



Kinematic evolution of a continental collision: Constraining the Himalayan-Tibetan orogen via bulk strain rates

Andrew V. Zuza^{a,*}, Yann Gavillot^b, Peter J. Haprock^c, Chen Wu^d

^a Nevada Bureau of Mines and Geology, University of Nevada, Reno, NV 89557, USA

^b Montana Bureau of Mines and Geology, Montana Tech University, Butte, MT 59701, USA

^c Department of Earth and Ocean Sciences, University of North Carolina Wilmington, Wilmington, NC 28403, USA

^d Key Laboratory of Continental Collisional and Plateau Uplift, Institute of Tibetan Plateau Research, Chinese Academy of Sciences, Beijing 100085, China

ARTICLE INFO

Keywords:

Himalaya-Tibetan orogen
Shortening strain rate
Crustal thickening
GPS data

ABSTRACT

We investigated temporal variations of the bulk strain rate across the Himalayan-Tibetan orogen, which impact the magnitudes, rates, and distribution of deformation across the orogen. The present-day strain rate estimated from geodesy is $\sim 5\text{--}6 \times 10^{-16} \text{ s}^{-1}$ and estimation of past rates depends on the orogenic shortening rate and effective orogen width. Plate-circuit reconstructions provide constraints on time-varying India-Asia convergence rates, showing a marked deceleration since initial collision at ca. 58 Ma. Geologic evidence suggests that most of the Tibetan crust started deforming shortly after initial collision, which simplifies estimation of the initial orogen width. We examined several kinematic models: (1) Greater Indian and Tibetan crust started deforming immediately after collision, (2) only Tibetan crust deformed initially, (3) plate convergence was decoupled from crustal shortening, with Tibetan crust deforming slower than plate rates, or (4) a hard India-Asia collision at ca. 45 Ma following closure of the Xigaze backarc. Models 1, 3, and 4 yield early Cenozoic bulk strain rates that match present-day rates ($5\text{--}6 \times 10^{-16} \text{ s}^{-1}$), whereas Model 2 yields faster rates ($1.2 \times 10^{-15} \text{ s}^{-1}$) that must have decreased through time to avoid shortening exceeding total plate convergence. Using these bounds, we constrain crustal shortening across the entire orogen and along its southern and northern margins, defined by the Himalaya and Qilian Shan, respectively. Application of these models to Tibetan crustal thickening suggests that the plateau reached its present-day value ($\sim 70 \text{ km}$) by ca. 15 Ma, which corresponds to the onset of lateral deformation (e.g., strike-slip and normal faulting). Lateral deformation in the Himalayan-Tibetan orogen may have resulted from progressive crustal thickening and reorientation of the intermediate principal stress axis. At the thrust-belt scale, our modeling yields shortening rates and magnitudes that match geologic observations. This approach provides testable external kinematic constraints to guide future geological and geophysical investigations.

1. Introduction

The plate convergence velocity between India and Asia incrementally decelerated during the Cenozoic, and the timing of this decrease roughly corresponds with collision-age estimates for the Himalayan-Tibetan orogen (e.g., Molnar and Tapponnier, 1975; Patriat and Achache, 1984; Molnar and Stock, 2009; Cande et al., 2010; Copley et al., 2010; van Hinsbergen et al., 2011a, 2011b, 2019). Various explanations for the causal response of this deceleration include removal of dense mantle lithosphere, an increase in gravitational potential energy related to the thickening of Tibetan crust and increased topography, Indian slab detachment, subduction of relatively buoyant Indian continent, and/or

viscous resistance of the mantle lithosphere (e.g., Flesch et al., 2001; Molnar and Stock, 2009; van Hinsbergen et al., 2011b; Clark, 2012; Webb et al., 2017). A slowdown in plate convergence has important kinematic implications for the spatial and temporal distribution of crustal deformation during the across-strike evolution of an orogen (Fig. 1A). Regardless of the driving mechanism, this paper focuses on the predicted effects required to accommodate a deceleration of plate convergence across a growing collisional orogen. Slowing convergence rates throughout the Cenozoic must have affected the spatially varying patterns and rates of deformation across the entire Himalayan-Tibetan orogen (HTO) that can be tested against observational and numerical experimental data to provide insight into the various debates of the

* Corresponding author.

E-mail address: azuza@unr.edu (A.V. Zuza).

<https://doi.org/10.1016/j.tecto.2020.228642>

Received 11 May 2020; Received in revised form 12 October 2020; Accepted 13 October 2020

Available online 19 October 2020

0040-1951/© 2020 Elsevier B.V. All rights reserved.

kinematic evolution of orogeny, heterogeneity in crustal strengths, and resulting responses to the deceleration of plate convergence between India and Eurasia.

