Rising Planetary Boundary Layer Height over the Sahara Deserts and Arabian Peninsula in a Warming Climate

Liming Zhou[1], Yuhong Tian[2], Nan Wei[3], Shu-peng Ho[4], Jing Li[5]

[1]*Department of Atmospheric and Environmental Sciences, SUNY at Albany, Albany, NY 12222, USA*
[2]*New York State Department of Environmental Conservation, Albany, NY 12222, USA*
[3]*School of Atmospheric Sciences, Sun Yat-sen University, No. 135, Xingang Xi Road, Guangzhou 510275, China*
[4]*Center for Satellite Applications & Research (STAR), NESDIS/NOAA, College Park, MD 20740, USA*
[5]*Department of Atmospheric and Oceanic Sciences, School of Physics, Peking University, Beijing, China*

Correspondence
Liming Zhou
Department of Atmospheric and Environmental Sciences
University at Albany, State University of New York
1400 Washington Avenue, Albany, NY 12222
Tel: (518) 442-4446; Fax: (518) 442-5825
Email: lzhou@albany.edu

Accepted for publication in
Journal of Climate
December 23, 2020

Abstract. Turbulent mixing in the planetary boundary layer (PBL) governs the vertical exchange of heat, moisture, momentum, trace gases, and aerosols in the surface-atmosphere interface. The PBL height (PBLH) represents the maximum height of the free atmosphere that is directly influenced by the Earth's surface. This study uses a multi-data synthesis approach from an ensemble of multiple global datasets of radiosonde observations, reanalysis products and climate model simulations to examine the spatial patterns of long-term PBLH trends over land between 60°S and 60°N for the period 1979-2019. By considering both the sign and statistical significance of trends, we identify large-scale regions where the change signal is robust and consistent to increase our confidence in the obtained results. Despite differences in the magnitude and sign of PBLH trends over many areas, all datasets reveal a consensus on increasing PBLH over the enormous and driest Sahara Desert and Arabian Peninsula (SDAP) and declining PBLH in India. At the global scale, the changes in PBLH are significantly correlated positively with the changes in surface heating and negatively with the changes in surface moisture, consistent with theory and previous findings in the literature. The rising PBLH is in good agreement with increasing sensible heat and surface temperature and decreasing relative humidity over the SDAP associated with desert amplification**,** while the declining PBLH resonates well with increasing relative humidity and latent heat and decreasing sensible heat and surface warming in India. The PBLH changes agree with radiosonde soundings over the SDAP but cannot be validated over India due to lack of good-quality radiosonde observations.

Key words: Global warming, desert amplification, planetary boundary layer, planetary boundary layer height

1. Introduction

The land surface has a pronounced diurnal cycle in solar insolation, surface temperatures, and atmospheric planetary boundary layer (PBL). Turbulent mixing in the PBL dominates the vertical exchange of heat, moisture, momentum, trace gases, and aerosols in the surface-atmosphere interface, and strongly influences the tropospheric temperature, humidity, and wind (Stull, 1988). One fundamental variable of the PBL is the PBL height (PBLH). PBLH represents the maximum height of the atmosphere that is directly influenced by the Earth's surface, sets limits for the mixing and dilution height of near-surface pollutants, and controls cloud formation and convection activity that affect the Earth's radiation budget and hydrological cycle (Ao et al., 2012; Chan and Wood, 2013; Ho et al., 2015).

PBLH displays substantial spatiotemporal variability under different surface and atmospheric conditions, ranging from a few hundred meters to several kilometers (Stull, 1988). The PBLH over land depends strongly on surface characteristics, including soil moisture, vegetation, land cover, terrain, and proximity to the sea (Talbot et al., 2007; Liu and Liang, 2010; Seidel et al., 2012; Zhang et al., 2013; Lee and De Wekker, 2016; Wei et al., 2017a; Sathyanadh et al., 2017). The growth of PBLH is driven primarily by surface heating and atmospheric stability (Chan and Wood, 2013; Lee and De Wekker, 2016; Ao et al., 2017; Brahmanandam et al., 2020). In the subtropics and tropics, PBLH is typically higher over drier regions and during drier seasons, because more surface sensible heat flux is available to drive vertical mixing due to less surface moisture and higher Bowen ratio. As expected, it maximizes over arid and semi-arid regions, in the afternoon, and during warm and dry seasons when land surface temperatures are warmest, sensible heat flux is most significant, and static stability is lowest. Hence, the global PBLH climatology shows the deepest daytime PBLH in the Sahara Desert and Arabian Peninsula (SDAP) (Gamo, 1996; Ao et al., 2012; Garcia-Carreras et al., 2015; Ao et al., 2017; Wei et al., 2017a).

Changes in near-surface atmospheric variables such as temperature and relative humidity (RH) are closely linked to changes in PBLH. Warmer and drier surfaces associated with higher temperatures and lower RH result in greater sensible heat flux and lower latent heat flux, leading to deeper convection and larger PBLH (Zhang et al. 2013; Darand and Zandkarimi, 2019). Hence, PBLH shows strong correlations positively with surface air temperature and negatively with surface RH in Europe (Zhang et al., 2013), China (Guo et al., 2016; Dang et al., 2016; Guo et al., 2019), East Asia and North Africa (Zhao et al., 2017), and Iran (Darand and Zandkarimi, 2019).

Global mean surface temperatures have increased since the late 19th century, and this warming has been spatially widespread and particularly marked since the 1980s, with the warming rate over land double that over the ocean (IPCC, 2013). Associated with this warming are the global increases in near-surface and tropospheric specific humidity of air (IPCC, 2013). Despite diverse and complex spatial patterns of RH changes at regional scales, recent studies using observations, reanalysis data, and general circulation models (GCMs) have suggested small increases in ocean RH but substantial decreases in land RH in recent years with global warming (e.g., Simmons et al., 2010; O'Gorman and Muller 2010; IPCC, 2013; Sherwood and Fu 2014; Willet et al., 2014; Byrne and O'Gorman, 2016; Vicente-Serrano et al., 2017). As increasing surface temperature and decreasing surface RH over land tend to deepen the PBLH, one scientific question is whether these changes in temperature and RH may have raised the PBLH over land (e.g., Zhang et al., 2013).

Due to limited spatial and temporal coverage of good-quality radiosonde measurements, there are only several observational studies on long-term PBLH trends at regional scales. Zhang et al. (2013) estimated the PBLH trends over Europe based on daily radiosonde observations at 25 stations during 1973-2010 and found statistically significant increases in daytime PBLH in all four seasons. Guo et al (2019) investigated the temporal trends of radiosonde derived PBLH from 1979 to 2016 in China and found a spatially uniform increasing trend from 1979 to 2003 but a trend shift thereafter. Li et al. (2020) calculated daily maximum PBLH globally using operational radiosonde and surface meteorological measurements from 219 carefully selected weather stations for the period 1973-2018. They found significant increasing (decreasing) trends over 74 (48) stations. However, these studies are inadequate to draw a broad conclusion about the large-scale patterns of PBLH trends because the radiosonde network is not evenly distributed globally and has data gaps in coverage over many regions.

The land surface has warmed rapidly in the past several decades but at different warming rates among different regions. Recent studies (Zhou et al., 2015; 2016; Cook and Vizy, 2015; Evan et al., 2015; Zhou, 2016) using observations, reanalysis data, and GCM simulations have found that surface air temperatures in the mid- and low- latitudes have warmed most over the SDAP. This warming amplification over deserts, which is termed desert amplification (DA), has intensified with increasing greenhouse gases (GHGs), particularly after the 1980s. The essential features of DA remained robust across all seasons, although the magnitude of DA was greater during warm seasons (Zhou et al., 2016; Vizy and Cook, 2017; Wei et al., 2017b). These results suggest that

DA is a fundamental large-scale feature of global warming patterns in the mid- and low- latitudes and will accelerate in a warming climate.

Deserts make up approximately 1/3 of the global land surface area (Zhou, 2016; Wei et al., 2017a). The SDAP is home to the two largest subtropical deserts in the world and covers a vast continental land area in the low latitudes. As the deserts are extremely dry, with limited soil moisture, vegetation and cloudiness, surface heating via sensible heat is documented as the dominant driver for the PBLH growth there (Zhao, 2011; Ao et al., 2017). Considering the amplified continent-scale surface heating associated with DA in a warming climate and constrained by limited moisture availability over the deserts, it is expected to observe widespread increases in temperature and decreases in RH over the SDAP. Another scientific question is whether the DA may have manifested its impact by deepening the PBLH at a much larger spatial extent over the SDAP than the other regions with spatially more heterogeneous RH changes.

Global reanalysis results in physically consistent estimates of past observations with complete spatial and temporal coverage and thus has greatly improved our ability to examine climate variability (Trenberth et al., 2008; IPCC. 2013). The reanalysis PBLH estimates have been used at regional to global scales with reasonable results (e.g., Ao et al., 2012; Von Engeln and Teixeira, 2013; Guo et al., 2016; Zhao et al. 2017; Darand and Zandkarimi, 2019). There are a couple of studies on the long-term trends of reanalysis PBLH over dry lands. Zhao et al. (2017) examined the inter-decadal variability of PBLH based on the ECMWF first atmospheric reanalysis of the 20th century (ERA-20C) over arid and semi-arid areas in East Asia and North Africa for the period 1900-2010 and found substantial spatiotemporal variations in the PBLH trends during the 111-

year period. This century-long reanalysis makes the investigation of multi-decadal variability possible, but it assimilated only surface observations and contains spurious long-term climate trends due to changes in the radiative forcing and the observing system throughout the century (e.g., Poli et al., 2013; Bloomfieldet al., 2018). Darand and Zandkarimi (2019) examined monthly PBLH data from the ERA-Interim reanalysis and revealed a significant increasing PBLH trend of ~31 m/decade at the country level over Iran for the period 1979-2016, with some seasonal differences and largest increases in the semi northern part of the country. However, these two regional studies are based only on one reanalysis product and have no validations against *in situ* observations and other datasets.

The availability of high-resolution reanalysis products and the newly available Coupled Model Intercomparison Project Phase 6 (CMIP6) simulations provide us an opportunity to address the above two questions by examining the PBLH changes at larger scales. This paper uses a multi-data synthesis approach from an ensemble of multiple global datasets of radiosonde observations, reanalysis products, and GCM simulations to detect and attribute the large-scale patterns of long-term PBLH trends over land in the mid- and low- latitudes between 60°S and 60°N. It focuses on the satellite era for the period 1979-2019 to maximize data coverage of measurements that are assimilated into reanalysis products and are used to drive GCMs. By considering both the sign and statistical significance of the trends, we identify large-scale regions where the change signal is robust and consistent to increase our confidence in the obtained results. The major objective is to test the hypothesis that the global warming signal is manifest most in the spatial extent of PBLH change over the SDAP where the amplified surface warming associated with DA enhances turbulent mixing and thus raise the PBLH height.

2. Data and Methods

2.1. Data sources

Here we used nine types of datasets, consisting of radiosonde observations, four reanalysis products, and 74 historical simulations from 27 CMIP6 models, for the period 1979-2019. To inter-compare the PBLH estimates from different methods, the bulk Richardson number (Ri) method (Vogelezang and Holtslag, 1996) was also chosen to consistently diagnose the PBLH, with a critical value of 0.25, directly from the atmospheric soundings for three datasets whose PBLHs were estimated using other methods. The Ri methods have proven to the most reliable approaches over a wide range of conditions for both stable and convective boundary layers and don't strongly depend on the sounding vertical resolutions (e.g., Seidel et al., 2012; Zhang et al., 2013; Davy, 2018). We followed exactly the steps detailed in Seidel et al. (2012) and Zhang et al (2013) to calculate the PBLH. Note that all PBLH estimates in this study are measured in meters above ground level (AGL). The data details are mostly listed in Tables 1 and 2, with some key information provided next.

2.1.1. Radiosonde measurements

Two observational PBLH datasets were derived from atmospheric soundings (mostly at 00 and 12 Coordinated Time Universal or UTC) in the updated Integrated Global Radiosonde Archive Version 2 (IGRA2) (Durre and Yin 2008, 2011). First, we used the Ri method to estimate the PBLH at 00 and 12 UTC for the period 1979-2019 (referred to as the IGRA2-RI, Table 2) based on the IGRA2 sounding-derived parameters (ftp://ftp.ncdc.noaa.gov/pub/data/igra/derived/).

Second, we used the daily maximum mixed layer height (MMLH) estimated by Li et al. (2020) via the parcel method (Holzworth, 1964) for the period 1979-2018 (referred to as the IGRA2-MMLH, Table 2) based on the IGRA2 and surface meteorological measurements. Because the daily MMLH usually occurs locally in the early to late afternoon, the standard two radiosonde observations each day cannot capture the fully developed mixed layer globally. The parcel method was proposed to calculate daily MMLH using radiosondes and diurnal potential temperature observations from morning to evening to characterize the convective mixing of the lower troposphere (Holzworth, 1964). It has the advantage of using twice-daily radiosonde soundings over most weather stations if the maximum virtual potential temperature coincides with the MMLH and thus has been adopted by many studies to estimate the MMLH thereafter (e.g., Seidel et al 2010, Li et al., 2020). Large uncertainties are expected over regions if this condition is not met. Due to the presence of missing data over many radiosonde stations, a set of carefully designed data selection and quality control criteria were developed to identify stations with good quality data for long-term trend analysis (section 2.2)

2.1.2. Reanalysis products

Two latest state-of-the-art reanalysis products provide physically consistent, global gridded hourly analysis fields at relatively high spatial and temporal resolutions. The second Modern-Era Retrospective analysis for Research and Applications (MERRA-2) is a NASA atmospheric reanalysis that begins in 1980 with the enhanced use of satellite observations (Gelaro et al., 2017**)**. ECMWF Reanalysis 5th Generation (ERA5) is the latest ECMWF atmospheric reanalysis of the recent global climate, produced based on historical observations since 1979 with advanced modeling and data assimilation systems (Hersbach et al., 2020). The reanalysis PBLH is estimated based on the Ri method with a critical value of 0.25 in the ERA5 (C3S, 2017) and the total eddy diffusion coefficient of heat with a threshold value of 2 m^2s^{-1} in the MERRA-2 (e.g., McGrath-Spangler and Molod 2014; Davy and Esau, 2014). The monthly means of hourly averaged PBLH, surface sensible heat flux (SHFX, W/m^2), surface latent heat flux (LHFX, W/m^2), surface skin temperature (T_S, K), 2m air temperature (T_{2m}, K), 2m dew point temperature (Td_{2m}, K), 2m specific humidity (q_{2m}, kg/kg), 2m air RH (RH_{2m}, %), and lifting condensational level (LCL, m or hPa), were analyzed for the MERRA-2 (1980-2019) and ERA5 (1979-2019). The LCL is provided by the MERRA-2 as the height of LCL (m). It is not provided in the ERA5 and so is computed as the pressure of LCL (hPa) by an iterative procedure described by Stipanuk (1973) based on surface pressure (P_s), T_{2m} and Td_{2m}. For both MERRA-2 and ERA5, RH_{2m} is not provided and so is calculated using T_{2m} and Td_{2m} based on the equation in Dutton (1976).

ERA5 also provides a 10-member reanalysis ensemble (referred to as ERA5-ensemble, Table 2) used for uncertainty estimation (CSC, 2020). The uncertainty as defined for ERA5 by the Ensemble of Data Assimilations (EDA) system only considered mostly random uncertainties in observations, sea surface temperatures (SSTs), and model physical parametrizations. Although not all uncertainties are accounted for, the mean and spread of the ensemble provide valuable information on the relative accuracy and reliability of the reanalysis data. The PBLH is estimated based on the Ri method as in the ERA5 reanalysis. The monthly means of daily mean PBLH for the 10 members were analyzed for the period 1979-2019.

We also used the Ri method to estimate the PBLH at 6-hourly intervals for the period 1980-2019 (referred to as the MERRA-2-RI, Table 2) based on the 6-hourly 3-d atmospheric instantaneous and hourly averaged 2-d near-surface fields for air temperature, humidity, geopotential height, and wind speed from the MERRA-2 output. Note that hourly averaged atmospheric analyzed fields are not available. The monthly means of daily mean PBLH (an average of the four 6-hourly values) were analyzed for the period 1980-2019.

2.1.3. CMIP6 simulations

The CMIP6 provides PBLH output available only from a subset of participating models at various spatial resolutions (Eyring et al., 2016). Different PBL schemes at different vertical resolutions are used in these models (Table 2). For example, the PBL scheme is based on the Ri number, mixing lengths, and moist non-local thermodynamic mixing in the Canadian Earth System Model (CanESM5, von Salzen et al., 2013), while the Community Earth System Model (CESM2) employed the so-called Cloud Layers Unified By Binormals (CLUBB) parameterization (Bogenschutz et al., 2018), one of the "assumed probability density function (PDF)" methods (Golaz et al., 2002; Larson et al., 2002). Unlike the traditional PBL schemes used in GCMs, the CLUBB is a third order turbulence closure that is centered around a multivariate PDF and represents a "unified" parameterization that is responsible for treating boundary layer clouds and shallow convection with one parameterization (Bogenschutz et al., 2018). Here we used two types of historical simulations from the CMIP6 archives (Eyring et al. 2016): historical (HIST) and Atmospheric Model Intercomparison Project (AMIP) runs, referred to as CMIP6-HIST and CMIP6-AMIP (Table 2), respectively. The CMIP6-HIST simulations (1850-2014) were forced with observed changes in anthropogenic and natural forcing. The CMIP6-AMIP run (1979-2014) was a standard global atmospheric general circulation model simulation for recent climate forced by observed SSTs/sea ice and prescribed external forcings. In addition, we also used the Ri method to estimate the PBLH at 6-hourly interval for the period 1980-2014 (referred to as the CMIP6-HIST-RI, Table 2) provided by Davy (2018) based on the 6-hourly 3-d atmospheric and 3-hourly 2-d near-surface instantaneous fields for air temperature, humidity, geopotential height, and wind speed from the CMIP6-HIST simulations (Table 1).

Temporal variations in the CMIP6 simulations are determined mainly by the externally imposed forcing, but also contain unforced internal variability (noise) within the atmosphere. To assess the internal variability, some models provide an ensemble of realizations with different initial conditions. For the CMIP6-AMIP and CMIP6-HIST, the available models with PBLH output at the time of analysis were chosen and only the first three realizations were obtained for the models with more than 3 realizations. For the CMIP6-HIST-RI, only the first realization for a subset of CMIP6 models, for which data were available on the model-level grid and at the output frequency needed, was chosen to estimate the PBLH as detailed in Davy (2018). In total, there are 74 simulations from 27 models: 26 AMIP runs from 12 models for CMIP6-AMIP, 35 runs from 15 models for CMIP6-HIST, and 13 runs from 13 models for CMIP6-HIST-RI. Each chosen model and its number of realizations for the CMIP6 simulations are listed in Table 1. The monthly mean of daily mean PBLH from these simulations were analyzed for the period 1979-2014.

2.2. Data processing

Here we examine the long-term PBLH trends over land between 60°N-60°S. The ocean is not considered as the major processes controlling PBLH differ primarily between land and ocean (Garratt, 1992; Seidel et al., 2012; Chan and Wood, 2013; Ho et al., 2015; Byrne and O'Gorman, 2016). The land beyond 60°N and 60°S is excluded because high-latitude continental interior regions and ice- and snow- covered surfaces have different PBLH characteristic from the mid- and low- latitudes, particularly during cold seasons (Liu and Liang, 2010; Seidel et al., 2012; Wei et al., 2017a; Davy, 2018). To attribute the changes in PBLH, we examine their statistical relationships with other key PBLH-related variables using the two high-resolution reanalysis data (i.e., ERA5 and MERRA-2) due to the complete spatial and temporal coverage.

