1. Introduction

Flat slab subduction generally refers to slabs that extend nearly horizontally for some distance, sometimes hundreds of kilometers, in the upper mantle (Cahill & Isacks, 1992). We distinguish flat subduction from under-thrusting by noting that an under-thrusted slab remains in contact with the overriding plate and no asthenosphere material separates them (Pennington, 1981). Flat slab subduction zones comprise about 10% of all subduction zones and have several features in common, including: (1) Increased upper plate seismicity compared to steep subduction (Gutscher et al., 2000); (2) Orogenesis hundreds of kilometers from the...
trench (Bird, 1988; Humphreys et al., 2003; Jordan & Allmendinger, 1986); (3) Volcanism that is either atypically far from the trench or absent altogether (e.g., Bishop et al., 2017; Eakin et al., 2014; Kay et al., 1987).

The southernmost edge of the Caribbean (CAR) plate subducts obliquely at a low angle beneath northwestern South America (NWSA) due to a buoyant large igneous province impinging on a trench that lies offshore northwestern Colombia (Figure 1) (Burke, 1988; Kellogg & Bonini, 1982; Pennington, 1981). Like other flat slab subduction zones, the overriding South American plate has lacked magmatism for at least 45 m.y., has mountains up to $\sim 500$ km from the trench and is seismically active over that same length (Figures 3 and S1) (e.g., Cardona, Valencia, Bayona, et al., 2011; Mora et al., 2017; Taboada et al., 2000). However, the subducting CAR has not been previously well imaged due to limited instrumentation, leaving open questions regarding how the flat slab contributes to the Laramide-style uplifts seen in NWSA (e.g., Audemard & Audemard, 2002; Monod et al., 2010).

Laramide-style uplifts are characterized by basement-involved foreland deformation giving rise to basement-cored mountains (Brown et al., 1988; Fan & Carrapa, 2014; Jordan & Allmendinger, 1986; Smithion et al., 1978). The commonly proposed mechanisms that led to the Laramide uplifts of the western US include the low-angle subduction of the Shatsky rise conjugate on the Farallon plate (Bird, 1988; English & Johnston, 2010; Fan & Carrapa, 2014; Humphreys et al., 2003; Liu et al., 2010). Some models invoke direct transfer of stress from the Farallon to the North American plate through mechanisms such as interplate coupling (Bird, 1988; Saleeby, 2003) or bulldozing of material (Axen et al., 2018; McQuarrie & Chase, 2000). Other models invoke dynamic response due to mantle upwelling as the slab retreats or sinks (Humphreys et al., 2003; Liu et al., 2010). Despite being well-studied, the often-referenced Laramide orogeny is still poorly understood in part because it is no longer an active subduction zone.

The basement-involved thrust systems that formed the SdP and MA are thought to be analogous to the thrust systems that formed the Rocky Mountains during the Laramide orogeny (De Toni & Kellogg, 1993;...
Duerto et al., 2006; Kellogg & Bonini, 1982). Like the Laramide, extended plate coupling due to flat subduction of the CLIP is one mechanism proposed to have driven the uplift of the MA (Duerto et al., 2006). Additional proposals include continental subduction (Chacín et al., 2005; Colletta et al., 1997) or an orogenic float (Audemard & Audemard, 2002; Cediel et al., 2003) whereby the MA are detached from the lithosphere and uplifted via distant stresses that are transferred from near the trench. Thus, the precise mechanism for the formation of the MA remains uncertain (Audemard & Audemard, 2002; English & Johnston, 2010; Erslev, 2005; Monod et al., 2010).

In addition to the question of the CAR's role in mountain building, the basic question of the geometry of the subducting CAR remains open. From north to south, the Caribbean plate starts as sea floor before under-thrusting beneath northern Colombia and Venezuela. In northern Colombia, the CAR transitions from under-thrusting to subducting beneath NWSA. Where this transition occurs is not resolved. It has been proposed to be along the Oca-Ancón fault near Lake Maracaibo (Masy et al., 2015) and south of the Oca-Ancón fault near the Santa Marta Massif (Mora et al., 2017). Farther south the subduction of CAR ends. Not only is the location of this boundary unclear, with various authors proposing a boundary between 5°N–9°N, but identifying the location is complicated by other slabs, notably the Nazca slab, possibly overlapping with the CAR (e.g., Chiarabba et al., 2015; Ídárraga-García et al., 2016; Kellogg et al., 2019; Malavé & Suárez, 1995; Sanchez-Rojas & Palma, 2014; Syracuse et al., 2016; Taboada et al., 2000).