Kinematic reconstructions of the HTO since the onset of collision are often presented as the summed bulk estimates of published Cenozoic crustal shortening, which are compared to estimates of plate convergence derived from paleomagnetic and marine geophysical data (e.g., van Hinsbergen et al., 2011a; Yakovlev and Clark, 2014). Using

relatively robust constraints from plate-circuit reconstructions, bulk strain rates can be derived across the HTO (e.g., Clark, 2012) to predict plate convergence rates and magnitudes across the entire orogen, or intra-plate shortening of individual thrust belts, as a function of time. The degree to which these predictions match geological or geophysical constraints of crustal deformation provides insights on the dominant tectonic processes that operate during the evolution of a continent-collision orogen, such as continental subduction, delamination,

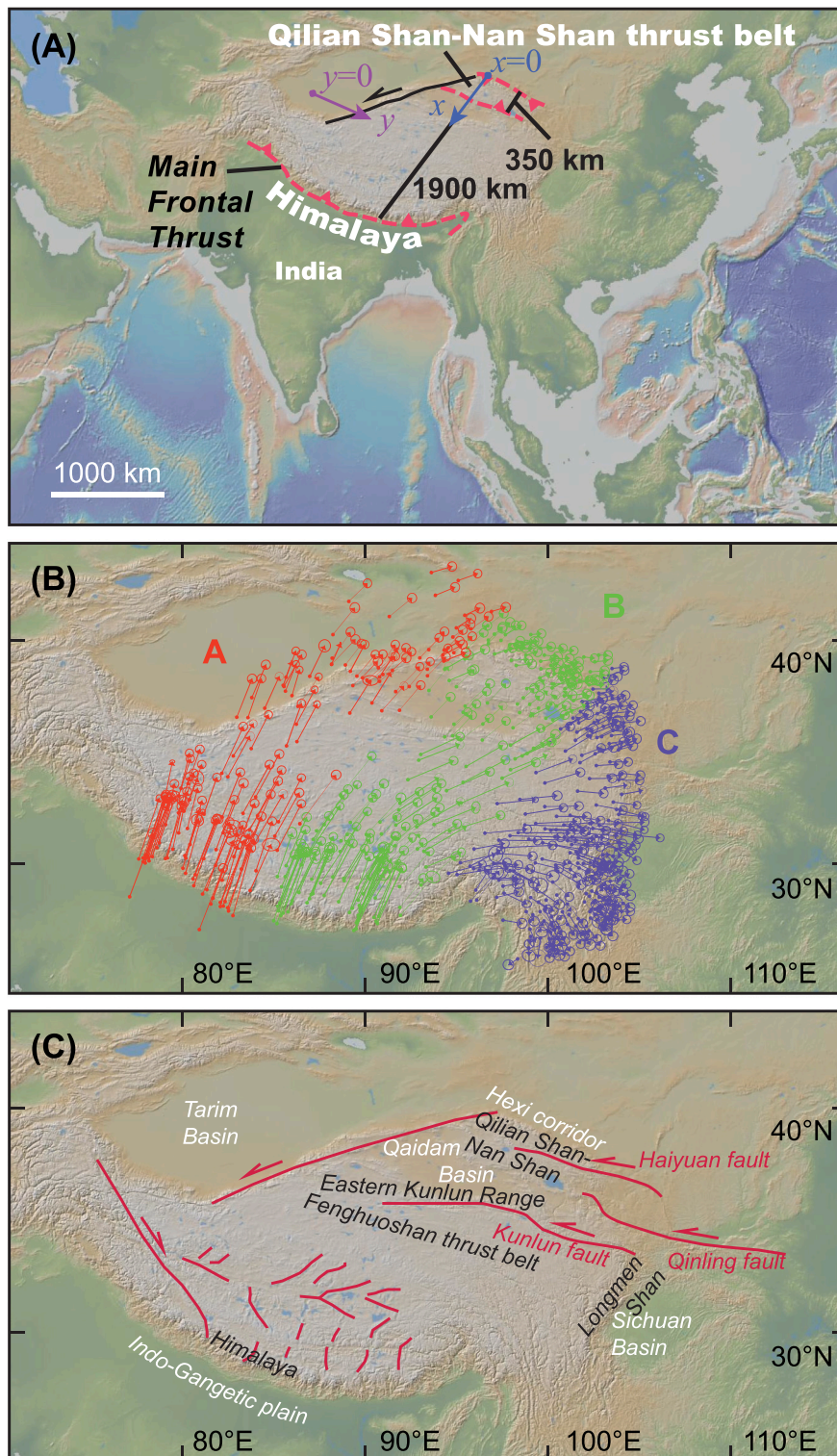


Fig. 1. (A) Map of the Himalaya-Tibetan orogen showing the present-day width of the entire orogen (~1900 km) between the Main Frontal thrust of the Himalaya in the south and the North Qilian Shan thrust in the north (geographically, between the Indo-Gangetic foreland basin and the Hexi Corridor, shown in panel C). The width of the Qilian Shan-Nan Shan thrust belt (~350 km) and the basic coordinate system used in this study for 1D analyses ($x=0$ at the Hexi Corridor and increases toward the Himalaya parallel to N23E, and $y=0$ in Tarim Basin and increases westward parallel to N113E) are also shown. (B) Eurasian reference frame velocity field for the Tibetan Plateau from Ge et al. (2015), color coded for three NNE-trending swaths (i.e., A, B, and C). (C) Simplified map of structures (from Taylor and Yin, 2009; Haproff et al., 2018) and locations discussed in text. Red fault lines show strike-slip or normal faults that activated in the middle Miocene (e.g., Duvall et al., 2013; Styron et al., 2015). Basemaps generated using GeoMapApp (Ryan et al., 2009).

extrusion, or distributed crustal shortening (e.g., McKenzie, 1969; Molnar and Tapponnier, 1975; England and Houseman, 1986; Tapponnier et al., 2001; Royden et al., 2008; Haproff et al., 2018).

