Synthetic Aperture Radar observations of resonantly generated internal solitary waves at Race Point Channel (Cape Cod)

J. C. B. da Silva¹ and K. R. Helfrich²

Received 1 July 2008; revised 22 August 2008; accepted 16 September 2008; published 20 November 2008.

[1] Synthetic Aperture Radar images revealed the two-dimensional propagation characteristics of short-period internal solitary waves in Race Point Channel in Massachusetts Bay. The images and in situ measurements of the flow in the channel are used to infer the likely generation mechanism of the waves. The solitary waves are generated during the ebb phase of the tide within the channel. On some occasions, two trains of internal waves are generated presumably at the same location but at slightly different phases of the ebb tide. The main characteristics of the (two-layer) flow are described based on the criticality of the Froude number. It is suggested that these two individual packets of waves result from flow passage through resonance (where the Froude number is one). One packet is generated as the flow passes through the transcritical regime during the acceleration phase of the (ebb) tidal current, and another packet is generated during the deceleration phase. Both packets propagate upstream when the tide slacks, but with slightly different propagation directions.


1. Introduction

[2] The generation of internal solitary waves in the ocean is a subject of active investigation and generally recognized as not well documented [Farmer and Armi, 1999]. High-resolution satellite images of the ocean have provided important insights to the study of the generation mechanisms of internal waves [e.g., Trask and Briscoe, 1983; New and da Silva, 2002; Zhao et al., 2004; Vlasenko and Alpers, 2005; Nash and Moum, 2005; da Silva et al., 2007]. In particular, Synthetic Aperture Radar (SAR) provides details of the two-dimensional spatial structure of horizontally propagating internal waves in the upper thermocline, which is difficult to obtain with other sensors. Satellite SARs have a quasi-instantaneous data-take character that lacks sufficient temporal resolution for studying the generation or evolution of a single internal wave packet. In most cases, however, the generation process repeats itself almost identically with tidal periodicity, so that if one has a large enough data set covering different phases of the complete (semi-diurnal) tidal cycle, it is possible to investigate and reconstruct the propagation “history” of waves generated by a given source. In fact it is possible, as we will show in this article, to identify the approximate position and time of internal wave generation using SAR and conveniently located current meter measurements.

[3] Massachusetts and Cape Cod Bays are situated in the western Gulf of Maine, but are partially isolated from the rest of the Gulf by Stellwagen Bank, which rises to within 30 m of the surface (see Figure 1). There are silled channels on either side of Stellwagen Bank. North Passage, the channel between Cape Ann and Stellwagen Bank, has a sill depth of 60 m, while South Passage, the channel between Race Point and Stellwagen Bank, has a sill depth of 50 m. The deepest portion of the Bays is Stellwagen Basin, with typical depths of 80 m. The North Passage has rough, irregular topography, while Cape Cod Bay and the South passage (also known as Race Point Channel) have smooth bottom topography. Trask and Briscoe [1983] analyzed the horizontal characteristics of the internal waves in Stellwagen Basin that were first reported by Halpern [1971]. Their study was based on the short-lived SEASAT SAR mission (the satellite’s radar worked for only three months, in 1978, and they only analyzed 8 images), and they identified Stellwagen Bank as the internal wave generation site. More recent studies of internal waves in Massachusetts Bay have also used satellite radar images [e.g., Butman et al., 2006], but again have concentrated on waves generated at Stellwagen Bank. In this article we revisit SAR image archives of Massachusetts Bay and the Gulf of Maine. We show and discuss, for the first time, internal solitary waves emanating from Race Point Channel and propagating into Cape Cod and Massachusetts Bays.

[4] Many environmental fluids are naturally forced close to a linear resonance condition. In the ocean, stratified water moves relative to topography with speeds that sometimes are close to the phase velocity of long, linear internal waves on the given density distribution. When this resonance occurs for tidal or long-shore flows over ridges for example,
upstream-propagating internal solitary waves have been observed [Cummins et al., 2003; Bogucki et al., 1997]. In such systems, upstream-propagating internal waves are formed when the current speeds are within a finite interval of the limiting long-wave phase speed of the system. This interval is called the transcritical regime and is generally described in terms of a Froude number. The Froude number, \( F = U/c \), is defined as the ratio of a current velocity, \( U \), to the phase speed, \( c \), of linear long waves in the ambient stratification. The flow is said to be critical when \( F = 1 \). For subcritical flows below the transcritical band \( (F < F_1 < 1) \) a dispersive wave train is formed in the lee of the forcing. For supercritical conditions \( (F > F_s > 1) \) above the transcritical band a stationary interfacial disturbance is confined to the immediate location of the forcing. Here \( F_1 (< 1) \) and \( F_s (> 1) \) are, respectively, the lower and upper bounds of the transcritical regime. A series of experimental and numerical modeling studies [Grimshaw and Smyth, 1986; Melville and Helfrich, 1987; Wu, 1987; Stastna and Peltier, 2005] have shown that steady forcing (i.e., mean flow \( U \)) within the transcritical regime \( (F_1 < F < F_s) \) leads to the generation of an upstream-moving train of solitary (or cnoidal) waves of uniform amplitude. In the linear regime, flow at the long-wave phase speed leads to an unbounded wave amplitude since the upstream propagating disturbance is unable to propagate away from the topography. This linear singularity is resolved by nonlinear amplitude dispersion which allows a finite amplitude disturbance to propagate upstream, into the oncoming flow. Weak nonhydrostatic dispersion then balances the nonlinearity to give the solitary waves. Weakly nonlinear and weakly non-hydrostatic (i.e., Korteweg-de Vries models) theory shows that the bounds \( F_1 \) and \( F_s \) are functions of the forcing (i.e., topography, stratification, etc.) [Miles, 1986; Grimshaw and Smyth, 1986]. The width of the resonant band grows with increased forcing.

