (Fe) have both advantages and disadvantages in their technologies, measurement capabilities, and operational complexities. The narrowband Na wind/temperature lidars based on the ring dye laser and pulsed dye amplifier have now reached technical maturity. They currently provide the highest resolution and most accurate temperature and wind measurements in the mesopause region. The dominant error source is the signal photon noise, which can be largely reduced by increasing the product of laser power and telescope aperture. Thus, coupling Na lidar with a large aperture telescope will significantly improve the temporal and spatial resolution, as demonstrated by the system developed by the University of Illinois at SOR and Maui. The full-diurnal-cycle Na wind/temperature lidar has been achieved at the Colorado State University by adding a Na FADOF narrowband filter in the lidar receiver. This system is making great contributions to the study of atmosphere dynamics. However, both the full diurnal cycle and large aperture Na wind/temperature lidars require sophisticated instrumentation, a well-controlled

laboratory working environment, and high-level operation/maintenance. These shortcomings preclude their application in aircrafts, on ships, or at remote sites.

The all-solid-state narrowband Na Doppler lidar based on the frequency mixing of two pulsed Nd:YAG lasers has proven to be rugged enough for deployment in Antarctica. However, it is currently only capable of temperature measurements, and unlikely to make wind measurements due to the lack of absolute frequency calibration. The solid-state Na wind/temperature lidar is a good alternative to the conventional Na wind/temperature lidar. It reduces the system complexity by replacing the ring dye laser with the frequency mixing of two cw Nd:YAG lasers while maintaining the measurement capability and accuracy. Further improvements in the stability of frequency mixing are needed and the possible replacement of pulsed dye amplifier (PDA) by optical parametric amplifier (OPA) is worthy of further investigation. Once the OPA successfully replaces the PDA, the Na wind/temperature lidar will be a full-solid-state system that would be more robust.

A possible solution to simplify the Na wind/temperature lidar is to develop interferometric receivers with broadband resonance fluorescence lidars. Traditional lidar receivers only collect and collimate the return signals onto a PMT, and then count the detected photons. The idea for interferometric receivers is to employ a high-resolution etalon in a lidar receiver. By measuring the Doppler broadening and Doppler frequency shift of the interference fringes, the atmospheric temperature and wind can be derived. This is basically the same as the interferometer principle for spectral imaging of the atmosphere (Shepherd, 2002). However, to obtain the range information of the returned photon signals, the CCD device needs to have the capability to rapidly switch between columns. These devices are being investigated and developed by Michigan Aerospace, but it may take years before they reach a usable state. Once this CCD device is practical for use, the lidar transmitters could be much simpler than their current narrowband system. Actually, a broadband lidar beam would be required to illuminate the Na Doppler broadened

line. Broadband lidar can also help avoid the Na layer saturation problem, which is experienced by the narrowband Na Doppler lidars. The price would be the much weaker return signals compared to narrowband systems. The high quantum efficiency of CCD devices (~90%) could partially compensate for the weak signals. Such broadband lidar systems may experience significant difficulties in daytime measurements, however.

The study of the middle and upper atmosphere requires more high-resolution lidar measurements at different locations all over the world. Thus, transportable lidar systems are desirable. The narrowband K Doppler lidar based on a solid-state ring alexandrite laser has been containerized, as demonstrated by the IAP, and is transportable for temperature measurements at different locations. The application of K FADOF makes the K lidar capable of temperature measurements during both day and night. Although it is possible to make wind measurements, the low K densities in the mesosphere result in relatively poor resolution and accuracy for wind measurements. Rayleigh Doppler lidar is being proposed for making wind measurements. It is either based on an iodine filter edge technique or on a Fabry–Perot etalon for measuring the Doppler frequency shift. Since the Rayleigh signals decrease exponentially as the altitude increases, a pulsed laser with high energy and short wavelength would be required to perform good wind measurements above 50 km.

The Fe Boltzmann temperature lidar based on the solid-state alexandrite lasers has proven to be robust enough to be deployed at remote sites and in aircraft. The lidar measurements of the MLT thermal structure, polar mesospheric clouds, and Fe layers from the North Pole to the South Pole have been accomplished by the University of Illinois with this type of lidar system. Because of its shorter wavelength, higher output power, and daytime capability, the Fe Boltzmann temperature lidar is very suitable for PMC/NLC detection in the polar region and can also provide Rayleigh temperature measurements over a wide range of altitudes. However, the biggest drawback to the Fe Boltzmann technique is the signal level on the weak 374-nm channel, which is about 30 times weaker

than the 372-nm channel at 200 K. This weak signal level limits the resolution of Fe Boltzmann lidar temperature measurements. Although the daytime measurement of MLT temperature is feasible, it requires long integration times.

To overcome this problem and maintain the large altitude coverage for temperature measurements as well as achieving wind measurements, a narrowband Fe Doppler lidar operating only at 372 nm would be attractive. This lidar combines the solid-state alexandrite ring laser technologies with existing acousto-optic modulation techniques to enable temperature and wind measurements in the mesopause region using Fe resonance fluorescence. Although the Fe atomic backscatter cross-section is about an order of magnitude smaller than that of Na, the Fe density is at least two times higher than the Na density. Since alexandrite ring laser pulses have long duration time (>200 ns) and saturation is not a problem for the Fe layer, the Fe lidar power could be many times higher than the Na lidar. Thus, the overall product of the backscatter cross-section, constituent density and laser power for the narrowband Fe Doppler lidar can be comparable to that of a Na wind/temperature lidar. Since the Fe lidar operates at UV wavelength and has high power, it will still have good altitude coverage for Rayleigh temperature measurements. Therefore, the narrowband Fe Doppler lidar is expected to have virtually continuous temperature coverage from 30 to 110 km. It will also be able to measure wind in the mesopause region. Such measurements will be available for both daytime and nighttime with Fabry–Perot etalons employed in the lidar receiver. Such an Fe Doppler lidar will be robust enough to be deployed in aircraft and at remote sites, and to be transportable for measurements of temperature and wind all over the globe.

The current and proposed resonance fluorescence lidars are all based upon the mesospheric metal layers Na, Fe, and K. These layers exist from approximately 75 to 120 km. Out of this range, the neutral Na, Fe, and K number densities are too low to make reliable temperature or wind measurements. Nevertheless, it is desirable to measure the temperature and species above 120 km for upper atmosphere studies.

Therefore, new atoms or molecules with appropriate number density distribution, and suitable energy level diagram for available laser wavelengths, need to be identified for developing new lidars. Several ideas have been proposed. One is the N_2^+ molecule when aurora occurs in the upper atmosphere (Garner and Dao, 1995; Collins et al., 1997). Another is atomic He in the thermosphere (Kerr et al., 1996; Gerrard et al., 1997). These molecules or atoms have a significant number density above 150 km, and it would be possible to detect them using resonance fluorescence lidar techniques. Atomic oxygen O and hydrogen H are also very important species in the upper atmosphere, especially for atmospheric chemistry. However, neither atomic O nor H has been measured with high range resolution. The importance of these constituents is the motivation for the investigation of far-ultraviolet (FUV) lidars. The lines at 121.6 nm for H and 130.5 nm for O are two potential candidates for measuring the densities with resonance fluorescence lidar techniques. The challenge is to build the FUV laser. Since the lower atmosphere strongly absorbs FUV light, such lidars would need to be space-borne. Other neutral and ionic molecules in the middle and upper atmosphere, such as N, NO, NO_2, CO, CO_2, O_2, N_2, N_2^+, O^+, He^+, are also new candidates for lidar detection.

Although lidars can provide high-resolution data, a lidar only probes a single point in the atmosphere. It would be ideal to have a wider detection range as well as high resolution. Correlated observations from ground-based clusters of collocated lidars, imagers, radars, and other remote sensing equipment provide unique vertical and horizontal measurement capabilities that can aid in resolving the outstanding fundamental issues in atmospheric science.

Resonance fluorescence lidar techniques are now an established measurement technology for high-resolution and high-accuracy studies of the middle and upper atmosphere. It is clear that given the important contributions these instruments have made to atmosphere science in their brief lifetime, many more new and significant results will be achieved by existing and new proposed lidars.