In this study, we use finite-frequency teleseismic P-wave tomography to resolve structures in the CAR-NWSA subduction zone and suggest how the flat subduction has mechanically resulted in the formation of Laramide-style uplifts in the Serrania de Perijá and Mérida Andes. The data were recorded by a 65-element portable broadband seismograph network deployed in NWSA and 39 broadband stations of the Venezuelan and Colombian national seismograph networks. We constrain the geometry and timing of the subduction system by identifying and mapping slab segments and using subduction rates by previous authors to deduce the subduction history of the CAR. We then compare the subduction history to the tectonic history of NWSA to hypothesize the role of the CAR.

### 1.1. Tectonic Background

In the hot spot reference frame, the CAR has remained roughly fixed at a constant longitude since 55 Ma as the Americas moved westward (Boschman et al., 2014; Jordan, 1975; Pindell et al., 1998). Along the Southern Caribbean Deformation Belt, where no physiographic trench is visible due to sedimentation into the Colombia and Venezuela basins, the CAR under-thrusts the northern edge of NWSA and subducts beneath the northwest margin of NWSA (Figure 1) (Miller et al., 2009; van der Hilst & Mann, 1994). The CAR began subducting beneath NWSA between 70 and 55 Ma, coinciding with magmatism along the margin (Cardona, Valencia, Bayona, et al., 2011; Mora et al., 2017). By 45–55 Ma magmatism abruptly ceased, likely due to the onset of low-angle subduction of the Caribbean Large Igneous Province (CLIP) (Montes et al., 2019; Mora et al., 2017; Taboada et al., 2000). The CLIP is thought to have been erupted from the Galápagos hot spot in the late Cretaceous and added 12 or more kilometers of buoyant crust atop the oceanic plate (Kerr et al., 1997). Since 45 Ma the CAR-NWSA subduction zone has lacked volcanism and seen uplift of the Santa Marta Massif (SMM), the Serrania de Perijá (SdP), and the Mérida Andes (MA)—the latter being ~500 km from the trench (Figure 1) (Kellogg & Bonini, 1982). Both the SdP and the MA were primarily uplifted since ~10 Ma (Bermudez et al., 2010; Cediel et al., 2003; Kellogg & Bonini, 1982).

NWSA is seismically active due to deformation in both the overriding and subducting plate (Figures 3 and S1). Plate boundary deformation is accommodated by slip on the Boconó fault, Santa Marta-Bucaramanga fault, as well as the El Pilar-San Sebastian fault system and several other secondary faults (Figure 1). The El Pilar-San Sebastian fault accommodates nearly all the CAR-SA strike-slip motion (Symithe et al., 2015). The Boconó fault accommodates 9–11 mm/yr of dextral shear and 1 mm/yr of compression (Kellogg & Bonini, 1982). The Santa Marta-Bucaramanga fault accommodates about 6 mm/yr of left-lateral slip (Symithe et al., 2015). The Oca-Ancón fault (OAF) and the many secondary faults slip at a rate of 2 mm/yr or less (Audemard & Audemard, 2002). Together, these three fault systems facilitate the escape of the Maracaibo Block (MB) over the CAR (Figure 1). Most seismicity in western Venezuela occurs shallower than 30 km (Malavé & Suárez, 1995), but Benioff seismicity is observed beneath the SdP and Santander Massif (Figures 3 and S1).
1.2. Previous Work

Several previous studies proposed geometries of the subducting CAR. As in the flat slab segments of western South America (e.g., Antonijevic et al., 2015; Bishop et al., 2017; Cahill & Isacks, 1992; Gutscher, Malavieille, et al., 1999; Gutscher, Olivet, et al., 1999; Jordan & Allmendinger, 1986; Ramos & Folguera, 2009), the CAR slab subducts with a shallow angle at shallow depths, then transitions to a much steeper angle some distance inland. Early studies used hypocenters to determine the shallow slab dipped from ~15° to 30° at an azimuth of 109° to a depth of 200 km beneath the Maracaibo basin (Kellogg & Bonini, 1982; Pennington, 1981). From tomography van der Hilst and Mann (1994) estimated the slab dipped at 17° ± 3° in a direction of 150° ± 20°. They also interpreted an overlap between two slabs at the Bucaramanga seismic cluster. Subsequent studies measured the flat segment dip at ~25° (Malavé & Suárez, 1995; Ojeda & Havskov, 2001; Zarni et al., 2007) to as much as 40° (van Benthem et al., 2013). Bernal-Olaya et al. (2015) found that the slab varied from north (as little as 8°) to south (as much as 35°). The steeply dipping segment was imaged to the 660 discontinuity beneath the MA by Bezdala et al. (2016) and van Benthem et al. (2013) and was interpreted to penetrate the 660 by Taboada et al. (2000).

