On the Diurnal Cycle of GPS-Derived Precipitable Water Vapor over Sumatra

GIUSEPPE TORRI
Department of Atmospheric Sciences, University of Hawai’i at Mānoa, Honolulu, Hawaii

DAVID K. ADAMS
Centro de la Ciencias de la Atmósfera, Universidad Nacional Autónoma de México, Mexico City, Mexico

HUIQUN WANG
Smithsonian Astrophysical Observatory, Cambridge, Massachusetts

ZHIMING KUANG
Department of Earth and Planetary Sciences, and John A. Paulson School of Engineering and Applied Sciences, Harvard University, Cambridge, Massachusetts

(Manuscript received 15 April 2019, in final form 19 July 2019)

ABSTRACT
Convective processes in the atmosphere over the Maritime Continent and their diurnal cycles have important repercussions for the circulations in the tropics and beyond. In this work, we present a new dataset of precipitable water vapor (PWV) obtained from the Sumatran GPS Array (SuGAr), a dense network of GPS stations principally for examining seismic and tectonic activity along the western coast of Sumatra and several offshore islands. The data provide an opportunity to examine the characteristics of convection over the area in greater detail than before. In particular, our results show that the diurnal cycle of PWV on Sumatra has a single late afternoon peak, while that offshore has both a midday and a nocturnal peak. The SuGAr data are in good agreement with GPS radio occultation data from the Constellation Observing System for Meteorology, Ionosphere, and Climate (COSMIC) mission, as well as with imaging spectrometer data from the Ozone Measuring Instrument (OMI). A comparison between SuGAr and the NASA Water Vapor Project (NVAP), however, shows significant differences, most likely due to discrepancies in the temporal and spatial resolutions. To further understand the diurnal cycle contained in the SuGAr data, we explore the impact of the Madden–Julian oscillation (MJO) on the diurnal cycle with the aid of the Weather Research and Forecasting (WRF) Model. Results show that the daily mean and the amplitude of the diurnal cycle appear smaller during the suppressed phase relative to the developing/active MJO phase. Furthermore, the evening/nighttime peaks of PWV offshore appear later during the suppressed phase of the MJO compared to the active phase.

1. Introduction
The island of Sumatra sits at the western edge of the Maritime Continent, an area situated between the Indian and Pacific Oceans which also includes the Malay Peninsula and large islands, such as Java, Borneo, Sulawesi, New Guinea, and the Philippines, as well as a galaxy of smaller islands. The combination of easterly trade winds over the tropical Pacific Ocean and the relatively shallow seabeds between many of the islands makes the waters around the Maritime Continent some of the warmest on Earth, with values of sea surface temperature (SST) persistently above 300 K (Locarnini et al. 2013). Thus, the ocean around the Maritime Continent is also known as the Tropical or Indo-Pacific Warm Pool.

Ramage (1968) was one of the first to recognize the importance of the Maritime Continent in the world’s climate. In essence, the abundance of atmospheric water vapor due to high SSTs facilitates the development of deep convection which, in turn, results in large amounts of latent heat release in the free troposphere, shaping the circulation in the tropics and beyond (e.g., Neale and Slingo 2003). Despite this importance, global climate...
models (GCMs) often display large errors when simulating the diurnal cycle in this region (Yang and Slingo 2001). The reasons may be ascribed to inadequately resolved complex topography and convective processes. Given the Maritime Continent’s remote impact over the planet, simulation errors can propagate on a global scale. A better understanding of the interaction among the diurnal cycle, convective activity, and large-scale circulation patterns, such as the Madden–Julian oscillation (MJO), is therefore highly desirable.

More generally, the diurnal cycle of convection in the tropics still represents a challenge. For example, GCMs tend to predict a peak of convection at midday over land, while observations suggest this happens in the late afternoon (e.g., Yang and Slingo 2001; Betts and Jakob 2002; Bechtold et al. 2004; Dai and Trenberth 2004). This problem can have a number of repercussions, including on the calculation of Earth’s radiation budget, given the importance of the relative phase between solar radiation and cloud cover. One reason for this early bias in precipitation could be that the onset of deep convection, or the transition from shallow to deep convection, in the models has inadequate dependence on midtropospheric humidity. From a conceptual point of view, the dynamical features that participate in the transition—updrafts, downdrafts, and cold pools—and how they interact are still the subjects of much study (e.g., Adams et al. 2013; Schlemmer and Hohenegger 2014; Feng et al. 2015a; Torri et al. 2015; Gentine et al. 2016; Schlemmer and Hohenegger 2016; Torri and Kuang 2016; Schiro et al. 2016; Grant and van den Heever 2016; Romps and Jeevanjee 2016; Adams et al. 2017; Drager and van den Heever 2017; Grant et al. 2018; Grant and van den Heever 2018; Haerter and Schlemmer 2018; Tian and Kuang 2019; Torri and Kuang 2019), and many questions still remain to be fully answered.

In the Maritime Continent, convection develops over the islands in response to diabatic heating at the surface and then propagates offshore (e.g., Houze et al. 1981; Yang and Slingo 2001; Qian 2008; Love et al. 2011; Peatman et al. 2014). In an observational study of convection in the northwest region of Borneo, Houze et al. (1981) hypothesized that the offshore propagation of convection was due to land–sea breezes. While this and similar ideas were considered by other authors (e.g., Zhou and Wang 2006; Qian 2008; Fujita et al. 2010, 2011; Wapler and Lane 2012), in a study on convection in the Panama Bight, Mapes et al. (2003) argued that sea breezes do not have the right speed nor the outward extent to explain the propagation of convection, and suggested that gravity waves were instead responsible. This hypothesis was used in the modeling study of Love et al. (2011) to explain the presence of nocturnal convection off the coasts of Sumatra, and was also considered by Hassim et al. (2016), who showed that gravity waves were important for the propagation of convection generated over land off the coasts of Papua New Guinea during nighttime.

