# Data processing and data products from 2017 to 2019 campaign of astronomical site testing at Ali, Daocheng and Muztagh Ata $^\star$

Zi-Huang Cao<sup>1,2</sup>, Jian Li<sup>1</sup>, Yong-Heng Zhao<sup>1,2</sup>, Zhao-Xiang Qi<sup>2,3</sup>, Hugh R. A. Jones<sup>4</sup>, Huang Shen<sup>5</sup>, Jian-Rong Shi<sup>1,2</sup>, Yuan Tian<sup>1</sup>, Zheng Wang<sup>1,2</sup>, Xia Wang<sup>1</sup>, Jun-Bo Zhang<sup>1</sup>, Jian-Ping Xiong<sup>2</sup>, Tai-Sheng Yan<sup>2</sup>, Xi Zhang<sup>6</sup>, Jin-Xin Hao<sup>1</sup>, Yan-Jie Xue<sup>1</sup>, Jing-Yao Hu<sup>1</sup>, Jian-Sheng Chen<sup>1</sup>, Lu Feng<sup>1</sup>, Jin-Liang Hou<sup>3</sup>, Jin-Min Bai<sup>8</sup>, Ji Yang<sup>9</sup>, Xu Zhou<sup>1</sup>, Yong-Qiang Yao<sup>1</sup>, Li-Yong Liu<sup>1</sup>, Jia Yin<sup>1</sup>, Ali Esamdin<sup>2,7</sup>, Jing Xu<sup>2,7</sup>, Guang-Xin Pu<sup>7</sup>, Peng Wei<sup>7</sup>, Chun-Hai Bai<sup>2,7</sup>, Guo-Jie Feng<sup>2,7</sup>, Lu Ma<sup>7</sup>, Xuan Zhang<sup>7</sup>, Yu Liu<sup>8</sup>, Teng-Fei Song<sup>8</sup>, Hua-Lin Chen<sup>6</sup>, Chong Pei<sup>6</sup>, Xiao-Jun Jiang<sup>1</sup>, Jian-Feng Wang<sup>1</sup>, Jian-Feng Tian<sup>1</sup>, Xian-Qun Zeng<sup>1</sup>, Zhi-Gang Hou<sup>1</sup>, Dong-Wei Fan<sup>1</sup>, Yun-Fei Xu<sup>1,2</sup>, Chang-Hua Li<sup>1</sup>, Yi-Han Tao<sup>1</sup>, Xu Yang<sup>1</sup>, Zhi-Xia Shen<sup>1</sup>, Yun-Ying Jiang<sup>1</sup> and Shi-long Liao<sup>2</sup>

- <sup>1</sup> National Astronomical Observatories, Chinese Academy of Sciences, Datun Road 20A, Beijing 100101, China
- $^2\,$  University of Chinese Academy of Sciences, Beijing 100049, China
- <sup>3</sup> Shanghai Astronomical Observatory, Chinese Academy of Sciences, Nan Dan Lu 80, Shanghai 200030, China
- <sup>4</sup> Centre for Astrophysics, University of Hertfordshire, Hatfield, UK
- <sup>5</sup> School of Journalism, Communication University of China, Beijing 100024, China
- <sup>6</sup> National Astronomical Observatories Nanjing Institute of Astronomical Optics & Technology, Chinese Academy of Sciences, Nanjing 210042, China
- <sup>7</sup> Xinjiang Astronomical Observatory, Chinese Academy of Sciences, 150 Science 1-Street, Urumqi 830011, China
- <sup>8</sup> Yunnan Observatories, Chinese Academy of Sciences, Yang Fang Wang 396, Kunming 650216, China
- <sup>9</sup> Purple Mountain Observatory, Chinese Academy of Sciences, Ban Cang Jie 188, Nanjing 210033, China

**Abstract** Based on previous site-testing and satellite cloud data, Ali, Daocheng and Muztagh Ata have been selected as candidate sites for a large optical/infrared telescope in China. We present the data collection, processing, management and quality

<sup>\*</sup> zhcao@bao.ac.cn

analysis for our site testing based on the use of similar hardware. We analyse meteorological data, seeing, background light, cloud and precipitable water vapour data from March 10, 2017 to March 10, 2019. An investigation has also been made into the relative usefulness of our all-sky camera data in comparison to that from the meteorological TERRA satellite data based on a night-by-night comparison of the correlation and consistency between them. We find a 6% discrepancy arising from a wide-range of factors.

Key words: LOT, Site-selection, data processing

#### **1 INTRODUCTION**

The topography of China is high in the west and descends to the east coast. The middle is dominated by basins, the Gobi desert, rolling plateaus, towering massifs and the highest plateau on Earth. These Western lands might be suitable for astronomical observations for astronomical observations in the optical and infrared bands due to their higher altitude, lack of water vapor and human activity as well as better atmospheric transparency. Similar to most cultures older observatories were located in or near existing large cities. After the 1980s, site-testing gradually turned to the mid-west of China and founded some astronomical observatory sites, such as Delingha (Tian et al. 2016), Nanshan (Liu et al. 2013), and Gaomeigu (Chen et al. 2003). Altitudes of these sites with 2067 to 3216 meters above sea level are systematically higher than the sites established before the 1960s, which are no more than 1000 meters. These new sites resulted in a significant improvement in the available astronomical conditions and accumulated more experience for the site testing in western China. The latest site-testing work has launched in late 2016. Combining experiences of previous site-testing and satellite cloud data (Cao et al. 2019), Ali (Liu et al. 2019), Daocheng (Song et al. 2019), and Muztagh Ata (Xu et al. 2019) were selected for long-term site data monitoring. The distribution of observatories and candidate sites is shown in Fig. 1 and briefly discussed below.

The Ali site is located in the southwest of the Tibetan province at the top of a mountain range where several measurement points have been investigated. The main instruments were deployed at the Point A1 (longitude: 80.02671° E, latitude: 32.32573° N, altitude: 5,040 m) see upper plots in Fig. 2.

The Daocheng site, located between Ganze county and Daocheng county in Sichuan province, with two measurement points, including Point 1 at Mountain Wumingshan (longitude: 100.1089° E, latitude: 29.10695° N, altitude: 4,739 m) and Point 2, the main site of inference for night observation at the same mountain (longitude: 100.10890° E, latitude: 29.10695° N, altitude: 4,739 m). The Daocheng instrument deployments are shown in the middle of Fig. 2.

The Muztagh Ata site is located in the Eastern Pamirs (longitude:  $74.89676^{\circ}$  E, latitude:  $38.33044^{\circ}$  N, altitude: 4,526 m). The instrument deployments are shown at the bottom of Fig. 2.



The candidate sites referred to in this

Fig. 1 Geographical distribution of observatories and candidate sites in China. The observatories marked in red were established before the 1960s. Observatories in yellow were established after the 1980s. The candidate sites referred to in this research are marked in blue.

The sites are spread between 75° E and 100° E which represent a latitude region relatively lacking in large optical and infrared telescopes and at altitudes higher than 4,500 meters. Thus it is probably that such a telescope should be in a competitive position to make useful contributions to time-sensitive monitoring observations and astronomical events of particular interest (e.g. Brown et al. 2013), A next generation optical/infrared telescope will likely require an adaptive optics system so that the full range of meteorological data deserves attention, including weather, night cloud cover, night seeing, night sky background and Precipitable Water Vapour (PWV). Furthermore, it is necessary to refer to the full range of 24-hour meteorological data for the construction and routine maintenance of the observatory infrastructure and the night-time parameters should be analyzed for the astronomical observations in optical and infrared bands. The meteorological measurement instruments deployed at each site started to collect data uniformly on 2017 March 10. The operation of each site and data processing are performed by independent teams.

This article introduces and discusses the detail of the site-testing data acquisition and processing, and the arrangement is shown as follows: In Section 2, we briefly introduce the acquisition and processing of weather station (WS) data and Precipitable Water Vapour data. In Section 3, we analyse the differential image motion monitor (DIMM) raw data acquisition and introduce our data processing. In Section 4, we describe the Sky Background Meter (SBM) data, All-sky Camera (ASCA) data and data processing. In Section 5, we will show our site-testing data management and quality control. In Section 6, we investigate the correlation and consistency of the ASCA and satellite data. In Section 7, we summarize our conclusions.