In some subduction zones the subducting plate is separated from nonsubducting plate by lithospheric tearing at a Subduction-Transform Edge Propagator (STEP) fault (Govers & Wortel, 2005). North of SA the CAR comprises the sea floor and, in some locations, under-thrusts SA to depths of ~150 km (Guedez et al., 2007; Masy et al., 2015; Sanchez et al., 2010; Schmitz et al., 2008; Van der Hilst & Mann, 1994). To accommodate the transition between where it is under-thrusting seafloor and where it is deeply subducting there must be a tear or major deformation of the plate. The northern edge of the CAR-NWSA subduction system has not been identified, so it is not clear whether a STEP fault accommodates a transition between subducting and nonsubducting/under-thrusting CAR or if subduction is continuous along the entirety of the NWSA margin. Authors have proposed both continuous deforming CAR (e.g., Sanchez et al., 2010; van der Hilst & Mann, 1994) and subducting/nonsubducting CAR separated by a tear (Masy et al., 2011; Mora et al., 2017; Taboada et al., 2000). CAR under-thrusting from the north has been partially imaged in several studies of varying resolution (Guedez et al., 2007; Masy et al., 2015; Miller et al., 2009; Sanchez et al., 2010; Schmitz et al., 2008).

The southern boundary of the subducting CAR slab is poorly defined. It has been interpreted as a continuous plate to at least as far south as the Caldas tear at 5°N (Cortes & Angelier, 2005; Kellogg et al., 2019; Pennington, 1981; Van der Hilst & Mann, 1994; Vargas & Mann, 2013; Yarce et al., 2014) or only as far south as the Bucaramanga seismic nest at 7°N (Figure 1) (Chiarabba et al., 2015; Idárraga-García et al., 2016; Sanchez-Rojas & Palma, 2014; Syracuse et al., 2016). At Bucaramanga, the relationship between the seismic cluster and plate configurations is also poorly understood. Some authors have interpreted the cluster as indicative of a boundary between two slabs (Sanchez-Rojas & Palma, 2014; Syracuse et al., 2016; Zarifi et al., 2007) while others see the cluster origin wholly within the Nazca plate (Chiarabba et al., 2015; Van der Hilst & Mann, 1994) or wholly within the Caribbean plate (Cortes & Angelier, 2005; Kellogg et al., 2019; Pennington, 1981; Taboada et al., 2000; Vargas & Mann, 2013; Yarce et al., 2014). Thus, the southern boundary of the Caribbean slab might not be a clear delineation, but rather a confluence of overlapping and interacting slabs that may include the Nazca plate or some other intermediate slab segment.

2. Data and Methods

2.1. Data

Data were collected during the CARibbean-Mérida Andes seismic experiment (CARMArray) (Levander, 2016) which comprised 65 IRIS-PASSCAL broadband (BB) stations deployed from April 2016 to March 2018, with 19 BB stations from the permanent Venezuela National Seismic Network, 12 stations from the GIAME project in Venezuela, and 8 BB stations from the permanent Colombian National Seismic Network (Table S1). Combined, the arrays extend from the Caribbean coast of northwest Colombia to 200 km east of the Mérida Andes and along the Venezuelan coast, with station spacing of ~40–75 km (Figure 1). Embedded in the layout are two dense lines with ~20 km station spacing that extend from the Colombian coast to the Mérida Andes in the northern and southern edges of the study area.
Our Vp tomography used 148 events with Mb ≥ 5.5 and epicentral distances of 30°–95° (113 events) and 150°–180° (35 events) (Figure 2). Waveforms were bandpass filtered into three narrow frequency bands centered at 1.0, 0.5, and 0.3 Hz. Relative traveltime delays for P and PKIKP phases were computed by cross-correlating the waveforms (VanDecar & Crosson, 1990) so that each waveform had a delay for each frequency band. The delays were then demeaned for each event. Waveforms that had delays greater than 3 s were discarded. After running the inversion several iterations, residuals were inspected. Waveforms with residuals greater than 5 standard deviations above the mean were removed, leaving a total of 23,675 teleseismic P-delays used in the inversion with 7,833 from the 1.0 Hz band, 7,948 from the 0.5 Hz band and 7,894 from the 0.3 Hz band.

