Atmospheric Characterization of the Space Environment: Unique Observations from Haleakala

Randall J. Alliss
Northrop Grumman Corporation

ABSTRACT

Many space-based applications are impacted by tropospheric atmospherics. For example, atmospheric optical turbulence distorts the wave front of incoming light. Water and ice clouds are often major contributors in the inability of ground based systems to observe their targets. These atmospheric disturbances can often times impact critical, time-sensitive missions from space imaging to optical communications, therefore the ability to detect these disturbances are vital.

Recently, a unique atmospheric monitoring station has been developed and deployed to the summit of Haleakala, HI for the purposes of characterizing the atmosphere above it. Many past observations of atmospheric seeing conditions have been conducted at the summit. Indeed, the site of the Inouye solar telescope was partially chosen due to its favorable atmospheric seeing conditions. However, atmospheric transmission characteristics are much less understood. A whole sky, Infrared Cloud Imager (ICI) has been developed and provides on board, calibrated radiances of the sky from horizon to horizon at intervals up to 20 seconds. The dynamic range of this instrument is from approximately 0 Wm^{-2} to nearly 30 Wm^{-2}. These calibrated radiances are post processed and reduced with the aid of a co-located laser ceilometer to provide atmospheric transmission loss at each field of view. A laser ceilometer provides back scatter profiles of the atmosphere from the ground up to 13 km above ground level at 6 second intervals allowing for near continuous monitoring of the atmosphere. This ceilometer is providing unique insight into the very local atmospherics above it and its potential impacts on Space Situational Awareness (SSA) and other applications. Nearly two years of continuous data have been collected and processed thus far and reveal unique characteristics of the atmosphere above Haleakala Summit. For example, water based clouds, although measured only 30% of the time, have transmission losses of less than 0.5 (3 dB for communication applications) approximately 50% of the time. These observations imply at a minimum, favorable outcomes for many space based applications.

With the assistance of recent computing advances at the Maui High Performance Computing Center (MHPCC), deep learning algorithms are applied to this data in order to learn how to predict when challenging atmospherics are likely to occur. Specifically, a perceptron model is trained from inputs from the ICI and ceilometer. Results from the training and evaluation along with these unique observations will be shared at the conference.

1. INTRODUCTION

The volume of space-based data being generated and transmitted by commercial, government and military has rapidly grown over the last decade. In addition, the number of satellites launched whether large or small is accelerating an even more crowded space environment. Therefore, the need for accurate space situational awareness (SSA) has never been more critical. As users continue to demand more data, the existing communications infrastructure will have to expand to meet the demands. Radio Frequency communications have been relied on exclusively for nearly 60 years but their limitations are becoming more obvious. These limitations include constrained bandwidth, spectrum allocation constraints and in some cases jamming. These technical and regulatory limitations may be alleviated, in part, by Free-Space Optical Communications (FSOC). There are several advantages to using FSOC to meet future communications requirements. In particular, data can be transmitted through free-space via lasers at very high data rates of multi-Gb/s over long distances. Optical beams are also very narrow making them much less susceptible to jamming than radio frequency signals. Additionally, unlike radio frequencies, the optical spectrum is currently unregulated. Finally, optical communications systems are relatively small and potentially much less expensive than comparable radio frequency systems, particularly for air and space missions.

The ultimate realization of practical, high-availability FSOC systems, however, will depend on how well they can mitigate the impacts of atmospheric effects, primarily cloud cover and optical turbulence (OT). Clouds are the largest source of atmospheric attenuation for space-to-ground optical communications, often producing transmission losses of several decibels (dBs) to many tens of dBs. Without impractically large link margins, most clouds are generally considered blockages to FSOC links. For SSA applications, similar impacts may also be found.
Since atmospheric phenomena such as tropospheric clouds are the major limiting factor to the success of FSOC, ground sites for FSOC systems are preferentially located in areas with little to no atmospheric interference (i.e. dry and largely cloud-free sites). Haleakala mountain on the Hawaiian island of Maui rises to a summit level of 10,000 feet. A semi-permanent low-level temperature inversion often traps clouds below the summit reducing the amount of cloud cover. Due to its favorable atmospherics Haleakala and Table Mountain Facility in Southern California are optical ground sites for NASA's upcoming Laser Communication Relay Demonstration (LCRD).

A unique opportunity exists at both sites to collect first ever measurements of atmospheric losses associated specifically with clouds. An array of instrumentation was deployed in 2017 to quantify the optical impacts. This paper summarizes efforts to date taken to characterize the atmosphere over the Haleakala summit. Section two describes the instrumentation deployed. Section three describes an analysis of some of the atmospheric parameters collected. Section four shows how a deep learning based perceptron model is being utilized to make shortterm predictions of cloud cover. A summary of the results and future directions are provided in Section 5.

2. INSTRUMENTATION

In designing an Atmospheric Monitoring Station (AMS), we desired several features. We wanted to collect information on clouds at very high frequency compared to standard meteorological observations which are typically collected on one hour intervals. We wanted to characterize clouds from horizon to horizon in order to understand the correlations of clouds both spatially and temporally and to predict their short term motion. Since not all clouds are optical thick we wanted to develop statistics on the frequency and duration that optically thin clouds occur. Therefore, we wanted to collect the necessary information required to estimate transmission loss. We wanted to design a system that worked autonomously with little intervention from an operator accept for periodic maintenance. We wanted to deploy a system with very little obstructions of the sky, for example, from nearby buildings, domes or antennas. And finally, we wanted to design, procure and deploy a system with a small spatial footprint on the ground without investing millions of dollars.