Initial analyses reveal some similar large-scale features in the PBLH trends across all seasons except boreal winter. To reduce redundancy and the number of plots by season for different datasets, we focus on the annual mean PBLH changes, which can capture well the major large-scale PBLH trends for most seasons while minimizing signals in PBLH associated with seasonal variations (e.g., insolation, clouds, soil moisture, and SSTs). This simplicity is reasonable as observational studies generally show consistent patterns of long-term PBLH trends for all seasons over 25 weather stations in Europe (Zhang et al., 2013) and over 219 carefully selected radiosonde stations globally (Li et al., 2020).

For every station in the two radiosonde-derived PBLH datasets, the observed daily estimates of MMLH and PBLH at 00 and 12 UTC were processed into the annual mean anomaly time series for the period 1979-2018 following four steps. First, the observed daily estimates were first averaged to calculate the monthly mean. For every month, at least 10 days of data were required for the monthly averaging. Second, the monthly anomaly data was created by subtracting the long-term monthly mean (climatology) from the monthly mean data, i.e., removing the climatological seasonality. Third, the monthly anomaly data were averaged to create the annual mean for every year, and at least 6 months of data were required for the annual averaging. Finally, the long-term annual mean anomaly time series was obtained for the period 1979-2018/2019, which required at least 28 years (or 70%) of data. The use of thresholds of 10 days per month (e.g., Li et al., 2020), 6 months per year (e.g., Wang and Wang, 2016), and 70% of the temporal coverage (e.g., Gertler and O'Gorman, 2019) is a reasonable compromise between the data length, completeness, and spatial coverage. In total, 147 and 192 stations with adequate observations were chosen following the above steps for the IGRA2-MMLH and IGRA2-RI PBLH datasets, respectively. The data coverage is reasonably good over most regions in the Northern Hemisphere but relatively poor over remote deserts and in the Southern Hemisphere.

It is well recognized that PBLH is strongly coupled with land-atmospheric interactions. Despite the complexity and uncertainty in representing land-atmosphere interaction in different weather and climate models, it was found that the ensemble mean (EM) of forecasts driven by different initial conditions can reduce forecast uncertainties that result from errors in initial conditions (e.g., Guo et al., 2007; Hofer et al., 2012; Potter et al., 2018). It is also found that the multi-model ensemble mean (MEM) often outperforms most individual models in simulating the land surface component of weather and climate systems (Kharin and Zwiers, 2002; Guo et al., 2007). In particular, different PBL schemes make varying assumptions about the transport of heat, moisture, and momentum within the PBL (Lee and De Wekker, 2016). Averaging over multiple members enhances the forcing signal and reduces noise from internal variability and errors from individual members or models (IPCC, 2013). For the ERA5-ensemble, the 10-members were averaged to calculate the EM (referred to as the ERA5-EM). For the CMIP6 models, the models of single realization and the multi-realization mean of each model with more than one realization were averaged to obtain the MEM. Here we do not simply take the mean of all realizations of the available CMIP6 models to avoid biasing the MEM toward the models with a higher number of realizations. The MEM for the 12 models in the CMIP6-AMIP, the 15 models in the CMIP6-HIST, and the 13 models in the CMIP6-HIST-RI is referred to as the CMIP6-AMIP-MEM, CMIP6-HIST-MEM, and CMIP6-HIST-RI-MEM, respectively. However, the EM and MEM average out internal variability and so have a smaller magnitude in variability and trend than individual members. The monthly mean PBLHs from the individual members of the ERA5 and CMIP6 ensembles were analyzed and used to calculate the ensemble distribution, include the ensemble mean and spread, for uncertainty estimate.

The global gridded reanalysis and CMIP6 data are monthly mean values with no missing data. All data at different spatial resolutions were spatially re-projected into the common 1° by 1° grid boxes using bilinear interpolation and then were processed into the annual mean anomaly time series for every grid box. Note that the PBLH exhibits a distinct diurnal cycle. The CMIP6 data only consists of monthly means of daily mean values, while the reanalysis data contain sub-daily values and so need further processing. The reanalysis MERRA-2 and ERA5 consist of 24 hourly averaged values every month. For every grid box, we first used the 24 hourly-averaged PBLH values to obtain the long-term climatology of the PBLH diurnal cycle and then identified the five consecutive hours with maximum and minimum climatological PBLHs. Then the monthly means of hourly data were aggregated to produce the monthly means of daily mean (an average of 24 hourly values), daily maximum (an average of 5 hourly values with maximum climatological PBLHs), and daily minimum (an average of 5 hourly values with minimum climatological PBLHs). We also created the monthly means of daily PBLHs at 00 and 12 UTC for validations against the radiosonde observations. For the ERA5-ensemble members, the monthly means of daily mean PBLH were calculated from the 24 hourly-averaged PBLH values. For the MERRA-2-RI, the monthly means of daily mean PBLH were calculated from the four 6-hourly instantaneous PBLH estimates. Finally, for every grid box, the monthly mean data were processed into the annual mean anomaly time series for each individual data period following the last three steps done above for the radiosonde data. In total, there are 12660 land grid boxes of 1° by 1° between 60°N-60°S across a wide range of atmospheric conditions and surface characteristics.

2.3. Methods of large-scale trend analysis

Large differences and uncertainties exist among different PBLH estimates and different datasets (e.g., Seidel et al., 2010; McGrath-Spangler and Molod, 2014; McGrath-Spangler et al., 2015; Wei et al., 2017b; Ao et al., 2017). In order to synthesize the differences and cope with the uncertainties, we use a multi-data synthesis approach to make inferences on the robustness and consensus of the long-term trends across the aforementioned different datasets. The emphasis is more on the sign and significance of the trend and less on the magnitude as the changes are more likely robust if more independent datasets agree on the direction and statistical significance of the changes (e.g., Power et al., 2012; IPCC, 2007; 2013; Dosio et al., 2019). We identify and focus on large-scale regions where the change signal is considered to be robust and consistent if all (or 100%) of the datasets show a statistically significant trend ($p<0.05$) and agree on the sign of the trend, to increase our confidence in the obtained results.

Two widely used methods were used to quantify the magnitude and significance of the trend of the annual mean anomaly time series processed above for any variable over the study period. The first was to use the ordinary least squares regression (OLS) to estimate the linear trend (i.e., slope) combined with a two-tailed student's *t* test for significance test, referred to as the OLS method. The second was to use the Theil-Sen slope estimate to assess the monotonic trend (linear or non-linear) combined with the Mann–Kendall test (Theil, 1950; Sen, 1968; Kendall 1970; Dytham, 2011) for significance test, referred to as the M-K method. The M-K method is a non-parametric (i.e., distribution-free) test and is much less sensitive to outliers and skewed distributions compared to linear regressions (IPCC, 2013). Trend analysis was performed for the annual mean anomaly time series for all data variables created in section 2.2.

We also performed a detailed time series analysis over the entire study domain (60°N-60°S) and two chosen regions (the SDAP and India) where a consensus on the PBLH trends was identified. To maximize large-scale PBLH change patterns and minimize local-scale variability, we aggregated the PBLH and related data via spatial averaging at two spatial scales: (1) station mean and (2) regional mean. The former is simply an arithmetic mean of individual station data and was applied to both the observational and reanalysis data. For the reanalysis, the station data were obtained from the grid boxes where the chosen stations are located based on their geographic location (latitude and longitude). The latter was applied only to the two reanalyses and was calculated using area-weighted averaging over the land grid boxes between 60°S and 60°N, SDAP (18°N-31°N, 5°W-50°E) and India (17°N-34°N, 68°E-96°E), depicted by the rectangle box in Fig. 3c. We calculated the Pearson's correlation coefficient (referred to as R) to quantify the temporal association between two times series or the spatial similarity between two variables.

Note that every variable analyzed in this study is an annual mean quantity. For brevity, the annual mean of daily mean PBLH, daily maximum PBLH, daily minimum PBLH, daily PBLH at 00 UTC, and daily PBLH at 12 UTC, five frequently used variables, are referred to as $PBLH_{mean}$, $PBLH_{max}$, $PBLH_{min}$, $PBLH_{00}$, and $PBLH_{12}$ hereafter, respectively. Also the term "annual mean" is often omitted for the remainder of this paper.

3. Results and Discussion

3.1. Climatology and trends in $PBLH_{mean}$ for the reanalysis and CMIP6

Figure 1 shows the spatial patterns of climatological $PBLH_{mean}$ (m) from the four reanalysis and three CMIP6 datasets. Overall, the PBLH climatology exhibits similar spatial patterns across the different datasets, with the deepest PBL over low latitude drylands and the shallowest in high latitudes and humid tropical regions. For the 12660 land grid boxes between 60°S-60°N, the spatial correlation (R) between the ERA5 and the other 6 datasets, MERRA-2, ERA5-EM, and MERRA-2-RI, CMIP6-AMIP-MEM, CMIP6-HIST-MEM, and CMIP6-HIST-RI-MEM, are 0.70, 0.99, 0.67, 0.76, 0.76 and 0.75, respectively. These coefficients are all statistically significant ($p < 0.0001$). At the grid box level, the minimum, maximum, mean and standard deviation (STD) of $PBLH_{mean}$ are 134.4, 1209.1, 585.1, and 155.1 meters for the ERA5, and 335.0, 2229.6, 937.5, and 336.8 meters for the MEAAR-2, respectively. These values in the ERA5 are almost identical to those in the ERA5-EM and mostly comparable to those in the MERRA-2-RI and three CMIP6 MEMs, while the MERRA-2 has much larger values than the MERRA-2-RI. The major differences among these seven datasets include: (1) the three CMIP6 MEMs exhibit a much smaller spatial range in $PBLH_{mean}$, due to a much larger minimum and a smaller maximum, than the ERA5, (2) the MERRA-2 has a much larger values in the minimum, maximum, mean and STD than any other datasets. These results are generally consistent with previous results (e.g., Svensson and Lindvall, 2015; Wei et al., 2017a; Davy, 2018; Zhou, 2020).

Figure 2 shows the spatial patterns of $PBLH_{mean}$ trend (m/decade) estimated using the M-K method from the reanalysis and CMIP6 datasets. Widespread positive trends are seen in the ERA5 over most land areas, except South Asia, Western Australia, and most of Canada, and Central and Eastern U.S. (Fig. 2a). In comparison, fewer grid boxes have significant trends in the MERRA-2, with strong and significant increasing trends in the Brazilian Highlands, most of Africa and the Middle East, and significant decreasing trends in the Indian subcontinent, Australia and Eastern China (Fig. 2b). The trend in the ERA5-EM (Fig. 2c) is very similar to that in the ERA5. The MERRA-2-RI (Fig. 2d) shares similar trend patterns to the MERRA-2 but in a much smaller magnitude and a slightly smaller spatial extent. The three CMIP6 MEMs (Fig. 2e-2f) share similar significant increasing trends over most areas in the Northern Hemisphere (e.g., the SDAP, continental U.S., and Europe), except West and South Asia, but widespread insignificant trends across the Southern Hemisphere. Note that the CMIP6 trends over coastal regions and islands differ from the reanalysis results due to the coarse spatial resolution of the models and the impacts of marine PBL. In general, the trends are largest between 30°N-30°S in the reanalysis datasets and between 30°N-60°N in the CMIP6 models. Among the 12660 land grid boxes between 60°S-60°N, 51.3% (4.1%), 18.8% (24.2%), 49.7% (4.3%), 20.5% (18.2%), 30.6% (12.8%), 36.0% (23.5%), and 33.4% (16.7%) exhibit a statistically significant increasing (decreasing) trend at the 5% level, for the ERA5, MERRA-2, ERA5-EM, MERRA-2-RI, CMIP6-AMIP-MEM, CMIP6-HIST-MEM, and CMIP6-HIST-RI-MEM, respectively. When the trend is estimated using the OLS method, the results remain very similar to these in the M-K method, except that the percentage of grid boxes with significant trends are slightly higher (Fig. S1). Evidently, the PBLH trend demonstrates large differences in the sign and magnitude over many areas, but significant trends are consistently seen at large scales over the SDAP and Indian subcontinent.

To identify the large-scale spatial patterns with a consensus on the $PBLH_{mean}$ trends among different datasets, we calculate a consistency index (CI) as the number of datasets with the same sign of trends that are statistically significant at the 5% level. The spatial patterns of the CI based on the M-K method are shown, respectively, for the four reanalysis (Fig. 3a), three CMIP6 (Fig. 3b), and seven reanalysis+CMIP6 (Fig. 3c) datasets. For example, an index value of +4 (-4) in Fig. 3a indicates that all four reanalyses share similar and significant upward (downward) trends. Figure 3a highlights the large-scale consensus on the significant increasing trends (in red, CI=4) over North Africa, West Asia, Central Africa, and Brazil, and along the Mongolia–Russian Siberia borders, but on the significant decreasing trends (in blue, CI=-4) over the Indian subcontinent. Note that the MERRA-2-RI (Fig. 2d) has a much smaller area in red (e.g., over the SDAP) than the other three reanalyses (Figs 2a-2b) because its $PBLH_{mean}$ is estimated from four 6-hourly instantaneous atmospheric fields (rather than from 24-hourly averaged fields for the other reanalyses), which have much larger interannual variability and thus fewer grid boxes with significant trends (Table 2). Evidently, positive trends dominate over a broad, contiguous swath of land covering the SDAP while negative trends are relatively smaller in spatial extent and cover mostly India. The three CMIP6 datasets (Fig. 3b) show consistent significant increasing trends (in red, CI=3) over many areas in North Africa, the contiguous U.S., and Eastern Canada, but significant decreasing trends (in blue, CI=-3) over the Indian subcontinent and China. Note that their area in red over the SDAP (Fig. 3b) is smaller than the entire desert areas with increasing

PBLH trend in the reanalysis (Figs 2a-2c) because of the impacts of marine PBL over coastal regions in the periphery of the deserts in coarse resolution GCMs. When both the four reanalysis and three CMIP6 datasets are considered together (Fig. 3c), only the SDAP and India stand out with consistent trends. To confirm this further, we also performed the same analysis to Fig. 3 but using the OLS method and obtain almost identical results (Fig. S2), indicating that the results of $PBLH_{mean}$ trends are independent of the methods used for the trend estimate. The seven datasets highlight a consensus and robustness on a large-scale pattern of rising PBLH over the enormous SDAP (18°N-31°N, 5°W-50°E) and decreasing PBLH over India (17°N-34°N, 68°E-96°E), the two rectangle boxes depicted in Fig. 3c. For brevity, our remaining paper will focus mostly on the PBLH estimated from the M-K method and the SDAP and India as two regional hotspots.

Figure 4 shows the probability distribution function (PDF), or frequency of occurrence, of $PBLH_{mean}$ trends (m/decade, Fig. 2) that are statistically significant at the 5% level over land between 60°S-60°N (in blue), the SDAP (in red) and India (in green). Despite the large differences in the trend magnitude, the reanalyses (Figs. 4a-4d) clearly show a tendency toward larger and more positive (negative) trends over the SDAP (India) than the entire region of 60°S-60°N. Such PDF differences are also evident in the CMIP6 MEMs (Figs. 4e-4g). Note that the EM and MEM enhance the forcing signal (e.g., global warming signal) and reduce internal variability and model uncertainty. All indicate that the SDAP (India) has a higher frequency of occurrence of large and positive (negative) PBLH trends than the entire study domain. We also calculate the PDF to demonstrate the ensemble spread of $PBLH_{mean}$ trend among the ensemble members in the ERA5 (Fig. 5a) and CMIP6 models (Figs. 5b-5d). The uncertainty associated with initial conditions, observations, SSTs, and model physical parametrizations in the ERA5-ensemble is small, and so all ensembles exhibit similar features to the ERA5-EM. The CMIP6 models exhibit large inter-model differences, but most models demonstrate a higher frequency of occurrence of increasing (decreasing) PBLH over the SDAP (India) than the entire study domain.

Figures 1-5 illustrate the $PBLH_{mean}$ climatology and trends at the grid box level. To focus on the large-scale features, Table 3 lists the climatology (m) and trend (m/decade) of regional mean $PBLH_{mean}$ for the reanalysis and CMIP6 datasets averaged over the entire study domain (60°S-60°N), SDAP, and India. The MERRA-2 has substantially larger climatological values, 966.9 (60°S-60°N), 1453.2 (SDAP), and 1265.2 (India) meters than the other 6 datasets, which range from 536.8-717.2, 691.3-843.2, and 594.4-761.7 meters, respectively. All datasets show consistent and significant trends, positively in the SDAP and negatively in India, and the resulting trends estimated using the M-K and OLS are very similar. Like the climatology, the MERRA2 has much larger trends than the other datasets. For the region between 60°S-60°N, only the two ERA5 reanalyses and one of the CMIP6 data show statistically significant upward trends. The lack of significant trend over the entire domain in the two MERRA-2 reanalyses is a result of smoothing the spatially heterogeneous trends (Figs. 2b and 2d).

3.2. Comparisons with radiosonde observations at daytime

The above results show the $PBLH_{mean}$ changes for the gridded reanalysis and CMIP6 datasets. Over land, the PBLH has a strong diurnal cycle. It is typically shallow and stable at night because of longwave radiative cooling but grows deep and unstable at daytime because of solar heating (Stull, 1988; Liu and Liang, 2010). Next, we use three compositing methods to validate the reanalysis results against the radiosonde measurements but focus on the daytime PBLH for two reasons. First, most PBLH changes are expected to occur at daytime due to solar heating (Seidel et al., 2012; Chan and Wood, 2013; Brahmanandam et al., 2020). Second, the PBLH is more difficult to quantify and have larger uncertainty for the stable than unstable regime and so is better estimated at daytime in terms of quality by current methods using radiosonde, reanalysis and GCM data (e.g., Seidel et al., 2012).

First, we compare the observed MMLH and reanalysis $PBLH_{max}$ (termed the compositing method A) which represents the daytime maximum PBLH and is used to approximate the MMLH estimated from the radiosonde observations. Note that temporal sampling may differ largely between these two datasets over some regions (section 2.1.1). Figure 6a shows the spatial patterns of MMLH trends estimated from the IGRA2-MMLH. Among the 147 stations between 60°S-60°N, 29.9% (8.2%) have a statistically significant increasing (decreasing) trend at the 5% level. In particular, the radiosonde data do exhibit coherent and large-scale spatial patterns of increasing trends over the SDAP. Similar increasing (decreasing) trends are also seen over Europe (India), consistent with previous observational studies (Zhang et al., 2013; Li et al., 2020). Figures 7a and 7b show the scatter plot of the MMLH trends and corresponding reanalysis $PBLH_{max}$ trends for 63 stations that have a statistically significant trend ($p<0.1$) in the IGRA2-MMLH. Evidently, the MMLH has much more stations with increasing trends than those with decreasing trends and it has much larger trends than the reanalyses. Its correlation R is 0.12 ($P=0.35$) with the ERA5 and 0.38 ($p<0.01$) with the MERRA-2, indicating the similar sign for most trends but large differences in the magnitude. This weak correlation is partially expected considering the sampling issue mentioned previously. However, at the regional scale, the MMHL could be useful over the SDAP where the sampling issue is minor (see more results later)

Second, we composite the $PBLH_{00}$ and $PBLH_{12}$ trends from the IGRA2-RI to better match the reanalysis and observations and to avoid the above temporal sampling issue. The daytime radiosonde trend is chosen from one of the two IGRA2-RI trends at 00 and 12 UTC (e.g., $PBLH_{00}$ and $PBLH_{12}$) whose climatological PBLH is larger (termed the compositing method B). The reanalysis daytime trend is determined from the chosen UTC accordingly. Figure 6b shows the spatial patterns of daytime PBLH trends estimated from the IGRA2-RI. Among the 192 stations between 60°S-60°N, 60.9% (4.7%) have a statistically significant increasing (decreasing) trend at the 5% level. The radiosonde data exhibit a much higher percentage of positive trends than Fig. 6a and also increasing trends over the dry Arabian Peninsula and Europe, similar to Fig 6a. Figures 7c and 7d show the scatter plot of the observed and reanalysis daytime PBLH trend for 133 stations with a statistically significant trend at the 10% level in the IGRA2-RI. The correlation R is 0.47 ($p<0.01$) for the ERA5 and 0.50 ($p<0.01$) for the MERRA-2, much stronger than the R values in Figs. 7a and 7b.