In this study, we first derive the bulk contractional strain rate across the HTO through time, which has been postulated to have remained constant throughout the Cenozoic (Clark, 2012). This rate depends heavily on the assumed geodynamics and paleogeography of the India-Asia collision, which vary significantly in the literature. For example, some kinematic models involve Greater India crust deforming since the onset of collision with Asia, whereas others suggest that the early stages of orogeny involved oceanic subduction and collisions of one or several island/terrane arcs, such that plate convergence and intra-continental shortening may have been decoupled (e.g., Roy, 1976; Aitchison and Davis, 2004; van Hinsbergen et al., 2011a, 2011b, 2019; Ingalls et al., 2016; Webb et al., 2017; Kapp and DeCelles, 2019 and others). To first order, various reconstructions support the HTO deforming at a nearly constant bulk strain rate throughout the Cenozoic (Clark, 2012). However, in detail this simplified scenario does not reproduce the observed deformation patterns at the scale of individual thrust belts across the HTO. Cenozoic fold-thrust belts are not evenly spaced between the Himalayan thrust front in the south and the North Qilian Shan thrust in the north (Fig. 1A). Instead, several discrete thrust systems characterize the HTO including the Himalaya (including the Tethyan Himalaya), Gangdese retroarc, Qiantang, Fenghuoshan, Kunlun-Qimen Tagh, Qaidam, and Qilian Shan-Nan Shan thrust belts from south to north respectively, all of which experienced variable shortening histories since the initial India-Asia collision. Deformation was variably partitioned among these intra-orogen thrust belts to accommodate the observed deceleration of India-Asia convergence.

Quantifying reasonable strain rates at the scales of both the entire HTO and individual thrust belts can provide a predictive kinematic framework to evaluate spatially and temporally varying shortening magnitudes and rates. Here we use two examples from the Himalaya and Qilian Shan-Nan Shan thrust belts, which bound the orogen in the south and north respectively, where modeled kinematics appear to match published geological observations over 10's of myr timescales. Furthermore, the models fill in data gaps in regions of limited or obscured kinematic observations. We suggest that this approach can yield reasonable bounds of predicted geological shortening magnitude and rates, which can help guide future geological investigations, and highlight that the inferred deformation kinematics across parts of the orogen should be consistent with the broader plate-motion framework. Most importantly, we argue that (1) intra-HTO shortening, at all scales, likely decelerated to match the net deceleration of India-Asia convergence rates during the Cenozoic and (2) deformation rates/magnitudes should be compatible across the orogen, such that significant deceleration or acceleration of local rates across a discrete fold-thrust belt must be partitioned elsewhere in the orogen.

2. Infinitesimal bulk strain rate across the Himalayan-Tibetan orogen

2.1. Present-day strain rate

Using the contemporary velocity field across the Himalayan-Tibetan orogen (compiled in Ge et al., 2015), we calculated the orogen-wide infinitesimal bulk strain rate (e.g., Clark, 2012) (Fig. 1). Because we are using velocity data on short timescales relative to the orogen's history, we refer to this as infinitesimal strain, which differs from using convergence/shortening magnitudes versus orogen width that represents finite strain. We project these data into two primary components: northeast-southwest contractional strain, which shows convergence parallel bulk shortening, and NW-SE stretching of the HTO.

The GPS data was divided into three NNE-SSW-trending swaths to examine how the velocities change from west to east along the HTO: (A) Western HTO, (B) Central HTO, and (C) Eastern HTO in Fig. 1B. A

transect oriented parallel to India-Asia convergence (N23E) across the Tibetan Plateau defines our x-axis (Fig. 1A), with $x=0$ at the northern margin of the Tibetan Plateau directly north of the Qilian Shan-Nan Shan thrust belt (Meyer et al., 1998; Zuza et al., 2016; Zuza et al., 2018), beyond which contractional deformation becomes negligible (<1 mm/yr) (Figs. 1 and 2). The positive x direction is toward the active southern margin of the HTO along the southern edge of the Himalayan accretionary wedge and the Indo-Gangetic foreland basin, located at $x = 1,900$ km (Fig. 1A). The width of the orogen, $L(t)$, has varied through time and we define today as $t = 0$ myr. Thus, $L(t=0 \text{ myr}) = 1,900$ km, which we refer to as the final length of the orogen, L_f . The initiation age of the India-Asia collision is debated, with most estimates falling between 40 Ma and 60 Ma (e.g., Besse et al., 1984; Leech et al., 2005; Dupont-Nivet et al., 2010; Najman et al., 2010; Wang et al., 2011a). The revival of double-subduction hypotheses (e.g., Roy, 1976; Aitchison and Davis, 2004) derived from paleomagnetic data offer a different model of the collisional history with an initial arc/microcontinent collision in the earliest Cenozoic followed by a more recent hard collision of India with Asia in the Oligocene-Miocene (e.g., Meng et al., 2017; van Hinsbergen et al., 2011a, 2011b, 2019). However, these alternative models have been contested based on geological and statistical arguments (e.g., Kapp and DeCelles, 2019; Rowley, 2019 and references therein). Kapp and DeCelles (2019) provided a kinematic model involving backarc opening and subsequent closure along the southern Asian margin, which we discuss in more detailed below. Herein, we primarily focus on the more traditional early Cenozoic single India-Asia collisional framework. Recent sedimentological analyses, coupled with detrital zircon geochronology, suggest a robust ca. 58 Ma collision age for the collision between India and Asia (DeCelles et al., 2014; Hu et al., 2015, 2016), which we set as $t = -58$ myr.