An important aspect of the diurnal cycle we will consider in part of this study is the complex interaction between the Maritime Continent diurnal cycle and the MJO (Madden and Julian 1971, 1972; Zhang 2005). The MJO is essentially a convectively active envelope that develops over the Indian Ocean and slowly propagates eastward. Because the MJO has such an important influence on the rainfall over the Maritime Continent (e.g., Tian et al. 2006; Suzuki 2009; Fujita et al. 2011; Oh et al. 2012; Hagos et al. 2016; Sakaeda et al. 2017; Kerns and Chen 2018), and, more generally, on the weather of the planet (see, among many others, Kiladis and Weickmann 1992; Mo 2000; Higgins et al. 2000; Higgins and Shi 2001; Bond and Vecchi 2003; Jones et al. 2004; Becker et al. 2011; Schreck et al. 2013; Thompson and Roundy 2013; Matsueda and Takaya 2015; Klotzbach et al. 2016; Zhou et al. 2016; Zheng et al. 2018; Tippett 2018; Barrett 2019), it is important to forecast the MJO accurately.

Upon reaching the Maritime Continent, some MJO events weaken and do not propagate farther (e.g., Rui and Wang 1990; Salby and Hendon 1994; Zhang and Hendon 1997; Hsu and Lee 2005; Oh et al. 2013; Feng et al. 2015a; Vincent and Lane 2017). With few exceptions, GCMs are known to have relatively low skill in simulating the MJO (Inness et al. 2003: Inness and Slingo 2006; Kim et al. 2009; Jiang et al. 2015; Kim et al. 2018), particularly when crossing the Maritime Continent. Many GCMs tend to produce events that collapse too frequently while propagating through the Maritime Continent, an issue often referred to as the “Maritime Continent prediction barrier” (e.g., Inness et al. 2003; Inness and Slingo 2006; Kim et al. 2016; Zhang and Ling 2017; S. Wang et al. 2019). Several studies have suggested that the strong diurnal cycle over the Maritime Continent might play an important role in creating this barrier (e.g., Neale and Slingo 2003; Peatman et al. 2014, 2015; Majda and Yang 2016; Hagos et al. 2016). Hagos et al. (2016), for example, showed that removing the diurnal cycle of insolation led to a more coherent propagation of the MJO across the Maritime Continent.

Even general conclusions about interactions between the MJO and the diurnal cycle over the Maritime Continent often diverge (e.g., Sui and Lau 1992; Chen and Houze 1997; Tian et al. 2006; Fujita et al. 2011; Rauniyar and Walsh 2011; Oh et al. 2012; Vincent et al. 2016). For example, Oh et al. (2012) used reanalysis data to show that the diurnal cycle of rainfall was weaker over the
islands of the Maritime Continent and stronger over the ocean during the mature stage of the MJO. On the other hand, Tian et al. (2006) analyzed data from the Tropical Rainfall Measurement Mission (TRMM) and reported an enhanced cycle of deep convective cloud amount during the mature stage of the MJO over both land and ocean. In an earlier paper, Sui and Lau (1992) reported that active periods of intraseasonal variability were associated with a diminished diurnal cycle of convection over the Maritime Continent. Similarly varying results were found examining the MJO impact on the phase of rainfall and cloudiness, with some authors documenting a change (e.g., Chen and Houze 1997; Fujita et al. 2011; Rauniyar and Walsh 2011; Oh et al. 2012), while others found no significant differences (e.g., Tian et al. 2006; Suzuki 2009).

Precipitable, or column, water vapor has proven to be a particularly useful variable for studying tropical deep convection given its strong relationship with precipitation (Zeng 1999; Bretherton et al. 2004; Holloway and Neelin 2009, 2010; Lintner et al. 2011; Schiro et al. 2016), as well as a critical metric for models to replicate (Adams et al. 2013, 2017; Lintner et al. 2017). The PWV–precipitation relationship, observed essentially only over tropical oceans, is shown to be nonlinear with precipitation, increasing exponentially above a certain PWV threshold value (see, e.g., Bretherton et al. 2004). A caveat should be kept in mind, however, that this relationship has been found to be weaker over coasts for rainfall associated with land–sea effects (Bergemann and Jakob 2016). Furthermore, PWV is readily available from radiosondes and different satellite platforms. However, these observations are often not adequate for representing the convective diurnal cycle given their temporal and/or spatial resolutions, availability only in clear-sky conditions, or being less reliable over land. Ground-based Global Navigation Satellite System (GNSS) or GPS-derived PWV is advantageous because it has high temporal resolution (~5 min) and is available under all-weather conditions (Bevis et al. 1992; Solheim et al. 1999). This manuscript examines the diurnal cycle of PWV obtained from a network of GPS stations situated on the west coast of Sumatra and on the archipelago of islands offshore.

Several previous studies of Maritime Continent convection have used GPS PWV data. For instance, field campaigns focusing on the diurnal cycle have been carried out (Wu et al. 2003, 2008; Realini et al. 2014). Fujita et al. (2011) conducted a longer term study, from 2005 to 2008, employing a network of 24 GPS stations and interpolated NCEP reanalysis meteorological data over Sumatra with the goal of assessing MJO effects on the diurnal cycle of convection in the Indian Ocean off the coasts of Sumatra. Compared with previous studies, the network employed here, the Sumatran GPS Array, is denser, with more stations over Sumatra. The data produced by this expanded network are better suited to shed light on the convective processes affected by the complex topography of Sumatra, on how convection propagates offshore over the ocean, and on how this is influenced by the MJO.

In section 2, we will introduce the datasets used for this study and discuss the model setup for the simulations. In section 3, we will present the main results from the SuGAr network, compare them with other datasets and model outputs, and discuss the differences revealed by the SuGAr data during different phases of the MJO. In section 4, we will synthesize observational and modeling results to interpret our findings, from the perspective of the impact of the MJO on the diurnal cycle of PWV over Sumatra.

2. Methods and data

In this section, we present the SuGAr network and the PWV obtained. Likewise, the PWV datasets employed for comparison purposes are described. Finally, the setup of the WRF numerical simulations and how the output was used to diagnose MJO phases are presented.

a. SuGAr data

The principal data employed here originate from 60 GPS stations, called the Sumatran GPS Array (SuGAr; Feng et al. 2015b), which is operated in a collaborative effort between the Indonesian Institute of Science (LIPI) and the Earth Observatory of Singapore, principally to monitor seismic and tectonic activity. The zenith total delay data produced (see following paragraph) are available with a temporal resolution of 5 min and can be converted to PWV. Although the majority of the stations are located on the archipelago to the west of Sumatra, many are at various heights on the island of Sumatra (Fig. 2a). The precise coordinates and the elevation of the stations can be found in Feng et al. (2015b). For this study, we will discuss results obtained with data from 45 stations from 2008 to 2013.

