Evaluation of ERA5-Land and HARv2 Reanalysis Data at High Elevation in the Upper Dudh Koshi Basin (Everest Region, Nepal)

Arbindra Khadka, Patrick Wagnon, Fanny Brun, Dibas Shrestha, Yves Lejeune, Yves Arnaud

Evaluation of ERA5-Land and HARv2 Reanalysis Data at High Elevation in the Upper Dudh Koshi Basin (Everest Region, Nepal)

ARBINDRA KHADKA,[a,b,c] PATRICK WAGNON,[a,b] FANNY BRUN,[a] DIBAS SHRESTHA,[c] YVES LEJEUNE,[d] AND YVES ARNAUD[a]

[a] *Université Grenoble Alpes, CNRS, IRD, IGE, Grenoble, France*
[b] *International Centre for Integrated Mountain Development, Kathmandu, Nepal*
[c] *Central Department of Hydrology and Meteorology, Tribhuvan University, Kirtipur, Nepal*
[d] *Université Grenoble Alpes, Université de Toulouse, Météo-France, CNRS, Centre d'Etudes de la Neige, Grenoble, France*

(Manuscript received 11 May 2021, in final form 5 April 2022)

ABSTRACT: We present a multisite evaluation of meteorological variables in the Everest region (Nepal) from ERA5-Land and High Asian Refined Analysis, version 2 (HARv2), reanalyses in comparison with in situ observations, using classical statistical metrics. Observation data have been collected since 2010 by seven meteorological stations located on or off glacier between 4260 and 6352 m MSL in the upper Dudh Koshi basin; 2-m air temperature, specific and relative humidities, wind speed, incoming shortwave and longwave radiations, and precipitation are considered successively. Overall, both gridded datasets are able to resolve the mesoscale atmospheric processes, with a slightly better performance for HARv2 than that for ERA5-Land, especially for wind speed. Because of the complex topography, they fail to reproduce local- to microscale processes captured at individual meteorological stations, especially for variables that have a large spatial variability such as precipitation or wind speed. Air temperature is the variable that is best captured by reanalyses, as long as an appropriate elevational gradient of air temperature above ground, spatiotemporally variable and preferentially assessed by local observations, is used to extrapolate it vertically. A cold bias is still observed but attenuated over clean-ice glaciers. The atmospheric water content is well represented by both gridded datasets even though we observe a small humid bias, slightly more important for ERA5-Land than for HARv2, and a spectacular overestimation of precipitation during the monsoon. The agreement between reanalyzed and observed shortwave and longwave incoming radiations depends on the elevation difference between the station site and the reanalysis grid cell. The seasonality of wind speed is only captured by HARv2. The two gridded datasets ERA5-Land and HARv2 are applicable for glacier mass and energy balance studies, as long as either statistical or dynamical downscaling techniques are used to resolve the scale mismatch between coarse mesoscale grids and fine-scale grids or individual sites.

KEYWORDS: Asia; Complex terrain; Glaciers; Automatic weather stations; Surface observations; Statistics; Reanalysis data; Mountain meteorology

1. Introduction

High-Mountain Asia (HMA), which comprises the Tibetan Plateau and its surrounding mountain ranges, is referred to as the water tower of Asia. The cryosphere (i.e., snow, ice, and permafrost) contributes to water storage in this region (Immerzeel et al. 2010). Observed and projected changes in the cryosphere will affect the magnitude and timing of streamflow, especially in the upstream area, with large socioeconomic impacts (Lutz et al. 2014; Pritchard 2019; Bolch et al. 2019; Immerzeel et al. 2020).

HMA contains the largest concentration of glacier ice outside of the polar regions, with nearly 100 000 km^2 of glacierized area (Pfeffer et al. 2014). Glaciers have been shrinking at least since the 1970s except for parts of Karakoram, eastern Pamir, and western Kunlun (e.g., Bolch et al. 2019; Berthier and Brun 2019). The sea level rise contribution from HMA glacier mass loss since 2000 is up to ~0.05 ± 0.01 mm sea level equivalent per year (Brun et al. 2017; Shean et al. 2020). The ice loss rates are the greatest across Nyainqentanglha, the Himalayas, and Tien Shan (Shean et al. 2020). Even if global warming is constrained to the most ambitious target of 1.5°C (United Nations 2015), more than 30% of HMA glacier ice will likely disappear by the end of this century (Kraaijenbrink et al. 2017; Bolch et al. 2019).

In the central Himalayas, the rate of mass loss is proportionally high (Shean et al. 2020). Several field-based glaciological studies in Nepal show negative mass balance years, especially during the last 5–10 years (e.g., Sunako et al. 2019; Sherpa et al. 2017; Wagnon et al. 2020; Stumm et al. 2021). Glacier mass changes result from climatic forcing and from the mass-balance sensitivity to meteorological variables such as air temperature or precipitation (Sakai and Fujita 2017). To examine the relationship between mass balance and climate, long-term high-quality meteorological datasets are required. In the upper Dudh Koshi basin, our region of interest, the first meteorological data ever recorded come from the historical mountaineering expeditions to Everest, back in the 1950s (Pugh 1954), but continuous and systematic records were not initiated before

Supplemental information related to this paper is available at the Journals Online website: https://doi.org/10.1175/JAMC-D-21-0091.s1.

Corresponding author: Arbindra Khadka, arbindra.khadka@icimod.org

DOI: 10.1175/JAMC-D-21-0091.1

the installation of Pyramid, at the beginning of the 1990s (Salerno et al. 2015). Since 2010, a network of automatic weather stations (AWSs) was installed and has been operating in the upper Dudh Koshi basin, with a focus on glacier areas and high elevation (>4260 m MSL) (Shea et al. 2015; Sherpa et al. 2017). The data from these AWSs contributed to better understand the glacier behavior locally (Sherpa et al. 2017) but they provide very local information and consequently cannot be used to analyze regional glacier evolution.

Climate reanalyses provide alternative data for climate and glacier studies, especially in regions where in situ data are scarce and discontinuous. The European Centre for Medium-Range Weather Forecasts (ECMWF) reanalysis product ERA-Interim (Dee et al. 2011) has been used for instance to reconstruct mass balance data in Nepal (Sunako et al. 2019). In 2017, ECMWF released a fifth generation of reanalysis product, ERA5 (Hersbach et al. 2020), with a higher spatial and temporal resolution than ERA-Interim. ERA5 reanalyses perform well in the Antarctic and Arctic regions (Wang et al. 2019; Tetzner et al. 2019). The performance of ERA5 temperature and wind data has been significantly improved in comparison with ERA-Interim in the Antarctic Peninsula where ERA5 data have been used to calibrate proxy records from ice cores (Tetzner et al. 2019). Similarly, in the Indus basin, ERA5 precipitation data are highly correlated with Asian Precipitation–Highly Resolved Observational Data Integration Towards Evaluation of Water Resources (APHRODITE) data (Baudouin et al. 2020). However, ERA5-Land (ERA5L), which is a reanalysis product derived by running the land component of ERA5 at increased resolution, is not well suited for permafrost studies because it overestimates soil temperature in high latitudes and underestimates it in mid–low latitudes (Cao et al. 2020). Also, the High Asian Refined Analysis, version 2 (HARv2), is another recently published high-resolution dataset generated by dynamical downscaling of ERA5 reanalysis data (Wang et al. 2020). Maussion et al. (2011, 2014) provided a first generation of HAR data and analyzed the precipitation seasonality and variability over the Tibetan Plateau and surrounding regions. They validated HAR precipitation data with both rain gauge observations unevenly distributed over the Tibetan Plateau, and satellite-based precipitation estimates from the Tropical Rainfall Measuring Mission. Based on their regional dataset, they proposed a classification of glaciers according to their accumulation regimes. Mölg et al. (2012, 2014) used this dataset and showed that the timing and amount of snowfall in the early ablation season (May–June) is a key process controlling the annual mass balance of glaciers on the southern Tibetan Plateau.

Because of the lack of long-term observation data, the use of high-resolution reanalysis data is promising in the Nepalese Himalayas, where in situ data collection is challenging given the complex topography, remoteness, and data gaps caused by infrequent maintenance and failure of AWSs. Nevertheless, even though those reanalysis products have the best spatiotemporal resolution in the central Himalayas, ERA5L and HARv2 data have not been validated yet using available in situ data from high-elevation AWSs located on and off glaciers although such products often need calibration.

In this study, we compare ERA5L and HARv2 reanalysis data with in situ data from seven AWSs from different glacierized and unglacierized areas since 2010 in the upper Dudh Koshi basin, which are not assimilated in the reanalysis process of ERA5L and HARv2. The main focus of this study is to assess the performance of ERA5L and HARv2 reanalysis data at glacier elevations in the central Himalayas. We qualitatively discuss whether such data are reliable for glacier mass and energy balance studies, and we discuss the relevance of substituting in situ data, if not available, with reanalysis data for glaciological or hydrological modeling purposes.

2. Study site and climatology

The Nepalese Himalayas are under the influence of the Indian monsoon originating from the Bay of Bengal. In summer, large amounts of humid air travel north to northwest and trigger orographic precipitation (e.g., Bookhagen and Burbank 2010; Perry et al. 2020). In the upper Dudh Koshi basin (Fig. 1), approximately three-quarters of the annual precipitation fall between June and September (Shea et al. 2015; Sherpa et al. 2017; Perry et al. 2020). Following previous studies (e.g., Bonasoni et al. 2010), we divide the year into four seasons: winter [December–February (DJF)], premonsoon [March–May (MAM)], monsoon [June–September (JJAS)], and postmonsoon [October–November (ON)].

3. Data and methods

a. Data

1) AWS DATA

Meteorological observations are collected at various AWSs off and on glaciers in the upper Dudh Koshi basin at elevations ranging from 4260 to 6352 m MSL. Precipitation data have been collected by all-weather rain gauges (Geonor T-200B) at only three sites (Pyramid and Pheriche since December 2012; Khare since November 2016). They have been corrected for undercatch following the method by Førland et al. (1996) and Lejeune et al. (2007) as a function of wind speed and precipitation phase (liquid or solid) depending on air temperature (section 1c in the online supplemental material). This correction results in an approximately 20% increase of the amount of precipitation originally measured by the gauge, and the error range is estimated to be ±15% (Sherpa et al. 2017). Records go back to October 2010 at the Changri Nup site, which makes this dataset one of the longest at high elevation (5350 m MSL) in Nepal (Fig. 1 and Table 1). Table 1 provides the range of accuracy of every sensor used on those AWSs, as specified by the manufacturer. However, in such a harsh environment, measurement errors may sometimes exceed this error range, even though all datasets have been quality controlled (see section 1 in the online supplemental material, where AWS photographs are also available). Temperature/humidity sensors are artificially ventilated (with an aspirated Atmos radiation shield, model 43502, maintaining a 5 m s^{-1} ventilation during daytime when clear sky), when installed over glacierized surfaces, except at Mera Summit. At this

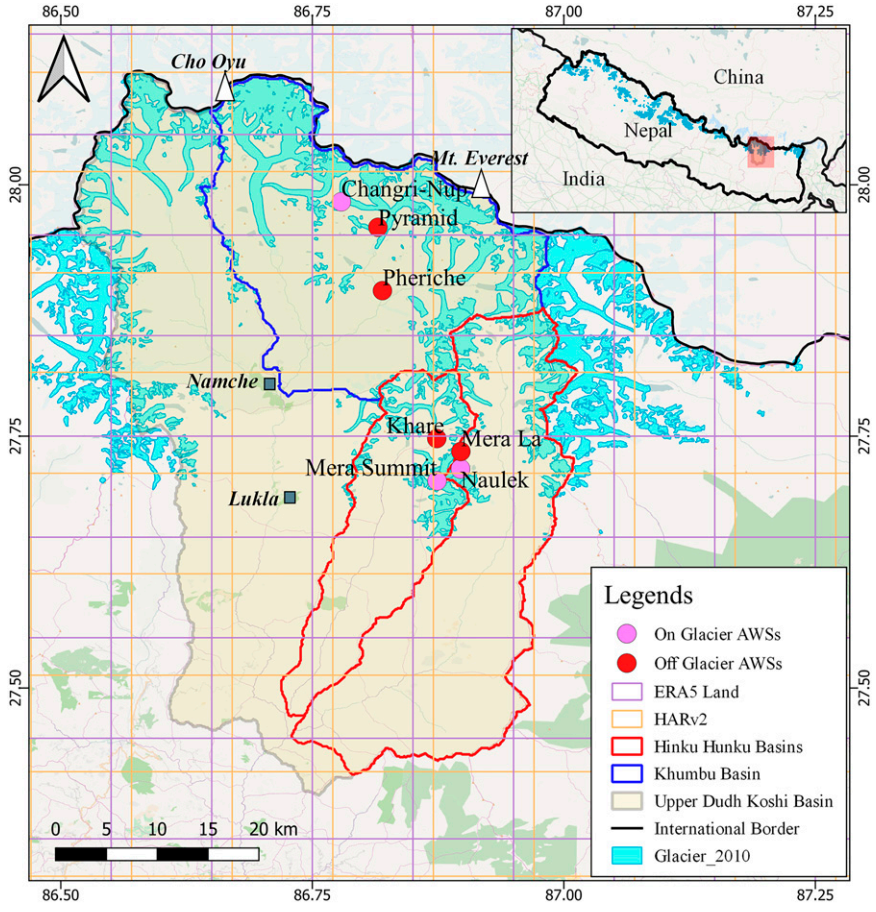

FIG. 1. Map of the study area with the AWSs located on glacier (pink dots) and off glacier (red dots). The glaciers are represented in blue (ICIMOD inventory 2010), and the grids represent the reanalysis grid cells (ERA5L in purple; HARv2 in orange). The Dudh Koshi basin is also delineated (yellow shading and gray line) in the main map and in the inset.

extremely high site, it is not possible to maintain artificial ventilation and the sensor is thus prone to overheating when shortwave radiation is large and natural ventilation low. Consequently, observed air temperature at 6352 m MSL is likely to be overestimated when wind speed is low, mainly during the monsoon. There are no data at Mera Summit during the postmonsoon because the AWS has been systematically buried by monsoonal snowfalls.

2) REANALYSIS DATA

ERA5L (1981–2020) is an ECMWF reanalysis dataset providing a consistent view of the evolution of land variables at a higher resolution (0.1° × 0.1°) than that of ERA5 (0.25° × 0.25°). ERA5L has been produced by replaying the land component of the ECMWF ERA5 climate reanalysis (Hersbach et al. 2020). The HAR (Maussion et al. 2011, 2014) is a regional atmospheric dataset generated by dynamical downscaling using the Weather Research and Forecasting (WRF) Model as regional climate model. HARv2 (1991–2020) is a refined version of HAR data with extended temporal and spatial coverage, using ERA5 as input data. It will be extended back to 1979 and continuously updated in the future. It provides gridded meteorological fields at 10-km resolution for the Tibetan Plateau and surroundings (Wang et al. 2020).

b. *Methods of evaluation*

The two reanalyses (ERA5L and HARv2) provided from the nearest grid points to each AWS site are compared with AWS data to evaluate their performance. In the Khumbu area (Changri Nup, Pyramid and Pheriche sites), AWSs are located on two distinct cells of ERA5L and HARv2, whereas Mera AWSs are located on one single cell of ERA5L and two neighboring cells of HARv2 (Fig. 1, Table 1). We compare the temporal variability of 2-m air temperature T, relative humidity RH, specific humidity q, wind speed u, shortwave (SWin) and longwave (LWin) incoming radiations, and total precipitation P recorded at the different AWSs with that of ERA5L and HARv2 reanalysis data. We evaluate the performance of HARv2 and ERA5L reanalysis datasets until November 2020, over the exact same periods corresponding to the periods when in situ data are available (Table 1). Consistently, the reanalysis data are converted into Nepal standard

TABLE 1. AWS locations, measurements, data gaps, sensor types, heights and accuracy, data periods, and morphological settings (T = air temperature, RH = relative humidity, u = wind speed, SWin = shortwave incoming radiation, LWin = longwave incoming radiation, P = total precipitation, and P_a = atmospheric pressure). For the sensors, the error range is according to the manufacturer.

