1. Introduction

In a stratified ocean, internal solitary waves (ISWs; internal solitons and nonlinear internal waves) occur when barotropic tidal currents interact with abruptly changing topography, such as sea mounts, underwater ridges and continental shelves. During the last several decades, ISWs have been ubiquitously observed in oceans worldwide via in situ measurements (Gargett 1976; Osborne and Burch 1980; Apel et al. 1985a; New and Pingree 1992; Duda et al. 2004; Alford et al. 2012) and in satellite images (Fett and Rabe 1977; Osborne and Burch 1980; New and Da Silva 2002; Jackson 2007; da Silva et al. 2011, 2012). Located in the northeastern Indian Ocean, the Andaman Sea is a marginal sea connected to the Bay of Bengal through multiple channels between the Andaman and Nicobar archipelagos (Fig. 1), and astronomical tides enter the Andaman Sea through those narrow, shallow channels, which are extraordinarily favorable for spawning internal waves. By analyzing satellite images, ISWs have been found to be emitted from those channels, and most of them propagate eastward and are distributed widely across the Andaman Sea (Apel et al. 1985b; Raju et al. 2020). Note that the multi-ISW-source Andaman Sea is different from most ISW hotspots, such as the northern South China Sea (SCS; Zhao et al. 2004; Alford et al. 2015) and the Sulu Sea (Apel et al. 1985a), where ISWs are mainly emitted from one source.

In the past, field observations of ISWs in the Andaman Sea were limited (Perry and Schimke 1965; Osborne and Burch 1980; Hyder et al. 2005), and the basic characteristics of ISWs were preliminarily revealed by these limited observations. These waves have been found to occur during semi-diurnal tidal periods and usually appear as packets of 5 or 6 solitons, whose amplitudes can reach as much as 80 m. Nevertheless, those observations are sporadic and short term, so the spatial and temporal variations in ISWs in the Andaman Sea, such as monthly, seasonal and interannual variations, still remain largely unclear. Satellite images are widely used to investigate ISWs in the Andaman Sea (Osborne and Burch 1980; Alpers et al. 1997; Jackson et al. 2012; da Silva and Magalhaes 2016; Magalhaes and da Silva 2018; Magalhaes et al. 2020; Raju et al. 2020; Tensubam et al. 2021). It has been determined that the ISWs are emitted from five source regions, including four water channels in the western Andaman Sea and the continental shelf off the country of Myanmar (Raju et al. 2020). However, the underwater structures of ISWs cannot be observed through satellite imagery. Therefore, the similarities and differences in ISWs generated from different source regions are largely unclear. Based on the nearly 22-month-long observations from
two subsurface moorings, this paper focused on the spatio-temporal variability of ISWs in the Andaman Sea, and the ISWs generated from different source regions were compared using mooring observations.

The generation mechanism of ISWs in the Andaman Sea is poorly understood. One of the reasons is that ISWs in the Andaman Sea are generated from multiple source regions, making it difficult to comprehensively investigate their generation mechanisms. Recently, Raju et al. (2021) investigated the generation of ISWs in the shallow channel south of Car Nicobar in the Nicobar Islands. They found that the waves are generated through lee wave mechanism in the north of the channel but evolve from a long wave disturbance induced by upstream influences in the southern portion of the channel. The waves generated in the north and south of the channel merge together at some distances away from their generation sites, forming the ISWs emitted from the channel. Nevertheless, the generation mechanisms of ISWs in other source regions have not yet been investigated. By examining the most common mechanisms driving the generation of ISWs using mooring observations, the generation mechanism of ISWs in the Andaman Sea was investigated in this study. Here, we mainly focused on the ISWs emitted from the submarine ridge northwest of Sumatra Island and discussed their potential generation mechanism.

In the Andaman Sea, dynamic processes are extremely abundant. Monsoons in the Andaman Sea, as in other parts of Indian Ocean, are prevalent, dominated by northeast monsoon in winter [November–January (NDJ)] and southwest monsoon in summer [May–July (MJJ)]. During the transition periods between monsoons, eastward propagating Kelvin waves are formed in the equatorial Indian Ocean. They can impinge upon the coast of Sumatra Island and one of the branches turns northward along the eastern rim of the Andaman Sea (Wyrtki 1973; Rao et al. 2016). Furthermore, the general circulation caused by monsoons, heat flux, and tides enters the Andaman Sea mainly through the Preparis Channel in the northwest and leaves mainly through the Great Channel in the southwest (Wyrtki 1961; Rizal et al. 2012). In addition, it is widely understood that mesoscale eddies are active in the Bay of Bengal (Hackert et al. 1998; Cheng et al. 2018). The Andaman Sea, located to the east of the bay, is a major source region of mesoscale eddies (Chen et al. 2012; Cui et al. 2016). Due to these active dynamic processes in the source region as well as the propagation pathways of ISWs, our study found that the generation and evolution of ISWs in the Andaman Sea were remarkably modulated by such background processes.

The remainder of this paper is organized as follows. In section 2, we introduce the data and methods used in this paper. The observational results, including wave appearances, wave arrivals, occurrence frequencies and wave intensities, are presented in section 3. In section 4, the generation mechanism of the observed ISWs and the factors leading to their temporal variability are discussed. The concluding remarks are presented in the last section of the paper.

