GNSS radio occultation in-filling of the vast radiosonde data gap of the UTLS over Africa reveals global, regional and coupled climate drivers of tropopause variability

Tong Ding[a], Joseph L. Awange [a], Barbara Scherllin-Pirscher[b], Michael Kuhn [a], Khandu [c], Richard Anyah [d], Ayalsew Zerihun [e], Luyen K. Bui [f,g]

[a]School of Earth and Planetary Sciences, Spatial Science Discipline, Curtin University, Western Australia, Australia

[b]Zentralanstalt für Meteorologie und Geodynamik, Graz, Austria

[c]Landgate (formerly the Department of Land Information), Western Australia, Australia

[d]Department of Natural Resources and the Environment, University of Connecticut, USA

[e]School of Molecular and Life Science, Curtin University, Australia

[f]Department of Civil and Environmental Engineering, University of Houston, USA

[g]Faculty of Geomatics and Land Administration, Hanoi University of Mining and Geology, Vietnam

Corresponding author: Tong Ding, `tong.ding@student.curtin.edu.au`

Abstract

Over much of Africa, radiosonde data are lacking; consequently, the African UTLS is understudied, and potential proxies such as climate models and reanalyses fail to capture the behaviour of the UTLS fully. This study pioneers the use of Global Navigation Satellite System Radio Occultation (GNSS-RO) data from 2001 to 2020 to address the radiosonde data gaps over Africa and contributes to a better understanding of the tropopause (TP) characteristics under the influence of multiple climate drivers. The analyses show that GNSS-RO data from CHAMP, GRACE, MetOp, COSMIC, and COSMIC-2 agree with radiosonde measurements with differences being smaller than 1 K in the UTLS; thereby enabling in-filling of 80% of the missing radiosonde data in Africa during 2001-2020. By contrast, the smoothed vertical temperature profiles of reanalysis products lead to a warm bias of +0.8K in ERA5 and +1.2K in MERRA-2, and these biases alter some vertical and temporal structure details, with possible implications on climate change detection and attribution. Furthermore, the analysis of GNSS-RO data over Africa revealed: 1) influences of global climate drivers on TP temperature, with QBO > ENSO > IOD > NAO > SAM > MJO, and on TP height with ENSO > QBO > NAO > MJO > IOD > SAM; 2) multiple coupled global climate drivers such as ENSO-MJO, ENSO-NAO etc.; 3) coupled global and regional climate drivers that influence the TP variability, e.g., ENSO-ITCZ; and 4), the deep convection associated with the Asian Summer Monsoon and Tropical/African Easterly Jet locally influence TP height.

1 Introduction

In Africa, which has high vulnerability to impacts of extreme weather events and climate change (IPCC, 2014), there is paucity of surface- and upper-air observations. The alarming state of the issue has been highlighted recently by the Wold Meteorological Organisation (WMO, 2020), which states that there is "a dramatic decrease of almost 50% from 2015 to 2020 in the number of radiosonde flights and/or observations, the most important type of surface-based observations for weather prediction and climate analysis" and furthermore "now has poorer geographical coverage". Over the last two decades, some 82% of the countries in Africa have experienced severe (57%) and moderate (25%) radiosonde data gap. For Africa, in particular, where many studies have indicated a possible increase in the frequency and severity of climate extremes (droughts and floods) (Agutu et al., 2017; J. Awange, 2021), increased water stress (Terink et al., 2013), and

food security risks (Gregory et al., 2005), the situation is dire, prompting a call by Peuch (2020) for urgent need to fill the data gap in Africa and globally.

Radiosondes from weather balloons have conventionally been used as means of measuring atmospheric profiles of humidity, temperature, pressure, wind speed and direction. High quality, spatially and temporally "continuous" data from upper-air monitoring along with surface observations are critical bases for understanding weather conditions and climate trends, and providing weather and climate information for the welfare of societies. Reliable and timely information underpin society's preparedness to extreme weather conditions, and to changing climate patterns. Africa, however, suffers from insufficient, non-consistent in-situ data, and poor data quality (Thomson et al. 2011). Moreover, the weather balloons rely heavily on the amount of gas fuel used and are usually launched at specific time intervals (e.g., every 6 to 12 hours) from airports or other approved locations. Additionally, the commonly used hydrogen generator is an expensive instrument that requires consistent maintenance and technical expertise, which is out of reach for most African countries. Other challenges pertaining to the African radiosonde data include systematic error in data and misaligned time series (see, e.g., J. Wang & Zhang 2008; Ramella Pralungo et al. 2014). These have led to the requirement of a re-evaluation of African radiosonde observation reliability (Lanzante et al., 2003).

Despite the establishment of the African Monsoon Multidisciplinary Analysis (AMMA) radiosonde program to address data quality of radiosonde stations over West Africa, 30% of the data are still lost due to persistent technical failures (Fink et al. 2011). Moreover, Thorne & Vose (2010) reported heterogeneity of the in-situ radiosonde record and its potential impact on long-term reanalyses, particularly over Africa (Lanzante et al., 2003). Further issues include temporal and spatial difficulties for long-term modeling (Seidel et al. 2004), limited data coverage in the southern hemisphere (J. Awange, 2018), challenges of operating over the ocean (J. L. Awange, 2012), and high operational costs (Parker et al., 2008). The vast data gap in such a large part the global landmass, home to some of the most vulnerable societies, the aforementioned call has galvanised a global effort (Taalas et al., 2021) to "plug the data gap" in the decade ahead and halt a further deterioration in the observation networks.

The use of global reanalysis products could alleviate these challenges. However, different reanalysis products exhibit considerable discrepancies in regional studies (Chen

et al., 2014). Additionally, differences in the responses of global temperature to volcanic eruptions (Fujiwara et al., 2015) and biases associated with the reanalyses' low vertical resolution (Meng et al., 2021) lead to different UTLS structures among reanalysis products and therefore to different tropopause characteristics (Homeyer et al., 2010). Recently, the fifth-generation Reanalysis (ERA5) of the European Centre for Medium-Range Weather Forecasts (ECMWF) has been added to the plethora of tropopause monitoring products offering higher vertical resolution than previous reanalysis data sets. However, its performance against in-situ observations over Africa's upper-air region has not been investigated yet.

In light of these shortcomings over Africa, Isioye et al. (2015) proposed the exploration of space-based techniques such as Global Navigation Satellite System (GNSS) - Radio Occultation (GNSS-RO) for addressing the radiosonde data gap. The inception of the GNSS-RO technique that utilizes GNSS signals onboard of LEO satellites to remotely sense the Earth's atmosphere has emerged as a state-of-the-art data set over the last two decades, providing high vertical resolution atmospheric profiles around the tropopause region (see, e.g., Scherllin-Pirscher et al. 2021 and reference therein). Examples include; the CHAllenging Mini-satellite Payload (CHAMP), the Gravity Recovery And Climate Experiment (GRACE), GRACE Follow-On (GRACE-FO), the Meteorological Operational (MetOp) satellites, and the Constellation Observing System for Meteorology, Ionosphere, and Climate (COSMIC). They have operated continuously around the globe under all-weather conditions, and offer millions of consistent measurements (Angerer et al., 2017) that are successfully assimilated in to global reanalyses to improve global weather predictions and climate modeling (Ho et al. 2020). J. Awange (2018) showed that GNSS-RO can improve the deficiency of high-quality data in remote areas, which cannot be achieved by radiosondes alone. The recently launched COSMIC-2 mission (Anthes & Schreiner, 2019; Schreiner et al., 2020) increased sampling density at low- and mid-latitudes and provides reams of near-real-time data for weather research (Lien et al., 2021). The innovation of COSMIC-2 has enormously extended the data coverage and thereby significantly filled the data gap over Africa. Nevertheless, the COSMIC-2 products have neither been tested nor explored over the African continent yet.

Clearly, infilling of the African continent's upper tropopause-lower stratosphere (UTLS) will benefit the understanding of its climate, which is influenced by numerous global, regional, and local climate variability modes, e.g., the Indian Ocean Dipole (IOD; Saji et

al. 1999), El Niño–Southern Oscillation (ENSO; Scherllin-Pirscher et al. 2012), North Atlantic Oscillation (NAO; Hurrell et al. 2003), Southern Annular Mode (SAM; Fogt et al. 2011), Madden–Julian Oscillation (MJO; Zhang 2005), Quasi-Biennial Oscillation (QBO; Baldwin et al. 2001), the Inter Tropical Convergence Zone (ITCZ; Basha et al. 2015), the Tropical Easterly Jet (TEJ; Nicholson & Klotter 2021), West African Monsoon (WAM; Sultan & Janicot 2003), growing ozone pollution in the Southern Africa (Thompson et al., 2014), and changing climate in the mountainous region (Kubokawa et al., 2016). Even though these climate drivers are known to trigger multi-effect severe weather and climate conditions (e.g., J. L. Awange et al. 2013), there are only a few studies that attempt to understand the African tropopause variability in relation to these climate drivers.

Indeed, that the tropopause, a transitional region (i.e., 9 km to 17 km) between the stratosphere and the troposphere could offer the possibility to quantitatively analyze upper level atmospheric characteristics in relation to regional climate variability and change over Africa is supported by regional studies that have been conducted, e.g., in Australia (Khandu et al., 2011), Ganges-Brahmaputra-Meghna basin (Khandu et al., 2016), South America (Nascimento et al., 2020), and the tropical and subtropical tropopause regions (Tegtmeier et al. 2020). These studies were motivated by Santer, Sausen, et al. (2003); Santer, Wehner, et al. (2003) who reported that the warming of the upper troposphere (UT, e.g., due to increased greenhouse gas emissions) and cooling of the lower stratosphere (LS, e.g., due to stratospheric ozone depletion) lead to a rise in tropopause height, indicating that climate change impacts can be quantified through UTLS monitoring. Subsequently, many studies on monitoring the global tropopause have explored various data sets ranging from in-situ radiosonde to high vertical resolution data acquired from Low Earth Orbit (LEO) satellites, i.e., satellite-based data (e.g., Shangguan et al. 2019), radiosonde data (e.g., Feng et al. 2012), and reanalysis products (e.g., Xian & Homeyer 2019). Nonetheless, such global tropopause assessments have limited physical implications and spatial importance at regional scales such as Africa (Franzke et al. 2020). For example, recent studies on global tropopause by Xian & Homeyer (2019) and tropical tropopause characterization by Tegtmeier et al. (2020) revealed large spatial inconsistencies between reanalyses and observations over equatorial Africa. A comprehensive and in-depth analysis of tropopause characteristics over Africa is, thus, still missing.

Despite almost two decades of advancement in the application of GNSS-RO globally and regionally, its use in Africa's climate-related studies has not been explored. This

study, therefore, takes advantage of the 20 years (2001–2020) of GNSS-RO data over Africa from multiple mission GNSS and LEO satellites to (i) statistically compare and validate GNSS-RO (including the latest COSMIC-2) profiles against radiosonde and reanalysis products (including the state-of-the-art ERA5), (ii) address the missing radiosonde data and assess reanalysis products of Modern-Era Retrospective analysis for Research and Applications (MERRA-2) and ERA5, and finally (iii), use the GNSS-RO and reanalysis products to assess the tropopause variability over Africa in relation to global (e.g., ENSO, IOD, NAO, SAM, MJO, and QBO), regional (e.g., ITCZ, TEJ, WAM, and regional ozone variability), and local climate drivers (e.g., impacts of local air-ocean interaction on tropopause).

2 Data and methods

2.1 Data

2.1.1 Radiosonde

Radiosonde data used in this study is obtained from Integrated Global Radiosonde Archive version 2 (IGRA v2), an improvement of IGRA version 1 (Durre et al. 2018; `https://www1.ncdc.noaa.gov/pub/data/igra/`). The IGRA data provides primary atmospheric variables (i.e., geopotential height, temperature, relative humidity, pressure, and wind speed, etc.) at the mandatory, significant, tropopause, and surface pressure levels (Durre et al. 2006) for each launch time. World Meteorological Organization (WMO) archives radiosonde station information across the globe and provides each radiosonde station's historical and operational status, including location, latitude, longitude, declared reporting status, type of station, WMO identifier, and elevation (World Meteorological Organization 2020; `https://www.wmo.int/datastat`). In this study, WMO data are used to reevaluate the African radiosonde network. Problematic radiosonde data are identified if the radiosonde profile has less than five temperature readings. Furthermore, low-quality data is identified if the balloon did not reach the desired pressure level (i.e., at least 30 to 70 hPa dependent on the region). Neither problematic nor low-quality radiosonde data are used for comparison with GNSS-RO and reanalyis data in this study.

2.1.2 GNSS-RO

GNSS-RO measurements used in this study are provided by the COSMIC Data Analysis and Archive Center (CDAAC; `https://cdaac-www.cosmic.ucar.edu/cdaac/doc/about.html`). Multiple products, including raw measurements and retrieved atmospheric profiles are made available. CDAAC is a processing and service center for data collected from various RO missions such as CHAMP (2001–2008), GRACE (2007–2017), COSMIC (2006–2020), MetOp-A (2016–2019), and COSMIC-2 (2019–2020). CHAMP, launched in July 2000 to succeed the first proof-of-concept occultation mission Global Positioning System / Meteorology (GPS/MET), collected approximately 650 vertical profiles over Africa per month, summing up to 59,620 profiles from May 2001 - October 2008 (see Figure 1). Note that CHAMP did not collect observations from July 3, 2006, to August 8, 2006 due to technical issues (Foelsche et al., 2008).

GRACE, a US/German collaborative mission, was launched in March 2002 and provided the first occultation observation on July 28, 2004 (Wickert et al., 2005). Before 2007, its data were not made available from CDAAC. It has provided approximately 530 profiles over Africa per month, with a total of 68,537 profiles from February 2007 to November 2017 (see Figure 1). The GRACE satellite mission ended in 2017 and was succeeded by the GRACE-FO mission launched on May 22, 2018, whose products were not available at CDAAC at the time of this study. The number of GNSS-RO profiles has grown considerably since the successful launch of the COSMIC constellation of six satellites in 2006 (see Figure 1). COSMIC is a joint project between the Taiwanese National Space Program Office (NSPO) and University Corporation for Atmospheric Research (UCAR). The constellation provided approximately 5,800 profiles over Africa per month, i.e., a total of 991,691 vertical profiles between April 2006 and April 2020. Note that COSMIC exhibits a data gap of RO profiles between April 1, 2019, and September 30, 2019, due to decommissioning of the mission. Therefore, post-processed MetOp-A data are used, which provided approximately 2,600 profiles per month and a total of 117,043 profiles from January 2016 to September 2019.

Following the successful operation of the COSMIC mission, the next generation COSMIC-2/FORMOSAT-7 was launched on June 25, 2019, providing approximately 30,000 high-quality profiles over Africa per month. As shown in Figure 1, COSMIC-2 has an increased number of profiles over Africa due to the low inclination of the six COSMIC-2 satellites,

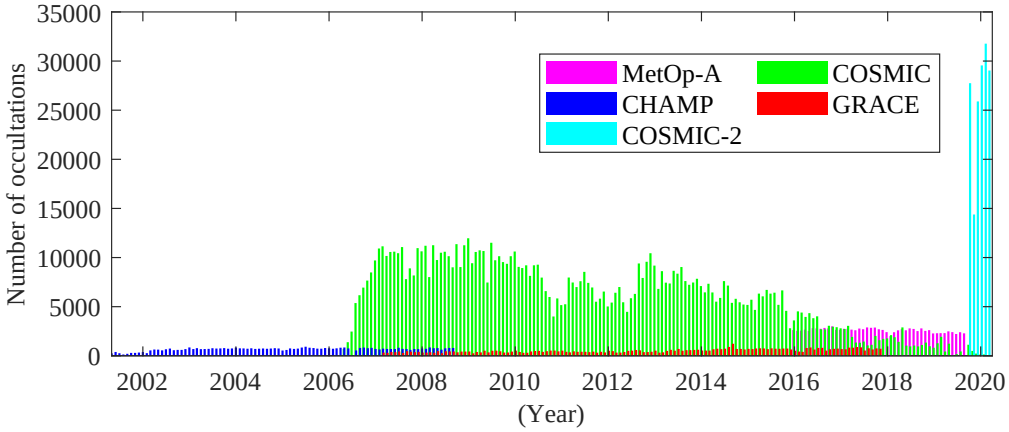


Figure 1: Temporal evolution of the number of GNSS-RO dry temperature profiles that were used to study tropopause variability and fill the radiosonde data gap over Africa from May 2001 to April 2020. The figure shows the number of monthly observations provided by CHAMP, GRACE, COSMIC, MetOp-A, and COSMIC-2.

which leads to a higher measurement density at low- and mid-latitudes (Schreiner et al., 2020) compared to other RO missions. Daily COSMIC-2 data are published by CDAAC at 02:00 UTC on the following day. The constellation has collected a total of 192,608 vertical atmospheric profiles over Africa from October 2019 to April 2020 (see Figure 1). In this study GNSS-RO profiles are excluded if the identified tropopause temperature and height exceed the range of 150 K to 250 K and 5 km to 25 km, respectively.

2.1.3 Reanalyses

The reanalyses data sets used in this study are temperatures, pressures, and geopotential heights provided by ERA5 and MERRA-2 for the period consistent with that of GNSS-RO data sets (May 2001 to April 2020). ERA5 is the 5th generation ECMWF reanalysis (Hersbach et al. 2020; `https://apps.ecmwf.int/data-catalogues/era5/?class=ea`), replacing the ERA-Interim reanalysis. It provides hourly atmospheric profiles at 137 model pressures levels on a global 0.25° × 0.25° grid. ERA5 combines model data with observations from across the world through data assimilation to produce the best estimates of the state of the atmosphere, and has an improved spatial and temporal resolution compared to MERRA-2 (see Table 1(I)). Note that the ERA 5.1 is employed in

this study for the period of 2001 to 2006 to avoid the tropopause cold bias (Simmons et al., 2020). ERA5 is used for the remaining period.

