Evaluation of fractional clear sky over potential astronomical sites

Shantikumar S. Ningombam,[1]⋆ H.-J. Song,[2]⋆ S. K. Mugil,[3] Umesh Chandra Dumka ®,[4]⋆
E. J. L. Larson,[5,6] Brijesh Kumar[4] and Ram Sagar[1]

[1]*Indian Institute of Astrophysics, Bangalore 560034, India*
[2]*National Institute of Meteorological Sciences, Seogwipo, Jeju 63568, South Korea*
[3]*Physics Department, St Joseph's College (Autonomous), Bangalore 560027, India*
[4]*Aryabhatta Research Institute of Observational Sciences (ARIES), Nainital 263 001, India*
[5]*CIRES, University of Colorado, Boulder, CO 80309, USA*
[6]*Chemical Sciences Laboratory, NOAA, Boulder, CO 80305, USA*

ABSTRACT

The estimation of the night-time cloud fraction (CF) is found to be one of the key parameters for evaluating the number of useful nights at an astronomical site. This work evaluates useful astronomical night-time observation over eight sites using a minimum threshold of CF from 21 years (2000–2020) of ground-based hourly visual and daily satellite data along with 41 years (1980–2020) of long-term hourly reanalysis data. The estimated number of photometric nights is underestimated by 8–24 per cent using the reanalysis data at Indian Astronomical Observatory-Hanle in comparison with the visual observations, while the estimated number of spectroscopic nights is 70–75 per cent per year and in good agreement with the visual observations. Among the astronomical sites, Paranal is found to be the best for astronomical observations, with 87 per cent spectroscopic nights per year. On the other hand, Hanle, Ali and Devasthal, located in the Himalayan region, exhibit an average of 68–78 per cent spectroscopic nights per year based on long-term reanalysis data, while Merak exhibits 61–68 per cent spectroscopic nights per year. Vertical profiles and global horizontal distributions for CF and related variable parameters are further compared among the sites. Global CF trends based on 41 years of reanalysis data show a decreasing tendency over most land regions and an increasing tendency over oceanic regions as well as over the Sahara desert, Middle East, and Indian subcontinent along the adjacent Tibetan Plateau. Such different CF trends between the ocean and land regions are thought to be the result of differential surface warming and water vapour changes associated with climate change.

Key words: atmospheric effects – site testing.

1 INTRODUCTION

The number of clear nights per year is one of the key characteristics of ground-based astronomical observatories operating at optical and near-infrared (NIR) wavelengths. A manual recording of the cloud coverage during the night has been a common practice at most observatories, although recently all-sky cameras and infrared radiometers have also been used. Ground-based observations with high temporal resolution provide an essential parameter in astronomical site characterization. However, such ground-based observational data have several limitations, mainly due to poor spatial coverage and short time periods. There is also the problem of the different criteria and methodologies used to define clear nights at different observatories located across the globe (Erasmus & van Rooyen 2006; Ningombam et al. 2020a). Moreover, such data are often not publicly available for analysis. In such cases, the use of model (reanalysis) or satellite data has many advantages because it provides a very large data coverage as well as homogeneous long-term consistency. Such a large homogeneous data set can

provide a better way to compare the cloud conditions at various potential astronomical sites across the globe. For satellite data, there are two kinds of data available: those from geostationary (e.g. *GOES*) and polar-orbiting (e.g. *Aqua*, *Terra*) satellites. Geostationary satellites orbit the Earth at an altitude of about 36 000 km above the equator with an orbital period of 24 h. Therefore, they appear to be stationary at a certain point above the equator, which enables observations with a high temporal resolution of several minutes (Thies & Bendix 2011). Polar-orbiting satellites, while still useful, have greater challenges quantifying or characterizing cloudiness over short temporal time-scales, as they take measurements over a given location just twice a day, as reported by Cavazzani, Zitelli & Ortolani (2015). Validation and comparative studies between the two types of satellites are described by Thies & Bendix (2011). Although the difference in cloud measurements between a geostationary satellite and a polar-orbiting satellite is small, as shown by Cavazzani et al. (2015) and references therein, the polar-orbiting satellite data is known to be more accurate because of the higher horizontal resolution.

The Indian Institute of Astrophysics (IIA), Bangalore, has carried out visual inspection of cloud data, meteorological parameters and other astronomical parameters since the 1990s, as part of an astronomical site survey program in the high-altitude Changthang

region of eastern Ladakh to set up astronomical facilities in the optical and NIR regions (Singh et al. 1989; HIROT Team 1996; Bhatt, Prabhu & Anupama 2000; Ye et al. 2016). After long-term astronomical site survey work had been carried out, the 2-m Himalayan Chandra Telescope (HCT), operating in the optical and NIR regions, was installed at the Indian Astronomical Observatory (IAO) Hanle ($32°46'46''$N, $78°57'52''$E; 4500 m amsl) in the year 2000. A visual inspection of hourly cloud data from IAO-Hanle has been performed since 1997 during the day and night (measured in octa, see below). The visual inspection of yearly cloud data reveals a consistency within 8–9 per cent with the number of clear nights given in the HCT observation log book (Ningombam et al. 2020a). Because of the promising nature of this site for ground-based optical and NIR astronomy, the IIA is also proposing a National Large Solar Telescope (NLST) at Merak, located about 118 km from Hanle and about 103 km from Leh town (aerial distance), the capital of Ladakh Union Territory. In addition, the IIA is also proposing a 8–10-m class National Large Optical Telescope (NLOT) at IAO-Hanle for night-time astronomical observation. Several other additional astronomical facilities have also been installed, and the details of these instruments and their scientific goals are described in Prabhu (2014). In this context, it is necessary to analyse the long-term (several decades) trends of cloud coverage, perceptible water vapour, and several other meteorological parameters over the high-altitude Ladakh region and compare them with trends from other similar astronomical observatories located across the globe using reanalysis and satellite data.

This paper is organized as follows: Section 2 describes the data sources, the locations of other observatories and the methodology to define and estimate photometric and spectroscopic nights from visual, satellite and reanalysis data; Section 3 presents monthly and yearly time series of photometric and spectroscopic nights; Section 4 describes the cloud fraction (CF) statistics during the satellite era (2000–2020), the long-term variation of CF during the reanalysis era (1980–2020), and vertical profiles; and Section 5 summarizes and discusses the implications of this work.

2 DATA AND METHODOLOGY

2.1 Data sources

This work makes use of CF, precipitable water vapour (PWV), surface temperature (ST), and relative humidity (RH) at 500 hPa from the global atmospheric reanalysis data of the Modern-Era Retrospective analysis for Research and Applications, version 2.0 (MERRA-2; Bosilovich, Lucchesi & Suarez 2016; Gelaro et al. 2017) and ERA5 (Hersbach et al. 2020) from 1980 to 2020. The MERRA-2 data is the latest atmospheric reanalysis of the modern satellite era produced by NASA's Global Modeling and Assimilation Office (GMAO) with a spatial resolution of 0.5° by 0.625° (∼50 km × 63 km on the Earth's surface) at 73 vertical levels, while ERA5 is the fifth generation reanalysis of the European Centre for Medium-Range Weather Forecasts (ECMWF) with a spatial resolution of 0.25° by 0.25° (∼25 km × 25 km on the Earth's surface).[1] Satellite data from the Moderate resolution Imaging Spectrometer (MODIS), onboard NASA's *Terra* launched on 1999 December 18, are also used in this work. The MODIS instruments provide data of the entire Earth's surface every 1 to 2 d in 36 spectral bands ranging from $\lambda = 0.4$ to

Table 1. Locations of the eight high-altitude astronomical observatories considered in this study.

Site	Elevation	Latitude	Longitude
SALT, Sutherland, South Africa	1798 m	32.38°S	20.81°E
Devasthal, Nainital, India	2450 m	29.36°N	79.68°E
Paranal, Chile	2635 m	24.63°S	70.40°W
NAO, Mexico	2800 m	31.04°N	115.46°W
Merak (NLST), Ladakh, India	4310 m	33.80°N	78.62°E
Hanle (IAO), Ladakh, India	4500 m	32.78°N	78.96°E
Ali, Shiquanhe Observatory, China	5100 m	32.63°N	80.00°E
TAO, Chile	5640 m	22.98°S	67.74°W

14.4 μm in the infrared region with a horizontal resolution of 1 km.[2] The MODIS data used in the present work are taken from MODATML2 (https://doi.org/10.5067/MODIS/MODATML2.061) and MOD08_M3 (https://doi.org/10.5067/MODIS/MOD08_M3.061). In this work, we have also used hourly visual cloud data expressed in octa (on a scale of 0 to 8), which are available only from IAO-Hanle during 2000–2020. An octa is defined as 1/8 of the sky covered by clouds, such that 8 octa indicates a completely overcast sky, and 0 octa indicates a completely clear sky.

2.2 Location of astronomical sites

Table 1 lists the locations of the eight high-altitude (> 1.7 km) astronomical observatories analysed in this work. The study region spreads over the globe, namely at IAO-Hanle, Merak, and Devasthal in India; Shiquanhe Observatory (Ali) in the Tibet Autonomous region in China; the South African Large Telescope (SALT) in Sutherland, South Africa; the University of Tokyo Atacama Observatory (TAO), Cerro Chajnantor and Paranal in Chile; and the National Astronomical Observatory (NAO) in Mexico, as shown in Table 1 and Fig. 1. However, we have not considered several other high-altitude oceanic sites where CF data are not significant owing to an inability to capture the CF precisely above the station.

2.3 Methodology

Methodologies of classifying the photometric and spectroscopic nights varied from an observatory to another (Erasmus & van Rooyen 2006; Zhang et al. 2015; Ningombam et al. 2020a). However, the methodologies of classifying photometric and spectroscopic nights have been practicing at IAO-Hanle since the late 1990s under the following conditions (HIROT Team 1996):

(i) Clear night: cloud cover is 0 octa for the entire night.

(ii) Partly clear night: clouds are seen at night, but the cloud cover is 0 octa for 4 or more hours continuously.

(iii) Partly cloudy night: cloud coverage is less than or equal to 3 octas for 4 or more hours in not more than two stretches.

(iv) Photometric night: a night is termed photometric if it satisfies (i) and (ii).

(v) Spectroscopic nights: a night is termed spectroscopic if it satisfies (i), (ii) and (iii).

(vi) Usable night: a night is termed usable if it satisfies (v).

(vii) Cloudy night: a night is termed cloudy if it does not satisfy (v).