2. Data and methods

a. Data

Two subsurface moorings (red stars in Fig. 1) referred to as QB1 and QB2 were deployed in the southern Andaman Sea from December 2016 to September 2018. At each mooring, an upward-looking and a downward-looking 75-kHz long-range acoustic Doppler current profilers (ADCPs) were mounted at a nominal depth of 500 m. Those ADCPs stopped working in July 2018 when the battery power was exhausted. The bin number and bin size of the ADCPs were 37 and 16 m, respectively, and they were set to narrow-bandwidth and low-power mode with sampling intervals of 3 min. The ping number of each ensemble was 7, and thus the standard deviation of velocity measurement was 2.86 cm s$^{-1}$. Temperature–salinity (T–S) chains consisting of dozens of temperature loggers and several conductivity–temperature–depth (CTD) recorders were mounted at each mooring to monitor the upper-layer T–S with a sampling interval of 3 min. All instruments operated adequately throughout their observation periods. The detailed configurations of the moorings are listed in Table S1 in the supplemental information.

Raw beam-coordinate velocities observed by the ADCPs were converted to Earth coordinates using the compasses
inside each ADCP, which were calibrated prior to deployment. Then the velocities were vertically interpolated into standard grid with an interval of 5 m, and the measurements in the upper 50 m were discarded due to the contamination of reflected beam signals from sea surface.

b. Wave identification

To determine the arrival times of the ISWs during the nearly 22-month observation period, we first calculated the depth of the 15°C isotherm throughout the whole observation period, and the isotherm was run through a high-pass filter at 4 h. Extreme spikes whose magnitudes were larger than 13 m were regarded as the initial representative ISWs. After that, we examined the isotherm and current velocity each day to verify the accuracy of the screening results and supplemented those ISWs that were missed in the digital filtering. In this way, we ultimately ascertained the arrival times of all ISWs.

3. Observation results

During the nearly 22-month observation period, 895 and 676 ISW episodes were captured at moorings QB1 and QB2, respectively, and the observed waves demonstrated notable spatial and temporal variability. Although many mode-2 ISWs were also captured by the moorings, they were not considered in this study.

a. Wave appearances

A MODIS image acquired at 0650 UTC 2 March 2018 was demonstrated (Fig. 2a), in which several ISW packets were evident and were numbered as packets 1–6. The underwater structures of packets 1 and 2 were observed by the moorings (Figs. 2b and 2c). Clearly, the waves arrived first at QB2 and then reached QB1 after traveling for another several hours. The time interval between packets 1 and 2 was about 12 h, and their separation distance illustrated in the satellite image was about 100 km, both demonstrating the characteristic of semi-diurnal motions. Moreover, the observed underwater structures of the two packets were very similar, which was in great contrast to the ISWs in the northern SCS, where ISWs also appeared twice a day but with different intensities and wave forms (Ramp et al. 2004, 2019). Based on the results of tidal analysis at the moorings (see appendix A), M2 tidal constituent was found to be the most prominent among all constituents. This could explain the similarity between packets 1 and 2 in the mooring measurements.

Clearly, the observed ISWs appeared mainly in the form of packets. For packets 1 and 2, there were about 5 solitons measured by QB1 and 3 solitons measured by QB2. During the whole observation period, ISWs that occurred in the form of packets accounted for 80% and 67% of the ISWs received by QB1 and QB2, respectively, and there was an average of 4.2 and 2.5 solitons in the packets received by QB1 and QB2, respectively. Generally, the leading soliton was the strongest among a packet, e.g., packet 1. However, for the packet 2 in Fig. 2b, the third soliton was the strongest among the packet. Throughout the whole observation period, wave packets whose strongest soliton was not in the leading position accounted for 40% and 33% at QB1 and QB2, respectively. Moreover, the strengths of ISWs at QB1 were comparatively larger. For example, the amplitude of packet 1 (the amplitude of the soliton...
that was the strongest in the packet) at QB1 was 75 m. Its horizontal kinetic energy (HKE) and available potential energy (APE) were 400 and 612 MJ m\(^{-2}\), respectively, with an APE-to-HKE ratio of 1.5 (the calculation methods of the energetics were shown in the supplemental information). The amplitude of packet 1 at QB2, however, was only 29 m. Its HKE and APE were 86 and 105 MJ m\(^{-2}\), respectively, with an APE-to-HKE ratio of 1.2. Specifically, the total energy (HKE plus APE) of packet 1 at QB1 was 5.3 times of that at QB2. A detailed analysis of wave amplitudes and energetics will be conducted in section 3d.

To calculate the propagation directions of packets 1 and 5 at QB1, their observed zonal and meridional currents were adopted (Fig. 3). First, the background currents half an hour prior to the arrival of ISW packets were removed. The direction of a packet was indicated by the core of water particles of the leading soliton in the upper layer, and the 10 water particles of the leading soliton with the fastest velocities were chosen as a representation of the core. The average zonal and meridional velocities of these 10 water particles were then used to calculate the direction of the packet. According to this method, packet 1 was found to propagate northeastward in a direction of 53° (clockwise from due north) at QB1. Tracing back its wave crest, we found these northeastward ISWs were emitted from the submarine ridge northwest of Sumatra Island, i.e., the location S in Figs. 1 and 2a. They were thus denoted as S-ISWs, which accounted for 82.7% of all episodes observed during the ADCP-occupied period at QB1, and their average propagation direction was 63 ± 14°. In addition, 17.3% ISWs at QB1, as packet 5, propagated southeastward, with an average direction of 114 ± 13°. These waves could be traced back to the submarine ridge south of Car Nicobar, i.e., the location

![Fig. 3. The underwater structures of the packets (left) 1 and (right) 5 observed at QB1. The upper and bottom panels indicate zonal and meridional velocities, respectively, and the gray lines denote isotherms with a vertical interval of approximately 100 m.](image)

![Fig. 4. (a) The waveform of the packet 1 observed at QB1 (15°C isotherm; black curve) and predicted using the dnoideal solutions to the KdV equation (gray curve). (b) As in (a), but for a C-ISW occurred at 1118 UTC 30 Jan 2017.](image)
Fig. 5. (left) Stack plots of the meridional velocity at QB1 during 25 Jan and 5 Feb 2017. The black lines show the 15°C isotherms and the arrival times of the S- and C-ISWs are indicated by red and purple arrows, respectively. (right) The corresponding zonal barotropic tidal currents predicted from TPXO 7.2 at the location S.