2.2. Finite Frequency Teleseismic Tomography

We used nonlinear 3-D finite-frequency teleseismic P-wave tomography code of Schmandt and Humphreys (2010) as modified by Bezada et al. (2013) and the details of the method are found therein. Here we provide details of our setup. Our model domain is ~2,150 × 1,550 km and centered at 8.8°N, 69.2°W. The grid size varies both laterally and vertically with lateral grid spacing varying from 42 km within the footprint of our array to 56 km outside of the array and vertical spacing increasing from 35 to 55 km with increasing depth. The vertical spacing of the upper 90 km varied from 5 to 10 km. The code uses an iterative solver and computes 1-D travel times using the AK135 Earth model (Kennett et al., 1995) from event source to the model domain, so-called source-to-box times, and the Stingray 3-D ray tracer (Hammond & Toomey, 2003; Toomey et al., 1994) to compute traveltimes from the edge of the domain through the model volume to the stations. Approximate Born sensitivity kernels were calculated after each iteration. Various damping and smoothing regularizations were tested to balance data variance reduction with model variance reduction (Figure S6). Values of 2 and 10 were chosen for damping and smoothing, respectively. Limited results for additional parameters are in Supplementary Information (Figure S7). We ran the code 6 iterations, which maximized variance reduction and minimized data misfit (Figure S8). Total variance was reduced by ~91.5% and the final rms data misfit was 0.195 s.

2.3. Inclusion of Shallow Structure in Tomography

To address unconstrained shallow layers and prevent unimaged shallow structure from being mapped to deeper parts of the model we created a 3-D local P-wave tomography model and inserted the upper 80 km of it into the upper 80 km of the starting teleseismic tomography model (Figures S1–S5). Details of the crustal model may be found in supporting information. The mean velocity at each depth was removed to produce anomalies compatible with the teleseismic tomography inversion. The local tomography used 1,369 local earthquakes with depths of 5–200 km and stations with predicted (based on the AK135 model) traveltime residuals less than 2 s (Kennett et al., 1995). Since the local tomography model is smaller in extent than the teleseismic model, values were extrapolated by the nearest-neighbor algorithm.

2.4. Model Resolvability

To understand our model resolvability, we utilize two tools—ray hit count quality and synthetic tomography. Hit quality is calculated as the weighted average of number of rays present in each of six 60° azimuthal bins at each node. Bins saturate at six rays so that a score of one means that at least six rays were present in every bin for a node. For each node the scores of the six bins were averaged to provide the final hit count quality. We consider values above 0.6 to be well-resolved. Areas with hit quality less than 0.6 and 0.3 are shaded light gray and dark gray, respectively (Figure 4).

Using the event-station geometry of our data set we created synthetic tomographic images to evaluate the details of the velocity model (Figures 3 and 7). We performed a checkboard test and two image recovery tests by creating slab structures within our model domain characterized by large fast anomalies. For the checkerboard test, we created synthetic traveltimes through a checkerboard velocity model and completed the inversion with the unconstrained upper 60 km highly damped. The checkerboard test reveals that within the footprint of the CARMArray ±5% velocity anomalies can be expected to be recovered as ±2–4% anomalies, but outside of the footprint anomalies will not be well recovered (Figures 3a–3g). The test
also revealed that lateral smearing is minimal but there is some vertical smearing, particularly toward the south-southeast where the largest concentration of events originates.

The first recovery test is a southeast-dipping continuous slab with a +5% anomaly that extends across the study area (Figures 3h and 3i). The inversion used identical damping and smoothing as the model. Like the checkerboard test, it shows that the slab-like anomaly is recovered within the footprint of the array. Velocity perturbations of 1%–2% are recovered at 125 km, but larger amplitudes are recovered at greater depth. There is some vertical smearing between 200 and 300 km and weak slow anomalies appear that were not part of the input. At 390 km depth there is a small break in the anomaly at its southern limit, which cautions against interpreting diverging plates at that depth in the model. The second recovery test is derived directly from the tomography: We created three surfaces from the fast anomalies and converted contours of those surfaces back into synthetic 5% fast anomalies (Figure S9). We ran two versions of this test: one with three slabs and the other with two slabs. We will discuss the second recovery test in the results section in the context of our tomographic images, as they provide us with a qualitative proxy for model resolution and robustness in our final tomographic model.

Overall, resolution is strongest within the footprint of the array. The southern edge of the study area, south of 7° near Bucaramanga, is complicated and not likely well resolved above 200 km or below 500 km. For shallow structure we use seismicity to improve our interpretations. The synthetic tests indicate that smearing, though present, does not obscure the imaged structures. We confine our interpretations to areas with a hit quality greater than 0.6.

3. Results

Here, we describe our tomographic model and compare it to the synthetic model described previously to determine the structure of the deeper subduction system. We observe three fast anomalies continuous through several depths, two of which connect at shallow depths but separate as depth increases. The third anomaly starts at ~200 km under the MA and appears to merge with another of the anomalies at greater depth. We also observe two fast anomalies identified only at depths less than 195 km at the northern and southern edges of the model.