Figure 1 shows the AMS built and deployed at NASA's Optical Ground Station – 2 (OGS2). The AMS consists of three main instruments, the Vaisala AWS 310 which measures temperature, pressure, wind speed direction, humidity, rain, and incoming solar radiation, the Vaisala Ceilometer 51 (CL51) which measures backscatter off of aerosols and clouds and an Infrared Cloud Imager (ICI) which is a passive infrared device. The ICI is a unique instrument that provides nearly horizon to horizon coverage of the sky. The AMS was deployed in the summer of 2017 with approximately 24 months of data collected as of this writing. Thus far the AMS has collected approximately 95% of

possible hours of data collection. This includes all the times where the site itself was down for maintenance. Other than the temperature sensor being replaced after one year, all other systems have functioned as intended. A more detailed description of each instrument package is provided below.

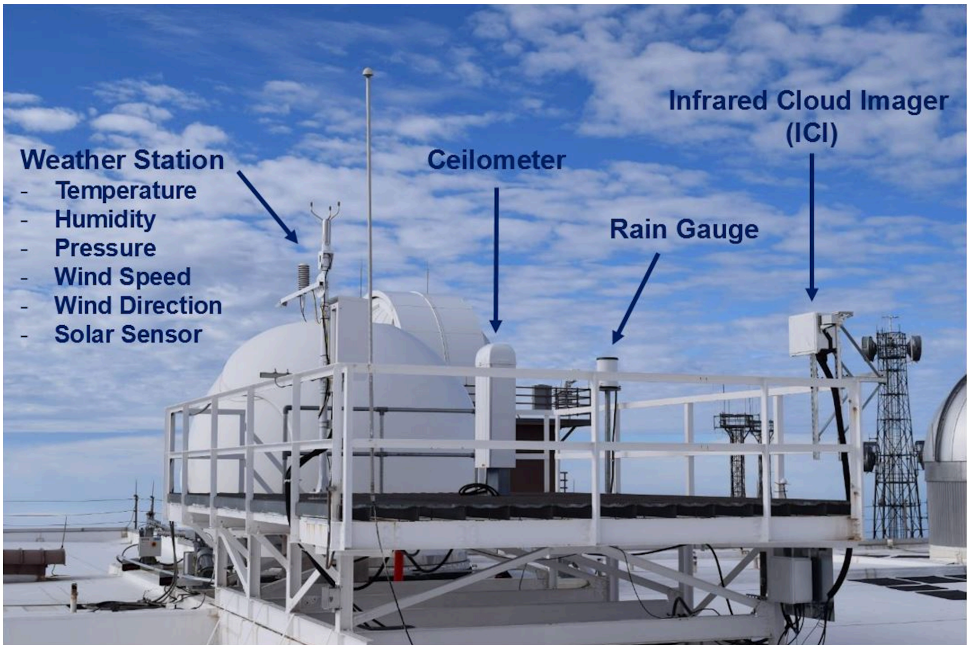

Figure 1. The Atmospheric Monitoring Station (AMS) at Haleakala, HI.

a. AWS310

The Vaisala Automatic Weather Station (AWS) is a standalone meteorological station that measures air temperature, humidity, pressure, wind speed and direction, precipitation amount and incoming solar radiation. The system runs autonomously collecting these parameters at sub one minute intervals. Data is ingested and archived for analysis on a combination of windows and linux servers. A pyranometer was also deployed to measure incoming solar radiation and to estimate the transmission loss due to clouds during the daytime. Figure 2 shows an example of data measured from the AWS310.

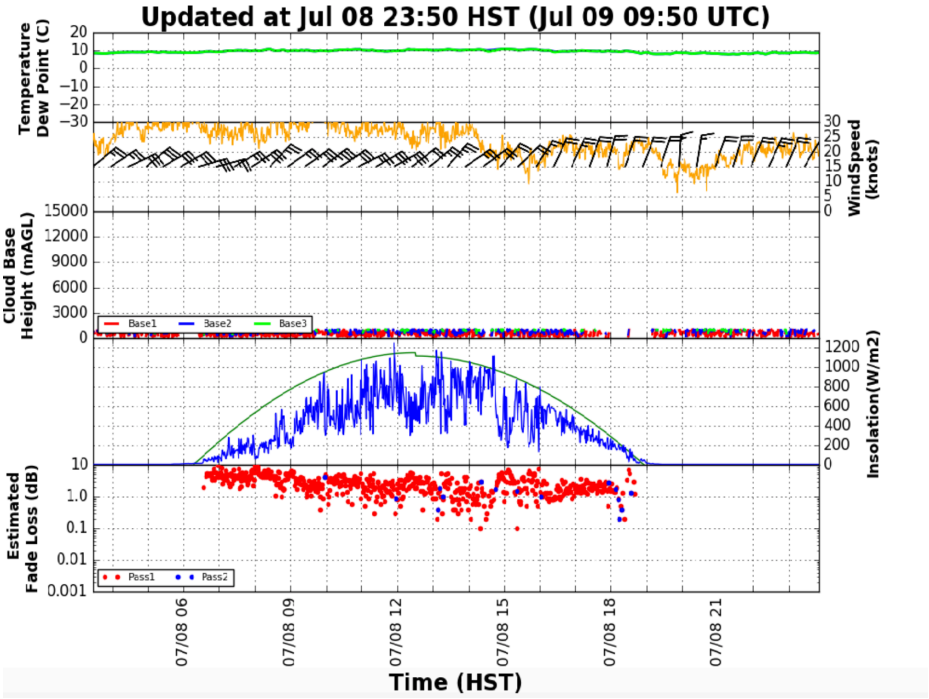

Figure 2. Time series of atmospheric data derived from the AWS310.