	Elev (m MSL) AWS/ERA5L/HARv2	Lat; Lon (°)	Variables (% missing data)	Sensors (height; error range)	Data period	Morphological setting
Changri Nup	5360/5579[a]/5217	27.983; 86.783	T and RH (5.9)	Vaisala-HMP45C[b] (2 m; ±0.2°C and ±2%)	Nov 2010–Nov 2020[d]	Flat debris-covered surface of the ablation zone of West Changri Nup Glacier
			SWin (9.3) and LWin (15.5)	Kipp and Zonen CNR4 (1.2 m; ±3%)		
			u (15.5)	Young 05103–5 (2.5 m; ±0.3 m s^{-1})		
Mera La	5350/4662[a]/5342	27.735; 86.898	T and RH (0)	Vaisala-HMP45C[b] (2.4 m; ±0.2°C and ±2%)	Nov 2013–Nov 2020	Rough rocky surface off glacier, near Mera Pass
			SWin and LWin (0)	Kipp and Zonen CNR4 (1.2 m; ±3%)		
			u (0)	Young 05103-5 (2.8 m; ±0.3 m s^{-1})		
			P_a (0)	CS100 (0.8 m; ±1.5 hPa)		
Mera Summit	6352/4662[a]/4637	27.706; 86.874	T and RH (10.9)	Vaisala-HMP45C (1.6 m; ±0.2°C and ±2%)	Nov 2013–Jul 2016	5° north-facing sloppy surface of the accumulation zone of Mera Glacier, near the south summit
			SWin (22.5) and LWin (18.3)	Kipp and Zonen CNR4 (1.3 m; ±3%)		
			u (12.2)	Young 05103-5 (3.2 m; ±0.3 m s^{-1})		
			P_a (10.9)	CS100 (0.8 m; ±1.5 hPa)		
Naulek	5360/4662[a]/5342	27.718; 86.897	T and RH (12.7)	Vaisala-HMP45C[b] (1.5 m; ±0.2°C and ±2%)	Dec 2012–Nov 2020[d]	Flat clean-ice surface of the ablation zone of Mera Glacier, on Naulek Branch
			SWin (19.3) and LWin (19.4)	Kipp and Zonen CNR4 (0.8 m; ±3%)		
			u (20.2)	Young 05103-5 (2 m; ±0.3 m s^{-1})		
Pyramid	5035/5579[a]/5217	27.958; 86.815	P (0)	GEONOR T-200BM (1.8 m; ±15%) (shielded and corrected[c])	Dec 2012–Nov 2019	Flat grassy surface, off glacier
Pheriche	4260/4922[a]/4723	27.895; 86.820	P (0)	GEONOR T-200BM (1.8 m; ±15%) (shielded and corrected[c])	Dec 2012–Nov 2019	Flat grassy surface, off glacier
Khare	4888/4662[a]/5342	27.748; 86.874	P (0)	GEONOR T-200BM (1.8 m; ±15%) (shielded and corrected[c])	Nov 2016–Dec 2020	Flat grassy surface, off glacier

[a] Calculated by dividing the geopotential height by the standard acceleration due to gravity $g_0 = 9.81$ m s^{-2}.
[b] Artificially ventilated during daytime.
[c] Precipitation data have been corrected for noise and undercatch (section 1 in the online supplemental material; Førland et al. 1996).
[d] For specific humidity q, the period is restricted to November 2013–November 2020, because P_a (Mera La) is used to compute q.

TABLE 2. Statistical metrics used to compare ERA5L and HARv2 reanalysis variables with observed meteorological variables.

Statistical metrics	Formula[a]	Perfect value
Correlation coef (r)	$\sum_{i=1}^{n}(x_i-\bar{x})(y_i-\bar{y})\Big/\left[\sqrt{\sum_{i=1}^{n}(x_i-\bar{x})^2}\sqrt{\sum_{i=1}^{n}(y_i-\bar{y})^2}\right]$	−1; 1
Bias	$(1/N)\sum_{i=1}^{N}(y_i-x_i)$	0
SD	$\sqrt{(1/N)\sum_{i=1}^{N}[(x_i-\bar{x})-(y_i-\bar{y})]^2}$	0
RMSE	$\sqrt{(1/N)\sum_{i=1}^{N}(y_i-x_i)^2}$	0

[a] The subscript i stands for daily means, obtained from hourly values. When the hourly values from the AWS are missing, the corresponding hourly values from ERA5L and HARv2 are discarded. The overbar (i.e., $\bar{x}$ and $\bar{y}$) stands for mean values over the entire studied period or over each season.

time (UTC + 5.75 h) to match the observed data and are removed when there is a gap in the measurements (section 1 in the online supplemental material).

Four different performance indicators are used to evaluate daily mean values, of the different reanalysis variables listed above: Pearson correlation coefficient r, bias, standard deviation (SD), and root-mean-square error (RMSE) (Table 2); r is a measure of linear correlation between reanalyzed and observed data, bias is the mean value of the difference between those two sets of data and is also called systematic error, SD is the dispersion of the error, and RMSE is the standard deviation of the error (Tetzner et al. 2019; Sanz Rodrigo et al. 2013). Also, the subdaily bias for different seasons and the mean seasonal bias are calculated using hourly data and daily data, respectively. To mitigate the artificial increase in r when comparing two series with a strong seasonal cycle, we also calculate series of anomalies as defined in section 3 in the online supplemental material.

To account for the elevation difference between reanalysis grid cells and AWS sites (Table 1), we apply various elevational gradients of air temperature above ground from Kattel et al. (2013) and Immerzeel et al. (2014). They are always shallower than the standard environmental lapse rate (-6.5×10^{-3}°C m^{-1}). The mean seasonal elevational gradients, derived by Kattel et al. (2013) provide the best match with measurements at each AWS and are considered in this study. Reanalyzed air temperature T_R (°C) at the elevation of the grid point z_{gp} is shifted to the elevation of the AWS, z_{AWS}, as

$$T_R(z_{AWS}) = T_R(z_{gp}) + LR \times (z_{AWS} - z_{gp}), \quad (1)$$

where LR is the elevational gradient of air temperature above ground from Table 3 of Kattel et al. (2013). The LR varies between minimal values of -5.9×10^{-3}°C m^{-1} and -5.3×10^{-3}°C m^{-1} during the premonsoon and the postmonsoon, respectively, and maximal values of -4.7×10^{-3}°C m^{-1} and -5.1×10^{-3}°C m^{-1} in winter and monsoon, respectively.

We calculate specific humidity both for the observation and reanalysis datasets. For the AWS measurements, air temperature T (°C), relative humidity RH (%), and atmospheric pressure P_a (hPa) are used to calculate specific humidity q (g kg^{-1}) for the Mera La and Mera Summit sites. At Naulek and Changri Nup AWSs where there is no barometer, we use the atmospheric pressure recorded at Mera La because all three sites are almost at the same elevation (less than 10 m of difference; Table 1), and horizontal gradients of atmospheric pressure are small over a few kilometers (Naulek and Changri Nup are 1.5 and 30 km away from Mera La, respectively). At each AWS site, q (g kg^{-1}) is calculated as

$$q = \frac{\varepsilon e}{P_a - (1-\varepsilon)e}\times10^{-3}, \quad (2)$$

where ε, the ratio of molar mass of water to that of dry air, is equal to 0.622. The actual vapor pressure e (hPa) is calculated from saturation vapor pressure e_s (hPa) and RH (%) as

$$e = \frac{RH\times e_s}{100}, \quad (3)$$

with

$$e_s = 6.1094\exp\left(\frac{17.625T}{243.04+T}\right). \quad (4)$$

For ERA5L, q is calculated from Eqs. (2)–(4), with RH (%) obtained following Alduchov and Eskridge (1996) using air temperature T (°C) and dewpoint temperature T_d (°C) from ERA5L:

$$RH = 100\left[\frac{\exp\left(\frac{17.625T_d}{243.04+T_d}\right)}{\exp\left(\frac{17.625T}{243.04+T}\right)}\right]. \quad (5)$$

For HARv2, q is obtained from the mixing ratio r_h (g kg^{-1}) as

$$q = \frac{r_h}{(1+r_h)} \approx r_h. \quad (6)$$

To convert the wind fields obtained from the reanalyses to those comparable to in situ measurements, we use the following equation providing the total wind speed u (m s^{-1}) as a function of the zonal (U) and meridional (V) wind velocities:

$$u = \sqrt{U^2+V^2}. \quad (7)$$

In addition, to directly compare both datasets, a height adjustment is applied to convert the 10-m reanalyzed total wind

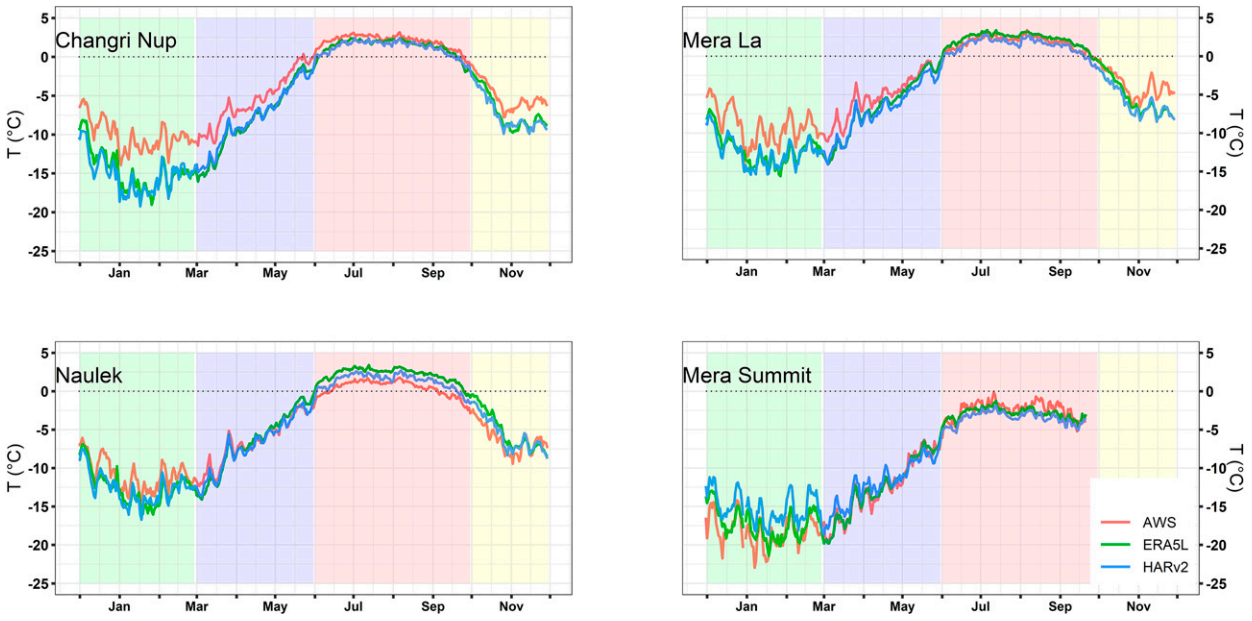

FIG. 2. Mean annual cycle of daily air temperature from ERA5L, HARv2, and AWS at different sites. The referred periods used to calculate the mean annual cycle are reported with the AWS site name in Table 1. Shaded areas correspond to seasons: winter (green), premonsoon (blue), monsoon (red), and postmonsoon (yellow).

speed to the AWS measurement height, assuming neutral atmospheric conditions (e.g., Oke 2002):

$$u_h = u_{10}\frac{\ln(z_h) - \ln(z_0)}{\ln(z_{10}) - \ln(z_0)}, \tag{8}$$

where u_{10} (m) is the reanalyzed wind speed at $z_{10} = 10$ m, u_h (m) is the wind speed at height z_h and z_0 is the aerodynamic roughness length [assumed to be 10^{-3} m above snow/ice surfaces and 10^{-2} m elsewhere, that is, debris cover or bare rocky ground (e.g., Miles et al. 2017)]. The assumptions of considering neutral conditions or using usual roughness lengths over AWS surfaces are questionable in a mountain environment because such conditions are seldom encountered or because roughness lengths may vary in space and time. Nevertheless, that is the best compromise we can find to adjust the reanalyzed wind speed to AWS heights. The resulting decrease in wind speed is weak (<23% of the 10-m wind speed).

4. Results

a. *Meteorology of the upper Dudh Koshi basin*

Our dataset corroborates the course of the seasons already observed previously (Shea et al. 2015; Sherpa et al. 2017). The winter is cold ($T < -8$°C, above 5350 m MSL) (Figs. 2 and 3), very dry ($q < 1.0$ g kg^{-1} and RH < 40% at all sites) (Figs. 4 and 5, along with Figs. S9 and S10 in the online supplemental material) and windy especially at high elevation ($u = 6.5$ m s^{-1} at 6352 m MSL) (Figs. 6 and 7). SWin and LWin have their minimum annual values (205–223 W m^{-2} and 148–182 W m^{-2}, respectively) (Figs. 8–11), in relation with low solar angle and the low cloudiness of this high-altitude dry and cold atmosphere, respectively (Table 3). At all AWS sites, less than 10% of precipitation fall during this season (Table 3; Fig. 12). At the beginning of March, as the atmosphere starts to slowly warm up, the moisture content increases, leading to progressively cloudier conditions and more frequent precipitation events until the monsoon starts in June. During the premonsoon, wind speed gradually decreases, and SWin is maximal (254–331 W m^{-2}), due to high solar angle and still frequently clear atmosphere, at least in the morning. The monsoon is warm (93% of the days have a positive daily air temperature at 5350 m MSL), constantly humid ($q > 6.8$ g kg^{-1} and RH > 85% at 5350 m MSL) and the wind speed remains low (u is close to 1 m s^{-1} at 5350 m MSL). Permanently overcast conditions prevail, reducing SWin to the benefit of LWin, which reaches its maximal values (LWin > 300 W m^{-2} at 5350 m MSL). It rains or snows on 4 of every 5 days. Daily precipitation exceeds 1 mm water equivalent (w.e.) on approximately 2 of every 3 days. It is noteworthy to mention that precipitation occurs mostly during the late afternoon and night (Perry et al. 2020). Suddenly, in less than 1 week usually at the end of September, the monsoon stops and conditions switch to the postmonsoon, warmer and slightly less dry but otherwise similar to winter, that is, quasi absence of precipitation and strengthening of western winds.