Third, we composite the $PBLH_{00}$ and $PBLH_{12}$ trends from the IGRA2-RI to match the reanalysis and observations following the same logic in the compositing method B. The daytime radiosonde trend is chosen from one of the two IGRA2-RI trends (e.g., $PBLH_{00}$ and $PBLH_{12}$) based on the local solar time as done in Wang and Wang (2016), and the reanalysis daytime data is chosen accordingly (termed the compositing method C). The spatial patterns of observed daytime PBLH trends and their percentages of the observed significant trends (Fig. 6c) and the scatter plot between observed and reanalyzed (Figs. 7e and 7f) show similar results to the compositing method B (Figs. 6b, 7c and 7d).

Next, we validate the two reanalysis results against the observations from available radiosonde stations at the regional scale as it is essential that the reanalyses can at least capture the major PBLH features observed over the SDAP and India where consistent trends are identified in section 3.1. The focus is on the station mean variability instead of individual stations to maximize large-scale patterns and minimize local influences. The daytime PBLHs for the reanalysis and radiosonde data are compared for the compositing method A and B. The compositing method C is not shown due to its similarity to the compositing method B.

Figures 8a and 8c show the station mean interannual variations in observed MMLH and reanalysis $PBLH_{max}$ anomalies (i.e., the compositing method A) from 5 stations over the SDAP. The PBLH exhibits a persistent and statistically significant ($p<0.01$) upward trend and shares similar interannual variability in the observed and reanalyzed data. The increasing trend is 98.4 m/decade for the observations, 31.2 m/decade for the ERA5, and 18.7 m/decade for the MERRA-2. Although underestimating the observed trend, the reanalysis PBLH shows a statistically positive correlation with the observed PBLH, with R=0.76 ($p<0.01$) in the ERA5, and 0.69 ($p<0.01$) in the MERRA-2. These results indicate that the observed long-term trend and interannual variability in daily MMLH are generally captured by the reanalyses reasonably well over the SDAP and the ERA5 is closer to the observations than the MERRA-2. Figures 8b and 8d are similar to Figs 8a and 8c but from 9 stations in India. The reanalysis PBLH exhibits a significant decreasing trend ($p<0.01$) while the observed data show negative but insignificant trends. In particular, the observed PBLH exhibits opposite trends between the first and last 20 years. It correlates significantly with the ERA5 (R=0.53, $p<0.01$) but insignificantly with the MERRA-2 (R=-0.10, p=0.59). It is well documented that Indian radiosonde data contain large inhomogeneities due to frequent instrument changes and other causes (e.g., Lanzante et al. 2003; Thorne et al. 2005; Zhou et a., 2020), which may help to explain the opposite trends and the poor correlation. In addition, the PBLH over the Indian subcontinent is characterized by complex topography and heterogeneous land surface, coupled with the Indian monsoon and various soil-vegetation-atmosphere interactions (e.g., Sathyanadh et al., 2017). This complex along with the data quality issues result in low confidence even in homogenized datasets because of the very poor quality and abnormally large variances in the raw data (Zhou et a., 2020).

Figure 9 shows the station mean interannual variations in the observed and reanalysis daytime PBLH anomalies (i.e., the compositing method B, Figs. 6b, 7c and 7d) from 5 stations available over the Arabian Peninsula and 1 station available in India from the IGRA2-RI. Again, significant increasing trends and strong correlations are evident in the dry Arabian Peninsula ($R=0.72-0.76$, $p<0.01$), while weak and insignificant correlations are seen in India, where missing data is evident in the radiosonde measurements.

We need to realize that the radiosonde data are point measurements, while the reanalysis values are averaged over a grid box at a much coarser resolution (Chan and Wood, 2013). It is difficult for the reanalysis data to match the observed PBLH trend that is localized in space and time. Also, radiosonde profiles are measured twice a day at specified synoptic times (00 and 12 UTC), have missing data over many stations, and are often insufficient in vertical resolution for most data, which can create large fluctuations in PBLH estimate (Liu and Liang, 2010). Atmospheric reanalyses produced at various institutes have substantially improved in quality as a result of better models, better input data, and better assimilation methods (Dee et al., 2011). Despite much progress made in reducing uncertainties in assimilating various types of observations, current reanalysis data still suffer from artifacts largely due to the global observing system changes and cannot well represent near-surface variables such as surface energy flux and moisture, which could produce complications on climate studies, especially regarding low-frequency trends at regional scales (Robertson et al., 2014; Gelaro et al., 2017; Bosilovich et al., 2017). Hence, large differences in PBLH trends between observed and reanalyzed exist at local to regional scales. Nevertheless, the reanalysis PBLHs reproduces the observed trend and interannual variability over the SDAP, while differ from observations over India due to radiosonde data quality issues.

3.3. Statistical relationships between PBLH and related variables

Previous studies have documented that changes in PBLH are strongly correlated with changes in surface sensible heat, temperature, and RH (Zhang et al., 2013; Zhao et al., 2017; Darand and Zandkarimi, 2019; Li et al., 2020). Here we perform similar statistical analyses between PBLH and several key PBLH related variables using the two high-resolution reanalysis datasets.

Figure 10 shows the scatter plots between the trends in daily maximum PBLH, SHFX, Ts and RH_{2m} for the ERA5 (left panels) and MERRA-2 (right panels). Only the grid boxes with a statistically significant trend in PBLH ($p<0.05$) over land between 60°N-60°S are included. As expected from theory, the PBLH trend is positively correlated with the trend in SHFX and T_s, and negatively correlated with the trend in RH_{2m}. For example, R is 0.77 with SHFX, 0.69 with T_s, and -0.72 with RH_{2m} for the 7590 grid boxes in the ERA5, and the corresponding value is 0.85, 0.91 and -0.91 for the 4696 grid boxes in the MERRA-2. Considering the large sample size, the R values are extremely strong and all statistically significant ($p<0.001$), indicating a strong spatial coupling between the paired trends.

To further examine the above relationships, we also calculate the correlation for three temporally averaged PBLHs (daily maximum, daily minimum, and daily mean) and consider more PBL related variables. Table 4 lists the R values for SHFX, LHFX, LCL, Ts, T_{2m}, Td_{2m} (q_{2m}), and RH_{2m} using both the M-K and OLS methods. Evidently, PBLH trends are correlated positively with the trends in variables related to surface heating (SHFX, T_{2m}, and Ts), and negatively with the trends in variables related to surface moisture (LHFX, Td_{2m}, q_{2m}, and RH_{2m}). Note that changes in LCL are related to both surface heating and humidity as lower surface RH (i.e., warmer and/or drier air) results in higher LCL (and PBLH as well). The correlation between PBLH and LCL is negative in the ERA5 as the LCL is expressed as the pressure (hPa), not the height (m) at the LCL that is used in other datasets. Overall, the R values are extremely strong and all statistically significant ($p<0.001$) for the daily maximum and daily mean PBLH considering the large sample size. For example, the 28 R values for the daily maximum PBLH range between 0.50 and 0.92, with 21 of them exceeding 0.70. In general, the R values are comparable for the daily maximum and daily mean PBLH, but are much weaker for the daily minimum PBLH at nighttime. Interestingly, the correlations are much stronger in the MERRA-2 than the ERA-5.

Figure 11 shows the regional mean interannual variations in daily mean PBLH, SHFX, T_s, and RH_{2m} for the ERA5 (left panels) and MERRA-2 (right panels) averaged over the entire study domain (60°S-60°N). The PBLH shows long-term upward trends consistent with the increase in sensible heat and surface warming and the decrease in RH_{2m} with time. At the interannual scale, the PBLH is correlated positively and significantly with SHFX (R=0.96) and T_s (R=0.88) and negatively with RH_{2m}. (R=-0.96) in the ERA5. The correlations in the MERRA-2 data are similar but slightly weaker. Unlike the persistent increasing trend in the ERA5, the MERRA-2 PBLH exhibits an increasing trend from 1980 to 2002 but a downward trend thereafter.

Figures 12 is similar to Fig. 11 but for the SDAP. Again, the PBLH shows positive trends consistent with the increase in sensible heat and surface warming and the decrease in RH_{2m} with time. At the interannual scale, the PBLH correlates positively and significantly with SHFX (R=0.75) and T_s (0.75) and negatively with RH_{2m} (R=-0.32) in the ERA5 and the correlations are slightly weaker in the MERRA-2. The weak correlation with RH_{2m} is expected given the limited moisture availability over the deserts. Figures 13 is similar to Fig. 11 but for the India. In contrast to the SDAP, the PBLH shows negative trends consistent with decreasing SFHX and T_s and increasing RH_{2m}. At the interannual scale, the PBLH correlates positively and significantly with SHFX (R=0.82) and T_s (0.28) and negatively with RH_{2m} (R=-0.92) in the ERA5 and the correlations are slightly weaker in the MERRA-2. The weak and insignificant R values with T_s are expected given increasing moisture availability (and thus latent heat) over India.

Besides the results for $PBLH_{mean}$ shown in Figs 11-13, we also calculate the correlation for $PBLH_{max}$ and $PBLH_{min}$ and consider more PBL related variables. Table 5 lists the R values for SHFX, LHFX, T_s, T_{2m}, and RH_{2m}. Evidently, PBLH is correlated positively and mostly significantly with the variables related to surface heating (SHFX, T_{2m}, and T_s), and negatively with the variables related to surface moisture (LHFX and RH_{2m}). Overall, the R value is extremely strong and all statistically significant ($p<0.001$) for $PBLH_{max}$ at daytime and $PBLH_{mean}$, while the

R value is much weaker for $PBLH_{min}$ at nighttime, consistent with the R values for the trends (Table 4). Note that the global mean time series (60°S-60°N) is mostly determined by large-scale external forcing and so has a much larger R value than the regional mean time series (e.g., SDAP and India) that is also affected by local to regional factors (e.g., clouds and SSTs).

3.4. Physical explanations for PBLH trends

The Earth is mainly warmed bottom up, as most solar radiation is absorbed at the surface and this energy is transmitted to the rest of the atmosphere through PBL processes. There exists a high level of complexity and heterogeneity of various factors in controlling the PBLH changes at multiple spatial and temporal scales under different surface and atmospheric conditions (Stull 1988). Inherently, the PBLH is particularly sensitive to soil moisture, vegetation, and terrain (Talbot et al., 2007; Liu and Liang, 2010; Seidel et al., 2012; Zhang et al., 2013; Lee and De Wekker, 2016; Wei et al., 2017a; Sathyanadh et al., 2017), and thus exhibits a far more heterogeneous picture than other variables such as temperature (e.g., Donat et al., 2014). For example, soil moisture determines the partitioning of net radiation between sensible and latent heat and thus the PBLH; its temporal change can significantly modify the PBLH from daily to interannual time scales and its spatial change can largely determine the spatial heterogeneity in PBLH (Guo et al., 2007; Lee and De Wekker, 2016). The spatiotemporal variations in these surface conditions can substantially affect both the magnitude and sometimes the sign of the PBLH trends at local to regional scales.

There exists a wide range of complexity and uncertainty in PBLH estimates among different datasets and methods. PBLH is one key measure of the strength of the PBL processes but lacks a unified definition. Hence a variety of methods have been used to estimate the PBLH, and different methods can produce substantially different values, even for the same atmospheric profile (e.g., Seidel et al., 2010; McGrath-Spangler and Molod, 2014). The radiosonde-based PBLH estimates have limited spatial and temporal coverage and suffer from inhomogeneities (Thorne et al., 2011; Haimberger et al., 2012). The reanalysis PBLH is a model-based estimate and so is prone to model deficiencies (e.g., McGrath-Spangler and Molod, 2014; McGrath-Spangler et al., 2015; Wei et al., 2017a; Zhou, 2020) and artifacts and non-physical trends largely due to the global observing system changes (e.g., Dee et al., 2011; Robertson et al., 2014; Gelaro et al., 2017; Bosilovich et al., 2017). It has been well documented that current weather and climate models have difficulties and large uncertainties in accurately representing key PBL processes (Garcia-Carreras et al. 2013; Holtslag et al. 2013; Wei et al., 2017b; Ao et al., 2017).

Our results show a large spread in the magnitude and sign of PBLH trends among different datasets over many regions. This is not surprising at the global scale considering the difficulty and uncertainty in estimating PBLH and the complexity and heterogeneity of PBLH changes discussed above. Thus, it is challenging and somewhat impossible to validate the global long-term PBLH trends in the reanalysis and GCM datasets using radiosonde observations with limited spatial and temporal coverage. In order to synthesize the differences and cope with the uncertainties, we use a multi-data synthesis approach from an ensemble of different datasets to identify large-scale regions where the change signal is considered to be robust and consistent by considering both the sign and statistical significance of the trends by all datasets (Power et al., 2012; Dosio et al., 2019).

We assume that the PBLH estimate methods are based on well-established physical principles and so are more likely to reach a consensus on the direction (or sign), than the magnitude, of the change.

Our results indicate strong spatial coupling in the long-term trends between PBLH and surface heating and moisture variables, particularly in the daytime. Over land, the PBLH growth is driven mainly by surface heating and static stability (Chan and Wood, 2013; Lee and De Wekker; 2016; Ao et al., 2017; Brahmanandam et al., 2020). Warmer and drier surfaces result in greater sensible heat flux and PBLH, and so PBLH is strongly correlated with changes in near-surface sensible heat, temperature, and RH (Zhang et al., 2013; Darand and Zandkarimi, 2019; Li et al., 2020). Our statistical analyses (Table 4, Fig. 10) show significant correlations in the long-term trend, positively between PBLH and variables related to surface heating (SHFX, T_{2m}, and T_s), and negatively between PBLH and variables related to surface moisture (LHFX, Td_{2m}, q_{2m}, and RH_{2m}). These correlations are stronger at daytime than nighttime because of the close daytime coupling between PBLH and solar heating (e.g., Liu and Liang, 2010; Zhang et al. 2013; Lee and De Wekker, 2016). Our reported relationships are consistent with previous studies (Zhang et al. 2013; Chan and Wood, 2013; Darand and Zandkarimi, 2019).

Our results highlight a consensus on increasing PBLH trends among different datasets over the SDAP. The SDAP is among the driest and hottest regions on Earth and has limited soil moisture, vegetation, and cloudiness. As discussed previously, surface heating via sensible heat is documented as the dominant driver in determining the convective PBLH growth over arid areas because of limited availability of surface moisture. It is well known that drier regions with less soil moisture and vegetation are associated with higher Bowen ratios and tend to experience larger warming rates due to more sensible heat flux and less local evaporative cooling (Zhou et al. 2007; 2009; 2010). Increased downwelling longwave radiation (DLR) associated with large-scale warming and moistening in response to increasing GHGs has been identified as the primary surface radiative forcing for the amplified surface warming associated with DA over the SDAP (Zhou et al., 2015; 2016; Cook and Vizy 2015; Zhou 2016; Evan et al., 2017; Wei et al., 2017b). This positive radiative forcing is converted mainly into sensible heat over the dry deserts, which enhances surface heating and deepens PBLH via elevated turbulent mixing in the PBL.

Our results also highlight a consensus on decreasing PBLH trends among different datasets over the Indian subcontinent. Indian monsoon precipitation has intensified over the past three decades, while drying trends are seen over the SDAP (e.g., Wang et al., 2012; Jin et al., 2014). This is consistent with the well-coupled monsoon-desert mechanism (e.g., Rodwell and Hoskins, 1996; Sun et al., 2019; Kim et al., 2019) and with GCM-based prediction of intensified monsoon in a warming climate (e.g., Chen et al., 2020; Wang et al., 2020). For example, Hoskins (1996) proposed that the drying trend in the arid regions resulted from the increased descent produced by the monsoon heating-induced Rossby waves that interact with subtropical westerly flows. The contrast changes in PBLH and T_{2m} between the SDAP and India seem to support this monsoon-desert coupling. In addition, our results in Fig. 13 and Table 4 also show close connections in the trend and interannual variation between PBLH and near-surface moisture variables (e.g., RH_{2m} and latent heat) over India, resonating with intensified monsoon precipitation. The reanalysis results in India, however, cannot be validated using reliable radiosonde observations.

Our results suggest that the reanalysis PBLH estimate might be more reliable in the ERA5 than the MERRA-2. The reanalysis PBLH is estimated based on the Ri method in the ERA5 and the total eddy diffusion coefficient of heat in the MERRA-2 (e.g., McGrath-Spangler and Molod 2014; Davy and Esau, 2014). Among various methods used to estimate the PBLH, the algorithms based on the Ri method were proved to be most suitable for application to large radiosonde, reanalysis, and GCM data sets (Seidel et al., 2012; McGrath-Spangler and Molod, 2014). Our PBLH estimates using the Ri method from the MERRA-2-RI is much smaller than the PBLHs from the MERRA-2 and comparable to the ERA5 estimate. Also, systematic biases were documented in the MERRA-2 PBLH, particularly at nighttime (e.g., McGrath-Spangler and Molod, 2014; Svensson and Lindvall 2015; Dang et al., 2016; Davy, 2018; Zhou, 2020). Furthermore, our validations using the radiosonde data show the ERA5 PBLH is closer to the observations than the MERRA-2. If this is the case, rising PBLH is likely more widespread spatially than that seen in the MERRA-2.

Our results indicate large discrepancies among different PBLH datasets, which are likely due to the differences in spatial resolution (point measurements versus coarse-resolution grid averaged data), observational uncertainties, and deficiencies in modeling the surface radiative forcing, surface energy partitioning, and PBL mixing. The land surface and PBL change in response to external forcings are a result of complex interactions among the atmosphere, PBL and land surface. Considering the complexity of turbulent mixing and the challenges in observing and modeling the PBL processes, it is very challenging to attribute the differences among different PBLH datasets in the fully coupled land-atmosphere system. For example, the systematic underestimated diurnal range in the PBLH and surface air temperature has been a long-standing issue in reanalysis and numerical models (e.g., Wei et al., 2017a; 2017b; Du et al., 2018; Davy, 2018). As the focus of the present study is the detection of convergent PBLH trends, further attribution of these differences is beyond the scope of this paper and will be explored in future studies.

4. Conclusions

This paper examines the large-scale patterns of long-term PBLH trends over land between 60°S and 60°N. Different nine types of datasets consisting of radiosonde observations, reanalysis products and climate model simulations are evaluated over the satellite era for the period 1979-2019. To synthesize the differences and cope with the uncertainties among different PBLH estimates, a multi-data synthesis approach is used to make inferences on the robustness and consensus of the long-term trends across different datasets. The emphasis is more on the sign and significance of the trend and less on the magnitude. We identify large-scale regions where all datasets (or 100%) show a statistically significant trend ($p<0.05$) and agree on the sign of trends, to increase our confidence in the obtained results. The testable hypothesis is that the global warming signal is manifest most in terms of the spatial extent of PBLH change over the SDAP where the amplified surface warming associated with DA enhances turbulent mixing and thus raise the PBLH height. Despite methodological uncertainties and data limitations, the main findings of this study are summarized as follows:

1. Large differences in long-term PBLH trends among different datasets are found over many regions – expressed in different magnitudes and/or signs of trends. This spread reflects the difficulty and uncertainty in estimating PBLH and the complexity and heterogeneity of various factors in controlling the PBLH changes under different surface and atmospheric conditions.

2. There is strong spatial coupling in the long-term trends between PBLH and related key variables, particularly in the daytime. There are statistically significant correlations in the trends, positively between PBLH and variables related to surface heating and negatively between PBLH and variables related to surface moisture. The reported relationships are consistent with theory and previous findings in the literature.

3. Different reanalysis and GCM datasets indicate consistently coherent and large-scale spatial patterns of rising PBLH over the enormous SDAP and declining PBLH in India. Consistent PBLH trends also exist in other regions but are much smaller in spatial extent than the SDAP and in a subset of the nine datasets used. The radiosonde data exhibit similar spatial features of increasing PBLH over the SDAP and the reanalysis data generally capture the observed regional mean long-term trend and interannual variability in PBLH reasonably well over the deserts. The PBLH changes over India cannot be validated due to lack of good-quality radiosonde observations.