Delays in GPS signals due to the presence of atmospheric water vapor provide the basis for calculating PWV. Furthermore, given the relative insensitivity of L-band radio signals to cloud droplets and rainfall, GPS can provide all-weather PWV data (Solheim et al. 1999). It should be noted that the requirements necessary for measuring millimetric displacements in tectonic studies can be relaxed with minimal influence on the accuracy of GPS PWV, allowing for the expansion of GPS meteorological sites into logistically difficult places as well as
decreasing costs (Rocken et al. 2005; Adams et al. 2011, 2015).

Specifically, PWV is calculated from the zenith total delay (ZTD), which represents the refractivity delay of the signal induced by constituent atmospheric gases, including water vapor (Bevis et al. 1992). PWV was calculated from ZTD using GPS-Inferred Positioning System and Orbit Analysis Simulation Software (GIPSY-OASIS; https://gipsy-oasis.jpl.nasa.gov/), obtained from NASA JPL. To convert ZTD to PWV, surface pressure and temperature are necessary. These were obtained from the Integrated Surface Hourly (ISH) dataset (NCDC 2005) at 3-h intervals. Given that these data do not necessarily coincide geographically with SuGAr stations, the nearest ISH site was chosen. To approximate changes in pressure and temperature given differences between the meteorological station height and the GPS antenna elevation, the hypsometric equation was used, and a lapse rate of 6.5 K km

$^{-1}$ was applied. The corresponding 3-h ZTD data were then converted into PWV with these adjusted pressure and temperature values. Results derived from these data were compared with the data sources described below.

b. COSMIC radio occultation data

The GPS radio occultation (RO) provides high-vertical-resolution, all-weather vertical profiles of atmospheric water vapor globally. Data from Constellation Observing System for Meteorology, Ionosphere, and Climate (COSMIC; Kuo et al. 2004; Anthes et al. 2008) are employed in this study as a point of comparison with the SuGAr data. Several factors may affect the comparison. First, COSMIC water vapor profiles have different horizontal and temporal coverage relative to ground-based GPS PWV, with typically one daily profile over a 400 km $\times$ 400 km area (Xie et al. 2018). Second, the pathlength representative of the water vapor is on average over approximately 200–300 km (Kursinski et al. 1997; Ao et al. 2012), compared with the $\sim$20-km-diameter viewing cone of ground-based GPS (Adams et al. 2013). Another important caveat when using COSMIC, or radio occultation data in general, is that the presence of a moisture inversion over marine boundary layers, particularly in subtropical regions, can cause superrefraction, which can bias the retrieved data (see, e.g., Xie et al. 2006; Schröder et al. 2018). Regardless, COSMIC RO data can be compared with SuGAr data to check their consistency (see section 3).

c. OMI

The Ozone Measuring Instrument (OMI) is on board the NASA EOS Aura satellite in a sun-synchronous orbit. It is an ultraviolet–visible imaging spectrometer that provides daily global observations around 1330 local time. Although OMI was primarily designed to observe ozone, PWV data have been retrieved from the blue spectral range at a resolution of 13 km $\times$ 25 km (Wagner et al. 2013; Wang et al. 2016). This paper uses the latest version 4 OMI PWV available at NASA’s Aura Validation Data Center. OMI observations have a similar sensitivity for land and ocean, making them attractive to study the Maritime Continent. However, errors due to clouds can propagate into large errors in PWV. Thus, we adopt the recommended data filtering criteria by requiring cloud fraction to be smaller than 15% and cloud-top pressure higher than 750 hPa, in addition to spectral fitting root-mean-square less than 0.005 (H. Wang et al. 2019). Due to the frequent deep convective cloudiness of the Maritime Continent, a large portion of retrieved PWV values are filtered out.

d. NVAP-M

The NASA Water Vapor Project (NVAP) is a global dataset created by combining water vapor retrievals obtained from different sources (Randel et al. 1996; Vonder Haar et al. 2012). Here, we utilize the data from the NASA Making Earth Science Data Records for Use in Research Environments (MEaSUREs) initiative, obtained by reprocessing and extending the original NVAP data (Vonder Haar et al. 2012), specifically, NVAP-M PWV, available from 1988 until 2009, with a horizontal and temporal resolution of 0.5° and 6 h, respectively.

e. WRF modeling

The ARW version of the Weather Research and Forecasting (WRF) Model (Skamarock et al. 2008), version 3.8.1, is used to compare its output with the PWV values retrieved from SuGAr, and to diagnose different phases of the MJO. WRF is a fully compressible nonhydrostatic model, offering many user-selected configurations and parameterization schemes, and is widely used in atmospheric process studies, such as the MJO (e.g., Wang et al. 2015; Ensembles: Wang et al. 2016; Hassim et al. 2016; Vincent et al. 2016; Vincent and Lane 2017, 2018; Tan et al. 2019). In this study, WRF is run in a two-way nested configuration, with an outer domain containing the tropical Indian Ocean and the Maritime Continent (20°N–20°S, 48°–120°E) at 9-km horizontal resolution, and an inner domain focused on Sumatra (6°N–6°S, 94°–107°E) at 3-km horizontal resolution. Variables are output every 3 h in the outer domain and every 30 min in the inner one. There are 45 vertical levels, 8 of which are in the lowest 1 km, and the nominal top is at 20 hPa. Our setup follows Wang et al. (2015), differing only in the nested domain and the integration period. We conducted two simulations: the first, started
on 1 January 2012 and run until 31 December 2012, to compare WRF with SuGaR; the second, from 1 September 2011 to 1 April 2012, mainly to compute the outgoing longwave radiation (OLR) field for MJO-index calculations. The second simulation allows us to investigate the effect of the MJO at temporal and spatial resolutions finer than those for SuGaR PWV. The shorter time period for the second simulation is because the MJO undergoes significant seasonal shifts in its location (Wheeler and Kiladis 1999).