Value	Sensor	Resolution	Range	Precision
August Temperature	HMP155(A)	0.01 $^{\circ}\mathrm{C}$	$-50 \sim +60 \ ^\circ \mathrm{C}$	$\pm 0.1~^\circ\mathrm{C}$
Relative Humidity	HMP155(A)	0.1~%	$0 \sim 100~\%$	$\pm 2~\%$
Pressure	EL15-1A	$0.1 \ \mathrm{hPa}$	$500 \sim 1100 \ \mathrm{hPa}$	$\pm 0.2~\mathrm{hPa}$
Wind speed	EL15-2D	$0.1 \mathrm{~m/s}$	$0\sim75~{\rm m/s}$	$\pm 0.3~{\rm m/s}$
Wind direction	PTB210	$3^{\circ}$	$0 \sim 360$ $^\circ$	$\pm5$ $^{\circ}$

Table 1 The critical parameters obtained from CAWS-600 Automatic weather station

# 2 WEATHER STATIONS AND PRECIPITABLE WATER VAPOUR DATA

## 2.1 Weather Stations Data

Each site is equipped with a unified Huayun CAWS-600 six-element automatic weather station<sup>4</sup> to ensure the precision and consistency of the meteorological data. This includes temperature, air pressure, relative humidity, wind speed, and wind direction, which can be measured, stored, and transmitted by WS. All WSs have passed a standard quality check at the factory. Their installation was made by the local operations team at each site. The WS data measurement interval can be set from 1 second to N minutes, and the output average values and min/max values are derived from all samples within the corresponding data measurement period. All key parameters of the WS are taken from the manual, and listed in Table 1.

We found the sampling frequencies of the raw data can be unstable due to the setting was not fixed after a power interruption. Following an agreement on the appropriate sampling interval of 10 minutes, we have re-sampled the original data of each site uniformly to produce median values. All data and inferences are based on this re-sampled data. We note that randomly missing data or data sampling rate instability has less impact on the statistics, but the loss of data for multiple days might result in deviation from annual trends and overall statistical results.

Meteorological data for the Ali site is from the Point A1 WS. Data of the Point B4 WS are used for the days without A1 data. The data from the Ali Point A1 and B4 WSs is shown in Fig. 3 (a). A total of 675 days of data was collected from March 10, 2017 to March 10, 2019. The main reason for lost data is power failures.

Meteorological data for the Daocheng site is from the Point 2 WS, covered from 6:00 PM to 8:00 AM (all times are Beijing time), which results in a little difference between 24-hour and nighttime statistic. After more than a year of operation, these WSs were moved to Point 1 at April 2018. A total of 559 days of data was collected from March 10, 2017 to March 10, 2019, shown as Fig. 3 (b). The main reason for the loss of data is power failures.

Meteorological monitoring of Muztagh Ata started March 10, 2017, however, initial observations were made with a protective plastic cover in place which does not give useful readings. Meorological monitoring of Muztagh Ata started March 10, 2017, however, initial observations were made with a

 $<sup>^{1}\</sup> https://earth.google.com/web/@32.32573,80.02671,5040a,1000d,35y,0h,0t,0r$ 

 $<sup>^{2}</sup>$  https://earth.google.com/web/@29.10695,100.10890,4739a,1000d,35y,0h,0t,0r

<sup>&</sup>lt;sup>3</sup> https://map.tianditu.gov.cn/

<sup>&</sup>lt;sup>4</sup> http://www.cnhyc.com/showcpzs.asp?id=489

protective plastic cover in place which does not give useful readings. Relative-humidity data before August 3, 2017 has been removed. After the analysis of the wind direction data, the Muztagh Ata operation team found the output of the wind direction sensor was fixed at 0. After several failed repairs, the wind direction data was accepted from the backup CAWS-600 automatic weather station, about 10 meters away with the same sensors. The statistic of the Muztagh Ata meteorological data is shown in Fig. 3 (c). A total of 675 days of data was collected from March 10, 2017 to March 10, 2019. The reasons for lost data are power failures, installation problems, and sensor failure.

# 2.2 Precipitable Water Vapour Data

A Low Humidity and Temperature Profiling microwave radiometer (LHATPRO), manufactured by Radiometer Physics GmbH (RPG) in Germany, is used to take measurements of PWV at the three candidate sites. The instrument was deployed at Ali site from March 19, 2018 to March 22, 2018, at the Daocheng site from February 10, 2018 to March 5, 2018, and at the Muztagh Ata site from January 22, 2018 to February 4, 2018. Since the instrument requires professional engineers to maintain and consumes liquid nitrogen, the number of days with observation data at each site is limited. Based on the operation journal and astronomical twilight, we divided the whole day (24 hours) data into "night" and "clear night" for detailed analysis.

# **3 DIFFERENTIAL IMAGE MOTION MONITOR DATA**

A differential image motion monitor is an instrument generally used for measuring seeing in the astronomical observatory, which was firstly proposed by Sarazin and Roddier (1990). Three types of DIMMs are used in this work, which are developed independently by ALCOR-SYSTEM<sup>5</sup> (French DIMM for short), Nanjing Institute of Astronomical Optics and Technology, Chinese Academy of Sciences<sup>6</sup>, (NIAOT DIMM for short) and Ali operation team from National Astronomical Observatories of China, Chinese Academy of Sciences<sup>7</sup> (Ali DIMM for short). The parameters of DIMMs are listed in the Table 2 and briefly described below.

The ALCOR-SYSTEM is a French company, focusing on the instruments and software in the field of Astronomy<sup>8</sup>. The DIMMs produced by ALCOR-SYSTEM are the standard instrument for the site testing projects. The main components of their DIMMs are from the reliable manufacturers, and the Integration and software were finished by ALCOR-SYSTEM.

The NIAOT-DIMM is low-temperature enhanced, designed for the seeing measurement in Antarctica. The DIMM is deployed at the Antarctic Kunlun Station in 2011 and an enhanced model deployed at the Antarctic Taishan Station in 2014<sup>9</sup>.

<sup>&</sup>lt;sup>5</sup> http://www.alcor-system.com/new/index.html

<sup>&</sup>lt;sup>6</sup> http://www.niaot.cas.cn/

<sup>&</sup>lt;sup>7</sup> http://http://www.bao.ac.cn/

<sup>&</sup>lt;sup>8</sup> http://www.alcor-system.com/new/Company/Whoweare.html

<sup>&</sup>lt;sup>9</sup> http://english.niaot.cas.cn/ns/201402/t20140228\_117038.html

Parameters	French DIMM	NIAOT DIMM	Ali DIMM
Focal Length (mm)	2400	1600	2500
Focal Ratio	8	8	10
Subaperture Number	2	2	2
Subaperture Diameters (mm)	51	50	50
Subaperture Distance (mm)	240	149	200
Wedge angle (arcsec)	Variable	50	60
Pixel Size( $\mu$ m)	5.86	5.5	7.4
Exposure Time (ms)	Dynamic	5  and  10	10
Center Wavelength (nm)	550	500	500
Sites Equipped	Ali, Daocheng,	Daocheng	Ali
	and Muztagh Ata	and Muztagh Ata	

Table 2 The parameters referred in this work of DIMMs

The Ali DIMMs were developed for the site testing in western China, which have been deployed at the Ali site for many years. The same type of DIMM was used for the seeing measurement of the LAMOST site (Liu et al. 2010).

The parameters of the DIMMS referred in this work are given below.