The MERRA-2 data have a spatial resolution of $0.5° \times 0.625°$ and a consolidated vertical grid at 72 model levels (Gelaro et al. 2017; `https://disc.gsfc.nasa.gov/datasets`). Model level products of both ERA5 and MERRA-2 are used from 2001 to 2020 since they offer higher vertical resolution profiles.

2.1.4 Climate variability indices

The global climate variability indices used in this study are ENSO (Trenberth & Stepaniak 2001), NAO (Hurrell & Deser 2010), IOD (Saji et al. 1999), QBO (Baldwin et al. 2001), MJO (Wheeler & Hendon 2004), and SAM (Marshall 2003). All climate indices are obtained from May 2001 to April 2020. Table 1(II) provides a summary of the indices used. MJO, SAM, and NAO are smoothed with a 3-month moving average to reveal mid- and long-term variability, while ENSO, QBO, and IOD indices are provided already having been smoothed. The smoothed indices are checked against the raw indices to avoid over-smoothing that can lead to loss of climate variability patterns (see e.g., Hansen et al. 1998). No over-smoothed indices are detected.

Indices of regional climate drivers are calculated from zonal wind at 150 hPa (TEJ). ITCZ is represented by Outgoing Longwave Radiation (OLR) while the WAM is captured through rainfall data provided by Precipitation Estimation from Remotely Sensed Information using Artificial Neural Networks (PERSIANN; Hsu et al. 1997). PERSIANN contains the highest signal-to-noise ratio (SNR) and is thereby a suitable product to be employed over Africa (J. L. Awange et al., 2016). The southern Africa total ozone index is based on monthly area-averaged ozone values derived from an assimilation product that contains 15 different types of satellites observations. This type of ozone data has very few data gaps, but is only available from May 2001 to Dec 2018 (Copernicus Climate Change Service, 2018). Zonal wind and OLR are obtained from the NCEP reanalysis provided by the NOAA Earth System Research Laboratories (ESRL; Kalnay et al. 1996). In this study, the NCEP reanalysis serves as an independent data source.

Table 1: Summary of data used in this study. (I) Radiosonde, GNSS-RO, and reanalysis products. (II) Climate indices.

(I) Data used for this study				
Source	**Temporal resolution**	**Spatial resolution**	**Vertical resolution at UTLS (m)**	**Period**
Radiosonde (IGRA V2)	6-hourly or 12-hourly	50 stations	N/A	May 2001 - April 2020
CHAMP	650 profiles/month		≈100-300 (Scherllin-Pirscher et al., 2021)	May 2001 - Oct 2008
GRACE	530 profiles/month			Feb 2007 - Nov 2017
COSMIC	5,800 profiles/month			April 2006 - April 2020
COSMIC-2	30,000 profiles/month			Oct 2019 - April 2020
MetOp-A	2,600 profiles/month			Jan 2016 - Sep 2019
ERA5	Hourly and monthly	0.25°×0.25°	≈350 (Hersbach et al., 2020)	May 2001 - April 2020
MERRA-2	3-hourly and monthly	0.5°×0.625°	≈500 (Gelaro et al., 2017)	May 2001 - April 2020

(II) Climate indices			
Climate driver	**Region**	**Variables**	**url**
NAO	20°-80°N; 90°W-40°E	Sea Level Pressure (SLP) anomaly	[a]
IOD	(i) 50°-70°E; 10°S-10°N (ii) 90°-110°E; 10°S-10°N	Sea Surface Temperature (SST) gradient	[b]
ENSO	5°N-5°S; 170°-120°W	SST anomaly (Niño 3.4)	[c]
QBO	10°N-10°S	zonal-mean winds at 50 hPa	[d]
SAM	40°-65°S	zonal-mean SLP	[e]
MJO	Equatorward of 30°N, 20°E	200-hPa velocity potential anomalies	[f]
ITCZ	30°-38°E; 13°-18°N	OLR	[g]
TEJ	5°W-15°E; 5°-15°N	zonal-mean wind at 150 hPa	
Total ozone	10° - 35°E; 25° - 40°S	area-averaged total ozone	[h]
WAM	5°W - 5°E; 6° - 8°N	Precipitation	[i]

[a] `https://climatedataguide.ucar.edu/sites/default/files/nao_pc_monthly.txt`

[b] `https://psl.noaa.gov/gcos_wgsp/Timeseries/DMI`

[c] `https://psl.noaa.gov/data/correlation/nina34.data`

[d] `https://www.cpc.ncep.noaa.gov/data/indices/qbo.u50.index`

[e] `https://legacy.bas.ac.uk/met/gjma/sam.html`

[f] `https://www.cpc.ncep.noaa.gov/products/precip/CWlink/MJO/mjo.shtml`

[g] `https://psl.noaa.gov/data/timeseries/`

[h] `doi.org/10.24381/cds.4ebfe4eb`

[i] `http://chrsdata.eng.uci.edu/`

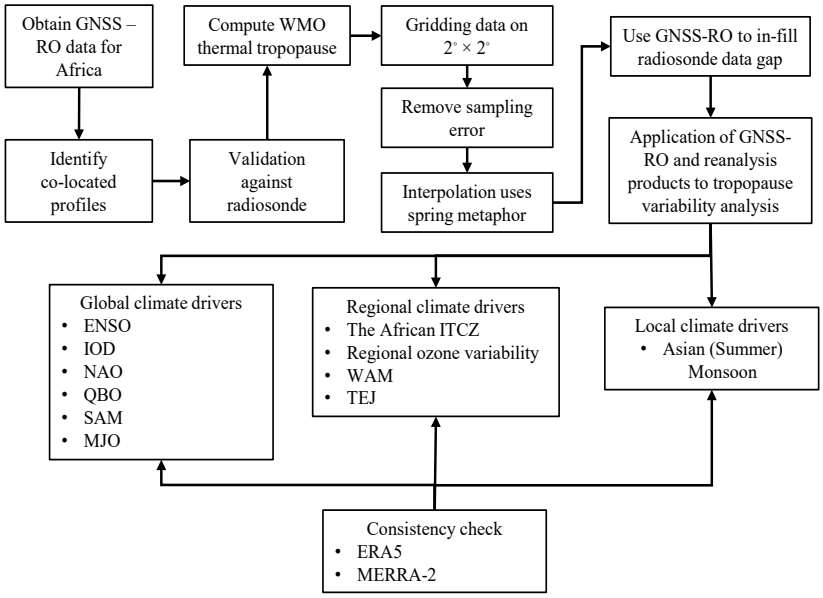


Figure 2: Schematic workflow of using GNSS-RO to address the radiosonde data gap and investigate tropopause variability over Africa in relation to climate drivers (global and regional) as well as local drivers.

2.2 Method

This study follows the workflow presented in Figure 2. First, GNSS-RO data are obtained from the various missions. Second, co-located (within 100 km and 3 h) GNSS-RO profiles, radiosonde observations, and reanalysis products are identified to conduct a thorough statistical validation and comparison. Then, tropopause temperature and height are obtained from individual profiles and gridded into 2 ° × 2° monthly climatological fields. The sampling error is estimated and subtracted from monthly-mean GNSS-RO fields (Pirscher et al., 2007; Foelsche et al., 2011). GNSS-RO are then interpolated using the spring analogy (D'Errico, 2004). Subsequently, the incomplete and inconsistent radiosonde record over Africa is filled with GNSS-RO data. Finally, GNSS-RO fields are used to characterize the African tropopause and analyze tropopause variability in relation to global, regional, and local climate drivers. Consistency checks between GNSS-RO and reanalysis products (i.e., ERA5 and MERRA-2) are carried out throughout.

2.2.1 Definition of the tropopause used in this study

The WMO (1957) thermal tropopause definition (also known as the lapse rate tropopause (LRT)) is applied to identify the tropopause from all data sets. WMO (1957) defines the tropopause as "the lowest level at which the lapse rate decreases to 2 °C km^{-1} or less, provided also the average lapse rate between this level and all higher levels within 2 km does not exceed 2 °C km^{-1}".

2.2.2 Validation method

To compare GNSS-RO with in-situ radiosonde data from across Africa, the mean temperature difference $\overline{\Delta T(l)}$ and the corresponding standard deviation $\sigma_{\Delta T}(l)$ are calculated at pressure levels l = 850, 700, 500, 400, 300, 250, 200, 150, 100, 70, 50, 30, and 20 hPa and at tropopause level by:

$$\overline{\Delta T(l)} = \frac{1}{M(l)} \sum_{i=1}^{M(l)} T_{\mathrm{GNSS}}(i,l) - T_{\mathrm{RS}}(i,l), \tag{1}$$

$$\sigma_{\Delta T}(l) = \sqrt{\frac{1}{M(l)-1} \sum_{i=1}^{M(l)} (T_{\text{GNSS-RS}}(i,l) - \overline{\Delta T(l)})^2}, \tag{2}$$

where $M(l)$ is the number of data points at the level l, T_{GNSS} and T_{RS} are the temperature from GNSS-RO and the radiosonde data, respectively (Wickert, 2004). $T_{\text{GNSS-RS}}$ are the actual temperature differences between GNSS-RO and radiosonde data. Comparisons between radiosonde and GNSS-RO data are made only for profiles less than 100 km and 3 hours apart from each other (Khandu et al., 2011). Data are discarded if they exhibit a temperature difference of more than 15 K at above 9 km for quality control. Differences caused by comparing GNSS-RO dry temperature and radiosonde physical temperature are usually less than 5 K above 9 km (see Scherllin-Pirscher et al., 2011). For all GNSS-RO and radiosonde profile pairs, their co-located ERA5 and MERRA-2 profiles are extracted as well and used for comparison. Kling-Gupta Efficiency (KGE; Gupta et al. 2009) is used to measure the agreement of GNSS-RO and reanalysis products against radiosonde.

$$KGE = 1 - \sqrt{(r-1)^2 + (\alpha-1)^2 + (\beta-1)^2} \tag{3}$$

where r represents the correlation between GNSS-RO and radiosonde data, α is the ratio between the standard deviation of GNSS-RO and the standard deviation of radiosonde

data, and β is the ratio between the mean of GNSS-RO and the mean of radiosonde data. The same procedure applies to reanalysis products for comparison. The KGE value equal to 1 indicates the model perfectly fits radiosonde data.

2.2.3 GNSS-RO climatological fields

All GNSS-RO data are monthly gridded at a spatial resolution of $2^\circ \times 2^\circ$, as an individual RO profile has on average a horizontal extension of 250 km along the ray (see, e.g., Scherllin-Pirscher et al. 2021). Due to the high consistency of GNSS-RO measurements (Angerer et al., 2017), high-quality profiles from several missions are merged for tropopause variability analysis. Tropopause parameters are obtained from dry atmospheric products, which accurately describe "real" physical conditions in regions where humidity is small (see, e.g., Kursinski et al. 1997; Scherllin-Pirscher et al. 2011). GNSS-RO sampling is inhomogeneous, and as such, it includes sampling errors due to unevenly distributed observations in both time and space (Pirscher et al., 2007; Foelsche et al., 2011). These sampling errors are estimated using co-located ERA5 reanalysis profiles (i.e., at four times 00:00, 06:00, 12:00, and 18:00 UTC) and monthly mean ERA5 fields. The climatological difference (i.e., the sampling error) of the reanalysis products was subtracted from the GNSS-RO climatology. More details on this method can be found in Pirscher et al. (2007) and Foelsche et al. (2008). Empty grid boxes are filled by using a spring analogy interpolation algorithm invented by D'Errico (2004). The sampling error-corrected spatially interpolated GNSS-RO fields are used for tropopause variability analysis.

2.2.4 Correlation analysis

The instantaneous Pearson correlation coefficient is used to analyze the relationship between global climate indices and tropopause variability, see, e.g., Scherllin-Pirscher et al. 2012; Anyah et al. 2018. First, low pass filters (i.e., order = 4; cutoff period = 12 months) are applied to African tropopause temperature and height fields to filter noise. Then, the spatial-temporal relationship between deseasonalized tropopause parameters (temperature and height) and climate indices is established. Statistically insignificant results ($P > 0.05$) are removed from the analysis.

2.2.5 Principal Components Analysis

Principal Components Analysis (PCA; Preisendorfer & Mobley 1988) is applied to extract the most dominant signals explaining the majority of variability in the two-dimensional atmospheric data sets. The input data are first deseasonalized and then decomposed into modes of variability ($D_{t,s}$). Each mode n is expressed by an Empirical Orthogonal Function (EOF) representing the spatial pattern s and a Principal Component (PC) describing the temporal component t:

$$D_{t,s} = \mathrm{PC}_{(t,n)}\mathrm{EOF}_{(s,n)}. \tag{4}$$

The eigenvalues' scree plot is applied to ensure that the retained EOFs represent more signal than noise, see e.g., Forootan et al. (2012, 2014).

2.2.6 Independent Component Analysis

While PCA does an excellent work in detecting major climate drivers over Africa, it lacks the robustness of the Independent Components Analysis (ICA; Hyvärinen & Oja 2000) needed to reveal hidden signals of the data that are not captured by PCA (see Westra et al. 2010). ICA is a higher order (fourth-order cumulant) statistics used to solve blind source separation by assuming the sources are statically independent in time, and the non-Gaussian distribution of data measured by the fourth-order cumulants at zero time lags (Forootan et al., 2012).

The kurtosis of time series k is used to compute the non-Gaussianity of time series (Comon, 1994)

$$k = \frac{E(x^4)}{E(x^2)^2} - 3, \tag{5}$$

where x represents each time series and E is the expectation function such as time average. Since regional climate drivers have assertive intraseasonal and interseasonal behavior that are not captured by PCA, ICA is applied to raw data to separate the seasonality link with the regional climate indices.

3 Results and Discussion

3.1 Potential of GNSS-RO to fill missing radiosonde data

3.1.1 Analysis the state of radiosonde data over Africa

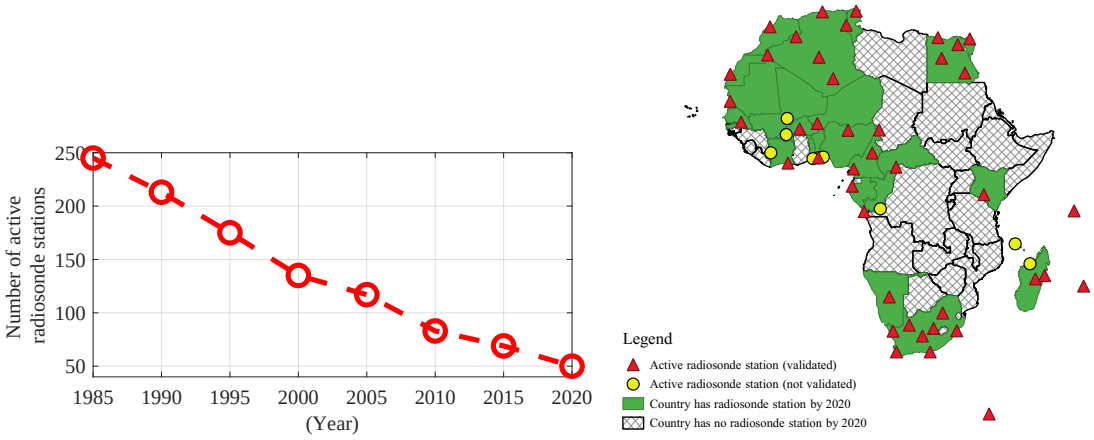
This study shows an alarming reduction in the number of radiosonde stations over the past decades in Africa (see Figure 3a). Since the 1980s, roughly 50 stations have been deactivated per decade, leading to a depletion of approximately 300,000 observations per decade. There are currently only 50 radiosonde stations left that operationally perform measurements in 26 African countries. The rest of the countries reported no active radiosonde stations by 2020 (see Figure 3b). Observations are predominately reported from specific regions (e.g., South Africa, the coastal rim of Northern Africa), but rarely from Central Africa and East Africa. Due to missing, problematic, and low-quality data, only 42 stations from 23 countries were used for validation in this study (see Figs. 3b and 3c). This means that approximately 57% of all African countries have no direct access to upper-air information from radiosonde stations. The minimum number of daily observations does not meet the WMO recommendations to study daily variation caused by diurnal and semidiurnal tides (see `https://library.wmo.int/doc_num.php?explnum_id=3158`).