The meteorology at Changri Nup AWS located over the debris-covered part of west Changri Nup Glacier, in the upper Khumbu Valley is very similar to that of Mera La or Naulek AWSs, located over bare ground or over the debris-free ablation zone of Mera Glacier, respectively, in Hinku/Hunku Valleys at the same elevation (~5350 m MSL) (Figs. 2–13, Tables 3 and 4). This means that meteorological conditions are homogeneous in both valleys at glacier elevations, more than 30 km

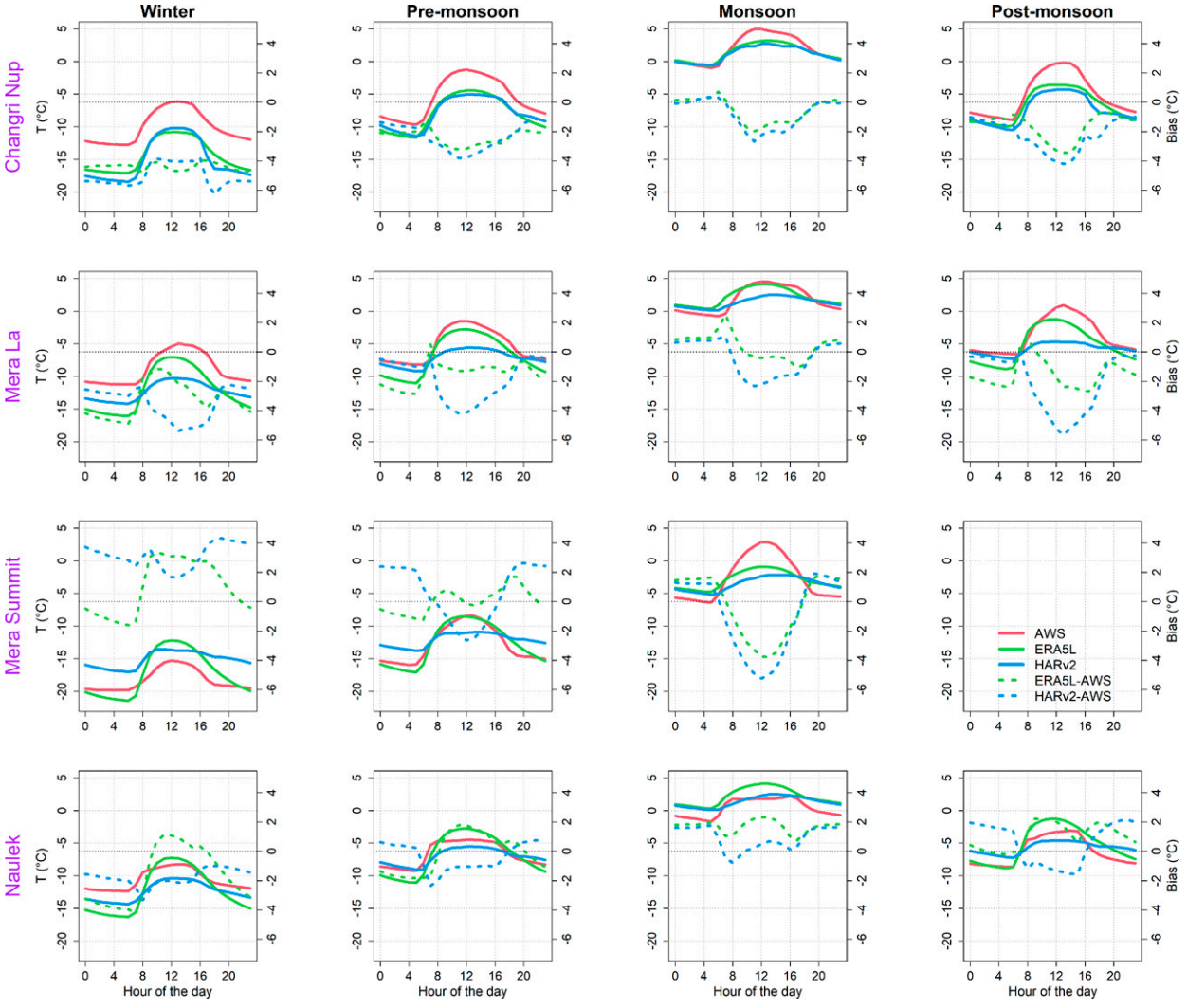

FIG. 3. Mean diurnal cycle of reanalyzed and observed 2-m air temperature (solid lines) and biases (dotted lines) at different AWS sites for different seasons.

apart. The main difference comes from the total amount of precipitation, ~30% higher in upper Hinku than in upper Khumbu because of the orographic effect leading to a strong negative horizontal gradient of annual precipitation in south-to-north direction across the range (Sherpa et al. 2017) (Table 4). However, we note that, at similar elevations, air temperature is on average 1.5°C lower and wind speed is 0.5 m s^{-1} higher over clean-ice glaciers (Naulek) than over debris-covered glaciers (Changri Nup) or bare rocky ground (Mera La) (Table 3).

Over the 7-yr (2012–19) period, annual precipitation is 591 and 540 mm w.e. at Pyramid and Pheriche, respectively, with 70% and 62% of precipitation occurring during the monsoon (Table 4). Between 2016 and 2020, precipitation is higher at Khare with a mean annual total of 818 mm w.e., 70% of which falling during the monsoon (Table 4). Pheriche always receives less precipitation than Pyramid during the monsoon as well as during the year except for 2012/13 and 2014/15, two exceptional years, the former impacted by Typhoon Phailin in mid-October 2013 (Shea et al. 2015), and the latter with exceptional precipitation events in winter and premonsoon 2014/15, for example, ~120 km farther west in the Langtang Valley as already noticed by Fujita et al. (2017). If we discard the three years 2012–15, three-quarters of annual precipitation falls during the monsoon at Pyramid and Pheriche. The depletion in precipitation between the upper Hinku Valley and the upper Khumbu Valley at similar elevations, that is, Khare (4888 m MSL) and Pyramid (5035 m MSL) sites, respectively, is 28% at annual scale as well as during the monsoon (Table 4).

The continuous record of data at Mera La, Pyramid, Pheriche and Khare without any gaps since November 2013 (November 2016 for Khare) allows for a year-by-year comparison. The years 2015/16 and 2013/14 are the warmest ($T = -3.0$°C) and coldest ($T = -4.1$°C), with 123 and 115 days of daily temperature above the freezing point at 5350 m MSL, respectively (Table 4).

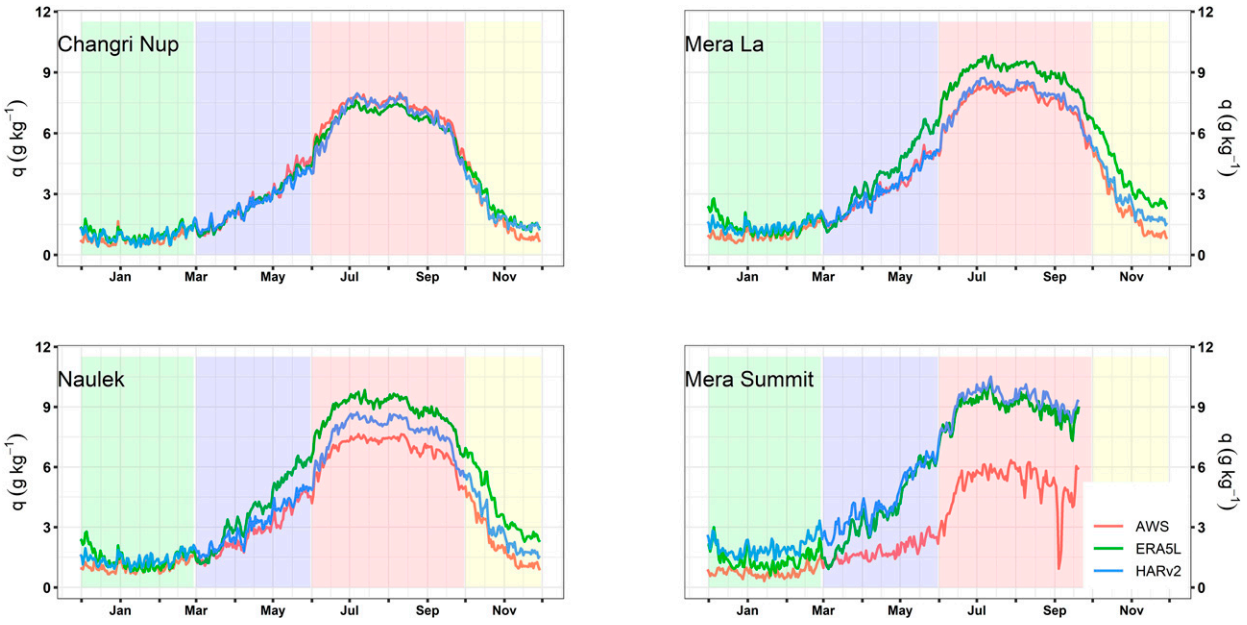

FIG. 4. As in Fig. 2, but for specific humidity.

The 2020 monsoon is the warmest, with a mean temperature of 2.3°C, 0.5°C above the 2013–20 mean. During the warmest year 2015/16, SWin is the lowest at annual and monsoon time scale (205 and 152 W m^{-2}, respectively) of the 7-yr studied period, and the annual and monsoonal precipitation is the highest (633 and 491 mm w.e. at Pyramid, respectively). The other variables are not significantly different than those of the other years. The year 2017/18 is the driest of the series, with only 539 mm w.e. at Pyramid (Table 4). Nevertheless, the 2015 monsoon is the driest of our 7-yr studied period, with only 332 mm w.e. at Pyramid. The year 2014/15 is unusual, with only 57% of the annual precipitation during the monsoon, due to exceptional events in winter and premonsoon 2014/15 (Fujita et al. 2017).

b. *Comparison between reanalysis and observed data in the upper Dudh Koshi basin*

Figures 2, 4, 6, 8, 10, and 12 show the mean annual cycle of daily in situ and reanalysis data of air temperature, specific humidity, wind speed, incoming shortwave and longwave radiations, and precipitation (at monthly scale for this latter variable), respectively. The mean annual cycle is obtained by averaging, for each date of the year, the daily values of this given date of each year of the complete measurement period. Figures 3, 5, 7, 9, 11, and 13 show the mean daily cycle during each season. Table 5 provides the values of each statistical metrics, r, bias, SD, and RMSE obtained by comparing in situ and reanalysis data at daily time scale, over the total measuring period for each AWS. Table S1 in the online supplemental material is the counterpart of Table 5 for the series of anomalies. Table 6 gives the seasonal bias between reanalysis and observed data of those six meteorological variables.

1) AIR TEMPERATURE

Over the whole measuring period, reanalyzed and observed daily temperatures are highly correlated [$r > 0.95$, $p < 0.001$; $r > 0.72$ for the anomalies, see section 3b(i) and Table S1, both in the online supplemental material] whichever sites or reanalyzed data are considered (Table 5; Fig. 2). Seasonal and daily cycles of temperature are well reproduced in both reanalysis datasets (Figs. 2 and 3). Reanalyzed temperatures are usually lower than in situ data during nonmonsoonal season (except Mera Summit). This cold bias is less important over debris-free glacier areas (Naulek and Mera Summit) than over debris-covered or rocky surfaces (Changri Nup and Mera La, respectively).

For the four seasons, the reanalyzed data are the closest to observations in monsoon and the farthest in winter, except at Naulek for ERA5L, where the reanalyzed data are the closest to observations during intermediate seasons (Table 6). As for annual values, the differences between reanalyzed and in situ temperatures are larger over rocky surfaces (debris-covered glacier or bare ground), with a negative bias always exceeding 3.2°C in winter for ERA5L or HARv2, than over clean glacier surfaces. The bias over clean-ice surfaces is mostly negative but does not exceed 1.8°C except for HARv2 in winter where it is positive (3.1°C at Mera Summit).

At subdaily time scale (Fig. 3), the differences between reanalyzed and observed temperatures are very dependent on whether it is night or daytime. The bias is usually slightly lower at night than during the day (i.e., monsoon for both reanalyzed data) but large biases, often negative, can still be observed at night, as in winter for ERA5L and HARv2. It is noteworthy to mention that at night, the bias may be positive for HARv2 and at the same time negative for ERA5L as in winter at Mera Summit (Fig. 3). We usually observe a mean

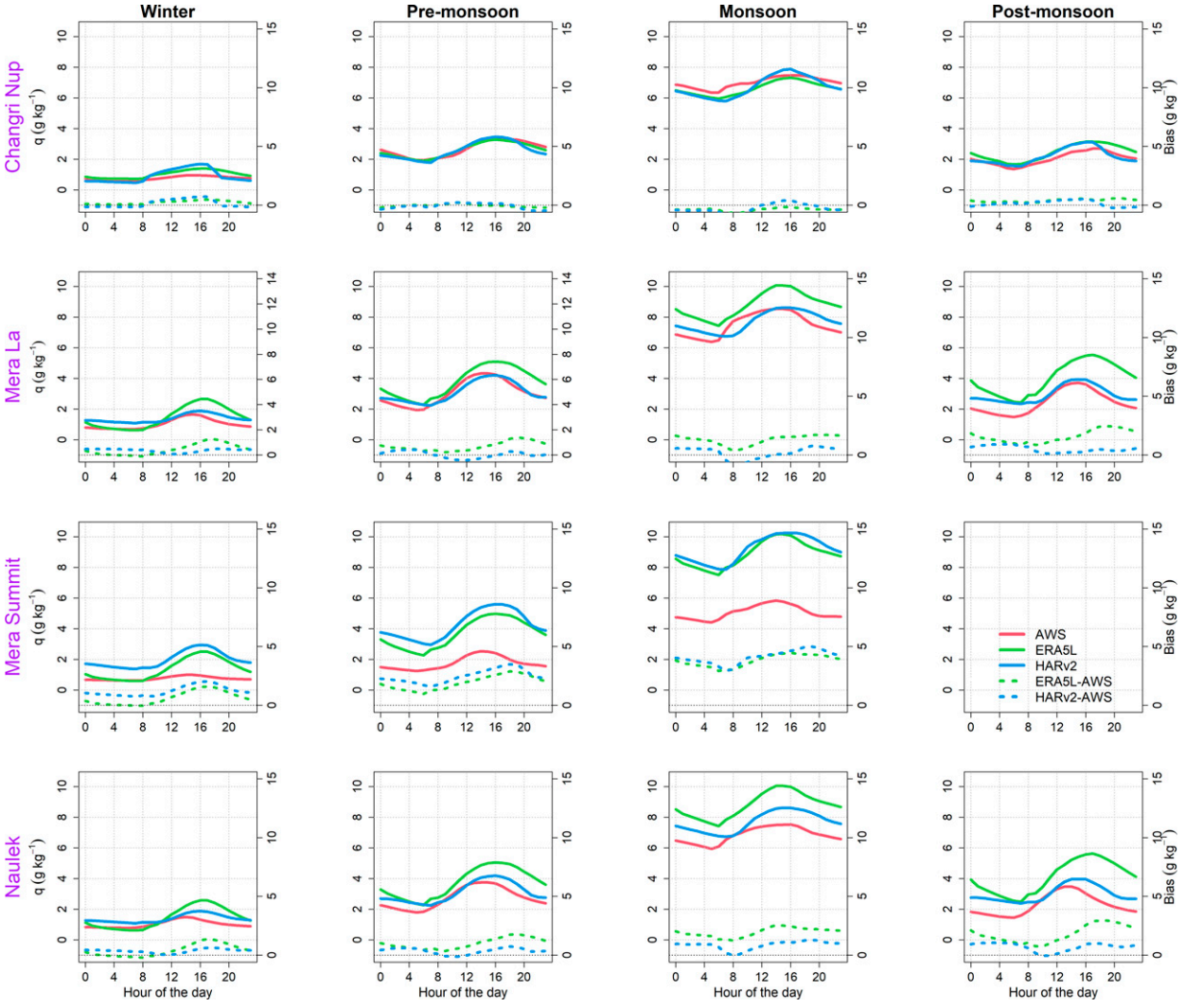

FIG. 5. As in Fig. 3, but for specific humidity.

daily cycle of the bias, with a minimum at midday especially for HARv2 data. When this bias is close to zero or negative at night, which is often observed, it is strongly negative at daytime. However, such mean daily cycle of the bias is not always observed, as for ERA5L in winter or premonsoon, where the bias passes through a maximum at daytime.

Considering all statistical metrics (Tables 5 and 6; Figs. 2 and 3; Table S1 in the online supplemental material), the performances of ERA5L and HARv2 reanalysis datasets are not significantly different, but large biases exceeding 4°C, mostly negative, are only encountered for HARv2 at any season and at all sites except Naulek. Both reanalyzed datasets still tend to underestimate air temperature at high elevation in the upper Dudh Koshi basin, especially over rocky surfaces.