While laboratory experiments and non-linear numerical solutions have satisfactorily reproduced internal wave generation in the transcritical regime for steady forcing, the situation in nature is more complicated due to the temporal variability of the background flow \( U \). There have been some successful attempts to model upstream generation of solitary waves by time-dependent flows that pass through the transcritical regime [e.g., Redekopp and You, 1995; Wang and Redekopp, 2001; Grue, 2005; Torsvik et al., 2006]. When the residence time in the transcritical band, for example of tidal flow across an obstacle, is comparable to, or longer, than the timescale required for soliton formation in the steady problem, the process of upstream generation may be effective, though only a finite numbers of solitary waves, or a nonlinear wave packet will be generated. In principle, the process of resonant generation may occur both during the acceleration phase, when passing through the transcritical band from the sub- to supercritical regimes, and during the deceleration phase when the flow passes through resonance back to a subcritical state [Wang and Redekopp, 2001; Torsvik et al., 2006]. This process is likely to occur in shallow seas with strong tidal currents forced by topography.

We now briefly review previous studies about generation of internal waves at Stellwagen Bank, as this provides a close (and important) reference for the object of the present study: generation of internal waves at Race Point Channel. A substantial body of research has been devoted to the generation of internal waves at Stellwagen Bank. Lee and Beardsley [1974] proposed that the waves were generated during the flood tide by the nonlinear steepening and evolution of a front that would presumably be formed upstream of Stellwagen Bank due to “blocking” during the ebb tide. Haury et al. [1979] discussed a series of observations using acoustic backscattering systems, and ascribed the generation of internal waves to lee waves generated on the eastern side of Stellwagen Bank during ebb tide. When the ebb (offshore) tidal flow over Stellwagen Bank turns to flood (onshore), the lee waves would propagate into the Bay, steepen nonlinearly, and develop into packets of short-period internal waves. Using laboratory experiments, Maxworthy [1979] showed that the internal waves observed in Stellwagen Basin could indeed be generated when the tide changed from ebb to flood, releasing the lee wave formed on the ebb tide to propagate back shoreward and evolve into a rank-ordered packet of internal solitary waves. Chereskin [1983] showed that the flow during the ebb tide is supercritical, and the formation of a single massive depression in the thermocline downstream of the sill is consistent with Maxworthy’s results [Farmer and Smith 1980] also observed a single depression in Knight Inlet when the Froude number was supercritical with respect to all wave modes). Hibiya [1986] argued that Maxworthy [1979] and Farmer and Smith [1980] generation mechanism is not entirely correct since they assume the process is quasi-steady. If the process were quasi-steady, as Maxworthy proposed, the lee wave should stay over the sill even while the tidal flow is decreasing (which is not the case, as the lee wave begins to leave the sill and advances upstream evolving into a train of internal solitary waves). Instead,
Hibiya [1986, 1988] modeled the generation as a transient process in response to a time varying tidal flow. In his model, infinitesimal internal waves are continuously generated during the tidal cycle and constructively superimposed during the supercritical phase of the ebb tide. Such upstream-moving internal waves are generated over the Bank and carried downstream, where a disturbance grows until their linear phase speed equals the opposing barotropic flow. The disturbance is further amplified until the tide slackens, resulting in a large amplitude internal wave that is released upstream when the tide turns from ebb to flood. This model captured the essential aspects of the internal wave generation observed on Stellwagen Bank [Matsura and Hibiya, 1990].

The present work has two aims: firstly to report the observations of the newly found solitary wave (ISW) hotspot, motivate future modeling and field work in the study area, and to demonstrate the usefulness of SAR imagery to investigate the generation mechanism of ISWs in Race Point (RP) Channel. This article is organized as follows. Section 2 presents a SAR image data set from 1992 to 2005 corresponding to all available ERS and ENVISAT images of the summer period (July, August and September). We show some propagation characteristics of the internal waves generated near Race Point, using data obtained after Low Water at Boston. Section 3 discusses the criticality of the flow through RP Channel based on available measurements, and describes the main characteristics of the (two-layer) flow. On the basis of this criterion we further investigate in section 4 the generation of internal waves by examining SAR images near Low Water (at Boston), corresponding to the tidal phases when the waves are expected to be generated. This “second” image data set was obtained at a later stage of our study, and also includes images from June (1992–2007), when internal waves were also found to be ubiquitous in the study region. We then discuss the observations and show that, on some occasions, two trains of internal waves are generated in Race Point Channel presumably at the same location but at different phases of the same tidal cycle. It is suggested that these two individual packets of waves result from passage through resonance. One packet is generated when the flow passes through the transcritical regime during the acceleration phase of the (ebb) tide, and another packet is generated during the deceleration phase. Both packets propagate upstream into Massachusetts Bay when the tide slackens, but with slightly different propagation directions. The article concludes with a summary of the main findings.

2. SAR Observations

A revisit to SAR image data of Massachusetts Bay clearly revealed that Stellwagen Bank (SB) was not the only generation source within the Bay. Several images showed what appeared to be a nearly point source of wave generation near Race Point, emanating waves toward the inner Bay in an oblique direction relative to SB (see Figure 2). In all, over 50 SAR images (from ERS-1 & 2 and ENVISAT missions) were analyzed, revealing RP Channel as a very effective generation spot since practically all images showed internal wave trains that appeared to emanate from the Channel. Figure 2 presents a composite constructed by mapping the locations of the ISWs (only leading crests for each individual packet are marked, for simplicity) in all available ENVISAT overpasses, in relation to the bathymetry of the region. This clearly shows that there is a high level of organized ISW activity here. We note that, as in the study of Trask and Briscoe [1983], all these images are from the summer period, when the thermocline stratification is best developed.