This study used a total of 432,239 observations since May 2001. From the active stations, 165,775 measurements were missing or incomplete, and 51,333 were of low quality, accounting for 38% and 12% of the data, respectively. Most high-quality measurements were provided by Algeria, Egypt, and South Africa. While some countries only provide very few high-quality measurements (e.g., Morroco, Nigeria, Seychelles, and Western Sahara), the percentage of high-quality measurements is low, e.g., in Benin, Congo Republic, and Madagascar (see Figure 3c). Moreover, this study found 38% of the radiosonde records include missing data, and approximately 57% of all African countries have had several years without any operational radiosonde launches. Despite there being 23 African countries that reported radiosonde measurements in 2020, 14 of them kept less than 50 % of radiosonde records from 2001 to 2020, accounting for 25 % of all African countries. Overall, 82 % of African countries are either experiencing a severe radiosonde data gap (57 %) and have had data gaps in the past two decades (25 %), i.e., Figure 3b and 3c, which are now in-filled in Fig. 3d. On the other hand, the use of GNSS-RO measurements from 2001 onward and, in particular, the high number of COSMIC-2 profiles

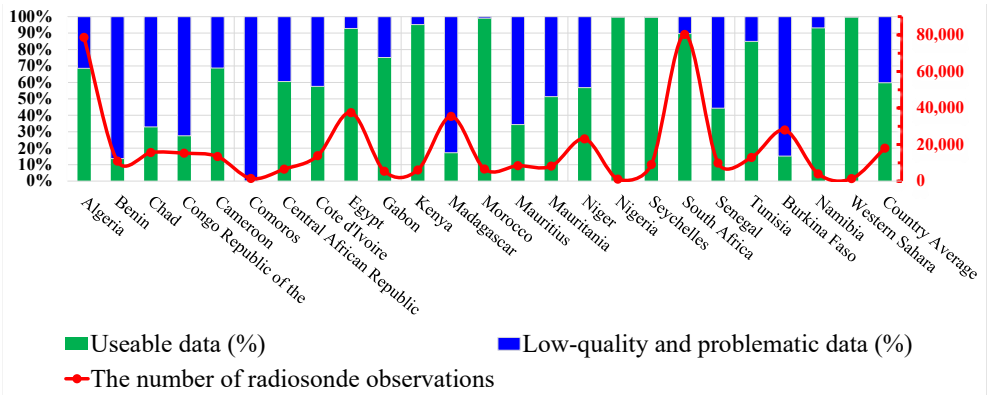
available after 2019 can mitigate the significant reduction of radiosonde data (see Figure 3d).

3.1.2 Statistical comparison and validation of GNSS-RO against radiosonde and reanalysis products

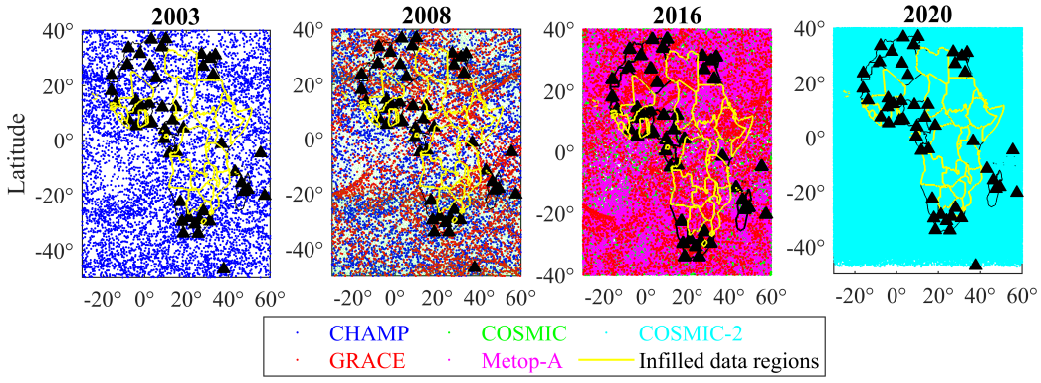
Figure 4 shows 12 radiosonde temperature profiles together with their co-located profiles from several GNSS-RO missions and reanalysis products. These profiles are selected to exemplify some characteristics representative of many other profiles. In general, the data sets show an almost identical behavior for the atmosphere between 8 km and 25 km, with the GNSS-RO products having a sharper, explicit feature in the upper troposphere and lower stratosphere (UTLS) region, where both ERA5 and MERRA-2 show a rather smooth behavior due to their lower vertical resolution. Profiles from Tamanrasset in Algeria (22.80°N, 5.43°E); Figure 4b, the Ivato International Airport in Madagascar (18.80°S, 47.48°E; Figure 4d), and Cape Town in South Africa (33.97°S, 18.60°E; Figure 4l) are examples that show the smooth spline used by the reanalysis products. In addition, Figure 4g and h show that both reanalyses sometimes misidentify the tropopause and in some cases, the reanalysis products do not to capture the tropopause inversion layer (e.g., Farafra in Egypt (Figure 4a), Nouakchott in Mauritania (Figure 4e), Pointe-Noire in Congo Republic (Figure 4h), and Tunis-Carthage in Tunis (Figure 4k)). On the other hand, GNSS-RO measurements from all satellites are in good agreement with radiosonde data. This is a promising result considering that GNSS-RO data can be used to fill the radiosonde data gap above Africa.



(a) (b)



(c)



(d)

Figure 3: (a) The number of operational radiosonde stations in Africa from 1985 to 2020, (b) availability of radiosonde stations over the African continent by 2020. Countries with operational radiosonde stations in 2020 (green) and those without operational radiosonde stations in 2020 (grey). (c) Percentage of high- (green) and low-quality (blue) measurements from 2000 to 2020 as well as total number of soundings per country (red line), and (d), GNSS-RO data coverage in 2003, 2008, 2016, and 2020.

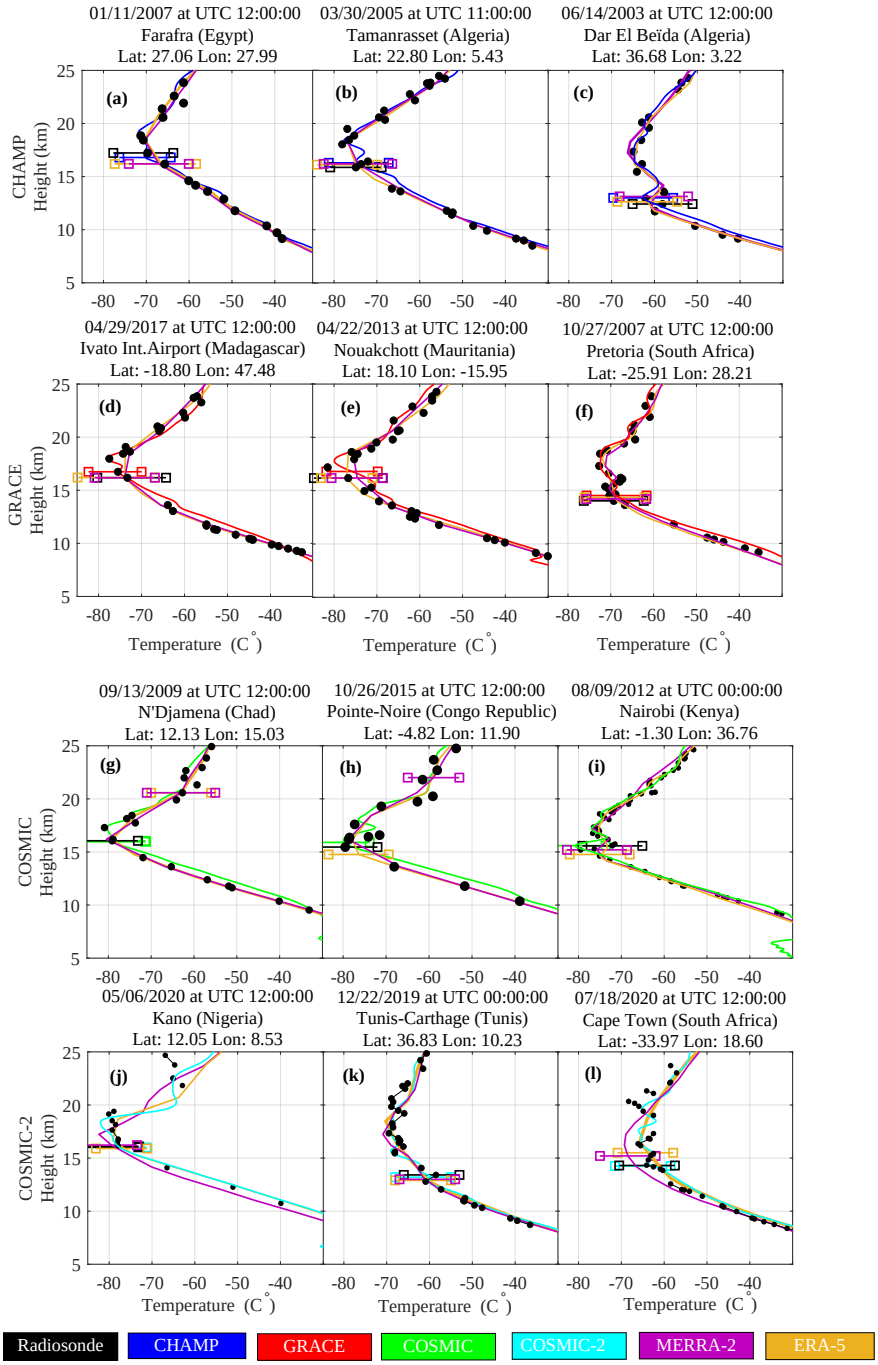


Figure 4: Radiosonde soundings, and co-located (within 100 km and 3 h) reanalysis profiles as well as GNSS-RO measurements from CHAMP (a-c), GRACE (d-f), COSMIC (g-i), and COSMIC-2 (j-l). Details about the radiosonde measurements, i.e., time of the observation, location with latitude and longitude, and provider country in the bracket are given in the panel titles. Horizontal lines indicate the location of the tropopause of each profile.

Figure 5 presents the overall statistical validation of the co-located GNSS-RO profiles with high-quality and complete vertical profiles from radiosonde soundings for the African continent for the period 2001 to 2020, including 109 CHAMP, 150 GRACE, 1,942 COSMIC, 91 MetOp-A, and 528 COSMIC-2 profiles. Between 10 km and 25 km the GNSS-RO data have a mean temperature difference to African radiosonde measurements of −0.68 K (CHAMP; Figure 5a), −0.41 K (COSMIC; Figure 5b), −0.70 K (GRACE; Figure 5c), −0.41 K (COSMIC-2; Figure 5d), and −0.64 K (MetOp; Figure 5e). The GNSS-RO and radiosonde data differ significantly below 10 km due to the presence of water vapour.

B.-R. Wang et al. (2013) reported that the Vaisala RS80 radiosonde is the main radiosonde type used in Africa, and the difference between radiosonde and GNSS-RO measurements is affected by the radiosonde bias that varies among different locations. Moreover, considering the Vaisala RS80 subject to a pressure bias (Inai et al., 2015) and knowing that GNSS-RO biases in the UTLS are small, it is reasonable to assume that GNSS-RO performs better than radiosondes and the bias is mainly attributable to radiosonde measurements rather than to GNSS-RO.

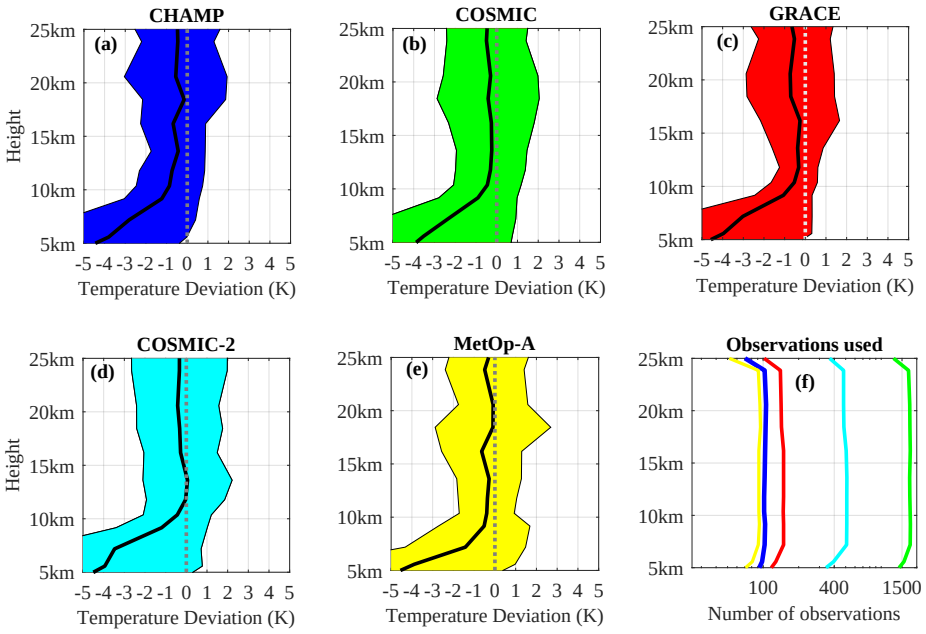


Figure 5: Mean temperature difference (thick black lines) and standard deviations (colored areas) between African radiosonde measurements and CHAMP (a; blue), COSMIC (b; green), GRACE (c; red) COSMIC-2 (d; cyan), and MetOp-A (e; yellow). (f) The number of observations used for validation at each geopotential height level.

3.1.3 The suitability of GNSS-RO data to fill radiosonde data gaps

Figure 6a-d shows examples of using GNSS-RO to infill the radiosonde data gap for some stations with long data gaps in the UTLS region due to problematic radiosonde measurements. The full time-series of GNSS-RO data are obtained from observation, which has been gridded using our method (see section 2.2.3). Radiosonde stations were launched decades ago (e.g., the South African station SFM68512 was launched in 1987 (Figure 6a), WAM68110 (Namibia; Figure 6b) in 1986, WIM60096 (Western Sahara; Figure 6c) in 1974, and NIM65202 (Nigeria; Figure 6d) in 1966). Tropopause height and temperature anomalies are calculated by subtracting the long-term mean of each data set. Overall, GNSS-RO performs well in addressing the radiosonde data gap. Figure 6e-h reveals issues regarding the use of reanalysis products to fill the radiosonde data gap over Africa. Both reanalysis products (MERRA-2 and ERA5) sometimes perform poorly in capturing the interannual variation of tropopause height, see Figure 6e and 6g. At selected stations, tropopause temperature differences between reanalyses and GNSS-RO are smaller than that between GNSS-RO and radiosonde measurements, indicating a potential radiosonde bias (Figure 6f and 6h).

Figure 7 shows the performance of real-time GNSS-RO vertical profiles and reanalyses simulated hourly vertical profiles relative to good quality radiosonde measurements. Towner et al. (2019) suggests that KGE values of 0.75 or higher indicate a good agreement between observations and simulated products, and KGE values lower than 0.5 are considered 'bad' (Andersson et al., 2017). In general, the GNSS-RO outperforms reanalysis products, with KGE values being higher than ERA5 and MERRA-2 at all height levels, indicating the advantage of GNSS-RO to in-fill the radiosonde data gap in the UTLS region. Figure 7b and 7c show the performance of decades-long time series derived from GNSS-RO data and reanalysis products relative to radiosonde records in both tropopause heights and temperatures. Given lower KGE values, the plot reveals that many countries suffer from poor long-term radiosonde records in the UTLS, e.g., BN (Benin), MR (Mauritania) and CD (Chad) (see Section 3.1.1). Noticeably, the data quality of tropopause heights is lower than that of tropopause temperature, which is possibly attributable to a radiosonde pressure bias discussed in Section 3.1.2. Higher KGE values in tropopause heights and temperatures are found in Algeria, Egypt, Morocco, and South Africa. Table 2 shows the Root Mean Sum Error (RMSE) of reanalysis products and GNSS-RO relative to consistent and good-quality radiosonde time-series from these four countries,

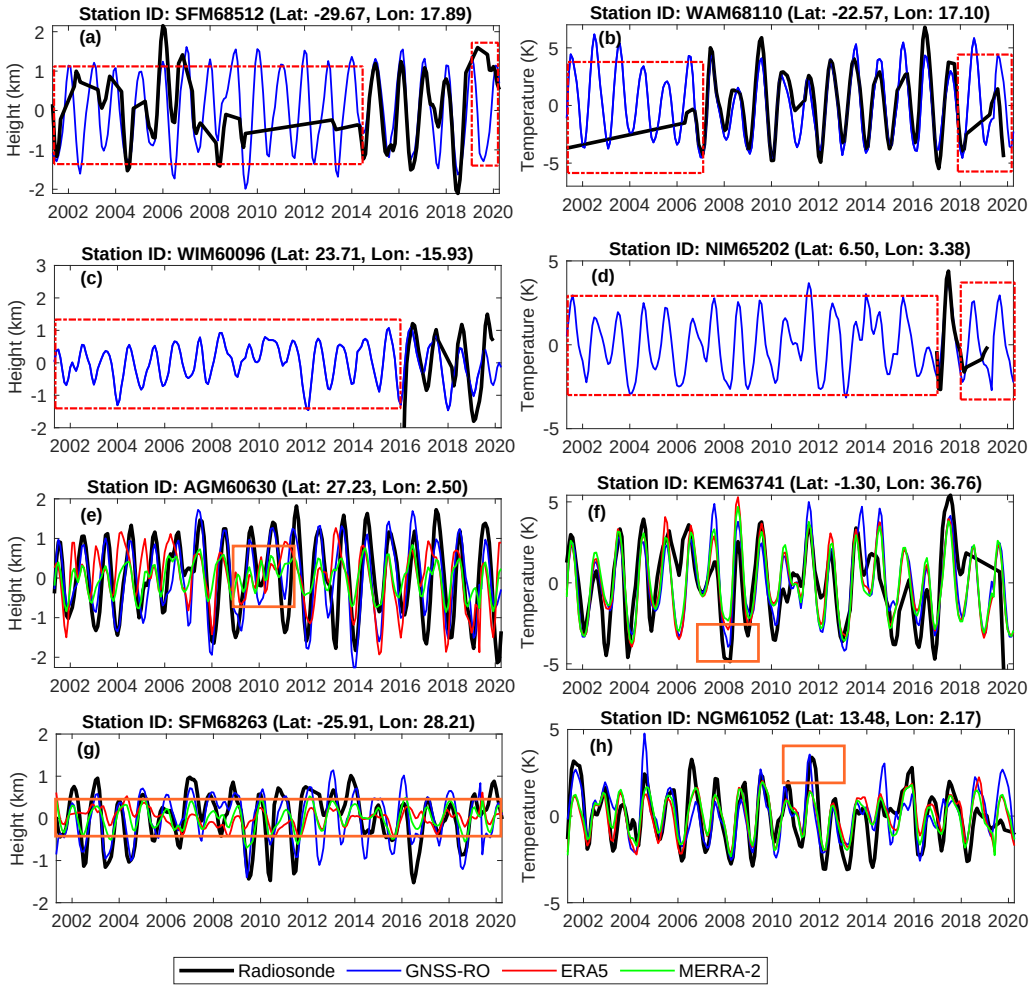


Figure 6: Time series of tropopause height anomalies (left panels) and tropopause temperature anomalies (right panels) from 2001 to 2020 for 8 African stations from radiosonde measurements (black), GNSS-RO (blue), ERA5 (red), and MERRA-2 (green). Red dashed boxes in a-d indicate long radiosonde data gaps due to missing, incomplete, or low-quality data. Orange boxes in e-h indicate time periods with larger differences between reanalyses and radiosondes. The panel titles contain radiosonde stations' country code, station code, latitude, and longitude.