Geodesy indicates that convergence-parallel velocities across the orogen, $v(x)$, in a Eurasian-reference frame, broadly decrease linearly from southern margin of the HTO in the south to the northern edge of the Tibetan Plateau (Fig. 2) (Zhang et al., 2004; Gan et al., 2007; Ge et al., 2015). Although the observed velocity field distribution is nonuniform within the HTO likely due to strain partitioning across rigid blocks and/or interseismic to seismogenic strain cycles of major bounding faults, there is a linear deceleration at the scale of the entire HTO. The rate at which the orogen width decreases, dL/dt , is equal to the rate at which India converges with stable Eurasia, presently $v(x = 1900 \text{ km}) = \sim 40$ mm/yr. To calculate the present-day bulk infinitesimal strain rate across the length of the HTO, the velocity $v(x = 1900 \text{ km})$ is divided by its width, $L(t = 0 \text{ myr})$ (Clark, 2012) (Fig. 2). Accordingly, $v(x = 1900 \text{ km})/L(t = 0 \text{ myr}) = (\sim 40 \text{ mm/yr}) / (1900 \text{ km})$, which leads to a bulk strain rate of $\dot{\epsilon} = 6.6 \times 10^{-16} \text{ s}^{-1}$. Other estimates of bulk contractional strain rate based on GPS data include $6.3 \times 10^{-16} \text{ s}^{-1}$ (Allmendinger et al., 2007) and $\sim 4.9 \times 10^{-16} \text{ s}^{-1}$ (Zhang et al., 2004). We explore this further by stacking the GPS velocity vectors of the Western and Central HTO swaths projected along the N23E transect (Fig. 2). A linear regression of the data, accounting for uncertainties, yields a negative slope (i.e., infinitesimal strain rate) of $5.5 \times 10^{-16} \text{ s}^{-1}$ (Fig. 2). Both of these calculations overlap with bulk strain rate values derived by Clark (2012).

The approach used throughout this paper is inherently one dimensional, along a vector parallel to India-Asia convergence. The rates derived above correspond to the NE-SW convergence-parallel bulk shortening of the entire HTO. Because this analysis is one dimensional, the inferred shortening of the orogen does not directly bear on crustal thickening, erosion, or out-of-plane lateral extrusion. However, we contend that 1D strain rates can be compared to geological observations, such as contractional strain values and strain rates derived from cross-section reconstructions made at the scale of detailed geologic maps or the entire orogen. Later in this paper, we use the 1D kinematics to extrapolate to multidimensional deformation.

Although the GPS velocity field broadly decelerates linearly from the Himalaya to the north (Fig. 2A), in detail there are slope variations within the data. In particular, significant velocity gradients and

