High-resolution mesospheric sodium observations for extremely large telescopes

Thomas Pfrommer and Paul Hickson

University of British Columbia, Department of Physics and Astronomy, 6224 Agricultural Road, Vancouver, BC, Canada;

ABSTRACT

Variations in density structure and altitude of mesospheric sodium impact the performance of adaptive optics systems employing sodium laser guide stars. The associated wave-front errors grow as the square of the telescope aperture and will be very significant for the next generation of large-aperture ground-based optical/infrared telescopes. To support the adaptive optics program for the Thirty Meter Telescope and European Extremely Large Telescope, we are conducting a program of sodium monitoring using a high-resolution sodium lidar system on the 6-meter Large Zenith Telescope (LZT). Located at 49°N latitude, the LZT lidar system provides density profiles with spatial and temporal resolution sampling of 4.8 m and 20 ms. In this paper we report highlights of results obtained over two years of observations.

Keywords: AO, LIDAR, TMT, E-ELT

1. INTRODUCTION

Next generation extremely-large ground-based telescopes with unprecedented light-gathering power will enable astronomers to address many open scientific questions such as the formation and evolution of galaxies. In addition to light-collecting area, another important parameter is resolution. With no atmosphere this is only a function of wavelength and entrance aperture diameter and diffraction-limited performance is achieved with space-based observatories. For ground-based telescopes, the atmosphere distorts incoming parallel wavefronts on time scales down to millisecond. These observatories mitigate atmospheric effects by employing adaptive optics (AO) techniques in which wavefront distortion is measured in real time and corrected by means of wavefront sensors (WFS), deformable mirrors (DM) and powerful computer systems. To match atmospheric time scales, information about the instantaneous wavefront error (WFE) requires bright guide stars. This restricts the use of AO to a small sky coverage because sufficiently bright natural guide stars (NGS) are not distributed densely enough. However, large scale distortions, comparable to the size of the entrance aperture occur on longer time scales and integration times of a few tens of seconds help in this context. To correct for higher order distortions, laser guide stars (LGS) are used. These artificial beacons are generated in the sodium layer, located in the atmospheric mesopause region at ∼90 km, by resonant excitation of Na-atoms. With sufficiently powerful and focused laser beams, tuned to the Na D₂ line, a column through the ∼15-20 km thick layer is illuminated and backscattered photons are created in the direction of the science target. Such LGS appear edge on and hence star-like as seen from directly below. However, with larger entrance apertures, LGS are seen as cylinders with an intensity structure that reflects the density structure of the sodium layer. Such elongation spreads the sodium light over more WFS pixels and therefore reduces signal-to-noise ratios. Sodium layer structure changes due to wind shear, gravity wave overturning and turbulence cause variations in the centroid altitude of the sodium layer. These variations cannot be distinguished from atmospheric focus variations, so the performance of the AO system is degraded.

This effect is proportional to the square of the telescope diameter and therefore, sodium layer dynamics is of great importance to the design of next generation AO systems for extremely large telescopes. Atmospheric lidar investigations study the dynamics of the sodium region for several decades. However, to directly investigate the AO regime, spatial and temporal resolution reported so far, was not sufficiently high. Performance estimates of AO systems for extremely-large telescopes have thus relied on uncertain extrapolation of low-frequency
measurements over four decades in frequency. This motivated us to design a sodium lidar (light detection and ranging) system for the Large Zenith Telescope (LZT), located near Vancouver, Canada at 49° latitude. By sending short (7 ns) powerful (5 W average power) laser pulses with a repetition rate of 50 Hz into the sodium layer, backscattered photon are collected and detected with a four-channel receiver system using high-efficiency photomultipliers (PMT) and fast counting electronics. Compared to conventional lidar systems, the 6-m liquid mirror of the LZT allows us to collect one to two orders of magnitude more photons per laser shot. This provides high resolution Na-profiles with 4 m height and 20 ms time resolution showing for the first time sodium layer dynamics on scales, important to AO systems. Our sodium profiles can be used directly on AO design benches to verify control algorithms and optimize their use for best performance right from first light on. Return flux simulations for existing and new laser systems for LGS AO systems profit from the high resolution data to generate reliable results that help in the design and decision process.