Overall, Figure 2 shows several ISW packets which, because of their curvature (concave when viewed from the East), appear to be generally traveling toward the West (into Massachusetts and Cape Cod Bays). Moreover, there appears to be two main propagation directions that fit reasonably well most of the ISW observations, namely 250°T and 275°T (see Figure 2). In order to demonstrate the existence of two individual packets of ISWs traveling toward the West with slightly different directions, an example is shown in Figure 3 from the image on 21 August 1994 (15:26 UTC ERS-1 SAR). Figure 3a shows the location of the SAR full resolution image centered at 42.06°N and 70.28°E, to the West of RP (Figure 3b). Two well-developed wave packets are observed in Figure 3b with typical wavelengths of up to 0.4 km, and along-crest lengths of 10–15 km, which are similar to the properties of the ISWs generated at SB [Halpern, 1971; Trask and Briscoe, 1983; Scotti et al., 2007]. Figure 3 is typical of most of the SAR images of the area in that there are two distinct ISW packets (wave packets WP1 and WP2), which appear to be traveling in slightly different directions. These wave packets are separated by a distance of about 2 km (along their propagation direction), and frequently intersect producing interference patterns.

The propagation direction of such wave packets is assumed to be perpendicular to the orientation of the wave crests and such that the wavelength, or separation between the solitons, decreases from the front to the rear of the packet, according to non-linear theory [e.g., Thompson and Gasparovic, 1986]. Figure 3b therefore indicates that the leading edges of the packets are toward the West, and hence that the packets are traveling toward the West (more accurately 250°T and 275°T). We also observe in Figure 3b that the radar signature is of “double sign” type, with each wave consisting of a bright band toward the West followed by a darker band on its East side (i.e., a positive backscatter variation followed by a negative one, relative to the local mean). Since the thermocline depth is relatively shallow compared to the total depth (as will be shown in the next section) we can assume that the ISWs are waves of depression, and the direction of propagation can be confirmed by considering hydrodynamic modulation theory [Alpers, 1985]. This theory states that the bright bands should appear ahead of (in the direction of wave propagation) the darker bands, in particular for the first wave of the packet, which is easily identified in Figure 3b. This therefore further implies that the ISWs are traveling toward the West, into Cape Cod and Massachusetts Bay and will shoal along the Massachusetts coast. The packet WP1 is seen to be traveling toward about 275°T, whereas the packet WP2 is traveling toward about 250°T. These findings are consistent with the main propagation directions previously identified in Figure 2, and indicate the possibility of two (or more) separate wave-trains generated at different times from...
the same source (or by two distinct sources) and traveling in slightly different directions. This is investigated in more detail in the sections to follow.

In individual SAR images the ISW wave packets along each of the paths are often spaced at about 11–15 km which, assuming a semi-diurnal generation periodicity, translates to an average propagation speed of up to 0.35 m/s. Figure 4 shows the regression curve obtained from plotting ISW positions observed in all ENVISAT images as a function of time after Low Water (LW) at Boston Harbor previous to the satellite overpass. One can see that the phase speed estimates from this method (0.33 – 0.36 m/s) are in close agreement with the above mentioned value obtained from individual images. Moreover, it can be seen that wave packets traveling in different directions (250°T and 275°T) have similar phase speeds, and if we extrapolate the SAR data to the origin, generation times (relative to the tidal phase) appear to be just prior to LW at Boston. Thus the SAR observations suggest that generation of these waves occurs during the ebb phase of the tide in Massachusetts Bay.

3. Criticality of the Flow at Race Point Channel

The long-shore bathymetry between RP and the outer Cape Cod is nearly equivalent to a vertical wall (see Figures 1 and 2). Given the equally abrupt southern slopes of SB, and the relatively smooth flat bottom of RP Channel, it seems reasonable to explain the internal wave generation in the context of models of stratified flows through a channel with a lateral contraction [Melville and Macomb, 1987; Clarke and Grimshaw, 1994; Baines, 1995]. The forcing in this case is different from that resulting of flow across a pronounced sill, as in the case of SB. In what follows we first investigate the criticality (with respect to the wave phase speed) of flow at the inner end of Race Point Channel (entrance to Cape Cod Bay) and relate the findings with the positions of wave fronts revealed by the SAR and the morphology of the “nascent” internal wave signatures found within the Channel.

The Massachusetts Bays Program made continuous high-frequency (a measurement every 3.75 min) water velocity and temperature observations in RP Channel and Cape Cod Bay during 1990 and 1991. These are available on line at the USGS Woods Hole Science Center Moored Time Series Data (http://stellwagen.er.usgs.gov/), and were published and discussed by Irish and Signell [1992]. Figure 3a shows the location (red square mark) where the current meter and temperature loggers were deployed (42.1138°N –70.2693°E at 5 meters depth, and simultaneous measurements were obtained at 25 and 51 m depth.
Figure 3. (a) Positions of two distinct wave packets (denominated WP1 and WP2) relative to the bottom topography of Massachusetts and Cape Cod Bays. The dash line represents the axis of Race Point Channel, and its heading is 250°TN. Green shows the current velocity vectors for the first 24 hours of measurements, scaled according to the arrow shown on the top left corner of the map. The site of the in situ measurements is also indicated by the red square in between RP and SB. The two red arrows pointing eastward represent the average current during the period of transcritical regime. (b) ERS-1 SAR image extract dated 21 August 1994 acquired at 15:26 UTC, 35 minutes before High Water at Boston Harbor. The image area is shown in Figure 2a by a dotted rectangle, and the image is 12.8 × 12.8 km². The center of the SAR image is located at −70.28°E, 42.06°N.
Figure 4. (a) Plot of time versus distance measured from RP for waves propagating along 250°TN direction. The reference for time is Low Water at Boston Harbor. The best linear fit to the satellite data is shown. (b) Plot of time versus distance measured from RP for waves propagating along 275°TN direction. The reference for time is Low Water at Boston Harbor. The best linear fit to the satellite data is shown.
nearly at the same position 42.1182°N −70.2680°E, in total water depth of 61 m. The position where the measurements were acquired is approximately half the way across the Channel, between RP and the shallowest part of SB. The measurements position is thus conveniently located in the center and along the Channel axis, which we defined as a line that intersects the measurements position and heading 250°TN (see Figure 3a). Note that when viewed from the West (Cape Cod Bay) and along the RP Channel axis, the measurements position corresponds approximately to the point of minimum breadth between Cape Cod and SB. This is an ideal position to study the characteristics of the flow.