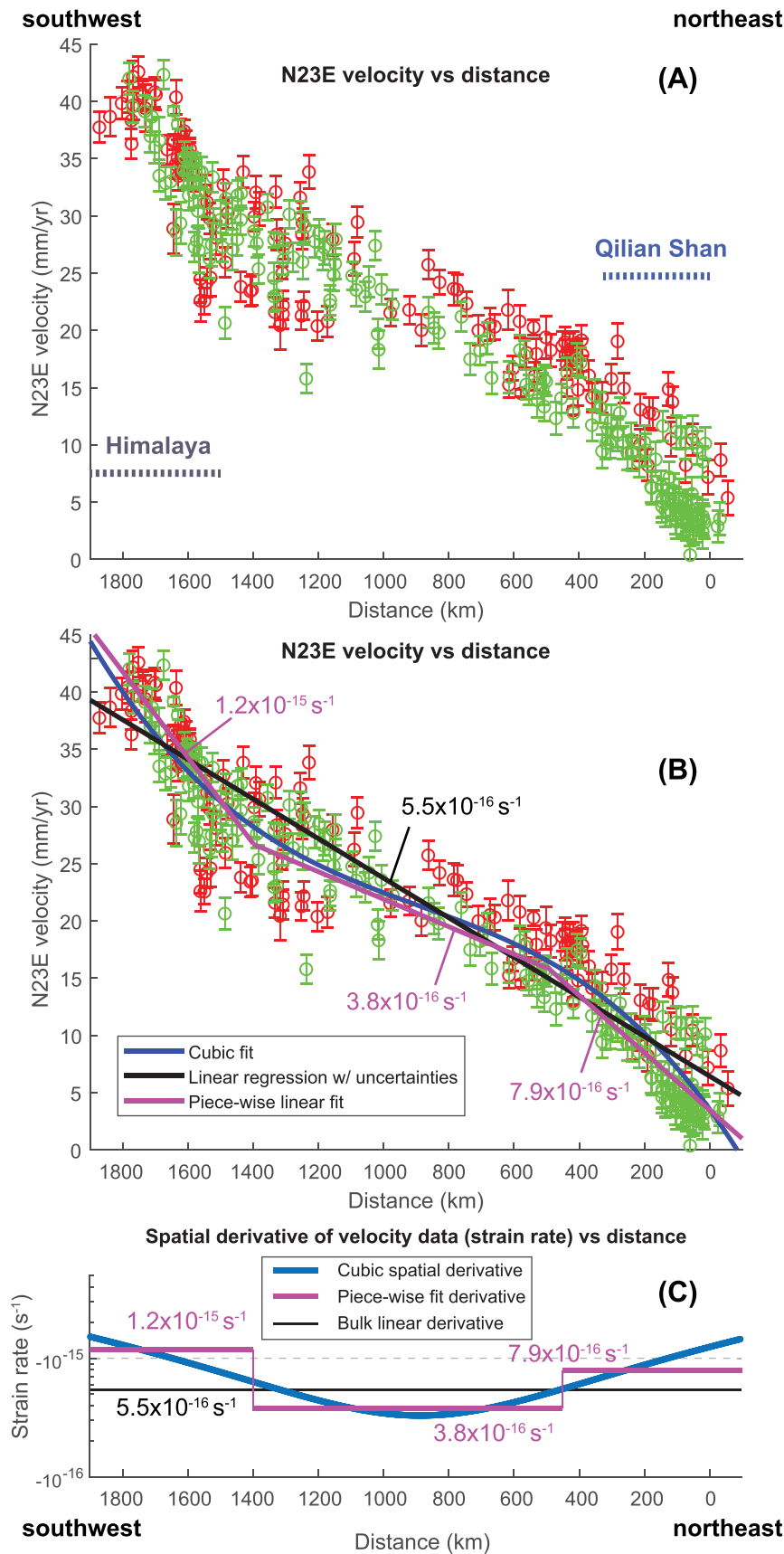


Fig. 2. (A) Convergence-parallel N23E GPS velocities from Western (red) and Central (green) HTO swaths (Fig. 1B) plotted versus distance. The spatial position of the Himalaya and Qilian Shan mountain belts are shown. (B-C) Best-fit regressions (B) using a cubic, linear, and piece-wise fit, which yield bulk (C) spatial derivatives or bulk strain rates, as a function of distance.

deceleration characterize regions across the Himalayan thrust belt in the south and the northern Tibet thrust systems in the north (i.e., the Eastern Kunlun Range, Qaidam, and Qilian Shan-Nan Shan thrust belts; Fig. 1C). We fit two additional regressions to the data: one as a cubic regression and one manual piece-wise fit. The spatial derivatives of these curves represent infinitesimal strain rates. These fits reveal higher strain rates in the Himalaya and northern Tibet compared to the central part of the Tibetan Plateau (Fig. 2). Specifically, this yielded strain rates of $\sim 1 \times 10^{-15} \text{ s}^{-1}$ for the Himalaya, $\sim 4 \times 10^{-16} \text{ s}^{-1}$ for the central Tibetan Plateau, and $\sim 8 \times 10^{-16} \text{ s}^{-1}$ for northern Tibet. These results imply strain rate variations across the orogen that contrast an assumed average bulk linear regression (e.g., Clark, 2012). These rates, applied over the length scales for which they were calculated for, imply present-day shortening rates across the Himalaya, central Tibet, and northern Tibet of 18 mm/yr, 12 mm/yr, and 13 mm/yr, respectively.

The HTO experiences bulk WNW-ESE stretching orthogonal to plate convergence, either via extrusion or extensional collapse (Tapponnier et al., 1982; Peltzer and Tapponnier, 1988; Blisniuk et al., 2001; Tapponnier et al., 2001; Liu and Yang, 2003; Bian et al., 2020). A plot of all GPS velocities projected along a N113 transect shows a linear increase in velocity towards the east (Fig. 3), which implies east-directed horizontal stretching of the crust. The slope of a best-fit linear regression suggests an extensional infinitesimal strain rate of $2.2 \times 10^{-16} \text{ s}^{-1}$, which is slower than the northeast-southwest contractional strain rate ($\sim 5\text{--}6 \times 10^{-16} \text{ s}^{-1}$). Allmendinger et al. (2007) derived a comparable extensional strain rate of $2.9 \times 10^{-16} \text{ s}^{-1}$ from GPS data. This strain rate applied over the $\sim 2,500$

width of the orogen equates to 17 mm/yr of west-east stretching of the orogen, which is less than half of the total ~ 40 mm/yr north-south plate convergence rate across the HTO. In the Discussion we expand on the implications of west-east stretching for our 1D modeling.