ACKNOWLEDGMENTS

The chapter of Resonance Fluorescence Lidar for Measurements of the Middle and Upper Atmosphere was finished in October 2003. Authors Chu and Papen have drawn on the experience of several years work at the Department of Electrical & Computer Engineering in the University of Illinois at Urbana-Champaign, and are grateful to all colleagues in the research team during this period, especially to Dr Chester S. Gardner for his leadership and encouragement. The authors are also indebted to Dr Chiao-Yao She, Dr Ulf von Zahn, Dr Franz-Josef Lübken, Dr Josef Höffner, Dr Jonathan Friedman, and Dr Takuya Kawahara, who have contributed material to this chapter. The authors gratefully acknowledge Dr Steve J. Franke, Dr Patricia M. Franke, Dr Erhan Kudeki, Dr Xing Gao, and Jing Tang for their fruitful discussion during the preparation of this manuscript and their proofreading of the manuscript.

REFERENCES

Alpers, M., J. Höffner, and U. von Zahn, Iron atom densities in the polar mesosphere from lidar observations, *Geophys. Res. Lett., 17*, 2345–2348, 1990.

Alpers, M., J. Höffner, and U. von Zahn, Sporadic Fe and E layers at polar, middle and low latitudes, *J. Geophys. Res. 99 (A8),* 14971–14985, 1994.

Alpers, M., J. Höffner, and U. von Zahn, Upper atmosphere Ca and Ca^+ at mid-latitudes: first simultaneous and common-volume lidar observations, *Geophys. Res. Lett., 23,* 567–570, 1996.

Alpers, M., M. Gerding, J. Höffner, and U. von Zahn, NLC particle properties from a five-color lidar observation at 54°N, *J. Geophys. Res., 105,* 12235–12240, 2000.

Alpers, M., M. Gerding, J. Höffner, and J. Schneider, Multiwavelength lidar observation of a strange noctilucent cloud at Kühlungsborn, Germany (54°N), *J. Geophys. Res., 106,* 7945–7953, 2001.

Andrews, D. G., J. R. Holton, and C. B. Leovy, *Middle Atmosphere Dynamics*, Academic, San Diego, CA, 1987.

Arimondo, E., M. Inguscio, and P. Violino, Experimental determinations of the hyperfine structure in the alkali atoms, *Rev. Mod. Phys., 49*, 31–75, 1977.

Arnold, K. S., and C. Y. She, Metal fluorescence lidar (light detection and ranging) and the middle atmosphere, *Contemp. Phys., 44*, 35–49, 2003.

Avaste, O., Noctilucent clouds, *J. Atmos. Terr. Phys., 55*, 133–143, 1993.

Backhouse, T. W., The luminous cirrus clouds of June and July, *Meteorol. Mag., 20*, 133, 1885.

Batista, P. P., B. R. Clemesha, D. M. Simonich, and V. W. J. H. Kirchhoff, Tidal oscillations in the atmospheric sodium layer, *J. Geophys. Res., 90,* 3881–3888, 1985.

Baumgarten, G., K. H. Fricke, G. von Cossart, Investigation of the shape of noctilucent cloud particles by polarization lidar technique, *Geophys. Res. Lett., 29 (13),* doi: 10.1029/2001GL013877, 2002a.

Baumgarten, G., F. J. Lübken, and K. H. Fricke, First observation of one noctilucent cloud by a twin lidar in two different directions, *Ann. Geophys., 20,* 1863–1868, 2002b.

Beatty, T. J., R. E. Bills, K. H. Kwon, and C. S. Gardner, CEDAR lidar observations of sporadic Na layers at Urbana, Illinois, *Geophys. Res. Lett., 15,* 1137–1140, 1988.

Berger, U., and U. von Zahn, The two-level structure of the mesopause: a model study, *J. Geophys. Res., 104 (D18)*, 22083–22093, 1999.

Berger, U., and U. von Zahn, Icy particles in the summer mesopause region: three-dimensional modeling of their environment and two-dimensional modeling of their transport, *J. Geophys. Res., 107* (A11), 1366, doi: 10.1029/2001JA000316, 2002.

Bills, R. E., and C. S. Gardner, Lidar observations of mesospheric Fe and sporadic Fe layers at Urbana, Illinois, *Geophys. Res. Lett., 17*, 143–146, 1990.

Bills, R. E., C. S. Gardner, and C. Y. She, Narrowband lidar technique for sodium temperature and Doppler wind observations of the upper atmosphere, *Opt. Eng., 30*, 13–21, 1991a.

Bills, R. E., C. S. Gardner, and S. J. Franke, Na Doppler/temperature lidar: initial mesopause region observations and comparison with the Urbana medium frequency radar, *J. Geophys. Res., 96 (D12)*, 22701–22707, 1991b.

Bills, R. E., and C. S. Gardner, Lidar observations of the mesopause region temperature structure at Urbana, *J. Geophys. Res., 98 (D1),* 1011–1021, 1993.

Blamont, J. E., M. L. Chanin, and G. Megie, Vertical distribution and temperature of the nighttime atmospheric sodium layer obtained by laser backscatter, *Ann. Geophys. 28*, 833–838, 1972.

Bowman, M. R., A. J. Gibson, and M. C. W. Sandford, Atmospheric sodium measured by a tuned laser radar, *Nature, 221*, 456–457, 1969.

Carbary, J. F., G. J. Romick, D. Morrison, L. J. Paxton, C. I. Meng, Altitudes of polar mesospheric clouds observed by a middle ultraviolet imager, *J. Geophys. Res., 104*, 10089–10100, 1999.

Cerny, T., and C. F. Sechrist, Jr., Calibration of the Urbana Lidar System, Aeronomy Report No. 94, University of Illinois, August 1, 1980.

Chapman, S., and P. C. Kendall, Noctilucent clouds and thermospheric dust: their diffusion and height distribution, *Q. J. R. Metorol. Soc., 91*, 115–131, 1965.

Charlson, R. J., Noctilucent clouds: a steady-state model, *Q. J. R. Metorol. Soc., 91*, 517–523, 1965.

Charlson, R. J., A simple noctilucent cloud model, *Tellus, 18*, 451, 1966.

Chen, H., C. Y. She, P. Searcy, and E. Korevaar, Sodium-vapor dispersive Faraday filter, *Opt. Lett., 18*, 1019–1021, 1993.

Chen, H., M. A. White, D. A. Krueger, and C. Y. She, Daytime mesopause temperature measurements using a sodium-vapor dispersive Faraday filter in lidar receiver, *Opt. Lett., 21,* 1093–1095, 1996.

Chen, S., Z. Hu, M. A. White, H. Chen, D. A. Krueger, and C. Y. She, Lidar observations of seasonal variation of diurnal mean temperature in the mesopause region over Fort Collins, Colorado (41°N, 105°W), *J. Geophys. Res., 105 (D10),* 12371–12379, 2000.

Chiu, P. H., A. Magana, and J. Davis, All-solid-state single-mode sum-frequency generation of sodium resonance radiation, *Opt. Lett., 19*, 2116–2118, 1994.

Chu, X., W. Pan, G. Papen, C. S. Gardner, G. Swenson, and P. Jenniskens, Characteristics of Fe ablation trails observed during the 1998 Leonid meteor shower, *Geophys. Res. Lett., 27,* 1807–1810, 2000a.

Chu, X., A. Z. Liu, G. Papen, C. S. Gardner, M. Kelley, J. Drummond, and R. Fugate, Lidar observations of elevated temperatures in bright chemiluminescent meteor trails during the 1998 Leonid shower, *Geophys. Res. Lett., 27*, 1815–1818, 2000b.

Chu, X., C. S. Gardner, and G. Papen, Lidar observations of polar mesospheric clouds at South Pole: seasonal variations, *Geophys. Res. Lett., 28*, 1203–1206, 2001a.

Chu, X., C. S. Gardner, and G. Papen, Lidar observations of polar mesospheric clouds at South Pole: diurnal variations, *Geophys. Res. Lett., 28,* 1937–1940, 2001b.