For both simulations, the initial and boundary conditions, as well as SSTs, were provided by ERA-Interim (Dee et al. 2011). Cloud microphysics were parameterized using the WRF Double-Moment (WDM) scheme (Lim and Hong 2010), while radiative transfer was parameterized using the GCM version of the Rapid Radiative Transfer Model (RRTM) longwave radiation scheme (Iacono et al. 2008) and the updated Goddard shortwave scheme (Chou and Suarez 1994). Surface fluxes were computed using the Revised MM5 Monin–Obukhov scheme (Jiménez et al. 2012). The Unified Noah land surface model (Chen and Dudhia 2001) was used to compute temperatures over land, the Yonsei University scheme PBL (YSU; Hong et al. 2006) to parameterize subgrid-scale vertical turbulent mixing, convective processes were resolved explicitly by the model. For the first three days, horizontal winds were relaxed to the reanalysis values using spectral nudging with zonal wavenumber 0–4 and meridional wavenumber 0–2. No nudging was imposed after day 3. Rayleigh-damping was imposed in the top 7 km of the model to prevent reflection of gravity waves (Klemp et al. 2008).

To ensure that the model reproduces the passage of the MJO in a reasonably accurate manner, we compared the modeled precipitation rates diagnosed in the outer domain with those estimated by the TRMM 3B42 V7 (Huffman et al. 2007). The TRMM data have a temporal resolution of 3 h and a spatial resolution of 0.25° × 0.25°. Figures 1a and 1b show, respectively, the Hovmöller diagrams of precipitation rate in WRF and TRMM data. Each panel was constructed by first averaging precipitation rates in the latitudinal band between 10°N and 10°S. In spite of some differences, the main features seen by TRMM, particularly the eastward propagation of a number of convective envelopes, are well reproduced. Together with a favorable comparison with SuGaR data to be discussed later, this provides confidence in the use of WRF to better understand how the MJO affects the diurnal cycle of PWV over the Maritime Continent.

d. Diagnosing MJO phases

Among the methods to diagnose MJO phases, we have opted for the all-season OLR-based index described in Kiladis et al. (2014). We computed the index using daily means of OLR simulated in the outer domain for the period from 1 October 2011 to 1 April 2012. To reduce the computational costs involved, we coarsened the model output by a factor of 4 in each horizontal direction. Following Kiladis et al. (2014), we also applied a 30–96-day bandpass filter to the eastward-only wavenumbers (including the zonal mean), and computed the two leading empirical orthogonal functions (EOFs) of the resulting field. Finally, the principal components associated with the EOFs were used to span a two-dimensional phase space where different octants represent eight phases of the MJO cycle. As the analysis was on a relatively short simulation, we did not use a 121-day sliding time window to compute the MJO phases as in Kiladis et al. (2014).

3. Results

We begin by presenting the average diurnal cycle of PWV obtained from SuGaR, discussing the daily mean and the diurnal amplitude at various GPS stations. We then compare SuGaR data with PWV values obtained from other datasets: for COSMIC and OMI, we compare single measurements, whereas for NVAP-M we compare the average diurnal cycles. We also compare SuGaR PWV data with the numerical simulations. Finally, we analyze SuGaR data under the context of WRF-derived MJO phases to examine how the MJO affects the diurnal cycle of PWV over Sumatra.

a. The diurnal cycle

Given that the diurnal cycle from stations offshore often reveal a double-peak structure of different magnitudes, we first discuss the daily mean (Fig. 2a) and the amplitude (Fig. 2b) of the average diurnal cycle of PWV for every SuGaR station analyzed. Typical daily mean SuGaR PWV values are mostly between 50 and 55 mm for near–sea level sites. The amplitude was computed as the difference between the maximum and minimum PWV of the averaged diurnal cycle. Stations along the coast of Sumatra have higher amplitudes, ranging between 2 and 5 mm, whereas those on the archipelago remain mostly under 2 mm. This result conforms with other studies, showing that the diurnal cycle is generally weaker over the ocean than over land (e.g., Gray and Jacobson 1977; Yang and Slingo 2001; Nesbitt and Zipser 2003).

Figures 2c and 2d show, respectively, the daily mean and the amplitude of the diurnal cycle as a function of station altitude. There appears to be some weak correlation with height, although the relationship is debatable. The colors of the dots in Fig. 2c represent the distance of the corresponding SuGaR station from the coast of Sumatra, with negative values to the east, and
positive to the west. While there is a clear difference in the amplitudes between stations along the west coast of Sumatra and those on the archipelago, the same does not hold for the daily means.

Figure 3a shows the normalized average diurnal cycles for all SuGAr stations. The normalization was conducted for each station by first subtracting the minimum value of the average diurnal cycle and then dividing by the amplitude. The resulting normalized diurnal cycle ranges from 0 to 1 for each station. The profiles are colored according to the distance from the west coast of Sumatra, with brighter yellows indicating that a station is on Sumatra, and darker blues indicating a station is on one of the islands in the offshore archipelago. For convenience, we have picked a distance of 20 km from Sumatra’s west coast and divided stations in two groups: the first, called “Archipelago,” contains all the stations lying west of the demarcating line; the second, called “Sumatra,” those to the east. The thicker navy- and mustard-colored lines indicate the normalized average diurnal cycle representative of the two groups.

The diurnal cycle of the Sumatra group has a unimodal shape with a late afternoon peak, while that of the Archipelago group is nearly semidiurnal, with one peak around midday and the other late at night (e.g., Wu et al. 2003; Mori et al. 2004; Sakurai et al. 2005; Wu et al. 2008; Sakurai et al. 2009; Wu et al. 2009; Fujita et al. 2011; Kamimera et al. 2012; Yokoi et al. 2017). It is possible that a semidiurnal pattern could result from averaging together a large number of diurnal cycles, where some
stations have a diurnal cycle that peaks at noon and the rest a diurnal cycle that peaks at night. To investigate this, the power spectral density for each year and station is computed and shown in Figs. 3b and 3c for the Sumatra and the Archipelago stations, respectively. The curves are normalized by the 3-h sampling frequency of the data. For Archipelago stations, the peaks in the spectrum at 24 and 12 h are comparable, suggesting that semidiurnal variations for the PWV are indeed present in the Archipelago PWV data.

To gain a better appreciation for how the diurnal cycle of PWV changes from the Sumatra to the Archipelago stations, we take advantage of the density of the SuGAr network by examining the transect SDKL-RNDG-PBBL-PBKR in the northern part of the domain (highlighted with red circles in Fig. 2a). The diurnal cycles averaged over the 5-yr period are shown in Fig. 4a. Due to the station elevation range—from 894.4 m for SDKL, near the city of Sidikalang, to −3.49 m for PBBL, located on one of the Banyak islands—the mean of each curve is subtracted for easier visualization. The figure illustrates that the diurnal cycle of PWV has a progressively smaller amplitude and a late afternoon/evening peak as one goes from stations on Sumatra to those offshore.