2.2. Constraints on strain rate at the time of India-Asia collision

Restoration of the effective pre-collisional orogen width—that is, the total restored orogen width affected by collision-related deformation since the onset of collision—compared with the Cenozoic India-Asia plate convergence rates can be used to estimate the temporal variability in infinitesimal strain rate. The northernmost extent of the HTO and Tibetan Plateau is currently defined by the active Qilian Shan-Nan Shan thrust belt and Haiyuan fault (Gaudemer et al., 1995; Zheng et al., 2010) bounded to the north by the Hexi Corridor (Fig. 1), which is the northern foreland of the HTO (Wang and Coward, 1993; Métivier et al., 1998). Stratigraphic, structural, and thermochronologic evidences suggest that northern Tibet, including the Qilian Shan-Nan Shan thrust belt, started deforming in the early Cenozoic, potentially almost immediately after initial India-Asia collision (e.g., Dupont-Nivet et al., 2004; Yin et al., 2008a, 2008b; Clark et al., 2010; Craddock et al., 2011; Li et al., 2020). This suggests that the region comprising the Qilian Shan-Nan Shan and Hexi Corridor (Fig. 1C) has remained relatively stationary as the northern edge of the HTO since initial India-Asia collision (e.g., Clark, 2012; Yakovlev and Clark, 2014; Zuza et al., 2019; An et al., 2020). The stationary location of the Hexi Corridor through time seemingly

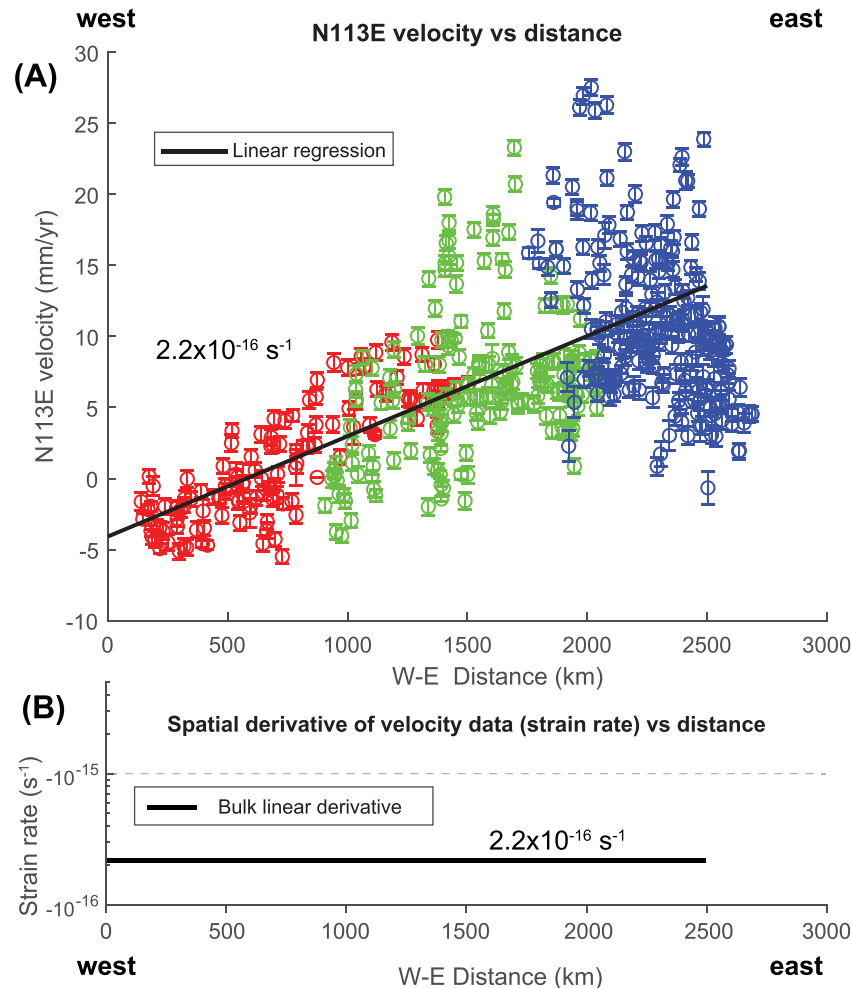


Fig. 3. (A) N113E velocity versus distance for the Western (red), Central (green), and Eastern (blue) swaths with best fit-linear regression. The colors are the same as Fig. 1B. (B) Spatial derivative or bulk strain rate versus distance.

contradicts its inferred role as a long-lived foreland basin, which typically should be time transgressive as the North Qilian Shan thrust system (Zheng et al., 2010) propagated northward (e.g., Meigs and Burbank, 1997; DeCelles, 2004). However, several studies show that the northern Tibet thrust systems involved significant out-of-sequence deformation (Wu et al., 2019, 2020a, 2020b; Zuza et al., 2018, 2019; Li et al., 2020), which may explain how the Qilian Shan-Nan Shan thrust belt and the Hexi Corridor basin system to the north remained relatively stationary (e.g., Clark, 2012; Yakovlev and Clark, 2014). Because of this, we treat the Hexi Corridor region as a fixed pinline in the analysis of this study.

India-Asia convergence rates have decreased throughout the Cenozoic (e.g., Molnar and Tapponnier, 1975; Molnar and Stock, 2009; Copley et al., 2010; van Hinsbergen et al., 2011a, 2012), and the HTO width decreased due to progressive convergence via shortening, underthrusting, subduction, extrusion, or a combination of mechanisms (e.g., Dewey and Bird, 1970; Dewey and Burke, 1973; Molnar and Tapponnier, 1975; England and Houseman, 1986; Peltzer and Tapponnier, 1988; Tapponnier et al., 2001; Royden et al., 2008). It is therefore possible, and perhaps likely, that bulk strain rates also varied during the development of the orogen. To test potential variability in bulk strain rate, we calculated the ratio of convergence-parallel velocities to orogen width (i.e., v/L) at the time of collision, $t = -58$ myr, using external constraints from kinematic plate reconstructions. At ca. 58 Ma, India converged with the Asian continent at a rate of $\sim 120 \pm 10$ mm/yr (Molnar and Stock, 2009; van Hinsbergen et al., 2011b) (Fig. 4). We use published datasets and models to speculate on the width of the deforming orogen at the time of India-Asia collision following four 1D framework kinematic models (Fig. 4).