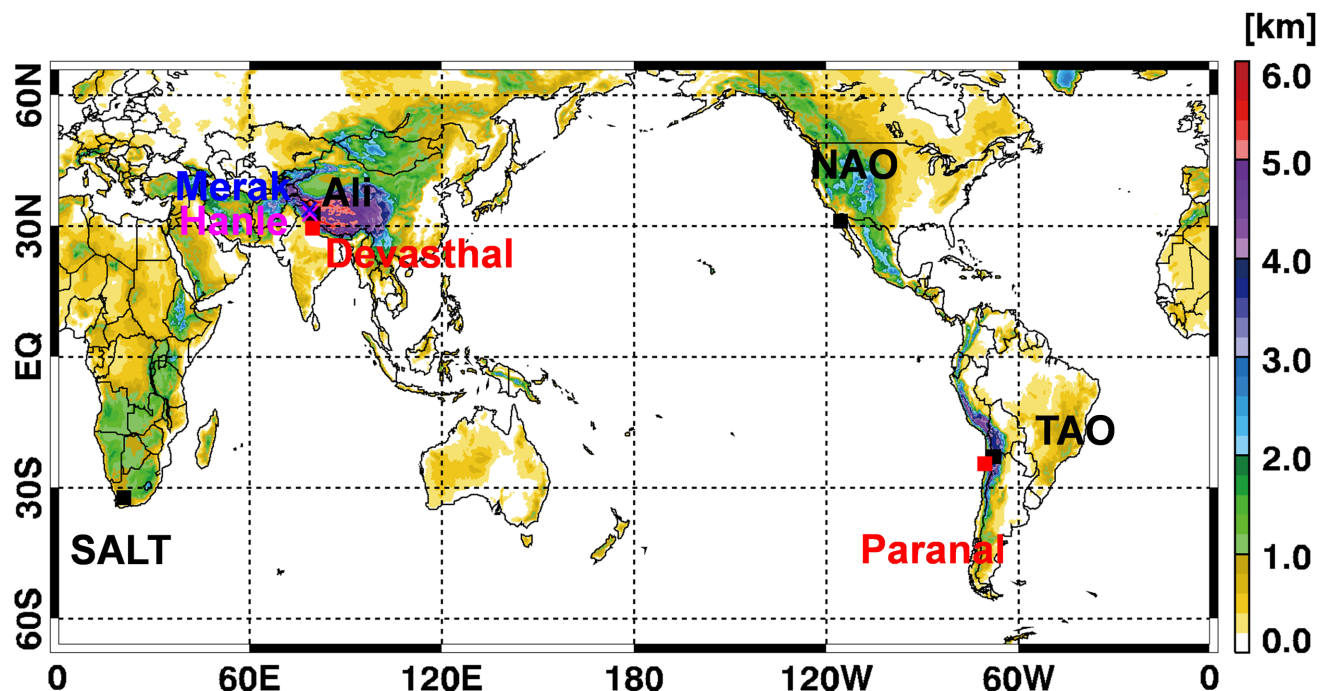**Figure 1.** The eight high-altitude astronomical observatories/sites considered in this study.

Table 2. Results of the root mean square deviation (RMSD), mean bias difference (MBD), absolute percentage difference (APD, %) and correlation coefficient (R) between the visual and reanalysis (ERA5 and MERRA-2) data, along with their monthly and yearly statistics of photometric and spectroscopic nights from the visual and the reanalysis data during 2000–2020.

Parameter	RMSD	MBD	APD	R	Monthly	Yearly
Photometric						
ERA5	− 4.25	6.78	33.96	0.69	13 ± 7	155 ± 21
MERRA-2	− 1.44	4.94	27.38	0.73	16 ± 7	188 ± 29
Visual	–	–	–	–	17 ± 6	205 ± 19
Spectroscopic						
ERA5	0.06	4.49	19.43	0.71	21 ± 6	257 ± 18
MERRA-2	1.42	4.85	21.79	0.62	23 ± 5	273 ± 21
Visual	–	–	–	–	21 ± 5	256 ± 14

The CFs estimated from MERRA-2 and ERA5 data are represented from 0 for a completely clear sky to ∼1 for a completely overcast sky. Following Erasmus & van Rooyen (2006), a CF ≤ 0.125 and CF > 0.125 but < 0.5 are considered as photometric and spectroscopic nights, respectively. However, in this work we modified the condition for a spectroscopic night to CF < 0.5, which also includes photometric nights (HIROT Team 1996; Ningombam et al. 2020a), while the the condition for a photometric night remains the same, namely with CF ≤ 0.125. In the case of high-resolution (hourly) reanalysis (MERRA-2 and ERA5) data, we defined a photometric night as a CF with ≤ 0.125 for 4 consecutive hours or more, and a spectroscopic night as CF < 0.5 for ≥ 4 hours in not more than two stretches.

In the data processing, we used a uniform rectangular grid of 0.5° × 0.5° from the centre of each station's latitude and longitude as listed in Table 1. We used 21 yr (from 2000 January to 2020 December) of high-resolution night-time (reanalysis: hourly; satellite: daily) data.

The duration of a night at a site is counted from an hour after sunset to an hour before sunrise. It is similar to the duration from evening twilight to morning twilight (Ningombam et al. 2020a). Furthermore, in order to study the long-term night-time changes of the CF, PWV, ST and RH values, we used daily and monthly night-time data from the reanalysis from 1980 to 2020 and from MODIS (*Terra*) from 2000 February to 2020 December.

In addition, the statistical technique of linear trend analysis is used in this work to study the trend in certain variables (Sohn et al. 2013, 2016; Yim et al. 2017; Song & Sohn 2020; Lee, Song & Sohn 2021). For monthly data, the linear trend is defined as the slope of the linear regression with time (1 to 492 months) as a predictor (or independent) variable and monthly anomalies for certain data (e.g. ΔCF) as a dependent variable. The linear equation can therefore be expressed as

$$\Delta CF = time \times slope + offset. \tag{1}$$

Here, ΔCF is defined as the difference between the original value and the climatological monthly mean value in order to remove the effect of the seasonal component on the linear trend. Consequently, the linear trend is obtained in units of per cent month^{-1}. It is converted into per cent decade^{-1} by multiplying by 120 months (10 years = 1 decade). In this study, the linear trend analysis is applied to all grid boxes over the globe using MODIS, ERA5, and MERRA-2 data. For example, the total number of grid boxes for 60°N–60°S is 43 200 (360 × 120) for the MODIS data with a horizontal resolution of 1° × 1°. The resulting positive and negative linear trends represent the increasing and decreasing tendency, respectively, for a certain grid box in terms of the long-term scale (i.e. 41 yr). The statistical confidence of the linear trend can be obtained by Student's *t*-test. For a regression problem, the *t*-value is calculated by

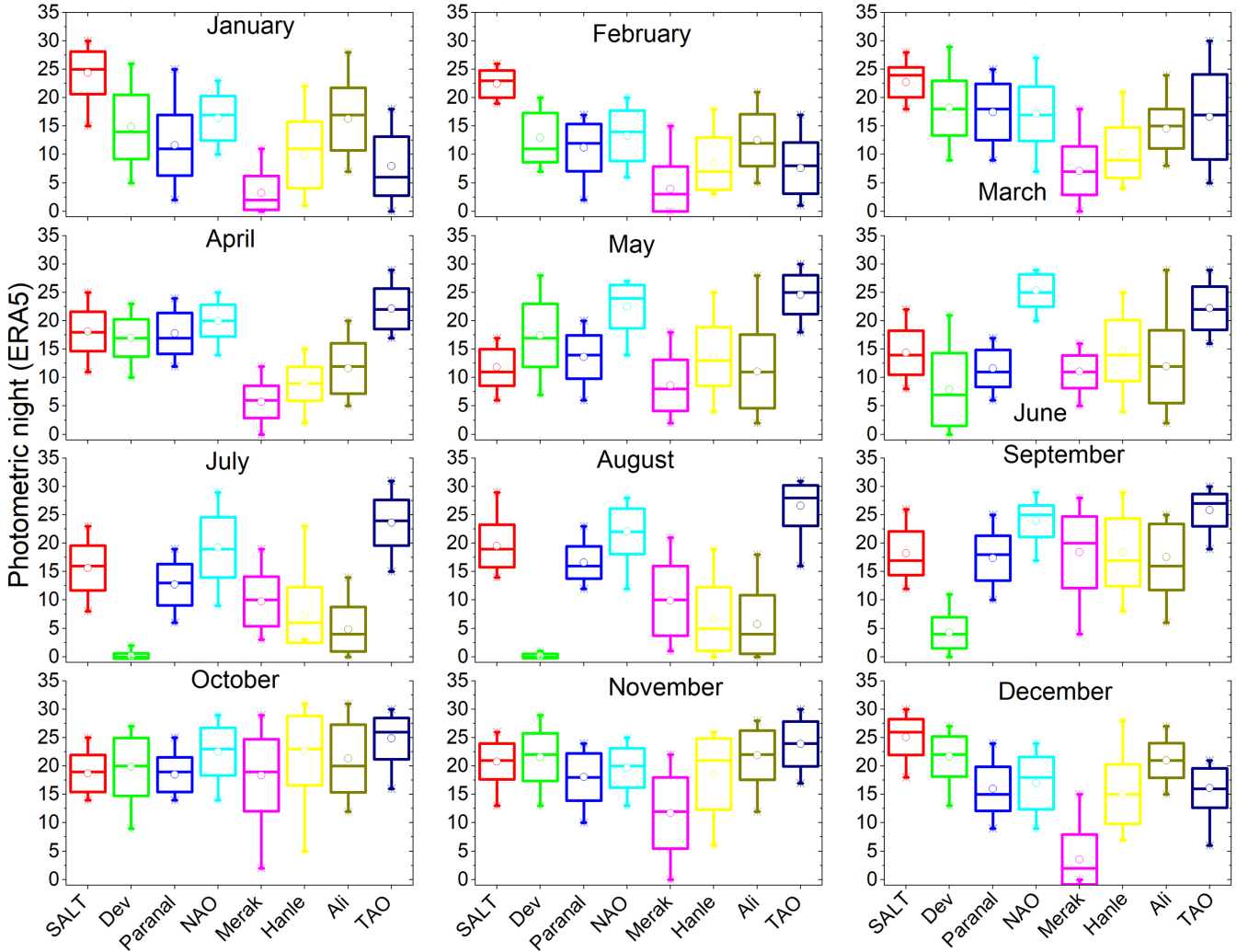$$t = R \frac{\sqrt{n-2}}{\sqrt{1-R^2}}. \tag{2}$$

Figure 2. Box and whisker plots for the climatological photometric nights of ERA5 during 2000–2020 at the eight astronomical sites. The central horizontal lines inside the boxes are medians; open circles inside the boxes are arithmetic means; bottoms and tops of the boxes are ± standard deviation; and the ends of the whiskers are the minimum and maximum values. The eight colours of the boxes represent the eight different sites considered in this study.

Here, n is the number of data (i.e. 492 months), and R is the Pearson correlation coefficient in the linear trend analysis (i.e. between time and ΔCF). When the slope and correlation coefficient of the linear regression are negative, then the t-value becomes negative. The linear trend at a given location is statistically significant at the 95 per cent and 90 per cent confidence levels if the corresponding absolute values of t are 1.965 and 1.648, respectively. As a result, higher positive or negative t-values indicate statistical robustness for the estimated linear trend, whereas t-values close to zero are not statistically significant.

2.4 Comparison of visual and reanalysis data

We compared the monthly visual observations of the CF at IAO-Hanle with the ERA5 and MERRA-2 (reanalysis) data from 2000 to 2020. The details of the statistics are presented in Table 2. Analyses of ERA5 and MERRA-2 data yield 155 (41 per cent) to 188 (52 per cent) photometric and 257 (70 per cent) to 273 (75 per cent) spectroscopic nights in a year, corresponding to 13 to 16 photometric and 21 to 23 spectroscopic nights per month. The number of yearly spectroscopic nights agrees well with the visual

data of 256 (70 per cent) at IAO-Hanle. Furthermore, these results are consistent with the earlier results of 254 spectroscopic nights in a year during the period 1997–2018 (Ningombam et al. 2020a). The visual data yield 205 photometric nights (56 per cent) per year at IAO-Hanle. These numbers derived from the reanalysis data are underestimated by 24 per cent for ERA5 and by 8 per cent for MERRA-2. The underestimation of photometric nights may be due to the different thresholds of CF adopted in the visual and reanalysis data (i.e. 0 octa versus 0.125 CF), as described in the Section 2.3.