Chu, X., W. Pan, G. C. Papen, C. S. Gardner, and J. A. Gelbwachs, Fe Boltzmann temperature lidar: design, error analysis, and initial results at the North and South Poles, *Appl. Opt., 41,* 4400–4410, 2002a.

Chu, X., and C. S. Gardner, Pole-to-Pole: lidar observations of middle and upper atmosphere temperature and polar mesospheric clouds over the North and South Poles, *Proc. SPIE, 4893*, 223–236, 2002b.

Chu, X., C. S. Gardner, and R. G. Roble, Lidar studies of interannual, seasonal, and diurnal variations of polar mesospheric clouds at the South Pole, *J. Geophys. Res., 108 (D8),* 8447, doi: 10.1029/2002JD002524, 2003.

Chu, X., G. J. Nott, P. J. Espy, C. S. Gardner, J. C. Diettrich, M. A. Clilverd, and M. J. Jarvis, Lidar observations of polar mesospheric clouds at Rothera, Antarctica (67.5°S, 68.0°W), *Geophys. Res. Lett*., *31*, L02114, doi: 10.1029/2003GL018638, 2004.

Clemesha, B. R., D. M. Simonich, P. P. Batista, and V. W. J. H. Kirchhoff, The diurnal variation of atmospheric sodium, *J. Geophys. Res., 87*, 181–185, 1982.

Clemesha, B. R., M. P. P. Martins Jorge, D. M. Simonich, and P. P. Batista, A new method for measuring the Doppler temperature of the atmospheric sodium layer, *Adv. Space Res., 19*, 681–684, 1997.

Clemesha, B. R., I. Veselovskii, P. P. Batista, M. P. P. M. Jorge, and D. M. Simonich, First mesopause temperature profiles from a fixed southern hemisphere site, *Geophys. Res. Lett., 26,* 1681–1684, 1999.

Clemesha, B. R., P. P. Batista, and D. M. Simonich, Tide-induced oscillations in the atmospheric sodium layer, *J. Atmos. Sol. Terr. Phys., 64*, 1321–1325, 2002.

Collins, R. L., D. Lummerzhein, and R. W. Smith, Analysis of lidar systems for profiling aurorally excited molecular species, *Appl. Opt., 36,* 6024–6034, 1997.

Collins, R. L., M. C. Kelley, M. J. Nicolls, C. Ramos, T. Hou, T. E. Stern, K. Mizutani, and T. Itabe, Simultaneous lidar observations of a noctilucent cloud and an internal wave in the polar mesosphere, *J. Geophys. Res., 108 (D8)*, 8435, doi: 10.1029/2002JD002427, 2003.

Debrestian, D. J., et al., An analysis of POAMII solar occultation observations of polar mesospheric clouds in the southern hemisphere, *J. Geophys. Res., 102*, 1971–1981, 1997.

Dick, D. J., and T. M. Shay, Ultrahigh-noise rejection optical filter, *Opt. Lett., 16,* 867–869, 1991.

Donahue, T. M., B. Guenther, and J. E. Blamont, Noctilucent clouds in daytime: circumpolar particulate layers near the summer mesopause, *J. Atmos. Sci., 29,* 1205–1209, 1972.

Drummond, J. D., B. W. Grime, C. S. Gardner, A. Z. Liu, X. Chu, and T. J. Kane, Observations of persistent Leonid meteor trails 1. Advection of the "Diamond Ring", *J. Geophys. Res., 106* (A10), 21517–21524, 2001.

Elterman, L. B., The measurement of stratospheric density distribution with the search light technique, *J. Geophys. Res., 56*, 509–520, 1951.

Elterman, L., A series of stratospheric temperature profiles obtained with the searchlight technique, *J. Geophys. Res., 58*, 519–530, 1953.

Elterman, L., Seasonal trends of temperature, density, and pressure to 67.6 km obtained with the searchlight probing technique, *J. Geophys. Res., 59*, 351–358, 1954.

Eska, V., and J. Höffner, Observed linear and nonlinear K layer response, *Geophys. Res. Lett., 25*, 2933–2936, 1998.

Eska, V., J. Höffner, and U. von Zahn, Upper atmosphere potassium layer and its seasonal variations at 54°N, *J. Geophys. Res., 103 (A12),* 29207–29214, 1998.

Eska, V., U. von Zahn, and J. M. C. Plane, The terrestrial potassium layer (75–110 km) between 71°S and 54°N: observations and modeling, *J. Geophys. Res., 104 (A8),* 17173–17186, 1999.

Evans, W. F., L. R. Laframboise, K. R. Sine, R. H. Wiens, and G. G. Shepherd, Observation of polar mesospheric clouds in summer, 1993 by the WINDII instrument on UARS, *Geophys. Res. Lett., 22,* 2793–2796, 1995.

Eyler, E. E., A. Yiannopoulou, S. Gangopadhyay, and N. Melikechi, Chirp-free nanosecond laser amplifier for precision spectroscopy, *Opt. Lett., 22*, 49–51, 1997.

Fee, M. S., K. Danzmann, and S. Chu, Optical heterodyne measurement of pulsed lasers: toward high-precision pulsed spectroscopy, *Phys. Rev. A, 45*, 4911–4924, 1992.

Felix, F., W. Keenliside, G. Kent, and M. C. W. Sandford, Laser radar observations of atmospheric potassium, *Nature, 246*, 345–346, 1973.

Fiedler, J., G. Baumgarten, and G. von Cossart, Noctilucent clouds above ALOMAR between 1997 and 2001: occurrence and properties, *J. Geophys. Res., 108 (D8)*, 8453, doi: 10.1029/2002JD002419, 2003.

Fleming, E. L., S. Chandra, J. J. Barnett, and M. Corney, Zonal mean temperature, pressure, zonal wind and geopotential height as functions of latitude, *Adv. Space Res., 10(12),* 11–59, 1990.

Fricke, K. H., and U. von Zahn, Mesopause temperatures derived from probing the hyperfine structure of the D_2 resonance line of sodium by lidar, *J. Atmos. Terr. Phys., 47*, 499–512, 1985.

Fricke-Begemann, C., J. Höffner, and U. von Zahn, The potassium density and temperature structure in the mesopause region (80–105 km) at a low latitude (28°N), *Geophys. Res. Lett., 29 (22),* doi: 10.1029/2002GL015578, 2002.

Friedman, J. S., S. C. Collins, R. Delgado, and P. A. Castleberg, Mesospheric potassium layer over the Arecibo Observatory, 18.3°N 66.75°W, *Geophys. Res. Lett., 29 (5),* doi: 10.1029/2001GL013542, 2002.

Friedman, J. S., Tropical mesopause climatology over the Arecibo Observatory, *Geophys. Res. Lett., 30 (12),* doi: 10.1029/2003GL016966, 2003a.

Friedman, J. S., C. A. Tepley, S. Raizada, Q. H. Zhou, J. Hedin, and R. Delgado, Potassium Doppler-resonance lidar for the study of the mesosphere and lower thermosphere at the Arecibo Observatory, *J. Atoms. Sol.-Terr. Phys., 65*, 1411–1424, 2003b.

Fritts, D. C., and R. A. Vincent, Mesospheric momentum flux studies at Adelaide, Australia: observations and a gravity wave-tidal interaction model, *J. Atmos. Sci., 44*, 605–619, 1987.

Fritts, D. C., and T. van Zandt, Spectral estimates of gravity wave energy and momentum fluxes: energy dissipation, acceleration, and constraints, *J. Atmos. Sci., 50,* 3685–3694, 1993.

Fritts, D. C., and Z. Luo, Dynamical and radiative forcing of the summer mesopause circulation and thermal structure 1. Mean solstice conditions, *J. Geophys. Res., 100 (D2)*, 3119–3128, 1995.

Gadsden, M., The silver-blue cloudlets again: nucleation and growth of ice in the mesosphere, *Planet. Space Sci., 29*, 1079–1087, 1981.

Gadsden, M., and W. Schröder, *Noctilucent Clouds*, Springer-Verlag, Berlin, 1989.

Gadsden, M., A secular change in noctilucent cloud occurrence, *J. Atmos. Terr. Phys., 52*, 247–251, 1990.