The simplest kinematic framework is that the Indian and Asian plates collided and converged since ca. 58 Ma, resulting in crustal shortening across the orogen that reflects bulk plate convergence. That is, decelerating plate convergence rate should be matched by decelerating

crustal shortening rate. Other kinematic scenarios are possible and have been published on, such that India-Asia convergence does not translate directly to continental shortening because of potential multiple arc-subduction systems that partition India-Asia convergence between oceanic subduction and continental collision/convergence. Here we outline three of these alternative scenarios: (1) the oceanic Greater India Basin hypothesis (van Hinsbergen et al., 2012), (2) the Xigaze backarc basin model (Kapp and DeCelles, 2019), and (3) a broad group of double-subduction models (e.g., Roy, 1976; Aitchison and Davis, 2004; Jagoutz et al., 2015). The Greater India Basin hypothesis suggests that a microcontinent rifted from northern India, forming an ocean basin between India and this microcontinent. This microcontinent then collided with India first, which was followed by subduction of the Greater India ocean basin and eventual collision of India and Asia in the early Miocene (van Hinsbergen et al., 2012, 2019). In contrast, the Xigaze backarc basin model proposes that the Xigaze arc system rifted from the southern Asian margin at ca. 90 Ma, opening the Xigaze backarc ocean basin. Closure of this backarc ocean basin started at ca. 70 Ma via north-dipping subduction beneath the Lhasa-southern Asia margin (Kapp and DeCelles, 2019). The northward moving Indian plate collided first with this rifted Xigaze arc. This “soft” collision was followed by closure of the backarc ocean basin and finally a “hard” collision between the composite India-Xigaze arc and Asia at ca. 45 Ma. Lastly, the third scenario encompasses a group of similar double-subduction models that suggest fast convergence between India and Asia during the Cretaceous-early Cenozoic was a result of the integrated effects of two arc-subduction systems, thus requiring lower subduction rates at each arc. The double-subduction models suggest that the final India-Asia collision occurred at ca. 40 Ma (e.g., Jagoutz et al., 2015) or later in the Cenozoic (e.g., Aitchison and Davis, 2004).

All of these alternative kinematic models imply that early Cenozoic India-Asia convergence was accommodated by some component of