The top panel in figure 2 shows a time series of temperature and dew point, the next panel shows wind direction and speed. The third panel shows the derived cloud base height from the CL51 (see more information below) and the fourth panel shows the incoming solar radiation from the pyronometers (blue). The bottom panel shows an estimate of transmission loss (expressed in dB fade loss). This data provides situational awareness on important issues like surface humidity saturation and wind speed for purposes of closing domes. It also shows how conditions can change rapidly on minute timescales. On this particular day clouds covered the summit throughout the daylight hours producing optical fades that varied from a dB to many dB's.

b. CL51

The Vaisala Ceilometer CL51 employs pulsed diode laser LIDAR technology, where short, powerful laser pulses are sent out in a vertical direction. The reflection of light, backscatter – caused by haze, fog, mist, precipitation, and clouds - is measured as the laser pulses traverse the vertical column above the site. The resulting backscatter profile, which is a measure of intensity as a function of height, is stored and processed at six second intervals to compute cloud base heights and transmission loss. The time it takes between the launch of the laser pulse and the detection of the backscatter signal defines the cloud base height. Backscatter profiles are available up to approximately 16 kilometers above ground level. To eliminate background and solar noise as well as shadowing above cloud decks and reverse shadowing above fog decks a two stage filter is applied. Stage one is simple, fast and time-of-day invariant and uses two Gaussian filters (one low and one high). Stage two uses vertical windowed means to model vertical backscatter bias from shadowing and rebounding. Bias is measured in both mean and variance and is computed in double-log display space. The mean is then subtracted from the original data and the variance is suppressed. The resulting filtered backscatter results in minimal signal being lost. Figure 3 shows an example of a time series of filtered backscatter. The cloud base in this case ranges between 2500 and 3000 meters above ground. In this form, the backscatter and resulting power is ready to be inverted to back out an estimate of transmission loss following Klett [2]. This work is currently under development.

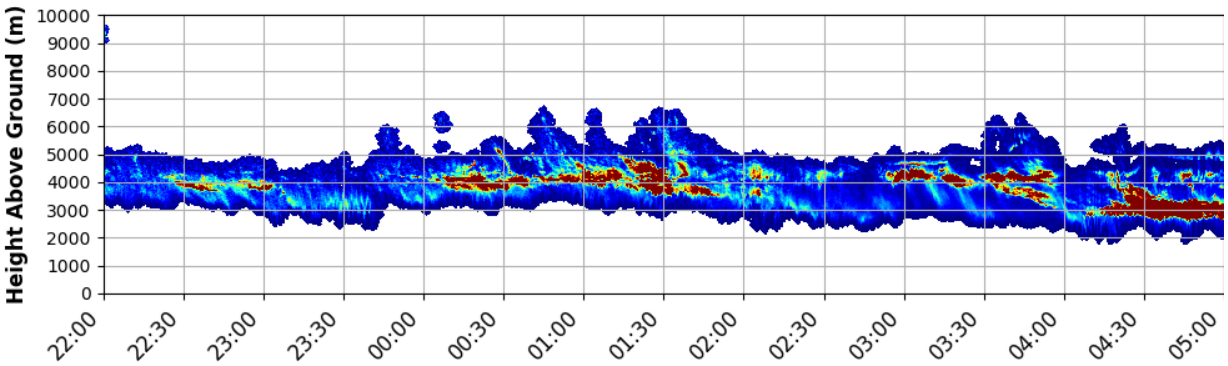

Figure 3. A seven hour time series of filtered backscatter from the CL51. Larger returns (shown in red) are associated with more concentrated cloud droplets and optical thick clouds. Note the variability in the intensity of the backscatter as a function of time.

c. Infrared Cloud Imager (ICI)

The Infrared Cloud Imager (ICI) was developed in conjunction with NASA and NWB Sensors/Montana State University and uses a FLIR Photon 640 camera [3]. Figure 4 shows the ICI system which consists of a camera and electronics (not shown) enclosure. The photon 640 comes mounted underneath a Stingray full sky lens. In order to protect the lens during inclement weather a hatch with rain sensor was designed. The ICI system was designed to capture calibrated IR images at up to 20 second intervals with the ability to shut down in the presence of rain in order to prevent water droplets from collecting on the lens. The shutdown procedure is initiated by the presence of a rain sensor which when triggered by any rain drops sends a signal to the hatch to close over top the lens. Only after the rain sensor no longer detects drops does the hatch re-open and image collection resumes. The ICI has the benefit of being able to collect images day/night in the presence of sun and/or moonlight. The images are stored as calibrated radiances with units of Watts/m^2/steradian. A cloud retrieval algorithm has been developed to interpret each image at the pixel level as cloud or no cloud. This algorithm uses the clear sky background (CSB) technique [4] which evaluates many sky radiance images as a function of time of day and identifies the 10^{th} percentile lowest values. The 10^{th} percentile is typically associated with an image absent of observable clouds based on coincident observations from the CL51. A given image is then compared to the current CSB and if the difference is more than 0.5 Watts then the pixel is labeled cloudy. Example sky radiance images under a variety of atmospheric conditions are shown in Figure 5. In the top left a virtually clear sky is shown whose typical radiance are around just a few Wm^{-2}. The top right image shows just a few clouds near the horizon (white shades) with radiances exceeding 15 Wm^{-2}. The lower left image shows a nearly cloudy sky but with the sun visible (nearly saturated dot near zenith). The lower right image shows a totally opaque sky condition in which not even the sun is visible. It should be pointed that a sun and horizon mask are applied to images during the cloud retrieval process.