Our aim is to quantify the difference between the visual and reanalysis data so that homogeneous and long-term (reanalysis and satellite) data with high temporal resolution (hourly) can be used to compare the number of useful nights at high-altitude astronomical sites located across the globe. To do this, we compared the visual and reanalysis data using the root mean square deviation (RMSD), mean bias difference (MBD), and absolute percentage difference (APD) of photometric and spectroscopic from a total of 252 months of data as:

$$RMSD = \left(\sum_{i=1}^{N} (X_i - M_i)^2 \times \frac{1}{N} \right)^{\frac{1}{2}}, \tag{3}$$
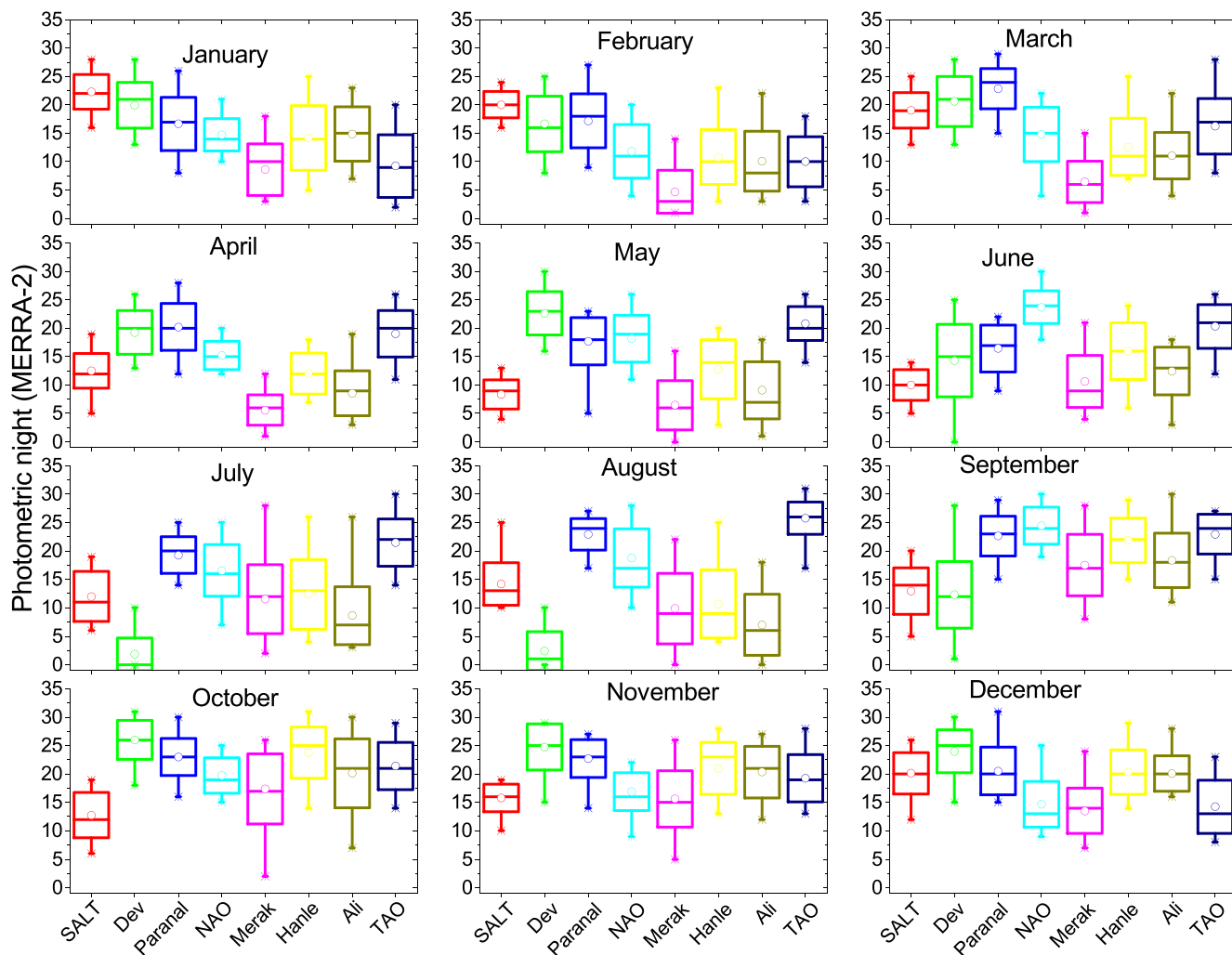
Figure 3. Box and whisker plots for climatological photometric nights of MERRA-2 during 2000–2020 at the eight astronomical sites considered in this study. The central horizontal lines inside the boxes are medians; open circles inside the boxes are arithmetic means; bottoms and tops of the boxes are ± standard deviation; and ends of the whiskers are the minimum and maximum values.

$$MBD = \left(\sum_{i=1}^{N} (X_i - M_i) \times \frac{1}{N} \right), \quad (4)$$

$$APD = 100 \times Abs \left(\frac{X_i - M_i}{M_i} \right). \quad (5)$$

Here, X_i and M_i represent the individual monthly data from the model (ERA5 and MERRA-2) and visual data, respectively, with $N = 252$. The estimated RMSD, MBD, APD and correlation coefficient (R) between the visual and reanalysis (ERA5 and MERRA-2) data are given in Table 2. The visual and reanalysis data for photometric and spectroscopic nights are more or less similar, with R varying from 0.69 to 0.73 and from 0.62 to 0.72 for photometric and spectroscopic nights, respectively. However, on climatological analysis, R varied from 0.76 to 0.85 during winter (December-January-February) as the best season, and from 0.24 to 0.58 during autumn (September-October-November) as the worst. For the spectroscopic nights, RMSD and MBD values derived from reanalysis data are found to be lower than the photometric nights. In contrast to the case for photometric nights, the number of spectroscopic nights derived from reanalysis data agreed very well with the number obtained from visual data, as shown in the Table 2.

3 PHOTOMETRIC AND SPECTROSCOPIC NIGHTS

3.1 Monthly statistics

The numbers of monthly photometric and spectroscopic nights estimated from long-term, high-resolution satellite and visual data are analysed at the various astronomical sites considered in this study. Figs 2 and 3 show the climatological number of monthly photometric nights estimated from ERA5 and MERRA-2 satellite data, respectively, while the corresponding numbers of spectroscopic nights are displayed in Figs 4 and 5, respectively. These numbers are the data obtained during the period 2000–2020. The linear correlations present between ERA5 and MERRA-2 monthly time-series photometric and spectroscopic nights at the locations considered in this study are shown in Table 3, along with the estimated average number of monthly photometric and spectroscopic nights. At IAO-Hanle, the monthly numbers of 13 to 16 photometric and 21 to 23 spectroscopic nights estimated from the reanalysis data agree fairly well with the corresponding numbers of 17 ± 6 and 21 ± 5, respectively, reported by Ningombam et al. (2020a) using visual data during the years 1997–2018. Among the sites, SALT, Paranal, NAO
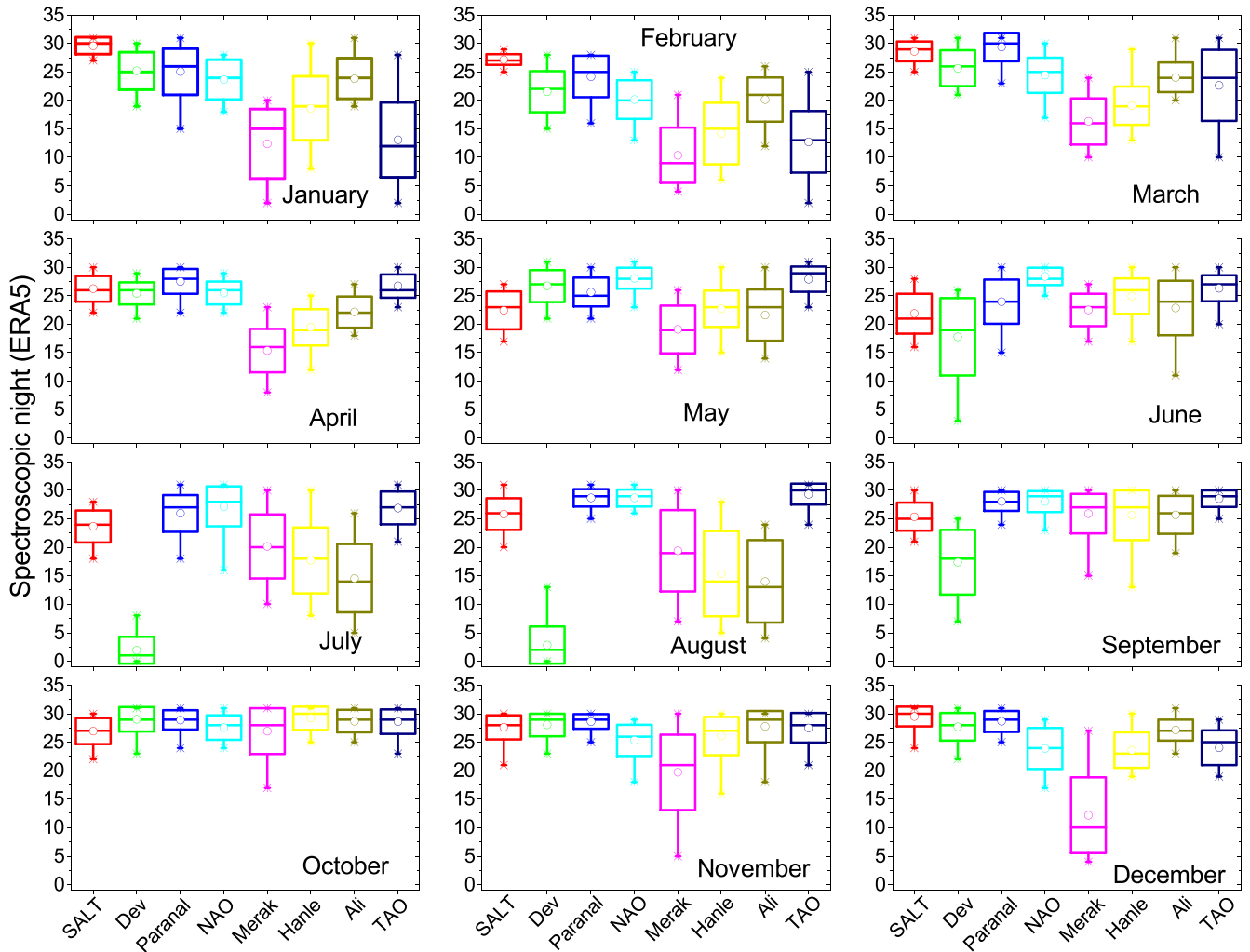
Figure 4. Box and whisker plots for climatological spectroscopic nights of ERA5 during 2000–2020 at the eight astronomical sites considered in this study. The central horizontal lines inside the boxes are medians; open circles inside the boxes are arithmetic means; bottoms and tops of the boxes are ± standard deviation; and ends of the whiskers are the minimum and maximum values.

and TAO, located in different parts of the globe, exhibit the highest number of 15–20 photometric nights per month, whereas Hanle, Ali and Devasthal, located in the Himalayan region, exhibit similar numbers of 13–17 photometric nights per month. Merak, which is also located in the region, exhibits only 9–11 photometric nights per month. The highest number of monthly spectroscopic nights is 26 to 27 at Paranal. There are 24 to 26 spectroscopic nights per month at SALT, NAO and TAO. This number is 21 to 24 at the Hanle, Ali and Devasthal sites located in the Himalayan region, while at Merak it is 18 to 21. Amongst the Himalayan sites, only Devasthal is affected by monsoon clouds, mostly during July and August (Sagar et al. 2000; Sagar, Kumar & Omar 2019), while other sites of this region are not affected by monsoon clouds as they are located in the rain-shadow of the Himalayas.