2) SPECIFIC AND RELATIVE HUMIDITIES

At the annual time scale, both reanalyzed and observed daily specific humidities are very well correlated ($r > 0.97$, $p < 0.001$; $r > 0.62$ for the anomalies) except at Mera Summit where the correlation is lower especially for the anomalies ($r = 0.92$ and 0.93 for ERA5L and HARv2, respectively, $p < 0.001$; $r = 0.27$ and 0.42 for the anomalies for ERA5L and HARv2, respectively) and the bias is large (bias > 2.1 g kg^{-1} for both reanalyses) (Table 5, along with Table S1 in the online supplemental material). The seasonal cycle of specific humidity is well reproduced by both reanalyses, with an almost perfect match at Changri Nup for both reanalyses and at Mera La for HARv2 (Fig. 4; Tables 5 and 6). Looking at the anomalies [sections 3b(ii) and 3b(iii) in the online supplemental material], HARv2 performs slightly better than ERA5L, but both reanalysis products are not very good at Mera Summit, where they predict saturation during the monsoon although this is not the case in reality (Fig. S9 in the online supplemental material), leading to an overestimation of specific humidity (Fig. 4). The bias, SD, and RMSE are higher at Mera Summit than at lower-elevation AWSs mainly because of the larger difference in elevation and

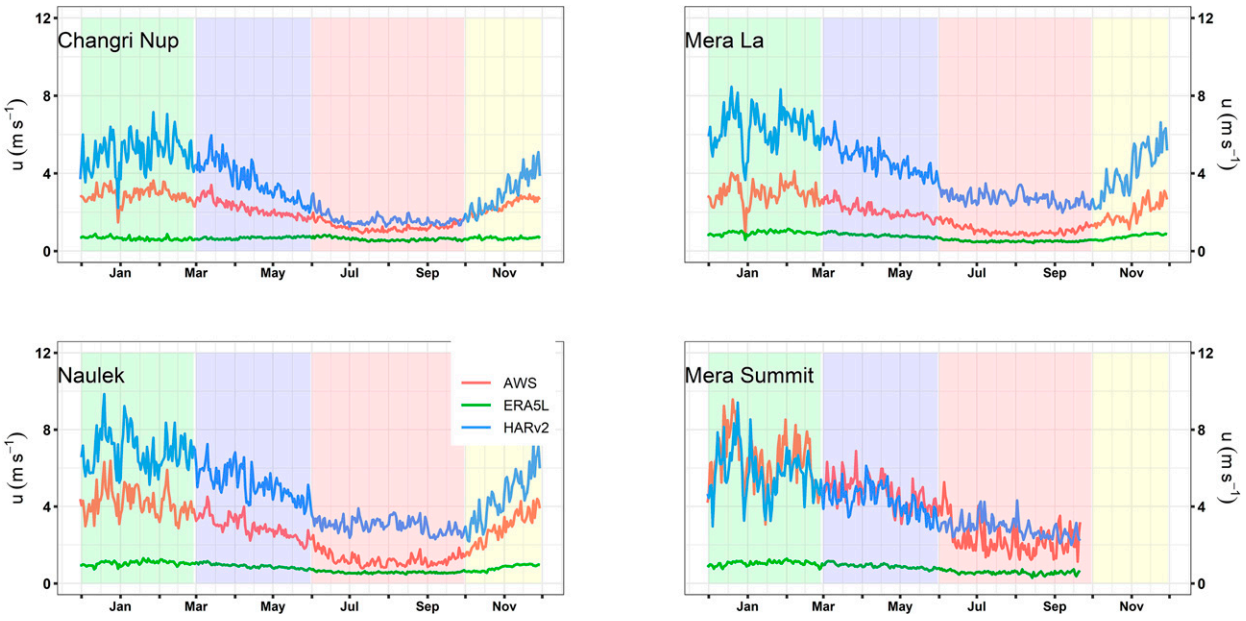

FIG. 6. As in Fig. 2, but for wind speed.

in turn in atmospheric pressure between the AWS and the grid cells of the respective reanalysis datasets. To a lesser extent, this is also the case at Mera La and Naulek sites for ERA5L data, because the gridcell altitude is approximately 700 m below that of the AWSs (Table 1), inducing a higher surface atmospheric pressure and, as a consequence, higher reanalyzed specific humidity relative to in situ values.

At the subdaily time scale, we observe a daily cycle of specific humidity, with most of the time a minimum value in the early morning, and a maximum in the early afternoon (Fig. 5). This cycle is well reproduced by the two gridded datasets, sometimes almost perfectly, as during premonsoon at Changri Nup. We often observe a slight positive humid bias usually a little more important in the afternoon, especially for ERA5L (Fig. 5).

Considering all statistical metrics (Table 5; Figs. 4 and 5; Table S1 in the online supplemental material), we can conclude that the seasonal and the daily cycles of specific humidity are well represented by both gridded datasets, with still a large humid bias at Mera Summit. Overall, for specific humidity, HARv2 is slightly closer to the observations than ERA5L.

We performed a similar analysis for relative humidity (see section 2 in the online supplemental material). Relative humidity is systematically slightly overestimated by reanalysis data, and better represented by HARv2 than ERA5L when considering the seasonal cycle of RH, and vice versa when looking at its daily cycle. Nevertheless, the agreement between observed and reanalyzed relative humidity is still not very good at high elevation (Mera Summit), except in winter.

3) WIND SPEED

Over the entire measuring period, the correlation between reanalyzed and observed daily wind speed is very variable, from low to high [$r = 0.15$–0.82, $p < 0.001$; $r = 0.13$–0.66 for the anomalies; see section 3b(iv) in the online supplemental material] (Table 5, along with Table S1 in the online supplemental material). Those correlations, however, do not reflect the fact that both reanalyses do not represent well the local wind speed at all sites, because the wind speed is mainly controlled by local topography and surface conditions that cannot be accounted for correctly in reanalysis products.

Overall, HARv2 overestimates wind speed by a factor of 1.5–3 for all sites located at ~5350 m MSL but is really close to the AWS measurements at Mera Summit, a high-altitude open site where conditions are likely close to the free atmosphere. However, even at this very high site, the wind speed is a little overestimated during the monsoon (bias = 0.5 m s^{-1}) when there are light winds ($u = 2.5$ m s^{-1}), and slightly underestimated the rest of the year (Tables 5 and 6; Fig. 6).

On the other hand, ERA5L strongly underestimates wind speed at all sites, except during the monsoon (bias ranging from −4.0 to −1.2 m s^{-1}; Table 5). Moreover, ERA5L shows neither any pronounced seasonality, as at Changri Nup, where wind velocity is almost the same year-round, nor any day-to-day variability, although this is often observed especially in winter at all sites (Fig. 6). The only season when ERA5L performs fairly well is the monsoon, at all sites except Mera Summit, even though it still fails to reproduce the day-to-day variability (Table 6; Fig. 6).

Both reanalysis datasets fail to reproduce the subdaily cycle of wind speed. We usually do not observe any clear daily cycle of wind speed, except a slight decrease of wind speed in the morning or during the daytime, over clean-ice areas, more pronounced at very high elevation (Mera Summit). This daily cycle is not reproduced by reanalysis data at all, and HARv2 data even show an opposite cycle, with increasing wind velocity during the day (Fig. 7).

FIG. 7. As in Fig. 3, but for wind speed.

In conclusion, none of the reanalysis datasets reproduce observed wind speed at high elevation in this region, partly because local wind speed is very dependent on local topography and surface conditions, but not only because the performance is barely better on high-altitude open sites. HARv2 reproduces well the seasonality of wind speed observed at AWSs but tends to overestimate it, and completely fails to reproduce its observed daily cycle. ERA5L systematically underestimates the wind speed at all elevations and does not even show any seasonality.

4) INCOMING SHORTWAVE RADIATION

At annual time scale, reanalyzed and observed daily values are fairly or well correlated, with r values higher than 0.62 and even often close to 0.77 (Table 5, except Mera Summit). Even when removing the seasonal cycle, r values stay high and close to 0.60 [section 3b(v) and Table S1, both in the online supplemental material]. However, these relatively high correlations should not hide the fact that the reanalyses do not represent the cloudiness very well, and, as a consequence, have degraded evaluation metrics during the monsoon. More precisely, during the winter and the postmonsoon, both reanalyzed incoming radiations have a negative bias not exceeding 16 W m^{-2} (winter) or 30 W m^{-2} (postmonsoon), for ERA5L at all sites (Table 5). However, as soon as it starts to be cloudy during the premonsoon and the monsoon, this bias tremendously increases especially for HARv2 data (Table 6; Fig. 8). This bias is most of the time negative and exceeds sometimes 100 W m^{-2} during the monsoon (Table 6; Fig. 8).

Both reanalyses fail to reproduce incoming shortwave radiation at Mera Summit except in winter, although this site receives the highest solar radiation of all, due to its very high altitude and large sky view factor. During the monsoon, HARv2 and ERA5L underestimate SWin by a factor of 6 and 2, respectively. This underestimation is also mainly due to atmosphere attenuation given that the elevations of the reanalysis grid cells (4662 m MSL for ERA5L, and 4637 m MSL for

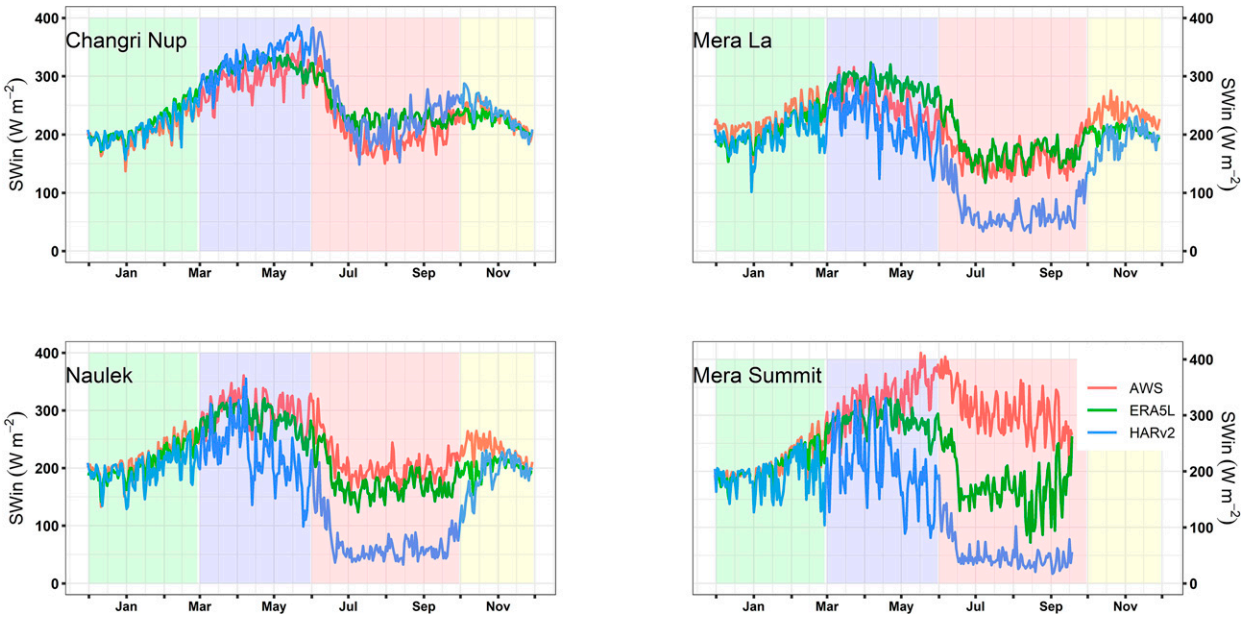

FIG. 8. As in Fig. 2, but for SWin.

HARv2) are much lower than the AWS site elevation (6352 m MSL for Mera Summit) (Table 1). The only site where both reanalyses are close to the observations is Changri Nup, although HARv2 is less good than ERA5L, and this time overestimates SWin during the premonsoon and monsoon (Tables 5 and 6; Fig. 8). Overall, the performance of ERA5L is slightly better than HARv2 at all sites but still poor at very high altitude when there are clouds (premonsoon and monsoon). This performance strongly depends on the elevation difference between the observed site and the corresponding reanalysis grid cell.

The mean daily cycle of the bias is sometimes asymmetrical, negative in the morning and positive in the afternoon, as in winter at all sites, or on Changri Nup and Naulek, because those sites are oriented east, and in the shade early in the afternoon (1400–1500 local time) (Fig. 9). The surrounding complex topography is responsible for some shading effects blocking direct sunlight at AWS sites, which explains why the bias between reanalyzed and observed data can be variable during the day.

5) INCOMING LONGWAVE RADIATION

Over the whole measuring period, daily reanalyzed and observed values are extremely well correlated, with r higher than 0.89 ($p < 0.001$) for all sites (Table 5). When removing the seasonal cycle, this correlation stays high, with r values higher than 0.70 except at Mera Summit where r drops to 0.55 for ERA5L data [section 3b(vi) and Table S1, both in the online supplemental material]. This is somehow expected because correlations for T and q are also high. Overall, ERA5L performs very well at all sites, as shown by the good performance of all statistical metrics (Table 5), with a mean bias usually not exceeding 16.1 W m^{-2} (except Mera Summit where the bias is as high as 40 W m^{-2}), negative at Changri Nup, positive elsewhere (Table 6; Fig. 10). As for HARv2 data, the bias is most of the time more important, usually exceeding 20 W m^{-2}, positive at Mera Summit, negative elsewhere. As for SWin, this bias strongly depends on the elevation difference between the observation site and the reanalysis grid cell. For instance, the positive bias at Mera Summit is mainly explained by elevation difference between the HARv2 or ERA5L grid point and the observation site located ~1700 m higher (Table 1).

The seasonality is also very well reproduced. The seasonal bias between ERA5L and observed LWin is reasonable, and in worst cases reaches 20–50 W m^{-2}, sometimes positive, sometimes negative, depending on the site and season, as well as the elevation difference between AWS and grid cell (Table 6; Fig. 10). For HARv2 data, the bias is usually higher, positive at Mera Summit mainly due to this elevation difference, and negative elsewhere.

The daily cycle of reanalyzed data agrees usually well with the observed daily cycle, still with a slight shift of the daily maximum observed at 1300 LT, but occurring later for ERA5L, and earlier for HARv2 (Fig. 11).

6) PRECIPITATION

Over the 2012–19 measuring period, daily reanalyzed and observed precipitation are well correlated ($r > 0.60$, $p < 0.001$) (Table 5). The seasonal distribution of precipitation is well reproduced. Indeed, observations show that over this 7-yr period, 70% of annual precipitation fall during the monsoon at Pyramid (62% at Pheriche), while ERA5L and HARv2 give 74% and 65%, respectively (76% and 72%, respectively, at Pheriche). Similarly, between 2016 and 2020, Khare AWS collected 28% more precipitation than at Pyramid, 70% of which falling during the monsoon, while ERA5L and HARv2 give

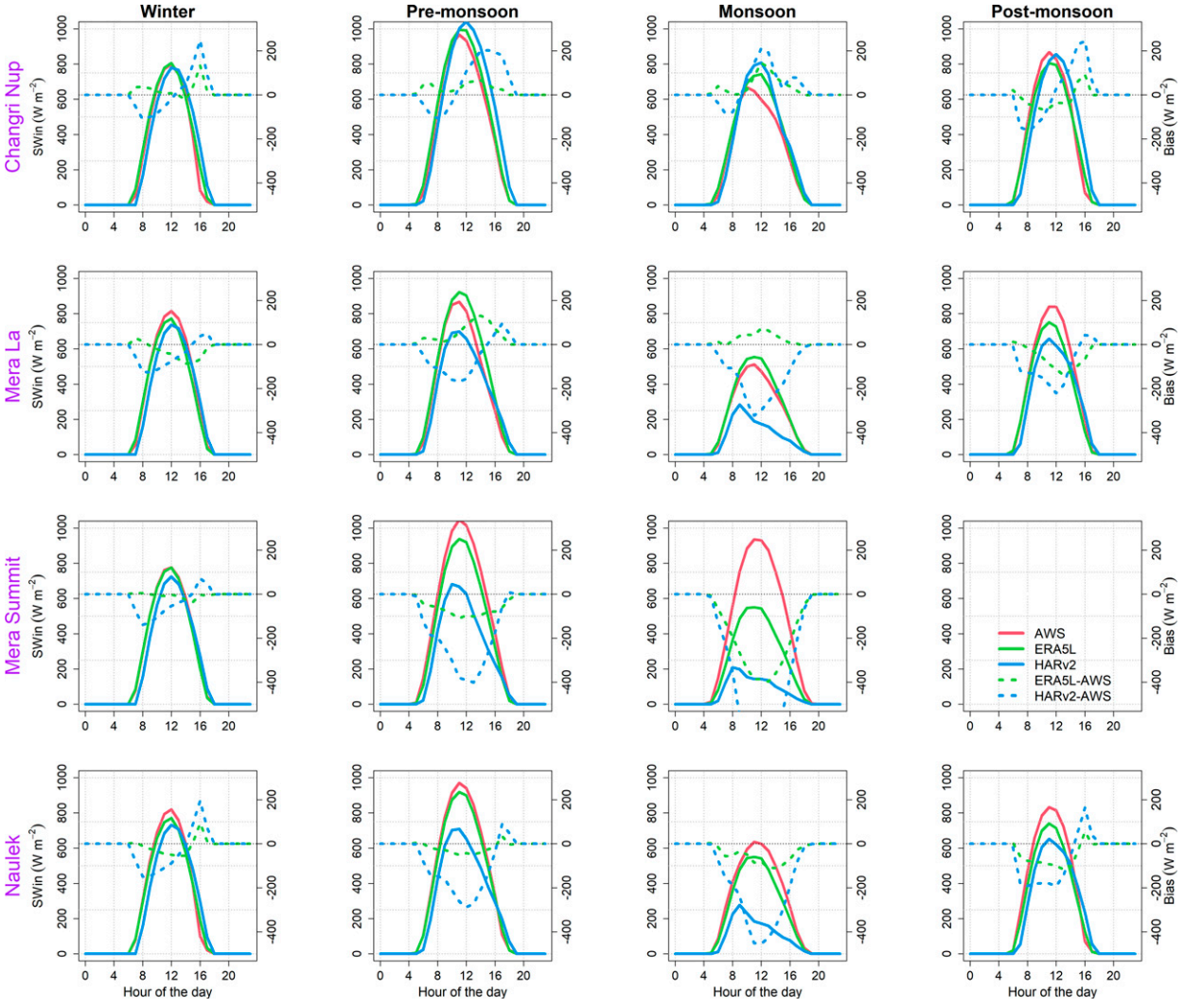

FIG. 9. As in Fig. 3, but for incoming shortwave radiation.