C in Figs. 1 and 2a, and were denoted as C-ISWs. Based on this classification, all waves observed at QB2 belonged to S-ISWs, propagating southeastward with an average direction of 111 ± 12°.

To characterize the waveforms of S- and C-ISW packets observed at QB1, the dnoideal solutions to the Korteweg–de Vries (KdV) equation were used. We followed the methods given by Huang et al. (2016) to calculate the dnoideal solutions and the details of the calculation were demonstrated in the supplemental information. For packet 1 (i.e., a S-ISW packet), its waveform could be generally well predicted by the dnoideal solution (Fig. 4a). The observed distances between successive solitons in the rear of the packet, nevertheless, were wider than those predicted by the dnoideal solution, which was likely due to the modulations of the bumpy topographies along the propagation path (Xie et al. 2019). In the dnoideal solution, the evolution time for the packet was 24 h, roughly two semiurnal tidal periods. This could also be indicated from the distribution of statistical crests (Fig. 1), in which the distance between QB1 and location S was roughly two mode-1 wavelengths for semiurnal motions. Because packet 5 (i.e., a C-ISW packet) observed at QB1 was very weak, a stronger C-ISW packet observed at 1118 UTC 30 January 2017 was used to fit the dnoideal solution (Fig. 4b). It was with an evolution time of 20 h that the dnoideal-predicted results could well match the observed waveform. However, the distance between QB1 and location C was more than three mode-1 wavelengths for semiurnal motions, indicating that the waves needed more than 36 h to travel such a distance. This suggested that the C-ISW packets received at QB1 were not formed immediately at the source region but during the propagation. This was probably because QB1 was located near the southern end of C-ISW crests but close to the center of S-ISW crests, and the evolution processes of ISWs
near crest ends were likely different from those close to crest center.

**b. Wave arrivals**

The timing of wave arrivals of S- and C-ISWs can be analyzed in greater detail using daily stack plots of meridional velocity at QB1. The velocities measured during 25 January and 5 February 2017 are shown in Fig. 5, and the 15°C isotherm is superimposed. S- and C-ISWs, appearing with northward and southward velocities in the upper layer, respectively, are indicated by red and purple arrows, respectively. Evidently, the S-ISWs arrived at QB1 with notable regularity. They occurred twice and arrived approximately one hour later each day, roughly following the M2 tidal period. However, on 30 January, the S-ISWs arrived at approximately the same time as they did one day before, and the first S-ISW packet measured on 1 February even arrived earlier than it did one day before. According to the observed background fields at the moorings (Fig. 6), there was an evident surface intensification of currents during this period, flowing westward at QB1. By examining the sea level anomalies (SLA) during this period (Fig. S1 in supplemental information), we found a cyclonic eddy around the moorings. This eddy gave rise to the strong westward background currents at QB1 and likely led to the observed irregular wave arrival patterns of S-ISWs. The occurrence of S-ISWs followed the spring–neap tidal cycle, and waves appeared nearly throughout the entire cycle. On the other hand, the C-ISWs first appeared on 26 January, and eight days later, no C-ISWs were observed, indicating that they were mainly concentrated during the period of spring tide. Similar to the S-ISWs, the C-ISWs also arrived following the M2 tidal period, and their arrival times were probably modulated by the mesoscale eddy likewise. In addition, C-ISWs often occurred as a single soliton, but S-ISWs always occurred as packets. During the whole ADCP-occupied period, there was a total of 3193 and 255 solitons belonging to S- and C-ISWs, respectively,
among which S- and C-ISWs occurring in the form of packets accounted for 87% and 43%, respectively. More comparisons between S- and C-ISWs were displayed in Table 1.

As demonstrated in Fig. 5, the C-ISWs generally arrived 1–2 h earlier than the S-ISWs. However, on 19 April 2018, the packets of S- and C-ISWs arrived at QB1 at approximately the same time, and the two packets were mixed together (Fig. 7). The S- and C-ISWs are indicated by red and purple arrows in Fig. 7a, respectively, and the wave type was mainly identified by the propagation directions of the solitons. In the rear of the mixed packet, although the waveforms (Fig. 7a) of the solitons were evident, their propagation directions could not be identified easily. The cores of the upper-layer horizontal velocities of these solitons were ambiguous, especially in terms of their meridional velocities. For the direction-identified solitons in the front of the mixed packet, the S- and C-ISWs arrived alternatively at approximately half-hour intervals. This alternating occurrence made the horizontal velocities of S- and C-waves extremely disordered (Figs. 7b,c), which did not follow the mode-1 structure in the vertical direction and probably resulted in the chaos of the horizontal velocities seen in the rear of the mixed packet. In addition, the amplitudes of the S-ISWs in the mixed packet were larger than those of the C-ISWs, with amplitudes of the leading solitons of the S- and C-ISW packets being as high as 92 and 41 m, respectively.

Similarly, we reported the timing of wave arrivals at QB2 using daily stack plots of zonal velocity during 10–21 January 2017 (Fig. 8). As demonstrated before, the ISWs captured at QB2 were all S-ISWs, and they were found to be concentrated during the period of spring tide. There were also two episodes each day, roughly following the M2 tidal period, and they appeared more regularly than those at QB1. This was likely because QB2 was located closely to the coast and hence the background processes there, as shown in Fig. 6, were steadier than those at QB1.

c. Occurrence frequencies

During the whole observation period, S-ISWs occurred an average of 1.2 and 1.0 times per day at QB1 and QB2, respectively, indicating that S-ISWs occurred more frequently in the north of their crests than in the south. More comparisons between the S-ISWs at QB1 and QB2 were shown in Table 1. The occurrence frequency of C-ISWs at QB1 was much lower, with an average frequency of 0.2 times per day. To further investigate the temporal variations in wave occurrence, the monthly averaged occurrence frequency of ISWs is displayed in Fig. 9.