Shown in Figure 4 are map views of the subduction system starting at 60 km depth down to 635 km depth. In each panel the anomaly label orientations are meant to match the strike of the anomalies. The fast anomalies at depths of 60–90 km under the SMM might be the flat portion of the Caribbean plate. The down-going Caribbean plate is the prominent fast (1%–3%), continuous feature at 125 km (labeled S1, for subduction) and follows the seismicity along the SdP and Santa Marta-Bucaramanga fault. The strike of S1 is concave at shallow depths but becomes more linear with increasing depth (Figures 4 and 5). From 125 to 195 km the S1 strike roughly parallels the Santander Massif and the SdP. Our synthetic image (Figure 7)
shows that this geometry is recovered except at the southwestern corner of the SdP. Also notable is that S1 extends to just south of the OAF where it connects to another anomaly, U (for under-thrust), that strikes mostly east-west (4). Segment U is not clearly identifiable deeper than 195 km. To the south at 125 km depth is a fast anomaly, S2, that we partition into shallow S2a and deeper S2b. The strikes of S2a and S2b also differ as S2a is generally north-south while S2b is northeast-southwest. S2 starts ∼300 km farther inland than S1 and maintain strikes distinct from S1. From 195-435 km depth S1 exhibits nearly constant slope from the SdP to the MA (Figure 6). Another anomaly, S3, appears under the MA at ∼160 km depth and merges with S1 by ∼400 km depth. East of the MA S3 dips steeply, nearly reaching the bottom of the MTZ where the model domain ends. At 310 km depth S2b and S3 anomalies appear continuous, but their strikes are clearly different.

In our synthetic image at 195 km we inserted an anomaly corresponding to S3 (Figure 7). We do not see vertical smearing of that anomaly. Also, when S3 is not included in the synthetic no fast anomalies appear in its place. At 350 km depth we model S1 and S3 with as little as 40 km of separation. As in the tomography,
the synthetic image demonstrates the two anomalies are distinguishable. At 435 km depth S1 has nearly completely merged with S3. At that depth S2 and S3 appear continuous in our model with slightly different strikes. We placed a gap between them in our synthetic, and what results are two distinct anomalies, unlike in the tomography model.

Profiles A-D (Figure 6) are roughly perpendicular to the MA and are sub-perpendicular to the SdP. They extend in the dip direction across the SdP and show the slab from ∼80 to 690 km depth. In profiles B and C, the CAR dip is nearly constant until the MA, where we observe a large anomaly atop the CAR that we identify as S3. We observe that S1 is trenchward of and has shallower dip than S3 in every profile where both segments appear. S1 does not reach the MTZ while S3 does. Profile A has a shallow, flat anomaly and a discontinuous steep anomaly. We identify the shallow structure as U, the under-thrusting CAR from the north. The deep structures S3. Profile D shows a steep, fast anomaly that is preceded by gently sloping seismicity. The steep anomaly appears as twice the width of a normal slab, similar to where S1 and S3 overlap, so it could represent two slabs overlapping as proposed by Taboada et al. (2000). Profile F is an east-west dip profile that runs oblique to A–D and intersects the corner of the SdP and SM. It shows S1 and S3 with a nearly identical configuration as seen in profiles B and C. We inserted a sub-vertical anomaly for S3 into...
our synthetic and found that it was well recovered (Figure 7). The synthetic S1 was well recovered below 100 km depth across the profiles.

Profile E is nearly in the strike direction of S1 and crosses the S1 and U anomalies. A gap between the anomalies aligns with intermediate depth seismicity. We do not include a gap in the synthetic Profile E (Figure 7), and no gap is present in the recovered synthetic. Profile G runs west to east and crosses the Eastern Cordillera slightly north of the Bucaramanga nest. It shows a high amplitude, steep anomaly oriented east. Profile H traces the western margin of the MA. It shows dipping seismicity that aligns with a fast anomaly, indicating the presence of a slab. We assign the dipping anomaly to S2 while the horizontal anomaly may comprise both S1 and S3 since the profile is taken where S1 and S3 converge. Similarly, in our synthetic profile H we model a dipping S2 converging with S1 and S3 (Figure 7). All anomalies are recovered.

4. Discussion

4.1. Interpretation of the Caribbean Plate

We interpret our observations of the CAR segments as follows: U is the under-thrusting segment of the CAR that transitions to sea floor north of Venezuela and Colombia. S1 is the primary, western-most CAR slab that is currently subducting. S2 is likely part of a southern segment of the CAR, and possibly also a segment of the Nazca plate, but it is less well resolved due to array size. S2a appears in the tomography before S2b.
and has a north-south strike. S2b appears to have a strike similar to S1 and perhaps is connected to S1 to the northwest, outside the resolution of our study. S3 is a broken CAR segment previously attached to S1 (and possibly S2) and appears as the anomaly beneath the MA. Our interpretations are summarized as a box model in Figure 8 and surfaces in Figure S9. From these surfaces segment lengths were calculated and related to subduction timing in the following subsections.