63% and 66% more precipitation than at Pyramid, with 79% and 75% falling during the monsoon, respectively. However, both reanalyses strongly overestimate precipitation at all sites, with a mean positive bias of 0.5–0.9 mm w.e. day^{-1} at Pyramid for ERA5L and HARv2, respectively, and 2–3 times higher at Pheriche and Khare, respectively (Table 5). Overall, the annual totals of precipitation estimated by both reanalysis products exceed by a factor of 2–3.5 the observed precipitation, except for ERA5L at Pyramid where the overestimation is only 30%. Those biases are even more important during the monsoon and exceed 1.1 mm day^{-1} at Pyramid or are even larger at Pheriche or Khare with values reaching 6.5 mm day^{-1} for HARv2 (Table 6).

The comparison at monthly time scale (Fig. 12) confirms that reanalysis products tend to systematically overestimate observed precipitation, but the pattern is very dependent on the site and on the season. At all sites, the agreement is good during the driest months, that is, during the postmonsoon and in winter, with still an overestimation of precipitation in January–February for HARv2 data (Table 6). During the premonsoon, in the driest part of our studied area (i.e., Pheriche and Pyramid), precipitation is well represented by both reanalysis products. At Khare, precipitation is overestimated by both reanalyses, more importantly by HARv2, when the monsoon gets closer (April and May). Monsoon is the season when the positive bias is systematic and the highest, with monthly precipitation being often overestimated by a factor of 1.6–3.8, especially in the wettest part of our studied area (Khare). Only ERA5L data represent fairly well precipitation recorded at Pyramid, with an overestimation of only 37% during the monsoon.

Figure 13 compares the mean daily cycle of observed precipitation with that of reanalyzed data, for each season. Days with less than 1 mm w.e. of observed daily precipitation are discarded in these daily cycles. As already observed in other studies (e.g., Ueno et al. 2008; Yamamoto et al. 2011;

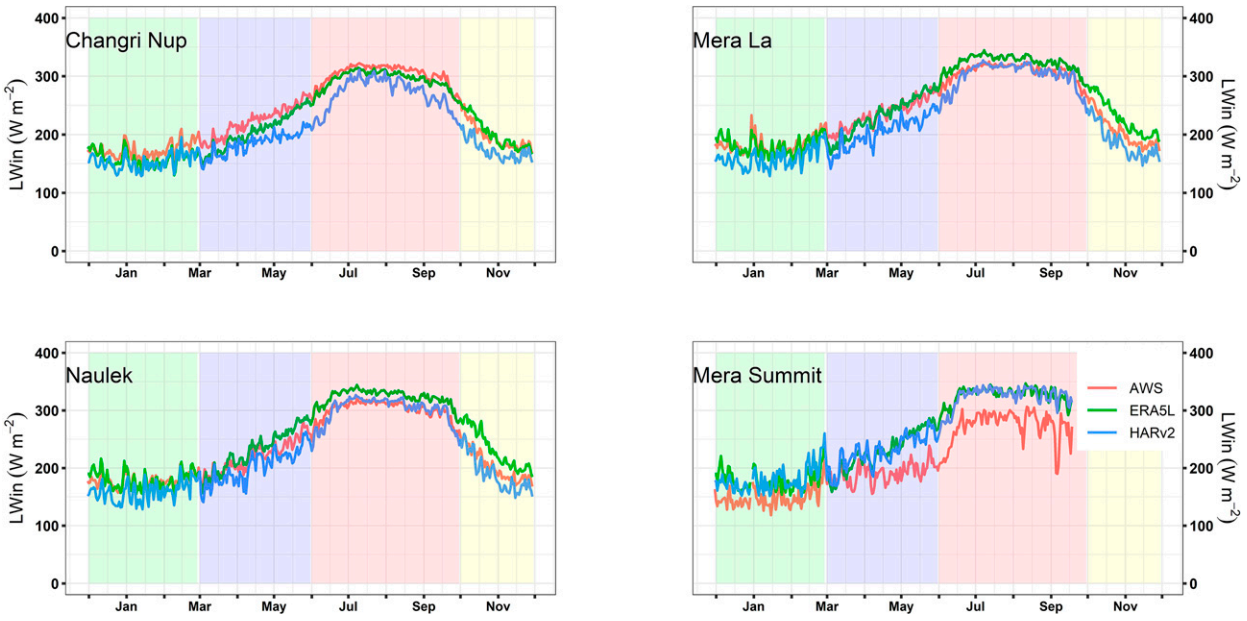

FIG. 10. As in Fig. 2, but for LWin.

Perry et al. 2020), during the monsoon, precipitation is more intense at night than during daytime, with a peak of precipitation at sunset at Pyramid or later during the night at the other sites. The pattern is similar during the premonsoon, but less clear. During the other seasons, given that there are only a few days with precipitation exceeding 1 mm w.e. day^{-1}, we cannot extract any clear and representative pattern of daily distribution of precipitation. Focusing on the two most humid seasons (i.e., monsoon and premonsoon), both reanalyses are unable to reproduce the daily cycle of precipitation and are most of the time out of phase with maximum precipitation intensity at daytime. This is especially true at Khare where strong precipitation maxima at noon, concomitant to the observed precipitation minimum, are responsible for large precipitation overestimations already reported.

Tables 7 and 8 compare the occurrence of precipitation at AWS sites with that of reanalysis products, at hourly and daily time scales, respectively, over the entire measuring period and over the monsoons only. Overall, the agreement between observed and reanalyzed precipitation occurrences is reasonable, slightly better for HARv2. Still, 13%–36% of the time [sum of the false counts in Tables 7 and 8, 13% being for HARv2 data at Pheriche at hourly time step for the entire measuring period (Table 7) and 36% being for ERA5L data at Khare at hourly time step for the monsoons only (Table 7)], depending on the site and on ERA5L or HARv2, reanalysis data fail to reproduce precipitation, mostly simulating precipitation while there is none in reality (false positive in Tables 7 and 8), but vice versa also, not only at the hourly time scale but also at the daily time scale. Looking only at the monsoons, the days with precipitation are fairly well reproduced with a percentage of total false values not exceeding 25% (Table 8). But the agreement is less good while considering the occurrences of precipitation at hourly time scale when the total percentage of false values is between 13% and 36% (Table 7).

We also look at the two most important events of our precipitation record, that is, Typhoons Phailin and Hudhud, which both occurred in mid-October, in 2013 and 2014, respectively (Shea et al. 2015; Sherpa et al. 2017). Even though both reanalyzed datasets tend to overestimate precipitation during those extreme events, especially for HARv2 data, the daily precipitation distribution is fairly well reproduced over the three days of each event. For both typhoons, 14 October was the wettest day, receiving 55%–70% of the total precipitation of each event (Table 9).

5. Discussion

a. *Performances of ERA5L and HARv2 reanalyses in the central Himalayas*

We present here a multisite evaluation of meteorological variables using ERA5L and HARv2 reanalyses in comparison with in situ observations from seven meteorological stations located in the upper Dudh Koshi basin, Nepal. It is noteworthy to mention that reanalysis data are spatially resolved data (0.1° × 0.1° for ERA5L and 10 km × 10 km for HARv2) whereas in situ observations are point-scale data at AWS sites. Because of this scale difference, we cannot expect a perfect match between datasets, especially when the topography is complex. The performance of reanalysis data depends on the meteorological variables, on the geographical context (Tetzner et al. 2019) and on the surface state. In the upper Dudh Koshi basin, the 2-m air temperature is the best captured (r = 0.95–0.97; $p < 0.001$; RMSE of 2°–3°C; r = 0.72–0.91 for the anomalies) among all variables by both reanalyses (ERA5L and HARv2), which are nevertheless usually cold

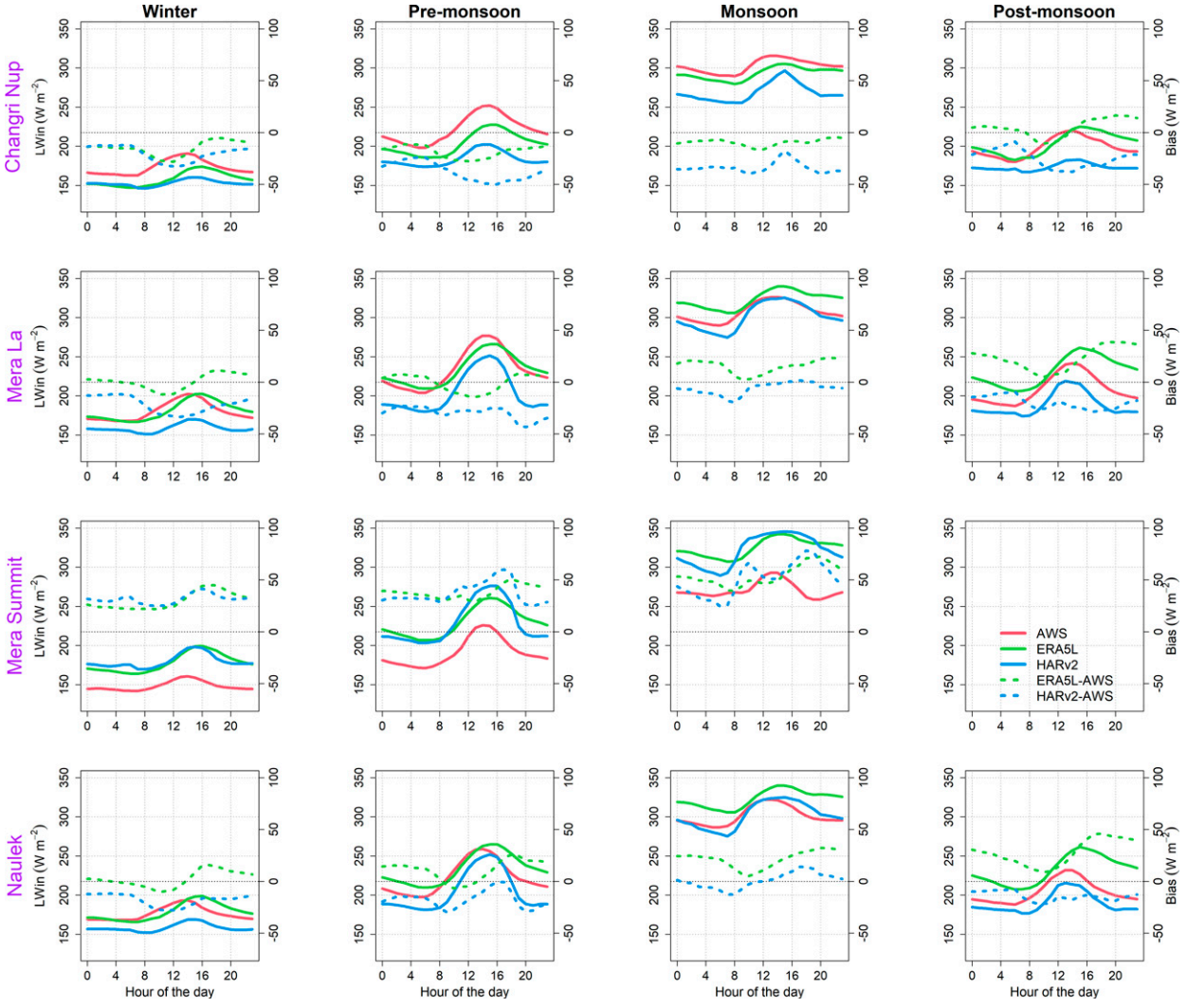

FIG. 11. As in Fig. 3, but for incoming longwave radiation.

biased especially over rocky surfaces (bias from −2.2° to −2.4°C at Changri Nup) (Tables 5 and 6). This cold bias, typical of most reanalysis products in this region (Orsolini et al. 2019), is attenuated over glacierized surfaces because they are themselves colder than the average surfaces. The variations between AWS and reanalysis data are likely to be attributed to the elevation difference between the AWS site and the grid point of the reanalysis (Tetzner et al. 2019). Even though both reanalysis temperature data are corrected using a seasonal elevational gradient of air temperature above ground (Kattel et al. 2013), the bias is stronger during the cold seasons of the year (winter and postmonsoon) than during the humid season and is also more important at daytime than during the night. During these cold seasons in this mountainous environment, an inversion layer may occur. In this case, the use of a vertical gradient of air temperature is inappropriate and may be responsible for this cold bias, especially when the altitude difference between the reanalysis grid cell and the AWS exceeds a few hundreds of meters. As already pointed out by Immerzeel et al. (2014) and Steiner and Pellicciotti (2016) in the nearby Langtang Valley, there is a great spatiotemporal variability of temperature lapse rates in Himalayan valleys due to extreme topography and the type of glaciers. Instead of a standard environmental lapse rate, it is strongly recommended to use observed lapse rates usually shallower, at high temporal resolution, that is, hourly time scale for instance (Ragettli et al. 2015; Mimeau et al. 2019). A standard correction method based on spatially and temporally variable lapse rates should be developed to correct surface temperature reanalysis data for the central Himalayas, accounting for not only the season and the hour of the day (Steiner and Pellicciotti 2016), but also the surface state and the location of the measurement site as a function of topography (leeward vs windward for instance).

Specific humidity is directly related to local convection of the region as well as the monsoonal activity. In the upper Dudh Koshi basin, the Indian summer monsoon has a significant effect from June to September, but local convection

TABLE 3. Seasonal and annual mean values of meteorological variables observed at different AWS sites, except for precipitation for which seasonal or annual totals are reported.