4. One robust signal across all datasets reveals a consensus on increasing (decreasing) PBLH trends over the SDAP (India). The ensemble distribution of reanalysis and GCM PBLH trends indicates a greater coherence and a higher frequency of occurrence of rising (declining) trends over the SDAP (India) than any other regions. The rising PBLH is in good agreement with amplified surface warming associated with DA, decreasing RH, and increasing sensible heat over the SDAP, while the declining PBLH is consistent with increasing RH and latent heat and decreasing sensible heat in India in the reanalysis data.

To the best of our knowledge, this work is the very first study to identify the large-scale patterns of long-term PBLH trends in a warming climate among different datasets and establish their relationships with several key PBLH related variables at the global scale. Climate models predict consistently that DA will accelerate over the arid and semi-arid regions in the context of global warming (Zhou et al., 2016; Zhou, 2016). Along with this amplified surface heating, the PBLH is expected to rise continuously over the SDAP. The PBLH represents how deep the free atmosphere is directly influenced by the Earth's surface and responds to surface impacts. Rising PBLH indicates deeper impacts of warming deserts on the free atmosphere. This finding has important implications as the Sahara and Arabian deserts are considered to be a hotspot in terms of climate change and impacts from regional to global scales through the influence of Saharan dust and atmospheric circulation (Knippertz and Todd, 2012; Vizy and Cook, 2017; Thomas and Nigam, 2018).

Data Availability Statement. All the observational, reanalysis and climate model datasets are publicly available.

Acknowledgements. We are grateful to three anonymous reviewers for constructive comments. We acknowledge the World Climate Research Programme, which, through its Working Group on Coupled Modelling, coordinated and promoted CMIP6. We thank the climate modeling groups (listed in Table 1) for producing and making available their model output, the Earth System Grid Federation (ESGF) for archiving the data and providing access, and the multiple funding agencies who support CMIP6 and ESGF. We thank Dr. Richard Davy for providing the PBLH dataset estimated using the bulk Richardson number method from the CMIP6 historical simulations. We also thank NASA's MERRA-2 team for making their product available. L.Z. was supported by National Science Foundation (NSF AGS-1952745 and AGS-1535426).

References

Ao C O, Waliser D E, Chan S K, Li J-L, Tian B, Xie F and Mannucci A J 2012 Planetary boundary layer heights from GPS radio occultation refractivity and humidity profiles *J. Geophys. Res*. **117** D16117

Ao Y, Li J, Li Z, Lyu S, Jiang C and Wang M 2017 Relation between the atmospheric boundary layer and impact factors under severe surface thermal conditions. *Adv. Meteorol.* 8352461

Bloomfield H., Shaffrey, L., Hodges, K. I., & Vidale, P. L. (2018). A critical assessment of the long term changes in the wintertime surface arctic oscillation and northern hemisphere storminess in the ERA20C reanalysis. *Environmental Research Letters*, 13, 94004. https://doi.org/10.1088/1748-9326/aad5c5

Bogenschutz P A, Gettelman A, Hannay C, Larson V E, Neale R B, Craig C and Chen C C 2018 The path to CAM6: Coupled simulations with CAM5.4 and CAM5.5 *Geosci. Model Dev.* **11** 235–255

Bosilovich, M. G., F. R. Robertson, L. Takacs, A. Molod, and D. Mocko, 2017: Atmospheric Water Balance and Variability in the MERRA-2 Reanalysis. *J. Clim.*, 30, 1177-1196, https://doi.org/10.1175/JCLI-D-16-0338.1.

Brahmanandam P S *et al* 2020 A few important features of global atmospheric boundary layer heights estimated using COSMIC radio occultation retrieved data *Indian J. Phys.* **94** 555–563

Byrne M P and O'Gorman P A 2016 Understanding decreases in land relative humidity with global warming: Conceptual model and GCM simulations *J. Clim.* **29** 9045–9061

Chan K M and Wood R 2013 The seasonal cycle of planetary boundary layer depth determined using COSMIC radio occultation data *J. Geophys. Res. Atmos.* **118** 12422–12434

Chen Z., Zhou, T., Zhang, L., Chen, X., Zhang, W., & Jiang, J. 2020 Global land monsoon precipitation changes in CMIP6 projections. *Geophysical Research Letters*, 47, e2019GL086902. https://doi.org/10.1029/2019GL086902

Cook K H and Vizy E K 2015 Detection and analysis of an amplified warming of the Sahara Desert. *J. Clim.* **28**, 6560–6580, doi:10.1175/JCLI-D-14-00230.1

Copernicus Climate Change Service (C3S) 2017 ERA5: Fifth generation of ECMWF atmospheric reanalyses of the global climate. Copernicus Climate Change Service Climate Data Store (CDS) (access through https://cds.climate.copernicus.eu/cdsapp#!/home)

Copernicus Climate Change Service (C3S) 2020 ERA5: uncertainty estimation (created by Karl Hennermann, last modified by Paul Berrisford) (access through at https://confluence.ecmwf.int/display/CKB/ERA5%3A+uncertainty+estimation)

Dang R, Li H, Li Z and Yang Y 2016 Statistical analysis of relationship between daytime lidar-derived planetary boundary layer height and relevant atmospheric variables in the semiarid region in northwest China *Adv. Meteorol.* 5375918

Darand M and Zandkarimi F 2019 Identification of atmospheric boundary layer height and trends over Iran using high-resolution ECMWF reanalysis dataset *Theor. Appl. Climatol.* **137** 1457–1465

Davy R 2018 The climatology of the atmospheric boundary layer in contemporary global climate models *J. Clim.* **31** 9151–9173

Davy R and Esau I 2014 Planetary boundary layer depth in global climate models induced biases in surface climatology. *Atmospheric and Oceanic Physics* (physics.ao-ph). arXiv:1409–8426.

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, doi:10.1002/qj.828.

Donat M G, Sillmann J, Wild S, Alexander L V, Lippmann T and Zwiers F W 2014 Consistency of Temperature and Precipitation Extremes across Various Global Gridded In Situ and Reanalysis Datasets *J. Clim.* **27** 5019–5035

Dosio A *et al* 2019 What can we know about future precipitation in Africa? Robustness, significance and added value of projections from a large ensemble of regional climate models *Clim. Dyn.* **53** 5833–5858

Durre I and Yin X 2008 Enhanced radiosonde data for studies of vertical structure *Bull. Am. Meteorl. Soc.* **89** 1257–1262

Durre I and Yin X 2011 Enhancements of the dataset of sounding parameters derived from the Integrated Global Radiosonde Archive *American Meteorological society Annual Meeting* abstract 6B.2

Dutton J A 1976 The ceaseless wind, an introduction to the theory of atmospheric motion (1st Edition) New York: McGraw-Hill 579 pp. ISBN 0070184070

Dytham C 2011 Choosing and using statistics: A biologist's guide (3rd edition) John Wiley and Sons pp 320 ISBN 978-1-4051-9839-4

Evan A T, Flamant C, Lavaysse C, Kocha C, and Saci A 2015 Water vapor–forced greenhouse warming over the Sahara Desert and the recent recovery from the Sahelian drought, *J. Clim.*, 28(1), 108– 123, doi:10.1175/JCLI-D-14-00039.1

Eyring V, Bony S, Meehl G A, Senior C A, Stevens B, Stouffer R J, and Taylor K E 2016 Overview of the Coupled Model Intercomparison Project Phase 6 (CMIP6) experimental design and organization, *Geosci. Model Dev.*, 9(5), 1937– 1958, doi:10.5194/gmd-9-1937-2016.

Gamo M 1996 Thickness of the dry convection and large-scale subsidence above deserts. *Bound-Lay. Meteorol.* **79** 265–278

Garcia-Carreras L, Parker D J, Marsham J H, Rosenberg P D, Brooks I M, Lock A P, Marenco F, McQuaid J B and Hobby M 2015 The turbulent structure and diurnal growth of the Saharan atmospheric boundary layer *J. Atmos. Sci.* **72** 693–713

Garratt J R 1992 The Atmospheric Boundary Layer. Cambridge University Press, 316 pp.

Gelaro R and coauthors 2017 The Modern-Era Retrospective Analysis for Research and Applications, Version 2 (MERRA-2), *J. Clim.*, 30, 14, 5419-5454, https://doi.org/10.1175/JCLI-D-16-0758.1.

Gertler C and O'Gorman P 2019 Changing available energy for extratropical cyclones and associated convection in Northern Hemisphere summer *Proc. Natl. Acad. Sci.* **116.** 4105-4110

Global Modeling and Assimilation Office (GMAO) 2015a MERRA-2 tavgU_2d_slv_Nx: 2d,diurnal,Time-Averaged,Single-Level,Assimilation,Single-Level Diagnostics V5.12.4, Greenbelt, MD, USA, Goddard Earth Sciences Data and Information Services Center (GES DISC), Accessed: [Data Access Date], 10.5067/AFOK0TPEVQEK

Global Modeling and Assimilation Office (GMAO) 2015b MERRA-2 tavgU_2d_flx_Nx: 2d,diurnal,Time-Averaged,Single-Level,Assimilation,Surface Flux Diagnostics V5.12.4, Greenbelt, MD, USA, Goddard Earth Sciences Data and Information Services Center (GES DISC), Accessed: [Data Access Date], 10.5067/LUHPNWAKYIO3

Global Modeling and Assimilation Office (GMAO) 2015c MERRA-2 instU_3d_ana_Np: 3d,diurnal,Instantaneous,Pressure-Level,Analysis,Analyzed Meteorological Fields V5.12.4, Greenbelt, MD, USA, Goddard Earth Sciences Data and Information Services Center (GES DISC), Accessed: [Data Access Date], 10.5067/TRD91YO9S6E7

Global Modeling and Assimilation Office (GMAO) 2015d MERRA-2 tavgU_2d_flx_Nx: 2d,diurnal,Time-Averaged,Single-Level,Assimilation,Surface Flux Diagnostics V5.12.4, Greenbelt, MD, USA, Goddard Earth Sciences Data and Information Services Center (GES DISC), Accessed: [Data Access Date], 10.5067/LUHPNWAKYIO3

Golaz J.-C., Larson, V. E., and Cotton, W R 2002 A pdf-based model for boundary layer clouds part I: Method and model description, *J. Atmos. Sci*., 59, 3540–3551.

Guo J, Li Y, Cohen J B, Li J, Chen D, Xu H, Liu L, Yin J, Hu K and Zhai P 2019 Shift in the temporal trend of boundary layer height in China using long-term (1979–2016) radiosonde data *Geophys. Res. Lett.* **46** 6080–6089

Guo J, Miao Y, Zhang Y, Liu H, Li Z, Zhang W, He J, Lou M, Yan Y, Bian L and Zhai P 2016 The climatology of planetary boundary layer height in China derived from radiosonde and reanalysis data *Atmos. Chem. Phys.* **16** 13309–13319

Guo Z, Dirmeyer P A, Gao X and Zhao M 2007 Improving the quality of simulated soil moisture with a multi-model ensemble approach *Q. J. R. Meteorol. Soc.* **133** 731–747

Haimberger L, Tavolato C and Sperka S 2012 Homogenization of the global radiosonde dataset through combined comparison with reanalysis background series and neighboring stations *J. Clim.* **25** 8108–8131

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

Ho S, *et al* 2020 The COSMIC/FORMOSAT-3 Radio Occultation Mission after 12 years: Accomplishments, Remaining Challenges, and Potential Impacts of COSMIC-2, *Bul. Amer. Meteor. Sci.,* DOI: 10.1175/BAMS-D-18-0290.1.

Ho S, Peng L, Anthes R A, Kuo Y, and Lin H, 2015. Marine boundary layer heights and their longitudinal, diurnal, and interseasonal variability in the Southeastern Pacific using COSMIC, CALIOP, and radiosonde data. *J. Climate*, **28**, 2856–2872, https://doi.org/10.1175/JCLI-D-14-00238.1.

Hofer M, Marzeion B and Mölg T 2012 Comparing the skill of different reanalyses and their ensembles as predictors for daily air temperature on a glaciated mountain (Peru). *Clim. Dynam.* **39** 1969–1980

Holtslag A A M, Svensson G, Baas P, Basu S, Beare B, Beljaars A C M, Bosveld F C, Cuxart J, Lindvall J, Steeneveld G J, Tjernström M and Van De Wiel B J H 2013 Stable atmospheric boundary layers and diurnal cycles—challenges for weather and climate models *Bull. Am. Meteorol. Soc*. **94** 1691-1706

Holzworth G C 1964 Estimates of mean maximum mixing depths in the contiguous United States *Mon. Weather Rev.* **92** 235–242

IPCC 2007 Climate change 2007: the physical science basis, the contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change Cambridge University Press Cambridge ISBN 978-0-521-88009-1

IPCC 2013 Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change Cambridge University Press, Cambridge ISBN 978-1-107-66182-0

Jin, Q., J. Wei, and Z.-L. Yang, 2014: Positive response of Indian summer rainfall to Middle East dust, *Geophys. Res Lett*., 41, 4068–4074, doi:10.1002/2014GL059980.

Kendall M G 1970 Rank Correlation Methods London: Griffin ISBN 0-85264-199-0

Kharin V V and Zwiers F W 2002 Climate Predictions with Multimodel Ensembles *J. Clim.* **15** 793–799

Kim G-U Seo K.-H., and Chen, D., 2019: Climate change over the Mediterranean and current destruction of marine ecosystem, *Scient. Rep.*, 9, 18813, https://doi.org/10.1038/s41598-019-55303-7.

Knippertz P and Todd M C 2012 Mineral dust aerosols over the Sahara: Meteorological controls on emission and transport and implications for modeling *Rev. Geophys*. **50** RG1007

Lanzante, J. R., S. A. Klein, and D. J. Seidel, 2003: Temporal homogenization of monthly radiosonde temperature data. Part I: Methodology. *J. Clim.*, **16,** 224-240.

Larson V.E., Golaz J.-C., and Cotton W R 2002 Small-scale and mesoscale variability in cloudy boundary layers: joint probability density functions, *J. Atmos. Sci*., 59, 3519–3539.

Lee T R and De Wekker S F J 2016 Estimating daytime planetary boundary layer heights over a valley from rawinsonde observations at a nearby airport: An application to the Page Valley in Virginia, USA *J. Appl. Meteorol. Climatol.* **55** 791–809

Li J, Chu Y, Li X and Dong Y 2020 Long-term trends of global maximum atmospheric mixed layer heights derived from radiosonde measurements *Environ. Res. Lett.* **15** 034054

Liu S and Liang X-Z 2010 Observed diurnal cycle climatology of planetary boundary layer height *J. Clim.* **23** 5790– 5809

McGrath-Spangler E L and Molod A 2014 Comparison of GEOS-5 AGCM planetary boundary layer depths computed with various definitions *Atmos. Chem. Phys*. **14** 6717–6727

McGrath-Spangler E L, Molod A, Ott L E and Pawson S 2015 Impact of planetary boundary layer turbulence on model climate and tracer transport *Atmos. Chem. Phys.* **15** 7269– 7286

Molod A., H. Salmun, and M. Dempsey 2015 Estimating planetary boundary layer heights from NOAA profiler network wind profiler data, *J. Atmos. Oceanic Technol.*, 32, 1545– 1561, doi:10.1175/JTECH-D-14-00155.1.

O'Gorman P A and Muller C J 2010 How closely do changes in surface and column water vapor follow Clausius–Clapeyron scaling in climate change simulations? *Environ. Res. Lett.* **5** 025207

Poli P., et al 2013 The data assimilation system and initial performance evaluation of the ECMWF pilot reanalysis of the 20th-century assimilating surface observations only (ERA-20C). ECMWF ERA Rep 14:59. http://www.ecmwf.int/en/elibrary/11699-data-assimilation-system-and-initial-performance-evaluation-ecmwf-pilot-reanalysis

Potter G L, Carriere L, Hertz J, Bosilovich M, Duffy D, Lee T and Williams D N 2018 Enabling reanalysis research using the collaborative reanalysis technical environment (CREATE) *Bull. Amer. Meteor. Soc.* **99** 677–687

Power S B, Delage F, Colman R and Moise A 2012 Consensus on Twenty-First-Century Rainfall Projections in Climate Models More Widespread than Previously Thought *J. Clim.* **25** 3792–3809

Robertson, F. R., M. G. Bosilovich, J. B. Roberts, R. H. Reichle, R. Adler, L. Ricciardulli, W. Berg, and G. J. Huffman, 2014: Consistency of Estimated Global Water Cycle Variations over the Satellite Era. *J. Clim.*, 27, 6135-6154, https://doi.org/10.1175/JCLI-D-13-00384.1.

Rodwell, M. J. and Hoskins, B.J., 1996: Monsoons and the dynamics of deserts, *Q. J. Roy. Meteorol. Soc.*, 122, 1385–1404, https://doi.org/10.1002/qj.49712253408.

Salmun H, Molod A and Collow A 2018 The use of MERRA-2 near surface meteorology to understand the behavior of observed planetary boundary layer heights over the US Great Plains. *American Geophysical Union Fall Meeting* abstract #A13M-2638

Sathyanadh A, Prabhakaran T, Patil C and Karipot A 2017 Planetary boundary layer height over the Indian subcontinent: variability and controls with respect to monsoon *Atmos. Res.* **195** 44–61

Seidel D J, Zhang Y, Beljaars A, Golaz J-C, Jacobson A R and Medeiros B 2012 Climatology of the planetary boundary layer over the continental United States and Europe *J. Geophys. Res.* **117** D17106

Seidel D J,Ao C O and Li K 2010 Estimating climatological planetary boundary layer heights from radiosonde observations: Comparison of methods and uncertainty analysis *J. Geophys. Res.* 115 D16113

Sen P K 1968 Estimates of the regression coefficient based on Kendall's tau, *J. Am. Stat. Assoc.* **63** 1379–1389

Sherwood S C and Fu Q 2014 A drier future? *Science* **343** 737–739

Simmons A J,Willett K M, Jones P D, Thorne P W, and Dee D P 2010 Low-frequency variations in surface atmospheric humidity, temperature, and precipitation: inferences from reanalyses and monthly gridded observational data sets, J. Geophys. Res.- Atmos., 115, D01110, https://doi.org/10.1029/2009JD012442

Stipanuk G S 1973 Algorithms for generating a skew-t, log-p diagram and computing selected meteorological quantities. Atmospheric Sciences laboratory, U.S. Army Electronics Command, White Sands Missile Range, New Mexico, 88002. pp. 33.

Stull R B 1988 An introduction to boundary layer meteorology. Kluwer, Dordrecht, Holland. pp. 680.

Sun W Wang, B., Zhang, Q., Pausata, F. S. R., Chen, D., Lu, G., et al., 2019 Northern Hemisphere land monsoon precipitation increased by the Green Sahara during middle Holocene. *Geophys. Res Lett*., *46*. https://doi.org/10.1029/2019GL082116.

Svensson G, and Lindvall J 2015 Evaluation of near-surface variables and the vertical structure of the boundary layer in CMIP5 models. *J Clim* 28(13):5233–5253

Talbot C *et al* 2007 Impact of a sea breeze on the boundary-layer dynamics and the atmospheric stratification in a coastal area of the North Sea *Bound-Lay. Meteorol.* **125** 133–154

Theil H 1950 A rank-invariant method of linear and polynomial regression analysis *Proceedings of the Royal Netherlands Academy of Sciences* **53** PART I 386–392 PART II 521–525 PART III 1397–1412

Thomas N and Nigam S 2018 Twentieth-century climate change over Africa: Seasonal hydroclimate trends and Sahara Desert expansion *J. Clim*. **31** 3349-3370

Thorne P W, Lanzante J R, Peterson T C, Seidel D J and Shine K P 2011 Tropospheric temperature trends: History of an ongoing controversy *WIREs Clim. Change* **2** 66–88

Thorne P W D. E. Parker, S. F. Tett, P. D. Jones, M. McCarthy, H. Coleman, and P. Brohan, 2005: Revisiting radiosonde upper air temperatures from 1958 to 2002. *J. Geophys. Res. D Atmos.*, **110,** D18105.