Gadsden, M., Polar mesospheric clouds seen from geostationary orbit, *J. Atmos. Sol. Terr. Phys., 62*, 31–36, 2000.

Gangopadhyay, S., N. Melikechi, and E. E. Eyler, Optical phase perturbations in nanosecond pulsed amplification and second-harmonic generation, *J. Opt. Soc. Am. B, 11*, 231–241, 1994.

Garcia, R. R., and S. Solomon, The effect of breaking gravity waves on the dynamics and chemical composition of the mesosphere and lower thermosphere, *J. Geophys. Res., 90*, 3850–3868, 1985.

Garcia, R.R., Dynamics, radiation and photochemistry in the mesosphere — implications for the formation of noctilucent clouds, *J. Geophys. Res., 94*, 14605–14615, 1989.

Gardner, C. S., and J. D. Voelz, Lidar measurements of gravity wave saturation effects in the sodium layer, *Geophys. Res. Lett., 12*, 765–768, 1985.

Gardner, C. S., Sodium resonance fluorescence lidar applications in atmospheric science and astronomy, *Proc. IEEE, 77*, 408–418, 1989.

Gardner, C. S., T. J. Kane, D. C. Senft, J. Qian, and G. C. Papen, Simultaneous observations of sporadic E, Na, Fe, and Ca^+ layers at Urbana, Illinois: three case studies, *J. Geophys. Res., 98*, 16865–16873, 1993.

Gardner, C. S., Diffusive filtering theory of gravity wave spectra in the atmosphere, *J. Geophys. Res., 99*, 20601–20622, 1994.

Gardner, C. S., and G. C. Papen, Mesospheric Na wind/temperature lidar, *Rev. Laser Eng., 23*, 131–134, 1995.

Gardner, C. S., Introduction to ALOHA/ANLC-93: the 1993 airborne lidar and observations of the Hawaiian airglow/airborne noctilucent cloud campaigns, *Geophys. Res. Lett., 22*, 2789–2792, 1995.

Gardner, C. S., 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, *J. Geophys. Res., 103 (D14)*, 16909–16927, 1998.

Gardner, C. S., K. Gulati, Y. Zhao, and G. Swenson, Measuring gravity wave momentum fluxes with airglow imagers, *J. Geophys. Res., 104 (D10)*, 11903–11915, 1999.

Gardner, C. S., G. C. Papen, X. Chu, and W. Pan, First lidar observations of middle atmosphere temperatures, Fe densities, and polar mesospheric clouds over the North and South Poles, *Geophys. Res. Lett., 28*, 1199–1202, 2001.

Gardner, C. S., Y. Zhao, and A. Z. Liu, Atmospheric stability and gravity wave dissipation in the mesopause region, *J. Atmos. Sol. Terr. Phys., 64*, 923–929, 2002.

Garner, R. C., and P. Dao, Molecular nitrogen fluorescence lidar for remote sensing of the auroral ionosphere, *J. Geophys. Res., 100*, 14131–14140, 1995.

Gelbwachs, J. A., Iron Boltzmann factor lidar: proposed new remote-sensing technique for mesospheric temperature, *Appl. Opt., 33*, 7151–7156, 1994.

Gerding, M., M. Alpers, J. Höffner, and U. von Zahn, Simultaneous K and Ca lidar observations during a meteor shower on March 6–7, 1997 at Kühlungsborn, Germany, *J. Geophys. Res., 104 (A11)*, 24689–24698, 1999.

Gerding, M., M. Alpers, U. von Zahn, R. J. Rollason, and J. M. C. Plane, Atmosphere Ca and Ca^+ layers: midlatitude observations and modeling, *J. Geophys. Res., 105 (A12)*, 27131–27146, 2000.

Gerrard, A. J., T. J. Kane, D. D. Meisel, J. P. Thayer, and R. B. Kerr, Investigation of a resonance lidar for measurement of thermospheric metastable helium, *J. Atmos. Sol. Terr. Phys., 59*, 2023–2035, 1997.

Gerrard, A. J., T. J. Kane, and J. P. Thayer, Noctilucent clouds and wave dynamics: observations at Sondrestrom, Greenland, *Geophys. Res. Lett., 25*, 2817–2820, 1998.

Gibson, A. J., and M. C. W. Sandford, The seasonal variation of the night-time sodium layer, *J. Atmos. Terr. Phys., 33*, 1675–1684, 1971.

Gibson, A. J., and M. C. W. Sandford, Daytime laser radar measurements of the atmospheric sodium layer, *Nature, 239*, 509–511, 1972.

Gibson, A. J., L. Thomas, and S. K. Bhattachacharyya, Laser observations of the ground-state hyperfine structure of sodium and of temperatures in the upper atmosphere, *Nature, 281*, 131–132, 1979.

Goldberg, R. A., et al., DROPPS: a study of the polar summer mesosphere with rocket, radar and lidar, *Geophys. Res. Lett., 28*, 1407–1410, 2001.

Gong, S. S., G. T. Yang, J. M. Wang, B. M. Liu, X. W. Cheng, J. Y. Xu, and W. X. Wan, Occurrence and characteristics of sporadic sodium layer observed by lidar at a mid-latitude location, *J. Atmos. Sol. Terr. Phys., 64*, 1957–1966, 2002.

Goody, R. M., and Y. L. Yung, *Atmospheric Radiation: Theoretical Basis*, 2nd edition, Oxford University Press, Oxford, 1989.

Grahn, S., and G. Witt, Photogrammetric Triangulation of Diffuse Objects in Space: Part II, Report AP-4, Institute of Meteorology, University of Stockholm, April 1971.

Granier, C., and G. Megie, Daytime lidar measurements of the mesospheric sodium layer, *Planet. Space Sci., 30*, 169–177, 1982.

Granier, C., J. P. Jegou, and G. Megie, Resonant lidar detection of Ca and Ca^+ in the upper atmosphere, *Geophys. Res. Lett., 12*, 655–658, 1985.

Granier, C., J. P. Jegou, and G. Megie, Atomic and Ionic Calcium in the Earth's upper atmosphere, *J. Geophys. Res., 94 (D7)*, 9917–9924, 1989a.

Granier, C., J. P. Jegou, and G. Megie, Iron atoms and metallic species in the Earth's upper atmosphere, *Geophys. Res. Lett., 16*, 243–246, 1989b.

Grime, B. W., T. J. Kane, S. C. Collins, M. C. Kelley, C. A. Kruschwitz, J. S. Friedman, and C. A. Tepley, Meteor trail advection and dispersion: preliminary lidar observations, *Geophys. Res. Lett., 26*, 675–678, 1999.

Grime, B. W., T. J. Kane, A. Liu, G. Papen, C. S. Gardner, M. C. Kelley, C. Kruschwitz, and J. Drummond, Meteor trail advection observed during the 1998 Leonid shower, *Geophys. Res. Lett., 27*, 1819–1822, 2000.

Gumbel, J., and G. Witt, *In situ* measurements of the vertical structure of a noctilucent cloud, *Geophys. Res. Lett., 25*, 493–496, 1998.

Hagan, M. E., J. M. Forbes, and F. Vial, On modeling migrating solar tides, *Geophys. Res. Lett., 22*, 893–896, 1995.

Hake, R. D. Jr., D. E. Arnold, D. W. Jackson, W. E. Evans, B. P. Ficklin, and R. A. Long, Dye-laser observations of the nighttime atomic sodium layer, *J. Geophys. Res., 77*, 6839–6848, 1972.

Hansch, T. W., I. S. Shahin, and A. L. Schawlow, High-resolution saturation spectroscopy of the sodium D lines with a pulsed tunable dye laser, *Phys. Rev. Lett., 27*, 707–710, 1971.

Hansen, G., M. Serwazi, and U. von Zahn, First detection of a noctilucent cloud by lidar, *Geophys. Res. Lett., 16*, 1445–1448, 1989.

Hansen, G., and U. von Zahn, Sudden sodium layers in polar latitudes, *J. Atmos. Terr. Phys., 52*, 585–608, 1990.

Hansen, G., and U. von Zahn, Simultaneous observations of noctilucent clouds and mesopause temperatures by lidar, *J. Geophys. Res., 99*, 18989–18999, 1994.