		Seasonal[a]				
	AWS site	DJF	MAM	JJAS	ON	Annual mean[a]/total
T (°C)	Changri Nup	−10.1	−5.5	2.0	−5.0	−3.9
	Mera La	−8.9	−5.3	1.8	−3.6	−3.5
	Mera Summit	−18.3	−12.8	−2.5	—	−11.2
	Naulek	−10.4	−6.8	0.3	−6.7	−5.4
RH (%)	Changri Nup	25	52	85	39	55
	Mera La	30	60	91	42	60
	Mera Summit	39	55	76	—	56
	Naulek	33	56	92	42	61
q (g kg^{-1})	Changri Nup	0.8	2.6	7.0	2.1	3.6
	Mera La	1.0	3.1	7.5	2.4	3.9
	Mera Summit	0.8	1.8	5.1	—	2.5
	Naulek	1.0	2.7	6.8	2.3	3.6
u (m s^{-1})	Changri Nup	2.9	2.1	1.2	2.3	2.1
	Mera La	3.0	2.0	1.1	1.9	2.0
	Mera Summit	6.5	4.8	2.5	—	4.6
	Naulek	4.6	3.0	1.4	3.3	2.6
SWin (W m^{-2})	Changri Nup	205	296	212	230	236
	Mera La	223	254	162	234	212
	Mera Summit	212	331	315	—	283
	Naulek	216	304	222	224	234
LWin (W m^{-2})	Changri Nup	173	223	303	199	233
	Mera La	182	234	308	209	243
	Mera Summit	148	191	272	—	203
	Naulek	175	219	300	197	240
P (mm w.e.)	Pyramid	39	108	413	32	591
	Pheriche	49	115	337	39	540
	Khare	78	142	570	28	818

[a] Seasonal or annual mean values are discarded when data gaps exceed 33% of the time.

already increases regularly during the premonsoon, leading to a gradual increase of specific humidity, though it is less pronounced at very high elevation (Mera Summit) (Fig. 4). At 5350 m MSL, q is very well represented by both reanalyses, although the difference in altitude and thus atmospheric pressure between the reanalysis grid and the AWS induces a wet bias as for both gridded datasets at Mera Summit, or for ERA5L at Naulek and Mera La (Fig. 4; Tables 5 and 6). The seasonality of q at Mera Summit is significantly different from what we observe at lower elevations. Indeed, at high elevation, the only humid season is the monsoon even though the atmosphere is far from being permanently saturated (Fig. S9 in the online supplemental material), the day-to-day variability is higher than at lower altitude, and the atmospheric water content slowly increases during the premonsoon (Fig. 4). It seems that there is a decoupling of the atmosphere at around 6000 m MSL during the premonsoon and the monsoon, the lower layers being much more affected by convection than the higher layers, more occasionally reached by convective clouds and in turn much drier. As a consequence, since the reanalysis grid points for ERA5L and HARv2 are located ~1700 m lower in elevation, they cannot reproduce the water content of the atmosphere at Mera Summit.

ERA5L completely fails to reproduce observed wind velocities at any elevation or over any surfaces. In addition to the large systematic underestimation of the wind velocity, this reanalysis product is not able to reproduce its large seasonal variability, characteristic of high-elevation areas of the central Himalayas. Probably due to the fact that this product is obtained by dynamical downscaling of ERA5, HARv2 reanalysis performs better, but tends to systematically overestimate wind speed at lower elevations all year round (Fig. 6). The only site where HARv2 wind speed agrees well with observations is Mera Summit, even though wind speed is overestimated during the monsoon, underestimated the rest of the year and the daily cycle is out of phase (Fig. 7). In mountain areas, wind results from the interplay of synoptic-scale circulations, like strong westerly winds in winter or during the postmonsoon, and local circulation, like katabatic winds over glaciers or valley breeze circulation systems (Bollasina et al. 2002). Consequently, 2-m wind speed is highly heterogeneous in the central Himalayas at small spatial scale due to the extreme topography, a competition between local and general circulation systems, and variable land-covered patterns. In such conditions, the poor performance of reanalysis products is expected, except at very high-elevation open sites like Mera Summit, where conditions are closer to the free atmosphere, that is, less affected by local circulation systems or surface types. In this study region, only HARv2 reanalysis is able to represent mesoscale wind speed, but downscaling techniques are necessary to resolve local-scale wind regimes like anabatic or catabatic winds. It is also noteworthy to mention that

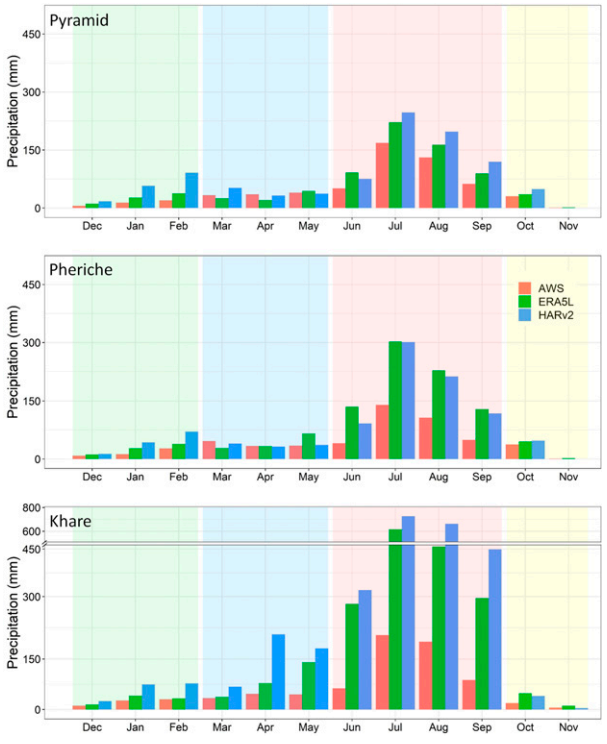

FIG. 12. Mean monthly precipitation (mm w.e. month^{-1}) at the Pyramid, Pheriche, and Khare sites. Monthly precipitation has been calculated over the period 2012–19 for Pyramid and Pheriche and 2016–20 for Khare. Note the broken scale of the *y* axis in the bottom panel, as shown by the double horizontal line.

applying a logarithmic profile to compute reanalysis wind speed at AWS measurement level from 10-m wind is likely to add further errors. Indeed, a stable stratification is common over glacier surfaces, which can result in a vertical wind gradient that is much steeper than a simple logarithmic profile, and in turn in a more pronounced reduction of wind speed. This could explain why HARv2 wind speed is often overestimated.

We observe a strong seasonality of incoming shortwave and longwave radiations, with maximal values of SWin during the premonsoon, and of LWin during the monsoon because of the presence of warm air and thick convective clouds. The seasonality of both incoming radiations is well captured by ERA5L and HARv2 reanalysis datasets (Figs. 8 and 10). However, there are sometimes large biases for some sites, mainly due to the elevation difference between the AWS and the reanalysis grid point. For some sites like Mera Summit, this elevation difference is as high as ~1700 m, explaining why reanalyzed SWin is negatively biased, especially during the premonsoon and monsoon when a large part of SWin is reflected by an extra 1700 m of cloud thickness, and why LWin is systematically positively biased, due an additional longwave emission of this extra 1700 m thick layer of atmosphere. When the reanalysis gridpoint elevation is close to that of the AWS, the agreement between reanalyzed and observed radiations is very good as at Naulek for HARv2 data. Nevertheless, on each site, ERA5L performs much better than HARv2. Indeed, HARv2 incoming shortwave radiation is abnormally depleted during the monsoon at all sites except Changri Nup, likely due to an overestimation or an exaggerated reflection of the cloud cover. In mountain areas with extreme topography, SWin and LWin are also strongly influenced by local conditions such as slope or aspect, which contribute to enhance biases. As a consequence, incoming shortwave and longwave radiations can be reconstructed from ERA5L reanalysis (and to a lesser extend also from HARv2 data but with a lower accuracy) as long as a correction method accounting for the elevation difference between the reanalysis grid point and the studied site and the local topography is considered.

Both reanalysis datasets are able to reproduce precipitation during the dry seasons (postmonsoon and winter, even though HARv2 precipitation is overestimated in January–February) but strongly overestimate precipitation when it starts to rain or snow, in premonsoon, and more evidently during the monsoon. Additionally, reanalysis products predict maximum precipitation intensity at daytime although it occurs at night. The performance of ERA5L is better than HARv2, but an overestimation by a factor of 2–3 is still usual, between June and September (Figs. 12 and 13). This is all the more problematic as more than 70% of the annual precipitation is concentrated in those 4 months. It is interesting to note that a recent study found different results when comparing HARv2 (at 10- and 2-km resolutions) and ERA5L with three stations located on the Tibetan Plateau (Hamm et al. 2020). They concluded that HARv2 had a better ability to represent the orographic effect on precipitation, and was in better agreement with station measurements than ERA5L. Multiple reasons could explain these differences: (i) there is a strong annual variability in precipitation, which might hinder systematic biases; (ii) the stations they investigate are located on the northern flank (leeward side) of the central Himalayas, and consequently the influence of topography is expected to be different; and (iii) the stations investigated in this present study are located at higher elevations than those in Hamm et al. (2020). The comparison with this study shows that it is difficult to draw general conclusions from site specific studies, and some conclusions apply only to specific settings.

It is well known that solid precipitation measurements are beset with significant inaccuracies, depending on the gauge catch efficiency (e.g., Kochendorfer et al. 2017). Nevertheless, our dataset has been corrected for undercatch [Førland et al. (1996) and section 1c in the online supplemental material] and the bias is strongest during the monsoon when conditions, that is, light winds and mostly rain below 5000 m MSL, are favorable for an efficient gauge catch. Consequently, we expect our precipitation measurements not to be underestimated. Anyway, with such a complex topography, the spatial variability of precipitation is extreme (Mimeau et al. 2019; Eeckman et al. 2017) and the spatial resolution of reanalysis products, as good as it is, is still not sufficient to reproduce the observations. The occurrence of precipitation is reasonably reproduced at least at daily time scale, but reanalysis products have still a tendency to simulate precipitation when there is not, mainly at hourly time scale (Tables 7 and 8). Promisingly, both reanalysis datasets have well captured the extreme precipitation events due to

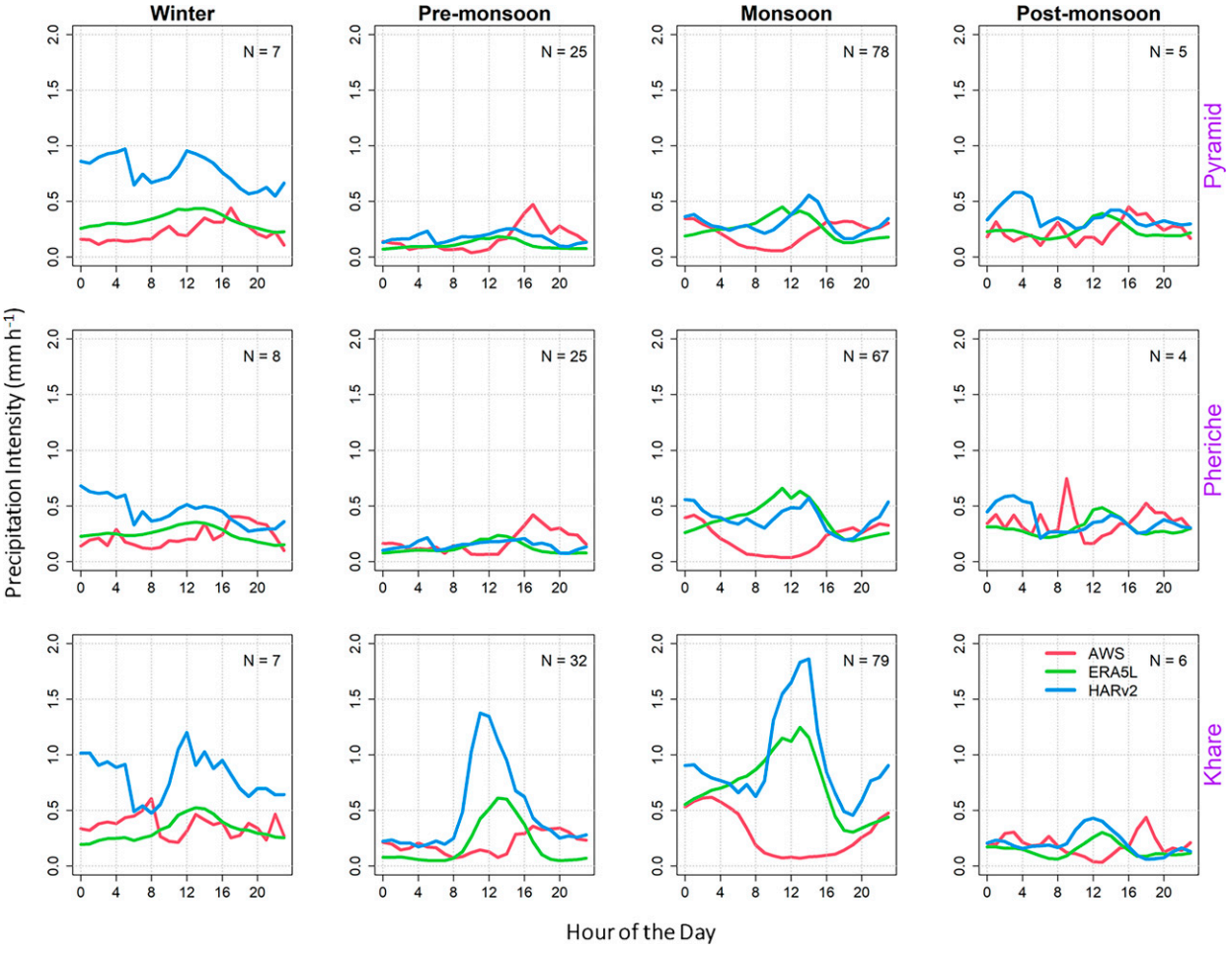

FIG. 13. Mean diurnal cycle of reanalyzed and observed total precipitation at AWS sites for different seasons. Only days with observed daily precipitation exceeding 1 mm w.e. day^{-1} are used in these plots. The mean number of days with daily precipitation exceeding 1 mm day^{-1} w.e. for each season at each site is reported in each panel (i.e., *N*).

Typhoons Phailin and Hudhud, mid-October 2013 and 2014, respectively, with a moderate overestimation especially for HARv2 (Table 9).

In general, both gridded datasets are able to well represent the seasonal cycle of all meteorological variables, except ERA5L for wind speed. This means that they are able to resolve the mesoscale atmospheric processes, but they cannot be directly used as surrogates for AWS based observations due to the scale mismatch. For instance, for most of the analyzed variables, the agreement between reanalysis data and in

TABLE 4. Monsoonal and annual (after the slash) mean meteorological variables at Mera La and total precipitation at Pyramid, Pheriche, and Khare. The annual values are calculated from 1 December of one year to 30 November of the following year. Also shown are the mean value over the entire measuring period and the SD.

Year	2012–13	2013–14	2014–15	2015–16	2016–17	2017–18	2018–19	2019–20	Mean	SD
T (°C)	—	1.4/−4.1	1.5/−3.8	1.7/−3.0	1.8/−3.3	2.0/−3.1	1.9/−3.9	2.3/−3.9	1.8/−3.5	0.3/0.5
RH (%)	—	91/59	89/60	93/59	92/61	92/60	90/61	93/63	91/60	1.4/1.6
q (g kg^{-1})	—	7.3/3.8	7.2/3.8	7.5/4.0	7.5/4.0	7.6/3.9	7.4/3.9	7.9/4.2	7.5/3.9	0.2/0.2
u (m s^{-1})	—	1.0/2.0	1.1/1.9	1.0/1.9	1.0/1.8	1.0/1.9	1.1/2.1	1.2/2.1	1.1/2.0	0.1/0.1
SWin (W m^{-2})	—	160/214	171/220	152/205	159/212	156/205	172/213	162/213	162/212	7.4/5.1
LWin (W m^{-2})	—	305/243	302/239	312/242	309/242	308/242	306/240	311/247	308/243	3.5/2.6
P_{Pyramid} (mm)	387/619	397/595	332/585	491/633	416/549	435/539	430/620	—	413/591	49/36
P_{Pheriche} (mm)	331/692	334/539	286/611	407/615	351/454	288/361	364/504	—	337/540	43/111
P_{Khare} (mm)	—	—	—	—	522/753	542/668	582/862	635/987	570/818	50/138

TABLE 5. Values of each statistical metric (r, bias, SD, and RMSE) obtained by comparing in situ and reanalyzed data at a daily time scale, over the total measuring period for each AWS. Measuring periods at each site are reported in Table 1. All r values reported in this table have $p < 0.001$.