Trenberth K E, Koike T and Onogi K 2008 Progress and Prospects for Reanalysis for Weather and Climate *Eos Trans. AGU* **89** 234-235

Vicente-Serrano S M *et al* 2017 Recent changes of relative humidity: Regional connection with land and ocean processes *Earth Syst. Dynam. Discussions* **9** 1–44

Vizy E K and Cook K H 2017 Seasonality of the observed amplified Sahara warming trend and implications for Sahel rainfall *J. Clim.* **30** 3073-3094

Von Engeln A and Teixeira J 2013 A planetary boundary layer height climatology derived from ECMWF reanalysis data *J. Clim.* **26** 6575–6590

von Salzen K *et al* 2013 The Canadian fourth generation atmospheric global climate model (CanAM4). Part I: Representation of physical processes *Atmos. Ocean* **51** 104-125

Vose R S, Easterling D R and Gleason B 2005 Maximum and minimum temperature trends for the globe: An update through 2004 *Geophys. Res. Lett.* **32** L23822

Wang X Y and Wang K 2016 Homogenized variability of radiosonde derived atmospheric boundary layer height over the global land surface from 1973 to 2014. *J. Climate*, **29**, 6893–6908, https://doi.org/10.1175/JCLI-D-15-0766.1.

Wang B J. Liu, , H.-J. Kim, P. J. Webster, and S.-Y. Yim, 2012 Recent change of the global monsoon precipitation (1979–2008). *Climate Dyn.*, **39**, 1123–1135, doi: 10.1007/s00382-011-1266-z.

Wang B Jin C.; Liu, J., 2020 Understanding Future Change of Global Monsoons Projected by CMIP6 Models. *J. Clim*. 33, 6471–6489. DOI: https://doi.org/10.1175/JCLI-D-19-0993.1

Wei N, Zhou L and Dai Y 2017a Evaluation of simulated climatological diurnal temperature range in cmip5 models from the perspective of planetary boundary layer turbulent mixing *Clim. Dynam.* **49** 1-22

Wei N, Zhou L and Dai Y 2017b Observational evidence for desert amplification using multiple satellite datasets *Sci. Rep-UK* **7** 2043

Willett K M, Dunn R J H, Thorne P W, Bell S, de Podesta M, Parker D E, Jones P D, and Williams Jr C. N. 2006 HadISDH land surface multi-variable humidity and temperature record for climate monitoring, Clim. Past, 10, 1983–2006. https://doi.org/10.5194/cp-10-1983-2014.

Zhang Y H, Seidel D J and Zhang S D 2013 Trends in planetary boundary layer height over Europe *J. Clim.* **26** 10071–10076

Zhao J 2011 A simulative study of the thermal mechanism for development of the convective boundary layer in the arid zone of northwest China *Acta Meteorologica Sinica* **69** 1029–1037

Zhao Y, Mao W, Zhang K, Ma Y, Li H and Zhang W 2017 Climatic variations in the boundary layer height of arid and semiarid areas in East Asia and North Africa *J. Meteorol. Soc. Jpn.* **95** 181–197

Zhou L 2016 Desert amplification in a warming climate *Sci. Rep-UK* **6** 31065

Zhou L, Chen H and Dai Y 2015 Stronger warming amplification over drier ecoregions observed since 1979 *Environ. Res. Lett.* **10** 064012

Zhou L, Chen H, Hua W, Dai Y and Wei N 2016 Mechanisms for stronger warming over drier ecoregions observed since 1979 *Clim. Dynam.* **47** 2955-2974

Zhou L, Dai A, Dai Y, Vose R S, Zou C-Z, Tian Y and Chen H 2009 Spatial dependence of diurnal temperature range trends on precipitation from 1950 to 2004 *Clim. Dynam.* **32** 429–440

Zhou L, Dickinson R E, Dai A and Dirmeyer P 2010 Detection and attribution of anthropogenic forcing to diurnal temperature range changes from 1950 to 1999: Comparing multi-model simulations with observations *Clim. Dynam.* **35** 1289-1307

Zhou L, Dickinson R E, Tian Y and Vose R S 2007 Impact of vegetation removal and soil aridation on diurnal temperature range in a semiarid region – Application to the Sahel *Proc. Natl. Acad. Sci.* **104** 17937-17942

Zhou, C., J. Wang, A. Dai, and P. W. Thorne, 2020: A new approach to homogenize global sub-daily radiosonde temperature data from 1958 to 2018. J. Clim., 1-64, doi: 10.1175/JCLI-D-20-0352.1.

Zhou L., 2020: Diurnal asymmetry of desert amplification and its possible connections to planetary boundary layer height: A case study for the Arabian Peninsula, *Climate Dyn.*, minor revision.

Table 1. List of 74 CMIP6 AMIP and HIST simulations from 27 models used in this study

Organization	Model	Number of realizations used[a]		
		CMIP6-AMIP	CMIP6-HIST	CMIP6-HIST-RI
Canadian Climate Centre for Modelling and Analysis, Canada	CanESM5	3	3	
Canadian Climate Centre for Modelling and Analysis, Canada	CanESM5-Can0E		3	
National Center for Atmospheric Research, USA	CESM2	3	3	
National Center for Atmospheric Research, USA	CESM2-FV2	3	3	
National Center for Atmospheric Research, USA	CESM2-WACCM	3	3	
National Center for Atmospheric Research, USA	CESM2-WACCM-FV2	3	3	
Centro Euro-Mediterraneo per I Cambiamenti Climatici, Europe	CMCC-CM2-SR5			1
Centre National de Recherches Météorologiques, France	CNRM-CM6-1			1
Centre National de Recherches Météorologiques, France	CNRM-ESM2-1			1
NOAA/Geophysical Fluid Dynamics Laboratory, USA	GFDL-CM4		1	1
NOAA/Geophysical Fluid Dynamics Laboratory, USA	GFDL-ESM4		1	
NASA/Goddard Institute for Space Studies, USA	GISS-E2-1-G	3	3	
NASA/Goddard Institute for Space Studies, USA	GISS-E2-1-H		3	
NASA/Goddard Institute for Space Studies, USA	GISS-E2-2-G	3		
Institute for Numerical Mathematics, Russian Academy of Science, Russia	INM-CM4-8	1	1	
	INM-CM5-0	1	3	
JAMSTEC/AORI/University of Tokyo/NIES, Japan	MIROC6			1
JAMSTEC/AORI/University of Tokyo/NIES, Japan	MIROC-ES2L			1
Max Planck Institute for Meteorology, Germany	MPI-ESM-1-2-HAM			1
Max Planck Institute for Meteorology, Germany	MPI-ESM1-2-HR			1
Max Planck Institute for Meteorology, Germany	MPI-ESM1-2-LR			1
Meteorological Research Institute, Japan	MRI-ESM2-0			1
Norwegian Climate Center, Norway	NorESM2-LM	1	3	1
Norwegian Climate Center, Norway	NorESM2-MM		1	1
Seoul National University, Seoul, Korea	SAM0-UNICON			1
Research Center for Environmental Changes, Academia Sinica, Taiwan	TaiESM1	1	1	
Met Office Hadley Centre, UK	UKESM1-0-LL	1		
Total models/simulations		12/26	15/35	13/13

Note: [a]To assess internal variability, some models provide an ensemble of realizations with different initial conditions. For the models with more than 3 realizations, only the first three realizations were obtained for each model. For the CMIP6-AMIP and CMIP6-HIST, the available models with PBLH output at the time of analysis were chosen. For CMIP6-HIST-RI, only the first realization for a subset of CMIP6 models for which data were available on the model-level grid and the output frequency needed was chosen to estimate the PBLH as detailed in Davy (2018).

Table 2. List of variable and method used to estimate the PBLH for the nine datasets used in this study

Dataset acronym	Variable used	PBLH estimate method	Time period
IGRA2-MMLH	Daily atmospheric profiles (pressure, temperature and dewpoint) at 00 and 12 UTC from the IGRA2 and 3-hourly surface measurements (pressure, temperature and dewpoint) from the NOAA's NCDC	The parcel method (Holzworth, 1964) to estimate MMLH as detailed in Li et al. (2020)	1979-2018
IGRA2-RI	Daily atmospheric and near surface virtual potential temperature, geopotential height, and wind speed at 00 and 12 UTC from the IGRA2 sounding-derived parameters from the NCDC	The bulk Richardson number method with a critical value of 0.25 following Seidel et al. (2012) and Zhang et al. (2013)	1979-2019
ERA5 ERA5-ensemble	Monthly mean of hourly mean PBLH provided by the reanalysis	The bulk Richardson number method with a critical value of 0.25 (C3S, 2017)	1979-2019
MERRA-2	Monthly mean of hourly mean PBLH provided by the reanalysis	The total eddy diffusion coefficient of heat with a threshold value of 2 m^2s^{-1} (Salmun et al., 2018)	1980-2019
MERRA-2-RI	6-hourly 3-d atmospheric instantaneous and hourly averaged 2-d near-surface fields for air temperature, humidity, geopotential height, and wind speed from the MERRA-2 output (GMAO; 2015a; 2015b; 2015c; 2015d)	The bulk Richardson number method with a critical value of 0.25 following Seidel et al. (2012) and Zhang et al. (2013)	1980-2019
CMIP6-AMIP CMIP6-HIST	Monthly mean of daily mean PBLH provided by available models for the CMIP6 AMIP and HIST simulations (Table 1)	Different PBL schemes and vertical resolution in different models (e.g., Svensson and Lindvall, 2015)	1979-2014
CMIP6-HIST-RI	6-hourly 3-d atmospheric and 3-hourly 2-d near-surface instantaneous fields for air temperature, humidity, geopotential height, and wind speed from the CMIP6 HIST simulations (Table 1)	The bulk Richardson number method with a critical value of 0.25 as detailed in Davy (2018)	1979-2014

Table 3: Climatology (m) and trend (m/decade) of regional mean $PBLH_{mean}$ for the reanalysis and CMIP6 datasets used in this study

Dataset	Type	60°S-60°N	SDAP	India
ERA5	Climatology	589.48	723.05	594.44
MERRA2	Climatology	966.91	1453.36	1265.16
ERA5-EM	Climatology	599.48	737.25	599.98
MERRA2-RI	Climatology	536.83	691.25	757.43
CMIP6-AMIP-MEM	Climatology	704.32	843.21	757.08
CMIP6-HIST-MEM	Climatology	668.44	776.98	761.7
CMIP6-HIST-RI-MEM	Climatology	717.19	754.27	751.73
ERA5	OLS	**7.59**	**14.51**	**-8.52**
MERRA2	OLS	1.45	**27.74**	**-25.7**
ERA5-EM	OLS	**7.66**	**16.43**	**-7.02**
MERRA2-RI	OLS	0.25	**8.71**	**-19.17**
CMIP6-AMIP-MEM	OLS	0.51	**3.26**	**-7.02**
CMIP6-HIST-MEM	OLS	0.22	**1.82**	**-9.83**
CMIP6-HIST-RI-MEM	OLS	**0.48**	**3.03**	**-12.33**
ERA5	M-K	**7.62**	**14.69**	**-9.02**
MERRA2	M-K	0.28	**26.92**	**-27.78**
ERA5-EM	M-K	**7.73**	**16.77**	**-7.65**
MERRA2-RI	M-K	0.09	**8.83**	**-22.62**
CMIP6-AMIP-MEM	M-K	0.46	**3.29**	**-6.47**
CMIP6-HIST-MEM	M-K	0.07	**1.79**	**-9.83**
CMIP6-HIST-RI-MEM	M-K	**0.42**	**2.86**	**-12.87**

Note: Column 2 (type): climatology - the climatology of regional mean $PBLH_{mean}$, OLS and M-K - two methods used to calculate the trend and its significance. Regional averaging is applied to the land grid boxes between 60°S and 60°N, SDAP (18°N-31°N, 5°W-50°E) and India (17°N-34°N, 68°E-96°E), depicted as the two rectangle boxes in Fig. 3c.

Table 4: Spatial correlations in annual mean trends between PBLH and related variables for the two reanalysis (ERA5 and MERRA-2) datasets

ERA5									
Method	PBLH	N	SHFX	LHFX	PLCL	Ts	T_{2M}	Td_{2M}	RH_{2M}
OLS	$PBLH_{mean}$	7137	**0.76**	**-0.52**	**-0.73**	**0.61**	**0.51**	**-0.50**	**-0.73**
OLS	$PBLH_{max}$	7693	**0.77**	**-0.64**	**-0.74**	**0.70**	**0.51**	**-0.57**	**-0.73**
OLS	$PBLH_{min}$	4580	**0.13**	**0.43**	**-0.16**	**0.04**	*-0.01*	**-0.16**	**-0.15**
M-K	$PBLH_{mean}$	7018	**0.76**	**-0.51**	**-0.72**	**0.60**	**0.51**	**-0.49**	**-0.72**
M-K	$PBLH_{max}$	7590	**0.77**	**-0.63**	**-0.74**	**0.69**	**0.50**	**-0.57**	**-0.72**
M-K	$PBLH_{min}$	4415	**0.12**	**0.43**	**-0.16**	**0.04**	*0.00*	**-0.15**	**-0.15**
MERRA-2									
Method	PBLH	N	SHFX	LHFX	ZLCL	Ts	T_{2M}	q_{2M}	RH_{2M}
OLS	$PBLH_{mean}$	5609	**0.65**	**-0.59**	**0.79**	**0.80**	**0.79**	**-0.68**	**-0.76**
OLS	$PBLH_{max}$	4857	**0.86**	**-0.76**	**0.92**	**0.91**	**0.91**	**-0.90**	**-0.92**
OLS	$PBLH_{min}$	6374	**0.18**	**-0.08**	**0.44**	**0.55**	**0.55**	**-0.17**	**-0.39**
M-K	$PBLH_{mean}$	5443	**0.63**	**-0.57**	**0.76**	**0.77**	**0.76**	**-0.66**	**-0.74**
M-K	$PBLH_{max}$	4696	**0.85**	**-0.74**	**0.90**	**0.91**	**0.90**	**-0.89**	**-0.91**
M-K	$PBLH_{min}$	5991	**0.16**	**-0.05**	**0.43**	**0.52**	**0.52**	**-0.18**	**-0.38**

Note: Column 1 (Method): Two methods (OLS and M-K) are used to calculate the trend (per decade) and its significance; Column 2 (PBLH): $PBLH_{mean}$ – daily mean of 24 hourly PBLH values, $PBLH_{max}$ – daytime mean of five hours with maximum climatological PBL values, $PBLH_{min}$ – nighttime mean of five hours with minimum climatological PBL values; Column 3 (N): the number of grid boxes (N) with a statically significant trend in PBLH ($p<0.05$) over land between 60°S and 60°N; Columns 4-10 (the corresponding PBLH related surface variables): SHFX – sensible heat flux (W/m^2), LHFX – latent heat flux (W/m^2), PLCL – the pressure at lifting condensation level (*hPa*), ZLCL – the height at lifting condensation level (m), T_s – surface skin temperature (K), T_{2m} – 2m surface air temperature (K), Td_{2m} – 2m surface dew point temperature (K), q_{2m} – 2m surface specific humidity (g/kg), RH_{2m} – 2m surface relative humidity (%). All correlation coefficients in bold except 2 *italicized* are statistically significant at the 1% level based on a two-tailed student's *t* test due to the large size of data samples (N).

Table 5: Temporal correlations in regional mean annual anomaly time series between PBLH and related variables for the two reanalysis (ERA5 and MERRA-2) datasets during the period 1979-2019

ERA5							
Region	PBLH	N	SHFX	LHFX	Ts	T_{2M}	RH_{2M}
60°S-60°N	$PBLH_{mean}$	41	**0.96**	-0.21	**0.88**	**0.88**	**-0.96**
60°S-60°N	$PBLH_{max}$	41	**0.98**	**-0.75**	**0.88**	**0.87**	**-0.98**
60°S-60°N	$PBLH_{min}$	41	**0.58**	**0.79**	**0.80**	**0.82**	**-0.67**
SDAP	$PBLH_{mean}$	41	**0.75**	0.04	**0.75**	**0.75**	**-0.32**
SDAP	$PBLH_{max}$	41	**0.55**	-0.15	**0.67**	**0.68**	-0.27
SDAP	$PBLH_{min}$	41	-0.02	**0.61**	**0.57**	**0.54**	-0.19
India	$PBLH_{mean}$	41	**0.82**	**-0.70**	0.28	-0.03	**-0.92**
India	$PBLH_{max}$	41	**0.86**	**-0.61**	**0.65**	0.16	**-0.93**
India	$PBLH_{min}$	41	**-0.62**	0.00	-0.13	-0.25	**-0.44**
MERRA-2							
Region	PBLH	N	SHFX	LHFX	Ts	T_{2M}	RH_{2M}
60°S-60°N	$PBLH_{mean}$	40	**0.81**	**-0.67**	**0.38**	**0.35**	**-0.88**
60°S-60°N	$PBLH_{max}$	40	**0.91**	**-0.69**	**0.51**	**0.42**	**-0.95**
60°S-60°N	$PBLH_{min}$	40	0.14	0.10	**0.47**	**0.47**	**-0.59**
SDAP	$PBLH_{mean}$	40	**0.43**	**-0.54**	**0.64**	**0.62**	**-0.38**
SDAP	$PBLH_{max}$	40	**0.48**	**-0.56**	**0.63**	**0.59**	**-0.34**
SDAP	$PBLH_{min}$	40	0.02	**-0.42**	**0.72**	**0.71**	**-0.41**
India	$PBLH_{mean}$	40	**0.88**	**-0.80**	0.24	0.08	**-0.86**
India	$PBLH_{max}$	40	**0.95**	**-0.88**	**0.75**	**0.57**	**-0.95**
India	$PBLH_{min}$	40	0.00	-0.04	-0.20	-0.19	-0.03

Note: All variables are defined in Tables 3 and 4. Correlation coefficients in bold are statistically significant at the 5% level based on the two-tailed student's *t* test.