S1 is imaged as shallow as 125 km and follows the curved curtain of seismicity along the SdP and Santander Massif at that depth. We interpret S1 as bowed concave east (Figures 4, 5 and S9). The edge of S1 coincides with the significant reduction in seismicity at intermediate depths. That site is likely the tear indicated by Mora et al. (2017), where S1 is separated from U. Figures 6a and 6b show profiles to the north and south of the proposed tear, respectively. Figure 6a lacks dipping seismicity. Figure 6b has both dipping seismicity and S1 that dips to ~400 km. Figure 6e transects the S1/U anomalies nearly along strike and shows a gap separating the anomalies where intermediate seismicity ends. We do not observe south-dipping structures originating north of the Oca-Ancón fault. Rayleigh wave tomography suggests that parts of the nonsubducting Caribbean plate under-thrust northern South America south of the OAF at depths of 80 km (Miller et al., 2009). The OAF generally separates the eastward subducting S1 from the southward under-thrusting U. We estimate the length of S1 to be ~760 km.

The southern segment, S2, appears as a fast anomaly near the Bucaramanga nest. Its strike is different than the other segments at depths shallower than 230 km; it parallels seismicity along the southern Santander Massif and the Andean Eastern Cordillera. Between 310 and 530 km depth S2 appears continuous with S3 (Figure 4). Profiles D and G (Figure 6) intersect east of the BN and transect S2. Both profiles show dipping anomalies following dipping seismicity. We suggest S2a is not part of the CAR while S2b is part of the CAR, which implies that the CAR’s southern boundary lies south of our study area. Many authors have proposed overlapping slabs in the Bucaramanga region (Cortes & Angelier, 2005; Prieto et al., 2012; Taboada et al., 2000; Vargas et al., 2007; Yarce et al., 2014) but the lack of resolution along the entirety of profile D prevents us from placing the Bucaramanga nest in one plate or another. Assuming that Profile D has at least
one slab continuous from the trench to the MTZ, we designate the entirety of that surface as S2 and estimate the length of the anomaly to be \(\sim 725\) km. Projecting the surface back to the Caribbean Sea adds \(\sim 500\) km for a total segment length of \(\sim 1,225\) km.

The broken slab, S3, lies east of the primary slab, S1. Its apex is at \(\sim 160\) km depth, which is \(\sim 175–225\) km above the deepest tip of S1 that subducts behind it. Whether S3 penetrates the 660 transition is unclear. The upper half of S3 is nearly vertical and was likely imaged by Bezada et al. \(2010\). Unlike S1, the strike of S3 parallels the Mérida Andes, suggesting there was a transition in the subduction history of the CAR. We estimate the length of S3 to be \(\sim 450\) km.

Figure 8 shows two potential geometries in the subduction system. The simplest geometry would be for S2 to comprise only the CAR (or Nazca?) (Figure 8a). In this configuration, S2 strike is parallel to the Eastern Cordillera. A challenge with this model is the discrepancy in depths where S1 and S2 meet. A more
complicated geometry has S1 overlying S2 (Figure 8b). In this scenario S3 could only have been attached to S1 and S2b, and S2a might be the Nazca as described by Taboada et al. (2000) and Londoño et al. (2020).

4.2. Subduction History of the Caribbean Beneath NWSA

Evidence of magmatism along northwest SA dated to $\sim$70–55 Ma indicates the normal subduction of CAR (Cardona, Valencia, Bayona, et al., 2011; Mora et al., 2017). By 45–55 Ma magmatism abruptly ceased, coinciding with major readjustment in the configuration of the South American, Caribbean, and North American Plates (Kellogg et al., 2019; Montes et al., 2019; Mora et al., 2017; Pindell et al., 1998). This magmatic cessation is interpreted as indicating the onset of low-angle subduction of the CLIP (Taboada et al., 2000). It is not clear how much of the subducted slab carries the CLIP. Assuming $\sim$15–25 m.y. of normal subduction followed by $\sim$45–55 m.y. of shallow subduction of CLIP at the current subduction rate of 20 mm/yr,
one could expect ∼300–500 km of subducted normal crust and ∼900–1,100 km of CLIP. Higher estimates of convergence rates prior to 45 Ma by Boschman et al. (2014) would increase the amount of normal CAR lithosphere subducted to 600–1,000 km, in line with estimates by Montes et al. (2019).