3.2 Yearly statistics

This work examined the yearly number of photometric and spectroscopic nights using daily CF data from MODIS (*Terra*) at three spatial resolutions, namely $0.2° \times 0.2°$, $0.1° \times 0.1°$ and $0.05° \times 0.05°$ during the years 2000–2020, along with hourly reanalysis data

from ERA5 ($0.25° \times 0.25°$) and MERRA-2 ($0.5° \times 0.625°$) during the years 1980–2020. Fig. 6(a,b) shows a bar chart of the long-term yearly averaged photometric and spectroscopic nights from the MODIS, MERRA-2 and ERA5 data at the eight astronomical sites under study. High-temporal-resolution data from ERA5 and MERRA-2 were combined to improve the statistics for photometric and spectroscopic nights. Figs 7 and 8 show box and whisker plots for photometric and spectroscopic nights during 2000–2020 for the eight astronomical sites. The combined data of ERA5 and MERRA-2 have 172 (47 per cent) photometric nights per year, against 205 (56 per cent) nights per year from the visual data at Hanle-IAO. The underestimation of photometric nights seems to be due to the different threshold of CF adopted between the visual and reanalysis data, as discussed earlier. At IAO-Hanle, the numbers of photometric and spectroscopic nights per year are 205 and 256 respectively using visual data during 2000–2020, while there are only 155–188 photometric and 257–273 spectroscopic nights per year using the reanalysis data. However, the spectroscopic nights obtained from the reanalysis data at IAO-Hanle are consistent with the visual data (see Table 2). This could be due to the similar methodology adopted for the different data sources. There are 257 to 273 (or 70 to 75 per cent) spectroscopic nights per year using reanalysis and visual data at
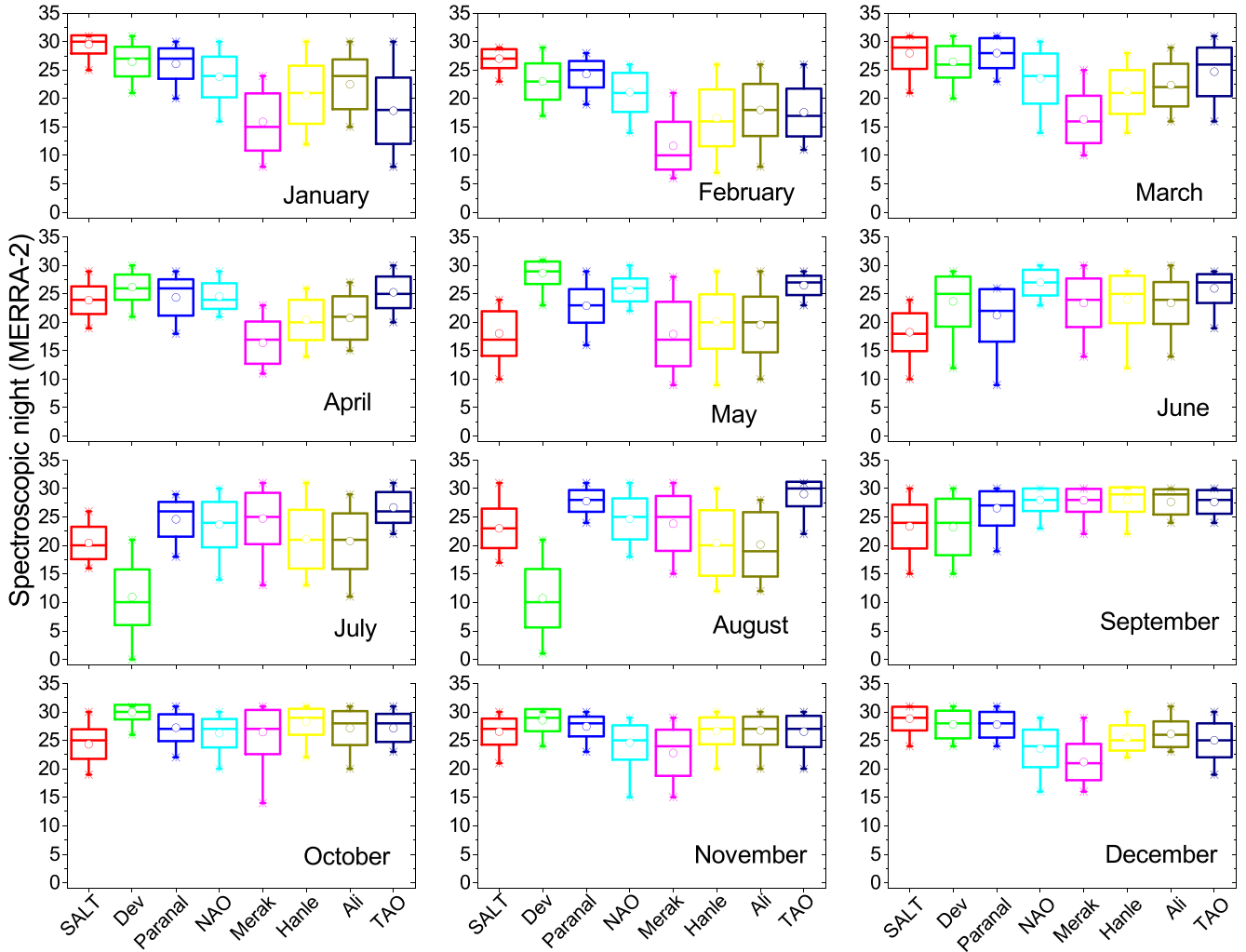
Figure 5. Box and whisker plots for climatological spectroscopic nights from hourly data of MERRA-2 during 2000–2020 at the eight astronomical sites considered in this study. The central horizontal lines inside the boxes are medians; open circles inside the boxes are arithmetic means; bottoms and tops of the boxes are ± standard deviation; and ends of the whiskers are the minimum and maximum values.

Table 3. Slope and correlation coefficient (*R*) between ERA5 (X-axis) and MERRA-2 (Y-axis), and monthly statistics of photometric and spectroscopic nights from monthly time-series ERA5 and MERRA-2 data during 2000–2020 at the eight astronomical sites.

Sites	Photometric				Spectroscopic			
	Slope	*R*	ERA5	MERRA-2	Slope	*R*	ERA5	MERRA-2
SALT	0.88 ± 0.04	0.84	19 ± 5	15 ± 5	1.16 ± 0.05	0.85	26 ± 3	24 ± 5
Devasthal	0.89 ± 0.03	0.88	13 ± 9	17 ± 9	0.67 ± 0.02	0.93	21 ± 9	24 ± 7
Paranal	0.62 ± 0.05	0.65	15 ± 5	20 ± 5	0.78 ± 0.05	0.73	27 ± 3	26 ± 4
NAO	0.87 ± 0.03	0.86	20 ± 5	17 ± 5	0.73 ± 0.04	0.74	26 ± 4	25 ± 4
Merak	0.66 ± 0.04	0.70	9 ± 7	11 ± 6	0.67 ± 0.04	0.75	18 ± 7	21 ± 6
Hanle	0.72 ± 0.04	0.78	13 ± 7	16 ± 7	0.67 ± 0.03	0.78	21 ± 6	23 ± 5
Ali	0.78 ± 0.03	0.85	14 ± 7	13 ± 7	0.60 ± 0.04	0.72	23 ± 6	23 ± 5
TAO	0.72 ± 0.03	0.86	20 ± 8	18 ± 6	0.63 ± 0.02	0.88	25 ± 7	25 ± 5

IAO-Hanle. There are 291–325 (or 80 to 89 per cent) spectroscopic nights per year at the sites located in the Atacama desert (TAO, Paranal), SALT, and NAO (see Fig. 6), and these results are consistent with earlier published work. For example, there are about 85–88 per cent clear nights in a year at Paranal as reported by Cavazzani

et al. (2011). Similarly, there are 63.1 per cent photometric and 80.8 per cent spectroscopic nights at NAO, Mexico during 1984 January to 2002 December as reported by Tapia (2003).

Among the Himalayan sites, Devasthal has 180 (49 per cent) photometric and 267 (73 per cent) spectroscopic nights in a year, and
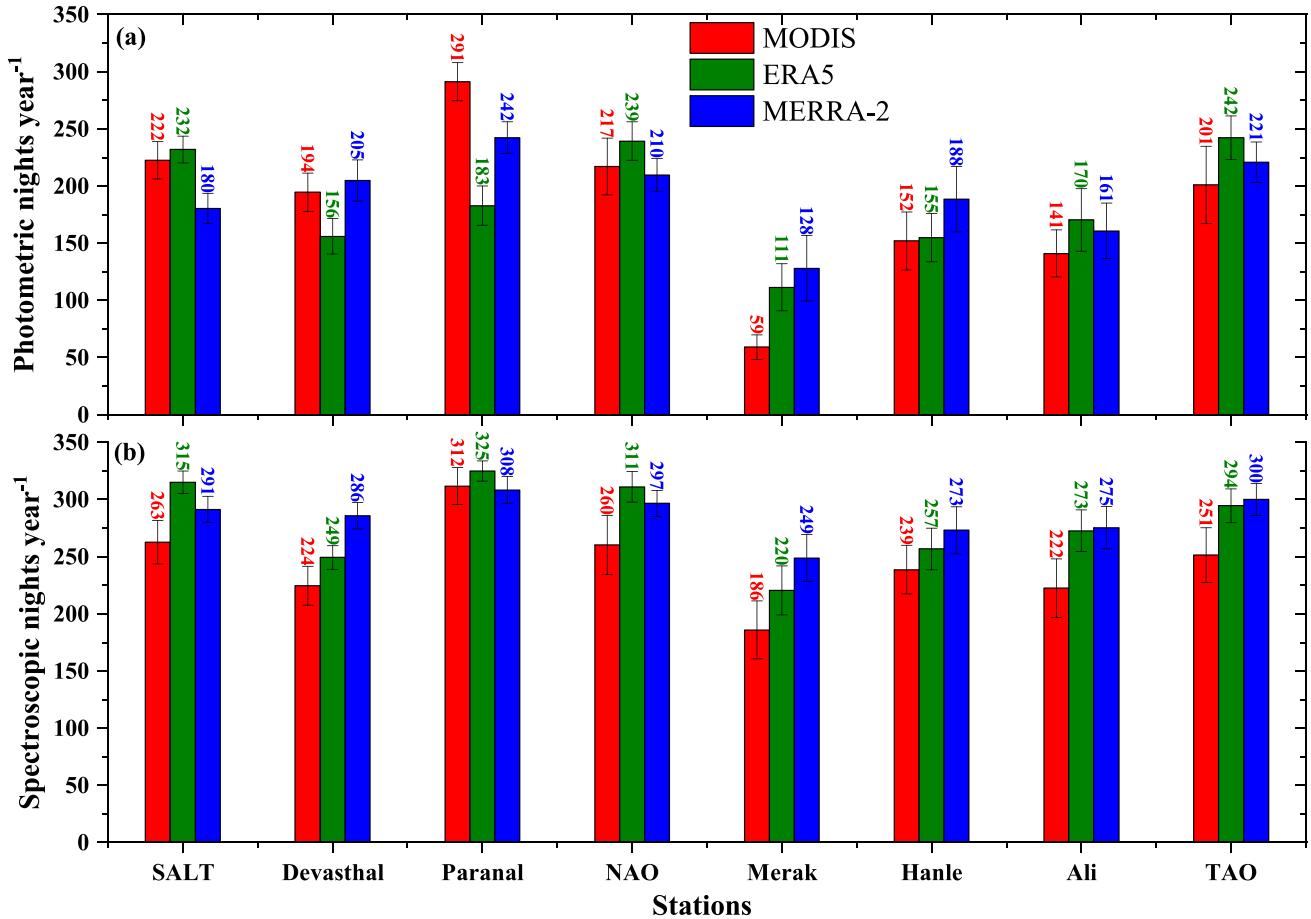
Figure 6. Bar chart of yearly photometric and spectroscopic nights estimated from hourly data of ERA5 and MERRA-2, and daily data from MODIS (*Terra*) during 2000–2020 at the eight astronomical sites. The numerical values by different colours (red, MODIS; green, ERA5; blue, MERRA-2) in (a) and (b) show the average total available nights for photometric and spectroscopic observations at each site.