Each of sites deployed 2 or 3 DIMMs, one for continuous measurement, the others acting as backup. An extra French DIMM was used for cross calibration between sites to ensure the seeing data is accurate and comparable. A consistency test between French DIMMs was conducted at the Xinglong Observatory by Wang (2019), and the average difference of measurements is 0.075 arcsec. The French and NIAOT DIMM consistency test was conducted at the top of 11 m tower of Muztagh Ata site by Xu (2019), and the median value of the French DIMM is a little higher than the NIAOT DIMM by 0.01 arcsec approximately within the measurement error. The French and Ali DIMM consistency test was conducted at the Ali site by Liu (2019), and the median value of French DIMM is lower than the Ali DIMM by 0.06 arcsec. The operation height of DIMM is important to avoid the near-ground turbulence, which has been tested by the operation team in Muztagh Ata and Ali independently. From June 23, 2017, to November 14, 2017, Xu (2019) deployed the NIAOT DIMM on the ground with the French DIMM on the 11 m tower at the Muztagh Ata site. The median value from NIAOT DIMM is 0.97 arcsec and French DIMM is 0.79 arcsec on the 11 m tower. Then, the NIAOT DIMM moved to the top of the 6 m tower and stay there from November 15, 2017 to September 20, 2018. The median values at 6 m and 11 m both were both found to be 0.87 arcsec. Similar to Muztagh Ata, two DIMMs were deployed at the 5 m and 11 m towers by Liu (2019) at the Ali site, and the median, 25% and 75% of the DIMMs comparison mentioned above are summarised in Table measurements are the same. The tests indicate that the operation height is important to avoid the near-ground turbulence, and the altitude of the near-ground turbulence is below 6 m at the Muztagh Ata and Ali sites.

DIMM operation is susceptible to environmental elements, e.g. wind speed, temperature. Higher wind speeds can cause vibrations in the DIMM tube and its tower, which could result in larger seeing measurements. The DIMM tube vibration is easily detected and can be solved by adding a Site Testing Campaign for the Large Optical/Infrared Telescope of China

Equipments	Sites	Period	Difference of medians
Difference of	Xinglong	March 3, 2017 –	0.075 arcsec
French DIMMs		March 12, 2017	
NIAOT-French	Muztagh Ata	April 3, 2017 –	0.04 arcsec
		April 12, 2017	
NIAOT-Ali	Ali	April 26, 2017 –	0.06 arcsec
		May 20, 2017	
French@11 m-	Muztagh Ata	June 23, 2017 –	0.18 arcsec
NIAOT@Ground		November 14, $2017$	
French@11 m-	Muztagh Ata	November 15, 2017 –	0.0 arcsec
NIAOT $@6 \text{ m}$		September 14, 2017	
French@11 m-	Muztagh Ata	September 21, 2017 –	0.01 arcsec
NIAOT@11 m		November 20, $2019$	

Table 3 The difference of medians of French, NIAOT, and Ali DIMMs at different positions.

wind shield. Tower vibrations are more difficult to detect, and the tower improvement requires more effort, sometimes even reconstruction. There is a rough seeing comparison of the impact of tower vibrations caused by wind in Fig. 4. The left-hand side of the figure is a histogram of the seeing measured by the French DIMM from June 23, 2017, to October 8, 2018, on top of the Muztagh Ata 11 m tower. The right-hand side of the figure is the histogram of the seeing measured by the same DIMM at the same place from October 8, 2017, to March 10, 2018, after the tower was reinforced on 2018 October 8. We can roughly estimate from the histograms that the reinforcement of the tower has a positive impact on the improvement of the seeing measurement. More detailed analysis about the relationship between the seeing and wind speed of prevailing wind direction in Muztagh Ata was finished by Xu (2019). He found the the median seeing value shows an initial drop until the wind speed reaches 3 m/s, then keeps growing. Xu suggests that the reason for this phenomenon could be a stable ground turbulence occurring at a specific wind speed. Another possibility is that the breeze accelerates the heat balance of the tube and tower cooling. Comparison between the DIMMs is shown in Table 3.

The formation of a temperature inversion layer can be important meteorological phenomenon that is beneficial to the seeing. Under normal circumstances, the atmospheric temperature decreases with the increase of altitude. However, under certain weather conditions, the atmospheric structure above the ground will have an abnormal phenomenon that the temperature rises with the increase of altitude, which is called a temperature inversion. Approximately one month after the establishment of the high-precision micro-temperature tower (Hu et al. 2019), it is confirmed that there is an apparent temperature inversion phenomenon in Muztagh Ata. The temperature gradient is more obvious below 6 meters (Hu et al. 2014). Thus in Fig. 5 we use the measurements at 2m and 6m and can see that 80.6% of the time there is a temperature inversion layer with an average of  $0.35^{\circ}$ C and that the temperature gradient is about  $0.08^{\circ}$ C/m. For the inversion layer formation, we assume that the Muztagh Ata site location in the Eastern Pamirs which is a series of high peaks including Kongur Tagh (7,719 m), Kongur Tiube (7,555 m), and Muztagh Ata (7,546 m) results in the cold air at night not passing over an upper layer of hot airflow. The "inversion layer" can make the atmosphere stable, reduce the phenomenon of turbulence in the atmosphere, and effectively ensure enhanced seeing, which is confirmed by Fig. 6. The seeing was measured by the French DIMM on the ground, and the micro-temperature tower is not far from the DIMM. So far, in western China, only the Muztagh Ata site has found this phenomenon and so is appropriate for more detailed analysis in Xu et al. (2019).

Temperature also has some negative effects on the measurement of seeing. Sunlight increases the temperature of the DIMM instrument and its tower in the daytime, then, the rapid drop in ambient temperature after sunset may cause turbulence inside the DIMM tube and local turbulence due to the exothermic heat of the tower, which results in an apparent increase in the seeing. Proper daytime sun protection for DIMMs and the addition of insulation to the tower can significantly improve this situation. During the development of the sites the DIMMs have moved from the ground to the tower and measurements have improved thus overall reported values slightly underestimate the site performance. A schedule for DIMM improvements has been planed by Ali and Muztagh Ata operation teams and is expected to be completed by 2020.

During the development of the sites the DIMMs have moved from the ground to the tower which have measurements upgraded, thus some reported values might underestimate the site performance to some extent.

The seeing data at the Ali site are mainly based on results measured by the Ali DIMM at the top of the 11 m tower. Due to poor initial tower strength the DIMM vibrated slightly following wind speed, causing measured values to be increased. After a few months, the Ali team reinforced the tower to reduce the negative effects on seeing. Meanwhile, another Ali DIMM monitored at the 3 m height ceiling of the observation building. We combined data sets from these two instruments, and the combined raw data includes a total of 457 days from March 10, 2017 to March 10, 2019.

The Daocheng site DIMM is on a 7.5 m observation platform at Point 2 and moved to a similar Point 1 after April 2018. The Daocheng seeing measurement is mainly based on the NIAOT DIMM, and the data of French DIMM were used when NIAOT DIMM data was not available. There is a total of 357 days of combined raw seeing data from March 10, 2017 to March 10, 2019.

The French DIMM was first commissioned on the ground at the Muztagh Ata site on April 15, 2017, and moved to a 11 m tower for observation since June 23, 2017, so as to avoid nearsurface turbulence. Reinforcing engineering for 11 m tower was conducted to increase the measured precision of DIMM on the 11 m tower was completed on October 8, 2017. The NIAOT DIMM worked on the ground at Muztagh Ata from March 12, 2017, then moved to 6 m tower from November 15, 2017, finally deployed on 11 m tower since September 21, 2018. The French DIMM on the 11 m tower is the main data source for the seeing at the Muztagh Ata site and the data of the NIAOT DIMM were used when French DIMM data were not available. There is a total of 389 days of combined raw seeing data from March 10, 2017 to March 10, 2019.

We calculated the average sample rate of each DIMM, and found French DIMM is much higher than other DIMM measurement frequency 1 time/minute. For consistency of DIMM data, we resampled the French DIMM to make a median before merging the different DIMMs data of each

**Table 4** Steps in the data processing of the DIMM data from March 12, 2017 to March 12,2019

Steps	Ali	Daocheng	Muztagh Ata
Original data	$178406 \ (457 \ days)$	290400 (357  days)	454877 (389 days)
After resample	$178406 \ (457 \ days)$	114316 (357  days)	117155 (389  days)
< 20  times/day	$156 \ (18 \ {\rm days})$	$165 \ (14 \ {\rm days})$	73 (8  days)
< 2 hours/day	1330 (36  days)	1113 (31  days)	$1981 \; (38 \; {\rm days})$
Processed data	$177013 \ (416 \ \mathrm{days})$	113168 (324  days)	$115161 \ (350 \ \mathrm{days})$

site. For continuous seeing data, we removed data with lower frequency than 20 times/day or 2 hours/day on the basis that this indicates underlying instrument issues usually due to power. The intermediate data statistics after each step are shown in Table 4, and the comparison of histograms before and after data processing are shown in Fig. 7. It should noted that the data "< 20 times/day" and "< 2 hours/day" at each site have overlap with each other and that given the relatively minor offsets found between DIMMs summarised in Table 3 we do not apply any corrections to the reported DIMM measurements.