The monthly averaged occurrence frequency of S-ISWs varied significantly at QB1 (Fig. 9a). In the first half of 2017,
it peaked in January with a magnitude of 1.6 episodes day\(^{-1}\) and gradually descended to 0.9 episodes day\(^{-1}\) in June. The variation in the latter half of 2017 was not that regular and showed an increasing tendency from August to November. Seasonally, the occurrence frequency in spring [February–April (FMA)] was relatively higher than that in other seasons. This was in stark contrast to other ISW hotspots, such as the northern SCS (Zheng et al. 2007) and the Bay of Biscay (New and Da Silva 2002), where the ISW occurrence frequency peaked in summer. The interannual variability was generally subtle, and the variation tendency in the first half of 2018 was close to that in 2017, with a maximum year-to-year increase in the occurrence frequency of 43\% in April.

The monthly averaged occurrence frequency of S-ISWs at QB2 (Fig. 9b) showed more subtle variability than that at QB1. In 2017, the monthly averaged occurrence frequency generally varied very little, rising and falling at approximately 1.0 episodes per day. Seasonally, the occurrence frequency also peaked in spring in 2017, while in 2018, the occurrence frequency in spring was relatively smaller than that in summer. In contrast, the interannual variation in the occurrence frequency of S-ISWs at QB2 was more obvious than that at QB1, with a maximum year-to-year increase in the occurrence frequency of 78\% in June.

The C-ISWs observed at QB1 were evidently concentrated during the period between January and March 2017 (Fig. 9c), during which the observed waves accounted for 44\% of the C-ISWs observed throughout the ADCP-occupied period. In 2017, the monthly averaged occurrence frequency peaked in January with a magnitude of 1.0 episodes per day, and in several months, no C-ISWs were observed at all. Seasonally, the occurrence frequency in spring was much higher than that in other seasons, and interannual variation was evident, with a maximum year-to-year decrease in the occurrence frequency of 90\% in January.

d. Wave intensities

Wave amplitude is often used to characterize wave intensity. During the whole ADCP-occupied period, the average

FIG. 8. As in Fig. 5, but for the zonal velocity at QB2 from 10 to 21 Jan 2017.
amplitude of the S-ISWs at QB1 was $56.3 \pm 22.3 \text{ m}$, 36% larger than that at QB2 with a magnitude of $41.4 \pm 12.6 \text{ m}$. This indicated that the northern portion of the S-ISWs was generally stronger than the southern portion. By examining the semidiurnal energy fluxes emitted from the Great Channel into the Andaman Sea (Fig. 8 in Mohanty et al. 2018), the semidiurnal energy propagated mainly toward the northern portion of S-ISW crests, which probably explained the stronger S-ISWs at QB1, as well as their high occurrence frequencies.

For the S-ISWs at QB1 (Fig. 10a), waves with amplitudes larger than 80 m accounted for 13.3% of all episodes. This indicated that the ISW amplitude in the Andaman Sea could be much larger than that estimated in early studies (Perry and Schimke 1965; Osborne and Burch 1980; Hyder et al. 2005) and comparable to that seen in the SCS (Alford et al. 2010; Huang et al. 2017). The variability of the wave amplitude was similar to that of the occurrence frequency, showing generally descending and ascending tendencies in the first and latter half of 2017, respectively. The monthly averaged amplitude (red circles with white filling) in 2017 peaked in February and reached the minimum in June, with magnitudes of 67.2 and 42.3 m, respectively. Seasonally, the averaged amplitudes (black circles with blue filling) in spring and autumn [August–October (ASO)] were comparatively stronger than those in summer and winter. This differs from other ISW hotspots, such as the SCS and Washington continental slope, where ISWs are strong in summer and weak in winter (Zhang et al. 2015; Huang et al. 2021, manuscript submitted to Prog. Oceanogr.). Reasons leading to this particular seasonal variation in the Andaman Sea will be discussed in section 4b. Moreover, the interannual variations in the amplitudes of the S-ISWs at QB1 were generally subtle, with a maximum year-to-year increase in the monthly averaged amplitude of 43% in April.

The variability in the amplitudes of S-ISWs at QB2 (Fig. 10b) was similar to that at QB1, and the amplitudes were also strong in spring and autumn. However, sometimes the monthly averaged amplitude at QB2 showed an antiphase tendency compared with that at QB1, as was seen in April and September 2017. The monthly averaged amplitude in 2017 peaked in April and reached the minimum in October with magnitudes of 51.4 and 34.6 m, respectively, which was also different from that seen at QB1. Generally, the variation in the amplitudes of S-ISWs at QB2 was relatively subtler than that at QB1, and the largest interannual difference in monthly averaged amplitude was only 14.1% in February.

For the C-ISWs observed at QB1 (Fig. 10c), their averaged amplitude during the whole ADCP-occupied period was $41.8 \pm 15.6 \text{ m}$, 26% smaller than the averaged amplitude of the S-ISWs at QB1 and close to the averaged amplitude of the S-ISWs at QB2. Excluding the period between January and March 2017, the waves were intermittently distributed during most of the ADCP-occupied period, and the variation in the monthly averaged amplitude was generally subtle. Based on

![Fig. 9. The monthly averaged occurrence frequency of (a) S-ISWs at QB1, (b) S-ISWs at QB2, and (c) C-ISWs at QB1. The dark and light gray bars indicate the results in 2017 and 2018, respectively.](image-url)
the limited C-ISWs measured, we found that the wave amplitude in spring was obviously larger than that in other seasons, and the interannual variation was generally subtle, with a maximum year-to-year increase in monthly averaged amplitude as high as 48% in April.