these results are consistent with the numbers reported earlier (Sagar et al. 2000), despite the region being affected by monsoon clouds, unlike Hanle, Merak and Ali. Merak exhibits 120 (33 per cent) photometric and 235 (64 per cent) spectroscopic nights, while Ali exhibits 166 (45 per cent) photometric and 274 (75 per cent) spectroscopic nights in a year, as seen in Figs 7 and 8. Similar to the present work, Ye et al. (2016) reported about 54 per cent photometric and 70 per cent spectroscopic nights at IAO-Hanle, using the Climate Forecast System Reanalysis (CFSR) data for 31 years. They also examined potential sites in Cerro Chajnantor and Cerro Tolonchar in the Atacama desert, as well as in the wide area of the high-altitude Tibetan Plateau including Ali Observatory. They claimed 57, 54 and 62 per cent photometric and 75, 70 and 79 per cent spectroscopic nights at Ali, Cerro Chajnantor and Cerro Tolonchar, respectively. These results on spectroscopic nights are consistent with those in this work reported at Ali (72 per cent), Hanle (70 per cent), and TAO (81 per cent, Cerro Chajnantor). However, the lower number of photometric nights from the reanalysis data in this work may be attributed to (i) a different grid area, (ii) a different minimum threshold of CF, and (iii) adopting a minimum of a continuous 4 hours or more. Furthermore, owing to poor spatial resolution, unable to capture the highly variable moisture parameters within the time constraints adopted in this work, as seen in the Fig. 7.

4 CLOUD FRACTION STATISTICS

Using the last 21 years (2000–2020) of data from satellite (MODIS, *Terra*) and reanalysis (ERA5 and MERRA-2), we have plotted the climatological mean of night-time CF in Fig. 9. It shows that the climatological mean distribution of CF derived from ERA5 and MERRA-2 reanalysis data is similar to the distribution obtained from MODIS, despite their different temporal resolutions. The CF values are small around the Sahara desert region, and large values are distributed over a wide range of high-latitude oceanic areas. Fig. 9 also indicates regions that are good for night-time astronomical observations from the lower mean CF values. However, both of the reanalysis data sets show a tendency to underestimate CF values in oceanic regions. More evident differences are found in supplementary Fig. S1(a–d), which shows the bias and RMSE of the reanalysis data from MODIS during the period under consideration (i.e. 2000 February to 2020 December). The CF value of the reanalysis data was derived from a numerical model based on data assimilation from various observational data sets and parameters (Gelaro et al. 2017; Hersbach et al. 2020). However, the data used in the numerical model over the oceanic regions are less reliable than those over land, especially for cloud calculations, where there are difficulties in using the data assimilation compared to a clear sky, and thus the uncertainty of CF in reanalysis data over oceans is thought to be high. The underestimation of CF is more pronounced for
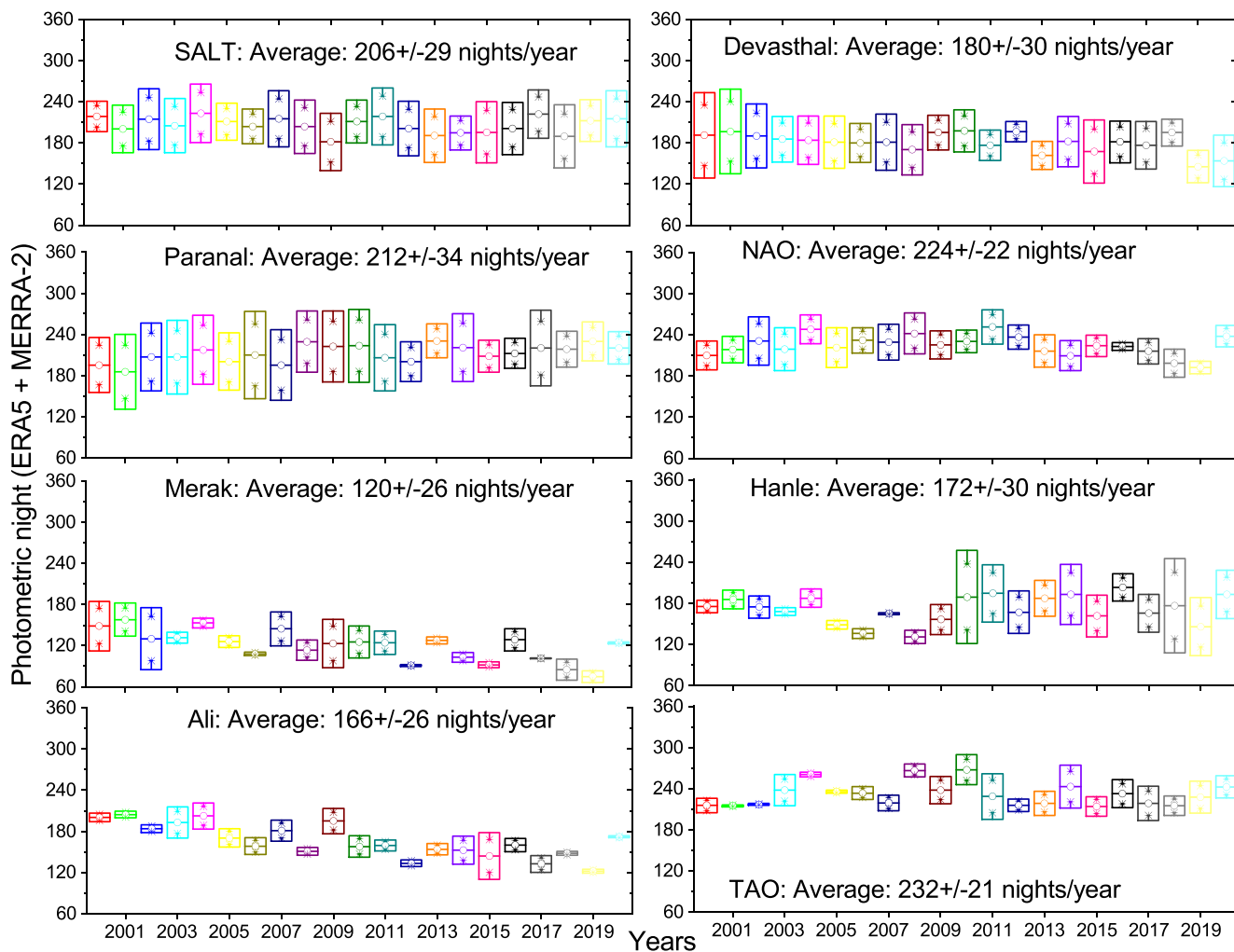
Figure 7. Box and whisker plots for yearly time-series photometric nights of ERA5 and MERRA-2 (combined) data during 2000–2020 at the eight astronomical sites. The central horizontal lines inside the boxes are medians; open circles inside the boxes are arithmetic means; and bottoms and tops of the boxes are the minimum and maximum values. The 21 different colours in each panel shows 21 different years of data, and the numerical values at each site in each panel shows the average with the standard deviation from the combined data.

MERRA-2, with a 40 per cent larger mean bias than for ERA5 (Fig. S1a,c). The RMSE of monthly CF for ERA5 was 18 per cent smaller than for MERRA-2 (Fig. S1b,d). The superior accuracy of CF for ERA5 over MERRA-2 is thought to result from the higher horizontal resolution associated with realistic terrain consideration (0.25° versus 0.5°–0.625°), and from adopting more advanced data assimilation as well as parametrization processes.

In order to investigate the long-term trend of CF, the existing observations of a 41-yr period (1980–2020) from both reanalysis data sets were further analysed, taking into consideration the fact that satellite data obtained before 2000 are less accurate. However, long-term results can be compared for the global locations listed in Table 1. The mean CFs shown in Fig. 9 from ERA5, MERRA-2 and MODIS data are at spatial resolutions of 0.25°, 0.5°–0.625°, and 1°, respectively. Among the sites under investigation, Paranal is found to be the best, with CF values of 8.4, 7.5, and 7.6 per cent for the three spatial regimes. The CF value of the Paranal site is also consistent with ERA5 and MERRA-2 (18 per cent and 24 per cent) reanalysis data for the last 41 years, although there is a difference in the analysis