Our interpretation accounts for ∼1,210 km of slab (S1 + S3), which falls on the lower end of the above estimate of normal crust + CLIP. Our estimates do not account for slab that might have flattened at or penetrated through the bottom of the MTZ. These estimates are greater than previous estimates by van Benthem et al. (2013), ∼900 km, and Mora et al. (2017), ∼1,050 km. Thus, assuming past subduction rates are close to current subduction rates, S1 accounts for ∼37.5 m.y. and S3 accounts for ∼22.5 m.y. for a total of ∼60 m.y. of subduction. Therefore, it is likely that S3 records a history of normal subduction of the CAR. From the tomography we cannot clearly distinguish between different thicknesses of slab to determine the transition from nonCLIP to CLIP.

### 4.3. Breakoff of the CAR

Why, when, and where under NWSA did the CAR break? We favor a lithospheric weak point where oceanic crust transitions to CLIP as the most likely site of the break. S3, considering its length and subduction rate, carries oceanic crust and likely carries some CLIP. Dynamic models from Arrial and Billen (2013) show that change in density—due to eclogitization of a subducted plateau in their case—can ultimately lead to slab breakoff in a flat slab system. The break off occurs as the slab rolls back due to its denser tip. For the CAR this corresponds to the eclogitization of enough CLIP combined with the dense, normal slab that would have also eclogitized to provide significant density/buoyancy contrast with the remaining subducted CLIP to lead to rollback and slab detachment. While we cannot narrowly constrain the timing of the break, we rule out the buoyant CLIP entering the trench as the initiation as it would leave S3 sub-vertical in the mantle for ∼45 m.y., an unlikely geometry for that length of time. Restoring S1 to its pre-break location using a sinking rate of 20 mm/yr places the breaking event under Lake Maracaibo.

The current length and position of S3 requires the slab to break no deeper than ∼200 km. Eclogitization reactions may occur as shallow as 50 km in the hottest subduction zones, but the full transformation of plateau crust to eclogite will be delayed due to thicker crust and the typically cooler temperatures above a flat slab (Arrial & Billen, 2013; Gutscher et al., 2000; Hacker, 1996; Pennington, 1984). Based on a sinking rate of 20 mm/yr and the assumption that S1 began rolling back soon after the break, we propose that the midrange of possible break times is 10 Ma. A delay in rollback would require the break time to occur later while a faster sink rate would require the break to occur earlier. In the conceptual model shown in Figure 9 we back out the slab from present day to 45 Ma assuming NWSA was moving west at 20 mm/yr relative to a stationary CAR within that period. We also assume 50 km of shortening between the SdP and MA from 15 Ma to present (Monod et al., 2010). The model shows that the once-continuous slab might have been flat as far east as the eastern banks of Lake Maracaibo. That S3 is currently at least as shallow as 200 km and is positioned upright east of S1 suggests to us that breakoff initiated later in the timeline. We therefore favor a model where the slab breaks between 5 and 15 Ma.

We cannot disqualify other factors contributing to slab break off. Coeval with our proposed timeline was the Panama arc collision, which might have initiated as late as 15 Ma, influencing regional tectonics, including the North Andean orogeny (Bermúdez et al., 2010; Dengo & Covey, 1993). Far-field stresses transmitted into the region could have exploited a weakness in the CAR, leading to detachment. Vargas and Mann (2013) proposed the Panama indenter wedged its way beneath NWSA, exploiting a previous lithospheric zone of weakness leading to the Caldas tear along its southern boundary. Conceptually, they propose that the Caribbean slab is currently breaking from the Caldas tear to the Bucaramanga nest, though this tear is south of the broken slab, S3, we observe.

### 4.4. Implications for the Overriding Plate

As with the Rocky Mountains (Bird, 1988; English et al., 2003) and the Sierras Pampeanas (Jordan & Allmendinger, 1986; Löbens et al., 2010), flat subduction is proposed to be responsible for the uplift of SMM, SdP and MA. In Figure 10, we align our proposed timeline for CAR subduction with local and regional events. While mechanisms have been proposed to explain the uplift of the SdP and MA (Audemard &
Audemard, 2002; Colletta et al., 1997; Duerto et al., 2006; Kellogg, 1984; Monod et al., 2010), here we offer additional possible contributions to uplift due to the CAR breakoff.