Hauchecorne, A., and M. L. Chanin, Density and temperature profiles obtained by lidar between 35 and 70 km, *Geophys. Res. Lett., 7*, 565–568, 1980.

Hauchecorne, A., M. L. Chanin, and R. Wilson, Mesospheric temperature inversion and gravity wave dynamics, *Geophys. Res. Lett., 14*, 935–939, 1987.

Hecht, J. H., J. P. Thayer, D. J. Gutierrez, D. L. McKenzie, Multi-instrument zenith observations of noctilucent clouds over Greenland on July 30/31, 1995, *J. Geophys. Res., 102*, 1959–1970, 1997.

Helmer, M., J. M. C. Plane, J. Qian and C. S. Gardner, A model of meteoric iron in the upper atmosphere, *J. Geophys. Res., 103 (D9)*, 10913–10925, 1998.

Hervig, M., R. E. Thompson, M. NcHugh, L. L. Gordley, J. M. Russell III, and M. E. Summers, First confirmation that water ice is the primary component of polar mesospheric clouds, *Geophys. Res. Lett., 28*, 971–974, 2001.

Hesstvedt, E., Note on the nature of noctilucent clouds, *J. Geophys. Res., 99*, 1985–1987, 1961.

Hesstvedt, E., On the possibility of ice cloud formation at the mesopause, *Tellus, 14*, 290–296, 1962.

Höffner, J., and U. von Zahn, Mesopause temperature profiling by potassium lidar: recent progress and outlook for ALOMAR, in *Proc. 12th ESA Symp. on European Rocket and Balloon Programmes and Related Research*, edited by B. Kaldeich-Schürmann, ESA SP-370, 403–407, 1995.

Höffner, J., U. von Zahn, W. J. McNeil, and E. Murad, The 1996 Leonid shower as studied with a potassium lidar: observations and inferred meteoroid sizes, *J. Geophys. Res., 104 (A2)*, 2633–2643, 1999.

Höffner, J., C. Fricke-Begemann, and F. J. Lübken, First observations of noctilucent clouds by lidar at Svalbard, *Atmos. Chem. Phys. Discuss., 3*, 521–549, 2003.

Hu, X., A. Z. Liu, C. S. Gardner, and G. R. Swenson, Characteristics of quasi-monochromatic gravity waves observed with Na lidar in the mesopause region at Starfire Optical Range, NM, *Geophys. Res. Lett., 29 (24)*, 2169, doi: 10.1029/2002GL014975, 2002.

Humphreys, W. J., Nacreous and noctilucent clouds, *Mon. Weather Rev., 61*, 228–229, 1933.

Jegou, J.-P., M.-L. Chanin, G. Megie, and J. E. Blamont, Lidar measurements of atmospheric lithium, *Geophys. Res. Lett., 7*, 995–998, 1980.

Jensen, E. J., and G. E. Thomas, A growth-sedimentation model of polar mesospheric clouds: comparisons with SME measurements, *J. Geophys. Res., 93*, 2461–2473, 1988.

Jensen, E. J., G. E. Thomas, and O. B. Toon, On the diurnal variation of noctilucent clouds, *J. Geophys. Res., 94*, 14693–14702, 1989.

Jensen, E. J., and G. E. Thomas, Numerical simulations of the effects of gravity waves on noctilucent clouds, *J. Geophys. Res., 99*, 3421–3430, 1994.

Jesse, O., Die Höhe der leuchtenden Nachtwolken, *Astron. Nachrichten, 140*, 161–168, 1896.

Jeys, T. H., A. A. Brailove, and A. Mooradian, Sum frequency generation of sodium resonance radiation, *Appl. Opt., 28*, 2588–2591, 1989.

Jones, K. M., P. S. Julienne, P. D. Lett, W. D. Phillips, E. Tiesinga, and C. J. Williams, Measurement of the atomic Na (3P) lifetime and of retardation in the interaction between two atoms bound in a molecule, *Europhys. Lett., 35*, 85–90, 1996.

Kane, T. J., and C. S. Gardner, Structure and seasonal variability of the nighttime mesospheric Fe-layer at midlatitudes, *J. Geophys. Res., 98*, 16875–16886, 1993a.

Kane, T. J., and C. S. Gardner, Lidar observations of the meteoric deposition of mesospheric metals, *Science, 259*, 1297–1300, 1993b.

Kasevich, M. A., E. Riis, S. Chu, and R. G. Devoe, RF spectroscopy in an atomic fountain, *Phys. Rev. Lett., 63*, 612–615, 1989.

Kawahara, T. D., T. Kitahara, F. Kobayashi, Y. Saito, A. Nomura, C. Y. She, D. A. Krueger, M. Tsutsumi, Wintertime mesopause temperatures observed by lidar measurements over Syawa station (69°S, 39°E), Antarctica, *Geophys. Res. Lett., 29*, doi: 10.1029/2002GL015244, 2002.

Kawahara, T. D., C. S. Gardner, and A. Nomura, Observed temperature structure of the atmosphere above Syowa Station, Antarctica (69°S, 39°E), *J. Geophys. Res.*, *109*, D12103, doi: 10.1029/2003JD003918, 2004.

Kelley, M. C., C. S. Gardner, J. Drummond, T. Armstrong, A. Liu, X. Chu, G. Papen, C. Kruschwitz, P. Loughmiller, B. Grime, and J. Engelman, First observations of long-lived meteor trains with resonance lidar and other optical instruments, *Geophys. Res. Lett., 27*, 1811–1814, 2000.

Kerr, R., J. Noto, R. S. Lancaster, M. Franco, R. J. Rudy, R. Williams, and J. H. Hecht, Fabry-Perot observations of helium 10830 Å emission at Millstone Hill, *Geophys. Res. Lett., 23*, 3239–3242, 1996.

Kirchhoff, V. W. J. H., and B. R. Clemesha, Atmospheric sodium measurements at 23°S, *J. Atmos. Terr. Phys., 35*, 1493–1498, 1973.

Klekociuk, A., M. M. Lambert, and R. A. Vincent, Characteristics of a polar mesospheric clouds observed by Rayleigh lidar at Davis, Antarctica, paper presented at IUGG Conference, Sapporo, Japan, 30 June–11 July 2003.

Klostermeyer, J., A simple model of the ice particle size distribution in noctilucent clouds, *J. Geophys. Res., 103*, 28743–28752, 1998.

Klostermeyer, J., Effect of tidal variability on the mean diurnal variation of noctilucent clouds, *J. Geophys. Res., 106*, 9749–9755, 2001.

Kudeki, E., and S. J. Franke, Statistics of momentum flux estimation, *J. Atmos. Sol. Terr. Phys., 60*, 1549–1553, 1998.

Kurzawa, H., and U. von Zahn, Sodium density and atmospheric temperature in the mesopause region in polar summer, *J Atmos. Terr. Phys., 52*, 981–993, 1990.

Kwon, K. H., C. S. Gardner, D. C. Senft, F. L. Roesler, and J. Harlander, Daytime lidar measurements of tidal winds in the mesospheric sodium layer at Urbana, Illinois, *J. Geophys. Res., 92*, 8781–8786, 1987.

Langer, M., K. P. Müller, and K. H. Fricke, Rayleigh lidar detection of aerosol echoes from noctilucent cloud altitudes at the Arctic Circle, *Geophys. Res. Lett., 22*, 381–384, 1995.

Leblanc, T., I. S. McDermid, P. Keckhut, A. Hauchecorne, C. Y. She, and D. A. Krueger, Temperature climatology of the middle atmosphere from long-term lidar measurements at middle and low latitudes, *J. Geophys. Res., 103*, 17191–17204, 1998.

Lindzen, R. S., Turbulence and stress owing to gravity-wave and tidal breakdown, *J. Geophys. Res., 86*, 9707–9714, 1981.

Liu, A. Z., and C. S. Gardner, Vertical dynamical transport of mesospheric constituents by dissipating gravity waves, *J. Atmos. Sol. Terr. Phys.*, *66*, 267–275, 2004.

Loudon, R., *The Quantum Theory of Light*, 2nd edition, Oxford University Press, New York, 1983.