period (21 versus 41 years). Furthermore, these results of low CF are consistent with work reported by several authors. For example, there are good clear fractions of nights over Aklim (Morocco), ORM (La Palma), Ventarrones (near Paranal) and Macon (Argentina) of 76, 84, 85 and 75 per cent, respectively, using more than 5 years of satellite data (Vernin et al. 2011). In fact, the Paranal site is located towards the coast, and there is a huge difference in CFs between the surrounding ocean and land regions. Therefore, different horizontal resolutions between reanalysis and satellite observations could be one possible reason for the difference of CF mean climatology (for example, a low horizontal resolution makes it difficult to distinguish between land and oceanic regions accurately). Furthermore, among the sites in the Himalayan region, Devasthal (35.8, 35.9 and 36.9 per cent) was found to have a slightly lower CF than Hanle (35.4, 38.5 and 39.9 per cent), followed by Ali (40.4, 40.3 and 39.9 per cent) and Merak (49.7, 50.2 and 51.6 per cent) during the night sky, as seen in Table 5 from the high-spatial-resolution MODIS data. Note that the results over the eight stations in this study were derived by linear interpolation from the surrounding four grid points.
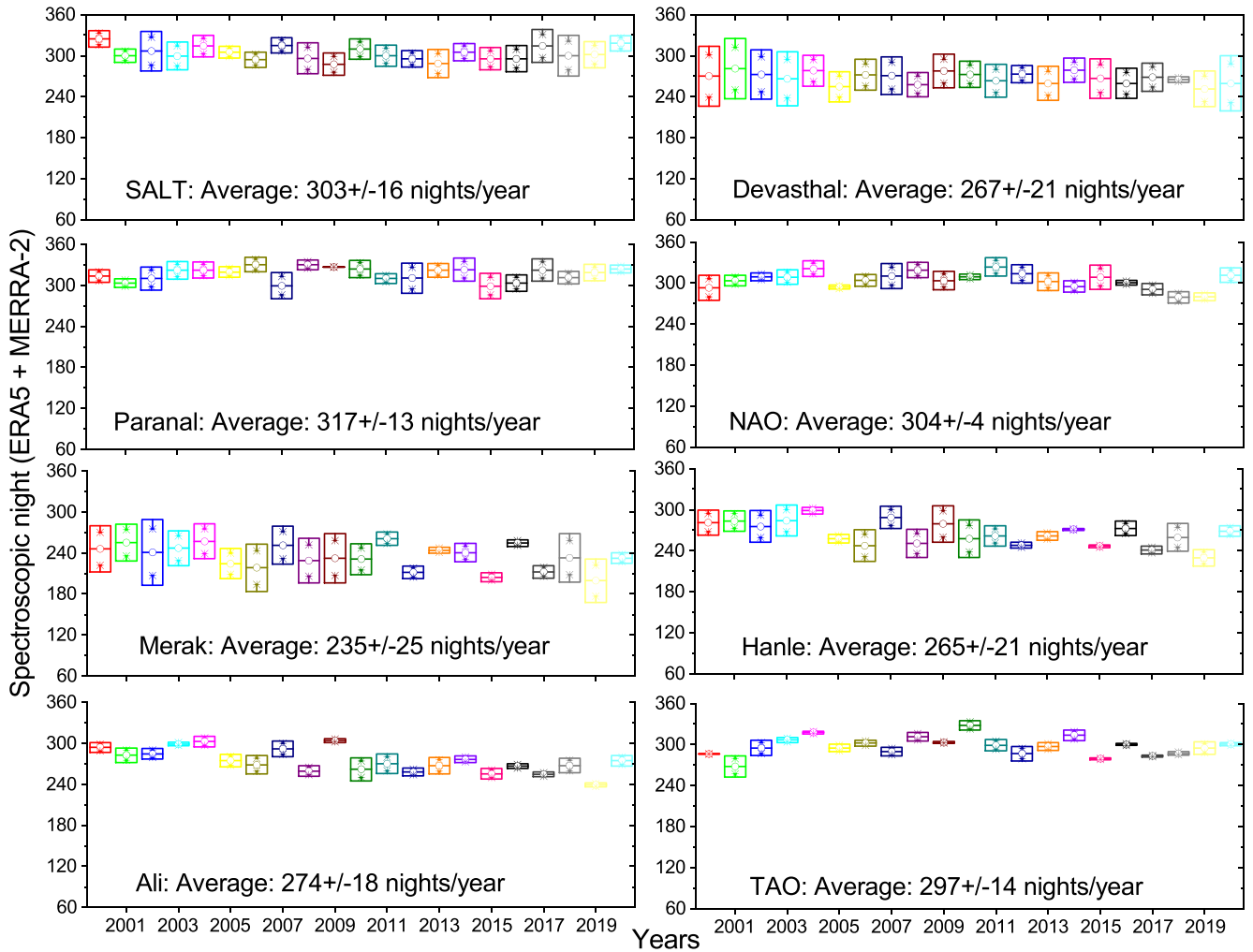
Figure 8. Box and whisker plots for yearly time-series spectroscopic nights of ERA5 and MERRA-2 (combined) data during 2000–2020 at the eight astronomical sites. The central horizontal lines inside the boxes are medians; open circles inside the boxes are arithmetic means; and bottoms and tops of the boxes are the minimum and maximum values. The 21 different colours in each panel show 21 different years of data, and the numerical values at each site in each panel show the average with the standard deviation from the combined data.

4.1 Long-term variation of CF

Based on the trend of CF changes over the past 41 years (Fig. 10b), there is an increasing trend in maritime regions, mainly in the eastern Pacific Ocean, the western Indian Ocean, and the tropical Atlantic Ocean. In contrast, the trend tends to decline in the central region of Africa, the Eurasian continent, and the American continents. Over land, the Sahara desert, Middle East, Indian subcontinent, Tibetan Plateau and some Southeast Asian islands show an increasing CF trend. Here, the increasing trend over the Southeast Asian islands can be regarded as the part of the changes in adjacent marine areas. Although the CF trends of ERA5 and MERRA-2 differ locally, they are similar in terms of large-scale changes (see Fig. 10b,c). However, with respect to accuracy, ERA5 is more reliable than MERRA-2, as shown in Fig. 10. The increasing trend over oceans and decreasing trend over land regions could be attributed to changes in temperature and water vapour associated with global warming (Held & Soden 2006). It is widely known that owing to the difference in specific heat between land and ocean, global warming is stronger in land areas than in the ocean (Boer 2011). With respect to global warming, the ocean's water vapour evaporates more, which increases the total amount of

water vapour in the atmosphere. Locally, there is also a tendency to increase water vapour over the western Pacific and decrease water vapour over the central-eastern Pacific, with the strengthened Walker circulation (Sohn et al. 2013; Chung et al. 2019). In fact, the RH is widely used to estimate CF in numerical models (Smith 1990; Song et al. 2020). The decreasing RH trend over land was also noted by Byrne & O'Gorman (2016), and its possible mechanism with ocean warming was explained by Byrne & O'Gorman (2018). The decreasing CF trend provides favourable conditions for astronomical observations, with more clear sky conditions. As shown in Table 4, the NAO (-0.68 per cent decade^{-1}), SALT (-0.69 per cent decade^{-1}), and TAO (-0.45 per cent decade^{-1}) sites correspond to the land regions with a decreasing CF trend. Tables 4–6 also support the mechanism of CF changes associated with land over these sites. However, with relatively low-temperature increases in the ocean regions, the total column water vapour increased, except for the central Pacific.

It can be noted that Paranal is located on the boundary (coastline) between the increasing ocean mode and decreasing land mode. In addition, the increasing CF trends over the Sahara desert, Middle
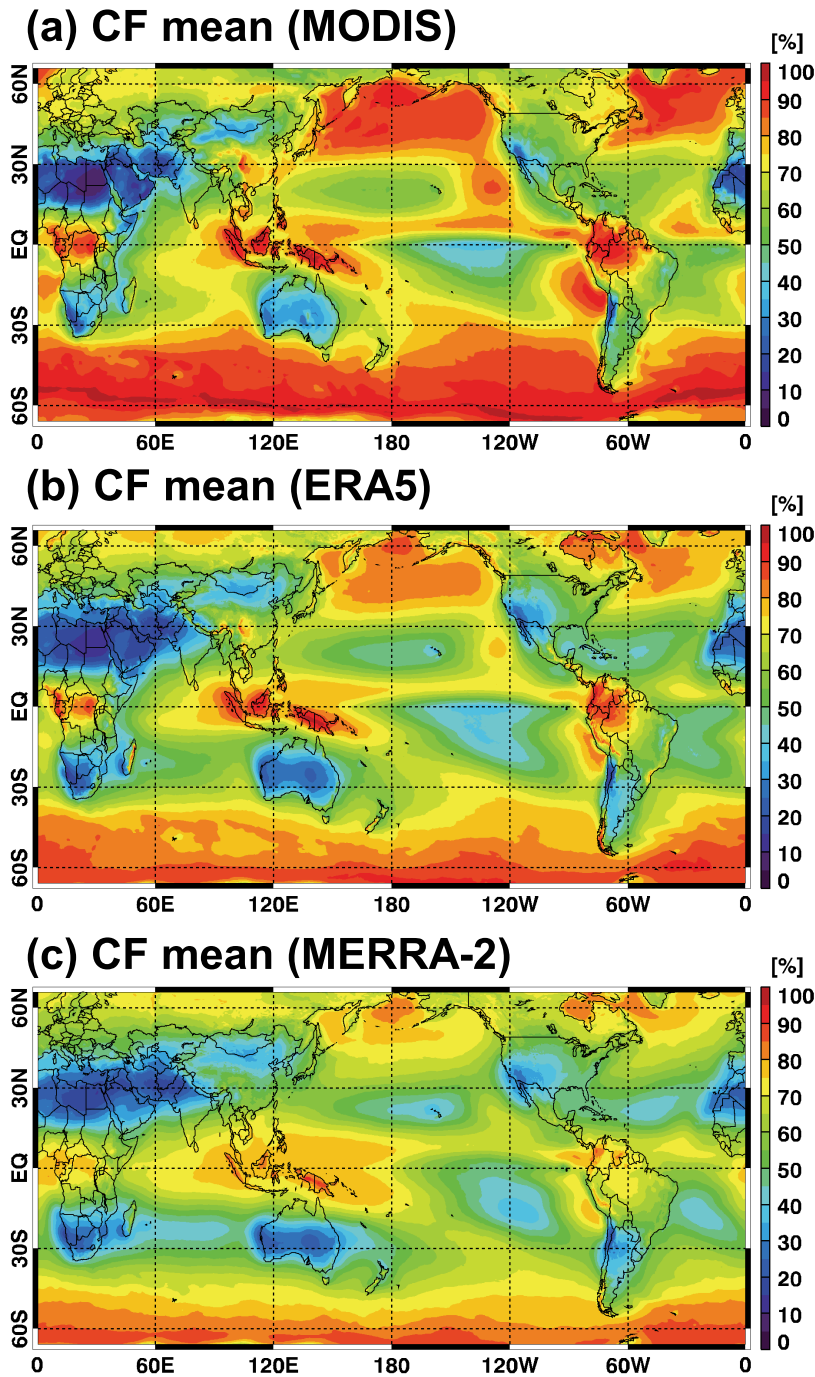
(a) CF mean (MODIS)

(b) CF mean (ERA5)

(c) CF mean (MERRA-2)

Figure 9. Climatological mean of night-time CF (per cent) during the recent 21 years (2000–2020) from (a) satellite (MODIS, *Terra*), and (b) ERA5 and (c) MERRA-2 reanalysis data.

East, Indian subcontinent, and Tibetan Plateau are thought to result from with tremendous supply of water vapour compared to the increase in temperature, as shown in Fig. 11(e,h). For example, the Ali (0.77 per cent decade^{-1}), Devasthal (0.63 per cent decade^{-1}), Hanle (0.43 per cent decade^{-1}) and Merak (0.05 per cent decade^{-1}) sites are subject to this effect.