The period of the data time series is from 12 September 1990 (start at 17 h 28 min 07 sec UTC) to 15 January 1991. Coincident temperature records at the above mentioned depths (5, 25 and 51 meters) confirm the existence of stratification in the Channel for at least one month after the start of the measurements, which disappears during the winter due to vertical mixing. The RP Channel is well known for the strongest currents in Massachusetts and Cape Cod Bays [Irish and Signell, 1992], which can exceed 0.8 m/s in magnitude at 5 m depth and 0.6 m/s at 51 m depth.

Figure 3a presents a subset of the current measurements for the 5 m depth mooring (in the form of scaled vectors with the origin at the mooring location) representing the period of the first 24 hours, starting on the 12 September 1990. The currents have maximum amplitudes for directions that are approximately aligned with the RP Channel axis (typically 70°T and 250°T), generally fitting an ellipse with high eccentricity and major axis approximately aligned with the axis of the Channel [see Irish and Signell, 1992]. We also note that the flow in the upper and lower layers (represented here by the 5 m and 51 m measurements, respectively) are approximately in the same direction. Figure 5a represents a fragment of the current measurements corresponding to a period of about 5 complete semi-diurnal tidal cycles after 12 September 1990. It shows the mean velocity component along the axis of RP Channel (70°T). Analysis of the current records also revealed that, even when the water column was well stratified, the velocity at 25 m and 51 m are very similar. In Figure 5a the continuous line gives the mean current at the depth of 5 m (upper layer) while the dash line is an average of current measurements at 25 m and 51 m (lower layer). Note that during the ebb phase of the tide (mostly positive velocity component values) there is significant current shear. In these cases, the mean ratio of the upper to the lower layer velocity is about 2, or higher. The vertical shear evident in the ebb flow phase in the tide (Figure 5a) is significant since resonant wave generation in a contraction requires either vertical shear in the along-channel velocity field, or sloping side-walls [Melville and Macomb, 1987; Clarke and Grimshaw, 1994].

In addition we have analyzed stratification conditions for the period of June 2001, and Figure 5b presents a density profile obtained in Cape Cod Bay (upstream of the RP Channel at 42.029°N and −70.31°E), for the 12 June 2001 at the depth of 53 meters. A two-layer model with the interface at 15 m depth (see Figure 5b) seems to be a good fit to the data, and we will use it for our analysis. Furthermore, as part of the analysis below we estimate the wave speeds in a vertically sheared stratified fluid. Since we only have records of the velocity at several vertical levels we are limited to a model with low vertical resolution. Thus, in what follows we will consider an idealized two-layer flow through a channel with variable breadth, uniform depth, and vertical side-walls. This seems adequate since we are interested in a physically consistent explanation of the wave generation phenomenon.

The near-resonance of the flow is measured by the closeness of the linear long-wave phase speed to zero. In the case of a two-layer flow the phase speed is given by [Baines, 1995]

$$c_+ = \frac{u_1 d_2 + u_2 d_1}{d_1 + d_2} \pm \left( \frac{g'd_2}{d_1 + d_2} \right)^{1/2} \left( 1 - \frac{(u_2 - u_1)^2}{g'(d_1 + d_2)} \right)^{1/2} = \bar{u} \pm \varepsilon_0$$

where $d_i$, $u_i$, and $\rho_i$ are the depth, speed, and density of the upper ($i = 1$) and lower layers ($i = 2$). The reduced gravity $g' = (\rho_2 - \rho_1)/\rho_2$. Taking the ebb flow to be in the positive (+) direction and the velocities $u_i$ to be parallel to the channel axis (a good approximation in light of Figure 3a), criticality of upstream moving waves occurs for $c_+ = 0$. The resonance band occurs for a range around $c_- = 0$. Criticality of the flow is usually expressed in terms of the Froude numbers $F_1$ and $F_2$, where $F_i^2 = u_i^2/g'd_i$ [Armi, 1986]. The composite Froude number of the two-layer flow is [see also Dalziel, 1991]

$$G^2 = F_1^2 + F_2^2 = 1 + \frac{d_1 + d_2}{d_1 d_2 g'} c_+ c_-$$

Criticality ($c_- = 0$) corresponds to $G^2 = 1$. Note, however, that while $G < 1$ and $G > 1$ correspond to sub- and super-critical flow, respectively, the depth of sub- and super-criticality with respect to the wave speed (and hence the transcritical forcing band) is measured by $c_-$. Thus we divide the first term on the right of (1) by the absolute value of the second, to quantify the depth of the transcritical band, and write

$$F = \frac{\bar{u}}{c_0}$$

To calculate the transcritical band from the weakly nonlinear theory [e.g., Clarke and Grimshaw, 1994] we would need to assume flow in a channel with only small variations in width and no variations in either the mean flow or waves in the cross-channel directions. Clearly, the RP topography and the SAR observations violate both these assumptions and will require a more elaborate model to make detailed tests of the resonant generation hypothesis. Our goal here is simply to establish that the observations are consistent with the resonant generation theory.