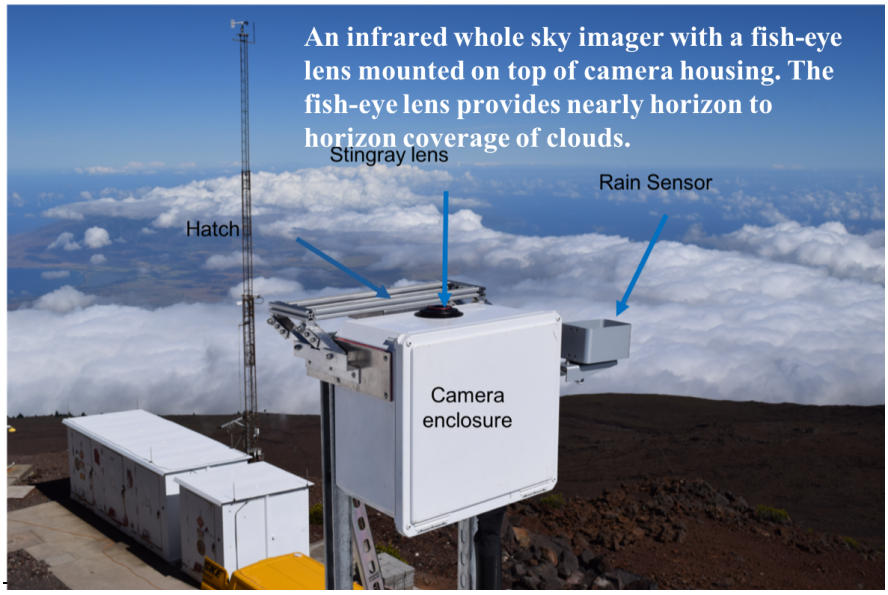

Figure 4. The infrared Cloud Imager (ICI) located above the clouds on the summit (10,000 feet Above Mean Sea Level) of Haleakala, HI.

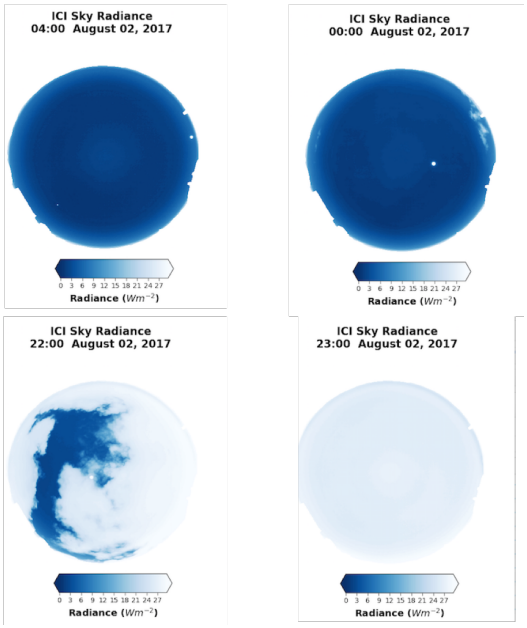

Figure 5. Example ICI sky radiance images (Wm-2). Small values of sky radiance of just a few Watts per meter squared are consistent with clear skies while large values are associated with varying thickness of cloud

3. ANALYSIS OF DATA

Nearly two years of data have been collected and analyzed at the Haleakala site. Figure 5 below shows a summary of the comparison between the GOES derived Cloud Free Line of Sight (CFLOS) and the ICI CFLOS at zenith [5]. The red dots represent the mean CFLOS or availability by month for each month of the year between 1997 and 2018. The blue x's and connected blue line represent the mean of the GOES derived CFLOS or availability over this twenty two year period. The green stars represent the GOES derived CFLOS for the year 2018, the year in which one minute ICI imagery and cloud retrievals are available. Finally, the black diamonds represent the mean CFLOS derived from the ICI cloud retrievals for the year 2018. The GOES footprint is approximately 4km resolution which does not resolve the summit of Haleakala, however, for comparison the pixel at zenith from the ICI was used to derive the CFLOS

estimate. In general, the two show reasonable agreement during 2018, however, the ICI indicates the presence of clouds more frequently than the satellite based record in some months (February-April 2018). This is likely due to the small scale nature of clouds often observed at Haleakala and not resolved by the 4km GOES footprint. It would also appear that April 2018 experienced the cloudiest month on record as denoted by both GOES and the ICI. Generally speaking the mean long term CFLOS according to these measures is greater than 70%. Although there is significant month to month variability (see red dots), the best CFLOS is found in the early winter months and during the summer. The overall correlation between the GOES and ICI CFLOS is approximately 0.65.

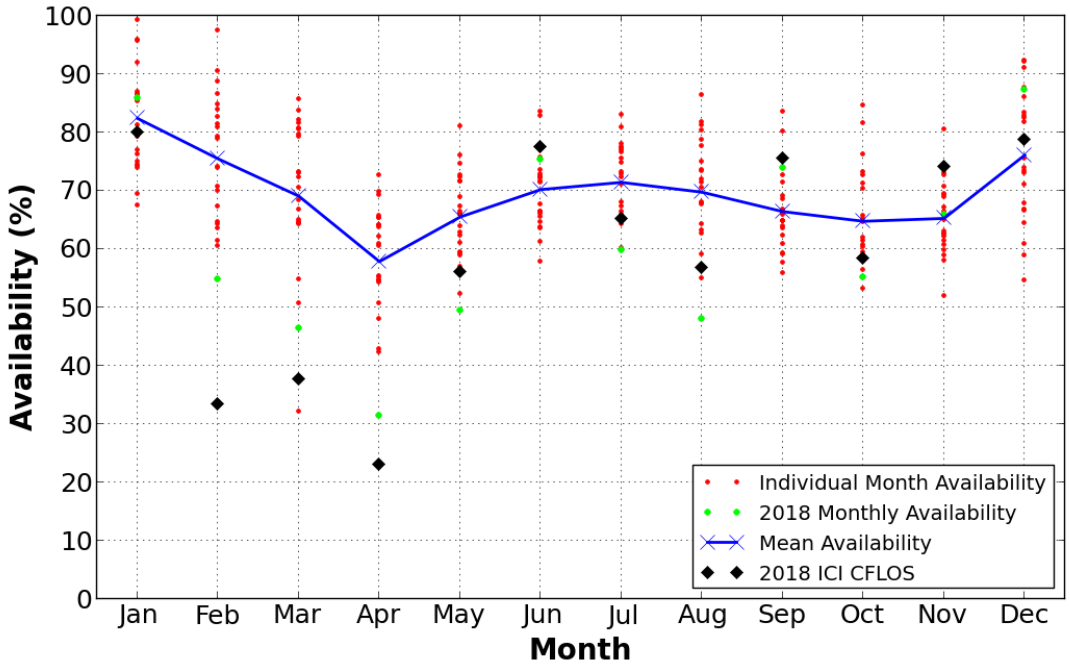

Figure 6. Cloud Free Line of Sight (CFLOS) Availability derived from the GOES climatology, and the ICI in 2018.