From another perspective, Fig. 4b shows the corresponding ZTD values for the four stations filtered to remove anomalous data below 2000 cm. The mean for each station has been removed as in Fig. 4a. Figure 4b shows that, even though the relationship between PWV and ZTD depends on local pressure and temperature in a nontrivial way, the amplitudes of ZTD become smaller and the cycles show later peaks as one goes westward away from Sumatra. Moreover, compared with PWV data, the higher temporal resolution of ZTD is helpful for better illustrating the changes in the timing of the diurnal cycle with geographical location.

b. Comparisons with other datasets

In this section, we compare the SuGAr PWV with other datasets. Apart from being important in data

Fig. 2. (a) Geographic location of SuGAr stations, with their colors representing the mean of the average diurnal cycle of PWV. The red-bordered markers in the top left identify a transect that will be used for some analyses (see text for more detail). (b) As in (a), but for the amplitude of the average diurnal cycle. (c) Mean of average diurnal cycle of PWV as a function of station’s height. (d) As in (c), but for amplitudes. The colors represent the distance of a SuGAr station from the coast, with positive values to the west and negative to the east.
validation, this procedure could be useful to calibrate datasets that have much broader spatial coverage and can be extended to the entire Maritime Continent.

1) COSMIC

COSMIC data have global coverage but are much sparser than SuGAr data. We compared the datasets point by point: every time a COSMIC data point was within a radius of $1.5^\circ$ in latitude and longitude from a SuGAr station, it was matched with the SuGAr PWV closest in time. The collocation radius was based on the 300-km spatial resolution of COSMIC data (e.g., Kursinski et al. 1997).

All points satisfying this condition were collected, and a density matrix was constructed with a resolution of 2.5 mm in PWV, which means that the $(i,j)$ entry of the matrix contains the density of matches within 2.5 mm between the datasets. To convert the total number of matches to density, we normalized the counts by the square of the resolution, $2.5\text{ mm}^2$. The result is

---

**Fig. 3.** (a) Average diurnal cycles of PWV as a function of local time and normalized by the amplitude. Colors range from bright yellow for stations to the east of Sumatra’s west coast to dark blue for stations west of the coast; the thick mustard (navy) line is the average of the stations in the Sumatra (Archipelago) group (see text for more information). (b) Average power spectral density for Sumatra stations normalized by the sampling frequency. (c) As in (b), but for Archipelago stations.

---

**Fig. 4.** (a) Average diurnal cycles of PWV for stations belonging to the transect (see text and Fig. 2). (b) As in (a), but for ZTD. In both panels, the mean of each cycle has been subtracted to make the comparison easier.
shown in Fig. 5a, where the one-to-one line is drawn for easier comparison. Overall, there is good agreement, in spite of the scatter that can be partly attributed to the different resolution of the two datasets, and partly to topography.

2) OMI

OMI PWV retrievals are highly sensitive to clouds, which, for this study, implies many missing data due to cloud filtering. For this reason, as well as its once-a-day availability, we used the same method as for the COSMIC dataset and directly compared “coincident” measurements. For each GPS station, the corresponding filtered version 4 OMI data (section 2c) within 0.25\° latitude and 0.25\° longitude was averaged. Because OMI measurements are made approximately at 1330 local time, for each SuGAr station we have matched its values at 1300 LST of the corresponding day, the closest to the OMI observational time. For points with matched local time, only OMI data with cloud fraction less than 15% were used to create Fig. 5b. OMI and SuGAr compare reasonably well, up to approximately 50 mm. The agreement degrades for higher values of PWV. Overestimation of PWV by OMI was noticed previously (Wang et al. 2016) and was attributed to the presence of clouds. A 5% cloud fraction cutoff is expected to decrease the discrepancy but leaves few data pairs to analyze.

3) NVAP-M

Available from 1988 until 2009, NVAP-M has at best only 2 years of data overlap with SuGAr, particularly given that not all the SuGAr stations go back to 2008 (Feng et al. 2015b). Due to this limitation, we compared the average diurnal cycles for two time intervals: 1988–2009 for NVAP-M, and 2008–13 for SuGAr, with the underlying assumption that no significant changes to the diurnal convective dynamics over Sumatra occurred between the two time intervals.

Because NVAP-M data has greater spatial and temporal coverage than COSMIC over the Sumatra region, we can reconstruct an average diurnal cycle for NVAP-M grids. As we did with SuGAr, we considered the daily mean as well as the amplitude of the time-averaged diurnal cycle. The results are shown in Figs. 6a and 6b, where the values derived from the SuGAr stations are represented by color points, and NVAP-M values are represented by the background colors; Figs. 6c and 6d show a point-to-point comparison between the SuGAr and NVAP-M daily means and amplitudes, respectively, for the grid box corresponding to each SuGAr station. The colors of the points in Figs. 6c and 6d correspond to the distance from the west coast of Sumatra, with positive values referring to points to the west of the coast and negative values to the east.

Figures 6a and 6c indicate that the daily mean values diagnosed with SuGAr express a much smaller range than those from NVAP-M. In particular, while points on the archipelago of smaller islands seem to compare reasonably well, measurements over Sumatra show significant discrepancies. This could be related to the coarse resolution of NVAP-M and the significant topography over Sumatra. The colors in Fig. 6a indeed show a relatively sharp increase of PWV values going from the
mountain chain in Sumatra toward the ocean, which suggests that these points are more prone to errors caused by coarse resolution, not properly representing sharp changes in elevation. Figure 6b compares the amplitudes derived from the two datasets. Figure 6d indicates that the amplitudes compare better for Archipelago stations, but, as with the daily mean values, there are significant discrepancies for Sumatra stations.

To better appreciate the differences in the average diurnal cycles derived from these two datasets, Fig. 7a shows a comparison between the PWV average diurnal cycles along the transect described at the beginning of this section, the continuous lines representing values from SuGAr, and the dashed lines the values from NVAP-M. The SuGAr values are the same as in Fig. 4. The values of NVAP-M are chosen from the grid box containing the SuGAr stations we wish to compare with. Because the SuGAr stations PBLI and PBKR correspond to the same NVAP-M grid box, only one curve (dashed violet) is shown for them. The NVAP-M results suggest a much smoother transition of PWV values from the mountainous zones toward the lower-lying points on the archipelago than SuGAr. Notice also that the phases of the diurnal cycles are not well represented by NVAP-M. The dashed blue line shows a peak 6 h later than the continuous, and the two-peak cycles in violet appear out of phase by 6 h. Given that the NVAP-M data have a temporal resolution of 6 h, such discrepancies in phase are not too surprising.