The energies of ISW packets, i.e., the summation of HKE and APE, were also calculated to indicate the variation of wave intensities (Fig. 11). In general, the energies showed nearly the same temporal variation tendency as wave amplitudes. Over the whole ADCP-occupied period, the average energy of the S-ISWs at QB1 was 1.18 ± 1.16 GJ m\(^{-2}\), 3.3 times of that at QB2 with a magnitude of 0.36 ± 0.35 GJ m\(^{-2}\). As shown in Table 1, the average APE of S-ISWs was larger than the average HKE at both QB1 and QB2, with an APE-to-HKE ratio of 2.3 and 2.6, respectively. For the C-ISWs at QB1, their average energy was 0.38 ± 0.53 GJ m\(^{-2}\), 67.8% smaller than that of the S-ISWs at QB1. It was worth mentioning that the intensity of the C-ISWs observed at QB1 was close to that of the S-ISWs observed at QB2, no matter for energetics or wave amplitudes. As suggested by Ramp et al. (2019), ISW intensity was nonuniform along its crest and became weaker toward the two ends of the wave crest. In our study, because QB1 and QB2 were located close to the southern end of C- and S-ISW crests, respectively, it probably led to the similarity of wave intensities between the C-ISWs at QB1 and S-ISWs at QB2. However, the detailed reasons leading to the similarities and differences between S- and C-ISWs were comprehensive, which should include both the generation and evolution processes.

4. Discussions

a. Generation mechanism of the observed ISWs

Among the two types of ISWs observed, the generation mechanism of C-ISWs has previously been investigated by Raju et al. (2021) using numerical simulations. To explore the generation mechanism of S-ISWs, we first show an Envisat Advanced Synthetic Aperture Radar (ASAR) image acquired at 0335 UTC 28 January 2005 around S (Figs. 12b and 12c). The eastward S-ISWs are obvious in this image. The zonal barotropic tidal current at S during 27–28 January 2005 predicted from TPXO 7.2 (Egbert and Erofeeva 2002) is shown in Fig. 12a. The moment at which the ASAR image is acquired is indicated as cyan triangle, close to the eastward current peak at 0230 UTC. Given that the generated S-ISW packet shown in the ASAR image is close to S (within 20 km), it needs less than 3 h for the packet to travel such a distance supposing a propagation speed of 2 m s\(^{-1}\). Accordingly, the generation of S-ISWs is closely related to the eastward barotropic tidal current at source region. This implies that the S-ISWs may not be generated through lee wave mechanism, in which the generation of S-ISWs should be associated with westward barotropic tidal current, as suggested by Maxworthy (1979).

To further verify this speculation, we try to find the phase of barotropic tidal current at which S-ISWs are generated based on the mooring measurements at QB2. Each S-ISW is supposed to travel from S directly to QB2, and the
propagation speeds along the path are calculated according to KdV theories based on the measured stratification and wave amplitude. The details about how to determine the propagation speeds are given in the supplemental information. Waves occurring during the period from 5 to 18 July 2017 are selected because there were no active mesoscale eddies in the southern Andaman Sea during this period. The calculated corresponding relationship is illustrated in Fig. 13. The upper panel shows the arrival times of waves and their corresponding amplitudes, and the lower panel shows the barotropic tidal currents at S. The wave arrival time is shifted early according to the propagation time of each S-ISW from S to QB2, and the shifted arrival time is connected to the barotropic tidal currents with dashed lines. For the 19 chosen episodes, the average travel time is 15.1 h, and all waves are associated with eastward tidal currents. This further verified the argument that S-ISWs are not generated through lee wave mechanism.

Another common mechanism for the generation of ISWs is internal tide steepening, through which the radiated internal tides gradually steepen and disintegrate into ISWs (Lee and Beardsley 1974; Gerkema and Zimmerman 1995). Here, we use the KdV equation (see appendix B) to simulate the evolution of internal tides along the section connecting S and QB2. We suggest that an M_2 internal tide with an amplitude of 14 m enters the simulation domain from the west, which resembles the situation in which internal tides are generated at S and then propagate eastward toward QB2. The simulation result is illustrated in Fig. 14. The initial sinusoidal wave experiences nonlinear effects during propagation and gradually steepens and disintegrates into solitary waves. It occurs at a distance of approximately 90 km away from S where the tidal trough begins to break (red dashed line). However, the S-ISW crests can appear at a distance within 20 km away from S (Fig. 12b), much smaller than the estimated breaking distance of the tidal trough. Thus, the generation mechanism of S-ISWs is likely not internal tide steepening.

New and Pingree (1990, 1992) argued that when an internal tidal beam hits a thermocline, it creates an interfacial wave, which may evolve into ISWs under the effect of nonlinearity. In this study, the M_2 internal tidal beams emitted from the source region along the section connecting S and QB2 are calculated (Fig. 15). The slope of the tidal beams is calculated via

\[ s_m = \sqrt{(\omega^2 - f^2)(N^2 - \omega^2)}, \]

where \( \omega \) is the M_2 tidal frequency and the buoyancy frequency \( N \) along the section is calculated based on data from the World Ocean Atlas 2013. The thermocline is assumed to be at a depth of 120 m, which is the average thermocline depth at QB2. The results reveal two upward tidal beams emitted from the critical slope of the ridge. The eastward tidal beam crosses the thermocline for the first time when propagating upward. After being reflected by the sea...
surface, the tidal beam crosses the thermocline for a second time when propagating downward. The two crossing positions are horizontally 23 and 47 km away from S, respectively, and the leading S-ISW crest shown in Fig. 12 appears closely to the position where the tidal beam first crosses the thermocline. However, Akylas et al. (2007) found that the beam-induced oscillation cannot form ISWs immediately after the impingement, and it still needs tens of kilometers for the oscillation to evolve into ISWs. In this case, S-ISWs can hardly be formed at a distance about 20 km away from source region. Therefore, we suggest that the S-ISWs are likely not generated through an internal wave beam mechanism either.