With this in mind, Figure 6 shows a plot of Froude number $F$ computed using (3) for the data discussed above, representing the first 62 h 32 min of the record starting on 12 September 1990 (corresponding to 2.5 days and exhibiting 5 complete semi-diurnal tidal cycles). In the calculations we have assumed the stratification upstream of RP Channel of 12 June 2001 (see Figure 5b). The flow is nearly critical ($0.9 < F < 1.1$) when the current direction is approximately 50°–90°T, corresponding to the ebb phase of the tide (this is
Figure 5. (a) Mean velocity component of currents along the RP Channel in the upper (continuous curve) and lower layers (dash curve). The upper layer is represented by the measurements at the depth of 5 m, while the lower layer is an average of the measurements at the depth of 25 and 51 m. The time of measurements (in hours) refers to 12 September 1990 at 0 h UTC, corresponding to roughly 60 hours (or approximately 5 semi-diurnal tidal cycles). The LW at Boston Harbor is indicated by triangles, and the HW is indicated by open squares. Note the offset of the two curves particularly for positive velocity values (in the ebb flow phase in the tide), which correspond to current shear. (b) Plot of density distribution used in this work for the study region. The density profile (continuous line) is based on CTD measurements obtained on 12 June 2001 at location −70.31°E, 42.029°N (Cape Cod Bay). The dash line indicates the two-layer system used to calculate the Froude number (see text and Figure 6), with the interface at 15 m depth. The dash-dot-dash curve is the square of the buoyancy frequency \( N^2 = -\left(\frac{g}{\rho_0}\right) \cdot \frac{d\rho}{dz} \) calculated from the data.
The choice of $0.9 < F < 1.1$ to delimit the transcritical band is clearly arbitrary and made just to illustrate that the flow passes through the band. We note that when the same observations are plotted using (2) to compute the composite Froude number $G$, the results are entirely consistent with Figure 6. The semi-diurnal character of the tide in Massachusetts Bay is evident from Figure 6. We further note that for the subset of the data that is shown (approximately 2.5 days) the flow is supercritical during the ebb phase of the tide and subcritical during the flood phase of the tide. This is also the dominant scenario for the rest of the data record when stratification persists (until the end of September).

We have also computed the time the system is in the transcritical regime, here taken to be the interval of time for each excursion across critical conditions, when the $F$ is in the interval $[0.9, 1.1]$. Here, it is named residence time in resonance, and typically varies from 30 min to over 2 hours (see Figure 6). It can be seen that the transcritical intervals often occur twice during the ebb phase of the tide, firstly during the acceleration phase of the ebb flow (usually 1–2 hours after HW) and secondly during the deceleration phase of the ebb flow (some 1–2 hours before LW). On some occasions however the flow may be in the transcritical regime for a single interval during the ebb, when the maximum of $F$ occurs within the resonant band (see, e.g., the periods 37–39 h and 62–64 h approximately). In these instances the residence time is usually longer (typically 2 hours). We reiterate that when the flow is in the transcritical regime the mean current direction is mainly in the interval $[50^\circ, 90^\circ]$ (see Figure 6). This is certainly consistent with the ebb phase of the tide, $20^\circ$ to each side of the RP Channel axis.

Figure 6 also displays the time record of the temperature at 25 m depth, where the variability is rather accentuated (almost $7^\circ$C from ebb to flood). The temperature is lowest during the ebb, generally near LW times, indicating that the thermocline is displaced upward during the ebb and downward during the flood. During the ebb phase of the tide the current velocity records at each fixed depth show that the upper layer current speed ($u_1$) is approximately twice the value of the lower layer current speed ($u_2$), revealing considerable shear (see Figure 5a). Baines [1995] showed that when the upper layer current speed exceeds the lower layer speed, in the case of steady two-layer flows through contractions, the interface is displaced upward when the flow is critical at the minimum breadth position. The current shear observed during the ebb, simultaneous with the upward shift of the thermocline, is thus consistent with Baines [1995] findings. On the other hand, during flood conditions the current speeds are equivalent ($u_1/\approx u_2$), revealing considerable shear in upper and lower layers, and thus not appropriate for resonant generation in a pure contraction even if the flow were critical [Melville and Macomb, 1987; Clarke and Grimshaw, 1994]. We remark that an increase in short-period temperature oscillations is sometimes observed near LW times. This is best seen in Figure 6 at time 28–30 hours and 78–80 hours, just after LW at Boston. These temper-
nature oscillations, which are in the form of “pulses” of warmer water, can be of 2°C or higher, with individual periods of less than 1 hour (see temperature record in Figure 6). We also note that, at RP Channel, the ebb tide is slightly longer than the flood.

4. Generation of Internal Solitary Waves

[22] The results presented so far support the hypothesis that the ISWs observed in the SAR in Cape Cod and Massachusetts Bays could be generated at RP Channel just prior to (or nearly at) LW (Boston harbor). However, the first set of images that were analyzed did not include any good examples of images acquired during the ebb phase of tide, when the ISWs are expected to be generated at RP Channel. We thus went back to the image archives and selected all available SAR data corresponding to the ebb tide from June (1992–2007), covering both the ERS and ENVISAT missions. In all, 6 ERS SAR and 3 ENVISAT ASAR images fulfilled the criterion and were analyzed, although only 5 were found to contain ISW signatures in RP Channel, those being images from 10 June 1992 (ERS-1 Orbit 4722 15:28 UTC); 30 June 1993 (ERS-1 Orbit 10233 15:28 UTC); 8 June 1996 (ERS-2 Orbit 5934 15:28 UTC); 28 June 1997 (ERS-2 Orbit 11445 15:28 UTC); 27 June 2005 (ENVISAT 14:55 UTC).