To date about 280 hours of data have been obtained within a 2 year time-span, ranging from February to October. To extract useful information for AO systems several analysis techniques have been investigated and the results are described together with preliminary statistics for the first year data of observations.

2. THE UBC LZT LIDAR SYSTEM

A water-cooled Spectra Physics Nd:YAG solid state laser pumps a Sirah Dye laser. With a 1.2 GHz bandwidth, it is tuned to the Na D$_{2a}$-line at 588.995 nm, the stronger hyperfine part of the D-line, which is in total about 3 GHz wide. The 99% linearly polarized light is directed through beam conditioning optics which widen the beam and collimate it by means of a 15 cm refractor launch telescope. After traveling horizontally to a 25 cm flat gimbal mirror, set to 45°, the beam is directed to the zenith. The gimbal mirror is automatically controlled and used to finetune the beam position on the sky. A beam sampler in the laser path on the optical table extracts a fraction (10$^\circ$) of the laser light and sends it to a high-resolution wavelength meter, used to check the outgoing wavelength. In addition at the beginning and the end of an observing run during one night, the laser power is measured to ensure maximum signal to noise. In order to maintain high power, the dye fluid is replaced after ~15 nights of operation. Due to the lack of running water at the telescope location an air-cooled chiller is used to maintain cool water for the pump laser.

The backscattered photons are by the parabolic rotating mercury surface of the f/1.5 LZT primary mirror, pass through a prime focus corrector, and enter the newly designed receiver. This receiver contains a mechanical chopper wheel (Boston Electronics), used to block Rayleigh backscattered light from the lower atmosphere, and four high-efficiency PMTs (Hamamatsu H7422P-40), each observing one quarter of the pupil in order to reduce coincidence losses. A spherical mirror, cut into four quadrants and slightly tilted outwards, divides the pupil. For each channel, 2-element beam conditioning optics ensure an angle of incidence of no more than ±2°, over the 1 arcmin FoV, on the narrow-bandpass filter. This interference filter has a passing width of 0.26 nm while still maintaining transmission of 90%. The photons are then focused by a concave lens onto the 5 mm diameter photosensitive area of the PMT. With a rise time of 1 ns, the PMTs response is very fast and the output pulse width is of the order 2 ns. The first generation counting system employed four Ortec constant fraction discriminators (CFD) and GHz amplifier electronics, driving high-quality coax cables delivering the signals to a 4-channel counting card (Fast Comtec P7888) located in the control room, about 15 m from the prime focus. The photon counting card used the 20 MHz PCI bus of a PC to transfer the arrival times of each photon to an on-board storage device. Limitations in on-board storage and data transmission resulted in occasional buffer overflows. Additionally, data needed to be transferred from the on-board storing buffer onto the PC hard disk, which would not be done with the system counting. This lead to duty cycles of about 90%, or in other words, every 3 minutes a 20 second gap in the data occurred.

To address the identified shortcomings for the data acquisition, in an upgrade design, a counting system was replaced with a remote system, which was not available at the time of the first design, and which was directly deployable at the prime focus near the receiver. This new Sigma Space counting system is a true 2 GHz system, which also contains a discriminator card with dead times of 2 ns compared to the Ortec CFD 10 ns dead times. With this compact solution photons are counted directly at the prime focus and already conditioned and digitized data is transferred via USB 2.0 connection onto the hard disk in the control room.
Additional components of the lidar system include a temperature-controlled and insulated laser room, an aircraft surveillance radar, an imaging CCD camera attached to a 15 cm reflecting telescope, which is used to control the laser beam position on the sky during data acquisition and a sophisticated trigger system.

To prevent interference, the laser is synchronized with the radar. 600 Hz radar trigger pulses enter a divide-by-twelve circuit that is used to synchronize the chopper at 50 Hz. Following a phase delay, the laser is triggered by the chopper wheel. The Q-switched pump laser triggers—with the respective additional phase delays—the wavelength meter as well as the counting system, which starts accepting pulses after a preset acquisition delay to ensure photons from the correct altitude range are counted. The jitter in the laser was measured with the highest resolution of the counting system (1 ns) and all pulses lied within the first counting bin.