# 4 SKY BACKGROUND METERS AND ALL-SKY CAMERA DATA

# 4.1 Sky Background Meter Data

The sky background meter manufactured by Unihedron (Unihedron, 2016) was deployed at each site for the measurement of night sky background with the unit of magnitudes per square arc second (mag  $\operatorname{arcsec}^{-2}$ ). The factory nominal accuracy of this device is  $\pm 0.1$  mag  $\operatorname{arcsec}^{-2}$  within the visible band. Before the deployment, three SBMs have been tested in Xinglong Observatory from February 12, 2018 to February 12, 2018 by Wang (2019). All test measurement differences are within 0.1 mag  $\operatorname{arcsec}^{-1}$ , and the absolute values of difference median are within 0.02 mag  $\operatorname{arcsec}^{-2}$ . A total of 351, 344 and 342 days of data was collected from March 10, 2017 to March 10, 2019 at the Ali, Daocheng, and Muztagh Ata sites respectively. The periodic trends in the original data indicate the impact of moonlight and cloudy. Since we point the SBM at the local zenith, the measurements are susceptible to moon altitude, moon phase, and moon azimuth. In order to avoid these effects, we removed the data when the moon is more than 6° above the horizon. Cloud might also dim the sky background to some extent. Therefore, we cross-matched the SBM and ASCA data and removed data classified as "covered" and "inner". Fig. 8 shows the differences between each step. The right-hand side of the histograms is for sky without moonlight and cloud interference.

#### 4.2 All-sky Camera Data

In order to detect local cloud each site is equipped with the same model of All Sky Camera developed by Wang (2019). The ASCA is a fisheye lens consisting of a Canon Digital Single Lens Reflex Camera with  $3456 \times 5184$  pixels. The camera captures images with different exposure time based on altitude of the sun and the moon phase. After the hardware updated, the exposure time

Class	Description
Clear	No cloud within the inner circle. Fig. 9 (a)
Outer	No cloud within the inner circle, has cloud within the outer circle. Fig. 9 (b)
Inner	No more than 50% cloud within both the inner and outer circle. Fig. 9 (c)
Covered	Coverage of cloud within inner+outer circle are over 50%. Fig. 9 (d)
None	Hard to determine cloudiness for the underexposure, dew, etc.

 Table 5
 The definition of cloud cover

was controlled by a new micro-controller reference to the night sky background measured by the local SBM.

We processed ASCA images following the Skidmore (2011) methodology. Images are divided into inner and outer circles with zenith angles of  $44.7^{\circ}$  and  $65^{\circ}$ , and each circle was marked by the bright stars. According to the distribution and thickness of the cloud on the image, cloud cover is determined as "clear", "outer", "inner", "covered", and "none". The definition of the classification is shown in Table 5, and the typical images are shown in Fig. 9. The interval of the images is about 5 minutes. We interpreted images into cloud-cover time slices and accumulated them during the night constrained by astronomical twilight. We note that Skidmore (2011) use a "Moon up" classification for moonlight-interfered images, but we treated the images with the moon as usual, because we don't want to lose the information from cloud-cover affected by the moonlight. The "none" classification we defined is equivalent to the questionable classification defined by Skidmore (2011). The images with "none" are excluded from the final result.

Fig. 10 shows our webpage of ASCA cloudiness interpretation to improve efficiency, facilitate collaboration within the team, and control the quality of image interpretation. Once ASCA image data are transmitted from each candidate site to our servers, original images are stored uniformly into the database.

Our webpage presents auto-scaled images and provides for further image enhancement but is not reliable in complex conditions, e.g., moon, thin cloud, and glare cause bias. Personnel was trained before engaging in data classification in order to try to produce uniform manual treatment. In order to verify the precision and stability of images classification, an independent team processed ASCA images obtained from January 27, 2017 to March 26, 2017 at the Daocheng site in the same way. The comparison dataset includes 58 nights, and the comparison result is shown in Table 6. The results of the two groups are generally close, especially for the classification of "covered". It shows that the "covered" is mostly obvious and can easily be recognized. The situation for clear nights is also relatively straight forward with good agreement. For the "clear", in the case of proper exposure time and no cloud, the classification is mainly based on a uniform skylight background and a bright foreground star. The discrepancy between the groups appeared to arise from the cases with dust, dew, and frost. This causes uneven background and star gloom, which makes people associate cloudy conditions. Also, in some time periods, there are fewer bright stars in the sky, or the exposure time is short, which makes people mistakenly think that there are thin clouds. Therefore, the judgment of "clear" is still dependent on human experience, especially in one case when the exposure time was not provided the teams. The interpretation of "inner" and

Class	Other team	Our team
Clear	65.34%	67.85%
Outer	2.28%	4.2%
Inner	9.67%	5.28%
Covered	22.5%	22.66%
None	0.21%	0.01%

**Table 6** Comparison of the results obtained by the interpretation of the Daocheng ASCA databy different teams from January 27, 2017 to March 26, 2017

**Table 7** The night cloud cover statistical data from March 10, 2017 to March 10, 2019 in Ali,Daocheng, and Muztagh Ata

Site	Processed	Clear	Clear+Outer	Clear+Outer	Clear+Outer	Clear+Outer+Inner
	Images			+Inner	>3 hrs	>3 hrs
Ali	92.7%	65.0%	70.8%	79.4%	71.9%	82.1%
Daocheng	81.8%	52.5%	58.2%	66.3%	60%	68.0%
Muztagh Ata	90.1%	53.7%	61.6%	70.6%	63.1%	73.1%

"outer" depends on the determination of the edge of the cloud and is less consistent between the teams. As the Fig. 11 shows, the "inner" is more complex than "outer" for including two different cloud distributions. For cases with a clear moonless night, the dark background and the presence of bright background stars make a good reference, and it is easy to make a distinction between the interpretation of "inner" and "outer". For data with moonlight, the glare between the inside of the lens or the lens and the acrylic cover outside can easily cause people to misinterpret "outer" or "clear" as "inner", or the opposite direction. By comprehensively interpreting multiple images, we can analyze the movement of the cloud and the movement of the glare with the moon, which can reduce the misjudgment of "inner" to some extent, but the discrepancies between the teams exist and further work is necessary. The example of outliers mentioned above is shown in Fig. 12.

The distribution of the night cloud cover at Ali, Daocheng, and Muztagh Ata are shown in Fig. 13 (a), Fig. 13 (b), and 13 (c) respectively. The Fig. 14 (a), Fig. 14 (b), and 14 (c) are Ali, Daocheng, and Muztagh Ata normalized night cloud cover by night-time. The Table 7 is the three sites' night cloud cover statistical data derived from ASCA from March 10, 2017 to March 10, 2019.

# **5 DATA MANAGEMENT AND QUALITY CONTROL**

The Ali site has existed for many years and has a relatively solid hardware foundation. Daocheng was originally a candidate site for solar telescopes and it began to accumulate data such as nighttime meteorology and seeing from 2015. Muztagh Ata is a new site and required setting up related hardware at the end of 2016. Therefore the infrastructure of each site is different, this can impact the data directly. For example, the power supply critically impacts the acquisition of observation data, and the communication situation affects the transmission of observation data. Except for the DIMM, each site is equipped with the same type of measuring instruments, and Ali and Daocheng's

Level	at each site: Each Identification	Production modes and significance of data at all levels
0	L0	Raw data uploaded by sites.
1	L1	L0 data are saved into the database upon data formatting or
		parsing to form L1 data.
2	L2	Normalize data format and adding observational days, L1 data are
		generated into L2 data with simple merger application.
3	L3	L2 data are generated into L3 data upon changing the original time
		scales (from minute to day, month and year), then cross-matching
		different data sources (including weather, seeing, etc.).
4	L4	Charts and statistics.
5	L5	Reports.