Table 2: Root Mean Square Error (RMSE) of GNSS-RO, ERA5, and MERRA-2 tropopause height and tropopause temperature relative to consistent record from high-quality African radiosonde measurements.

Root Mean Square Error (RMSE)			
Parameters	**GNSS-RO**	**ERA5**	**MERRA-2**
Height (in km)	0.37	0.45	0.67
Temperature (in K)	0.85	0.98	1.03

and the GNSS-RO has the best RMSE value of tropopause heights (0.37 km) and tropopause temperature (0.85 K), followed by ERA5 and MERRA-2. In contrast, radiosonde stations from other countries lack quality. As a whole, Figure 7b and 7c indicate that about 78% of the countries had their height profiles' KGE and 57% of temperature profiles' related KGE values of less than 0.5. These corroborate the findings in Section 3.1.1 that indicated that about 80% of the African countries lacked radiosonde data while 57% suspended their radiosonde stations prior to 2020 and 25% of countries have more than 50 % missing radiosonde data, underlying the importance of GNSS-RO to the data deficient Africa. In general, the ongoing developments in GNSS-RO remote sensing to obtain full coverage over the continent—by providing consistent measurements simultaneously—are essential to preserving the valuable source of atmosphere observations for Africa. More importantly, GNSS-RO provides real-time observations, and data become available in near-real-time, while ERA5 is updated daily but with a latency of 5 days. GNSS-RO can provide more than 1.5 million accurate and consistent atmospheric vertical profiles for data-deficient Africa, which is not achievable using radiosonde measurements alone.

3.2 Spatio-temporal characteristics of the African tropopause

Figure 8 presents the area-averaged temporal variability of the tropopause height (Fig. 8a-d) and temperature (Fig. 8e-h) of GNSS-RO measurements and reanalysis products in four geographical regions. A consistent seasonal variation is found for all data sets. However, the reanalysis products uniformly show a lower tropopause height than GNSS-RO (see Fig. 8a-d) with largest differences at tropical latitudes. Tropopause temperature is, in general, in better agreement among the data sets with a small positive bias of the reanalysis products (see Fig. 8e-h). This bias is possibly caused by the re-

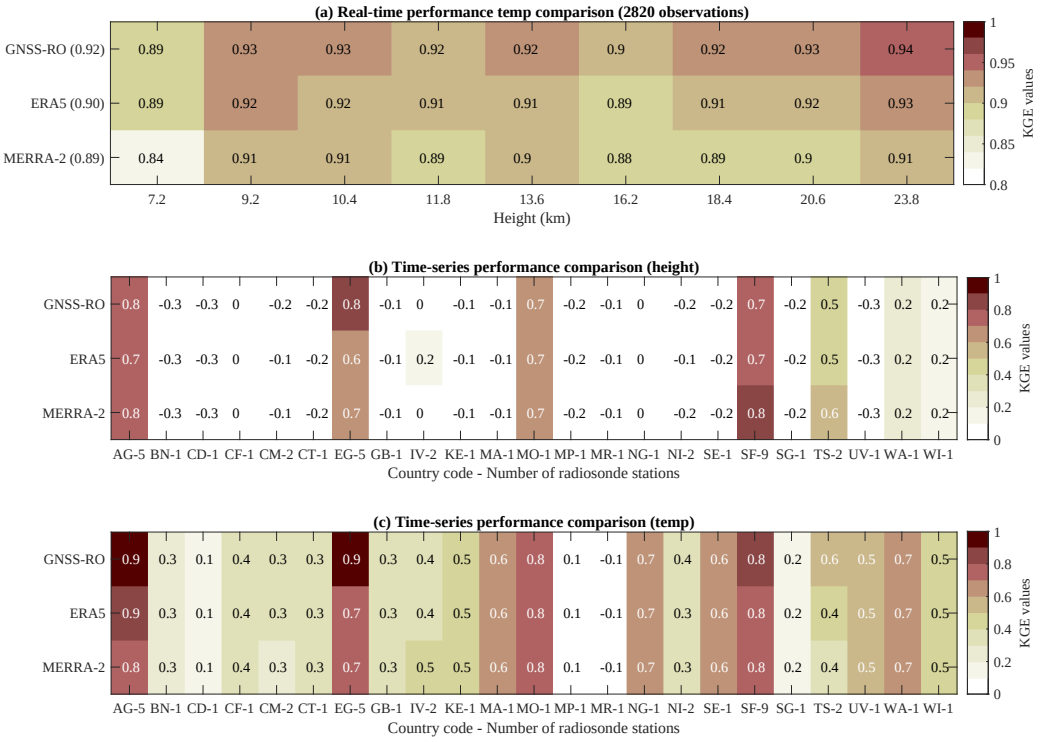


Figure 7: (a) Evaluate the Kling-Gupta Efficiency (KGE) values for the performance of real-time vertical profiles from reanalyses and GNSS-RO against good radiosonde observations in different heights, and the mean KGE value is given in the bracket. Evaluate the KGE values for time-series of tropopause heights (b) and temperatures (c) based on reanalysis products and GNSS-RO against radiosonde records. The X-axis shows the country code and the number of radiosonde stations used. Information for country code is provided in the supplementary material. KGE values lower than 0.5 indicate countries with unreliable radiosonde data (e.g., Figure 3c; Andersson et al. (2017)).

analyses coarser vertical resolution. This behavior is consistent with a recent study by Meng et al. (2021), who found a global mean pressure bias of −0.9 hPa in MERRA-2, attributable to a large bias over the subtropics in DJF and JJA. Similarly, Tegtmeier et al. (2020) showed that MERRA-2 and ERA5 exhibit a warm bias ranging from 0.2 K to 1.5 K over the tropics. In this study, the tropopause temperature based on ERA5 and MERRA-2 indicates a warm bias of 0.8 K and 1.2 K, respectively. ERA5 and MERRA-2 tropopause heights are negatively biased by 0.3 km and 0.5 km, respectively. Thus, ERA5 is generally in better agreement with GNSS-RO than MERRA-2. The reanalysis products indicate clear improvements after being deseasonalized and detrended. The mean difference between GNSS-RO and reanalyses reduces to less than 0.02 km and 0.02 K for height and temperature, respectively, and the primary cause of the seasonal bias is due to underestimation of seasonal amplitude.

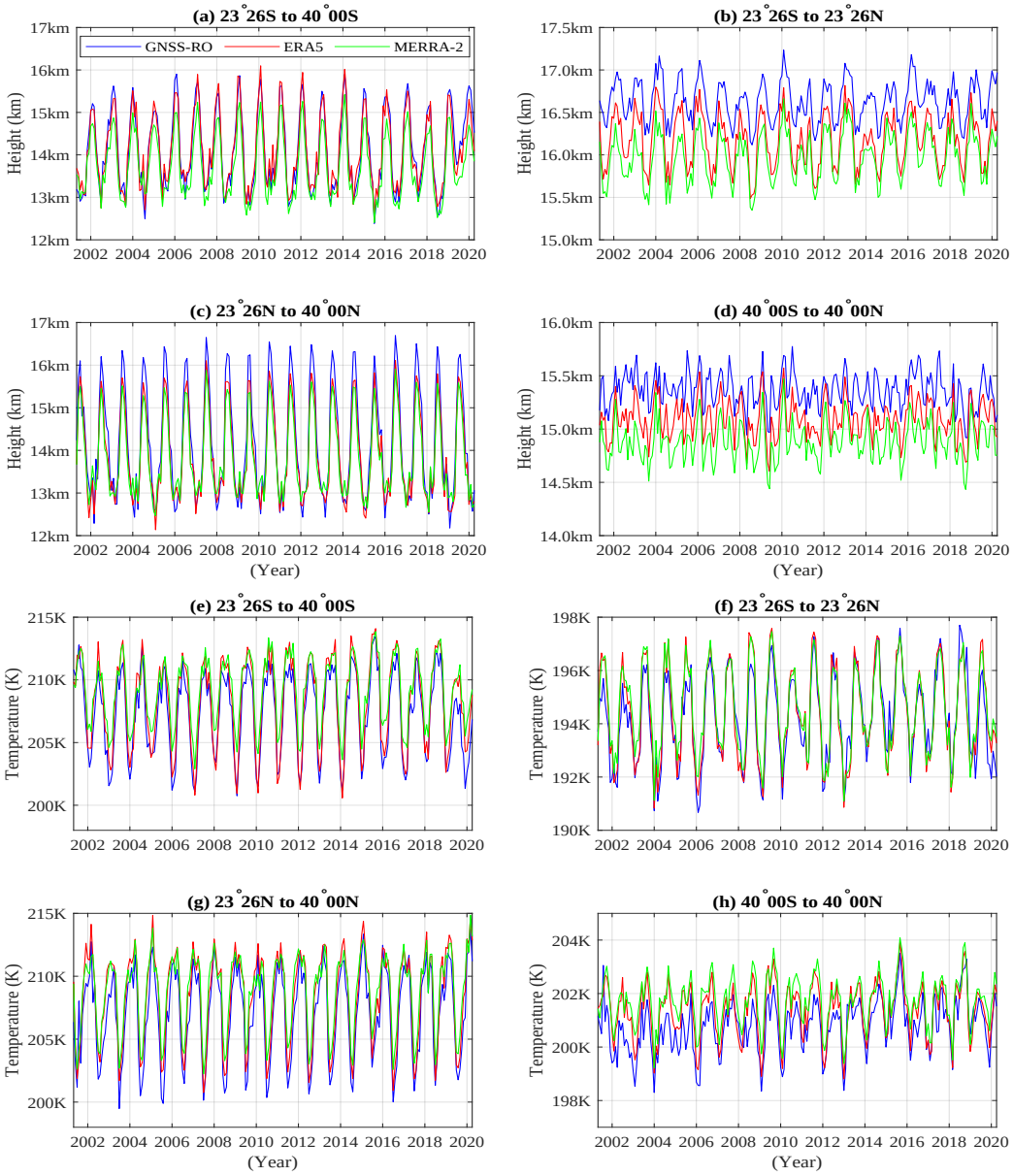


Figure 8: Temporal evolution of tropopause heights (a-d) and temperatures (e-h) from GNSS-RO and reanalysis products in southern subtropics (23°26'S to 40°00'S), tropics (23°26'N to 23°26'S), northern subtropics (23°26'N to 40°00'N), and over entire Africa (40°00'S to 40°00'N). The east-west extension is the same for all areas and corresponds to the extend of the study area.

The long-term mean tropopause height and temperature with respect to latitude are presented in Figure 9. The bias of the reanalyses tropopause heights are again clearly seen at tropical latitudes (23°S to 23°N). From 10°S to 10°N, this bias amounts to 0.5–

0.6 km. Polewards of 30°S, ERA5 is in good agreement with GNSS-RO but MERRA-2 tropopause height is negatively biased (0.4 km). Polewards of 30°N, however, GNSS-RO tropopause height is higher than that of ERA5 and MERRA-2 by 0.2 km and 0.3 km, respectively. The temperature bias of MERRA-2 increases from 0.7 K at tropical latitudes (10°S to 10°N) to 1.8 K at extratropical latitudes (polewards of 30°N/S). Differences between ERA5 and GNSS-RO remain within 0.8 K at tropical latitudes. Polewards of 30°N/S, ERA5 tropopause is biased by 1.5 K.

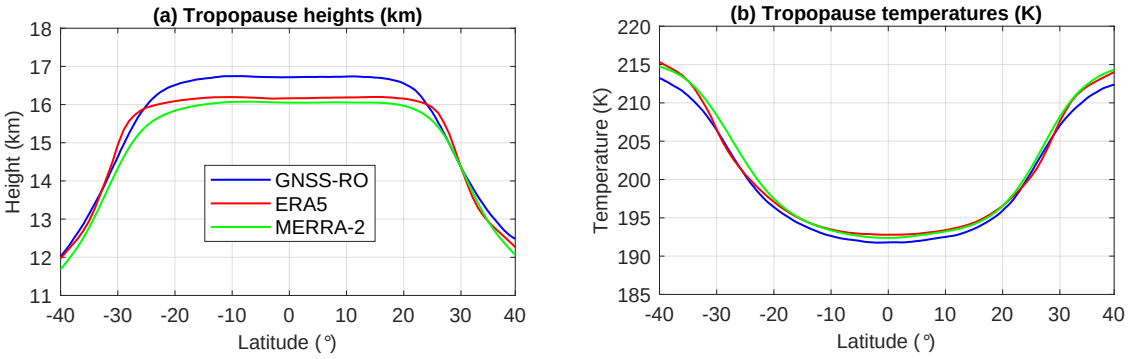


Figure 9: Latitudinal variation of the long-term mean (2001–2020) tropopause height (a) and temperature (b) from GNSS-RO, ERA5, and MERRA-2.

3.3 Variability of the African tropopause in relation to climate drivers

3.3.1 Variability in relation to global climate drivers

Figure 10 shows the correlation between six global climate driver indices (i.e., ENSO, MJO, NAO, SAM, IOD, and QBO) and deseasonalized timeseries of tropopause heights and temperatures for the period of 2001 to 2020 estimated with GNSS-RO and reanalysis products. Deseasonlisation is to highlight long-term variations. The table embedded in Figure 10 summarizes the relationship between global climate drivers and tropopause parameters for several African regions.

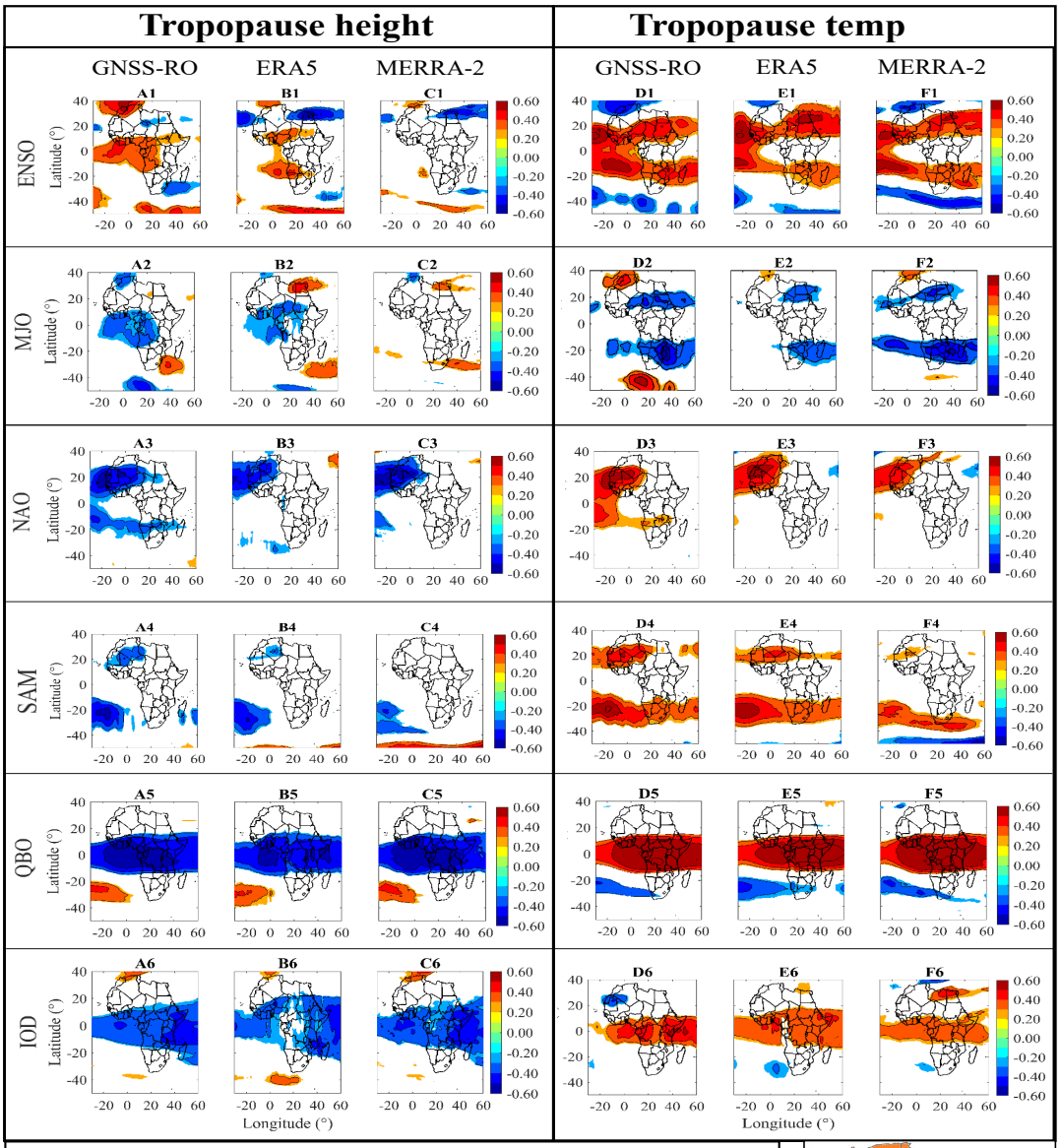
In general, reanalyses and GNSS-RO reveal similar correlation patterns with these six climate drivers. A connection between tropopause parameters and ENSO is found over most of the African continent. The atmospheric ENSO signal is stronger in tropopause temperature than in tropopause height. Both exhibit a rather zonally-symmetric struc-

ture with that of tropopause temperature being strongest over the subtropical regions (i.e., 10°S/N to 28°S/N). Tropopause height/temperature shows a negative/positive correlation over Northern Africa, Western Africa, and Southern Africa (see Figure 10(A1-F1). Other strong ENSO signals with opposite signs are observed over the Southern Ocean and close to Marocco and Algeria. These results are in good agreement with previous studies, see e.g., Rieckh et al. (2014).