4) WRF

We supplement our data analysis with a comparison between the SuGAr dataset and WRF. Following the same recipe as that for the NVAP-M data, we present the results in Fig. 8: Figs. 8a and 8c show a comparison of the daily means, while Figs. 8b and 8d compare the amplitudes. The top panels, Figs. 8a and 8b, show the values distributed geographically, with SuGAr stations represented by dots and WRF values as background colors; Figs. 8c and 8d show a point-to-point comparison,
where values from each SuGAr station are compared with WRF values for the grid box corresponding to the position of the SuGAr station. The colors of the scatterpoints in the bottom plots represent distance from Sumatra’s west coast, as in Fig. 3.

Figures 8a and 8c suggest that the daily means diagnosed from SuGAr are systematically higher than those simulated with WRF. The difference is approximately 2 mm for most stations, although for some it reaches 4 mm, particularly over Sumatra. Figures 8b and 8d suggest that the diurnal cycle amplitudes simulated by WRF compare well with SuGAr observations. At larger values the difference is greater, but it is hard to discern given the fewer stations in those regimes.

Figure 9 is constructed in the same way as Fig. 3, but for WRF results. The agreement between the two figures is quite good, although for Archipelago stations the peaks at 24 and 12 h in the power spectral density of SuGAr are of comparable magnitude, while those from WRF differ by a factor of 3. Nevertheless, the average shapes of the diurnal cycles in Fig. 9a are very similar to those in Fig. 3a. Good agreement in the shapes of the diurnal cycles between SuGAr and WRF can also be seen when comparing the values in the transect, as shown in Fig. 7b. SuGAr stations are moister than the corresponding grid boxes in WRF, amplitudes for some stations appear greater than WRF, and phases seem overall well reproduced. This provides additional confidence when using WRF to study the processes that influence the diurnal cycle of PWV revealed by SuGAr.

Figure 10 represents the average diurnal cycle diagnosed using WRF along a portion of a longitudinal slice at latitude 1.48°S. The black line at the bottom represents the topography of Siberut island—scaled by a factor of 100—that the slice cuts through. The figure shows a series of peaks over Siberut between 1300 and 1600 LST, possibly the result of thermally forced convection over the island (e.g., Cronin et al. 2015; Wang and Kirshbaum 2015; Wang and Sobel 2017), while peaks of PWV over the ocean to the east and to the west of Siberut tend to happen late at night. Notice, however, that convection has a diurnal periodicity almost everywhere, except for those grid points along the coast of the island. Thus, it would seem intuitive to think that the coarser the resolution, the greater the area along the coast where a double-peaked diurnal cycle could be observed. This suggests that caution should be exercised before generalization of the finding that PWV has a semidiurnal cycle for Archipelago stations: the semidiurnal feature could alternatively be the result of the locations of the SuGAr stations and the effective resolution of 20 km—coarser than the WRF simulation—associated with the GPS data.

c. The diurnal cycle and the MJO

As explained in section 2, we derived the all-season index of Kiladis et al. (2014) from the simulated OLR from 1 October 2011 to 1 April 2012. Figure 11 shows the filtered OLR anomalies at different phases of the MJO. Given that the MJO is a continuous process, there is some ambiguity in identifying precisely which is the active and which is the suppressed phase. For the purpose of this paper, we consider phases 4 and 5 to be the active phase stages, and phases 6 and 7 to be the suppressed phases. We will refer to the phases in between as decaying (phases 6 and 7) or developing (phases 2 and 3).
For every station, we average all the diurnal cycles for the days within the same MJO phase. Figures 12a and 12b show, respectively, the results for the Sumatra and Archipelago groups, with different colors indicating different phases of the MJO. For both groups, the largest values of PWV are observed during phase 4 (active phase), and the lowest values are in phase 8 (suppressed phase).

Figure 12c shows the difference between the daily mean values during phase 8 and those during phase 4, and Fig. 12d shows the difference in amplitudes. The amplitude of the diurnal cycle is greater during the developing and active phases, and smaller during the suppressed phase. Stations along the west coast of Sumatra exhibit a particularly large change, from an average amplitude difference of almost 6 mm to less than 2 mm. Among these stations, the change is somewhat gradual as the convective envelope arrives and passes by.

The amplitude of the PWV diurnal cycle for Archipelago stations is also affected, although in a more complex way. As we have previously discussed, the amplitudes for Archipelago stations tend to be rather small, making disentangling the MJO signal from noise difficult. Comparing some of the developing/active phases with the suppressed phases, it appears that the nighttime peaks (0100–0400 LST) are larger in the former than in the latter scenarios. The peaks at midday appear less sensitive to MJO phases.

Finally, changes in the timing of PWV peaks are difficult to decipher given the relatively coarse resolution of 3 h. For Sumatra stations, the peak of the diurnal cycle appears to shift from 2200 LST in the suppressed phase to 1900 LST in developing/active phases. Something similar seems to hold for Archipelago stations, whereby the nighttime peak tends to happen later in the day during the suppressed rather than the active MJO phase, while the midday peak does not appear to vary much, apart from some slight changes during phases 7 and 8. Because the changes are of the same order of magnitude as the temporal resolution of the SuGaR dataset, we follow up with a further test. The same analysis is repeated for ZTD which has a higher temporal resolution: Figs. 13a and 13c confirm the shift forward of a few hours in evening peaks for the Sumatra stations and nighttime peaks for the Archipelago stations.
peaks for the Archipelago stations as conditions advance from a developing/active MJO phase to a suppressed one; however, the midday peaks for Archipelago stations vary little throughout the different MJO phases.

Encouraged by the favorable comparison between SuGAr data and the WRF simulations, we use WRF to investigate how the MJO affects the diurnal cycle of PWV. The panels of Fig. 14 represent the diurnal cycles of the simulated PWV, averaged in the direction parallel to the Sumatra coast, as a function of distance from the west coast of Sumatra, positive values being to the west, negative to the east, averaged over the days of each MJO phase. For simplicity, we have subtracted the diurnal cycles of PWV averaged over all simulated days in each panel. As seen for SuGAr, the highest values of PWV occur during the development of phases 4 and 5, whereas the lowest values are seen during phases 7 and 8. There are some slight differences in the other phases, but otherwise WRF is able to capture the observed variations in PWV.