As shown in Figs. 12 and 13, the S-ISWs are associated with eastward tidal currents at source region, which meets the conditions of the recently proposed generation mechanism of ISWs, the internal tide release mechanism (Buijsman et al. 2010). In this mechanism, strong tidal currents create an elevation wave with a high energy density on the upstream side of a sill, and as soon as the currents slacken, the wave is released upstream and evolves into ISWs. Note that the southern portion of C-ISWs are generated due to upstream influences (Raju et al. 2021), which is similar to that of the internal tide release mechanism; Raju et al. (2021) further found that with stronger tidal currents, ISWs are formed closer to the source region. However, the barotropic tidal current at S is not as strong (about 0.2 m s$^{-1}$). Thus, attempts to ascertain the relationship between the closer formation of S-ISWs and the strength of source-region barotropic tidal currents yield undesired results. However, da Silva et al. (2015) found that ISWs that are generated through an internal tide release mechanism are formed closer to the source region when adding a background current around the source region propagating against the direction of ISWs. In the Andaman Sea, the general circulation mainly flows out through the Great Channel, forming a strong westward background current at S with a magnitude reaching as high as 0.2 m s$^{-1}$ (Rizal et al. 2012). This is further confirmed by the westward upper-layer background currents observed at QB2 (Fig. 6), which is located not far from S. Accordingly, we suggest that the S-ISWs were likely generated through an internal tide release mechanism, and the generation process is described as follows. Around the submarine ridge northwest of Sumatra Island, the westward barotropic tidal current accumulates energy on the eastern side of the ridge and creates an elevation wave, which is released eastward when the current slackens and gradually evolves into ISWs. When the outflow of general circulation around S is strong, the accumulated energy on the eastern side of the ridge is enhanced, and S-ISWs, as shown in Fig. 12, are thus formed closely to S.

b. Factors leading to the temporal variations in ISWs

In this subsection, the factors influencing the temporal variations in ISWs are discussed in different time scales,
ranging from hourly to yearly. According to the measurements, S- and C-ISWs appeared twice a day, corresponding to the period of semidiurnal tide, and they appeared in clusters in spring–neap tidal cycles. This indicated the vital role of astronomical tides in determining these short-period behaviors of ISWs. Meanwhile, the observed ISWs also showed long-period, i.e., monthly, seasonal and interannual, variability, but there were no obvious long-term variations in the barotropic tidal currents at source region (not shown). Accordingly, some other factors may play important roles in modulating the long-term behaviors of ISWs.

Stratification is one of the main factors that determine the behaviors of ISWs, and a strong shallow thermocline is favorable for the generation of ISWs (Shaw et al. 2009). In Fig. 16a, the observed thermocline depth over the whole observation period is demonstrated, and the results obtained from the two moorings showed similar variation tendencies. Generally, the thermocline depth showed a good correspondence with the occurrence frequencies and wave intensities. Over the whole observation period, there were evident subseasonal signals in thermocline depth, appearing as remarkable troughs with periods ranging from 1 to 2 months, like in early May, June, and December 2017. Those troughs indicated the deepening of thermocline, which inhibited the formation of ISWs, as illustrated in Figs. 9–11. Meanwhile, the thermocline in spring was relatively shallower than that in autumn, and the thermocline in spring and autumn were both shallower than that in summer and winter. This was in good agreement with the observation results of ISW behaviors, i.e., the shallower thermocline in spring and autumn leading to larger wave intensity and higher occurrence frequency. Moreover, the interannual difference of thermocline depth was modest, which led to the overall similarity of wave behaviors between 2017 and 2018. In summary, the long-period variations of ISW behaviors can be mostly explained by the variations in thermocline.

A variety of elements can affect the stratification in the Andaman Sea, like monsoons, remote wind forcing from equatorial Indian Ocean and mesoscale eddies (Rao and Sivakumar 2000; Girishkumar et al. 2013; Cheng et al. 2018). First, the local wind speed and wind vectors in the southern Andaman Sea were shown (dark gray curve and arrows in Figs. 16b and 16c, respectively). Southwesterly summer and northeasterly winter monsoons were evident, being prevalent during June–September and November–February, respectively, as revealed by former study (Shankar et al. 2002). The summer and winter monsoons could lead to respectively high and low sea levels in the southern Andaman Sea (Shankar et al. 2002). There was a correlation coefficient of 0.72 between the local zonal wind speed and sea surface height anomaly (SSHA; black line in Fig. 16b) in the southern Andaman Sea, when the SSHA lagged the zonal wind speed by 30 days, which was roughly the period of Ekman pumping penetrating from sea surface to the thermocline or vice versa. Since the positive and negative SSHA was closely connected to the deepening and shallowing of thermocline, respectively (Girishkumar et al. 2013), the local wind forcing could generally explained the fact that the observed thermocline in early spring (i.e., in February and March) was shallower than that in autumn.

The peaks of eastward equatorial winds (light gray curve in Fig. 16b) during monsoon transition periods (in early April, late May, and November 2017) could form Kelvin waves. These waves propagated eastward and impinged
upon the coast of Sumatra Island, bifurcating into two branches propagating northward and southward, respectively (Wyrtki 1973; Rao et al. 2016). The north branch propagated along the eastern coast of Andaman Sea, passing exactly by the moorings. It took about 3–4 weeks for the Kelvin waves to travel from the equator to southern Andaman Sea and the waves could deepen the thermocline during propagation (Rao et al. 2010), which corresponded well with the observed subseasonal deepening of thermocline (in early May, June, and December 2017). This likely gave rise to relatively deeper thermocline in summer and winter.