	ERA5L vs AWS				HARv2 vs AWS			
Metric	r	Bias	SD	RMSE	r	Bias	SD	RMSE
Changri Nup								
T (°C)	0.97	−2.2	2.2	3.1	0.96	−2.4	2.3	3.3
RH (%)	0.90	12.2	13.2	18.0	0.90	4.2	12.8	13.6
q (g kg^{-1})	0.98	0	0.4	0.6	0.98	−0.1	0.4	0.5
u (m s^{-1})	0.15	−1.4	0.9	1.7	0.64	1.1	1.6	2.0
SWin (W m^{-2})	0.76	15.9	46.9	49.6	0.78	21.1	51.1	55.3
LWin (W m^{-2})	0.95	−10.7	19.0	21.8	0.93	−30.1	22.9	37.8
Mera La								
T (°C)	0.95	−1.3	2.1	2.4	0.97	−1.8	2.1	2.4
RH (%)	0.86	6.4	15.2	16.5	0.93	10.8	11.8	16.0
q (g kg^{-1})	0.97	1.0	0.8	1.3	0.99	0.2	0.5	0.5
u (m s^{-1})	0.69	−1.2	0.9	1.5	0.82	2.4	1.4	2.8
SWin (W m^{-2})	0.71	4.0	49.4	49.6	0.77	−52.3	63.6	82.3
LWin (W m^{-2})	0.94	7.9	22.1	23.3	0.94	−19.0	23.8	30.5
Mera Summit								
T (°C)	0.97	0.3	1.8	1.8	0.97	1.0	2.5	2.5
RH (%)	0.59	7.7	22.1	23.4	0.69	15.2	18.3	23.6
q (g kg^{-1})	0.92	2.1	1.7	2.7	0.93	2.6	1.4	3.0
u (m s^{-1})	0.68	−4.0	2.3	4.6	0.80	−0.3	1.5	1.5
SWin (W m^{-2})	0.47	−59.8	72.7	94.1	−0.01	−134.6	123.7	182.8
LWin (W m^{-2})	0.89	40.5	30.5	50.7	0.92	39.6	26.8	47.9
Naulek								
T (°C)	0.95	0.3	2.1	2.1	0.97	−0.2	1.7	1.7
RH (%)	0.87	5.0	14.4	15.0	0.91	10.3	12.1	15.9
q (g kg^{-1})	0.97	1.3	0.9	1.6	0.99	0.6	0.5	0.8
u (m s^{-1})	0.71	−1.8	1.4	2.3	0.81	2.1	1.3	2.5
SWin (W m^{-2})	0.73	−22.2	49.3	54.0	0.62	−79.3	77.6	110.9
LWin (W m^{-2})	0.95	16.1	19.9	25.6	0.96	−10.4	21.5	23.8
Pyramid								
P (mm day^{-1})	0.63	0.5	3.7	3.7	0.60	0.9	5.2	5.3
Pheriche								
P (mm day^{-1})	0.62	1.4	4.6	4.8	0.67	1.1	5.4	5.5
Khare								
P (mm day^{-1})	0.68	1.8	6.0	6.3	0.66	2.8	8.8	9.2

situ data decreases when the altitude difference between the reanalysis grid point and that of the AWS increases. It is thus impossible to use the same grid point for different AWSs concurrently, without applying either dynamical or statistical techniques to downscale the gridded data to local scale or individual sites.

b. *Qualitative relevance for mass and energy balance studies over glaciers*

For energy and mass balance studies over glaciers in the central Himalayas, several point-scale surface energy balance studies, mostly conducted in the ablation area, have revealed that net all-wave radiation is the main energy flux controlling the melt of clean-ice glaciers (e.g., Kayastha et al. 1999; Litt et al. 2019) while turbulent fluxes play a secondary role, still relatively important in the accumulation zone (e.g., Stigter et al. 2018) or over debris-covered areas (e.g., Steiner et al. 2018; Giese et al. 2020). Solid precipitation with its impact on albedo and air temperature, with its control on the rain–snow limit, are therefore very important meteorological variables influencing the glacier mass balance in the ablation area, while wind speed, which governs sublimation, is more important in the accumulation area. When in situ measurements are not available close to glaciers, SWin and LWin from ERA5L data should be preferred to those from HARv2, as long as the elevation difference between the reanalysis grid point and the studied point is accounted for. Reanalyzed air temperature, from ERA5L or HARv2, are both suited to control the elevation of the rain–snow limit when using adequate elevational gradients of air temperature above ground, preferentially calibrated with observed records at different elevations in the studied area. However, both reanalysis datasets largely overestimate precipitation records at high elevation and should be bias corrected, also using in situ measurements. Additionally, a special attention should be paid for the occurrence of precipitation, not always in phase with observations at high temporal resolution, that is, hourly time scale. Finally, ERA5L wind speed totally fails to capture wind speed at high elevation, and

TABLE 6. Mean seasonal bias between reanalyzed daily data (ERA5L and HARv2) and observed data over different seasons (DJF = winter, MAM = premonsoon, JJAS = monsoon, and ON = postmonsoon).

		ERA5L vs AWS				HARv2 vs AWS			
	AWS site	DJF	MAM	JJAS	ON	DJF	MAM	JJAS	ON
T (°C)	Changri Nup	−4.4	−2.6	−0.7	−2.0	−5.0	−2.3	−0.8	−2.2
	Mera La	−3.2	−1.5	0.4	−1.5	−3.4	−1.9	−0.4	−2.0
	Mera Summit	1.0	0.1	−0.4	—	3.1	0.7	−1.0	—
	Naulek	−1.7	−0.1	1.7	1.0	−1.8	−0.4	0.8	0.6
RH (%)	Changri Nup	18.3	13.2	4.6	17.0	13.0	3.1	−2.2	7.0
	Mera La	8.6	3.3	−0.6	22.5	19.0	8.9	3.3	16.6
	Mera Summit	−1.0	6.6	15.4	—	11.4	15.5	19.4	—
	Naulek	4.6	3.7	1.4	20.1	15.3	9.2	3.6	12.5
q (g kg^{-1})	Changri Nup	0.2	0.0	−0.4	0.4	0.1	−0.1	−0.3	0.1
	Mera La	0.4	0.7	1.3	1.6	0.4	0	0.2	0.5
	Mera Summit	0.6	1.9	3.8	—	1.2	2.5	4.1	—
	Naulek	0.4	1.0	2.0	1.8	0.4	0.4	0.9	0.7
u (m s^{-1})	Changri Nup	−2.3	−1.5	−0.6	−1.6	2.2	1.4	0.4	0.6
	Mera La	−2.0	−1.3	−0.6	−1.2	3.4	2.6	1.7	2.1
	Mera Summit	−5.8	−4.2	−2.1	−3.6	−0.7	−0.4	0.5	−1.1
	Naulek	−3.8	−2.3	−1.0	−2.7	2.7	2.3	1.8	1.6
SWin (W m^{-2})	Changri Nup	12.3	19.6	27.4	−5.0	5.2	33.1	30.2	7.9
	Mera La	−15.9	30.9	15.0	−30.4	−24.5	−33.8	−88.8	−45.2
	Mera Summit	−3.5	−40.6	−138.7	—	−21.5	−128.1	−257.2	—
	Naulek	−6.2	−12.0	−32.4	−29.1	−16.5	−75.7	−137.6	−45.2
LWin (W m^{-2})	Changri Nup	−15.0	−18.2	−9.4	4.6	−20.3	−38.5	−33.4	−24.6
	Mera La	−1.3	−2.8	15.5	22.8	−23.2	−31.4	−6.2	−19.7
	Mera Summit	29.8	38.4	53.3	—	31.6	37.4	49.9	—
	Naulek	2.8	12.2	21.3	29.7	−18.2	−17.7	0.6	−12.9
P (mm w.e. day^{-1})	Pyramid	0.4	−0.2	1.1	0.1	1.2	0.1	1.6	0.3
	Pheriche	0.3	0.1	3.3	0.1	0.5	−0.1	2.8	0.1
	Khare	0.1	0.7	4.5	0.3	0.3	1.7	6.5	0.2

HARv2 wind velocity, even though it is slightly overestimated, should be considered preferentially to assess sublimation over glaciers, especially at high elevation where sublimation is an important component of glacier mass balances. Litt et al. (2019) have shown that temperature index models, including or not a shortwave radiation scheme, are suitable to assess melt in the ablation zone of glaciers in the central Himalayas, during the monsoon, but are unable to quantify ablation higher in altitude or during the other seasons, because sublimation prevails and is unresolved by such models. In turn, air

TABLE 7. Comparison between the occurrence of precipitation observed at AWS and in reanalysis datasets, at hourly time scale. True positive means that there is precipitation at AWS and in the reanalysis dataset. True negative means that there is no precipitation at AWS and in the reanalysis dataset. False positive means that the reanalysis dataset simulates precipitation although precipitation is not observed at AWS. False negative means that the reanalysis dataset does not simulate precipitation although precipitation is observed at AWS. Values are percentages of hours as a function of the total number of hours. On the one hand, we consider the entire measuring periods (see Table 1 for those periods), and, on the other hand, we consider only the monsoons. Thus, for each entry in the table, the number in front of the slash considers the total number of hours of the entire measuring period and the number after the slash considers the total number of hours of the monsoons only To discriminate hours with or without precipitation, a threshold of 0.1 mm w.e. h^{-1} is used for AWS data, below which we consider that there is no precipitation. The thresholds for reanalysis data are 1.3–3.5 higher than that of AWS based on the overall overestimation of the precipitation by reanalysis.

	True positive	True negative	False positive	False negative
Pyramid				
AWS vs ERA5L	5/9	79/65	11/19	5/7
AWS vs HARv2	4/6	83/72	7/11	6/10
Pheriche				
AWS vs ERA5L	4/8	79/64	12/22	5/7
AWS vs HARv2	3/5	84/73	7/12	6/9
Khare				
AWS vs ERA5L	7/12	71/52	15/25	7/11
AWS vs HARv2	7/11	74/57	12/20	8/12

TABLE 8. As in Table 7, but at daily time scale. To discriminate days with or without precipitation, a threshold of 0.5 mm w.e. d^{-1} is used for AWS data, below which we consider that there is no precipitation.

	True positive	True negative	False positive	False negative
Pyramid				
AWS vs ERA5L	30/53	53/30	10/13	7/4
AWS vs HARv2	23/39	57/37	5/6	15/18
Pheriche				
AWS vs ERA5L	28/49	50/26	15/22	7/2
AWS vs HARv2	22/37	60/41	6/7	13/15
Khare				
AWS vs ERA5L	35/57	40/17	22/24	3/1
AWS vs HARv2	32/53	52/29	10/12	6/5

temperature, from both reanalysis datasets, can be used as input variables for those models to assess melt in the ablation area of glaciers in the monsoon, as long as site- and season-specific ablation factors are considered (Litt et al. 2019). For physically based surface energy balance modeling, we suggest paying more attention to the nonstationary biases in all variables.

In conclusion, the two gridded datasets ERA5L and HARv2 are applicable for glacier mass and energy balance studies, as long as either statistical or dynamical downscaling techniques are used to resolve the scale mismatch between coarse mesoscale grids to fine-scale grids or individual sites. We found biases for some variables, like temperature and precipitation, and these biases may be nonstationary and vary in the different seasons. If no AWS measurements are available in the neighborhood of the studied glacier, the temperature and precipitation biases can still be corrected by indirect methods (i.e., Immerzeel et al. 2014; Kraaijenbrink et al. 2017). However, the indirect bias corrections remain less accurate than direct comparisons with AWS, because of equifinality in the glacier mass and energy balance models.

6. Conclusions

Because of the scarcity and discontinuity of long-term observational data at high elevation in the Himalayas, the main sources of data for studying climate-glacier relationship at regional scale are the reanalysis datasets. Here, ERA5L and HARv2 reanalyses, having the best spatial and temporal resolution available, are compared with in situ meteorological data to evaluate their performance on the basis of classical statistical metrics. Observation data have been collected since 2010 by seven AWSs located on or off glaciers above 4260 m MSL in the upper Dudh Koshi basin (Everest region, Nepal). 2-m air temperature, specific and relative humidities, wind speed, incoming shortwave and longwave radiations, and precipitation are considered in this study. Climate reanalyses show different levels of performance depending on the meteorological variables, the geographical context, and the surface state. Because of the complex topography of this high-altitude basin, the spatial resolution of reanalysis products is still a strong limitation to reproduce the observations, especially for highly spatially variables such as precipitation or wind speed that need local-scale spatial resolution obtained by further downscaling (e.g., Mölg et al. 2012). Air temperature is the best captured by reanalyses, as long as an appropriate elevational gradient of air temperature above ground, spatiotemporally variable and preferentially assessed by local observations, is used to extrapolate it vertically. A cold bias is still observed but less important over clean-ice glaciers than over rocky surfaces. For relative and specific humidities, both reanalysis products perform well except at Mera Summit, but still have a moderate humid bias, especially during the driest months, that is, in winter and postmonsoon. ERA5L totally fails to reproduce wind speed, with a systematic underestimation and an absence of seasonality. HARv2 performance is better, with high wind speeds in winter and postmonsoon, and calm conditions during the monsoon in agreement with observations.

TABLE 9. Two 3-day precipitation events captured at Pyramid and Pheriche AWSs during Typhoon Phailin in October 2013 and Typhoon Hudhud in October 2014, compared with reanalysis data (mm w.e.).

Date	Pyramid AWS	Pheriche AWS	Pyramid ERA5L	Pheriche ERA5L	Pyramid HARv2	Pheriche HARv2
Typhoon Phailin						
13 Oct 2013	13.7	21.4	13.1	13.7	31.1	30.3
14 Oct 2013	58.2	91.1	54.5	57.8	95.1	96.4
15 Oct 2013	14.7	24.4	28.0	32.8	8.8	6.8
Total	86.6	136.9	95.7	104.3	135	133.5
Typhoon Hudhud						
13 Oct 2014	5.8	5.9	9.3	9.0	16.4	10.5
14 Oct 2014	27.9	21.9	18.1	19.0	44.9	25.6
15 Oct 2014	5.7	14.7	3.1	3.5	1.2	0.5
Total	39.4	42.5	30.5	31.5	62.5	36.5

Nevertheless, HARv2 wind speed is systematically overestimated, and its daily cycle is out of phase with the observed one. The performance of reanalysis datasets for shortwave and longwave incoming radiations is highly dependent on the elevation difference between the reanalysis grid point and the observation site. When the elevation of the grid point is lower than that of the observation site, the discrepancy between reanalyzed and observed incoming radiation (underestimation for shortwave radiation and overestimation for longwave radiation) increases as the additional layer of atmosphere to be crossed is thicker. A correction procedure to account for the elevation difference between the reanalysis grid point and the observation site should be considered to properly reproduce incoming shortwave and longwave radiations. The seasonality and the horizontal south-to-north gradient of precipitation are reasonably captured by these two reanalyses. Nevertheless, they tend to highly overestimate precipitation up to a factor of 3 during the monsoon, ERA5L performing better than HARv2. The occurrence of precipitation is reasonably reproduced at daily time scale but less well at hourly time scale. In conclusion, the two gridded datasets ERA5L and HARv2 cannot be directly used as surrogates for AWS based observations. Nevertheless, they are applicable for glacier mass and energy balance studies, as long as either statistical or dynamical downscaling techniques are used to resolve the scale mismatch between coarse mesoscale grids to fine-scale grids or individual sites.

As long as high-quality long-term meteorological records are not available at high elevation in the Himalayas, reanalysis datasets will not be able to do otherwise than assimilate data only at lower altitudes. This is problematic for any studies focusing on climate at high elevation such as glacier studies. In that case, assessing the performance and adequately correcting the used reanalysis datasets is a prerequisite and implies having a weather station running at glacier elevations for at least one year. In remote mountain areas like the Himalayas, due to access difficulties, extreme weather, and complex topography, maintaining continuous high-quality meteorological records over a long-term is almost impossible, and therefore combining in situ data and various high-resolution reanalyses is highly recommended.

Acknowledgments. This work has been supported by the French Service d'Observation GLACIOCLIM, now called CRYOBS-CLIM and part of IR OZCAR, the French National Research Agency (ANR) through ANR-13-SENV-0005-04/05-PRESHINE, and a grant from Labex OSUG@2020 (Investissements d'avenir—ANR10 LABX56). This work would not have been possible without the support of the JEAI HIMALICE (principal investigators D. Shrestha, and Y. Arnaud) and all the efforts from people in the field: porters, students, and helpers who are greatly acknowledged here. This study was carried out within the framework of the Ev-K2-CNR Project in collaboration with the Nepal Academy of Science and Technology and Tribhuvan University. Author Khadka benefits from a Ph.D. fellowship from the National Geographic Society and ARTS IRD, and he benefited from the support of the International Centre for Integrated Mountain Development (ICIMOD), which is funded in part by the governments of Afghanistan, Bangladesh, Bhutan, China, India, Myanmar, Nepal, and Pakistan. We thank the Nepal Mountaineering Association, which provides free permits to conduct field work on Mera peak.