[23] Of these we have selected three examples to show here: 30 June 1993, 8 June 1996, 28 June 1997. The high-resolution images are displayed in Figures 7–9 along with maps with location and interpretation of the observed solitary waves. Figure 7b shows an image obtained 3 h 18 min before LW at Boston, acquired at very low wind speed conditions and showing surface signatures of (at least 4) “nascent” ISWs in RP Channel, with crest lengths of some 1400 m and average wavelength of about 200 m (the ISWs show as bright bands on a dark background; for explanation of the imaging mechanism see, e.g., da Silva et al. [1998]). Figure 7a shows the position and orientation of the ISWs relative to RP Channel bathymetry. Figure 8b shows the image dated 28 June 1997 acquired 1 h before LW showing a ISW packet with 4 waves at approximately the same location (see also Figure 8a). In addition to these 4 waves, it also reveals a signature of a “nascent” soliton some 2500 m to the SW of the previous packet, adjacent to RP and very near to the tip of Cape Cod. The crest lengths of these waves are about 3000 m and the wavelengths are again approximately 200 m on average (see Figure 8a for an interpretative map of the observation). Figures 9a and 9b show two striking signatures of packets of “nascent” ISWs in the image dated 8 June 1996, nearly at LW time at Boston Harbor. In this case the orientation of the easternmost wave packet is distorted compared to the previous examples, as if the waves were swept eastward down the Channel and refracted. Each packet contains again at least 4 individual waves, with crest lengths 2.3 to 6.5 km and wavelengths of 200–300 m.

[24] In addition to these examples of nascent solitons, we requested a tandem mission satellite acquisition in the summer 2007, when ERS and ENVISAT flew in tandem and overpassed the study area with a time difference of approximately 30 min. Figure 10a shows a map with the interpretation of the images (see also Figure 10b), which were acquired just 4 min before LW and 24 min after LW on 21 July 2007. For simplicity only the leading crests of the two wave packets that were observed (Figure 10b) are marked in Figure 10a. The propagation of the two wave packets during the time interval between the two satellite overpasses can be clearly seen in Figure 10a. One can see that the wave propagation directions are distinct with the westernmost waves moving toward 250°T, while the easternmost wave packet propagates along an axis that is shifted northward, close to the previously proposed propagation direction of 275°T (see section 2) marked in Figure 2. Again, the ISWs observed in the images of 21 July 2007 have similar crest lengths (3.3 and 2.8 km) and wavelengths (150–250 m) to those previously discussed in this section. Given the times of observation (ebb tide and near Low Water) and orientation of the crests of these wave packets, it is reasonable to assume that these waves might be nascent nonlinear internal waves that later evolve and propagate into Massachusetts and Cape Cod Bays. We hypothesize that they might develop into waves with similar characteristics of WP1 and WP2 discussed in section 2 (see Figure 3). In fact, the pattern of these small-scale ISWs observed within RP Channel is consistent with the general pattern of wave groups presented in Figure 2 and discussed in section 2.

[25] In order to test if the observed pairs of ISWs at RP Channel could be generated by passage through resonance of the two-layer flow, we investigate the speed of linear interfacial waves generated at the measurements point when the flow is nearly critical. Equation (1) gives the dispersion relation of linear interfacial waves in a two-layer system with a rigid upper boundary in the long-wave limit [Baines, 1995]. Equation (1) may be solved to determine the position of linear internal waves as a function of time, once we know the generation position and time. Here we will consider only the negative root of (1) which corresponds to waves propagating upstream (against the flow). We have used the 5 m and 51 m level current speeds in (1) as representative of the upper and lower layers (section 3), respectively. The observed average wavelengths, estimated from the SAR images (Figures 7–10), are all greater than 150 m. For such wavelengths (1) is a good approximation for the wave speed. Then, assuming that internal waves are generated at the same location when the flow is critical, we can estimate the propagation of the newly born waves in the vicinity of the generation site. In doing so we assume that the stratification and currents are spatially uniform in the generation near field, which, strictly speaking, might not be the case. We further assume that the background currents are varying slowly in time compared to the wave phase. This is also required in the weakly-nonlinear passage through resonance analysis in the study of Wang and Redekopp [2001]. Lastly we ignore nonlinear effects on the wave phase speeds. We are mainly interested in qualitative results to explore the possibility of resonant generation, both when the flow passes from subcritical to supercritical speeds (acceleration), and from supercritical to subcritical speeds (deceleration). We seek the general trajectories in space and time of the internal waves in the generation near field, rather than the detailed propagation characteristics of nonlinear waves.

[26] Figure 11 presents predictions from (1) of the position of the internal waves after generation (assumed to occur
Figure 7. (a) Location map for ISWs observed near RP on 30 June 1993 at 15:28 UTC, 3 h 18 minutes before LW at Boston Harbor. (b) The area covered by the SAR image extract shown in Figure 7b is marked with a dotted rectangle. Also marked by a dash line is the RP Channel axis, heading 250°TN. The parallel (dotted) line represents the idealized propagation path of ISWs used to construct Figure 11. (b) The ERS-1 SAR image extract is 6.4 × 6.4 km², and its center is at 42.0802°N, −70.2318°E. It shows at least 4 wave crests as bright bands on a dark background.
Figure 8. (a) Location map for ISWs observed near RP on 28 June 1997 at 15:28 UTC, 1 h 00 minutes before LW at Boston Harbor. The area covered by the SAR image extract shown in Figure 8b is marked with a dotted rectangle. Also marked by a dash line is the RP Channel axis, heading 250°TN. The parallel (dotted) line represents the idealized propagation path of ISWs used to construct Figure 11. (b) The ERS-2 SAR image extract is 12.8 × 12.8 km², and its center is at 42.1053°N, −70.2573°E. It shows two groups of wave packets, one to the east of RP and considerably developed with at least 4 wave crests and another just starting to develop to the west of RP.
Figure 9. (a) Location map for ISWs observed near RP on 8 June 1996 at 15:28 UTC, 02 minutes before LW at Boston Harbor. The area covered by the SAR image extract shown in Figure 9b is marked with a dotted rectangle. Also marked by a dash line is the RP Channel axis, heading 250°TN. The parallel (dotted) line represents the idealized propagation path of ISWs used to construct Figure 11. (b) The ERS-2 SAR image extract is 12.8 × 12.8 km², and its center is at 42.0758°N, −70.2453°E. It shows two groups of wave packets, one with crests extending to the northeast of RP and with at least 4 well-developed wave crests and another starting to develop to the west of RP with 4 smaller wave crests.
when the flow is critical). Two solutions are represented: the first solution (bold curve) represents the distance internal waves travel from the origin where they are generated during the acceleration phase of the transcritical speed excursion, while the second solution (dashed curve) represents the distance traveled by internal waves generated at the origin during the deceleration phase of the flow passing through resonance. Both curves are plotted as a function of
Figure 11. Dispersion diagram for long, linear internal waves propagating upstream from (1) applicable to the waves observed with SAR and shown in Figures 7–10. The two curves (continuous and dotted lines) are two distinct solutions of internal wave propagation from the same origin generated by passage through resonance. The bold continuous curve shows distance traveled by internal waves generated during the acceleration phase of the tidal flow, while the dash curve gives the trajectory for waves generated during the deceleration phase of the flow. Black dots represent the positions of the leading crests of the SAR observed internal wave packets shown in Figures 7–10.