 Table 8
 Data level identification

original observation equipment was used as backup. The teams at each site are responsible for the installation, commissioning, operation of the equipment and the quality of the raw data. The raw data was transmitted to NAOC and processed on a daily basis. The daily report is generated and sent to some astronomers, then the data quality is fed back to each site maintaince team to improve the data production. This procedure sometimes improved the original data quality, e.g. improving the DIMM measurement software for the manufacturer and changing the DIMM running parameters.

#### 5.1 Transmission, Initial Processing and Backup

Data transmitted daily includes WS data, SBM data, DIMM data, and ASCA images. Each of the sites transmits nearly four gigabytes of data through various methods. The transmission was affected by local meteorological, communication, traffic and other conditions, and often not carried out reliably e.g. due to power and communication failures. The data transmission, timely archiving, and appropriate initial processing of these important data were challenging. The quantity and quality of the site data can directly reflect the installation, operation, and maintenance of the relevant hardware, and so information must be fed back to the relevant person in charge of each site in a timely manner so that each site can quickly adjust the hardware parameters and improve the operation and maintenance.

The disaster recovery of precious site data is also one of our primary concerns. We need to apply mature technologies and establish a reliable and effective data backup system. It is often necessary to accumulate data for a certain period of time before proceeding to the next step. Further processing of the data involves site data of different types and time spans. In order to facilitate the management and use of the data, we abstracted the original and derived site data into six levels. Then, we stored and managed the data in a layered manner. The higher the level the higher the degree of processing reached in Table 8.

Based on hierarchical data management, we have established a standardized data processing method. Operation and maintenance teams from each of the three sites are in charge of equipment operation and data acquisition including site transmits daily data (L0) to our server; those data are decoded and saved into the database (L1) upon parsing; after normalizing the data format, adding observational days and processing a simple data merger, data are generated into L2 data; and then, some of the L2 data (including weather and seeing) is processed into various charts(L4), which are issued in a form of daily report (L5) with the purpose of ensuring the ongoing operation of each site and checking its data quality. If abnormal data are found, feedback will be sent to the site as soon as possible. Data of three sites are also treated and analyzed on a regular basis (L3) to generate statistical results (L4), charts (L4), and relevant results of various reports (L5). The whole process of data transmission and initial data treatment at each site is shown in Fig. 15.

Due to the communication conditions and hardware restrictions, the data transmission methods of the three sites are throughout (1) Real time data are transmitted to our servers via communication modules with related equipment on site, e.g. Muztagh Ata WS. (2) TCP protocols are utilized to synchronize data from site servers to our servers, e.g. all Ali data. (3) FTP protocol is utilized to manually upload data to our servers, e.g. Muztagh Ata and Daocheng ASCA images. (4) Data can also be sent by email manually, e.g. data synchronization following a network failure.

After the original data was transmitted to our servers, preliminary data processing began immediately. Depending on the device and the mode of transmission, the raw data (L0) may use various formats such as CSV, TXT, and XLS. Also, a set of scripts has been compiled to decode and process various kinds of files in order that they can be saved in the MySQL database (L1). EXIF information is extracted from image data (L0) of ASCA, which will be saved into the database (L1) together with file information (L1). On the basis of saving original images, ASCA TIF/RAW from the three sites are converted into PNG format with a unified image resolution.

In order to find problems in a timely manner and feedback to the teams at each site, we check the site data daily. The inspection of data is divided into two aspects: transmission checks and data quality checks. The transmission check is to make sure the daily data is transmitted at each site. We have a webpage to interactively check data quantity and generate a data exchange daily report. Raw data quality checks and further data analysis is based on data visualization. Hence, three types of reports are generated: daily report of daily meteorological statistics; technological report of certain topic, eg. DIMM comparison; and summary report of showing all data inferred in the site-testing. As Fig. 16 shown, a website has been developed for information inquiry and report download.

In terms of the daily report, only meteorological and seeing scatter diagrams of each site are plotted. Fig. 17 is an example of the daily report on March 19, 2018 in Ali site. The technological report only concerned about one instrument in a short period for specific engineering or scientific propose. The time span of the summary report generally ranges from several months to several years and all instruments involved. The data are presented in the summary report in a variety of ways, including scatter diagrams, histograms, compressed monthly statistical charts, and rose diagrams.

The cloud platform of the China Virtual Observatory (Cui et al. 2015) has been adopted to back up raw data (L0) from each site, intermediate results (L1, L2, L3) and products (L4, L5).

Specifically, data receiving and treating systems are deployed by cloud computing nodes; structured data are saved in the (RDS), while unstructured data are stored in the NAS. All data will be preserved in the long term. Incremental backup data will be pushed actively from the server to the online storage deployed in National Astronomical Observatories of China (NAOC) underground server room every half month. At the same time, an audit report will be automatically generated in the system and manually reviewed upon the increment of data backup each time.

# 6 CORRELATION AND CONSISTENCY OF THE ASCA AND SATELLITE DATA

Correlation and consistency of the cloud inferred from ASCA is close to the view of the potential telescope but due to recent installation the time span of the available data is short. Meteorological satellites could potentially give useful long-time span data for the study of site cloud cover and satellite data are essential for site-testing research (e.g., Cao et al. 2019). The cloud condition from the ASCA and the physical cloud cover measured by ASCA and satellite will be different (Schöck M., et al. 2009).

From a physical definition, cloud cover from ASCA and satellite are different. Clouds are generally encountered over a range of altitudes varying from sea level to the top of the troposphere and can be vertically divided into three levels<sup>10</sup> by the cloud altitude. The levels overlap and their limits vary with latitude. The camera points to the zenith direction and its field of view is  $180^{\circ} \times 180^{\circ}$  region with a fixed size. The cloud cover condition is classified by the cloud distribution and thickness in the images, referring two circles stack on the image, with a zenith angle of 44.7° and 65°. As Fig. 18 shown, clouds of different altitudes have different diameters in the sky for different circles. On the other hand, one pixel of a polar-orbiting satellite (MODIS) corresponds to the specific region with a fixed size. The satellite measurements of cloudiness are dependent on the observational bands and processing algorithm for different altitude clouds. Therefore, the consistency of the cloud measurement between the satellite and ASCA data is affected by many factors, and we can only analyze the relevance of these two types of cloud data from a statistical perspective without reference to other data.

Cloud cover in western China has been examined using MODIS data by Cao (2019). MODIS data is from the Earth Observing System (EOS) project, which was started by the United States National Aeronautics and Space Administration (NASA) in 1999, including EOS/TERRA and EOS/AQUA satellites. These polar-orbiting satellites observe the earth four times a day at local times around 10:30 AM, 1:30 PM, 10:30 PM, and 1:30 AM (Platnick et al. 2003). The resolution at nadir is 1 km. We compared the MODIS data with ASCA results at Xinglong observatory (40.39555° N, 117.57744° E) in 2015. We processed the images closest to 10:30 PM, corresponding the TERRA satellite observation time at night, and 354 images were processed following the work-flow mentioned above. We note that we choose the 10:30 PM TERRA data due to a relatively

	Clear	Outer	Inner	Covered	None	
Total	90	33	54	146	31	
Percentage	25.4%	9.3%	15.3%	41.2%	8.8%	

Table 9 The ASCA cloud cover at Xinglong in 2015.

In 2017 and 2018 instrument teams purchased, calibrated and compared high-precision and uniform measurement instruments for three potential telescope sites in Western China.

Table 10 The outliers in "Clear+Outer" and "Inner+Covered"

	Jan.	Feb.	Mar.	Apr.	May	Jun.	Jul.
Clear+Outer	1	0	0	0	1	3	0
Inner+Covered	0	0	4	4	0	0	2
	Aug.	Sept.	Oct.	Nov.	Dec.	Total	Percentage
Clear+Outer	0	0	0	1	5	11	8.9%
Inner+Covered	0	1	2	0	0	13	6.5%

smaller number of ASCA images being available for Xinglong around 1:30 AM apparently arising from power outages. The ASCA-based cloud coverage at Xinglong in 2015 is shown in Table 9.