An estimate of the cloud transmission loss is derived from an algorithm derived from data collected by the AWS310, ICI and CL51. Figure 7a shows the distribution of power loss expressed in terms of the log base 10 of the transmission or decibels (dB). Although considered very preliminary at this point, when clouds are observed (~27% of the time), the percent of the time they have a fade loss of 3dB or less is approximately 50%. The implication here is with 3 dB of margin for clouds, an optical link could be maintained. Note that fades much greater than 7 dB are considered to be not reliable and come with additional as of yet unquantifiable errors. However, visual inspection of the Cl51 and ICI indicate these clouds are obviously optically thick and therefore any space imaging or optical communications would be quite challenging. Given that the cloud loss is 3 dB or less we investigate the probability it will remain 3dB or less as a function of time to understand the temporal nature of *thin clouds*. Figure 7b indicates that there is a fifty-fifty chance the streak will last greater than four minutes and a 10% probability the streak will last longer than 30 minutes. This may be useful when planning missions and or data collects.

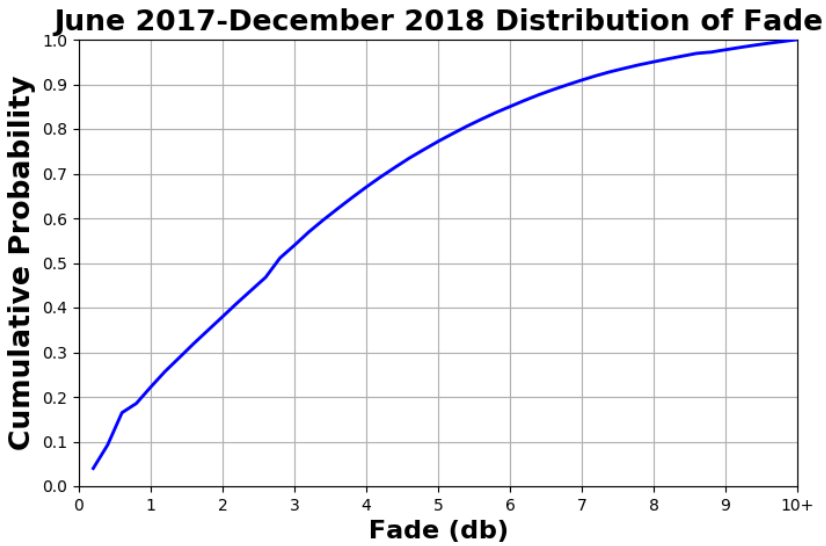

Figure 7a. Cumulative Probability of clouds of different optical thicknesses expressed in decibels (dB) fade.

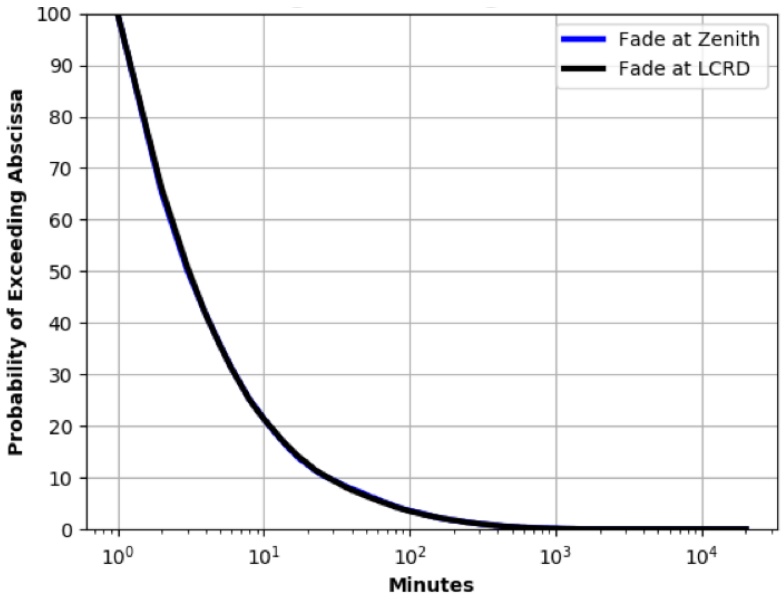

Figure 7b. Exceedance of the conditional probability that a cloud loss is 3dB and remains less than 3dB as a function of time.

To get a handle on where clouds occur in the atmosphere and how frequently they are observed data from the CL51 backscatter profiles are compiled in a way as to determine the cloud base heights. Figure 8 below shows a time history of the frequency of occurrence of clouds as a function of height and time. The figure shows data from June 2017 and May 2019. Vertical profiles of backscatter are translated into base heights and aggregated in the vertical and in time to yield a frequency plot. Areas in white are levels in the atmosphere where no clouds are observed. This could be either because there were never clouds in these layers or that low opaque clouds beneath those levels prevented the CL51 from seeing the layers above them. The majority of the figure shows two layers of clouds, one immediately on or above the summit or 3000 meters above mean sea level (AMSL) or in the 7-12km layer AMSL. The majority of the time clouds are found right on or above the summit, depending on time of day and year. Typically, these clouds are a result of the trade wind inversion layer either being non-existent or weak allowing the convective clouds that form along the slopes of Haleakala during the day to build up and fill in the Haleakala crater and subsequently take over the summit. This is a semi-regular occurrence that typically begins in mid-late morning and extends through the early to mid afternoon hours on those days where the inversion is not suppressed. Often times these types of clouds

are optically thick as visibility at the summit may approach zero. However, the CL51 and ICI also indicate times where these clouds are optically thin and more transient in nature.