Lübken, F. J., and U. von Zahn, Thermal structure of the mesopause region at polar latitudes, *J. Geophys. Res., 96 (D11)*, 20841–20857, 1991.

Lübken, F. J., K. H. Fricke, and M. Langer, Noctilucent clouds and the thermal structure near the Arctic mesopause in summer, *J. Geophys. Res., 101,* 9489–9508, 1996.

Martin, W. C., and R. Zalubas, Energy levels of sodium Na i through Na xi, *J. Phys. Chem. Ref. Data*, 10, 153–196, 1981.

Measures, R. M., *Laser Remote Sensing — Fundamentals and Applications*, Wiley-Interscience, New Year, 1984.

Measures, R. M. (ed.), *Laser Remote Chemical Analysis*, Wiley-Interscience, New York, 1988.

Megie, G., and J. E. Blamont, Laser sounding of atmospheric sodium interpretation in terms of global atmospheric parameters, *Planet. Space Sci., 25*, 1093–1109, 1977.

Megie, G., F. Bos, J. E. Blamont, and M. L. Chanin, Simultaneous nighttime lidar measurements of atmospheric sodium and potassium, *Planet. Space Sci., 26*, 27–35, 1978.

Melikechi, N., S. Gangopadhyay, and E. E. Eyler, Phase dynamics in nanosecond pulsed dye laser amplification, *J. Opt. Soc. Am. B, 11*, 2402–2411, 1994.

Meriwether, J. W., and C. S. Gardner, A review of the mesosphere inversion layer phenomenon, *J. Geophys. Res., 105 (D10),* 12405–12416, 2000.

Milonni, P. W., R. Q. Fugate, and J. M. Telle, Analysis of measured photon returns from sodium beacons*, J. Opt. Soc. Am. A, 15*, 217–233, 1998.

Milonni, P. W., H. Fearn, J. M. Telle, and R. Q. Fugate, Theory of continuous-wave excitation of the sodium beacon*, J. Opt. Soc. Am. A, 16*, 2555–2566, 1999.

Mlynczak, M. G., and S. Solomon, Middle atmosphere heating by exothermic chemical reactions involving odd-hydrogen species, *Geophys. Res. Lett., 18,* 37–40, 1991.

Mlynczak, M. G., and S. Solomon, A detailed evaluation of the heating efficiency in the middle atmosphere, *J. Geophys. Res., 98 (D6),* 10517–10541, 1993.

Neuber, R., P. von der Gathen, and U. von Zahn, Altitude and temperature of the mesopause at 69°N latitude in winter, *J. Geophys. Res., 93*, 11093–11101, 1988.

Nussbaumer, V., K. H. Fricke, M. Langer, W. Singer, and U. von Zahn, First simultaneous and common volume observations of noctilucent clouds and polar mesosphere summer echoes by lidar and radar, *J. Geophys. Res., 101,* 19161–19167, 1996.

Oates, C. W., K. R. Vogel, and J. L. Hall, High precision linewidth measurement of laser-cooled atoms: resolution of the Na 3p $^2P_{3/2}$ lifetime discrepancy, *Phys. Rev. Lett., 76,* 2866–2869, 1996.

Pan, W., C. S. Gardner, and R. G. Roble, The temperature structure of the winter atmosphere at South Pole, *Geophys. Res. Lett., 29,* doi: 10.1029/2002GL015288, 2002.

Pan, W., and C. S. Gardner, Seasonal variations of the atmospheric temperature structure from 3 to 110 km at South Pole, *J. Geophys. Res., 108 (D18)*, 4564, doi: 10.1029/2002JD003217, 2003.

Papen, G. C., W. M. Pfenninger, and D. M. Simonich, Sensitivity analysis of Na narrowband wind–temperature lidar systems, *Appl. Opt., 34*, 480–498, 1995a.

Papen, G. C., C. S. Gardner, and W. M. Pfenninger, Analysis of a potassium lidar system for upper-atmospheric wind–temperature measurements, *Appl. Opt., 34*, 6950–6958, 1995b.

Papen, G. C., and D. Treyer, Comparison of an Fe Boltzmann temperature lidar with a Na narrow-band lidar, *Appl. Opt., 37*, 8477–8481, 1998.

Plane, J. M. C., The chemistry of meteoric metals in the Earth's upper atmosphere, *Int. Rev. Phys. Chem., 10*, 55–106, 1991.

Plane, J. M. C., R. M. Cox, J. Qian, W. M. Pfenninger, G. C. Papen, C. S. Gardner, P. J. Espy, Mesospheric Na layer at extreme high latitudes in summer, *J. Geophys. Res., 103 (D6),* 6381–6389, 1998.

Plane, J. M. C., C. S. Gardner, J. R. Yu, C. Y. She, R. R. Garcia, and H. C. Pumphrey, Mesospheric Na layer at 40°N: modeling and observations, *J. Geophys. Res., 104 (D3),* 3773–3788, 1999.

Plane, J. M. C., B. J. Murray, X. Chu, and C. S. Gardner, Removal of meteoric iron on polar mesospheric clouds, *Science*, *304*, 426–428, 2004.

Portman, R. W., G. E. Thomas, S. Solomon, and R. R. Garcia, The importance of dynamical feedbacks on doubled CO_2-induced changes in the thermal structure of the mesosphere, *Geophys. Res. Lett., 22*, 1733–1736, 1995.

Qian, J., and C. S. Gardner, Simultaneous lidar measurements of mesospheric Ca, Na, and temperature profiles at Urbana, Illinois, *J. Geophys. Res., 100 (D4),* 7453–7461, 1995.

Raizada, S., and C. Tepley, Iron Boltzmann lidar temperature and density observations from Arecibo — an initial comparison with other techniques, *Geophys. Res. Lett., 29 (12)*, doi: 10.1029/2001GL014535, 2002.

Raizada, S., and C. A. Tepley, Seasonal variation of mesospheric iron layers at Arecibo: first results from low-latitudes, *Geophys. Res. Lett., 30 (2),* doi: 10.1029/2002GL016537, 2003.

Rapp, M., F.-J. Lübken, A., Müllemann, G. E. Thomas, and E. J. Jensen, Small-scale temperature variations in the vicinity of NLC: experimental and model results, *J. Geophys. Res., 107 (D19)*, 4392, doi: 10.1029/2001JD001241, 2002.

Reid, G. C., Ice clouds at the summer polar mesopause, *J. Atmos. Sci., 32*, 523–535, 1975.

Reise, M., D. Offermann, and G. Brasseur, Energy released by recombination of atomic oxygen and related species at mesopause heights, *J. Geophys. Res., 99,* 14585–14594, 1994.

Roble, R. G., and R. E. Dickinson, How will changes in carbon dioxide and methane modify the mean structure of the mesosphere and thermosphere? *Geophys. Res. Lett.,* 16, 1441–1444, 1989.

Roble, R. G., Energetics of the mesosphere and thermosphere. The upper mesosphere and lower thermosphere: a review of experiment and theory, *Geophys. Monogr., 87*, 1–22, 1995.

Rodgers, C. D., F. W. Taylor, A. H. Muggeridge, M. Lopez-Puertas, M. A. Lopez-Valverde, Local thermodynamic equilibrium of CO_2 in the upper atmosphere, *Geophys. Res. Lett., 19,* 589–592, 1992.

Rollason, R. J., and J. M. C. Plane, The reactions of FeO with O_3, H_2, H_2O, O_2, and CO_2, *Phys. Chem., 2*, 2335–2343, 2000.

Sandford, M. C. W., and A. J. Gibson, Laser radar measurements of the atmospheric sodium layer, *J. Atmos. Terr. Phys., 32*, 1423–1430, 1970.

Schmidlin, F. J., Temperature inversion near 75 km, *Geophys. Res. Lett., 3,* 173–176, 1976.

Schmitz, S., U. von Zahn, J. C. Walling, and D. Heller, Alexandrite lasers for temperature sounding of the sodium layer, in *Proc. 12th ESA Symp. on European Rocket and Balloon Programmes and Related Research*, edited by B. Kaldeich-Schürmann, ESA, SP-370, 395–402, 1995.