We matched the percentage of cloud from MODIS and ASCA rresults. The distribution of satellite data based on percentage cloudiness along with ASCA data classified by "clear", "outer", "inner" and "covered" scenario and are shown in Fig. 19. and on a night-by-night basis in Fig. 20. In the construction of Fig. 19, we assume that satellite cloud cover greater than 80% in the "clear" and "outer" scenarios and less than 20% in the "covered" and "inner" scenarios are outliers. All outliers were checked manually, and three images were moved from "inner" to "none" based on underexposure. The remaining images are listed in Table 10 grouped by month. We note the outliers in the "clear+outer" and "covered+inner" are distributed into different months without overlap. There are 24 outliers in Table 10 and thus we find a 6.8% (24/354 images) discrepancy between ASCA and satellite classification for the Xinglong site. This is reasonably consistent with TMT's 3–6% (Schöck et al. 2009).

#### 7 SUMMARY

The operation team from Ali, Daocheng, and Muztagh Ata installed all instruments, and completed the construction of necessary power supply facilities, communication facilities, observation towers, as well as other site infrastructure improvements. The achievement of this task in sparsely populated areas with average site altitudes over 4,500 m where the stability of power supply and communication cannot be guaranteed and the measurement of data is accompanied by the construction of the station was a challenging task. With regular data analysis, the independent teams established site testing data management guidelines and a data processing flow to ensure data quality. We have provided detailed information about the raw data processing, the utility of related data products, and shared problems with data management which might provide useful material for those who make and use site-testing results.

#### ACKNOWLEDGEMENTS

The research is partly supported by the Operation, Maintenance and Upgrading Fund for Astronomical Telescopes and Facility Instruments, budgeted from the Ministry of Finance of China (MOF) and administrated by the Chinese Academy of Sciences (CAS). The research is also supported by the National Natural Science Foundation of China (NSFC) under No. 11573054, 11703065, 11603044, and 11873081. HRAJ acknowledges support from a CAS PIFI and UK STFC grant ST/R006598/1. The research is also partly supported by the International partnership program's Key foreign cooperation project, Bureau of International Cooperation, Chinese Academy of Sciences 114A32KYSB20160049

# References

- Brown, T. M., Baliber, N., Bianco, F. B., Bowman, M., eet al. 2013, Publications of the Astronomical Society of the Pacific, 125, 1031B
- Chen, D., Wang, J. C., Xu, J., Luo, G. Q., Tan, H. S., Liu, Z., Zhang, B. R. 2003, Publications of the Yunnan Observatory, 95, 1-7.
- Cui. C. Z., Yu, C., Xiao, J., He, B. L., et al. 2015, Bulletin of the Chinese Academy of Sciences, 60(Z1), 445-449
- Hu, Y., Shang Z. H., et al, 2014, Publications of the Astronomical Society of the Pacific, 126, 868-881
- Hu, Y., Shang Z. H., et al, 2019, Publications of the Astronomical Society of the Pacific, 131, 015001
- Liu, L. Y., Yao, Y. Q., Wang, Y. P., Ma, J. L., He, B. L., & Wang, H. S. 2010, Research in Astronomy and Astrophysics, 10(10), 1061
- Liu, J., Zhang, Y., Feng, G., Bai, C. 2013, Proceedings of the International Astronomical Union, 9(S298), 427-427
- Sarazin, M., and Roddier, F., 1990, Astronomy and Astrophysics, 227, 294-300
- Schöck M., et al. 2009, PASP, 121, 384
- Skidmore, W., Riddle, R. Schöck M., et al. 2011, Revista Mexicana de Astronomía y Astrofísica (Serie de Conferencias), 41, 70-73
- Platnick, S., King, M. D., Ackerman, S. A., et al. 2003, IEEE Transactions on Geoscience and Remote Sensing, 41(2), 459-473
- Tian, J. F., Deng, L. C., Zhang, X. B., Lu, X. M., et al. 2016, Publications of the Astronomical Society of the Pacific, 128(968), 105003

This manuscript was prepared with the RAA  $IAT_EX$  macro v1.2.

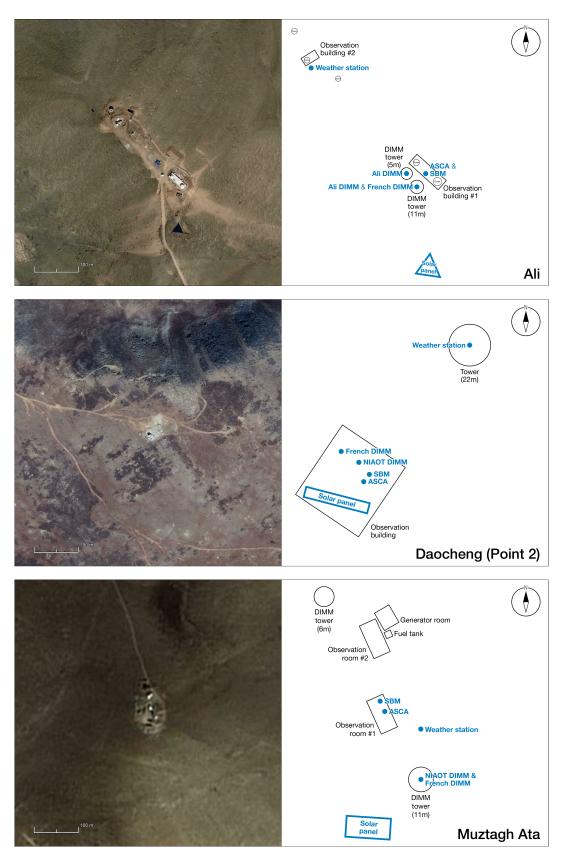


Fig. 2 The top row is the satellite map<sup>1</sup> of Ali Point A1 and the arrangement of the primary instruments. The middle row is the satellite map<sup>2</sup> of Daocheng Point 2 and the arrangement of the primary instruments. The bottom row is the satellite map<sup>3</sup> of Muztagh Ata and the arrangement of the primary instruments. The diagrams of the primary instruments on the right-hand side do not have an accurate scale.

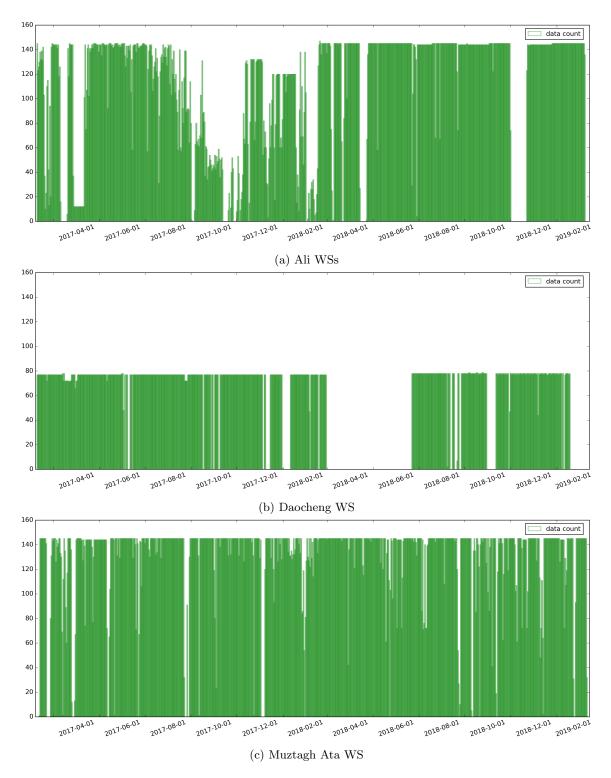
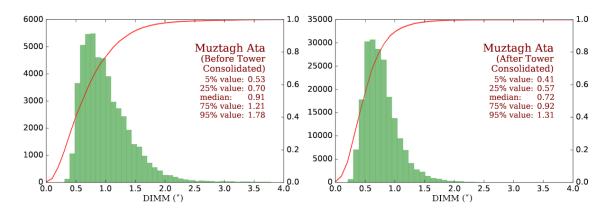
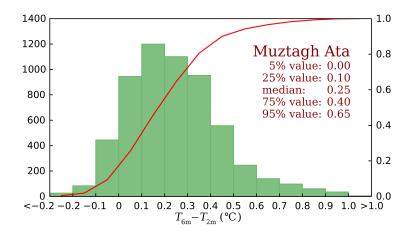


Fig. 3 The statistical data of Ali A1 and B4 WSs, whose measurement every ten minutes and can give a maximum of  $6 \times 24 = 144$  measurements per day. The statistical data for the Daocheng (Point 2) WS, whose measurement from 6:00 PM to 8:00 AM every ten minutes give the maximum of  $6 \times 13 = 78$  measurements per day. The statistical data of Muztagh Ata WS, whose measurement every ten minutes and a maximum of  $6 \times 24 = 144$  measurements per day.