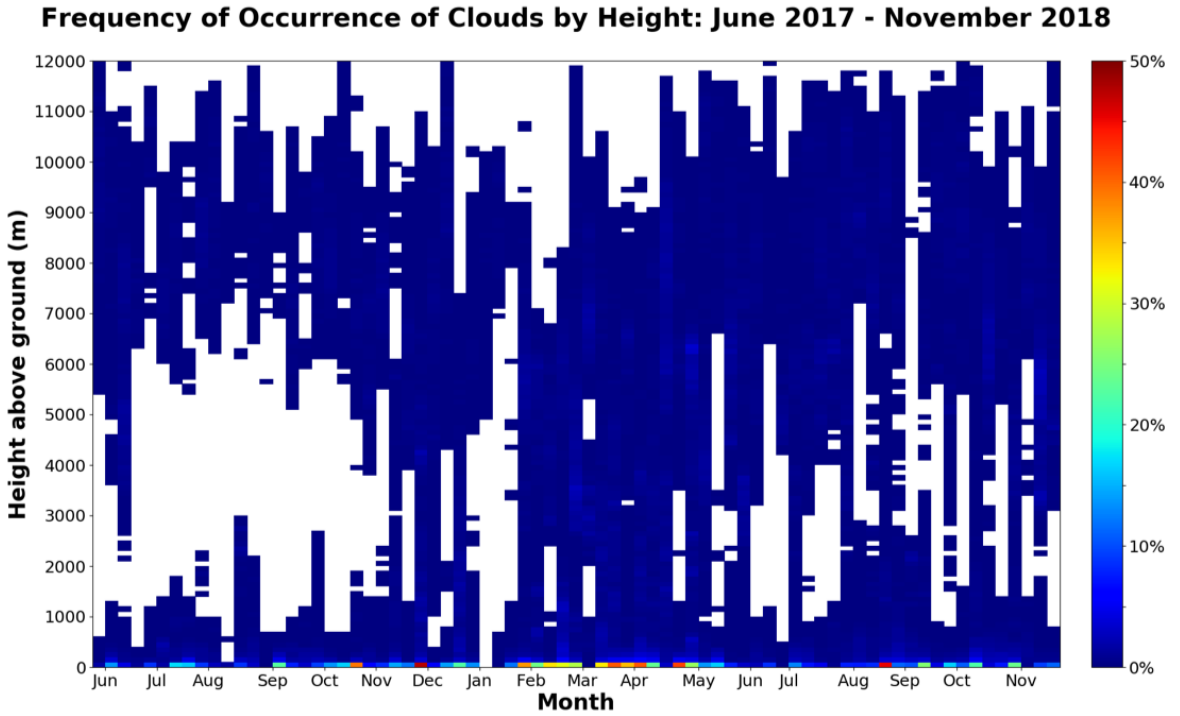

Figure 8. Frequency of occurrence of clouds as a function of time and height above the Haleakala, HI summit.

Remote sensing of clouds from space based instruments such as the GOES over the summit of Haleakala has its limitations (see Figure 6), however, the ICI with its high spatial and temporal resolution allows for the study of the cloud correlations. The correlations are important to understand because they inform us about the predictability of clouds over short time periods of a few minutes to an hour. Figure 9 below shows the spatial correlation of clouds referenced to the location of where the LCRD satellite will reside within the ICI skydome. In this figure, North is up and east is on the left, therefore, LCRD will be located in the East-Southeast portion of the skydome at an elevation angle of greater than 40°. The correlations are based on over one year of cloud retrievals. The derived cloud correlations are very high across the majority of the skydome often exceeding 0.7 down to an elevation angle of just 10-20°. At elevation angles less than 5° the correlations drastically drop to less than 0.2. Based on these observations, we expect clouds upwind of LCRD to be highly correlated with the actual satellite position. In the following section we will explore the use of a predictive algorithm to determine whether it can outperform a simple persistence approach.

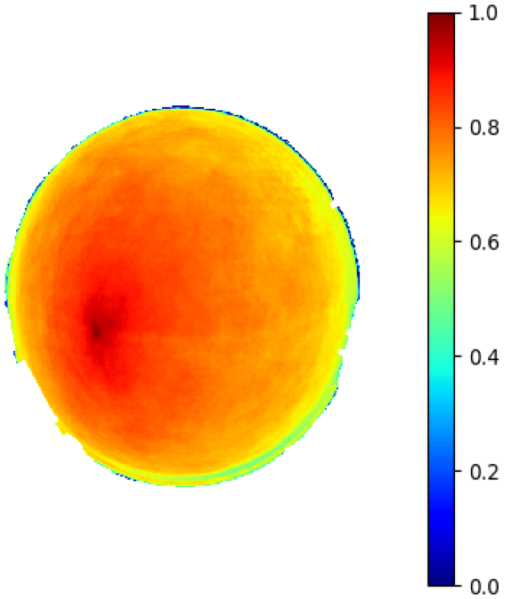

Figure 9. Spatial correlation of clouds across the skydome of the ICI.

4. SHORTTERM PREDICTIONS USING DEEP LEARNING

The characterization of the atmospherics above Haleakala has been demonstrated using retrievals from the AMS. In addition, it is desired to determine whether any predictive capability could be demonstrated using this same data. The predictive bar, so to speak, is quite high at Haleakala since the site is cloud free over 70% of the time. Therefore, a simple persistence forecast would work very well. By definition, however, persistence does not predict change and therefore, during times when clouds are moving *in and out* of a LOS this technique would fail. Figure 11 below shows a Roebber plot indicating the performance of a persistence forecast using clouds retrieved at zenith from the ICI as a function of forecast length. A Roebber plot shows the Probability of Detection, False Alarm Rate, bias and critical skills index (CSI) as a function of forecast length. In this case we show 10 forecast lengths (0,1,2,3,4,5,10,15,30,60), [6].