The quality of an astronomical site is decided from several astroclimatological parameters in addition to the clear sky condition or minimum CF throughout the year. Furthermore, PWV is a critical parameter for infrared astronomy. In this respect, IAO-Hanle and Merak in the Ladakh region are found to be as dry as TAO, the driest astronomical site on the planet as reported by Ningombam et al. (2020b). In addition, these sites are located in the rain-shadow area of the Himalayas, so observations can be continued throughout the year. Owing to its isolation from any source of anthropogenic activity, the ambient dust content at IAO-Hanle means that it is
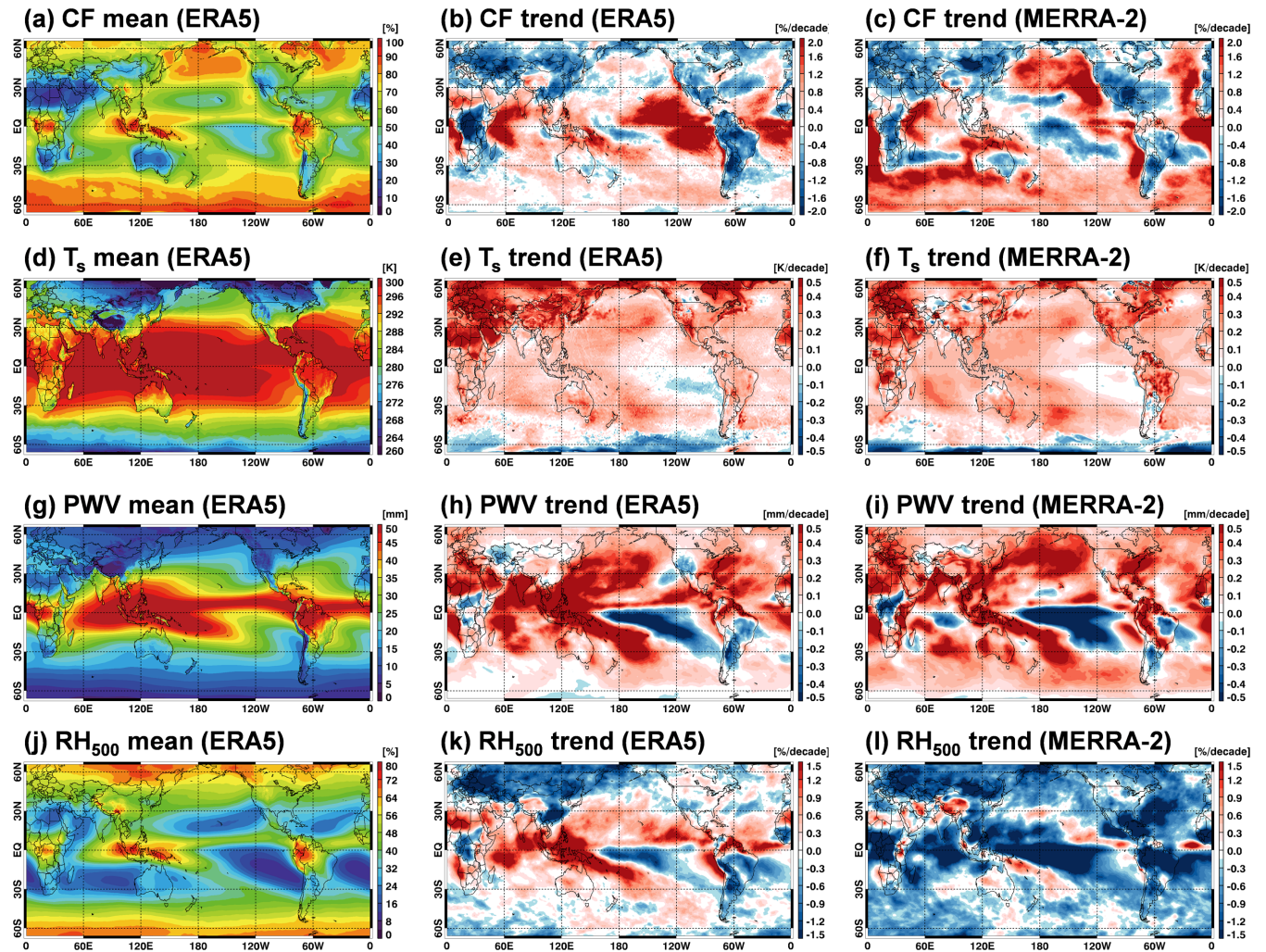
Figure 10. Global means of cloud fraction (CF, per cent), surface temperature (T_s, K), PWV (mm) and RH at 500 hPa (RH_{500}, per cent), and their long-term trends from reanalysis data (ERA5 and MERRA-2) during the 41 years of the study period (1980–2020).

Table 4. Mean night-time cloud fraction (per cent) and long-term trend (per cent decade^{-1}) using ERA5 and MERRA-2 data during 1980–2020 at the eight high-altitude astronomical observatories.

Site	ERA5			MERRA-2		
	Mean	Trend	t-value	Mean	Trend	t-value
SALT	31.36	− 0.69	− 2.32	36.31	− 0.24	− 0.89
Devasthal	40.63	0.63	1.51	30.35	0.27	0.68
Paranal	18.32	0.11	0.21	24.30	0.42	1.31
NAO	25.24	− 0.68	− 1.92	32.22	− 0.79	− 2.16
Merak	54.15	0.05	0.11	40.85	1.00	2.38
Hanle	38.05	0.43	0.95	35.17	1.51	3.62
Ali	30.07	0.77	1.88	34.03	1.29	3.14
TAO	24.91	− 0.45	− 1.20	24.71	0.09	0.24

characterized as a pristine location, and it is free from dust storms, despite the surrounding desert dust. It can be mentioned that the ambient dust concentration of particle sizes from 0.3 to 25 μm was measured at IAO-Hanle during 2015 November. The study found that the average number of particles of size 0.5 μm was about 10 000 per cubic foot at the hill top (Mt Saraswati) of the surrounding ($\sim$50–100 m) observatory, which is equivalent to ten thousand class of clean room standard protocol. During the measurement, the surface wind speed varied from 0 to 20 ms^{-1}, while the estimated aerosol optical depth (AOD) using a sky radiometer (Prede, POM-01) varied from 0.02 to 0.04 at 500 nm, which indicates that the observing site is characterized by clean environmental conditions and also free from desert dust (Ningombam et al. 2017).

The statistical significances of PWV for these Himalayan sites are higher than that of relative humidity at 500 hPa (RH_{500}) (Tables 6 and 7). The CF trend over the Merak site was found to be weak and also statistically insignificant. Owing to the high variation of moisture content in the atmosphere, the mean PWV is used from the high-resolution MODIS data. Among the sites, Merak exhibits the lowest PWV, with 1.37, 1.31 and 1.17 mm at 20 × 20 km, 10 × 10 km and 5 × 5 km spatial resolution, respectively, as seen in Table 5 from data for the last 21 years (2000–2020). The statistical significance levels for surface temperature and PWV trends are higher than that for CF. For temperature, most astronomical sites, except for TAO (ERA5) and Merak (MERRA-2), are significant at the 90 per cent confidence level. PWV trends are significant at the 95 per cent confidence level over the Ali, Devasthal, Hanle and Merak (ERA5) sites. Similar to CF trends, the statistical significance for linear trends of RH_{500} is relative weak. The linear trends of RH_{500} at Paranal and SLAT for ERA5, in addition to at Ali, Hanle, Merak, NAO and SALT for MERRA-2, are significant at the 90 per cent confidence level.

Table 5. Long-term, high-resolution (0.20° to 0.05°) MODIS *Terra*) night-time cloud fraction (CF, per cent), photometric (Photo) and spectroscopic (Spec) nights (per cent yr^{-1}), and PWV (mm) at the eight high-altitude astronomical observatories. The numbers of data available (*N*) during 2000 March to 2020 December are also shown in per cent.

Site	0.20° × 0.20°					0.10° × 0.10°					0.05° × 0.05°				
	CF	Photo	Spec	PWV	*N*	CF	Photo	Spec	PWV	*N*	CF	Photo	Spec	PWV	*N*
SALT	26.6	63.0	74.2	9.62	97.9	27.1	64.9	73.2	9.65	85.5	27.8	65.7	72.9	9.56	50.3
Devasthal	35.8	55.5	63.9	15.78	97.1	35.9	57.4	64.4	15.61	85.6	36.9	58.4	64.2	15.73	52.1
Paranal	8.4	86.6	92.8	10.21	92.7	7.5	89.7	93.0	10.18	80.6	7.6	90.3	92.7	10.20	50.5
NAO	27.5	61.4	73.5	14.54	96.8	31.5	57.9	70.1	14.22	82.4	33.3	56.3	68.0	14.15	50.3
Merak	49.7	16.5	52.3	1.37	97.8	50.2	22.4	50.4	1.31	86.6	51.6	27.2	49.7	1.17	51.1
Hanle	35.4	42.6	67.5	1.98	97.7	38.5	37.1	66.1	2.08	85.3	39.9	37.5	57.0	1.78	52.0
Ali	40.4	39.4	62.5	1.75	97.8	40.3	44.3	61.2	1.76	86.7	39.9	47.8	62.0	1.78	50.9
TAO	27.1	59.8	74.9	2.33	92.4	32.2	51.8	70.6	2.33	80.8	36.2	46.1	67.4	2.29	51.8

Table 6. Night-time long-term trends of surface temperature (K decade^{-1}) and PWV (mm decade^{-1}) using ERA5 and MERRA-2 data during 1980–2020 at the eight high-altitude astronomical observatories.

Site	ERA5				MERRA-2			
	T trend	*t*-value	PWV trend	*t*-value	T trend	*t*-value	PWV trend	*t*-value
SALT	0.20	4.97	0.00	0.01	0.21	5.31	0.03	0.55
Devasthal	0.15	4.79	0.34	3.61	0.21	5.83	0.89	7.27
Paranal	0.17	6.70	− 0.04	− 0.72	− 0.03	− 1.84	0.04	0.63
NAO	0.26	5.48	− 0.08	− 1.07	0.15	2.89	− 0.10	− 1.09
Merak	0.40	4.62	0.06	2.81	0.00	0.11	0.16	6.64
Hanle	0.43	3.71	0.08	3.02	0.16	2.02	0.20	7.30
Ali	0.21	1.75	0.10	3.58	0.44	8.46	0.21	7.24
TAO	0.12	1.25	0.00	0.15	0.27	7.30	0.04	0.96

4.2 Vertical profiles of CF

Because CF is directly related to the temperature and moisture content profiles in the atmosphere, it is interesting to examine such variations at various heights. A total of six variables (air temperature, specific humidity, RH, CF, cloud liquid water, cloud ice water) were analysed using the MERRA-2 monthly mean data with 73 vertical levels. Among them, according to the American Meteorology Society (AMS) glossary (https://glossary.ametsoc.org/wiki/Welcome), specific humidity is defined as the ratio of the mass of water vapour to the total mass of the system; relative humidity is the ratio of vapour pressure to the saturation vapour pressure with respect to water; cloud liquid and ice water (or the liquid and ice water mixing ratio) are the ratios of the mass of liquid and ice water to the mass of dry air in a unit volume of air. Fig. 11 shows the vertical profiles for (a) air temperature, (b) specific humidity (kg/kg), (c) RH, (d) CF, (e) cloud liquid water (CLW), and (f) cloud ice water (CIW) at various pressure levels (800 to 50 hPa) from the monthly MERRA-2 data during 1980–2020. Note that these results are based on a vertical cut-off from the nearest station pressure estimated from the geopotential height at each site listed in Table 1. First of all, the combination of temperature and specific humidity can determine the RH, and it can be directly related to CF. Here, RH is determined by the combination method of saturation vapour pressures over liquid and ice, such as described by Simmons et al. (1999) and Song et al. (2020). According to air temperature in the middle level around 300–500 hPa, the Ali, Hanle, Merak, NAO and SALT sites are characterized by a lower temperature than the Devasthal, Paranal and TAO sites, as shown in Fig. 11(a). In addition, specific humidity profiles over the Ali, Devasthal, Hanle and Merak sites are higher than at other sites (Fig. 11b). The lower temperature and higher specific humidity over the Ali, Hanle and Merak sites are related to the higher RH (Fig. 11c) and resulting higher CF (Fig. 11d) around 300–500 hPa. In the upper level around

100–200 hPa, TAO is characterized by the highest RH and CF group, while Paranal and NAO are in the second highest group. The CF result can be further divided into CLW and CIW in terms of mass fraction. It is interesting that high CFs in the middle troposphere over the Ali, Hanle and Merak sites are related to the high CLW, whereas the CLW shows an increasing tendency with decreasing height such as in specific humidity, whereas most CF profiles in the upper troposphere are closely related to the CIW pattern. These results will contribute to a better understanding of the vertical distribution of CF.