For safe operations, the alarm trigger from the aircraft surveillance radar is connected into the laser interlock path and in case of an aircraft entering the surveillance volume the laser shuts down automatically and would need to be restarted manually. A marine radar (Furuno 1942Mark-2) was modified by replacing the 1.2 m rotating horizontal bar with a feedhorn, whose -3dB lobes fulfill a cone with an opening angle of ±15°. The upper alarm lower range was set between just above the observatory and 14 km. Proper operations were tested with fly-over tests of a single engine air-craft, which verified the theoretically determined cone. The sensitivity is high enough that the system deflects aircraft at all altitudes and even shuts down if a larger bird, such as an eagle, enters the alarm volume. A more detailed description of the UBC LZT lidar system can be found elsewhere.

3. THE SODIUM VARIABILITY

While in previous publications, we showed data, taken with the old counting system, this example is from September 2009 using the new system. The two distinct black bars indicate small observing problems that were resolved during the night. The upper panel shows the sodium structure with a vertical bin size of 75 m and temporal bin size of 2.5 s, the lower panel uses the same data, but with the mean removed to highlight some of the interesting dynamical features. The ordinate represents the absolute altitude above sea level and the abscissa is the time. The color code indicates number of photons per bin. The data has intentionally not been corrected for the geometric effect that photons from higher altitudes subtend a smaller solid angle; hence their density is underrepresented. The reason to not correct for this effect is that unlike for atmospheric physics lidar studies, our aim is to study the sodium layer impact on AO systems. These systems also intrinsically don’t correct for this effect and hence our data matches their need.

By visual analysis of all nights when such density structures have been obtained (see also the panel with more sodium maps in figure 1 of Pfrommer et al. 2009), the only statement about a general behavior that immediately becomes obvious is: There is no “general” sodium structure. Each night reveals different structures on all time scales. As non-conclusive such a statement is, as important it is for AO, if one is to choose a “general” data set to test the behavior of an AO-system on a bench in a laboratory environment. At the altitude of the sodium layer, the atmosphere is about factor of 10^6 less dense than at sea level. At such densities, upwards traveling gravity waves from the troposphere reach amplitudes of order unity, overturn and dissipate. Global energy transport determines wind speed and direction, and, being affected from different locations on earth, strong wind shear occurs. The upper limit of the sodium layer is defined by the recombination of meteoric ablated sodium ions and free electrons from the ionosphere and the plasma, generated during the air-braking process from entering meteors. In addition to the meteor influx function, the source function is also dependent on space weather which affects the electron density, and is therefore not constant. With the sodium atoms being about a factor of 10^8 less abundant than air molecules, the sodium acts as tracer particle to highlight the underlying dynamic phenomena.

If overturning gravity waves dominate the structure in the region of interest, theory predicts distinct peaks in the height profile that appear as sodium clouds containing higher density in the structure maps. Such gravity waves exhibit a downward phase velocity, while their group velocity is directed upwards. Most detected sodium clouds show evidence of a general downward drift with time, also seen in Fig. 1, however other effects such as slowly sinking atoms as an ensemble may play a role, dependent on the source rate. Especially the latter explanation is to be preferred if the cloud originates from a single mg to g-sized meteoroid, whose neutralized ablation plume has been sheared by strong wind shear to the detected sodium cloud. In the sodium layer

---

Proc. of SPIE Vol. 7736 773620-3
multiple clouds exist whose structure evolves with time. It sometimes appears the entire layer is built up out of these clouds rather than one or two clouds being superimposed on a smooth sodium background. The lifetime of an individual cloud is typically a few hours. Some clouds appear rapidly and can reach densities that are a factor of several greater than the mean sodium background. These sporadic sodium layers (SSL) can cause shifts in the mean sodium altitude of several km in just a few minutes. Gravity waves do not only create overturning long range waves that appear in thin distinct clouds, they also create what looks like ripples on time scales with periods of several minutes in our maps (for example around 8:30 UTC in the range of 80 to 90 km in Fig. 1). Such ripples might originate from weather events such as storms from the tropopause that create wavelike structure on small spatial scales. Such buoyancy waves imprint their structure on the sodium layer, which, while being advected across the laser beam, is subsequently detected. Such oscillations extend sometimes coherently throughout the sodium region or are phase shifted, dependent on the gravity wave direction and mesospheric wind environment. On even smaller time scales, less than a few minutes, we often see turbulent structure such as Kelvin Helmholtz instabilities\textsuperscript{11} originating from the stratified shear flow, also seen around 8:00 to 8:30 UTC at \(\sim 92\) km in Fig. 1.