**Fig. 4** A seeing comparison illustrating the impact of tower vibrations caused by wind. The left plot is seeing measured on top of the Muztagh Ata 11 m tower prior to reinforcement. The right plot is measured after the tower has been reinforced.



**Fig. 5** The histogram of temperature difference at 6 m and 2 m  $(T_{6m}-T_{2m})$  in Muztagh Ata from May 7, 2017 to June 4, 2017.

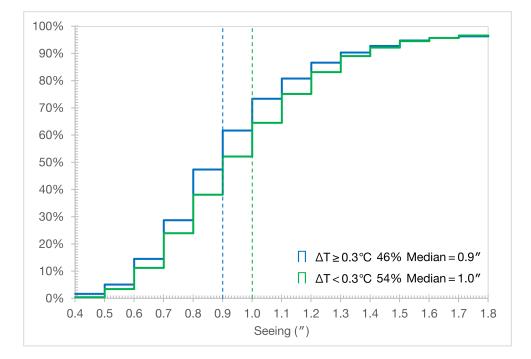


Fig. 6 The relationship between the temperature and the seeing measured at Muztagh Ata from May 7, 2017 to June 4, 2017. Different colors represent different growth curves corresponding to the  $T_{6m}-T_{2m}$  range and the curves were cut off at 1.8 arcsec seeing. 46% of seeing was measured during the period of  $T_{6m}-T_{2m} \ge 0.3^{\circ}$ C and 54% during the period of  $T_{6m}-T_{2m} < 0.3^{\circ}$ C. In order to divide seeing data in to two equal-size datasets, we selected the  $0.3^{\circ}$ C due to it being close to the mid-frequency point in Fig. 5. The seeing measurements were from the French DIMM deployed on the ground.

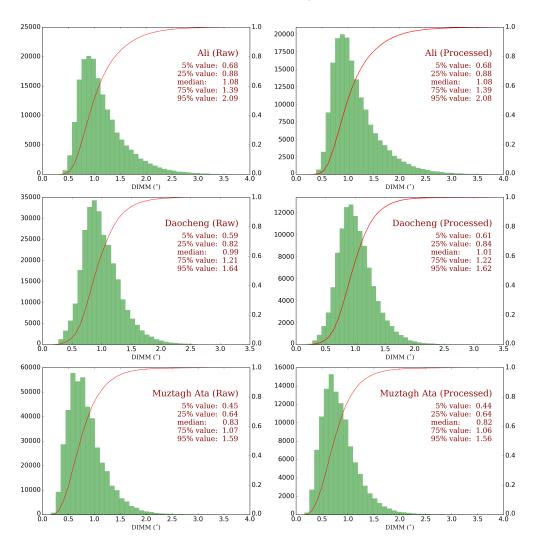


Fig. 7 The histograms of seeing before processing (left column) and after processing (right column) for Ali (top row), Daocheng (middle row), and Muztagh Ata (button row).

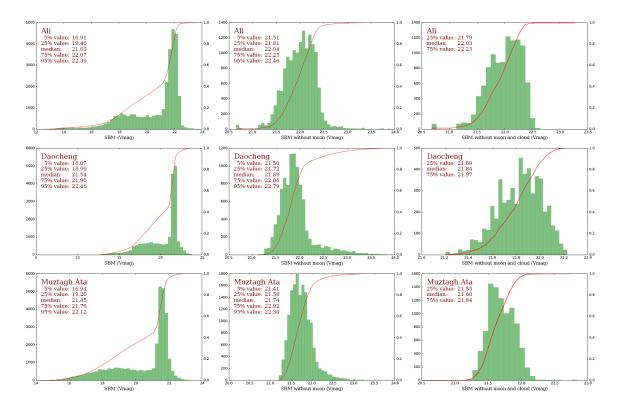
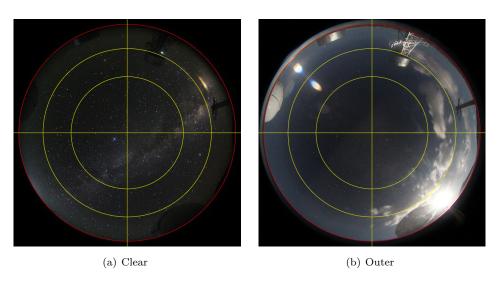
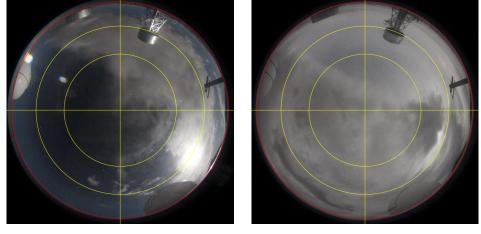


Fig. 8 The histograms compare original data, without moon and without moon+cloud for Ali, Daocheng, and Muztagh Ata at the top, middle, and bottom row respectively.

Site Testing Campaign for the Large Optical/Infrared Telescope of China





(c) Inner

(d) Covered

Fig. 9 Example ASCA images representing "clear" (a), "outer" (b), "inner" (c), and "covered" (d).

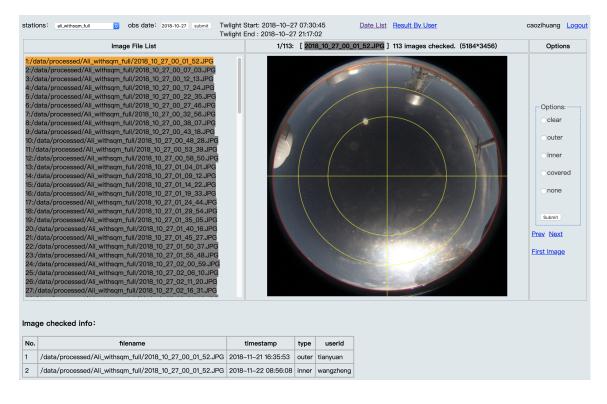
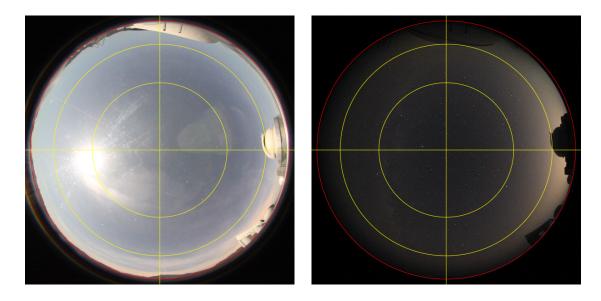


Fig. 10 ASCA image processing webpage with the image classification log at the bottom of the webpage, in which the moonlight caused a difference in cloud interpretation.



**Fig. 11** Example images of "inner" mixed with a bit cloudy in  $44.7^{\circ}$  circle (left) and full of thin cloud in the field of view (right).

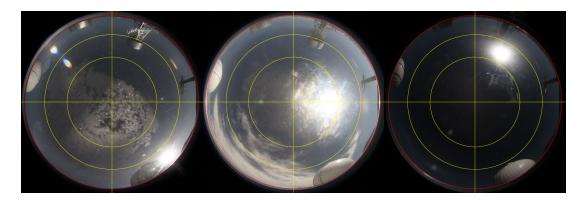


Fig. 12 Images with frost (left), glare (middle), and underexposure (right). Condensation and frosting in the moon-free nights are more difficult to judge than shown in this figure. When the moon altitude is high, the glare on the acrylic cover and lens is very close to the cloud. The ASCA exposure time is controlled by the SBM measurement results, so the moon near the north pole could result in underexposure.