Schoeberl, M. R., D. F. Strobel, and J. P. Apruzese, A numerical model of gravity wave breaking and stress in the mesosphere, *J. Geophys. Res., 88*, 5249–5259, 1983.

Senft, D. C., and C. S. Gardner, Seasonal variability of gravity wave activity and spectra in the mesopause region at Urbana, *J. Geophys. Res., 96*, 17229–17264, 1991.

Senft, D. C., G. C. Papen, C. S. Gardner, J. R. Yu, D. A. Krueger, and C. Y. She, Seasonal variations of the thermal structure of the mesopause region at Urbana, IL (40°N, 88°W) and Ft. Collins, CO (41°N, 105°W), *Geophys. Res. Lett., 21*, 821–824, 1994.

She, C. Y., H. Latifi, J. R. Yu, R. J. Alvarez II, R. E. Bills, and C. S. Gardner, Two-frequency lidar technique for mesospheric Na temperature measurements, *Geophys. Res. Lett., 17*, 929–932, 1990.

She, C. Y., J. R. Yu, J. W. Huang, C. Nagasawa, and C. S. Gardner, Na lidar measurements of gravity wave perturbations of wind, density, and temperature in the mesopause region, *Geophys. Res. Lett., 18*, 1329–1331, 1991.

She, C. Y., J. R. Yu, H. Latifi, and R. E. Bills, High-spectral-resolution fluorescence light detection and ranging for mesospheric sodium temperature measurements, *Appl. Opt., 31*, 2095–2106, 1992.

She, C. Y., J. R Yu, and H. Chen, Observed thermal structure of a midlatitude mesopause, *Geophys. Res. Lett.,* 20, 567–570, 1993.

She, C. Y., and J. R. Yu, Simultaneous three-frequency Na lidar measurements of radial wind and temperature in the mesopause region, *Geophys. Res. Lett., 21,* 1771–1774, 1994.

She, C. Y., and J. R. Yu, Doppler-free saturation fluorescence spectroscopy of Na atoms for atmospheric application, *Appl. Opt., 34*, 1063–1075, 1995.

She, C. Y., J. R. Yu, D. A. Krueger, R. Roble, P. Keckhut, A. Hauchecorne, and M.-L. Chanin, Vertical structure of the midlatitude temperature from stratosphere to mesopause (30–105 km), *Geophys. Res. Lett., 22*, 377–380, 1995.

She, C. Y., and U. von Zahn, Concept of a two-level mesopause: support through new lidar observations, *J. Geophys. Res., 103 (D5)*, 5855–5863, 1998.

She, C. Y., S. Chen, Z. Hu, J. Sherman, J. D. Vance, V. Vasoli, M. A. White, J. R. Yu, and D. A. Krueger, Eight-year climatology of nocturnal temperature and sodium density in the mesopause region (80 to 105 km) over Fort Collins, CO (41°N, 105°W), *Geophys. Res. Lett., 27*, 3289–3292, 2000.

She, C. Y., S. Chen, B. P. Williams, Z. Hu, and D. A. Krueger, Tides in the mesopause region over Fort Collins, Colorado (41°N, 105°W) based on lidar temperature observations covering full diurnal cycles, *J. Geophys. Res., 107 (D18),* 4350, doi: 10.1029/2001JD001189, 2002a.

She, C. Y., J. D Vance, B. P. Williams, D. A. Krueger, H. Moosmüller, D. Gibson-Wilde, and D. Fritts, Lidar studies of atmospheric dynamics near polar mesopause, *EOS, 83 (27),* 289–293, 2002b.

She, C. Y., J. Sherman, J. D. Vance, T. Yuan, Z. Hu, B. P. Williams, K. Arnold, P. Acott, D. A. Krueger, Evidence of solar cycle effect in the mesopause region: observed temperatures in 1999 and 2000 at 98.5 km over Fort Collins, CO (41°N, 105°W), *J. Atmos. Sol. Terr. Phys., 64,* 1651–1657, 2002c.

She, C. Y., J. Sherman, T. Yuan, B. P. Williams, K. Arnold, T. D. Kawahara, T. Li, L. F. Xu, J. D. Vance, P. Acott, and D. A. Krueger, The first 80-hour continuous lidar campaign for simultaneous observation of mesopause region temperature and wind, *Geophys. Res. Lett., 30 (6)*, 1319, doi: 10.1029/2002GL016412, 2003.

She, C. Y., and D. A. Krueger, Impact of natural variability in the 11-year mesopause region temperature observation over Fort Collins, CO (41°N, 105°W), *Adv. Space Res*., *34*, 330–336, 2004.

She, C. Y., Initial full-diurnal-cycle lidar observations: variability in mesopause region diurnal-means and diurnal tides over Fort Collins, CO (41°N, 105°W), *J. Atmos. Sol. Terr. Phys., 66*, 663–674, 2004.

Shepherd, G. G., *Spectral Imaging of the Atmosphere*, Academic Press, New York, 2002.

States, R. J., and C. S. Gardner, Influence of the diurnal tide and thermospheric heat sources on the formation of mesospheric temperature inversion layers, *Geophys. Res. Lett., 25*, 1483–1486, 1998.

States, R. J., and C. S. Gardner, Structure of the mesospheric Na layer at 40°N latitude: seasonal and diurnal variations, *J. Geophys. Res., 104 (D9),* 11783–11798, 1999.

States, R. J., and C. S. Gardner, Thermal structure of the mesopause region (80–105 km) at 40°N latitude. Part I: seasonal variations, *J. Atmos. Sci., 57,* 66–77, 2000a.

States, R. J., and C. S. Gardner, Thermal structure of the mesopause region (80–105 km) at 40°N latitude. Part II: diurnal variations, *J. Atmos. Sci., 57,* 78–92, 2000b.

Stebel, K., V. Barabash, S. Kirkwood, J. Siebert, and K. H. Fricke, Polar mesosphere summer echoes and noctilucent clouds: simultaneous and common-volume observations by radar, lidar and CCD camera, *Geophys. Res. Lett., 27*, 661–664, 2000.

Stevens, M. H., R. R. Conway, C. R. Englert, M. E. Summers, K. U. Grossmann, and O. A. Gusev, PMCs and the water frost point in the Arctic summer mesosphere, *Geophys. Res. Lett., 28,* 4449–4452, 2001.

Stroud, W. G., W. Nordberg, W. R. Bandeen, F. L. Bartman, and R. Titus, Rocket-grenade observations of atmospheric heating in the Arctic, *J. Geophys. Res., 64*, 1342–1343, 1959.

Sullivan, H. M., and D. M. Hunten, Lithium, sodium, potassium in the twilight airglow, *Can. J. Phys., 42*, 937–956, 1964.

Summers, M. E., R. R. Conway, C. R. Englert, D. E. Siskind, M. H. Stevens, J. M. Russell III, L. L. Gordley, and M. J. McHugh, Discovery of a water vapor layer in the Arctic summer mesosphere: implications for polar mesospheric clouds, *Geophys. Res. Lett., 28*, 3601–3604, 2001.

Tao, X., and C. S. Gardner, Heat flux observations in the mesopause region above Haleakala, *Geophys. Res. Lett., 22*, 2829–2832, 1995.

Taylor, M. J., M.Gadsden, R. P. Lowe, M. S. Zalcik, and J. Brausch, Mesospheric cloud observations at unusually low latitudes, *J. Atmos. Sol. Terr. Phys., 64*, 991–999, 2002.

Tepley, C. A., S. Raizada, Q. Zhou, and J. S. Friedman, First simultaneous observations of Ca^+, K, and electron density using lidar and incoherent scatter radar at Arecibo, *Geophys. Res. Lett., 30 (1),* doi: 10.1029/2002GL015927, 2003.

Thayer, J. P., N. Nielsen, and J. Jacobsen, Noctilucent cloud observations over Greenland by a Rayleigh lidar, *Geophys. Res. Lett., 22,* 2961–2964, 1995.

Thayer, J. P., M. Rapp, A. J. Gerrard, E. Gudmundsson, T. J. Kane, Gravity-wave influences on Arctic mesospheric clouds as determined by a Rayleigh lidar at Sondrestrom, Greenland, *J. Geophys. Res., 108 (D8)*, 8449, doi: 10.1029/2002JD002363, 2003.