Data availability statement. All AWS data are available in the GLACIOCLIM database (https://glacioclim.osug.fr/Donnees-himalaya).

REFERENCES

Alduchov, O. A., and R. E. Eskridge, 1996: Improved Magnus form approximation of saturation vapor pressure. *J. Appl. Meteor.*, **35**, 601–609, https://doi.org/10.1175/1520-0450(1996)035<0601:IMFAOS>2.0.CO;2.

Baudouin, J. P., M. Herzog, and C. A. Petrie, 2020: Cross-validating precipitation datasets in the Indus river basin. *Hydrol. Earth Syst. Sci.*, **24**, 427–450, https://doi.org/10.5194/hess-24-427-2020.

Berthier, E., and F. Brun, 2019: Karakoram geodetic glacier mass balances between 2008 and 2016: Persistence of the anomaly and influence of a large rock avalanche on Siachen Glacier. *J. Glaciol.*, **65**, 494–507, https://doi.org/10.1017/jog.2019.32.

Bolch, T., and Coauthors, 2019: Status and change of the cryosphere in the extended Hindu Kush Himalaya region. *The Hindu Kush Himalaya Assessment*, Springer, 209–255, https://doi.org/10.1007/978-3-319-92288-1_7.

Bollasina, M., L. Bertolalani, and G. Tartari, 2002: Meteorological observations at high altitude in the Khumbu Valley, Nepal Himalayas, 1994–1999. *Bull. Glaciol. Res.*, **19**, 1–12.

Bonasoni, P., and Coauthors, 2010: Atmospheric brown clouds in the Himalayas: First two years of continuous observations at the Nepal climate observatory-pyramid (5079 M). *Atmos. Chem. Phys.*, **10**, 7515–7531, https://doi.org/10.5194/acp-10-7515-2010.

Bookhagen, B., and D. W. Burbank, 2010: Toward a complete Himalayan hydrological budget: Spatiotemporal distribution of snowmelt and rainfall and their impact on river discharge. *J. Geophys. Res.*, **115**, F03019, https://doi.org/10.1029/2009JF001426.

Brun, F., E. Berthier, P. Wagnon, A. Kääb, and D. Treichler, 2017: A spatially resolved estimate of high mountain Asia glacier mass balances from 2000 to 2016. *Nat. Geosci.*, **10**, 668–673, https://doi.org/10.1038/ngeo2999.

Cao, B., S. Gruber, D. Zheng, and X. Li, 2020: The ERA5-land soil-temperature bias in permafrost regions. *Cryosphere*, **14**, 2581–2595, https://doi.org/10.5194/tc-14-2581-2020.

Dee, D. P., and Coauthors, 2011: The ERA-Interim reanalysis: Configuration and performance of the data assimilation system. *Quart. J. Roy. Meteor. Soc.*, **137**, 553–597, https://doi.org/10.1002/qj.828.

Eeckman, J., P. Chevallier, A. Boone, L. Neppel, A. De Rouw, F. Delclaux, and D. Koirala, 2017: Providing a non-deterministic representation of spatial variability of precipitation in the Everest region. *Hydrol. Earth Syst. Sci.*, **21**, 4879–4893, https://doi.org/10.5194/hess-21-4879-2017.

Førland, E. J., and Coauthors, 1996: Manual for operational correction of Nordic precipitation data. Nordic Working Group on Precipitation DNMI-Klima Rep. 24/96, 66 pp.

Fujita, K., and Coauthors, 2017: Anomalous winter-snow-amplified earthquake-induced disaster of the 2015 Langtang avalanche in Nepal. *Nat. Hazards Earth Syst. Sci.*, **17**, 749–764, https://doi.org/10.5194/nhess-17-749-2017.

Giese, A., A. Boone, P. Wagnon, and R. Hawley, 2020: Incorporating moisture content in surface energy balance modeling of a debris-covered glacier. *Cryosphere*, **14**, 1555–1577, https://doi.org/10.5194/tc-14-1555-2020.

Hamm, A., and Coauthors, 2020: Intercomparison of gridded precipitation datasets over a sub-region of the central Himalaya and the southwestern Tibetan Plateau. *Water*, **12**, 3271, https://doi.org/10.3390/w12113271.

Hersbach, H., and Coauthors, 2020: The ERA5 global reanalysis. *Quart. J. Roy. Meteor. Soc.*, **146**, 1999–2049, https://doi.org/10.1002/qj.3803.

Immerzeel, W. W., L. P. Van Beek, and M. F. P. Bierkens, 2010: Climate change will affect the Asian water towers. *Science*, **328**, 1382–1385, https://doi.org/10.1126/science.1183188.

——, L. Petersen, S. Ragettli, and F. Pellicciotti, 2014: The importance of observed gradients of air temperature and precipitation for modeling runoff from a glacierized watershed in the Nepalese Himalayas. *Water Resour. Res.*, **50**, 2212–2226, https://doi.org/10.1002/2013WR014506.

——, and Coauthors, 2020: Importance and vulnerability of the world's water towers. *Nature*, **577**, 364–369, https://doi.org/10.1038/s41586-019-1822-y.

Kattel, D. B., T. Yao, K. Yang, L. Tian, G. Yang, and D. Joswiak, 2013: Temperature lapse rate in complex mountain terrain on the southern slope of the central Himalayas. *Theor. Appl. Climatol.*, **113**, 671–682, https://doi.org/10.1007/s00704-012-0816-6.

Kayastha, R. B., T. Ohata, and Y. Ageta, 1999: Application of a mass-balance model to a Himalayan glacier. *J. Glaciol.*, **45**, 559–567, https://doi.org/10.1017/S002214300000143X.

Kochendorfer, J., and Coauthors, 2017: Analysis of single-alter-shielded and unshielded measurements of mixed and solid precipitation from WMO-SPICE. *Hydrol. Earth Syst. Sci.*, **21**, 3525–3542, https://doi.org/10.5194/hess-21-3525-2017.

Kraaijenbrink, P. D. A., F. P. Bierkens, A. F. Lutz, and W. W. Immerzeel, 2017: Impact of a global temperature rise of 1.5 degrees Celsius on Asia's glaciers. *Nature*, **549**, 257–260, https://doi.org/10.1038/nature23878.

Lejeune, Y., P. Wagnon, L. Bouilloud, P. Chevallier, P. Etchevers, E. Martin, J. E. Sicart, and F. Habets, 2007: Melting of snow cover in a tropical mountain environment: Processes and modeling. *J. Hydrometeor.*, **8**, 922–937, https://doi.org/10.1175/JHM590.1.

Litt, M., J. M. Shea, P. Wagnon, J. Steiner, I. Koch, and E. Stigter, 2019: Glacier ablation and temperature indexed melt models in the Nepalese Himalaya. *Sci. Rep.*, **9**, 5264, https://doi.org/10.1038/s41598-019-41657-5.

Lutz, A. F., W. W. Immerzeel, A. B. Shrestha, and M. F. P. Bierkens, 2014: Consistent increase in High Asia's runoff due to increasing glacier melt and precipitation. *Nat. Climate Change*, **4**, 587–592, https://doi.org/10.1038/nclimate2237.

Maussion, F., D. Scherer, R. Finkelnburg, J. Richters, W. Yang, T. Yao, D. Scherer, and W. Yang, 2011: WRF simulation of a precipitation event over the Tibetan Plateau, China—An assessment using remote sensing and ground observations. *Hydrol. Earth Syst. Sci.*, **15**, 1795–1817, https://doi.org/10.5194/hess-15-1795-2011.

——, ——, T. Mölg, E. Collier, J. Curio, and R. Finkelnburg, 2014: Precipitation seasonality and variability over the Tibetan Plateau as resolved by the High Asia Reanalysis. *J. Climate*, **27**, 1910–1927, https://doi.org/10.1175/JCLI-D-13-00282.1.

Miles, E., J. F. Steiner, and F. Brun, 2017: Highly variable aerodynamic roughness length (Z_0) for a hummocky debris-covered glacier. *J. Geophys. Res. Atmos.*, **122**, 8447–8466, https://doi.org/10.1002/2017JD026510.

Mimeau, L., M. Esteves, I. Zin, H.-W. Jacobi, F. Brun, P. Wagnon, D. Koirala, and Y. Arnaud, 2019: Quantification of different flow components in a high-altitude glacierized catchment (Dudh Koshi, Himalaya): Some cryospheric-related issues. *Hydrol. Earth Syst. Sci.*, **23**, 3969–3996, https://doi.org/10.5194/hess-23-3969-2019.

Mölg, T., F. Maussion, W. Yang, and D. Scherer, 2012: The footprint of Asian monsoon dynamics in the mass and energy balance of a Tibetan glacier. *Cryosphere*, **6**, 1445–1461, https://doi.org/10.5194/tc-6-1445-2012.

——, ——, and D. Scherer, 2014: Mid-latitude westerlies as a driver of glacier variability in monsoonal High Asia. *Nat. Climate Change*, **4**, 68–73, https://doi.org/10.1038/nclimate2055.

Oke, T. R., 2002: *Boundary Layer Climates.* Routledge, 464 pp.

Orsolini, Y., and Coauthors, 2019: Evaluation of snow depth and snow cover over the Tibetan Plateau in global reanalyses using in situ and satellite remote sensing observations. *Cryosphere*, **13**, 2221–2239, https://doi.org/10.5194/tc-13-2221-2019.

Perry, L. B., and Coauthors, 2020: Precipitation characteristics and moisture source regions on Mt. Everest in the Khumbu, Nepal. *One Earth*, **3**, 594–607, https://doi.org/10.1016/j.oneear.2020.10.011.

Pfeffer, W. T., and Coauthors, 2014: The Randolph Glacier Inventory: A globally complete inventory of glaciers. *J. Glaciol.*, **60**, 537–552, https://doi.org/10.3189/2014JoG13J176.

Pritchard, H. D., 2019: Asia's shrinking glaciers protect large populations from drought stress. *Nature*, **569**, 649–654, https://doi.org/10.1038/s41586-019-1240-1.

Pugh, L., 1954: Notes on temperature and snow conditions in the Everest region in spring 1952 and 1953. *J. Glaciol.*, **2**, 363–365, https://doi.org/10.1017/S0022143000025284.

Ragettli, S., and Coauthors, 2015: Unraveling the hydrology of a Himalayan catchment through integration of high resolution in situ data and remote sensing with an advanced simulation model. *Adv. Water Resour.*, **78**, 94–111, https://doi.org/10.1016/j.advwatres.2015.01.013.

Sakai, A., and K. Fujita, 2017: Contrasting glacier responses to recent climate change in high-mountain Asia. *Sci. Rep.*, **7**, 13717, https://doi.org/10.1038/s41598-017-14256-5.

Salerno, F., and Coauthors, 2015: Weak precipitation, warm winters and springs impact glaciers of south slopes of Mt. Everest (central Himalaya) in the last 2 decades (1994–2013). *Cryosphere*, **9**, 1229–1247, https://doi.org/10.5194/tc-9-1229-2015.

Sanz Rodrigo, J. S., J. M. Buchlin, J. van Beeck, J. T. Lenaerts, and M. R. van den Broeke, 2013: Evaluation of the Antarctic surface wind climate from ERA reanalyses and RACMO2/ANT simulations based on automatic weather stations. *Climate Dyn.*, **40**, 353–376, https://doi.org/10.1007/s00382-012-1396-y.

Shea, J. M., P. Wagnon, W. W. Immerzeel, R. Biron, F. Brun, and F. Pellicciotti, 2015: A comparative high-altitude meteorological analysis from three catchments in the Nepalese Himalaya. *Int. J. Water Resour. Dev.*, **31**, 174–200, https://doi.org/10.1080/07900627.2015.1020417.

Shean, D. E., S. Bhushan, P. Montesano, D. R. Rounce, A. Arendt, and B. Osmanoglu, 2020: A systematic, regional

assessment of high mountain Asia glacier mass balance. *Front. Earth Sci.*, **7**, 363, https://doi.org/10.3389/feart.2019.00363.

Sherpa, S. F., and Coauthors, 2017: Contrasted surface mass balances of debris-free glaciers observed between the southern and the inner parts of the Everest region (2007–15). *J. Glaciol.*, **63**, 637–651, https://doi.org/10.1017/jog.2017.30.

Steiner, J. F., and F. Pellicciotti, 2016: On the variability of air temperature over a debris-covered glacier, Nepalese Himalaya. *Ann. Glaciol.*, **57**, 295–307, https://doi.org/10.3189/2016AoG71A066.

——, M. Litt, E. E. Stigter, J. Shea, M. F. P. Bierkens, and W. W. Immerzeel, 2018: The importance of turbulent fluxes in the surface energy balance of a debris-covered glacier in the Himalayas. *Front. Earth Sci.*, **6**, 1–25, https://doi.org/10.3389/FEART.2018.00144.

Stigter, E. E., M. Litt, J. F. Steiner, P. N. J. Bonekamp, J. M. Shea, M. F. P. Bierkens, and W. W. Immerzeel, 2018: The importance of snow sublimation on a Himalayan glacier. *Front. Earth Sci.*, **6**, 108, https://doi.org/10.3389/feart.2018.00108.

Stumm, D., S. P. Joshi, T. K. Gurung, and G. Silwal, 2021: Mass balances of Yala and Rikha Samba glaciers, Nepal, from 2000 to 2017. *Earth Syst. Sci. Data*, **13**, 3791–3818, https://doi.org/10.5194/essd-13-3791-2021.

Sunako, S., K. Fujita, A. Sakai, and R. B. Kayastha, 2019: Mass balance of Trambau Glacier, Rolwaling Region, Nepal Himalaya: In-situ observations, long-term reconstruction and mass-balance sensitivity. *J. Glaciol.*, **65**, 605–616, https://doi.org/10.1017/jog.2019.37.

Tetzner, D., E. Thomas, and C. Allen, 2019: A validation of ERA5 reanalysis data in the southern Antarctic Peninsula—Ellsworth Land region, and its implications for ice core studies. *Geosciences*, **9**, 289, https://doi.org/10.3390/geosciences9070289.

Ueno, K., K. Toyotsu, L. Bertolani, and G. Tartari, 2008: Stepwise onset of monsoon weather observed in the Nepal Himalaya. *Mon. Wea. Rev.*, **136**, 2507–2522, https://doi.org/10.1175/2007MWR2298.1.

United Nations, 2015: Adoption of the Paris Agreement. Framework Convention on Climate Change Doc. FCCC/CP/2015/L.9/Rev1, 32 pp., https://unfccc.int/sites/default/files/resource/docs/2015/cop21/eng/l09r01.pdf.

Wagnon, P., and Coauthors, 2020: Reanalysing the 2007–19 glaciological mass-balance series of Mera Glacier, Nepal, central Himalaya, using geodetic mass balance. *J. Glaciol.*, **67**, 117–125, https://doi.org/10.1017/jog.2020.88.

Wang, C., R. M. Graham, K. Wang, S. Gerland, and M. A. Granskog, 2019: Comparison of ERA5 and ERA-Interim near-surface air temperature, snowfall and precipitation over Arctic Sea ice: Effects on sea ice thermodynamics and evolution. *Cryosphere*, **13**, 1661–1679, https://doi.org/10.5194/tc-13-1661-2019.

Wang, X., V. Tolksdorf, M. Otto, and D. Scherer, 2020: WRF-based dynamical downscaling of ERA5 reanalysis data for high mountain Asia: Towards a new version of the high Asia refined analysis. *Int. J. Climatol.*, **41**, 743–762, https://doi.org/10.1002/joc.6686.

Yamamoto, M. K., K. Ueno, and K. Nakamura, 2011: Comparison of satellite precipitation products with rain gauge data for the Khumb Region, Nepal Himalayas. *J. Meteor. Soc. Japan*, **89**, 597–610, https://doi.org/10.2151/jmsj.2011-601.