The CAR currently subducts at a low angle ~50–100 km beneath SMM and ~125–175 km beneath SdP. SMM was uplifted in phases, with the most recent initiating ~15 Ma (Cardona, Valencia, Weber, et al., 2011). SdP was also uplifted in phases, but the main uplift occurred in the last 3 m.y., well after the arrival of the CLIP into the subduction zone (Kellogg, 1984; Kellogg & Bonini, 1982). The main uplift of the SdP coincides with the proposed timing of S1 rollback. As modeled by Buiter et al. (2001), slab roll back may induce uplift over...
the hinge line and this effect may be greater for a more buoyant slab—up to 4 km of uplift. The hinge line is currently under the western margin of the SdP and would have been under the eastern margin at 3 Ma. Their model also indicates a weak contribution to subsidence of a land-locked basin from rollback, which suggests a small signal of S1 rollback may be recorded in the subsidence history of the Maracaibo basin.

Thus, shallow subduction of the CAR may have contributed to the initial phases of uplift of the SdP and subsidence of the Maracaibo basin.

How did shallow subduction of the CAR contribute to the rapid uplift of the MA? While we cannot definitively attribute the topography of the MA with slab breakoff, we consider three mechanisms that could contribute. First, the CAR remains flat beneath NWSA as far inboard as the location of the MA until S3 breaks off S1 at 10–15 m.y. ago. A sinking rate of 40 mm/yr allows S3 to be restored to the surface while still under the MA. A descent velocity of 40 mm/yr is reasonable given the plate age (Goes et al., 2011). The first contribution to the uplift of the MA is the CAR breaking into S1 and S3 and S1, with minimally eclogitized CLIP, buoyantly rising beneath the MA at ∼10–12 Ma. Elastic rebound might also contribute to flattening S1 beneath NWSA. S1 would recouple with South America to transfer NW–SE directed shear stresses into the overriding plate (Figure 9c), shortening the upper plate as Jurassic-aged rift structures are reactivated (Duerto et al., 2006). A similar mechanism was proposed by Fan et al. (2014) to account for deformation during the Laramide orogeny. In their model moderate uplift was initially achieved through basal traction of the flat Farallon slab with North America that reactivated preexisting structures. A second contribution could be shear drag on the upper plate from asthenospheric flow over S3 induced by S1 rollback (Figure 9b) producing compressive stress in the overriding plate that would lead to shortening. Similarly, Fan et al. (2014) proposed that the Laramide orogeny experienced accelerated uplift when the slab rolled back due to its negative buoyancy relative to the surrounding mantle. The third contribution to the uplift of the MA is the replacement of denser oceanic lithosphere of S1 with less dense asthenospheric material, causing an isostatic response.
5. Conclusions

Our results highlight a complicated subduction system with at least three subducted CAR segments with varying strikes and dips. The main segment, S1, strikes along the seismicity under the SdP and Santander Massif. It remains flat until the Santander Massif and SdP where it dips steeply and terminates in the upper MTZ. A broken segment, S3, appears under the MA at \( \sim 160 \) km depth and dips steeply into the MTZ. A southern segment, S2, dips steeply under the EC and is likely flat prior to that. We assigned S2 as part of the CAR, but two seemingly distinct anomalies with different strikes suggest there might be two slabs overlapping each other. We found that we imaged \( \sim 1,210 \) km of slab (S1+S3), which accounts for \( \sim 60 \) m.y. of subduction.

The northern boundary of subduction lies south of the OAF and between the SMM and the SdP. It is marked by a gap in the tomography and a cessation of intermediate seismicity. We conclude that the southern boundary of the subducting CAR lies south of our study area.

The segmentation of the CAR opens questions of why it happened and what impact it had on the overriding plate. We hypothesize that the break between S1 and S3 occurred along a zone of lithospheric weakness or a strength contrast between CLIP and normal oceanic lithosphere. We used conceptual models to restore the CAR to a configuration just prior to break off. Depending on the sinking rate the CAR could have broken 5–15 m.y. ago. The breaking of the slab is coeval with the Panama arc collision, so it is possible the collision contributed to the break and, consequently, the uplift of the MA. The break is also coeval with the uplift of the MA. The impact of the CAR breaking on the topography of NWSA is uncertain. We hypothesize that when S3 broke from S1, S1 recoupled with the overriding plate, reactivated preexisting structures, and uplifted them. The rollback of S1 might have induced trenchward horizontal asthenospheric flow along the base of the South American lithosphere, causing northwest-directed shortening that contributed to the rapid uplift of the MA. The replacement of dense oceanic lithosphere with less dense asthenosphere may have also contributed an isostatic response in the overriding plate.

Data Availability Statement

Data collected on the CARMA array are available through the IRIS Data Management Center (https://doi.org/10.7914/SN/YU_2016). Data collected on the instruments of FUNVISIS are available by request through the agency (servicios@funvisis.gob.ve). Data collected on Servicio Geológico Colombiano instruments are freely available and accessible to the scientific community through written request and justification addressed to the Colombian Geological Survey (sismologo@sgc.gov.co).

References


References From the Supporting Information