In contrast to the long-lasting SSL, our high spatio-temporal resolution reveals events, lasting from sub-second timescales to several 10s of seconds and extending to a maximum of a few 100s of meters in height. They originate from incoming meteors whose ablated material is sheared and advected across the laser by the prevailing mesospheric wind.\textsuperscript{19,20}
4. TRANSIENTS IN THE SODIUM LAYER

In previous publications\textsuperscript{,11, 14} we reported subsecond meteor trails and discussed their impact on the mean altitude time series. An example of such an event is shown in Fig. 2, where the meteor trail and its effect on the mean altitude time series is easily seen.

![Meteor trail example](image)

Figure 2. Strong meteor trail ∼10 km away from mean altitude and its effect on the mean altitude (in red), shown in the lower half. Superimposed on the mean altitude is a low-pass filtered mean altitude time series (blue dashed line). Such changes in centroid height of more than 1 km over less than a second will cause performance impacts on an AO system if not treated properly. While this event only lasts for 1.5 sec, other events last up to half a second.

We detect events that caused the mean altitude to change by as much as 1 km over less than 1 sec time span. In order to study these events in detail and draw conclusions about their statistical occurrence, automatic meteor-detection programs have been developed to extract these events and measure their size, intensity, current signal-to-noise ratio, location and distance to the mean altitude. To speed up the analysis process, it was decided to use a low-level programming language (C) over more convenient higher-level options.

The raw data count map was Fourier transformed, and an appropriate kernel with matching size was calculated. To maintain zero mean, a narrow positive 2D Gaussian was added to a wide negative 2D Gaussian. The former mimics the meteor content and its standard deviations were chosen to be 30 m/400 ms. To eliminate small-featured background noise, the negative Gaussian was chosen to have standard deviations of 400 m/5 sec. The kernel is normalized so that its integral over both dimensions results in 1. This choice of kernel parameters emphasizes features that are significantly smaller than the negative Gaussian and smoothes over all features, larger than this. Of course, a more thorough survey will be carried out with different settings to check for meteor trails and other larger scale features. However, dependent on the steepness of their borders, even large scale features may be detected with such a kernel, if their edges are distinct, compared to the surrounding background and match the sizes of the positive Gaussian in the kernel. This kernel therefore is also known as edge finder because it enhances edges in the original map. Fourier transforming the kernel and multiplying the Fourier-transformed original map reveals, after transforming back into real space, the meteor trail features, clearly visible in the resulting filtered image, as seen in Fig. 3a). The background with structure larger than the above given scale...
Figure 3. Illustration of the meteor finding technique. a) Extract of filtered sodium density map, using Fourier transform techniques and a zero mean double Gaussian filter. Details about the kernel and the technique are described in the text. The ordinate represents raw pixels on the altitude axis, the abscissa raw pixels on the time axis and color coded are the filtered photon counts with a zero mean kernel. This color code is then represented in the b) histogram of the fourier transformed density map from a) as abscissa. The meteor content shows clearly as the positive tail. c) and d) The corresponding original sodium maps. Especially the color code shows the sensitivity this technique provides towards very faint almost dissipated meteor events. The ordinate represents altitude, the abscissa local time and color coded are photon counts per 60 ms/14.4 m bin

is smoothed out and shows the expected scatter around zero mean. In order to automatically detect meteor trails and measure their parameters, histograms of the entire map are generated, as seen in Fig. 3b). In case no meteor is present, a symmetric function around 0 is expected. In the present case, however a few meteor trails rise above the noise and are clearly detectable in the histogram, that automatically sets the cut at which contour lines are drawn in the filtered image (Fig. 3a)). With the table of contour coordinates, we have all information to output images with the real meteor trails from the original data map and measure all necessary parameters. At the moment this analysis for all recorded data is underway and will be published soon. For the July 2008 data set, preliminary data shows that meteor rates are varying from 20 to a few per hour, and only a small fraction is strong and far enough away to significantly change the mean altitude time series. By this we mean that the mean altitude time series, must follow the meteor intensity and show at least a 2σ dip or bump. As shown in Fig. 4, about 1 to 2 meteors per hour are significant on average, however some nights show strong deviation from this average value and as many as six events per hour could happen.