Figure 10. (a) Location map for ISWs observed near RP on 21 July 2007 at 14:58 and 15:26 UTC, 4 minutes before LW and 24 minutes after LW at Boston Harbor, respectively. The observations correspond to an overpass sequence of ERS-2 (tandem ERS-2/ENVISAT mission). We first note that the SAR data points are arranged in two groups: the first group corresponding to positive (downstream) distances and the second, corresponding to negative (upstream) distances from the origin (the assumed generation site). We also note that, in general, the data points are in agreement with the curves representing the two solutions that are being considered: generation during the acceleration and deceleration phase of the (critical) flow at the same location. Note that waves generated during the acceleration (about 1 hour after HW) are first advected eastward (downstream) for approximately 2 km, and then are able to propagate upstream after about 4 hours. Internal waves generated during the deceleration phase of the flow (about 5 hours after HW) may begin propagating upstream soon after generation time, having propagated some 900 m by the time of LW. We reiterate that this plot relies on several strong assumptions. Nevertheless it provides a valuable representation of the timing of ISW generation and the distance traveled by the waves in the early period of their “lives”. It supports the hypothesis that the two wave packets are generated as the ebb tide first accelerates and then decelerates through the transcritical regime.

5. Summary and Discussion

[27] The SAR observations can also be directly compared with the internal wave predicted trajectories. The selected SAR observed ISW positions of each individual packet (leading waves of each individual packet) in Figures 7–10 have been plotted in Figure 11 together with the solutions from (1) (negative root). These include the observations of 30 June 1993, 8 June 1996, 28 June 1997, and 21 July 2007 (tandem ERS-2/ENVISAT mission). We first note that the SAR data points are arranged in two groups: the first group corresponding to positive (downstream) distances and the second, corresponding to negative (upstream) distances from the origin (the assumed generation site). We also note that, in general, the data points are in agreement with the curves representing the two solutions that are being considered: generation during the acceleration and deceleration phase of the (critical) flow at the same location. Note that waves generated during the acceleration (about 1 hour after HW) are first advected eastward (downstream) for approximately 2 km, and then are able to propagate upstream after about 4 hours. Internal waves generated during the deceleration phase of the flow (about 5 hours after HW) may begin propagating upstream soon after generation time, having propagated some 900 m by the time of LW. We reiterate that this plot relies on several strong assumptions. Nevertheless it provides a valuable representation of the timing of ISW generation and the distance traveled by the waves in the early period of their “lives”. It supports the hypothesis that the two wave packets are generated as the ebb tide first accelerates and then decelerates through the transcritical regime.

[28] In this article we have investigated available ERS SAR and ENVISAT ASAR imagery in Massachusetts and Cape Cod Bays (41.5°–42.5°N), and have shown the presence of packets of ISWs propagating toward the West from the northern tip of Cape Cod. The waves emanate from RP Channel, and we therefore identify the RP region as a “hotspot” for internal wave generation. In more detail, the ISWs fall into two groups: wave group 1 (WP1), which is composed of the ISWs that initially travel toward 275°T, and wave group 2 (WP2), which is composed of ISWs traveling toward 250°T. These observations, together with the analysis of the criticality of the flow in the Channel, indicate the likelihood of resonant generation. The possibility of separate generation times during the ebb tide for these two groups has been investigated and is found to be broadly consistent with the resonant flow hypothesis.
Theoretical model studies have demonstrated that upstream generation of ISWs is a possibility even when the flow has variable Froude dependence with time provided that the flow is within the transcritical band for a sufficient length of time (e.g., Wang and Redekopp, 2001). Further, both the acceleration and deceleration passages through the transcritical band can generate nonlinear waves. We have shown in section 3 that the flow in RP Channel can be supercritical during the ebb tide, thus producing wave packets in a resonant flow from a bottom bump. Their method yields a generation time of 1 hour for conditions similar to this case study. An estimate for a lateral contraction gives a similar figure. The transversal regime time intervals at the site of the measurements, therefore, seem adequate for generation of nonlinear internal waves. In addition, Torsvik et al. [2006] have estimated the generation period for ISWs by resonant flow forced by a bottom bump. Their method yields a generation time of approximately 1 hour for conditions similar to this case study. An estimate for a lateral contraction gives a similar figure. The transversal regime time intervals at the site of the measurements, therefore, seem adequate for generation of nonlinear internal waves. In addition, Torsvik et al. [2006] have found the magnitude of the acceleration to be important for the amplitude and number of solitary waves generated. The smaller the magnitude of the acceleration, the more efficient the generation should be. The acceleration can be quantified by $\gamma = g^{-1} \frac{du}{dt}$, i.e., the ratio of the current acceleration to the reduced acceleration of gravity. This parameter gives values of about 0.001 to 0.002 for the cases presented in Figure 6. Interestingly, these values correspond to the same values found by Torsvik et al. [2006] in their model simulations of the acceleration and deceleration of ships in shallow water channels and the generation of upstream propagating solitary waves. In fact, Torsvik et al. [2006] found that $\gamma = 0.002$ was an optimum value for generating a train of solitary waves propagating upstream, both in the case of positive and negative accelerations.