An investigation is underway to determine if sequences of ICI images can be utilized in such a way to predict the probability of detecting clouds in the future. In this case future is defined as one as long as 60 minutes. The metric for success is, *does it outperform persistence?* A multi-layer perceptron model (MLP) is developed to generate such forecasts that can be compared to the persistence forecast. Simply put, a perceptron is an algorithm used to perform binary classifications: cloudy or clear? It produces a single output based on inputs by forming a linear combination using input weights. A MLP is a deep artificial neural network composed of more than one perceptron. It contains an input layer that receives the signal, an output layer that makes the prediction, and hidden layers in between that do the computational work. The MLP is trained on input-output pairs and models the relationship between the two by adjusting weights and biases of the model to minimize errors. The adjustments are made via back propagation relative to the error (mean squared error, MSE, in this case).

The training and testing set of predictors for the MLP used in this study consists of cloud cover at zenith from the ICI: pixel only, a 10 degree ring around the zenith pixel, and a skydome covering much of the ICI field of view (30 degrees elevation angle and higher). The model is trained with 8 months of data spanning August 2017 – March 2018, and it is tested over a 9 month independent time period spanning April 2018 through December 2018. The model uses the zenith ICI pixel as truth for the model training and output statistics. After the data is read into the model and normalized, missing data is removed. The data is separated into training and testing blocks based on the length of the database, the data lookback window, and the number of forecasts (in this case 10). The lookback window is defined by the period of time leading up to the current time and has been optimized at 60 minutes. The model is not forecasting clouds at the forecast time but forecasting the binary state of presence of any clouds, or absence of all clouds from the current time and out to the forecast time. For example, the 60 minute forecast is not for the single moment 60 minutes out in the future, but over the entire span from the current time out until 60 minutes in the future.

The model uses an MLP with 5 layers. The first layer contains 18 nodes and the last layer contains 15 nodes. The number of nodes in the middle layers are geometrically interpolated from the first and last layers. The internal nodes use a leaky ReLU (Recitified Linear Unit) activation function with batch normalization to converge to a solution. When the MSE has been minimized, a solution is reached. A softmax activation function is also used in the last layer to represent the solution as a probability vector for the two binary states for each forecast length. The result is the probability of "red" (any cloud in the forecast interval) or "green" (for entirely clear for the forecast interval). The final output rests on which probability is higher. Figure 10 shows a graphical representation of the deep learning model.

Figure 11a shows the results achieved from the MLP on a Roebber plot. These results outperform a persistence forecast in terms of probability of detection, bias and false alarm rate (Figure 11b). Accuracy values of the MLP exceed 95% out to a 15 minute forecast. A 30 and 60 minute forecast have an accuracy of 92.0% and 90.1%, respectively. Forecasts on the x=y diagonal have no bias; forecasts below and to the right of the diagonal have a Green (clear) bias, and forecasts above and to the left of the diagonal have a Red (cloudy) bias. Forecast lengths of greater than two minutes have a slight Green bias, which gets stronger with forecast length, but remains very small. The Green/clear bias is due to the way truth is classified. If any one minute of a 60 minute period is cloudy, the truth is considered to be Red/cloudy. This stringent truth requirement is to measure whether a site would be available for the mission over the entire 60 minute forecast interval, rather than at just a single time step 60 minutes from now. The persistence forecasts have a much stronger green/clear bias than the MLP forecasts as shown in Figure 11b. For persistence, this means that relative to the natural frequency of clear/cloudy of the last recorded pixel state, the future states will be increasingly red/cloudy for longer forecast lengths. Thus, the frequency of Green/clear forecasts from

persistence will be greater for longer forecasts. The ability to achieve and exceed the persistence forecast with a simple MLP is encouraging and increases our confidence of mission success.

Future work involves modifying and expanding on the model predictors. Additions may include further sectorizations of the ICI skydome, CL51 cloud heights and cloud decisions, atmospheric parameters from the AWS, and using short term averages of ICI cloud amount.

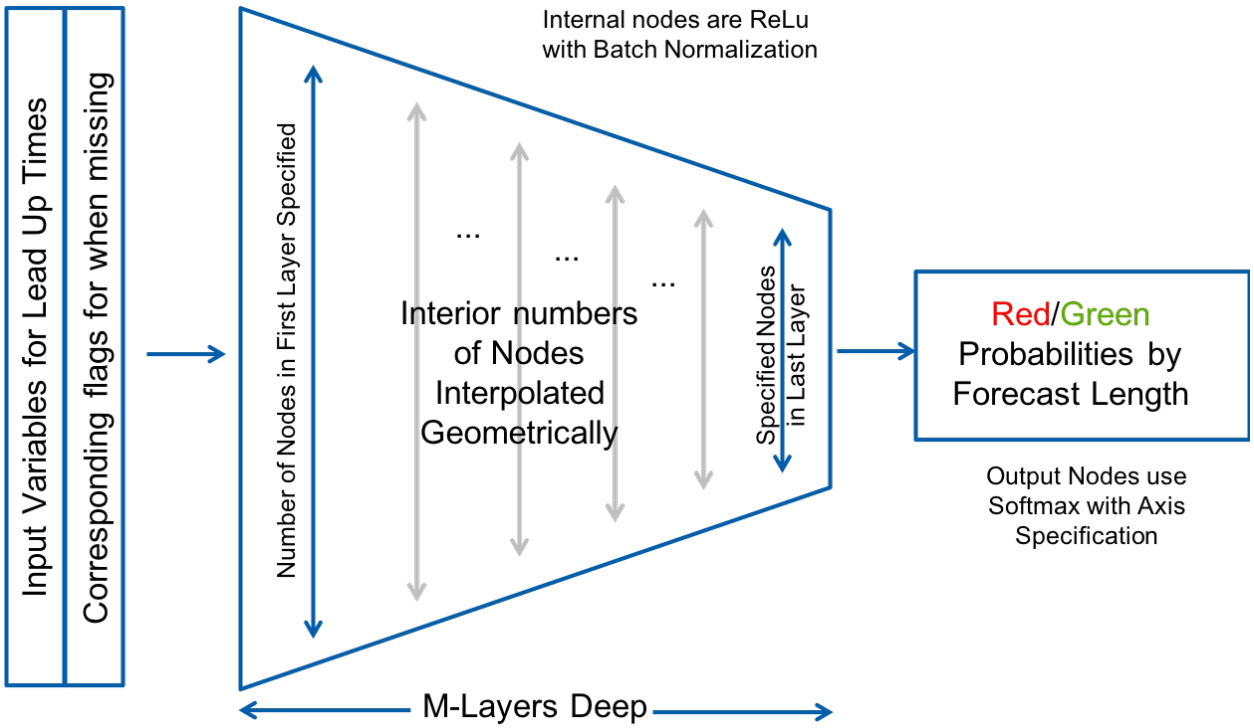

Figure 30. Schematic of Multi-layer Perceptron Model used to generate cloud predictions. Inputs include ICI Line of Sight cloud determination, cloud amount for a small ring around the LOS, and an ICI skydome derived cloud amount.