In addition to the intrinsic statistical analysis, correlation studies are underway to compare our data with meteor radar surveys that are sensitive to mg (100 μg) sized meteors. Apart from the fact that we are sensitive to smaller sized particles, but also detect larger events, it is possible that we can derive predictions, which nights are not preferable for AO observations in case too many events are scheduled. If such action is necessary, depends, if correlation between the two data sets can be established, and, if a detailed analysis with our profiles on AO test benches shows that meteor trails do affect AO systems significantly.
5. STATISTICS OF LIDAR DATA

In summer and fall 2008, 110 hours of sodium layer data were recorded. During the winter months maintenance work at the LZT as well as non-favorable weather prevented further data collection until June 2009. After 16 nights in July 2009 the counting system was changed and data collection resumed at the beginning of August until October, when bad weather began. As this counting system was a custom-designed product a few design flaws were identified and Sigma Space carried out refurbishment during December. Since then data collection continues whenever possible throughout the nights.

The sodium data was investigated by analyzing several different parameters, directed towards the needs of AO system designs that are summarized in the following.

Besides mean altitude time series and the respective mean altitude variance time series, the skewness and kurtosis of the profiles have also been calculated. The idea is to find a parameter that characterizes the time evolution over an entire night, and which can be used to determine necessary update rates for a matched filter algorithm. In this way the update rate could be matched to the current sodium layer conditions to optimize resulting wavefront errors (WFE) due to defocus effects. While WFE take time to calculate, a fast indicator would help one to adjust the wavefront reconstruction to current conditions, especially during the design phase. To find correlations between sodium layer parameter time series and the WFE, a principle component analysis (PCA) is under way that will investigate correlations in multidimensional parameter space, whose components are highlighted in italic throughout this section.

To link sodium layer variations to more physically based parameters, the spatial power spectrum in vertical direction has been calculated, and the slope and normalization time series are analyzed. The slope of the spatial power spectrum is an indicator of the strength of gravity waves, compared to intrinsic turbulent behavior. Power spectrum indices typically vary between -3.5 and -2.23, 24

The RMS of the difference between subsequent mean-altitude corrected sodium profiles that are separated by a certain time range can serve as a useful tool to indicate dynamic changes in the sodium layer. Their close proximity to the WFE, which also takes subsequent sodium profiles as input variable, is thought to be an indicator to relate sodium layer dynamics with AO characteristics. The first few moments of the RMS difference are also under investigation to complete the analysis.

Fig. 5 shows the resulting statistics for the night of 20080716, where profiles were integrated over 1.27 sec and the RMS differences are plotted for different separations, on the right panel the histograms, on the left panel their respective mean as well as the 1σ width.

No significant change in RMS error no matter how far apart the profiles are separated, is obvious. Different integration than the chosen 1.27 s is investigated as next step. The time is chosen because the LGS matched
Figure 5. Change in mean and standard deviations for distributions of RMS mean-altitude-corrected photon count differences for different separations of 1.27s integrated sodium profiles. While the first two boxes on the left panel show smaller RMS differences, this becomes clear as overlaps between successive profiles are present. The larger error bar results from the larger tail in the histograms towards smaller RMS differences, as seen on the right panel. With larger separations between successive profiles, the RMS difference increases slightly, as does the error.

Figure 6. Same as for Fig. 5, however this time with 312 ms integrated profiles (left panel) and 2.54 s integrated profiles

filter update rate for NFIRAOS, the AO facility for TMT, is currently set around 1 sec. If the derived WFE from a 1 sec integration is acceptable, this study shows that the update rate does not necessarily need to be as fast. If a profile 20 sec later is not significantly worse, a matched filter, derived 20 sec earlier, may still be satisfactory. Fig. 6 shows the same RMS differences as discussed for Fig. 5, however this time profiles are integrated over 312 ms (left panel) and 2.54 sec. Note the two different scales on the ordinate, as well as the evolution of the standard deviation, which is flat for the case of 0.3 s integrated profiles in all computed separations, but shows distinctly increasing evolution with increasing separations, after the overlapping profiles are over.