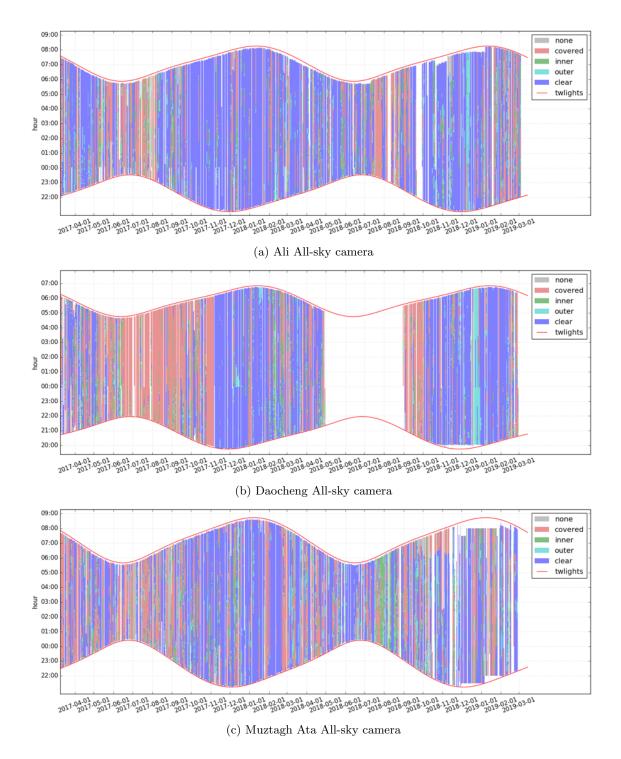


Fig. 13 The distribution of cloud cover obtained from ASCA in Ali, Daocheng, and Muztagh Ata from March 10, 2017 to March 10, 2019.

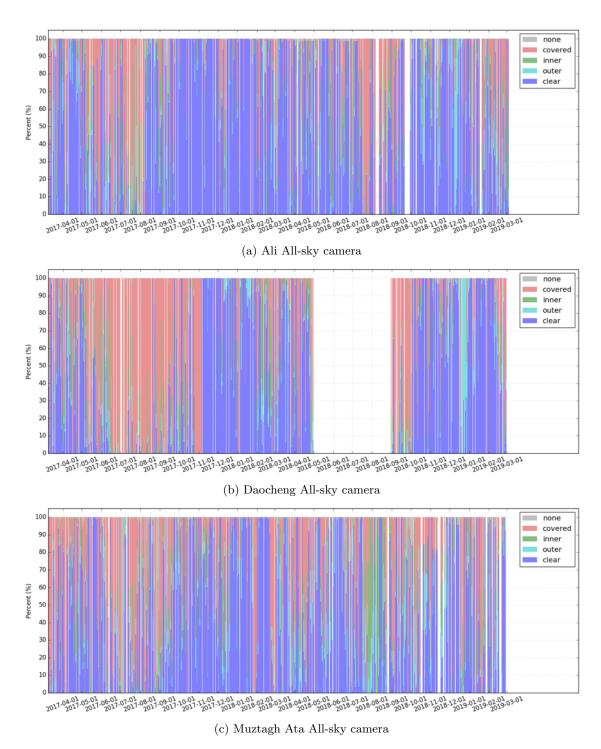


Fig. 14 The distribution of normalized cloud cover obtained from ASCA in Ali, Daocheng, and Muztagh Ata from March 10, 2017 to March 10, 2019.

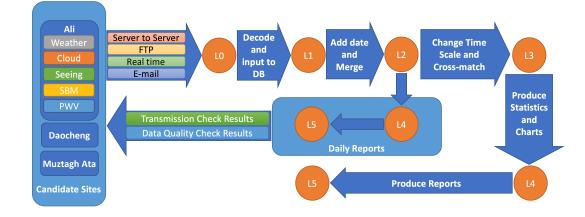
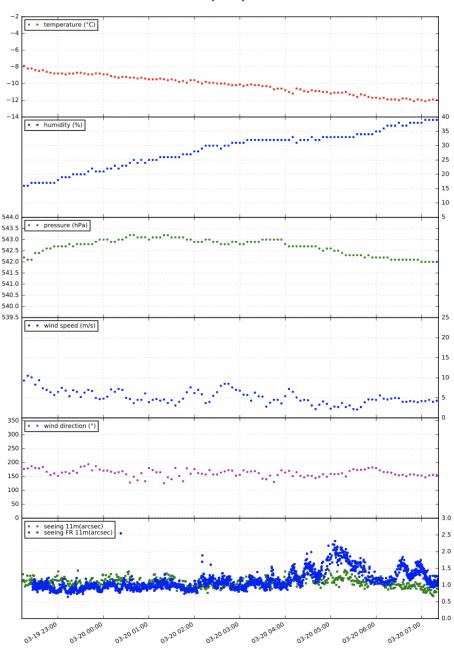


Fig. 15 The schematic diagram for data processing from transmission and initial data at each site.



Fig. 16  $\,$  A screen capture of the home page of the distribution and data quality webpage data.



WEATHER, SEEING 2018-03-19 ( Ali )

Fig. 17 An example of a daily report including from top to bottom plots of temperature, humidity, pressure, wind speed, and seeing with different DIMMs.

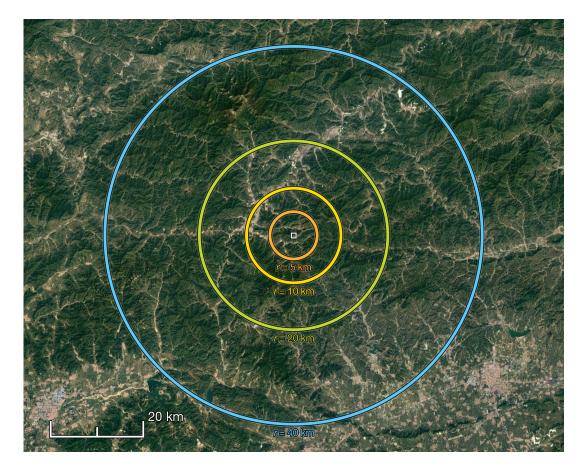


Fig. 18 The projected coverage of satellite and ASCA data at the Xinglong Observatory. The central white square is the satellite's single pixel (TERRA's  $1 \text{km} \times 1 \text{km}$  resolution). The low cloud (0–2 km) at the temperate region in a 44.7° FOV within the orange circle and 65° FOV within the yellow circle. The middle cloud (2–7 km) at in a 44.7° FOV around the orange circle and 65° FOV around the yellow circle. The high cloud (5–13 km) in a 44.7° FOV approximately between the orange and green circles and 65° FOV approximately between the green and blue circles.

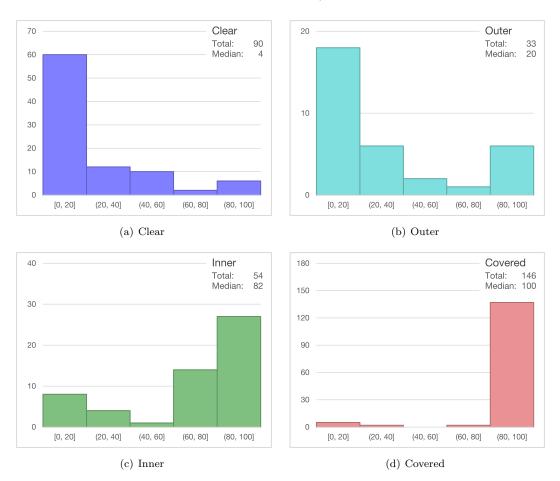


Fig. 19 For a particular ASCA classification at Xinglong at 10:30PM based on (a) "clear", (b) "outer", (c) "inner" and (d) "covered". The TERRA data is plotted as a histogram based on the percentage cloudiness. The figure caption records the number of available ASCA images (ASCA Total) and the median of the TERRA percentage cloudiness (TERRA Median).

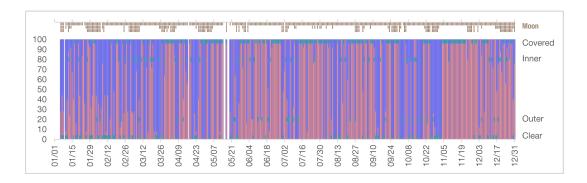


Fig. 20 Time series of the cloud cover from TERRA and ASCA at Xinglong. The impact of moonlight is shown at the top of the plot and is the combination of the moon phase and altitude, and shows the negative impact of moonlight on processors by three levels, marked by brown squares. The left-hand y-axis is the cloud percentage plotted in red from the TERRA satellite (or alternatively the clear-sky in blue). The right-hand y axis is the cloud cover as measured by ASCA and denoted by green squares.