The MJO signal has a similar pattern to ENSO but with an opposite sign (e.g. anti-correlation) and weaker correlation (see Figure 10(A2-F2)). The correlation signal between tropopause temperature and MJO appears to prorogate eastward, which is consistent with Virts & Wallace (2014). Zeng et al. (2012) found that cooling of the cold-point tropopause and lowering the tropopause height leads to MJO-related enhanced convection. Over Africa, this negative correlation between MJO and tropopause temperature possibly suggests that the changes in the tropopause temperature likely modulate the MJO-related activities.

Tropopause height/temperature and NAO are negatively/positively correlated in West Africa (see Figure 10(A3-F3)). The correlation between SAM and tropopause parameters indicates a weak coherence among them over Western Africa (see Figure 10(A4-F4)). Stronger correlations with temperature are observed over the South Atlantic and Southern Africa regions, where SAM shows a positive correlation with tropopause temperature (see Figure 10D1 and 10D4). The spatial overlap of SAM and ENSO correlation with tropopause temperature suggests that both SAM and ENSO influences the tropopause over Southern Africa. This finding was also observed by previous studies (i.e., Fogt & Bromwich 2006; L'Heureux & Thompson 2006). Fogt et al. (2011) suggests SAM modulates the strength of ENSO in Southern Hemisphere extratropics.

Over the tropics, QBO and tropopause height/temperature reveal a high negative/positive correlation (see Figure 10(A5-F5)). This signal is consistent with previous studies (Randel et al. 2000; Rieckh et al. 2014). Another weaker signal with opposite polarity is found in the Pacific Ocean west of South Africa. The IOD signal has a very similar pattern to the QBO signal. Over equatorial Africa, it reveals a strong negative correlation with tropopause height (Figure 10(A6-C6)) and a positive correlation with temperature (Figure 10(D6-F6)).



G

Drivers	ENSO		MJO		NAO		SAM		QBO		IOD	
Parameters	H	T	H	T	H	T	H	T	H	T	H	T
Northern Africa	0.4	−0.4	−0.3	0.3			−0.3	0.3				
Western Africa		0.5			−0.4	0.4			−0.4	0.4	−0.3	
Central Africa	0.4		−0.3						−0.5	0.6	−0.3	0.4
Eastern Africa		0.4							−0.4	0.6	−0.4	0.4
Southern Africa	−0.3	0.5	0.3	−0.4				0.4				

Figure 10: Correlation between six global climate drivers (ENSO, MJO NAO, SAM, QBO, and IOD) and tropopause heights (left column) and temperatures (right column) from GNSS-RO (columns A and D), ERA5 (columns B and E), and MERRA-2 (columns C and F) for 2001-2020. The correlations are presented if the region's tropopause variability responds to the climate driver with statistically significant results ($p = 0.05$). The table beneath the figure shows the linkage between global climate drivers and tropopause height (H) and temperature (T) from GNSS-RO data above Africa. These regions are defined in Figure (bottom right).

The interannual variability of the tropopause and its relation to global climate drivers are further studied by decomposing the tropopause temperature and height fields using PCA. PCA derives the most dominant global climate drivers over Africa by finding the greatest amount of variance within the original data set. The oceanic regions are masked to capture only the influence of climate variability on tropopause characteristics over land. The correlation between PC time series and climate indices are only discussed for statistically significant results ($p = 0.05$). The first three dominant EOFs and PCs are presented in Figs. 11 and 12; the remaining are not shown they likely contain more noise than signal. Figure 13 summarizes correlation coefficients and lags where maximum correlation is obtained.

ENSO is the primary driver of tropopause height variability as revealed by the correlation between the first PC time series and the ENSO index. EOF1 accounts for 39%, 47%, and 40% of variance for GNSS-RO, ERA5, and MERRA-2, respectively (Figure 11a-c) and shows a homogeneous positive pattern over Africa. The time-lagged correlation coefficients are -0.49 (at a lag of 6 months), -0.57 (6 months), and -0.47 (6 months) for the three data sets, see Figure 13, indicating higher tropopause height over Africa during El Niño events. The ENSO signal is weaker during tropopause temperature variability. Highest correlation coefficients are found for PC2 with correlations of 0.58 (at a lag of 6 months), 0.53 (6 months), and 0.52 (5 months) for GNSS-RO (28%), ERA5 (29%), and MERRA-2 (30%), indicating lower tropopause temperature over central and eastern Africa as being associated with El Niño, see Figure 12. The spatial pattern is symmetric about the equator and reveals lower tropopause temperature at equatorial latitudes and higher temperature polewards of $\sim$15°S/N.

QBO is another dominant driver of the African tropopause variability. Due to the interaction between ENSO and QBO (Geller et al., 2016), EOF1/PC1 of tropopause height also contain some QBO signal. The correlation between PC1 and the QBO index amounts to 0.41, 0.46, and 0.43 for the three data sets and is statistically significant. Maximum correlation is obtained at a lag of 1 month (see Fig. 13). A stronger signal is found in tropopause temperature with correlation coefficients (time lags of maximum correlation) between the QBO index and the first PC of tropopause temperature being 0.69 (2 months), 0.46 (1 month), and 0.67 (2 months), for GNSS-RO (35% of variance), ERA5 (34%), and MERRA-2 (43%). Corresponding EOFs reveal a negative zonally-symmetric low-latitude signal.

PC1 of tropopause temperature is also correlated with the ENSO and IOD index (see Figs. 12.2 and 13). Coupling of ENSO and IOD was strong in 2016-2017 (indicated by the red arrow in Fig. 12.2). Lim & Hendon (2017) found that the IOD event was a key promoter of the ENSO event.

MJO shows a moderate impact on African tropopause variability, mainly over north-east and southern Africa (Figure 10) and is captured by PC2 in both, tropopause height and tropopause temperature, and by PC3 of tropopause temperature with correlation coefficients of 0.37 (Fig. 13).

SAM is found in the second PC of tropopause temperature from GNSS-RO and ERA5, with a positive correlation coefficient ranging from 0.3 to 0.4, at lags less than 1 month. NAO is captured in PC2 in tropopause height and temperature, and in PC3 of tropopause temperature, with correlation coefficients between 0.3 and 0.5 at a time lags between 0 and 2 months (Fig. 13).

Ranking the influence of global climate drivers on tropopause variability by product of the PCs' variances and the magnitude of the correlation values, yields (1) ENSO, (2) QBO, (3) NAO, (4) MJO, (5) IOD, (6) SAM for the tropopause height, and (1) QBO, (2) ENSO, (3) IOD, (4) NAO, (5) SAM, and (6) MJO for tropopause temperature.

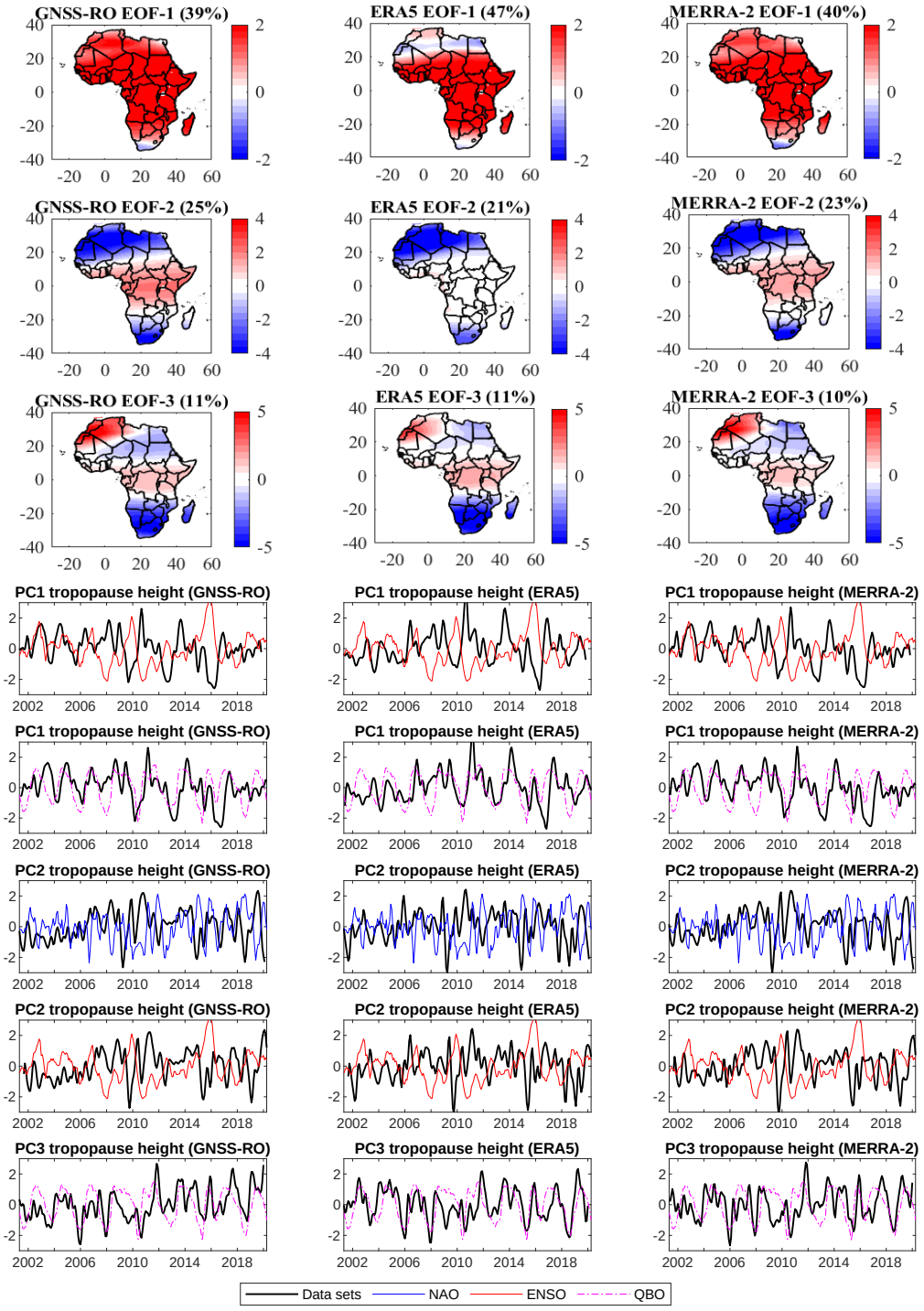


Figure 11: The first three leading EOFs (spatial patterns) of tropopause heights based on GNSS-RO (first column), ERA5 (second column), MERRA-2 (third column), and their corresponding PCs (lower panels), which are plotted together with some representative global climate drivers' indices that have statistically significant correlations ($p < 0.05$). The PCs are plotted with lag-time to match the climate driver indices. Details for correlation coefficients between each PC and global climate drivers indices are shown in Figure 13.

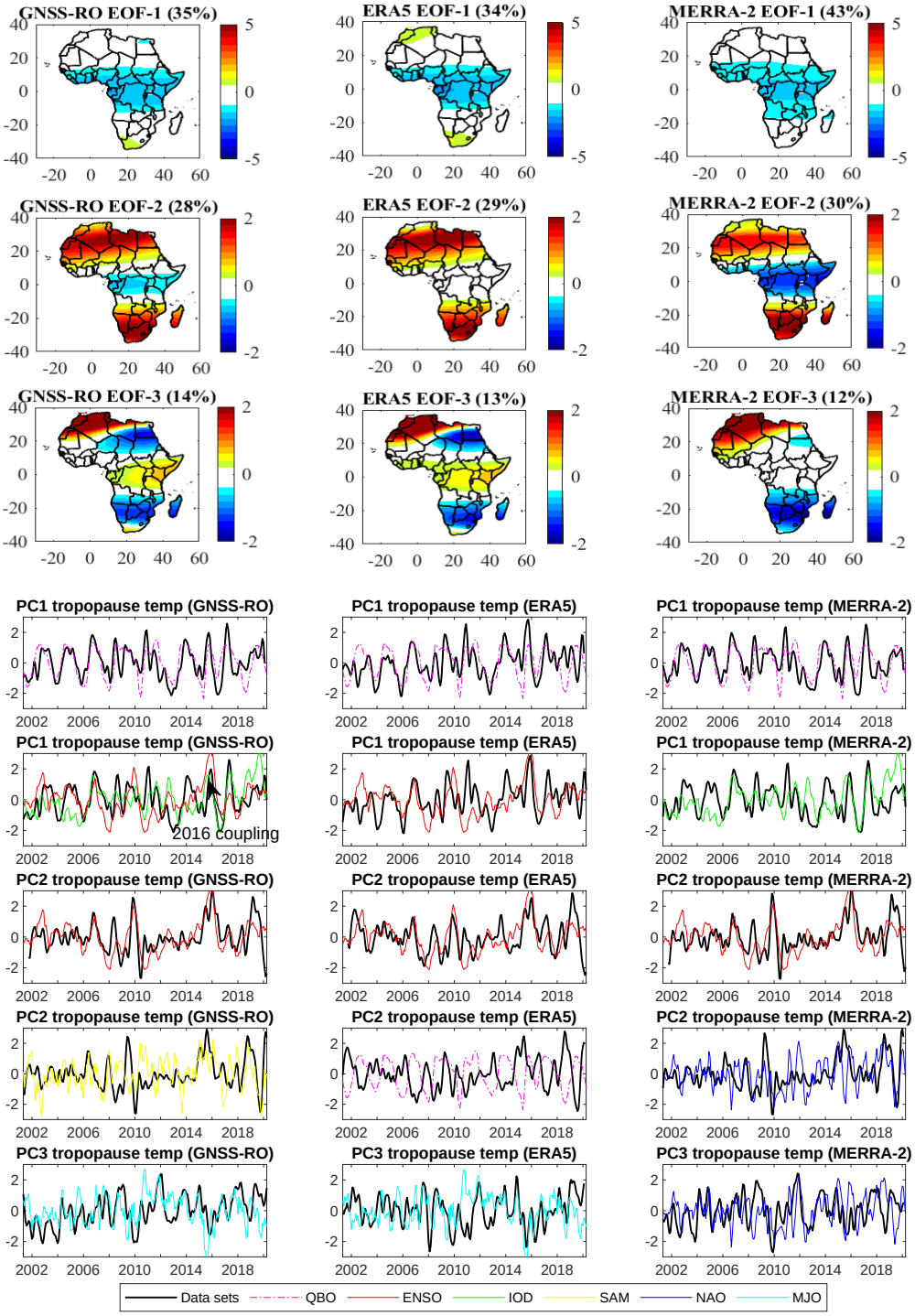


Figure 12: The first three leading EOFs (spatial patterns) of tropopause temperatures based on GNSS-RO (first column), ERA5 (second column), MERRA-2 (third column), and their corresponding PCs (lower panels), which are plotted together with some representative global climate drivers' indices that have statistically significant correlations ($p < 0.05$). Red arrow indicates coupling of ENSO and IOD in 2016-2017. The PCs are plotted with lag-time to match the climate driver indices. Details for correlation coefficients and lag time between each PC and global climate drivers indices are shown in Figure 13.