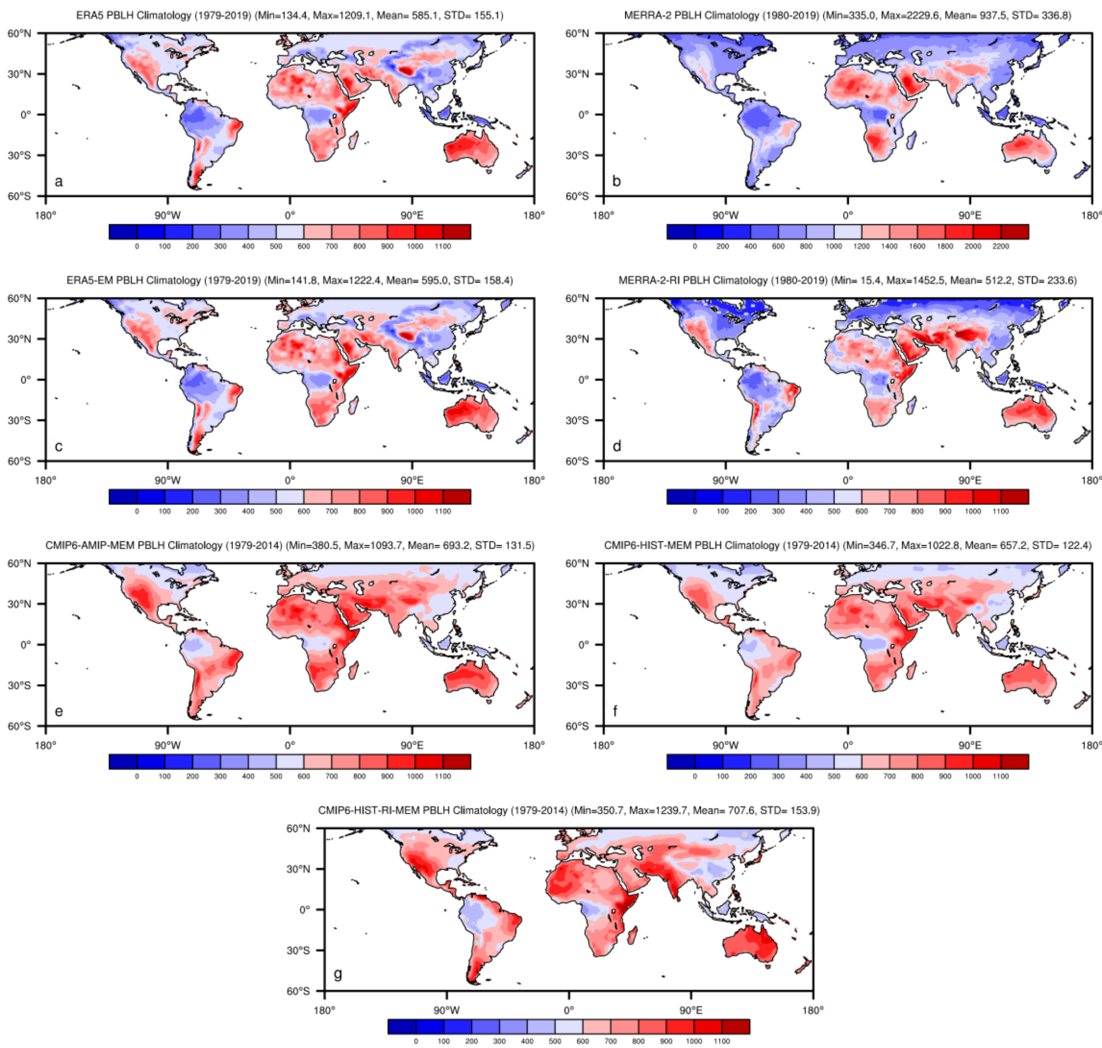

Fig 1. Spatial patterns of climatological $PBLH_{mean}$ (m) from the seven reanalysis and CMIP6 datasets at 1° x 1° resolution: (a) ERA5 (1979-2019), (b) MERRA-2 (1980-2019), (c) ERA5-EM (1980-2019), (d) MERRA-2-RI (1980-2019), (e) CMIP6-AMIP-MEM (1979-2014), (f) CMIP6-HIST-MEM (1979-2014), and (g) CMIP6-HIST-RI-MEM (1979-2014). The maximum, minimum, mean and standard deviation (STD) of PBLH over all land grid boxes between 60°S-60°N, are listed on the top of each panel.

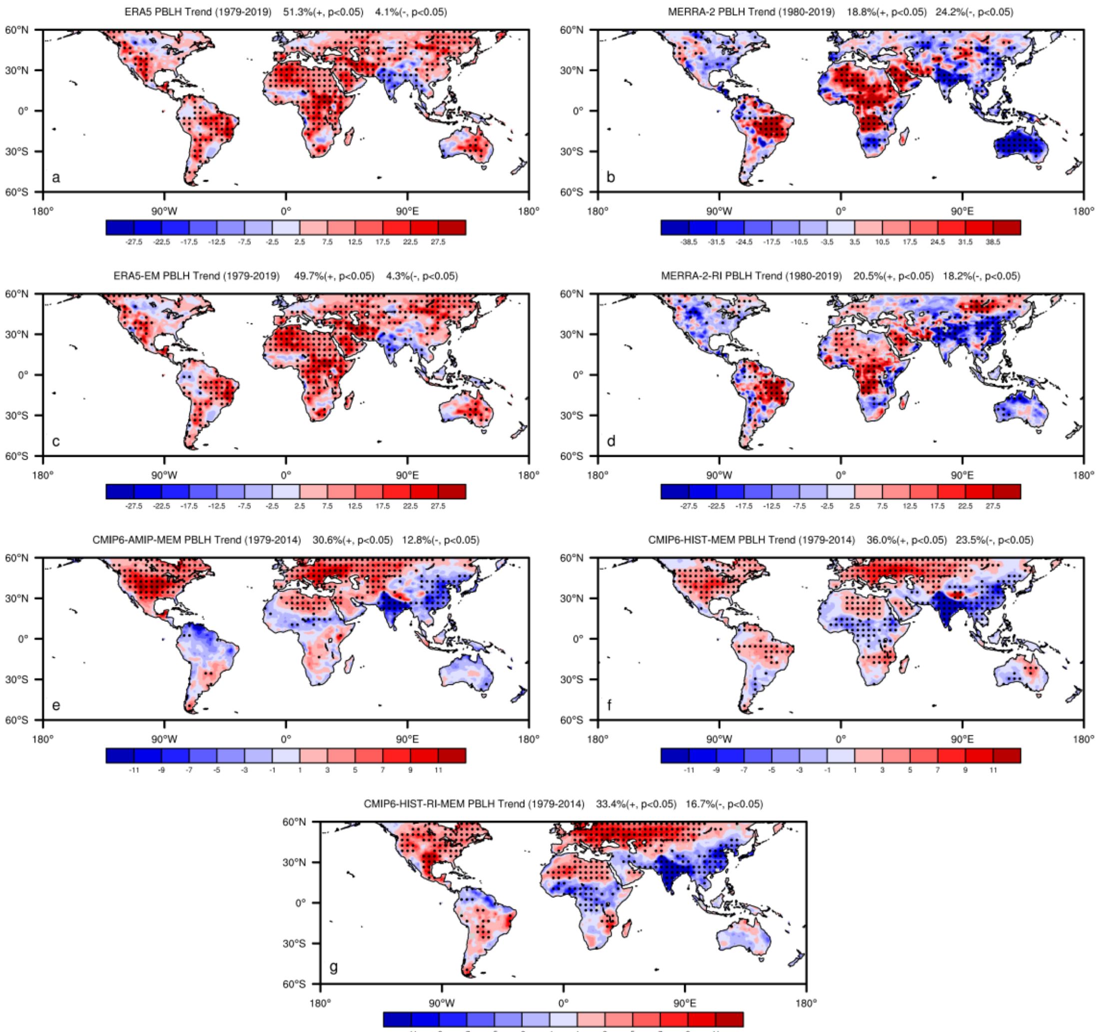

Fig 2. Spatial patterns of $PBLH_{mean}$ trend (m/decade) estimated using the M-K method from the seven datasets at 1° x 1° resolution: (a) ERA5 (1979-2019), (b) MERRA-2 (1980-2019), (c) ERA5-EM (1980-2019), (d) MERRA-2-RI (1980-2019), (e) CMIP6-AMIP-MEM (1979-2014), (f) CMIP6-HIST-MEM (1979-2014), and (g) CMIP6-HIST-RI-MEM (1979-2014). The percentage of positive (+) and negative (-) trends that are statistically significant at the 5% level over all land grid boxes between 60°S-60°N, are listed on the top of each panel.

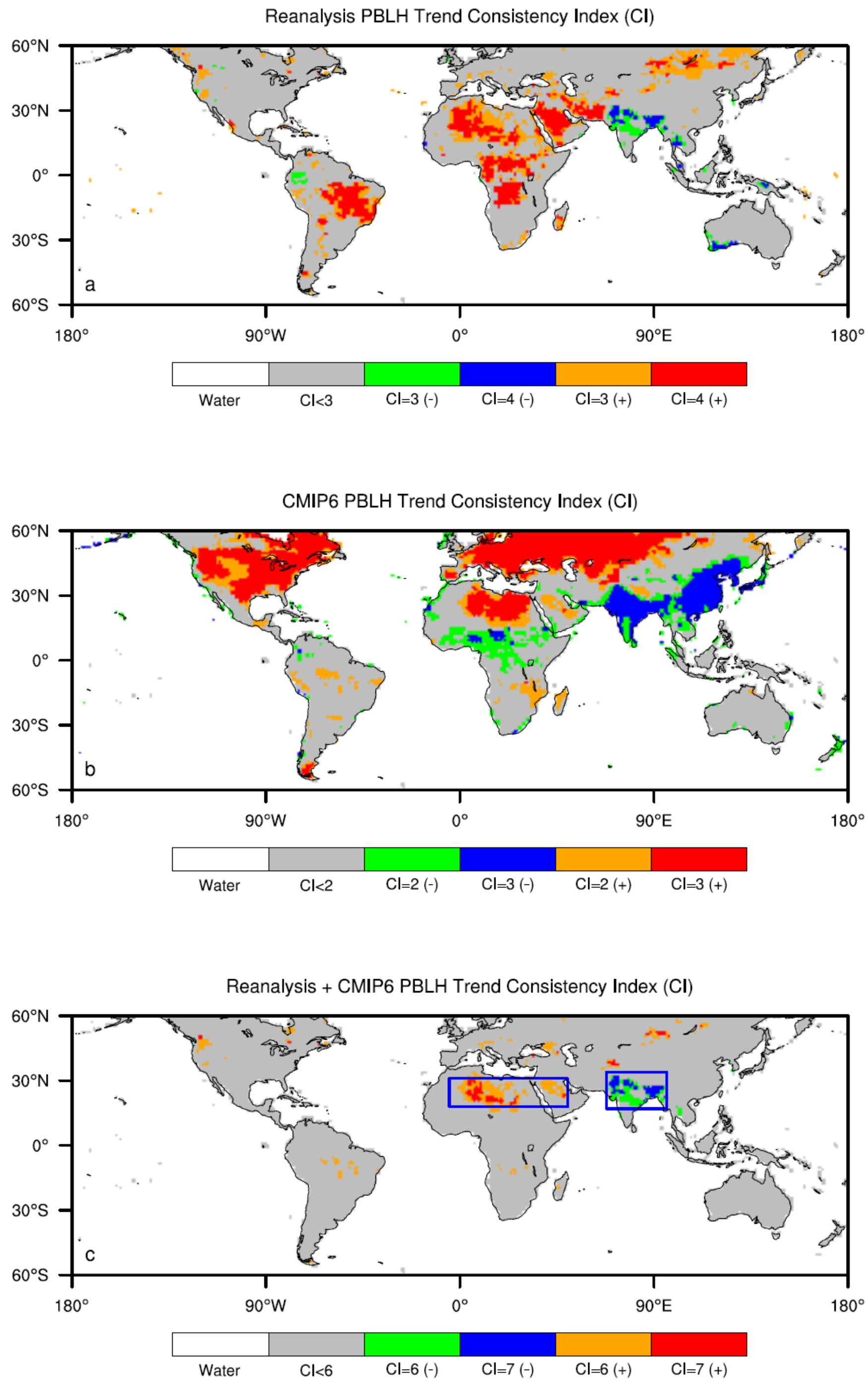

Fig 3. Spatial patterns of consistency index (CI) for $PBLH_{mean}$ trend (m/decade) estimated using the M-K method from (a) the four reanalysis datasets (Figs. 2a-2d), (b) the three CMIP6 datasets (Figs. 2e-2g), and (c) all the seven datasets (Figs. 2a-2g). The value of consistency index is defined as the number of the datasets with the same sign of trends that are statistically significant at the 5% level. The rectangle boxes in blue (18°N-31°N, 5°W-50°E) and (17°N-34°N, 68°E-96°E) in Fig. 3c depict the area over which the data is averaged for the regional analysis over the SDAP and India, respectively.

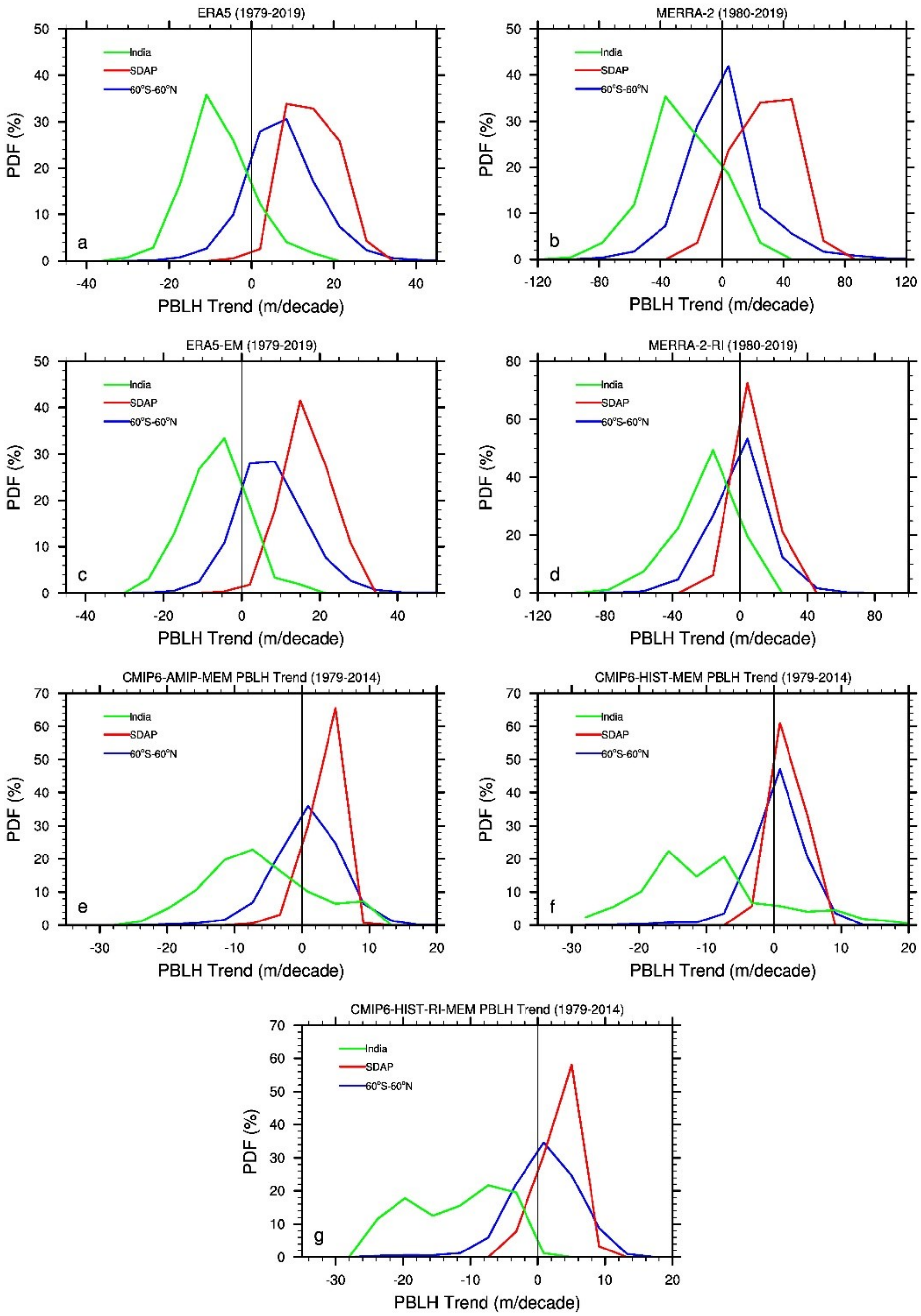

Fig 4. The probability distribution function (PDF) or frequency of occurrence of $PBLH_{mean}$ (m/decade) that are statistically significant at the 5% level over land averaged over three regions: the zonal mean between 60°S-60°N (in blue), the SDAP (in red), and India (in green) for the seven datasets (Fig. 2). The geographic domain for the SDAP and India are depicted in Fig. 3c.

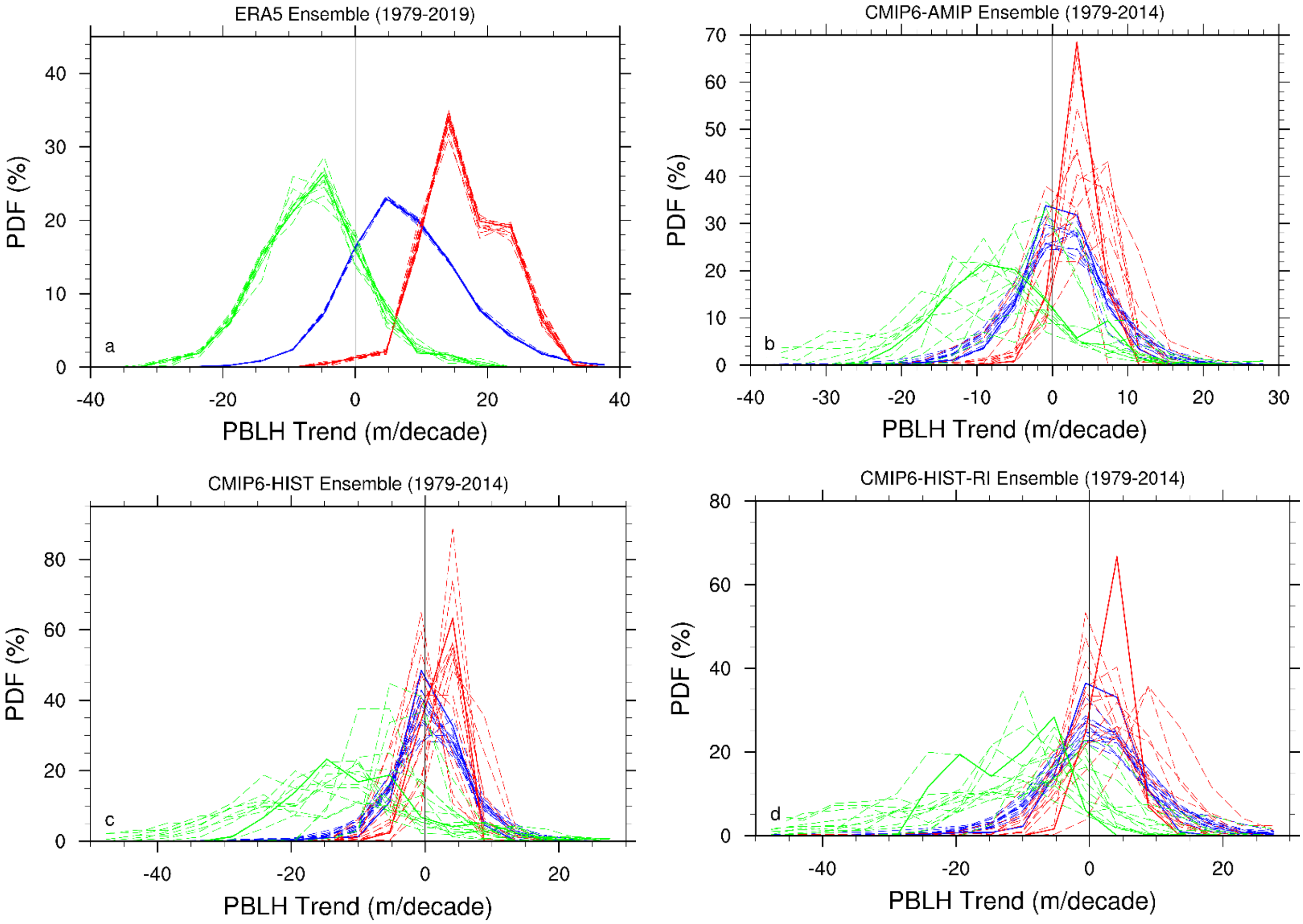

Fig 5. Same as in Fig. 4 but for the four ensemble datasets: (a) ERA5 10-member ensemble (1980-2019), (b) CMIP6-AMIP multi-model ensemble (1979-2014), (c) CMIP6-HIST multi-model ensemble (1979-2014), and (d) CMIP6-HIST-RI multi-model ensemble (1979-2014). The dashed and solid curves denote the individual ensemble members and the ensemble mean, respectively.