However, the relationship between the ISW behaviors and thermocline depth was not always as strong. For example, in 2017, the thermocline depth at QB2 peaked in March while the ISW intensity peaked in April, in which the wave intensity at QB1, in contrast, showed a remarkable drop. By scrutinizing the SLA (Fig. 17), we found a cyclonic mesoscale eddy formed around S at the end of March. The eddy could uplift the thermocline, which was favorable for the occurrence of thermocline peak in March. In the middle of April, an anticyclonic eddy was formed around S. The energies of ISWs in April could probably be scattered from their north portion to south portion by the anticyclonic eddy, as suggested by Xie et al. (2015). This likely explained the evident peak of ISW intensity in April 2017 at QB2. Similar situation was also identified in September 2017.

5. Conclusions

In this study, data collected during a nearly 22-month-long observation period from two moorings referred to as QB1 and QB2 were used to investigate the basic appearances and spatiotemporal variability of ISWs in the Andaman Sea. Two types of ISWs generated from respectively the submarine ridges northwest of Sumatra Island and south of Car Nicobar were captured by the moorings and were identified as S- and C-ISWs, respectively (Fig. 18a).

Both S- and C-ISWs were observed at QB1. According to the dnoideal solutions to the KdV equation, the observed S-ISW packets were found to be formed near their source region while the C-ISW packets were formed during the propagation toward QB1. The S- and C-ISWs that occurred in the form of packets accounted for 87% and 43%, respectively, and they generally arrived at QB1 separately within an interval of several hours. However, sometimes the solitons of the S- and C-ISW packets arrived in an alternating fashion and formed a mixed packet, leading to chaos of horizontal velocities. On average, the S- and C-ISWs occurred 1.2 and 0.2 times per day, respectively, and their average amplitudes (energies) were 56.3 m (1.18 GJ m$^{-1}$) and 41.8 m (0.38 GJ m$^{-1}$), respectively. This indicated that the observed S-ISWs occurred much more frequently and were generally stronger than the C-ISWs.

The ISWs in the Andaman Sea demonstrated obviously spatial variability. The simultaneous observations of moorings QB1 and QB2 along the crests of S-ISWs revealed an average of 4.6 and 2.5 solitons in the S-ISW packets received by QB1.
and QB2, respectively, which was schematically shown in Fig. 18b. In general, S-ISWs arrived at QB2 following semidiurnal tidal period with remarkable regularity, but the waves at QB1 often arrived early or late, likely due to the influences of background processes. On average, the S-ISWs at QB2 occurred 1.0 times per day, and their average amplitude (energy) was 41.4 m (0.36 GJ m$^{-2}$), indicating that the S-ISWs on the north side occurred more frequently and were generally stronger than those on the south side.

The temporal variations in ISWs in the Andaman Sea were complicated. The waves occurred twice a day and appeared in the spring–neap tidal cycle, indicating the significant role of astronomical tides in determining these short-period characteristics. Moreover, the monthly and seasonal variations in the ISWs were evident, characterized by relatively stronger S-ISWs in spring and autumn, while the interannual variations were generally subtle. These long-period, i.e., monthly, seasonal, and interannual, variations in S-ISWs showed good correspondence with the variability in stratification, which could be modulated by the monsoons, the Kelvin waves resulted from the winds in equatorial Indian Ocean and the mesoscale eddies in the Andaman Sea.

The generation mechanism of the observed ISWs is discussed in this paper. From careful analyses based on the long-term mooring measurements, we suggest that the observed ISWs are likely generated through an internal tide release mechanism, which is schematically depicted in Fig. 18b. The generation process of S-ISWs, which is intensively investigated in this paper, is described as follows. Westward barotropic tidal currents accumulate energy and create an elevation wave on the eastern side of the submarine ridge northwest of Sumatra Island. When the current slackens, the accumulated energy is released eastward, and the elevation wave gradually evolves into ISWs. In addition, the general circulation in the Andaman Sea flows out through the Great Channel, forming a strong westward background current around the source region. This background current enhances the accumulated energy on the eastern side of the ridge, and ISWs are formed closer to the source region.

Acknowledgments. We acknowledge the use of imagery from NASA’s Worldview application, and this study has been conducted using E.U. Copernicus Marine Service Information. This paper was supported by the National Natural Science Foundation of China (Grants 91858203, 91958205, 41976008, 41676011 and 41706005), the National Key Research and Development Program (Grant 2016YFC1402605), the 2019 Research Program of Sanya Yazhou Bay Science and Technology City (Grant 120LH018), the Foundation for Innovative Research Groups.
of the National Natural Science Foundation of China (Grant 41521091) and the Fundamental Research Funds for the Central Universities (Grant 201861006).

Data availability statement. ASAR images were downloaded from https://esar-ds.eo.esa.int/oads/access/ and the NCEP re-analysis data were provided by the NOAA/OAR/ESRL PSL (https://psl.noaa.gov/). If anyone needs the processed mooring data used in this study for research, please contact the corresponding author.

APPENDIX A

Tidal Current Ellipses at the Moorings

Based on the measurements from ADCPs, the current velocities in the upper 900 m were obtained. We averaged the velocities vertically to represent the local barotropic currents. Tidal ellipses of the barotropic currents were then analyzed in MATLAB using T_TIDE (Pawlowicz et al. 2002), and the elliptical properties at the two moorings were shown in Tables A1 and A2.