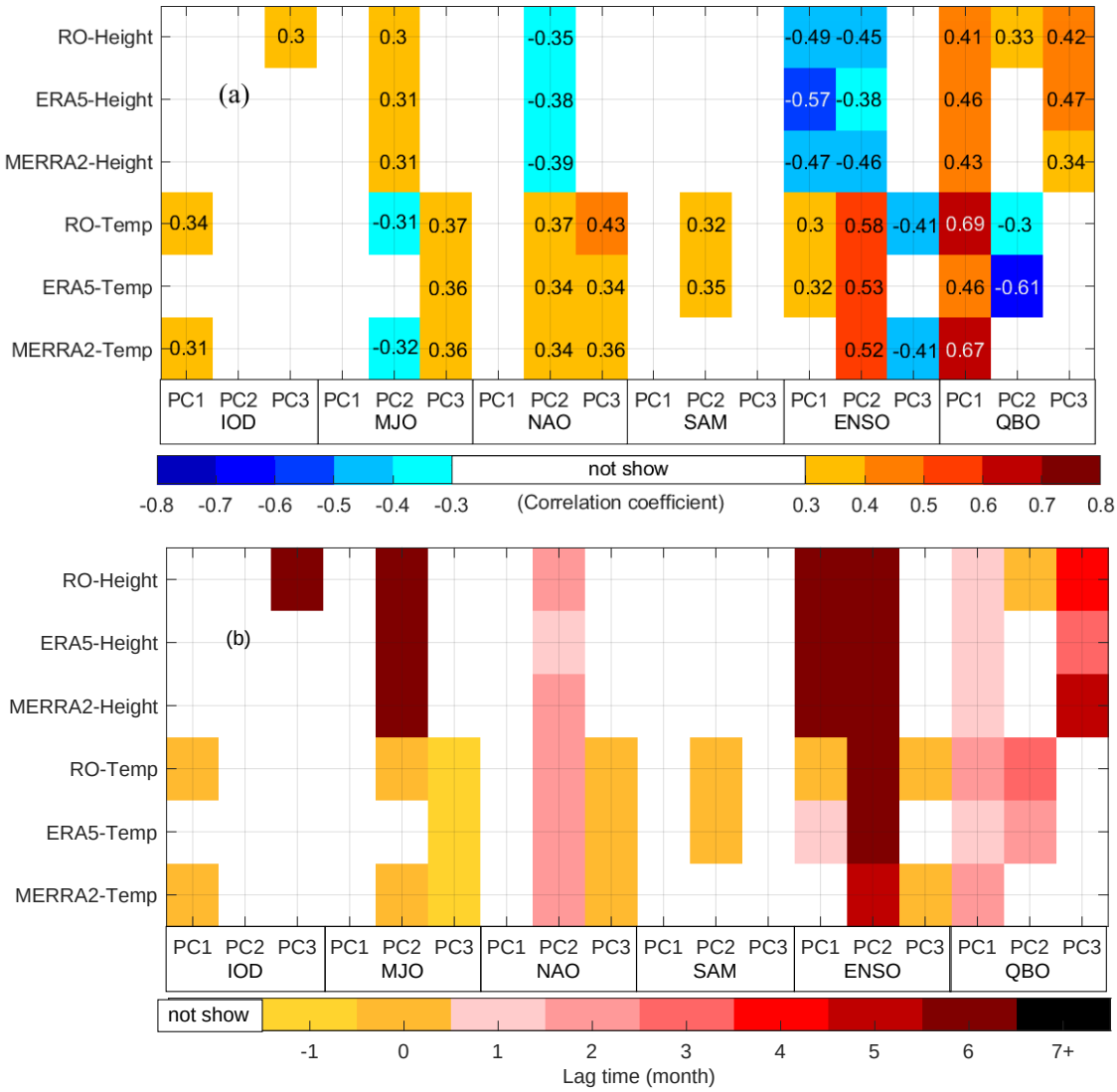


Figure 13: (a) Correlation coefficients between global climate indices and PC time series of tropopause height and temperature from GNSS-RO and reanalysis products. Only results that are statistically significant and having a correlation higher than 0.3 are shown. (b) Lag time (in months) where the maximum correlation is obtained.

3.3.2 Variability in relation to regional climate drivers

We further employ ICA to reveal some of the signals that cannot be detected by the PCA, such as the Tropical Easterly Jet (TEJ), Intertropical Convergence Zone (ITCZ), and Western African Monsoon (WAM), total ozone in Southern Africa. PCA is a pre-step for constructing ICA. The first five PCs (not shown), which represent more than

90% of the total variance in the tropopause variability, are considered to be sufficient to describe the African tropopause variability. Only interpretable independent components (i.e., the spatial and temporal patterns from ICA that match the behavior of regional climate drivers) are shown in Figures 14 and 15. Results are shown only for tropopause temperatures, although similar patterns are obtained for tropopause height. Figure 14 also shows some results from ERA5 and MERRA-2 data for comparison purposes. The spatial patterns are split into positive and negative parts for better interpretability.

A statistically significant correlation of 0.78 - 0.87 ($p = 0.05$) are observed between the ITCZ and tropopause variability based on GNSS-RO (IC-1), ERA5 (IC-2), and MERRA-2 (IC-1), see Figure 14g. The signs of GNSS-RO is flipped to remain consistent with re-analyses. The seasonal amplitude of the ITCZ are captured well. The strong negative signal near West Africa (Figure 14b,d,f) represents the relationship between the ITCZ and tropopause temperature in July, suggesting that the tropopause temperature and ITCZ have a positive relationship, which both are highest in this month as discussed in Lau & Tao (2020). ITCZ is also strongly correlated with the IC-2 derived from GNSS-RO tropopause temperature (r = 0.83; $p = 0.05$, see Figure 14j). The change in the location of ITCZ (see Figure 14h) coincides with the occurrence of the El Niño event. The time series suggests a possible relationship with strong ENSO events. The downward shift in ITCZ's position from 2002–2009, corresponds to a northward shift in ITCZ. In 2009/2010, a strong El Niño shifted ITCZ back to early 2000 conditions. It then resumed its northward shift from 2013–2015 until another major El Niño event happened in late 2015. This signal suggests a cooling tropopause temperature might be associated with a southward displacement of ITCZ. Pausata et al. (2020) also simulated a similar interactive pattern of ENSO-ITCZ based on a climate model in the UTLS and suggested that El Niño induced hemispheric asymmetric cooling of the upper troposphere shifted the ITCZ towards the Southern Hemisphere. This signal is consistent with Broccoli et al. (2006), who suggested that the ITCZ tends to approach the warmer hemisphere during the positive ENSO phase.

Figure 15a-c shows the relationship between the total ozone column and tropopause temperature in Southern Africa. In this study, a positive correlation of 0.80 is observed between the tropopause temperature and Southern African total ozone from 2001 to 2018. Steinbrecht et al. (1998) found that the tropopause height is negatively correlated with total ozone. Because tropopause height goes hand in hand with tropopause temperature,

it is reasonable to suggest that a high tropopause temperature is associated with high total ozone and a low tropopause temperature with low total ozone. Both total ozone values and tropopause temperature are the lowest in January over Southern Africa.

Figure 15d-f reveals that IC5 of tropopause temperature variability and TEJ have a statistically significant correlation of 0.79 ($p = 0.05$). The results are consistent with those of Nicholson & Klotter (2021), who showed that TEJ sustains a maximum strength at 150 hPa during August over the regions of 10°W to 30°E and 5°N to 15°N. Moreover, the temporal patterns indicate a relationship between tropopause temperature and TEJ (Figure 15i). Grist & Nicholson (2001) showed that the tropospheric temperature gradient has linkage with the intensity of TEJ in the latitudinal region higher than 10°N.

The positive and negative spatial patterns and time series representing the annual cycle of WAM is shown in Figures 15g-i. WAM is highly correlated with IC4 of tropopause temperature variability (r = -0.72; $p = 0.05$). Positive increase in the temporal patterns is followed by the evolution of the pre-onset phase (November–March; Figure 15i). A northward migration of the rainy season, which starts in March from approximately 5°N, marks the end of the pre-onset phase corresponding to the positive decrease side in the temporal patterns, which is followed by positive spatial patterns. GNSS-RO observes a strong and explicit signal near Sahel Africa (Figure 15h), and it tends to become more assertive from March until it is fully developed in May. This corresponds to increased convective activities related to ITCZ during the onset phase. Subsequently, the spatial signal transitions to the next phase, which is the monsoon jump, as the heavy rain gradually shifts northward and the rainy season begins in the Sahel region (12.5°N to 20°N) during July and August. This stage indicates that the monsoon is fully developed, corresponding to the negative decrease in the temporal pattern and the positive part of the spatial pattern (Figure 15g). The negative increase over time and transiting to negative spatial pattern, corresponding to the southward migration, mark the second rainfall season in West Africa. The protruding corner (see the northernmost domain in Figure 15h) possibly reveals a sharpened monsoon jump (Figure 15h). An abrupt shift in ITCZ leading to the monsoon jump could be due to an interaction with the topography of the Ahaggar Mountains (Sultan & Janicot, 2003). The negative correlations between WAM and tropopause temperature suggests a cool tropopause temperature is correlated with a strong WAM. Climate models and reanalysis products have revealed these interactions between global

and regional climate drivers but not yet studied through observational records over Africa. Overall, this study shows GNSS-RO can also observe these signals.

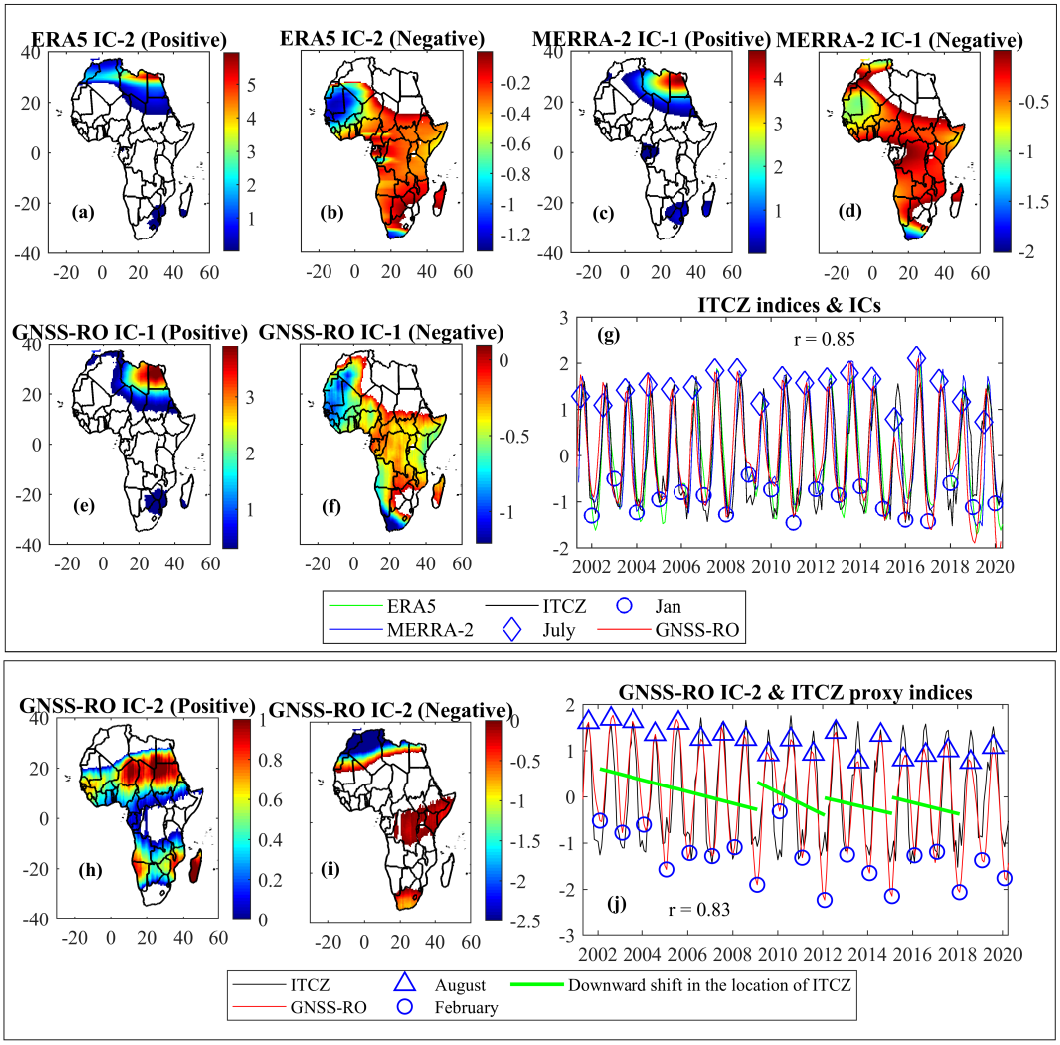


Figure 14: Independent Components derived from GNSS-RO, ERA5, and MERRA-2. Black time-series represent the proxy ITCZ index. Each spatial IC pattern is split into positive and negative parts for a better interpretability. The positive spatial pattern has a negative amplitude and vice versa. Bottom right panel: The correlation coefficients (r) between ICs and regional climate proxy indices are provided in each time series plot. The green line in 14(j) indicate the downward shift in the ITCZ's position in relation to the ENSO episodes. The correlation coefficients are at 95 % confidence level.

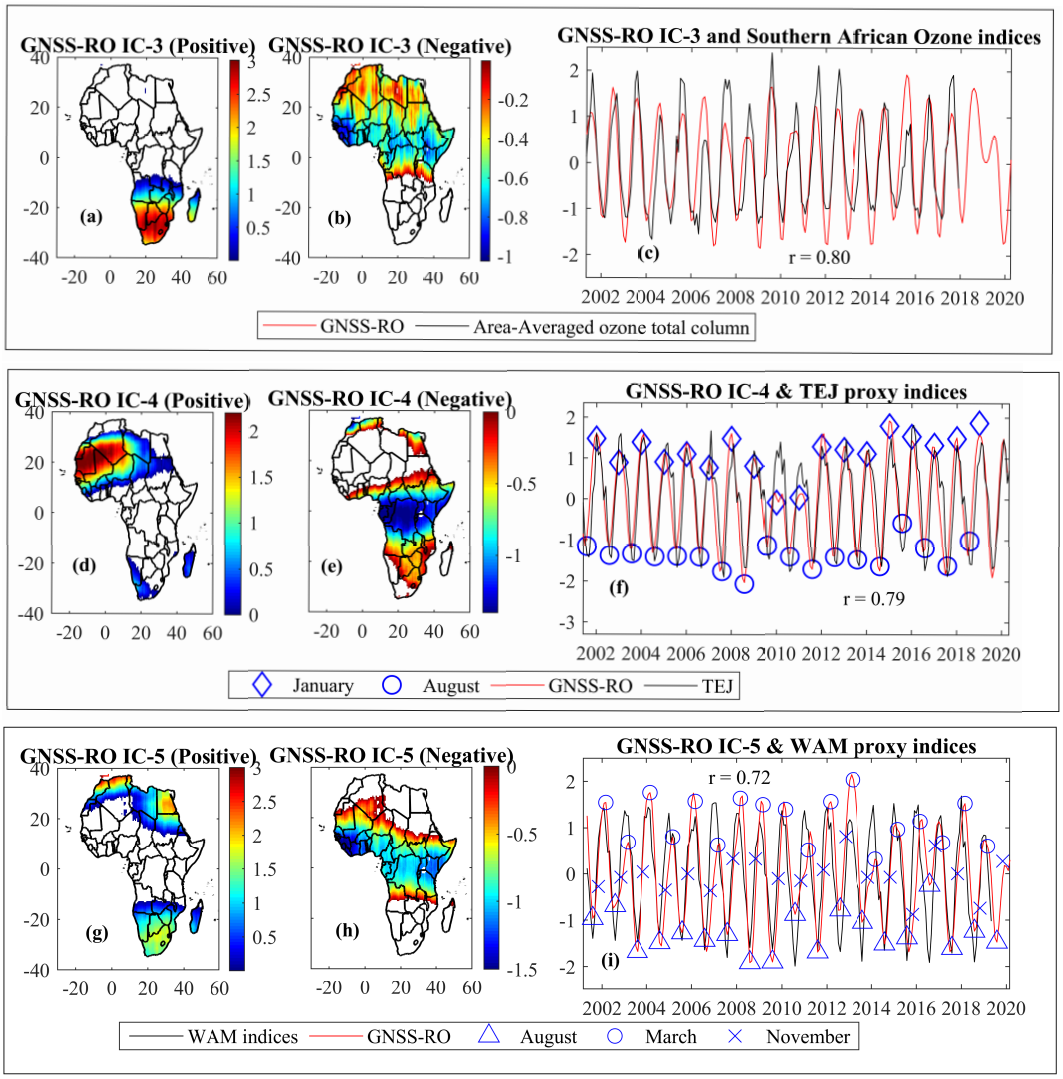


Figure 15: Overview of the spatial and temporal relationship between the African independent tropopause variability patterns derived from ICA of GNSS-RO data and Southern African total ozone column (a-c), Tropical Easterly Jet (TEJ; d-f), and West African Monsoon (WAM; g-i). The correlations are computed at 95% confidence level.

3.3.3 Variability in relation to local drivers

Figure 16 shows patterns of the long-term monthly mean tropopause heights based on GNSS-RO observations which are also revealed from tropopause temperature. The sub-plot in July shows a relationship between tropopause in northeast Africa and Asian summer monsoon. In June-July-August (JJA), tropopause heights tend to be highest in the northeastern part of the domain, which is possibly attributed to the tropopause

structure over the Middle East and the Tibetan Plateau to the Asian summer monsoon convective heating (Highwood & Hoskins, 1998). Tropopause in other seasons and regions reveal smaller differences between the two regions. For instance, the tropopause height in July reveals a asymmetric tropical tropopause height between the East African coast and the Iranian Plateau-Arabian Sea being highest over the Northeast African coast and lowest over the Western Indian ocean. Such a unique pattern is also observed from MERRA-2 and ERA5 and consistently occurs during June–August (most pronounced in July) and corresponds to the interannual variation of Asian summer monsoon (Highwood & Hoskins, 1998; Randel & Park, 2006). Tropopause in other seasons and regions reveal smaller differences between the two regions. In addition, the tropopause height tends to show little north-south gradient across the continent in other months and seasons, but tends to be more uniformly distributed, unlike the pattern in June-August, when it is skewed to the Northeast Africa coast. This pattern could be attributed to two drivers. First, June-August period coincides with the peak of the Asian Summer Monsoon Anticyclone within the UTLS. This tends to influence horizontal mass distribution stretching from west Africa coast to the Pacific (as also demonstrated in Basha et al. 2021). Second, during northern summer, the Asian monsoon interacts with the Tropical Easterly Jet(TEJ), with TEJ potentially inducing additional circulation and convection over south Asia (due to warming over eastern Indian Ocean), while suppressing vertical flow over equatorial Africa (due to cooling over western Indian Ocean).