To shed light on the diurnal cycle during distinct MJO phases, Fig. 15 compares the daily mean and amplitude as in Figs. 8c and 8d, where Figs. 15a and 15c refer to phase 4 of the MJO, while Figs. 15b and 15d refer to phase 8. The panels concur with our previous results with mean values and amplitudes generally being larger during the active phase 4, and smaller during the suppressed phase 8. As in Fig. 12, the simulated amplitudes for Sumatra stations are more affected by the MJO relative to those of the Archipelago. While there appears to be reasonable agreement for both diurnal cycle means and amplitudes during phase 8, the panels for phase 4 suggest the simulated daily mean PWV values are lower than SuGAr PWV while the amplitudes are larger. Compared with Figs. 8c and 8d, there is greater scatter around the one-to-one line, although one should keep in mind that Fig. 8 was constructed using data from an entire year, whereas Fig. 15 for single MJO phases during a 6-month period.

FIG. 10. Average diurnal cycle of PWV diagnosed from WRF on a longitudinal slice that cuts through the island of Siberut. The black line represents the topography of Siberut, scaled by a factor of 100.
To further illustrate the changes in simulated PWV values between different MJO phases, Fig. 16a shows the amplitudes diagnosed with WRF as a function of the distance from the coast and the different MJO phases. For stations within 50 km of the coast in both directions, amplitudes tend to maximize during the active phase 4, consistent with the SuGAr observations. Between 50 and 100 km to the west of Sumatra’s west coast, where many Archipelago stations are located, the amplitudes display fairly large values also during phase 3.

Changes in the timing of the diurnal cycle peaks can be seen in Fig. 16b, where the PWV values are plotted as a function of local time and distance for phase 4 in color, while those for phase 8 in contours. Because the topography creates a strong gradient in PWV values, we limit the contours at 46 mm and the color bar at 52 mm for the figure. As we saw from the SuGAr data, the diurnal cycle of PWV appears to peak earlier in the day during the enhanced phase—the exact time depending on the location—than it does during the suppressed phase. At a distance between 0 and 100 km, the difference seems large enough that it could also be seen using a 3-h-resolution dataset. This supports our earlier finding that the peak time in the diurnal cycle of PWV changes with different phases of the MJO.

Figures 16c and 16d represent the diurnal cycle of precipitation rate in WRF for phases 4 and 8, respectively. The main difference is the presence of a coherent band of precipitation that starts propagating westward in the late afternoon and reaches a distance of 200 km from the coast at approximately 0300 LST. This convective system also appears in phase 3, but not in phase 8, when precipitation appears to propagate westward more slowly and less coherently, possibly as a result of unorganized convection over the ocean. This feature is also present in Fig. 16c, although the different range of the color bar makes it harder to appreciate.

4. Discussion

Given the complex topography of the Maritime Continent, a dense network of stations like SuGAr that can provide high-temporal-resolution, all-weather PWV is highly valuable for understanding the interaction between convection and the larger-scale circulation (e.g., Bretherton et al. 2004; Peters and Neelin 2006; Holloway and Neelin 2009, 2010; Schiro et al. 2016; Schiro and Neelin 2018). In this paper, we illustrate how SuGAr data can be used for this purpose, using a transect of stations from the Barisan mountains to the small islands offshore of Sumatra almost perpendicularly to the shoreline. The GPS network data are used to construct the average diurnal cycles of PWV.

While Fig. 4a shows how the diurnal cycle changes progressively as we move offshore, it is potentially influenced by the relatively low temporal resolution.
ZTD data are available every 5 min, and, being proportional to PWV, they are analyzed to boost the confidence in the conclusion drawn from the PWV data. The 3-hourly ISH meteorological variables limit the GPS PWV temporal resolution. One way to address this problem is to interpolate the surface temperature and pressure data used in the algorithm to convert ZTD to achieve a much higher resolution. However, although PWV values retrieved from ZTD are not particularly sensitive to surface temperature, they are more sensitive to surface pressure: for example, a 10-K temperature change would give a difference roughly between 0.5 and 1.5 mm in PWV, a 1-hPa pressure change can lead to changes in derived PWV on the order to 3.5 mm. Thus, the result would likely depend on the interpolation method.

Our comparisons between datasets are encouraging. Both COSMIC and OMI PWV are linearly related to SuGAr PWV, although OMI data diverge at larger PWV due to cloud contamination. On the one hand, the comparison provides confidence in SuGAr PWV; on the other hand, it permits assessment of other datasets over the complex topography and island configuration of the Maritime Continent. Because both OMI and COSMIC have much broader spatial coverage than SuGAr, our study increases confidence in those datasets for regions broader than the one we have focused on. Our results suggest that caution should be exercised when using NVAP-M to study diurnal cycle in the Maritime Continent region, as NVAP-M PWV values are significantly lower than SuGAr’s for Sumatra stations, and higher for Archipelago stations. The time of peak of the diurnal cycles also appears to differ from SuGAr. The coarser spatial and temporal resolutions of NVAP-M may have contributed to the discrepancies.

The SuGAr PWV contains important information about how the MJO interacts with the diurnal cycle of convection over the Maritime Continent. To our knowledge, many previous observational studies on the topic, with the exception of Fujita et al. (2011), either focused on a larger area, such as the entire Maritime Continent, or on specific
campaigns that were limited in time to only a few MJO cycles. Given its density and long-term coverage, SuGAr PWV has the potential to provide the much needed metrics for models to replicate, as well as intensive case studies for single sites or transects, as we have demonstrated. This paper serves to motivate more in-depth future studies using the SuGAr data. In particular, ongoing work with the SuGAr data and WRF simulations is being conducted to understand the underlying physical mechanisms that explain why the diurnal cycle of PWV is significantly modulated by the MJO.

In comparing our results with the previous literature, one should keep in mind that Maritime Continent studies range widely in the observational techniques used, time period examined, and the geographical area studied. To our knowledge, Fujita et al. (2011) is the one most comparable to this paper in terms of the data used and the study area. Our work differs from Fujita et al. (2011) in several aspects. We used more GPS stations in the area, compared against other datasets, analyzed the details of the shape and timing of the diurnal cycles, raised/tested a couple of hypotheses about the diurnal cycles. Both our papers and their papers examine the effect of MJO on diurnal cycle.