What seems to be remarkable and unique about the ISWs generated at RP Channel, is that they might be generated on both excursions through the transcritical interval during a single ebb tide period. As far as we know, this has not previously been observed in the ocean. Cummins et al. [2003] have shown evidence of upstream influence in Knight Inlet, but the existence of two distinct wave packets generated on the acceleration and deceleration phases does not seem to exist. Similarly, ISWs generated at SB by the more usual lee-wave mechanism only produce a single packet each tidal cycle [Chereskin, 1983].

Such “double-generation” of ISWs in one single tidal cycle is supported by the SAR observations discussed in section 4. We note that the small-scale wave packets shown in Figures 7–10 are clearly observed in RP Channel during the ebb phase of the tide. In particular, on two of these dates, two distinct, relatively well-developed wave packets can be clearly identified within RP Channel during the ebb tide. The easternmost wave packet, located some 2 km (or more) to the east of RP, sometimes shows wave crests that are shifted some 30° (or more) in the clockwise direction relative to the westernmost wave packets. This can be best observed in Figures 9 and 10, where the westernmost wave packets are found to propagate toward 250°T while the easternmost wave packets propagate toward 275°T. Note that these two propagation directions are in close agreement with the two groups of ISW packets that have been previously identified in Massachusetts and Cape Cod Bays (see section 2 and also Figures 2 and 3). One possible explanation for the difference in wave propagation directions and locations may be the following. Wave group 1 (above called easternmost) would be generated first during the acceleration phase of the currents excursion across the transcritical regime. Although these wave packets are propagating upstream (westward), they would be initially advected eastward (downstream) by the (supercritical) flow. Thus the location of the wave crests would first move to the east. The speed of the currents is expected to be higher toward the center of the Channel, along the Channel axis, and lower near the coast due to friction. That would explain why the wave crests seem to be curling clockwise, since refractive effects would cause the waves to be displaced further east near the axis of the Channel where the currents are stronger. Wave group 2 (also called westernmost waves) would be generated at the same location as the first group, but at a later stage, during the deceleration phase of the transcritical excursion to subcritical speeds. When group 2 waves are generated they immediately escape upstream, and are less affected by refraction than are group 1 waves. This may explain why the general propagation direction of group 2 waves is approximately aligned with the axis of the Channel (250°T).

An alternative explanation for the two sets of ISW packets generated at RP Channel would be that these were generated at the same time (in the tidal phase) but at different locations (approximately 2.5 km apart according to the plot of the SAR observations in Figure 11). Note however, for the SAR image shown in Figure 7, corresponding to generation during the acceleration phase of the transcritical excursion, only one wave packet is observed. If the two sets of wave packets were generated at the same time at two different locations, in Figure 7 there would be a “second” wave packet 2.5 km to the west of the wave packet in the image. This is not the case. Since the SAR overpass occurs prior to the generation of group 2 waves that occurs during the deceleration of the flow to subcritical speeds it is consistent with the resonant generation at the same site, but at different times. We further recall the good agreement between wave packet trajectories calculated from the dispersion relation (1) and the SAR observations (see Figure 11) under this hypothesis.

We finally comment on the sharp warm water pulses that are sometimes observed in the temperature records (at 25 m depth) nearly at LW, and that have been mentioned in section 3. Although these short-period temperature oscillations have a lower frequency than the expected frequency of the ISWs, the timing of their appearance is consistent with the arrival of ISWs at the measurements location. The present data is inadequate to determine if the ISWs have actually been captured in the temperature records.

In the present study, we have identified the Cape Cod region off the coast of Massachusetts (USA), and in particular RP Channel, as a newly discovered “hotspot” of short-period internal waves in Massachusetts Bay. Here, the tidal flow is resonant with respect to long small-amplitude internal waves on the given (two-layer) density distribution. Pairs of packets of ISWs are observed by SAR
within RP Channel during the ebb tide. These are assumed to be generated during the ebb phase of the tide by the resonant flow. The first of these pairs of wave packets is generated during the acceleration phase of the flow passing through the transcritical speeds, while the second is generated when the flow decelerated from supercritical to subcritical rates. They both propagate upstream into Massachusetts and Cape Cod Bays, but with slightly different directions. We believe this may be the first time that resonant generation of ISWs by accelerating and decelerating flows through the transcritical regime has been observed in the ocean. The observations are consistent with the theories of resonant generation. However, it will take additional modeling that accounts for the specifics of the RP Channel geometry (i.e., lateral variations, etc), the local flows (e.g., lateral shear) and continuous stratification before a final conclusion can be made.

[35] Acknowledgments. We would like to thank Rich Signell from USGS for calling our attention to the existence of available in situ measurements in the study region and Brian White for stimulating discussions about the subject of this article. J. C. B. da Silva would like to thank Diego Almeida and his team for hosting his sabbatical leave at Woods Hole Oceanographic Institution. We thank Jorge Magalhães for helping in the preparation of figures. J. C. B. da Silva is also grateful to FCT for sabbatical leave support (BSAB/610/2006) and the Calouste Gulbenkian Foundation for partial support. J. C. B. da Silva was supported by FCT projects “SPOTIWA VE-II” (project code POCI/MAR/57836/2004) and “AMAZING” (project code PDCTE/CTA/49953/2003). K. R. Helfrich was supported by ONR grant N000140610798. Image data were provided by ESA (AOTP-2423 and AOE563). This research was partially supported by the Woods Hole Sea Grant Program (2008–2010 cycle), under a grant from the U.S. National Oceanic and Atmospheric Administration (NOAA), U.S. Department of Commerce, Grant No. NA06OAR4170021, project number R0-45.

References


Haury, L. R., M. G. Briscoe, and M. H. Orr (1979), Tidally generated internal wave packets in Massachusetts Bay, Nature, 278(5702), 312–317.