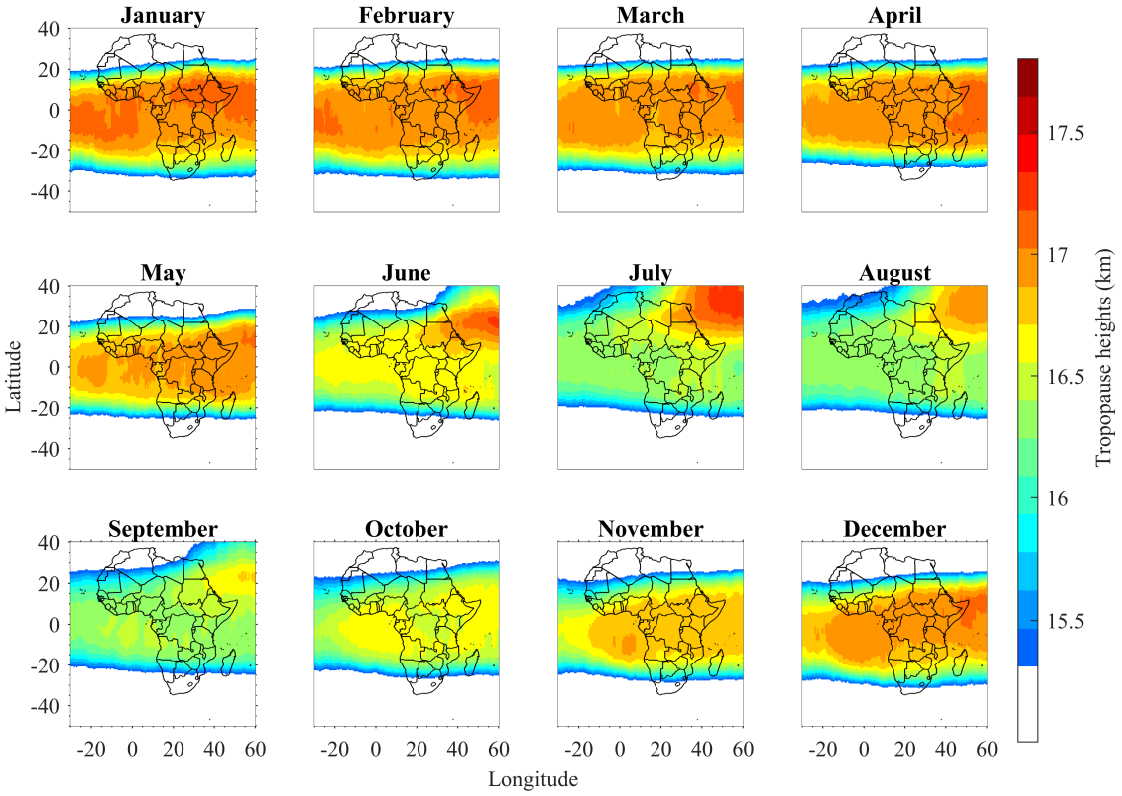


Figure 16: Monthly long-term means for tropopause height (January to December) based on GNSS-RO observations.

4 Summary and conclusions

This study addressed the radiosonde data gap in Africa and demonstrated that can be infilled with GNSS-RO data from various missions, including CHAMP, GRACE, COSMIC, COSMIC-2, and MetOp-A for the time period of 2001–2020. Altogether, GNSS-RO provides more than 1.5 million accurate, consistent, high-vertical resolution atmospheric profiles over the African continent, which has typically had deficient, inconsistent, or poor-quality upper-air measurements. Over recent decades, the African radiosonde network has been shrinking at an estimated rate of 50 stations/decade and 300,000 observations/decade.Today, only 50 stations remain active. Overall, 50% of the African radiosonde data are usable, while the rest are either problematic or missing. A statistical comparison and validation between radiosonde measurements and data from various RO missions revealed good agreement with differences being smaller than 1 K across the UTLS region (10–25 km). Furthermore, this study revealed potential issues with using reanal-

ysis products to infill the radiosonde gap over Africa, i.e., (i) are rather smooth profiles with only little vertical variability, limiting their ability to detect tropopause inversion layers, (ii) they sometimes underestimate the seasonal variations and suffer from temporal inhomogeneity of tropopause heights and (iii), tropopause temperatures/heights of ERA5 and MERRA-2 are biased at low latitudes by 0.8 K/0.3 km and 1.2 K/0.5 km, respectively. Statistical comparisons between radiosondes, GNSS-RO, and reanalysis products demonstrated best agreement between radiosonde measurements and GNSS-RO have the lowest RMSE and the highest KGE value. GNSS-RO's data availability only began in the early 2000s, limiting its use for long-term climate applications. Nevertheless, this study showed that GNSS-RO outperforms reanalysis products in addressing the radiosonde data gap over Africa.

The analysis of GNSS-RO data also revealed a better understanding of tropopause variability over Africa for the period 2001-2020 in relation to global, regional climate, and local drivers. The following results were obtained:

- Statistically significant signals of ENSO, QBO, NAO, MJO, IOD, and SAM were found in de-seasonalized time series of tropopause temperatures and heights based on correlation analysis as well as PCA. The influence of these global climate drivers on African tropopause variability was ranked as: (1) ENSO, (2) QBO, (3) NAO, (4) MJO, (5) IOD and (6) SAM for tropopause height, and (1) QBO, (2) ENSO, (3) IOD, (4)NAO, (5) SAM, and (6) MJO for tropopause temperature. Several coupled global climate drivers, e.g., ENSO-IOD, ENSO-QBO, ENSO-NAO, QBO-IOD, ENSO-MJO, and ENSO-SAM-MJO were also detected.
- Statistically significant signals of ITCZ, TEJ, WAM, and ozone variation in the Southern Africa were detected using ICA. GNSS-RO observations suggest that the tropopause height increases and temperature decreases due to global natural variability associated with El Niño events, which cause hemispheric asymmetry in the upper tropospheric temperature in the tropics, which cause the ITCZ to move southward.
- For the local driver, the relationship between local air-sea interaction and tropopause variability has been studied. During July, the Asian summer monsoon is sufficient to affect the zonal symmetry behaviour and mean latitudinal gradients of the tropical tropopause in Africa.

This study has therefore pioneered the use of GNSS-RO to investigate the African atmospheric climate change. Future studies will investigate the complex interaction of various climate drivers and their roles in the changing climate. While many climate drivers have shown their footprint in the tropopause layer, their contribution and interactive mechanism to atmospheric climate change are still not well understood in this study. Reanalyses assembles various data sets, which to some extent compromise their accuracies due to data fusion. On the flip side, the GNSS-RO relies on a single data source, making it a powerful tool for studying atmospheric climate change behavior. With more publicly available satellites, the GNSS-RO may, in future, play a significant role in estimating regional warming, monitoring atmospheric climate change, and regional climate forecasts. In particular for Africa, the GNSS-RO will be a primary data source for more than 80% of countries.

5 Acknowledgment

The authors thank the NOAA for kindly providing Integrated Global Radiosonde Archive (IGRA) data. Thanks to NOAA, US Air Force (USAF), NSPO, UCAR for making available the COSMIC-2 data. The authors appreciate Global Modeling and Assimilation Office (GMAO) and ECWMF for providing MERRA-2 and ERA5 data.

6 Data Availability Statement

All GNSS-RO data used during this study are openly available from the COSMIC Data Analysis and Archive Center at `https://doi.org/10.5065/qan9-we09` (CHAMP), `https://doi.org/10.5065/chk8-fx07` (GRACE), `https://doi.org/10.5065/789w-m137` (MetOp-A), `https://doi.org/10.5065/ZD80-KD74` (COSMIC-), and `https://doi.org/10.5065/t353-c093` (COSMIC-2). All MERRA-2 data are available at `https://doi.org/10.5067/WWQSXQ8IVFW8`. All ERA5 data are available at `https://apps.ecmwf.int/data-catalogues/era5/?class=ea` (Hersbach et al., 2020). All radiosonde station information provided by WMO are available at `https://www.wmo.int/datastat`. All radiosonde data are openly available at `https://doi.org/10.7289/V5X63K0Q`.

References

Agutu, N., Awange, J., Zerihun, A., Ndehedehe, C., Kuhn, M., & Fukuda, Y. (2017). Assessing multi-satellite remote sensing, reanalysis, and land surface models' products in characterizing agricultural drought in East Africa. *Remote Sensing of Environment*, *194*, 287–302. doi: 10.1016/j.rse.2017.03.041

Andersson, J. C., Arheimer, B., Traoré, F., Gustafsson, D., & Ali, A. (2017). Process refinements improve a hydrological model concept applied to the Niger River basin. *Hydrological Processes*, *31*(25), 4540–4554. doi: 10.1002/hyp.11376

Angerer, B., Ladstädter, F., Scherllin-Pirscher, B., Schwärz, M., Steiner, A. K., Foelsche, U., & Kirchengast, G. (2017). Quality aspects of the Wegener Center multi-satellite GPS radio occultation record OPSv5.6. *Atmospheric Measurement Techniques*, *10*(12), 4845–4863. doi: 10.5194/amt-10-4845-2017

Anthes, R. A., & Schreiner, W. S. (2019). Six new satellites watch the atmosphere over Earth's equator. *EOS*, *100*. Retrieved from `https://doi.org/10.1029/2019EO131779` doi: 10.1029/2019EO131779

Anyah, R., Forootan, E., Awange, J. L., & Khaki, M. (2018). Understanding linkages between global climate indices and terrestrial water storage changes over Africa using GRACE products. *Science of the Total Environment*, *635*, 1405–1416. doi: 10.1016/j.scitotenv.2018.04.159

Awange, J. (2018). *GNSS environmental sensing*. Heidelberg: Springer.

Awange, J. (2021). *Lake Victoria Monitored from Space*. Springer Nature.

Awange, J. L. (2012). *Environmental monitoring using GNSS: Global navigation satellite systems*. Springer International Publishers.

Awange, J. L., Anyah, R., Agola, N., Forootan, E., & Omondi, P. (2013). Potential impacts of climate and environmental change on the stored water of Lake Victoria Basin and economic implications. *Water Resources Research*, *49*(12), 8160–8173. doi: 10.1002/2013WR014350

Awange, J. L., Ferreira, V. G., Forootan, E., Andam-Akorful, S., Agutu, N. O., & He, X. (2016). Uncertainties in remotely sensed precipitation data over Africa. *International Journal of Climatology*, *36*(1), 303–323. doi: 10.1002/joc.4346

Baldwin, M., Gray, L., Dunkerton, T., Hamilton, K., Haynes, P., Randel, W., ... others (2001). The Quasi-Biennial Oscillation. *Reviews of Geophysics*, *39*(2), 179–229. doi: doi.org/10.1029/1999RG000073

Basha, G., Kishore, P., Ratnam, M. V., Ouarda, T. B., Velicogna, I., & Sutterley, T. (2015). Vertical and latitudinal variation of the intertropical convergence zone derived using GPS radio occultation measurements. *Remote Sensing of Environment*, *163*, 262–269. doi: 10.1016/j.rse.2015.03.024

Basha, G., Ratnam, M. V., Jiang, J. H., Kishore, P., & Ravindra Babu, S. (2021). Influence of Indian Summer Monsoon on Tropopause, Trace Gases and Aerosols in Asian Summer Monsoon Anticyclone Observed by COSMIC, MLS and CALIPSO. *Remote Sensing*, *13*(17), 3486. doi: 10.3390/rs13173486

Broccoli, A. J., Dahl, K. A., & Stouffer, R. J. (2006). Response of the ITCZ to Northern Hemisphere cooling. *Geophysical Research Letters*, *33*(1). doi: 10.1029/2005GL024546

Chen, G., Iwasaki, T., Qin, H., & Sha, W. (2014). Evaluation of the warm-season diurnal variability over East Asia in recent reanalyses JRA-55, ERA-Interim, NCEP CFSR, and NASA MERRA. *Journal of Climate*, *27*(14), 5517–5537. doi: 10.1175/JCLI-D-14-00005.1

Comon, P. (1994). Independent component analysis, a new concept? *Signal Processing*. doi: 10.1016/0165-1684(94)90029-9

Copernicus Climate Change Service. (2018). *Ozone monthly gridded data from 1970 to present derived from satellite observations.* ECMWF. Retrieved from `https://cds.climate.copernicus.eu/cdsapp#!/dataset/10.24381/cds.4ebfe4eb?tab=overview` doi: 10.24381/cds.4ebfe4eb

Durre, I., Vose, R. S., & Wuertz, D. B. (2006). Overview of the integrated global radiosonde archive. *Journal of Climate*, *19*(1), 53–68. doi: 10.1175/JCLI3594.1

Durre, I., Yin, X., Vose, R. S., Applequist, S., & Arnfield, J. (2018). Enhancing the data coverage in the Integrated Global Radiosonde Archive. *Journal of Atmospheric and Oceanic Technology*, *35*(9), 1753–1770. doi: 10.1175/JTECH-D-17-0223.1

D'Errico, J. (2004). Inpaint nans. *MATLAB Central File Exchange*. Retrieved from `\url{https://www.mathworks.com/matlabcentral/fileexchange/4551-inpaint_nans}`

Feng, S., Fu, Y., & Xiao, Q. (2012). Trends in the global tropopause thickness revealed by radiosondes. *Geophysical Research Letters*, *39*(20). doi: 10.1029/2012GL053460

Fink, A. H., Agustí-Panareda, A., Parker, D. J., Lafore, J.-P., Ngamini, J.-B., Afiesimama, E., ... others (2011). Operational meteorology in West Africa: observational networks, weather analysis and forecasting. *Atmospheric Science Letters*, *12*(1), 135–141. doi: 10.1002/asl.324

Foelsche, U., Borsche, M., Steiner, A. K., Gobiet, A., Pirscher, B., Kirchengast, G., ... Schmidt, T. (2008). Observing upper troposphere–lower stratosphere climate with radio occultation data from the CHAMP satellite. *Climate Dynamics*, *31*(1), 49–65. doi: 10.1007/s00382-007-0337-7

Foelsche, U., Scherllin-Pirscher, B., Ladstädter, F., Steiner, A. K., & Kirchengast, G. (2011). Refractivity and temperature climate records from multiple radio occultation satellites consistent within 0.05%. *Atmospheric Measurement Techniques*, *4*, 2007–2018. doi: 10.5194/amt-4-2007-2011

Fogt, R. L., & Bromwich, D. H. (2006). Decadal variability of the ENSO teleconnection to the high-latitude South Pacific governed by coupling with the southern annular mode. *Journal of Climate*, *19*(6), 979–997. doi: 10.1175/JCLI3671.1

Fogt, R. L., Bromwich, D. H., & Hines, K. M. (2011). Understanding the SAM influence on the South Pacific ENSO teleconnection. *Climate Dynamics*, *36*(7), 1555–1576. doi: 10.1007/s00382-010-0905-0

Forootan, E., Awange, J. L., Kusche, J., Heck, B., & Eicker, A. (2012). Independent patterns of water mass anomalies over Australia from satellite data and models. *Remote Sensing of Environment*, *124*, 427–443. doi: 10.1016/j.rse.2012.05.023

Forootan, E., Rietbroek, R., Kusche, J., Sharifi, M. A., Awange, J. L., Schmidt, M., ... Famiglietti, J. (2014). Separation of large scale water storage patterns over iran using grace, altimetry and hydrological data. *Remote Sensing of Environment*, *140*, 580–595. doi: 10.1016/j.rse.2013.09.025

Franzke, C. L., Barbosa, S., Blender, R., Fredriksen, H.-B., Laepple, T., Lambert, F., ... others (2020). The structure of climate variability across scales. *Reviews of Geophysics*, *58*(2), e2019RG000657. doi: 10.1029/2019RG000657

Fujiwara, M., Hibino, T., Mehta, S. K., Gray, L., Mitchell, D., & Anstey, J. (2015). Global temperature response to the major volcanic eruptions in multiple reanalysis data sets. *Atmospheric Chemistry and Physics*, *15*(23), 13507–13518. doi: 10.5194/acp-15-13507-2015

Gelaro, R., McCarty, W., Suárez, M. J., Todling, R., Molod, A., Takacs, L., ...

others (2017). The Modern-Era Retrospective Analysis for Research and Applications, Version 2 (MERRA-2). *Journal of Climate*, *30*(14), 5419–5454. doi: 10.1175/JCLI-D-16-0758.1

Geller, M. A., Zhou, T., & Yuan, W. (2016). The QBO, gravity waves forced by tropical convection. *Journal of Geophysical Research: Atmospheres*, *121*, 8886–8895. doi: 10.1002/2015JD024125

Gregory, P. J., Ingram, J. S., & Brklacich, M. (2005). Climate change and food security. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *360*(1463), 2139–2148. doi: 10.1098/rstb.2005.1745

Grist, J. P., & Nicholson, S. E. (2001). A study of the dynamic factors influencing the rainfall variability in the West African Sahel. *Journal of Climate*, *14*(7), 1337–1359. doi: 10.1175/1520-0442(2001)014⟨1337:ASOTDF⟩2.0.CO;2

Gupta, H. V., Kling, H., Yilmaz, K. K., & Martinez, G. F. (2009). Decomposition of the mean squared error and NSE performance criteria: Implications for improving hydrological modelling. *Journal of Hydrology*, *377*(1-2), 80–91. doi: 10.1016/j.jhydrol.2009.08.003

Hansen, J., Sato, M., Glascoe, J., & Ruedy, R. (1998). A common-sense climate index: Is climate changing noticeably? *Proceedings of the National Academy of Sciences*, *95*(8), 4113–4120. doi: 10.1073/pnas.95.8.4113

Hersbach, H., Bell, B., Berrisford, P., Hirahara, S., Horányi, A., Muñoz-Sabater, J., ... others (2020). The ERA5 global reanalysis. *Quarterly Journal of the Royal Meteorological Society*, *146*(730), 1999–2049. doi: 10.1002/qj.3803

Highwood, E., & Hoskins, B. (1998). The tropical tropopause. *Quarterly Journal of the Royal Meteorological Society*, *124*(549), 1579–1604. doi: 10.1002/qj.49712454911

Ho, S.-P., Anthes, R. A., Ao, C. O., Healy, S., Horanyi, A., Hunt, D., ... others (2020). The COSMIC/FORMOSAT-3 radio occultation mission after 12 years: accomplishments, remaining challenges, and potential impacts of COSMIC-2. *Bulletin of the American Meteorological Society*, *101*(7), E1107–E1136. doi: 10.1175/BAMS-D-18-0290.1

Homeyer, C. R., Bowman, K. P., & Pan, L. L. (2010). Extratropical tropopause transition layer characteristics from high-resolution sounding data. *Journal of Geophysical Research: Atmospheres*, *115*(D13). doi: 10.1029/2009JD013664

Hsu, K.-l., Gao, X., Sorooshian, S., & Gupta, H. V. (1997). Precipitation estimation from remotely sensed information using artificial neural networks. *Journal of Applied Meteorology*, *36*(9), 1176–1190. doi: 10.1175/1520-0450(1997)036⟨1176:PEFRSI⟩2.0.CO;2

Hurrell, J. W., & Deser, C. (2010). North Atlantic climate variability: the role of the North Atlantic Oscillation. *Journal of Marine Systems*, *79*(3-4), 231–244. doi: 10.1016/j.jmarsys.2008.11.026

Hurrell, J. W., Kushnir, Y., Ottersen, G., & Visbeck, M. (2003). An overview of the North Atlantic Oscillation. *Geophysical Monograph-American Geophysical Union*, *134*, 1–36. doi: 10.1029/134GM01

Hyvärinen, A., & Oja, E. (2000). Independent Component Analysis: algorithms and applications. *Neural Networks*, *13*(4-5), 411–430. doi: 10.1016/S0893-6080(00)00026-5

Inai, Y., Shiotani, M., Fujiwara, M., Hasebe, F., & Vömel, H. (2015). Altitude misestimation caused by the Vaisala RS80 pressure bias and its impact on meteorological profiles. *Atmospheric Measurement Techniques*, *8*(10), 4043–4054. doi: 10.5194/amt-8-4043-2015

IPCC. (2014). *Impacts of 1.5° C global warming on natural and human systems* (Tech. Rep.). The Intergovernmental Panel on Climate Change.