APPENDIX B

KdV Equation

The KdV equation,

$$\frac{\partial \eta}{\partial t} + c_0 \left( \frac{\partial \eta}{\partial x} + \alpha \frac{\partial^2 \eta}{\partial x^2} + \beta \frac{\partial^3 \eta}{\partial x^3} \right) = 0,$$

was employed to simulate the disintegration of linear internal tides into ISWs in the Andaman Sea. Here $\eta$ is vertical displacement, $c_0$ is linear phase speed, $\alpha$ and $\beta$ are nonlinear and dispersion parameters, respectively, and $t$ and $x$ are time and horizontal space coordinates, respectively. A two-layer assumption was adopted in the calculation, and the initial wave form was set to have an $M_2$ internal tidal profile, whose amplitude was reasonably selected such that the simulated ISW form could approximately meet the results of mooring observations. The environmental parameters were calculated as follows:

$$c_n = \sqrt{\frac{g \Delta \rho}{\rho_m} \frac{h_1 h_2}{h_1 + h_2}}.$$
Here $g$ is the gravitational acceleration; $h_1$ is the thickness of the upper layer and was set to 150 m, which is roughly the thermocline depth observed by the moorings; $h_2$ is the thickness of the lower layer, which was determined by the real topography measured along the propagation path; $\Delta \rho$ is the density difference between the upper and lower layers and was set to 2 kg m\(^{-3}\);

\[
\alpha = \frac{3h_1 - h_2}{2h_1h_2}, \quad \text{(B2b)}
\]

and

\[
\beta = \frac{h_1h_2}{6}. \quad \text{(B2c)}
\]

**TABLE A1.** The elliptical properties of the observed barotropic tidal currents at QB1. Here only the results with signal-to-noise ratio (SNR) larger than 5 were displayed.

<table>
<thead>
<tr>
<th>Tidal constituents</th>
<th>Period (h)</th>
<th>Major semiaxis (cm s(^{-1}))</th>
<th>Minor semiaxis (cm s(^{-1}))</th>
<th>Inclination (°)</th>
<th>Phase (°)</th>
<th>SNR</th>
</tr>
</thead>
<tbody>
<tr>
<td>O(_1)</td>
<td>25.82</td>
<td>0.15</td>
<td>0.03</td>
<td>29</td>
<td>278</td>
<td>20</td>
</tr>
<tr>
<td>P(_1)</td>
<td>24.07</td>
<td>0.18</td>
<td>-0.02</td>
<td>23</td>
<td>51</td>
<td>24</td>
</tr>
<tr>
<td>K(_1)</td>
<td>23.93</td>
<td>0.37</td>
<td>0.09</td>
<td>21</td>
<td>210</td>
<td>121</td>
</tr>
<tr>
<td>N(_2)</td>
<td>12.66</td>
<td>0.62</td>
<td>-0.16</td>
<td>37</td>
<td>98</td>
<td>44</td>
</tr>
<tr>
<td>H(_1)</td>
<td>12.44</td>
<td>0.20</td>
<td>-0.08</td>
<td>57</td>
<td>219</td>
<td>6</td>
</tr>
<tr>
<td>M(_2)</td>
<td>12.42</td>
<td>3.01</td>
<td>-0.73</td>
<td>35</td>
<td>215</td>
<td>1170</td>
</tr>
<tr>
<td>MKS(_2)</td>
<td>12.39</td>
<td>0.25</td>
<td>-0.11</td>
<td>153</td>
<td>319</td>
<td>7</td>
</tr>
<tr>
<td>S(_2)</td>
<td>12.00</td>
<td>1.46</td>
<td>-0.32</td>
<td>30</td>
<td>333</td>
<td>239</td>
</tr>
<tr>
<td>K(_2)</td>
<td>11.97</td>
<td>0.33</td>
<td>-0.09</td>
<td>19</td>
<td>312</td>
<td>16</td>
</tr>
</tbody>
</table>
and \( \rho_m \) is the mean density of the fluid and was set to 1024 kg m\(^{-3} \).

REFERENCES


Hyder, P., D. R. G. Jeans, E. Cauquil, and R. Nerzic, 2005: Observations and predictability of internal solitons in the

<table>
<thead>
<tr>
<th>Tidal constituents</th>
<th>Period (h)</th>
<th>Major semiaxis (cm s(^{-1}))</th>
<th>Minor semiaxis (cm s(^{-1}))</th>
<th>Inclination (°)</th>
<th>Phase (°)</th>
<th>SNR</th>
</tr>
</thead>
<tbody>
<tr>
<td>( O_1 )</td>
<td>25.82</td>
<td>0.11</td>
<td>5.43 \times 10^{-4}</td>
<td>164</td>
<td>317</td>
<td>6</td>
</tr>
<tr>
<td>( P_1 )</td>
<td>24.07</td>
<td>0.23</td>
<td>0.01</td>
<td>176</td>
<td>95</td>
<td>28</td>
</tr>
<tr>
<td>( K_1 )</td>
<td>23.93</td>
<td>0.68</td>
<td>−0.04</td>
<td>177</td>
<td>249</td>
<td>208</td>
</tr>
<tr>
<td>( N_2 )</td>
<td>12.66</td>
<td>0.79</td>
<td>−0.12</td>
<td>172</td>
<td>22</td>
<td>46</td>
</tr>
<tr>
<td>( M_2 )</td>
<td>12.42</td>
<td>3.73</td>
<td>−0.38</td>
<td>176</td>
<td>144</td>
<td>888</td>
</tr>
<tr>
<td>( S_2 )</td>
<td>12.00</td>
<td>1.81</td>
<td>−0.10</td>
<td>176</td>
<td>250</td>
<td>208</td>
</tr>
<tr>
<td>( K_2 )</td>
<td>11.97</td>
<td>0.30</td>
<td>−0.10</td>
<td>8</td>
<td>72</td>
<td>6</td>
</tr>
</tbody>
</table>