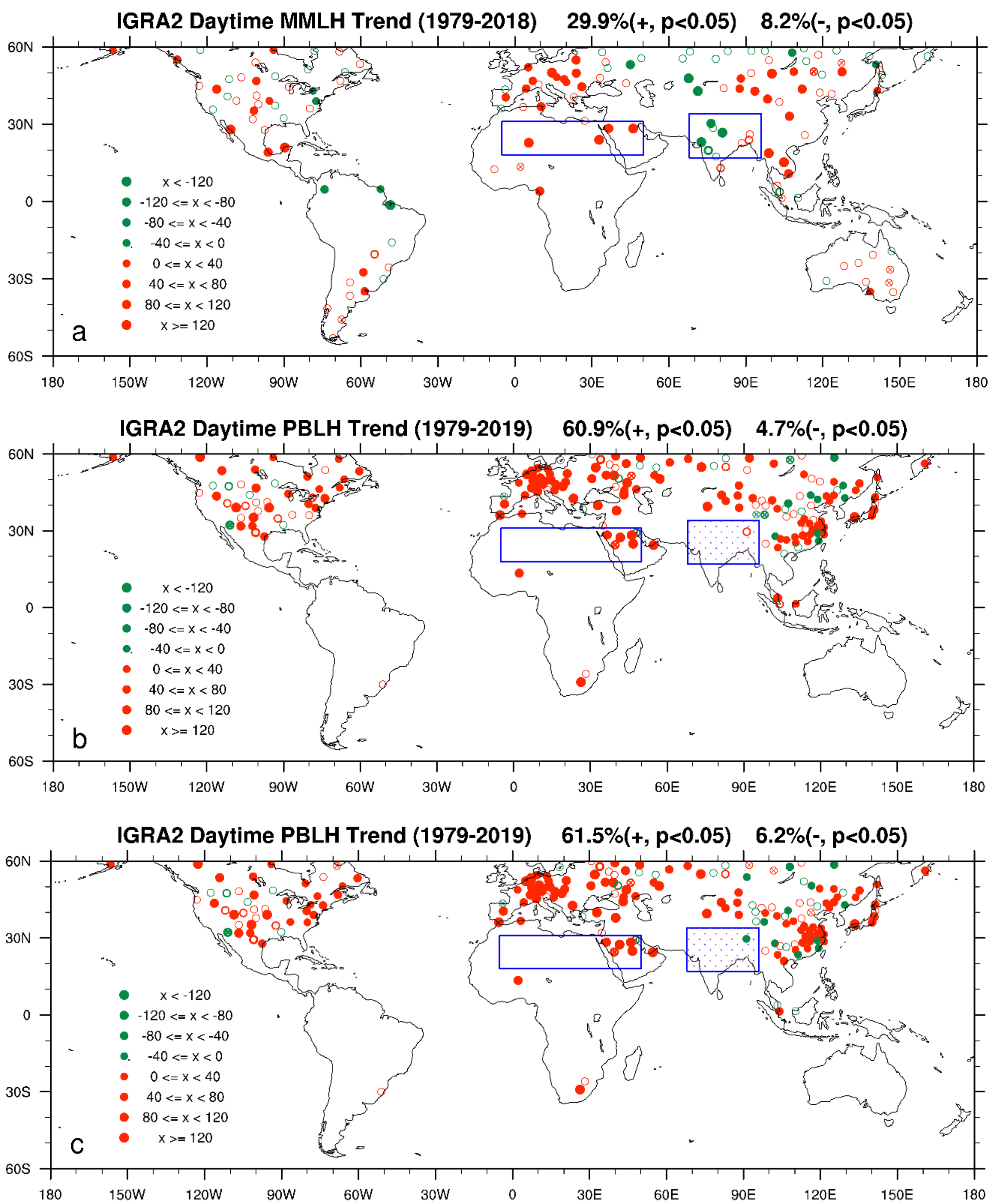

Fig 6. Spatial patterns of PBLH trend (m/decade) estimated using the M-K method from the IGRA2 radiosonde data: (a) daytime MMLH (the compositing method A), (b) daytime PBLH (the compositing method B), and (c) daytime PBLH (the compositing method C). The percentage of positive (+) and negative (-) trends that are statistically significant at the 5% level over land between 60°S-60°N is listed on the top of each panel. The rectangle boxes in blue are depicted in Fig. 3c.

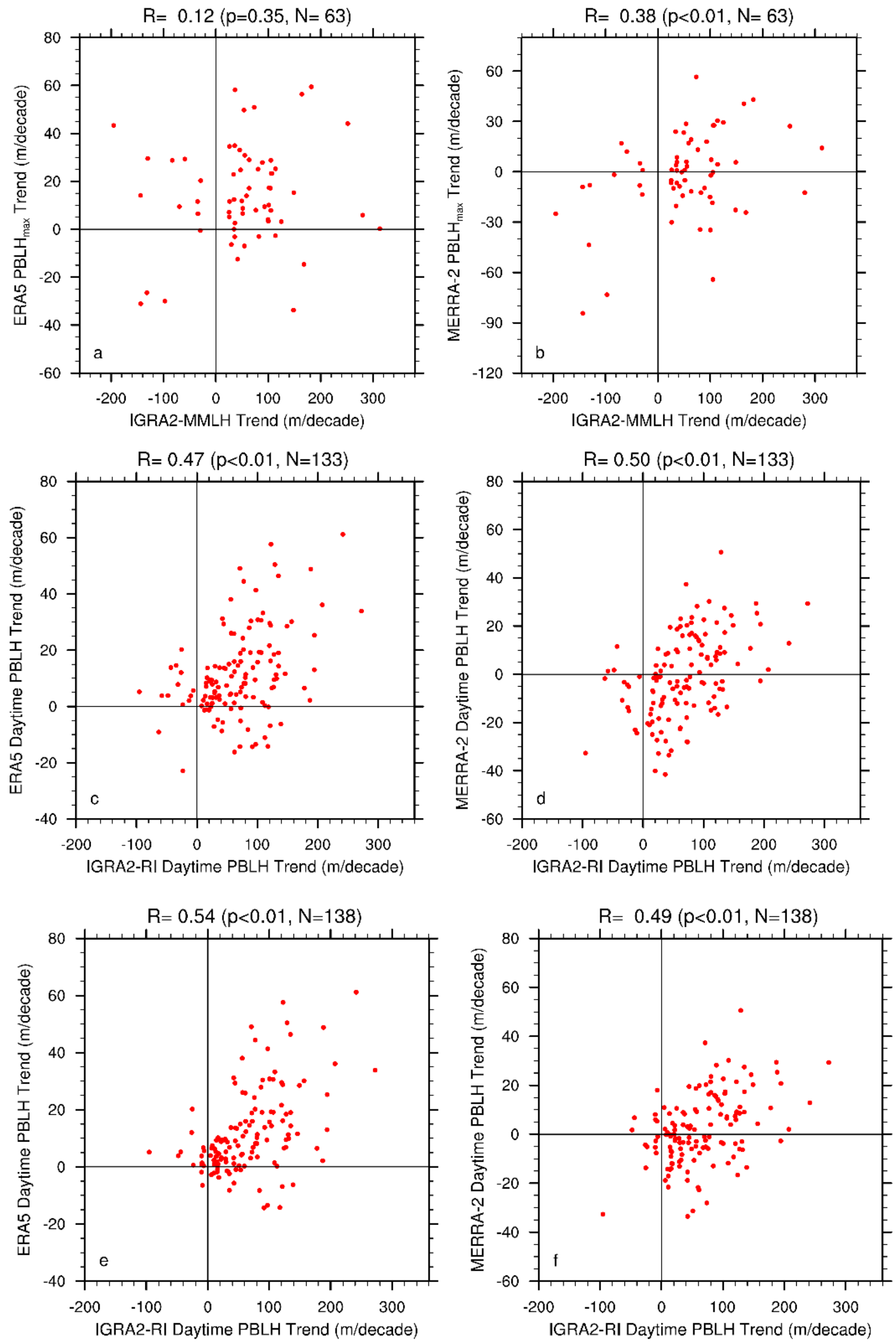

Fig 7. Comparison of PBLH trend (m/decade) estimated using the M-K method from the IGRA2 radiosondes and corresponding ERA5 and MERRA-2 datasets: (a,b) daytime MMLH from 63 stations (Fig. 6a, the compositing method A), (c,d) daytime PBLH from 133 stations (Fig. 6b, the compositing method B), and (e,f) daytime PBLH from 138 stations (Fig. 6c, the compositing method C). The correlation coefficient (R) and its significance (p value) and sample size (N) are listed on the top of each panel. Only the stations with a trend that is statistically significant at the 10% level is used.

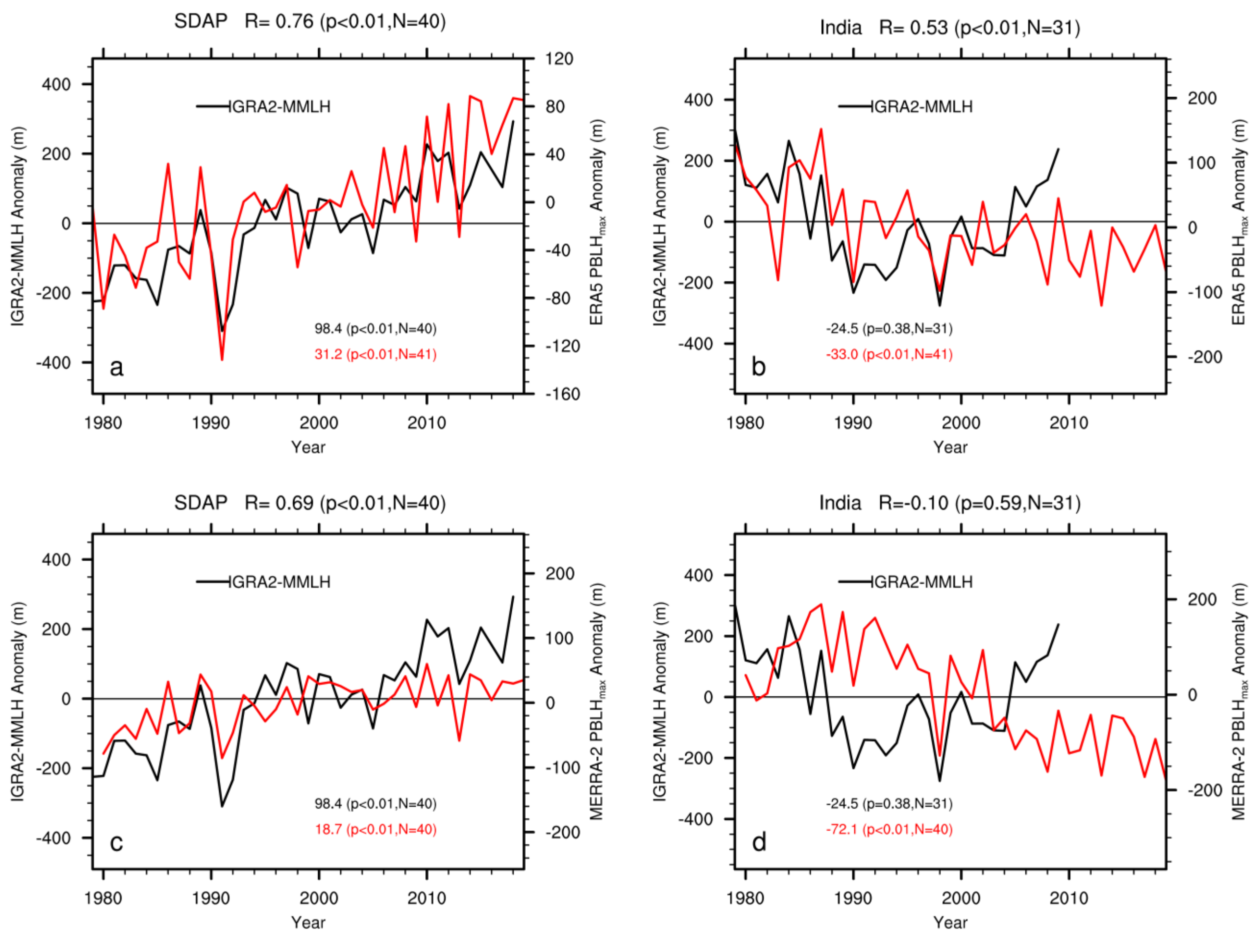

Fig 8. Station-mean interannual variations in daytime MMLH anomalies from the IGRA2-MMLH (left y-axis, m) and corresponding reanalysis daytime $PBLH_{max}$ anomalies (right y-axis, m) from (a, b) ERA5 and (c, d) MERRA-2 averaged over 5 radiosonde stations in (a, c) the SDAP and 9 statins in (b, d) India for the period 1979/1980-2019. These 14 stations are depicted in Fig. 6a. The correlation coefficient (R) and its significance (p value) and sample size (N) are listed on the top of each panel. The trend using the M-K method for the time series and its significance (p value) and N are shown within each panel as well. The compositing method A is used to produce the daytime PBLHs.

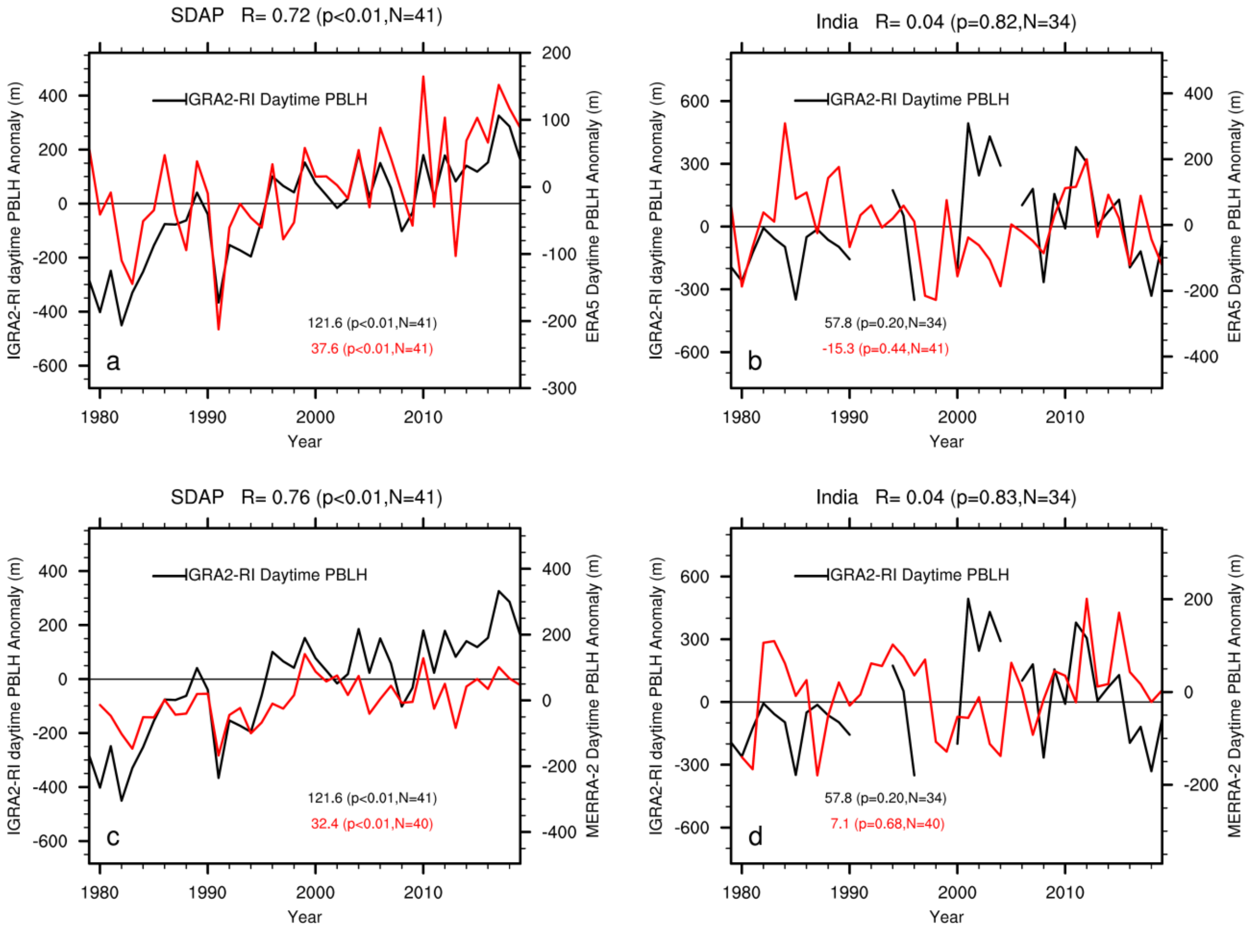

Fig 9. Station-mean interannual variations in daytime PBLH anomalies from the IGRA2-RI (left y-axis, m) and corresponding reanalysis daytime PBLH anomalies (right y-axis, m) from (a, b) ERA5 and (c, d) MERRA-2 averaged over 5 radiosonde stations in (a, c) the SDAP and 1 statin in (b, d) India for the period 1979/1980-2019. These 6 stations are depicted in Fig. 6b. The correlation coefficient (R) and its significance (p value) and sample size (N) are listed on the top of each panel. The trend using the M-K method for the time series and its significance (p value) and N are shown within each panel as well. The compositing method B is used to produce the daytime PBLHs.

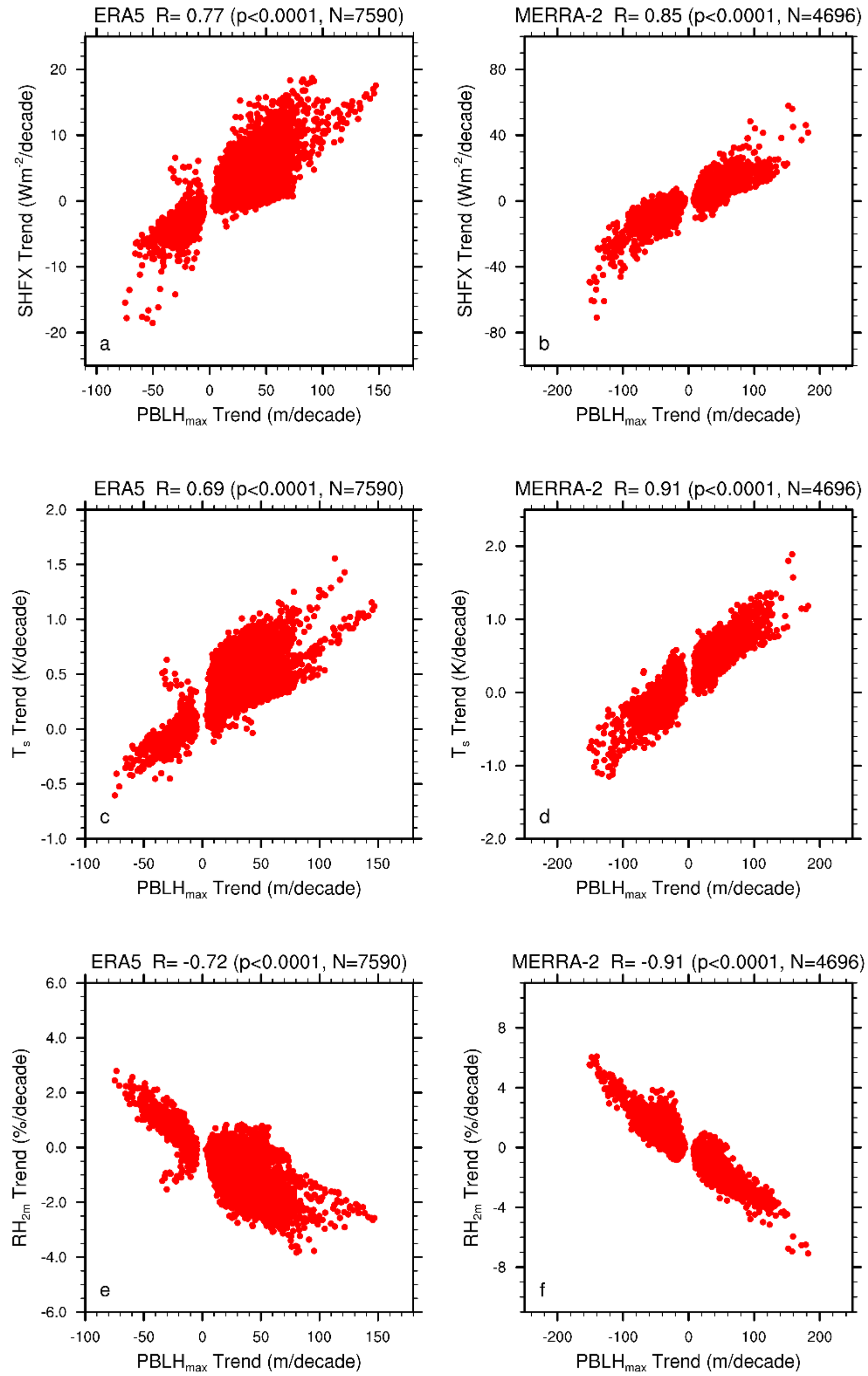

Fig 10. Scatter plots between $PBLH_{max}$ trend (x-axis) and corresponding trends in sensible heat (SHFX, W/m^2), surface skin temperature (Ts, K), and 2m relative humidity (RH_{2m}, %) for the ERA5 (left panels) and MERRA-2 (right panels). Only the grid boxes with a statistically significant trend in PBLH ($p<0.05$) over land between 60°S-60°N are included. The correlation coefficient between the two trends and the sample size (N) in grid boxes are listed on the top of each panel. It is statistically significant ($p<0.0001$) in all plots.