Isioye, O. A., Combrinck, L., Botai, J. O., & Munghemezulu, C. (2015). The potential for observing African weather with GNSS remote sensing. *Advances in Meteorology*, *2015*. doi: 10.1155/2015/723071

Kalnay, E., Kanamitsu, M., Kistler, R., Collins, W., Deaven, D., Gandin, L., ... others (1996). The NCEP/NCAR 40-year reanalysis project. *Bulletin of the American meteorological Society*, *77*(3), 437–472. doi: 10.1175/1520-0477(1996)077⟨0437:TNYRP⟩2.0.CO;2

Khandu, Awange, J., Forootan, E., et al. (2016). Interannual variability of upper tropospheric and lower stratospheric (UTLS) region over Ganges-Brahmaputra-Meghna basin based on COSMIC GNSS-RO data. *Atmospheric Measurement Techniques*, *8*(9). doi: 10.5194/amtd-8-9399-2015

Khandu, Awange, J., Wickert, J., Schmidt, T., Sharifi, M., Heck, B., ... others (2011). GNSS remote sensing of the Australian tropopause. *Climatic Change*, *105*(3-4), 597–618. doi: 10.1007/s10584-010-9894-6

Kubokawa, H., Satoh, M., Suzuki, J., & Fujiwara, M. (2016). Influence of topography on temperature variations in the tropical tropopause layer. *Journal of Geophysical Research: Atmospheres*, *121*(19), 11–556. doi: 10.1002/2016JD025569

Kursinski, E., Hajj, G., Schofield, J., Linfield, R., & Hardy, K. R. (1997). Observing Earth's atmosphere with radio occultation measurements using the Global Positioning System. *Journal of Geophysical Research: Atmospheres*, *102*(D19), 23429–23465. doi: 10.1029/97JD01569

Lanzante, J. R., Klein, S. A., & Seidel, D. J. (2003). Temporal homogenization of monthly radiosonde temperature data. Part I: Methodology. *Journal of Climate*, *16*(2), 224–240. doi: 10.1175/1520-0442(2003)016⟨0224:THOMRT⟩2.0.CO;2

Lau, W. K., & Tao, W. (2020). Precipitation-radiation-circulation feedback processes associated with structural changes of the ITCZ in a warming climate during 1980–2014: An observational portrayal. *Journal of Climate*, *33*(20), 8737–8749. doi: 10.1175/JCLI-D-20-0068.1

Lien, G.-Y., Lin, C.-H., Huang, Z.-M., Teng, W.-H., Chen, J.-H., Lin, C.-C., ... others (2021). Assimilation impact of early FORMOSAT-7/COSMIC-2 GNSS Radio Occultation data with Taiwan's CWB Global Forecast System. *Monthly Weather Review*. doi: 10.1175/MWR-D-20-0267.1

Lim, E.-P., & Hendon, H. H. (2017). Causes and predictability of the negative Indian Ocean Dipole and its impact on La Niña during 2016. *Scientific Reports*, *7*(1), 1–11. doi: 10.1038/s41598-017-12674-z

L'Heureux, M. L., & Thompson, D. W. (2006). Observed relationships between the El Niño–Southern Oscillation and the extratropical zonal-mean circulation. *Journal of Climate*, *19*(2), 276–287. doi: 10.1175/JCLI3617.1

Marshall, G. J. (2003). Trends in the Southern Annular Mode from observations and reanalyses. *Journal of Climate*, *16*(24), 4134–4143. doi: 10.1175/1520-0442(2003)016⟨4134:TITSAM⟩2.0.CO;2

Meng, L., Liu, J., Tarasick, D. W., & Li, Y. (2021). Biases of global tropopause altitude products in reanalyses and implications for estimates of tropospheric column ozone. *Atmosphere*, *12*(4), 417. doi: 10.3390/atmos12040417

Nascimento, A., Awange, J., Gonçalves, R., et al. (2020). South America's tropopause variability in relation to global teleconnection (2001–2017): A GNSS-radio occultation assessment. *Journal of Atmospheric and Solar-Terrestrial*

Physics, *209*, 105379. doi: 10.1016/j.jastp.2020.105379

Nicholson, S. E., & Klotter, D. (2021). The Tropical Easterly Jet over Africa, its representation in six reanalysis products, and its association with Sahel rainfall. *International Journal of Climatology*, *41*(1), 328–347. doi: doi.org/10.1002/joc.6623

Parker, D. J., Fink, A., Janicot, S., Ngamini, J.-B., Douglas, M., Afiesimama, E., ... others (2008). The AMMA radiosonde program and its implications for the future of atmospheric monitoring over Africa. *Bulletin of the American Meteorological Society*, *89*(7), 1015–1028. doi: 10.1175/2008BAMS2436.1

Pausata, F. S., Zanchettin, D., Karamperidou, C., Caballero, R., & Battisti, D. S. (2020). ITCZ shift and extratropical teleconnections drive ENSO response to volcanic eruptions. *Science Advances*, *6*(23), eaaz5006. doi: 10.1126/sciadv.aaz5006

Peuch, V.-H. (2020). What are the atmospheric observation data gaps and what should WMO do to close them? In *WMO Data Conference: Exchange of Earth System Data in the 21st Century*. ECMWF. (16-11-2020/19-11-2020)

Pirscher, B., Foelsche, U., Lackner, B., & Kirchengast, G. (2007). Local time influence in single-satellite radio occultation climatologies from Sun-synchronous and non-Sun-synchronous satellites. *Journal of Geophysical Research: Atmospheres*, *112*(D11). doi: 10.1029/2006JD007934

Preisendorfer, R. W., & Mobley, C. D. (1988). Principal Component Analysis in meteorology and oceanography. *Bulletin of the American Meteorological Society*, *17*.

Ramella Pralungo, L., Haimberger, L., Stickler, A., & Brönnimann, S. (2014). A global radiosonde and tracked balloon archive on 16 pressure levels (GRASP) back to 1905–Part 1: Merging and interpolation to 00: 00 and 12: 00 GMT. *Earth System Science Data*, *6*(1), 185–200. doi: 10.5194/essd-6-185-2014

Randel, W. J., & Park, M. (2006). Deep convective influence on the Asian summer monsoon anticyclone and associated tracer variability observed with Atmospheric Infrared Sounder (AIRS). *Journal of Geophysical Research: Atmospheres*, *111*(D12). doi: 10.1029/2005JD006490

Randel, W. J., Wu, F., & Gaffen, D. J. (2000). Interannual variability of the tropical tropopause derived from radiosonde data and NCEP reanalyses. *Journal of Geophysical Research: Atmospheres*, *105*(D12), 15509–15523. doi: 10.1029/2000JD900155

Rieckh, T., Scherllin-Pirscher, B., Ladstädter, F., & Foelsche, U. (2014). Characteristics of tropopause parameters as observed with GPS radio occultation. *Atmospheric Measurement Techniques*, *7*(11), 3947–3958. doi: 10.5194/amt-7-3947-2014

Saji, N., Goswami, B., Vinayachandran, P., & Yamagata, T. (1999). A dipole mode in the tropical Indian Ocean. *Nature*, *401*(6751), 360–363. doi: 10.1038/43854

Santer, B. D., Sausen, R., Wigley, T., Boyle, J. S., AchutaRao, K., Doutriaux, C., ... others (2003). Behavior of tropopause height and atmospheric temperature in models, reanalyses, and observations: Decadal changes. *Journal of Geophysical Research: Atmospheres*, *108*(D1), ACL–1. doi: 10.1029/2002JD002258

Santer, B. D., Wehner, M. F., Wigley, T., Sausen, R., Meehl, G., Taylor, K., ... others (2003). Contributions of anthropogenic and natural forcing to recent tropopause height changes. *Science*, *301*(5632), 479–483. doi: 10.1126/science.1084123

Scherllin-Pirscher, B., Deser, C., Ho, S.-P., Chou, C., Randel, W., & Kuo, Y.-H. (2012). The vertical and spatial structure of ENSO in the upper troposphere and lower stratosphere from GPS radio occultation measurements. *Geophysical Research Letters*, *39*(20). doi: 10.1029/2012GL053071

Scherllin-Pirscher, B., Kirchengast, G., Steiner, A. K., Kuo, Y.-H., & Foelsche, U. (2011). Quantifying uncertainty in climatological fields from GPS radio occultation: an empirical-analytical error model. *Atmospheric Measurement Techniques*, *4*(9), 2019–2034. Retrieved from `https://amt.copernicus.org/articles/4/2019/2011/` doi: 10.5194/amt-4-2019-2011

Scherllin-Pirscher, B., Steiner, A. K., Anthes, R. A., Alexander, M. J., Alexander, S. P., Biondi, R., ... others (2021). Tropical temperature variability in the UTLS: New insights from GPS radio occultation observations. *Journal of Climate*, *34*(8), 2813–2838. doi: 10.1175/JCLI-D-20-0385.1

Schreiner, W., Weiss, J., Anthes, R., Braun, J., Chu, V., Fong, J., ... Zeng, Z. (2020). COSMIC-2 Radio Occultation Constellation: First Results. *Geophysical Research Letters*, *47*(4), e2019GL086841. Retrieved from `https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2019GL086841` doi: https://doi.org/10.1029/2019GL086841

Seidel, D. J., Angell, J., Christy, J., Free, M., Klein, S., Lanzante, J. R., ... others

(2004). Uncertainty in signals of large-scale climate variations in radiosonde and satellite upper-air temperature datasets. *Journal of Climate*, *17*(11), 2225–2240. doi: 10.1175/1520-0442(2004)017⟨2225:UISOLC⟩2.0.CO;2

Shangguan, M., Wang, W., & Jin, S. (2019). Variability of temperature and ozone in the upper troposphere and lower stratosphere from multi-satellite observations and reanalysis data. *Atmospheric Chemistry and Physics*, *19*(10), 6659–6679. doi: 10.5194/acp-19-6659-2019

Simmons, A., Soci, C., Nicolas, J., Bell, B., Berrisford, P., Dragani, R., ... others (2020). Global stratospheric temperature bias and other stratospheric aspects of ERA5 and ERA5. 1. *ECMWF*. doi: 10.21957/rcxqfmg0

Steinbrecht, W., Claude, H., Köhler, U., & Hoinka, K. (1998). Correlations between tropopause height and total ozone: Implications for long-term changes. *Journal of Geophysical Research: Atmospheres*, *103*(D15), 19183–19192. doi: 10.1029/98JD01929

Sultan, B., & Janicot, S. (2003). The West African monsoon dynamics. Part II: The "preonset" and "onset" of the summer monsoon. *Journal of Climate*, *16*(21), 3407–3427. doi: 10.1175/1520-0442(2003)016⟨3407:TWAMDP⟩2.0.CO;2

Taalas, P., Steiner, A., & Andersen, I. (2021, 10). *COP26: How plugging data gaps will transform our response to climate change.* Retrieved from `https://www.scmp.com/` (`https://www.scmp.com/` (visited: 2021-12-10))

Tegtmeier, S., Anstey, J., Davis, S., Dragani, R., Harada, Y., Ivanciu, I., ... others (2020). Temperature and tropopause characteristics from reanalyses data in the tropical tropopause layer. *Atmospheric Chemistry and Physics*, *20*(2), 753–770. doi: 10.5194/acp-20-753-2020

Terink, W., Immerzeel, W. W., & Droogers, P. (2013). Climate change projections of precipitation and reference evapotranspiration for the Middle East and Northern Africa until 2050. *International Journal of Climatology*, *33*(14), 3055–3072. doi: 10.1002/joc.3650

Thompson, A. M., Balashov, N. V., Witte, J., Coetzee, J., Thouret, V., & Posny, F. (2014). Tropospheric ozone increases over the southern Africa region: bellwether for rapid growth in Southern Hemisphere pollution? *Atmospheric Chemistry and Physics*, *14*(18), 9855–9869. doi: 10.5194/acp-14-9855-2014

Thomson, M. C., Connor, S. J., Zebiak, S. E., Jancloes, M., & Mihretie, A. (2011).

Africa needs climate data to fight disease. *Nature*, *471*(7339), 440–442. doi: 10.1038/471440a

Thorne, P., & Vose, R. (2010). Reanalyses suitable for characterizing long-term trends. *Bulletin of the American Meteorological Society*, *91*(3), 353–362. doi: 10.1175/2009BAMS2858.1

Towner, J., Cloke, H. L., Zsoter, E., Flamig, Z., Hoch, J. M., Bazo, J., ... Stephens, E. M. (2019). Assessing the performance of global hydrological models for capturing peak river flows in the Amazon basin. *Hydrology and Earth System Sciences*, *23*(7), 3057–3080. doi: 10.5194/hess-23-3057-2019

Trenberth, K. E., & Stepaniak, D. P. (2001). Indices of El Niño evolution. *Journal of Climate*, *14*(8), 1697–1701. doi: 10.1175/1520-0442(2001)014⟨1697:LIOENO⟩2.0.CO;2

Virts, K. S., & Wallace, J. M. (2014). Observations of temperature, wind, cirrus, and trace gases in the tropical tropopause transition layer during the mjo. *Journal of the Atmospheric Sciences*, *71*(3), 1143–1157. doi: 10.1175/JAS-D-13-0178.1

Wang, B.-R., Liu, X.-Y., & Wang, J.-K. (2013). Assessment of COSMIC radio occultation retrieval product using global radiosonde data. *Atmospheric Measurement Techniques*, *6*(4), 1073–1083. doi: 10.5194/amt-6-1073-2013

Wang, J., & Zhang, L. (2008). Systematic errors in global radiosonde precipitable water data from comparisons with ground-based GPS measurements. *Journal of Climate*, *21*(10), 2218–2238. doi: 10.1175/2007JCLI1944.1

Westra, S., Brown, C., Lall, U., Koch, I., & Sharma, A. (2010). Interpreting variability in global SST data using independent component analysis and principal component analysis. *International Journal of Climatology*, *30*(3), 333–346. doi: 10.1002/joc.1888

Wheeler, M. C., & Hendon, H. H. (2004). An all-season real-time multivariate MJO index: Development of an index for monitoring and prediction. *Monthly Weather Review*, *132*(8), 1917–1932. doi: 10.1175/1520-0493(2004)132⟨1917:AARMMI⟩2.0.CO;2

Wickert, J. (2004). *Comparison of vertical refractivity and temperature profiles from CHAMP with radiosonde measurements* (Tech. Rep.). GFZ German Research Centre for Geosciences.

Wickert, J., Beyerle, G., König, R., Heise, S., Grunwaldt, L., Michalak, G., ...

Schmidt, T. (2005). GPS radio occultation with CHAMP and GRACE: A first look at a new and promising satellite configuration for global atmospheric sounding. *Annales Geophysicae*, *23*(3), 653–658. doi: 10.5194/angeo-23-653-2005

WMO. (1957). Meteorology-A three-dimensional science: Second session of the commission for aerology. *WMO Bulletin*, *4*(4), 134-138.

WMO. (2020). *The gaps in the Global Basic Observing Network (GBON)* (Tech. Rep.). Geneva, Switzerland: World Meteorological Organization.

World Meteorological Organization. (2020). *WMO Codes Registry.* `http://codes.wmo.int/`. WMO. (Accessed: 2020-06-06)

Xian, T., & Homeyer, C. R. (2019). Global tropopause altitudes in radiosondes and reanalyses. *Atmospheric Chemistry and Physics*, *19*, 5661–5678. doi: 10.5194/acp-2018-945

Zeng, Z., Ho, S.-P., Sokolovskiy, S., & Kuo, Y.-H. (2012). Structural evolution of the Madden-Julian Oscillation from COSMIC radio occultation data. *Journal of Geophysical Research: Atmospheres*, *117*(D22). doi: 10.1029/2012JD017685

Zhang, C. (2005). Madden-Julian Oscillation. *Reviews of Geophysics*, *43*(2). doi: 10.1029/2004RG000158