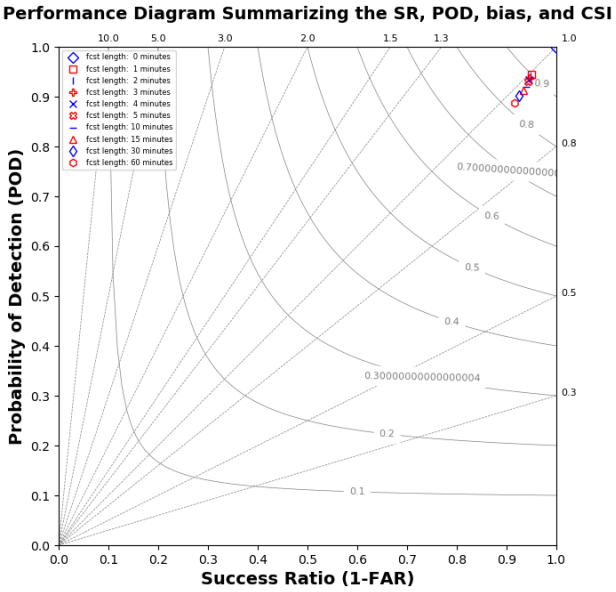

Figure 11a. Roebber plot showing results from MLP model for forecasts intervals of 0 minutes through 60 minutes. Markers in the right corner indicate a perfect forecast. Markers along the diagonal indicate no bias. A green/clear bias is below and to the right of the diagonal and forecasts above and to the left of the diagonal indicate red/cloudy bias. Forecasts have a slightly green bias due to the strict definition of green/clear conditions requiring no clouds be detected over the forecast period.

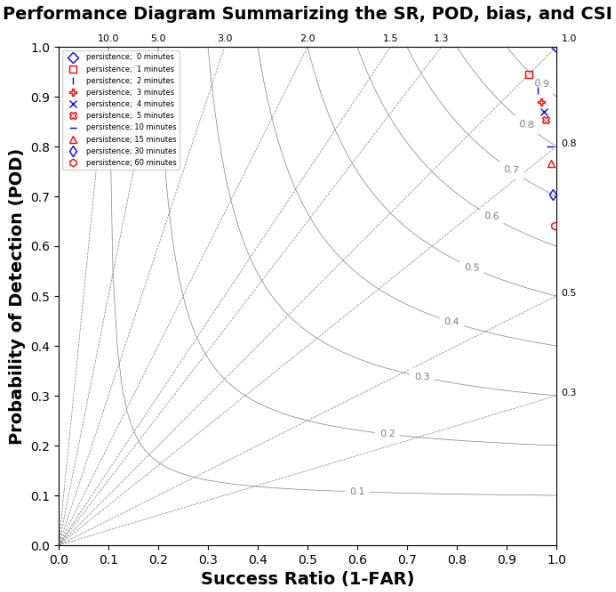

Figure 11b. Roebber plot showing persistence for forecasts intervals of 0 minutes through 60 minutes. Markers in the right corner indicate a perfect forecast. Markers along the diagonal indicate no bias.

5. SUMMARY AND FUTURE WORK

This paper has presented the details and analysis of an inexpensive Atmospheric Monitoring Station designed and deployed to Haleakala, HI. With over two years of data collection, the AMS is providing unique observations over the summit that are assisting in understanding the impact of clouds on space based applications such as optical communications and Space Situational Awareness. We presented data from three instruments including the AWS310 meteorological observation system, the CL51 ceilometer and the Infrared Cloud Imager (ICI). We showed that thin clouds, as defined by those clouds with atmospheric fade loss of less than 3dB, are present nearly 50% of the time. This may have important implications on various space applications including optical communications and space debris tracking. A deep learning algorithm was developed to make short term predictions of clouds based on data from the ICI. These results are encouraging as they outperform a persistence forecast out to 60 minutes. Future work will focus on extending the short term predictions of clouds using additional Deep Learning approaches.

6. REFERENCES

[1] Vaisala Corporation. *Vaisala Cl51 and AWS310 Data Sheets.* Published by Vaisala | B210861EN-E © Vaisala 2018.

[2] Klett, James D. *Lidar Inversion with Variable/backscatter Extinction Ratios, Applied Optics, Vol 24, Number 11, 1985.*

[3] Alliss, R.J., and B.W. Felton: *The mitigation of Cloud Impacts on Free Space Optical Communications*, *SPIE Proceedings, Vol. 8380, 83819 (2012).*

[4] Nugent P.W., Shaw J.A., and Piazzolla S. *Infrared Cloud Imaging in Support of Earth-space Optical Communication. Optical Express, Vol 17, No. 10, 2009.*

[5] Alliss, R.J., Kiley H.S., Mason M.L., Felton, B.D., and Craddock, M.E. *Optimizing the Performance of Space to Ground Optical Communications", Proceedings Volume 10910, Free-Space Laser Communications XXXI; 1091006 (2019).*

[6] Roebber, P.J., Visualizing Multiple Measure of Forecast Quality, *Weather and Forecasting*, Vol. 24, 601–608, 2009.