Additionally, the width of the sodium layer is an important quantity during the AO design process as it determines the aperture stops and size of sub apertures and pixel size on the CCD of wavefront sensors. The question arises now how to define “width”. Taking a certain percentile of enclosed intensity is one way, another is a certain percentile of the peak intensity. Fig. 7 shows cumulative distributions for two different nights with quite different behavior. While on the left side (Fig. 7a1 and a2) the sodium profile was relatively smooth and thin throughout the night, the right side (Fig. 7b1 and b2) represents a more dynamic and wider sodium evolution. The upper panels represent enclosed intensity up to the given percentile; the lower panels show the distribution if the thickness is calculated using the given fractions of peak intensity. A statistical analysis for all nights regarding the 50% and 90% distribution of thicknesses is underway and will be presented in a subsequent publication.

From these plots it becomes clear that for at least 90% of intensity, one needs to make sure a WFS should be able to acquire a sodium layer width of 20 km. However this number is still preliminary and needs verification by analysis of all nights accumulated so far. The different behavior of the two nights, shown in Fig. 7, is also
clearly visible in Fig. 8, which shows the mean altitude distribution for the same two nights from Fig. 7.

The previously mentioned principle component analysis is now treated in a descriptive manner. With PCA, correlations in a multidimensional parameter space can be investigated as this method searches for the best-aligned principle axis and computes the variances of all parameter groups from these axes. Eigenvalues for the individual principle components are computed and contributions of the individual parameters are given as values in the normalized eigenvectors. This analysis is possible with the help of L.Gilles (TMT), who provided the code that calculates WFE for the given sodium profiles.

All previously italicically highlighted parameter time series are listed for convenience in the following: Focus and piston removed wavefront error, mean altitude, mean altitude variance, sodium layer 90 percentile thicknesses, calculated after both previously mentioned methods, spatial vertical power spectrum index and normalization, skewness, kurtosis, total intensity and rms difference between successive profiles. With these parameters as input, the first 4 to 5 eigenvalues show significant correlation contributions. While the first 3 eigenvalues and their corresponding eigenvectors reveal correlation contributions from many of the parameters, the WFE is not under these parameters. The eigenvalue and eigenvector for principle component 4 however shows correlation between the wavefront error and the mean altitude variance, the sodium layer thickness and the spatial power spectrum index. On a slightly weaker basis, the rms difference and for some nights and the skewness show correlation signs to the WFE. This will be investigated further in the future and details and results will be given in a subsequent publication. It is however reassuring that correlations from sodium layer parameters to the WFE are present in the current data.
Figure 8. Mean altitude distributions for different nights (left 20080718 ↔ right 20080806). A Gaussian is fit using a Markov Chain Monte Carlo (MCMC) algorithm that also calculates the error bars. While a nice fit is obtained with the data shown in the left panel and therefore points to the statement that the mean altitude is distributed according to the central limit theorem, the right panel shows two distinct peaks that indicate more complex dynamics at that night.

6. CONCLUSION

In the present paper, we reviewed the UBC LZT lidar system with its implemented upgrades and discussed the general dynamical structure of the sodium layer, which shows highly dynamic and complex features on all time scales. A technique to extract automatically meteor trails is presented. Preliminary results show meteor trail rates at about 6 to 20 per hour, however only about 1 to 2 per hour change the mean altitude significantly over the course of less than a second. On nights that experience stronger meteor influx, this rate can go up to 6 significant meteor trails per hour.

ACKNOWLEDGMENTS

Funding for this work was provided by grants from the Natural Sciences and Engineering Research Council of Canada, the Canada Foundation for Innovation, and the British Columbia Knowledge Development Fund. TP gratefully acknowledges support from the German Academic Exchange Service (DAAD). We have benefited from frequent interaction and discussion with the TMT, Herzberg Institute of Astrophysics and University of Victoria AO teams.

REFERENCES