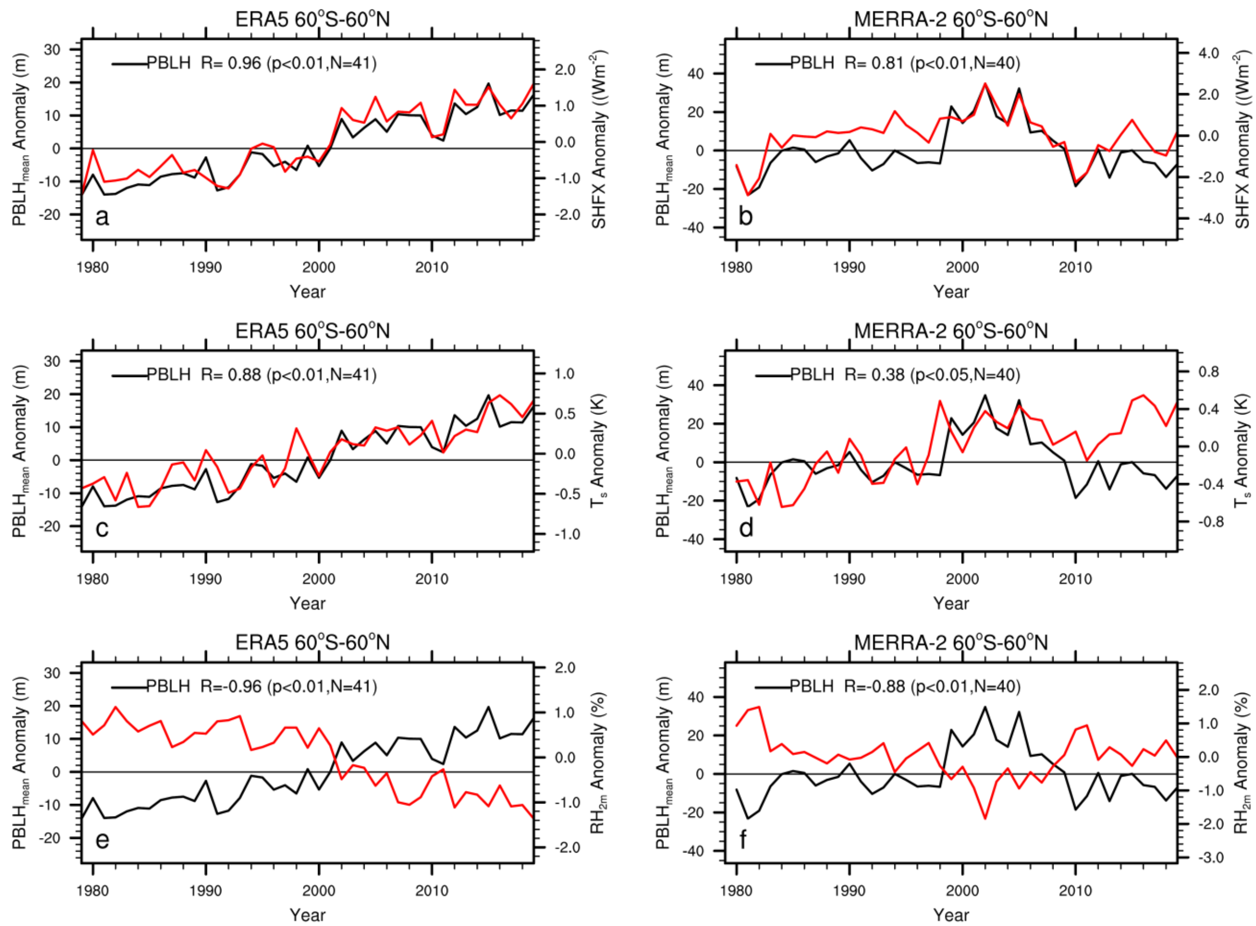

Fig 11. Regional mean interannual variations in $PBLH_{mean}$ anomaly (in green) and three PBLH related variables (in red): sensible heat (SHFX, W/m^2), surface skin temperature (Ts, K), and 2m relative humidity (RH_{2m}, %), for the ERA5 (left panels) and MERRA-2 (right panels) averaged between 60°S-60°N for the period 1979-2019. The correlation coefficient (R) and its significance (p value) and sample size (N) are shown within each panel.

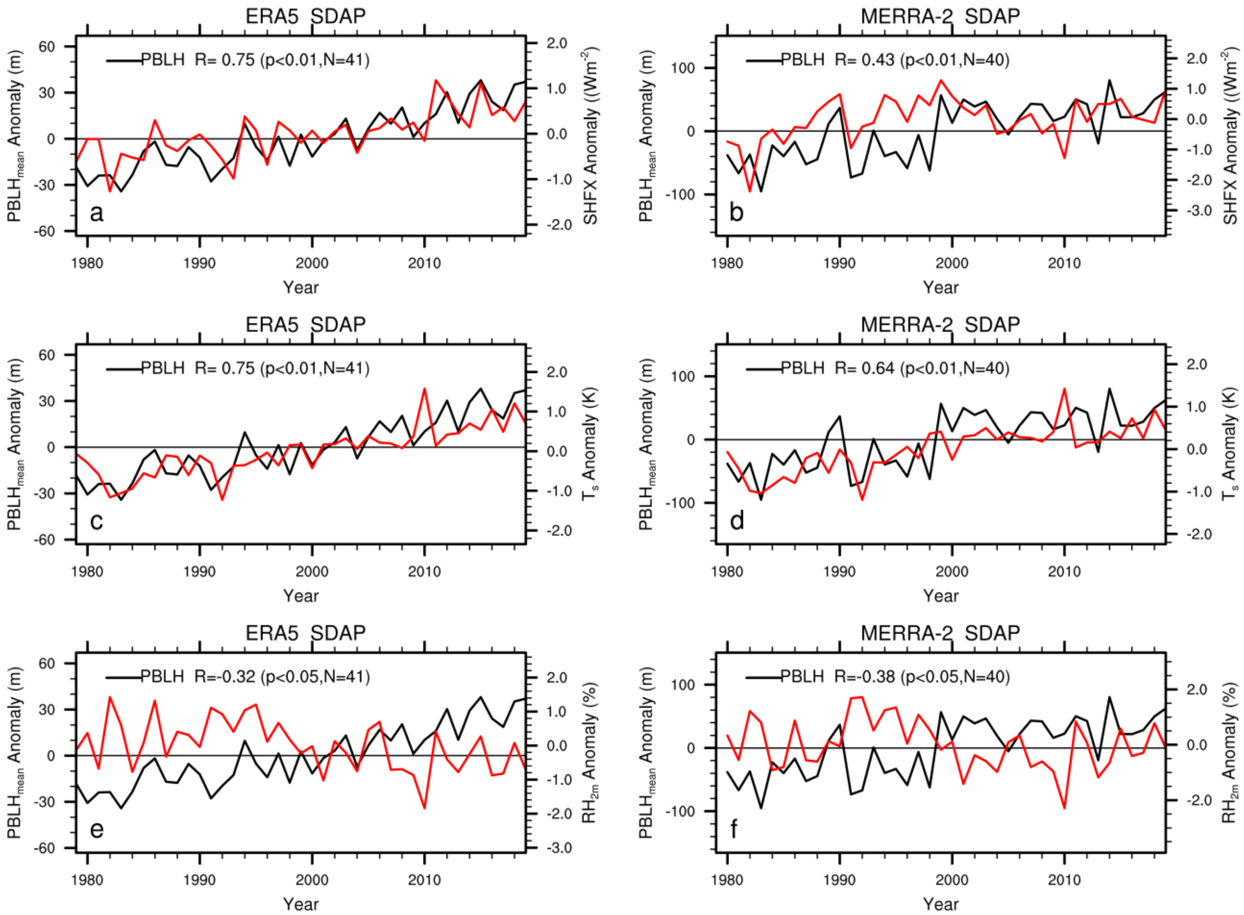

Fig 12. Same as in Fig 11 but is averaged over the SDAP.

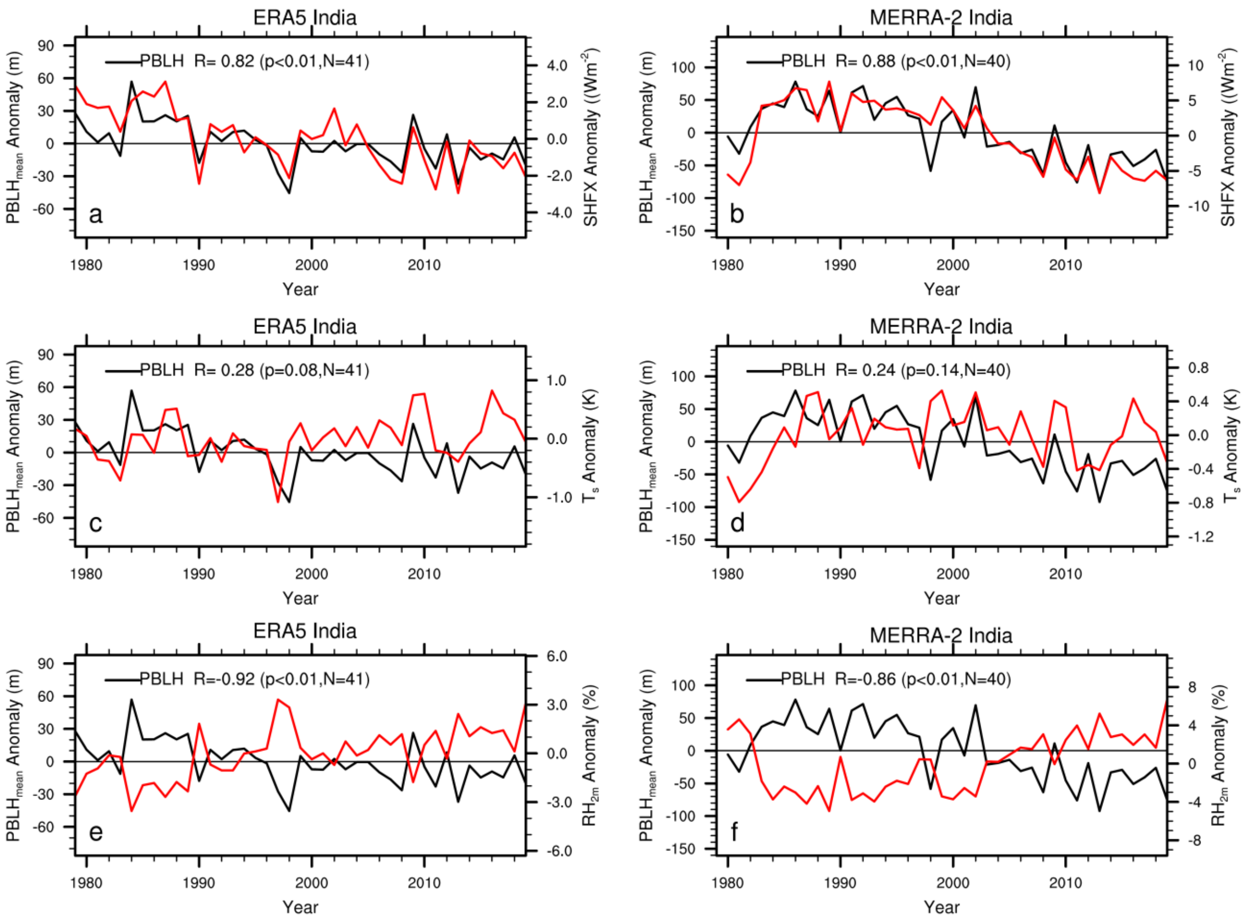

Fig 13. Same as in Fig 11 but is averaged over India.

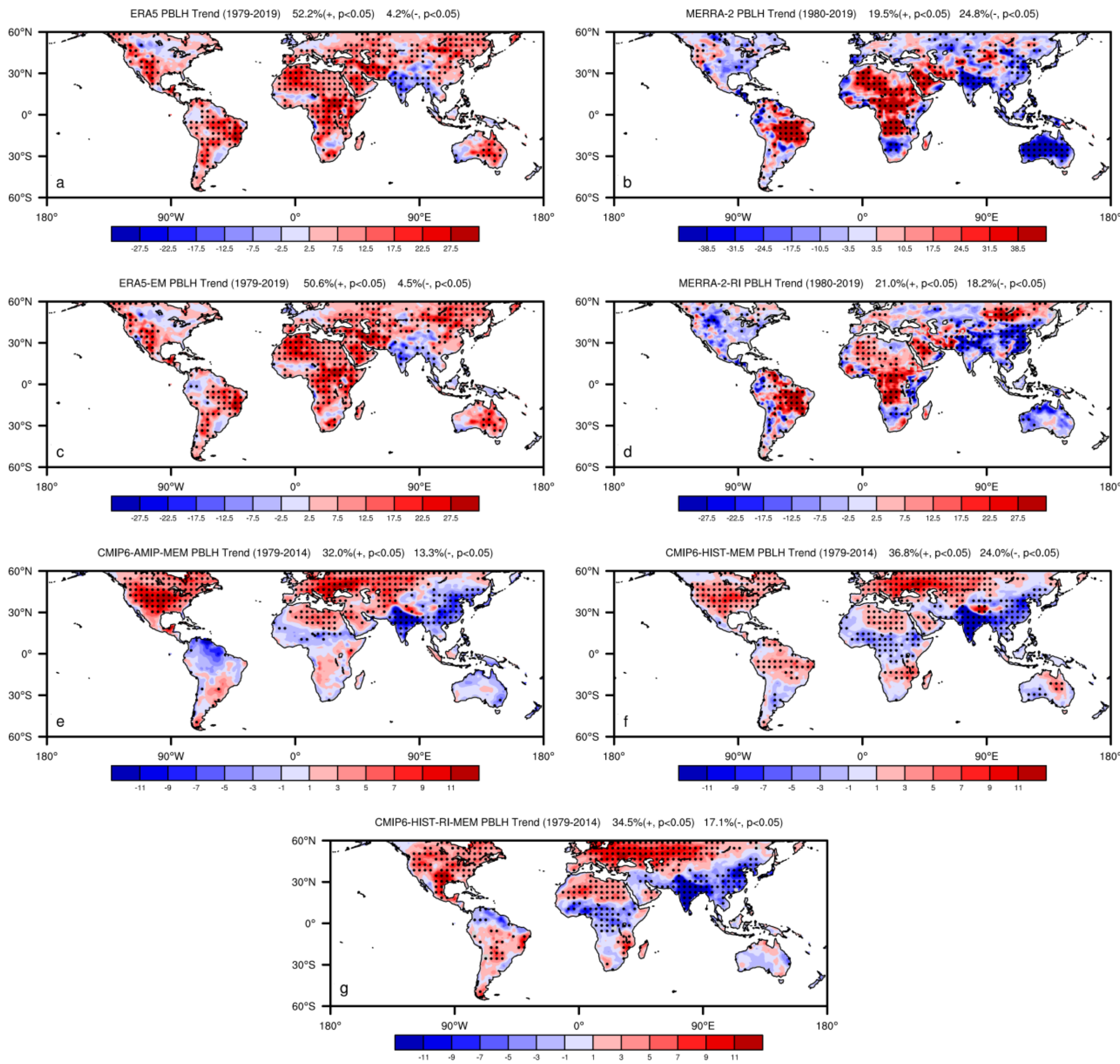

Fig S1. Same as Fig. 2, but the trend is estimated using the OLS method.

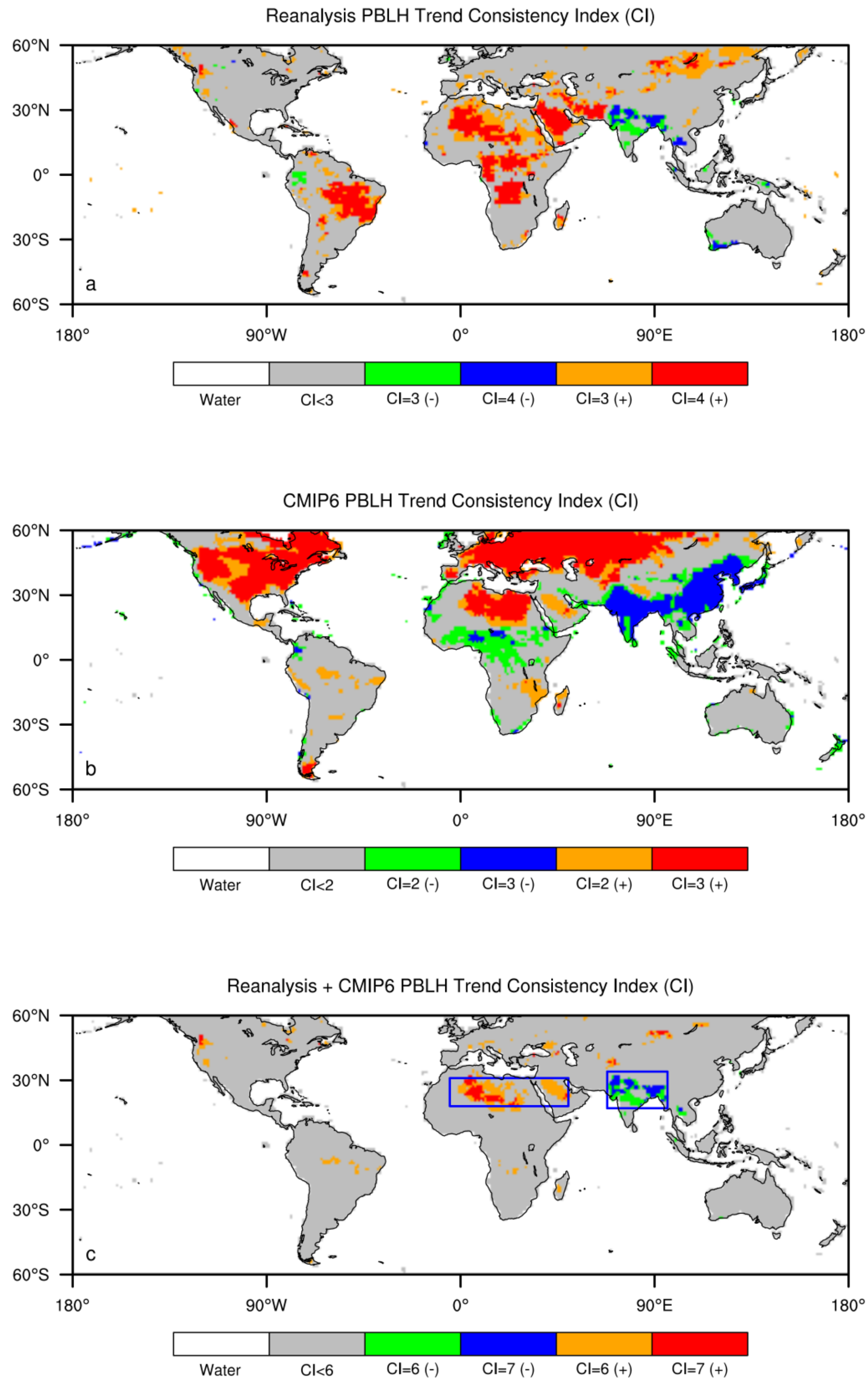

Fig S2. Same as Fig. 3, but the trend is estimated using the OLS method.