Thomas, G. E., and J. J. Olivero, Climatology of polar mesospheric clouds 2. Further analysis of Solar Mesosphere Explorer data, *J. Geophys. Res., 94*, 14673–14681, 1989.

Thomas, G. E., R. D. McPeters, and E. J. Jensen, Satellite observations of polar mesospheric clouds by the solar backscattered ultraviolet spectral radiometer: evidence of a solar cycle dependence, *J. Geophys. Res., 96,* 927–939, 1991.

Thomas, G. E., Mesospheric clouds and the physics of the mesopause region, *Rev. Geophys., 29*, 553–575, 1991.

Thomas, G. E., Is the polar mesosphere the miner's canary of global change?, *Adv. Space Res., 18*, 149–158, 1996.

Thomas, L., and S. K. Bhattacharyya, Mesospheric temperatures deduced from laser observations of the Na D_2 line profile, *Proc. 5th ESA-PAC Symp. Europ. Rocket & Balloon Programmes & Related Res.* (ESA SP-152), 49–50, 1980.

Thomas, L., A. K. P. Marsh, D. P. Wareing, and M. A. Hassan, Lidar observations of ice crystals associated with noctilucent clouds at middle latitudes, *Geophys. Res. Lett., 21,* 385–388, 1994.

Thomas, L., Lidar methods and applications, in *Spectroscopy in Environmental Science*, Chapter 1, edited by R. J. H. Clark and R. E. Hester, Wiley & Sons Ltd, New York, 1995.

Turco, R. P., O. B. Toon, R. C. Whitten, R. G. Keesee, and D. Hollenbach, Noctilucent clouds: simulation studies of their genesis, properties and global influences, *Planet. Space Sci., 30,* 1147–1181, 1982.

Vance, J. D., C. Y. She, and H. Moosmüller, Continuous-wave, all-solid-state, single-frequency 400-mW source at 589 nm based on

doubly resonant sum-frequency mixing in a monolithic lithium niobate resonator, *Appl. Opt., 37,* 4891–4896, 1998.

Vincent, R.A., and I. M. Reid, HF Doppler measurements of mesopheric gravity wave momentum fluxes, *J. Atmos. Sci., 40*, 1321–1333, 1983.

Volz, U., M. Majerus, H. Liebel, A. Schmitt, and H. Schmoranzer, Precision lifetime measurements on NaI 3p $^2P_{1/2}$ and 3p $^2P_{3/2}$ by beam-gas-laser spectroscopy, *Phys. Rev. Lett., 76*, 2862–2865, 1996.

von Cossart, G., P. Hoffmann, U. von Zahn, P. Keckhut, and A. Hauchecorne, Mid-latitude noctilucent cloud observations by lidar, *Geophys. Res. Lett., 23,* 2919–2922, 1996.

von Cossart, G., J. Fiedler, U. von Zahn, G. Hansen, U. P. Hoppe, Noctilucent clouds: one- and two-color lidar observations, *Geophys. Res. Lett., 24,* 1635–1638, 1997.

von Cossart, G., J. Fiedler, and U. von Zahn, Size distributions of NLC particles as determined from 3-color observations of NLC by ground-based lidar, *Geophys. Res. Lett., 26,* 1513–1516, 1999.

von der Gathen, P., Saturation effects in Na lidar temperature measurements, *J. Geophys. Res., 96 (A3),* 3679–3690, 1991.

von Zahn, U., P. von der Gathen, and G. Hansen, Forced release of sodium from upper atmospheric dust particles, *Geophys. Res. Lett., 14*, 76–79, 1987.

von Zahn, U., and T. L. Hansen, Sudden neutral sodium layers: a strong link to sporadic E layers, *J. Atmos. Terr. Phys., 50*, 93–104, 1988.

von Zahn, U., and J. Höffner, Mesopause temperature profiling by potassium lidar, *Geophys. Res. Lett., 23*, 141–144, 1996.

von Zahn, U., J. Höffner, V. Eska, and M. Alpers, The mesopause altitude: only two distinctive levels worldwide? *Geophys. Res. Lett., 23*, 3231–3234, 1996.

von Zahn, U., G. von Cossart, J. Fiedler, and D. Rees, Tidal variations of noctilucent clouds measured at 69°N latitude by ground-based lidar, *Geophys. Res. Lett., 25,* 1289–1292, 1998.

von Zahn, U., and J. Bremer, Simultaneous and common-volume observations of noctilucent clouds and polar mesospheric summer echoes, *Geophys. Res. Lett., 26,* 1521–1524, 1999.

von Zahn, U., J. Höffner, and W, J. McNeil, Meteor trails as observed by lidar, in *Meteors in the Earth's Atmosphere*, edited by E. Murad and I. P. Williams, Cambridge University Press, Cambridge, UK, pp. 149–187, 2002.

von Zahn, U., and U. Berger, Persistent ice cloud in the midsummer upper mesosphere at high latitudes: three-dimensional modeling and cloud interactions with ambient water vapor, *J. Geophys. Res., 108 (D8),* 8451, doi: 10.1029/2002JD002409, 2003a.

von Zahn, U., and U. Berger, The altitude of noctilucent clouds: groundbased observations and their interpretation through numerical modeling, *Proc. 16th ESA Symp. European Rocket & Balloon Programmes and Relat. Res.,* St. Gallen, Switzerland, June 2003b.

Walterscheid, R. L., Dynamical cooling induced by dissipating internal gravity waves, *Geophys. Res. Lett., 8,* 1235–1238, 1981.

Welsh, B. M., and C. S. Gardner, Nonlinear resonant absorption effects on the design of resonance fluorescence lidars and laser guide stars, *Appl. Opt., 28,* 4141–4152, 1989.

White, M. A., A Frequency-Agile Na Lidar for the Measurement of Temperature and Velocity in the Mesopause Region, Ph.D. thesis, Colorado State University, Fort Collins, Colorado, 1999.

Wickwar, V. B., M. J. Taylor, J. P. Herron, B. A. Martineau, Visual and lidar observations of noctilucent clouds above Logan, Utah, at 41.7°N, *J. Geophys. Res., 107 (D7),* doi: 10.1029/2001JD001180, 2002.

Witt, G., The nature of noctilucent clouds, *Space Res. IX,* 157–169, 1969.

Yariv, A., *Quantum Electronics*, 3rd edition, John Wiley & Sons, New York, 1988.

Yariv, A., *Optical Electronics in Modern Communications*, 5th edition, Oxford University Press, New York, 1997.

Yeh, P., Dispersive magnetooptic filters, *Appl. Opt., 21*, 2069–2075, 1982.

Yi, F, S. Zhang, H. Zeng, Y. He, X. Yue, J. Liu, H. Lv, and D. Xiong, Lidar observations of sporadic Na layers over Wuhan (30.5°N, 114.4°E), *Geophys. Res. Lett., 20 (9),* doi: 10.1029/2001GL014353, 2002.

Yin, B., and T. M. Shay, Theoretical model for a Faraday anomalous dispersion optical filter, *Opt. Lett., 16*, 1617–1619, 1991.

Yu, J. R., and C. Y. She, Climatology of a midlatitude mesopause region observed by a lidar at Fort Collins, Colorado (40.6°N, 105°W), *J. Geophys. Res., 100 (D4),* 7441–7452, 1995.

Yu, J. R., R. J. States, S. J. Franke, C. S. Gardner, and M. E. Hagan, Observations of tidal temperature and wind perturbations in the mesopause region above Urbana, IL (40°N, 88°W), *Geophys. Res. Lett., 24,* 1207–1210, 1997.

Zhao, Y., Stability of the Mesopause Region: Influence of Dissipating Gravity Waves on the Transport of Heat, Momentum and Constituents, Ph.D. thesis, University of Illinois at Urbana-Champaign, Urbana, Illinois, 2000.

Zhao, Y., A. Z. Liu, and C. S. Gardner, Measurements of atmospheric stability in the mesopause region at Starfire Optical Range, NM, *J. Atmos. Sol. Terr. Phys., 65*, 219–232, 2003.