5 SUMMARY AND DISCUSSIONS

This work has examined the number of clear nights at the locations of eight high-altitude astronomical sites using 21 years (2000–2020) of visual and satellite data along with 41 years (1998–2020) of hourly reanalysis data. However, the visual data used were available only at IAO-Hanle. In order to quantify the number of useful photometric and spectroscopic nights, the methodology used for the visual data was also applied to the satellite and reanalysis data. The number of photometric nights from reanalysis data at IAO-Hanle is underestimated by 8–24 per cent in comparison to that from the visual data, while the number of spectroscopic nights estimated from reanalysis data agreed reasonably well with the visual observations of 70–75 per cent nights per year. This highlights the limitations of the models and satellite data. Results may differ from observational data as a result of low temporal or spatial resolution, especially under particular topographical conditions. This study found that there are 68–78 per cent spectroscopic nights per year at the three potential sites located in the Himalayan region, namely Hanle, Ali and Devasthal, from the study of ERA5 and MERRA-2 long-term (2000–2020) reanalysis data, while Merak exhibits 61–68 per cent spectroscopic nights per year. Among the eight astronomical sites,
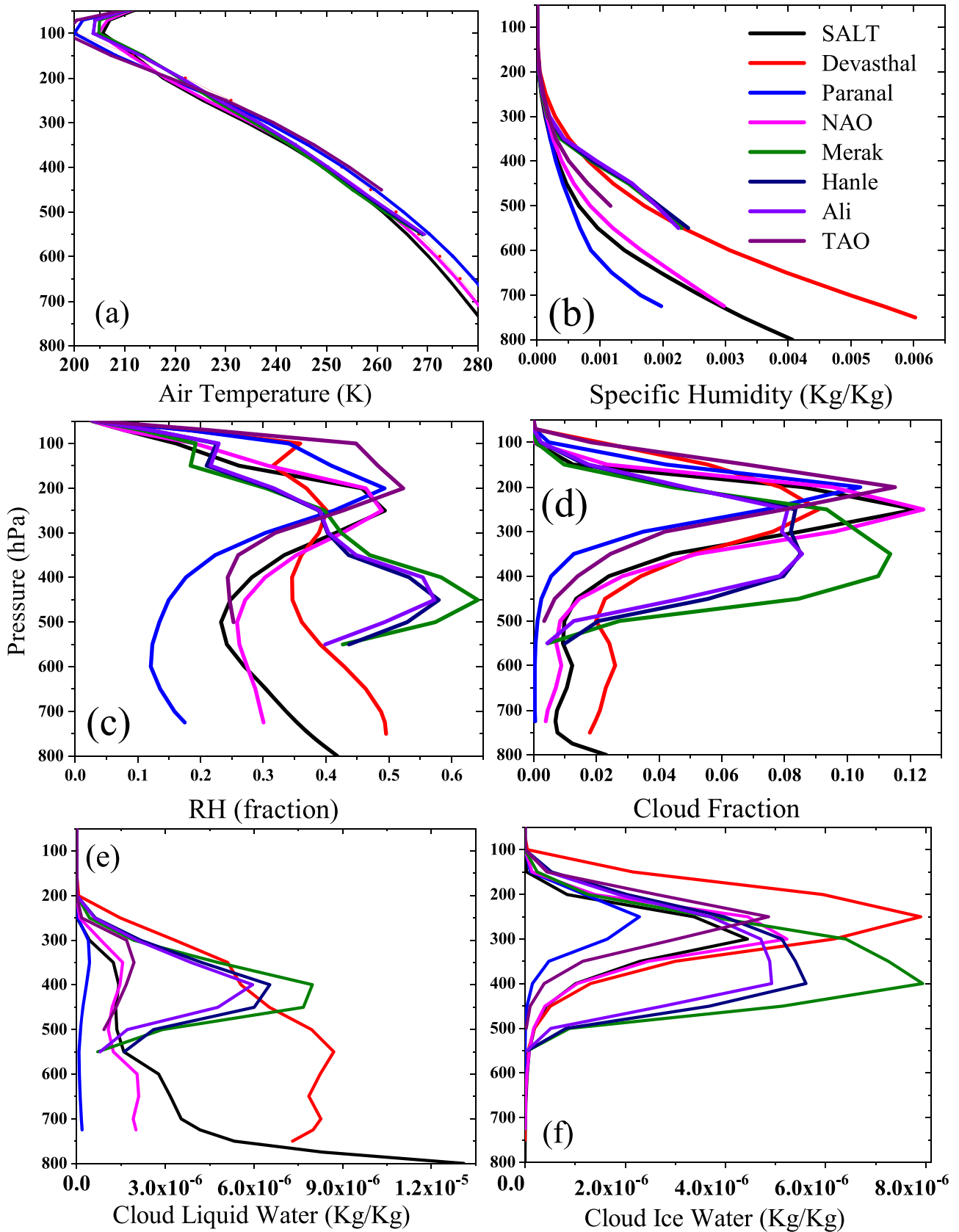
Figure 11. Vertical profiles of (a) air temperature (K), (b) specific humidity (kg/kg), (c) relative humidity (RH, fraction), (d) cloud fraction, (e) cloud liquid water, and (f) cloud ice water at the lowest pressure level observed from geopotential height (above the elevation height) using MERRA-2 data during 1980–2020 at the eight astronomical sites.

Table 7. Mean night-time RH at 500 hPa (per cent) and long-term trend (per cent decade^{-1}) using ERA5 and MERRA-2 data during 1980–2020 at the eight high-altitude astronomical observatories considered in this study.

Site	ERA5			MERRA-2		
	Mean	Trend	*t*-value	Mean	Trend	*t*-value
SALT	22.86	− 0.30	− 1.71	23.78	− 0.56	− 3.00
Devasthal	38.95	0.06	0.20	36.89	0.55	1.61
Paranal	13.71	0.46	2.49	12.81	0.09	0.48
NAO	25.42	− 0.30	− 1.32	25.98	− 0.64	− 2.53
Merak	60.51	− 0.07	− 0.23	61.89	0.93	3.05
Hanle	55.59	− 0.07	− 0.23	57.53	1.36	4.34
Ali	52.21	0.14	0.45	53.59	1.31	4.15
TAO	28.37	− 0.14	− 0.44	27.00	− 0.37	− 1.09

Paranal is characterized as the best site for astronomical observation, with 87 per cent spectroscopic nights per year.

Global CF trends over the past 41 years (1980–2020) showed decreasing trends over the central region of Africa, the Eurasian continent, and the American continents, but increasing trends in maritime regions, mainly in the eastern Pacific Ocean, western Indian Ocean, and tropical Atlantic Ocean. The increasing trend in oceans and decreasing trend over land regions is thought to be the result of global warming and associated water vapour changes. Among land regions, the Sahara desert, Middle East, Indian subcontinent, Tibetan Plateau and some Southeast Asian islands show an increasing CF trend.

The vertical profiles of CF and related variables over eight astronomical sites were also examined using the MERRA-2 data during 1980–2020. The combination of temperature and specific humidity can determine RH, and it can be directly related to CF. According to the air temperature in the middle level around 300–500 hPa, the Ali, Hanle, Merak, NAO and SALT sites were characterized by a lower temperature than other sites, whereas the specific humidity profiles over the Ali, Devasthal, Hanle and Merak sites were higher than over other sites. As a result, the lower temperature and higher specific humidity over the Ali, Hanle and Merak sites produced higher RH and CF around 300–500 hPa. In the upper level, around 100–200 hPa, TAO is characterized by the highest RH and CF group, while Paranal and NAO are in the second highest group. In addition, the mid-level CFs over the Ali, Hanle and Merak sites were related to the high CLW, whereas most CF profiles in the upper troposphere were closely related to the CIW pattern.

ACKNOWLEDGEMENTS

DATA AVAILABILITY

Data available on request: The data underlying this article will be shared on reasonable request to the corresponding author.

REFERENCES

Bhatt B. C., Prabhu T. P., Anupama G. C., 2000, Bull. Astron. Soc. India, 28, 441

Boer G. J., 2011, Climate Dynamics, 37, 2253

Bosilovich M. G., Lucchesi R., Suarez M., 2016, MERRA-2: File specification. GMAO Office Note No. 9 (Version 1.1). Available at: http://gmao.gsfc.nasa.gov/pubs/office_notes

Byrne M. P., O'Gorman P. A., 2016, J. Climate, 29, 9045

Byrne M. P., O'Gorman P. A., 2018, Proc. Nat. Acad. Sci., 115, 4863

Cavazzani S., Ortolani S., Zitelli V., Maruccia Y., 2011, MNRAS, 411, 1271

Cavazzani S., Zitelli V., Ortolani S., 2015, MNRAS, 452, 2185

Chung E.-S., Timmermann A., Soden B. J., Ha K.-J., Shi L., John V. O., 2019, Nat. Climate Change, 9, 405

Erasmus D. A., van Rooyen R., 2006, in Stepp L. M., ed., Proc. SPIE Conf. Ser. Vol. 6267, Ground-based and Airborne Telescopes. SPIE, Bellingham, p. 62671O

Gelaro R. et al., 2017, J. Climate, 30, 5419

Held I. M., Soden B. J., 2006, J. Climate, 19, 5686

Hersbach H. et al., 2020, Q. J. R. Meteorological Soc., 146, 1999

HIROT Team, 1996, Bull. Astron. Soc. India, 24, 859

Lee H., Song H.-J., Sohn B.-J., 2021, Asia-Pacific J. Atmos. Sci., 57, 51

Ningombam S. S. et al., 2017, Experimental Astronomy, 43, 145

Ningombam S. S., Narendra A., Bhatt B. C., Prabhu T. P., Anupama G. C., Angchuk D., Jorphail S., 2020a, Advances Space Res., 66, 826

Ningombam S. S., Sethulakshmy E. S., Jade S., Shrungeshwara T. S., Vivek C. G., Angchuk D., Prabhu T. P., Mahay T. T., 2020b, J. Atmos. Solar-Terrestrial Phys., 208, 105404

Prabhu T. P., 2014, Proc. Indian Nat. Sci. Acad. Part A, 80, 887

Sagar R. et al., 2000, A&AS, 144, 349

Sagar R., Kumar B., Omar A., 2019, Current Sci., 117, 365

Simmons A. J., Untch A., Jakob C., Kållberg P., Undén P., 1999, Q. J. R. Meteorological Soc., 125, 353

Singh J., Bhattacharyya J. C., Babu G. S. D., Ashoka B. N., Appakutty M., 1989, Bull. Astron. Soc. India, 17, 83

Smith R. N. B., 1990, Q. J. R. Meteorological Soc., 116, 435

Sohn B. J., Yeh S.-W., Schmetz J., Song H.-J., 2013, Climate Dynamics, 40, 1721

Sohn B.-J., Lee S., Chung E.-S., Song H.-J., 2016, J. Climate, 29, 2765

Song H.-J., Sohn B.-J., 2020, Int. J. Climatology, 40, 2173

Song H.-J., Kim S., Lee H., Kim K.-H., 2020, Atmosphere, 11, 704

Tapia M., 2003, in Cruz-Gonzalez I., Avila R., Tapia M., eds, Rev. Mex. Astron. Astrofis. Conf. Ser., 19, 75

Thies B., Bendix J., 2011, Meteorological Applications, 18, 262

Vernin J. et al., 2011, PASP, 123, 1334

Ye Q.-Z., Su M., Li H., Zhang X., 2016, MNRAS, 457, L1

Yim B. Y., Yeh S.-W., Song H.-J., Dommenget D., Sohn B. J., 2017, Geophys. Res. Lett., 44, 5854

Zhang J.-C., Ge L., Lu X.-M., Cao Z.-H., Chen X., Mao Y.-N., Jiang X.-J., 2015, PASP, 127, 1292

SUPPORTING INFORMATION

Supplementary data are available at *MNRAS* online.

Figure S1. Bias and RMSE of reanalysis data (ERA5 and MERRA-2) from MODIS (*Terra*) during the 20 continuous years (2000 February to 2020 December) of the study period.

Please note: Oxford University Press is not responsible for the content or functionality of any supporting materials supplied by the authors. Any queries (other than missing material) should be directed to the corresponding author for the article.

This paper has been typeset from a TEX/LATEX file prepared by the author.