Figures 13b and 13d show the anomalies of ZTD computed with respect to the mean of each MJO phase. The plots are similar to Fig. 6 of Fujita et al. (2011), where the deviations from the mean PWV for each MJO phase were presented, the main differences being that we are considering a different variable that also has higher temporal resolution.

Figure 13d is possibly the most immediately comparable to Fujita et al. (2011) since it is based on GPS stations placed in similar locations in the archipelago to the west of Sumatra. The midday/early afternoon positive anomalies have a similar behavior to those observed by Fujita et al. (2011). They tend to occur earlier in the day during the active phase, and later during the suppressed phase of the MJO. Likewise, the troughs appear earlier during the active phase, and there are no significant variations across MJO phases. A similar behavior is observed for stations on Sumatra, with ZTD peaking earlier in the day during the active phase of the MJO. Finally, Fujita et al. (2011) showed that the late afternoon/evening peaks in PWV happen later in the day during the active phase, while we observe only a slight shift of the peaks of ZTD and earlier in the day.

5. Conclusions

In this study, we have presented a new dataset of PWV measurements obtained from a GPS network called SuGAr, whose stations are located along the west coast of Sumatra and on the smaller islands offshore. This dataset could be considered as an expansion of that used in Fujita et al. (2011), with the addition of many stations on the island of Sumatra, and an overall greater density of stations. Interested in how this dataset could be used...
to better understand the diurnal cycle of PWV, we have examined the daily means and amplitudes of the diurnal cycle for each station between 2008 and 2013. As expected, the means correlate well with the altitude of the station, while the amplitudes decrease progressively from Sumatra to the islands offshore. We have then discussed the peaks of the average diurnal cycles, observing that the average profile for stations over land tended to have a single peak at 1900 LST, while the stations offshore tended to have two peaks, one at 1300 LST, and the other at 0100 LST.

We also compared the values of PWV obtained from SuGAr with those retrieved from other datasets for validation. Radio occultation data acquired from COSMIC were matched with the corresponding values from the closest SuGAr station. We found relatively good agreement, in spite of the scarcity of COSMIC data over Sumatra for the period of interest. We conducted a similar comparison with PWV values obtained from the visible spectra of OMI, and found reasonable agreement with SuGAr when PWV is below about 50 mm. The discrepancy at higher values is likely due to the high sensitivity of OMI to clouds.

A comparison followed with NVAP-M PWV. In this case, the spatial coverage was much larger than for the other two missions, but NVAP-M only has data from 1988 until 2009, so we compared average diurnal cycles in the same location as SuGAr stations from different time periods. Agreement between the two datasets was poorer than other two comparisons. The daily means diagnosed from NVAP-M have a much wider range than those from SuGAr, from which they diverged substantially for coastal stations on Sumatra. Moreover, amplitudes from NVAP-M were much larger than those calculated from SuGAr, especially for stations over Sumatra. We suspect the poor agreement between the two datasets is related to the coarse resolution of NVAP-M, which, combined with the

![Fig. 14](https://example.com/fig14.png)

**Fig. 14.** (clockwise from top left) Anomalies of PWV compared to the mean computed with WRF during the simulated period as a function of time and distance from the west coast of Sumatra across phases 1–8 of the MJO. Positive anomalies are drawn with continuous contours at an interval of 1 mm, while negative anomalies are drawn with a dashed–dotted contour line. The 0-mm value is marked with a thicker black line.
complex topography of Sumatra, could lead to large discrepancies.

We further analyzed the observed diurnal cycle with the aid of WRF. The model provides accessibility to nonobservables on a regular grid at higher resolutions than the data, opening up the possibility of gaining a deeper understanding of the underlying processes. Overall, the simulations capture the most pronounced features in the observed diurnal cycle for the purpose of this study. Specifically, although the simulated mean values are lower by a few millimeters for most stations, the simulated amplitudes aligned well with the observations, especially for stations on the offshore archipelago.

The phases of the diurnal cycles were also rather well captured by WRF: on average, values collocated with SuGAr stations over Sumatra had a single peak at about 1900 LST, while the averaged diurnal cycle for the locations offshore had double-peaked structure as SuGAr stations also exhibited. A closer examination of the island of Siberut led us to speculate that the double peak could be a combination of having stations along the coast of the smaller islands to the west of Sumatra and of having a coarse spatial resolution of 20 km: the peak at 0100 LST could be due to the offshore propagation of the moist convection developed over Sumatra, while the peak at 1300 LST could be a result of convection over Siberut.

We investigated how the MJO modulates the diurnal cycle of PWV for the period between 1 October 2011 and 1 April 2012. Starting from the OLR field computed through a WRF simulation for the time period of interest, we distinguished eight MJO phases, and we averaged diurnal cycles of PWV for each phase. Our results show that the largest values of PWV are obtained in the developing and active phase of the MJO, while the lowest PWV are seen during suppressed phases. Stations over the island of Sumatra experience large variations of amplitudes, with the highest during the active phases, and the smallest during the suppressed ones. Amplitudes seemed affected also for stations on the archipelago, although to a lesser extent given that these stations have smaller amplitudes. In the suppressed phase, there also seems to be a shift in the peak: for the stations

![Fig. 15. (a),(b) As in Fig. 8c, but for phases 4 and 8, respectively. (c),(d) As in Fig. 8d, but for phases 4 and 8, respectively.](image-url)
over Sumatra, the peak moves from 1900 LST in the active phase to 2200 LST in the suppressed phase. Similar changes seemed to happen for stations on the archipelago.

Acknowledgments. The data are available upon request. We are grateful to the Earth Observatory of Singapore at Nanyang Technological University and the Indonesian Institute of Science (LIPI) for operating the SuGAr network and sharing the ground GPS data, and to Ding Ma for providing the MATLAB code to diagnose MJO phases from model output. GT and ZM were supported by NOAA Climate Program Office Grant NA17OAR4310260 and a subcontract of NSF Grant 1649819 to the University of Hawaii. David Adams contributed to this study during a sabbatical at Rutgers University with funding from Programa de Apoyos para la Superación del Personal Académico de la UNAM (PASPA). Huiqun Wang is supported by NASA’s ACMAP program.

REFERENCES