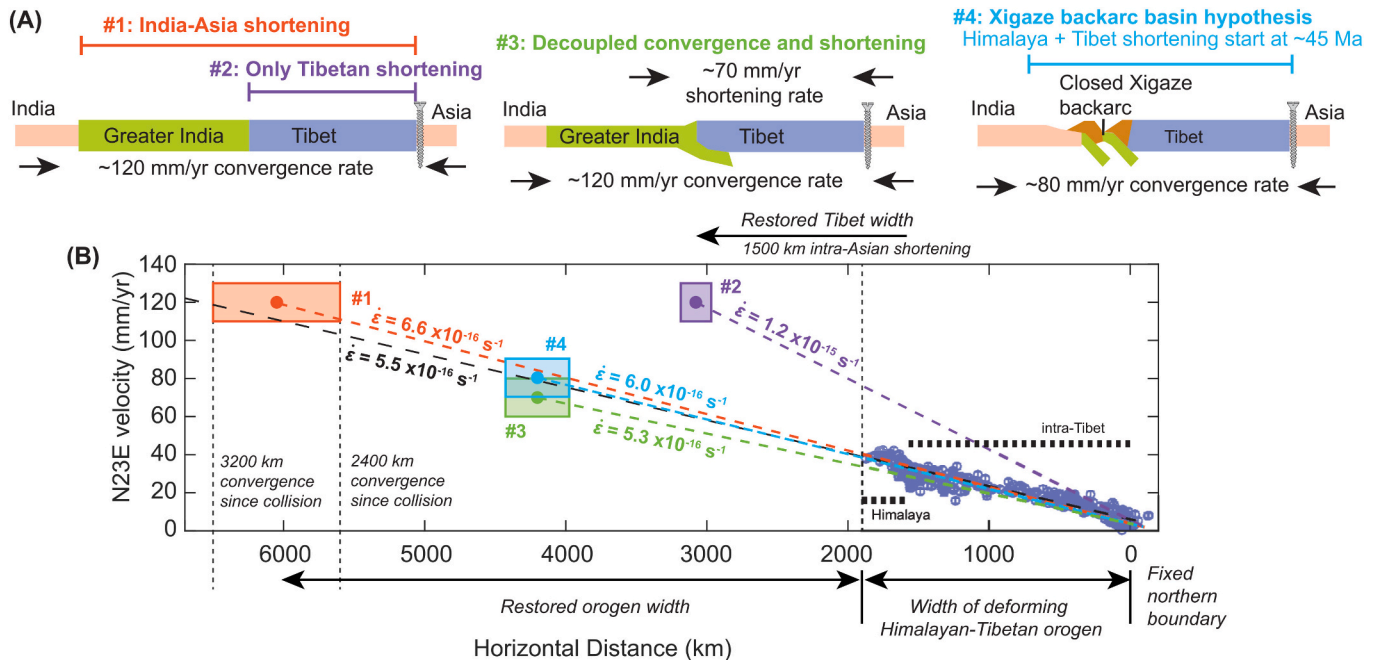


Fig. 4. (A) Four kinematics models of India-Asia convergence rate and orogen width at the time of initial collision (i.e., ca. 58 Ma for all models except #4). The deforming orogen may have consisted of the entirety of Greater India and Tibet or just Tibet north of the India-Asia suture (Models #1 and #2, respectively). Total orogenic shortening may have been lower if convergence was only partially coupled with Himalaya and intra-Asian (Tibetan) shortening (Model #3). Model #4 is based on the Xigaze backarc basin model of Kapp and DeCelles (2019), which involves collision of the composite India-Xigaze arc and Asia at ca. 45 Ma. See text for more explanation. (B) Plot of present-day GPS velocities projected on a N23E transect as shown in Fig. 2a, with predicted convergence rates ($v/L(t = -58 \text{ myr})$) and orogen widths ($L(t = -58 \text{ myr})$) at the time of collision (Molnar and Stock, 2009; van Hinsbergen et al., 2011a; Huang et al., 2017) at ca. 58 Ma for the four models (shaded boxes; see Fig. 4A). The bulk orogen shortening rate divided by the orogen width equates to the infinitesimal strain rate. Note that $x=0$ is the northernmost margin of the Tibetan Plateau, and the southern limited of the HTO is $x = -1,900$ km. GPS velocity data is from Ge et al. (2015). The present-day bulk strain rate is a best-fit line to the GPS velocity data (black dashed line), which has been extrapolated to the left.

oceanic subduction, and consequently India-Asia plate convergence required less intracontinental shortening recorded in the geological data. Therefore, these models predict lower bulk orogenic strain rates. However, in the interest of focusing on first-order deformation trends and due to the recent controversy of some of these alternative models (e.g., Rowley, 2019), we concentrate primarily on the simplest kinematic model where the HTO was generated by India-Asia continental collision at ca. 58 Ma and sustained convergence throughout the Cenozoic. To explore some of the alternatives, we also provide a model assuming the framework of the Xigaze backarc basin model (Kapp and DeCelles, 2019). Four framework models are discussed below. In a 1D framework, we assume that the present-day total orogen width is 1,900 km, partitioned into ~325 km of the Himalaya-Tethyan fold-thrust belt south of the suture zone and 1,575 km for the rest of the intra-Tibet thrust belts to the north (Fig. 4).

Model #1 assumes that total India-Asia convergence since collision translated to crustal shortening across the entire orogen, including Greater India (i.e., the Himalaya and Tethyan Himalaya fold-thrust belts; Ratschbacher et al., 1994; Murphy and Yin, 2003; Laskowski et al., 2017) and Tibet. We use the paleomagnetic convergence magnitude discussed in van Hinsbergen et al. (2011a, 2011b) and Huang et al. (2017) of $4,150 \pm 450$ km, which can be added to the present-day orogen width of ~1,900 km. This yields a restored orogen width of $6,050 \pm 450$ km that has been involved in collision-related deformation at ca. 58 Ma. The ~120 mm/yr of India-Asia convergence at this time was distributed across the entire orogen. This yields a bulk strain rate was $\sim 6.6 \times 10^{-16} \text{ s}^{-1}$ (Fig. 4), which overlaps the present-day strain rate across the orogen (Clark, 2012; Fig. 2). Extrapolation of the linear regression to the GPS data shows that the best-fit line intercepts the convergence rate and predicted orogen width at the onset of collision (Fig. 4B), further suggesting a potential constant strain rate throughout the Cenozoic. The key assumption for this model is that deformation involved the entirety of Greater India and Tibet since the onset of collision, such that the total width of the deforming orogen effectively involved ~6,050 km of crust since ca. 58 Ma (Fig. 4A). This condition was implied in the model of Clark (2012) and represents an end-member case of potential 1D orogen kinematics. Model #1 does not specifically differentiate between Himalaya and intra-Tibet deformation, although shortening rate and magnitude of these domains would scale according to their present-day length scales. If the total deforming orogen at ca. 58 Ma was narrower and accommodating the same convergence rate, the modeled strain rate would increase. This is represented by Model #2, discussed next.

The kinematics of Model #2 confine the orogen at the onset of collision between the Greater India-Asia suture zone and the northern edge of the Tibetan Plateau ($x = 0$ km). Thus, the total effective orogen width at ca. 58 Ma was the sum of the present-day orogen width north of the suture zone (i.e., ~1575 km) and the Cenozoic intra-Asian shortening magnitude of ~1,500 km (e.g., Zuza et al., 2018, 2019; Huang et al., 2017; van Hinsbergen et al., 2019; cf. Yakovlev and Clark, 2014). Summing these values yields a restored orogen width of roughly 3,075 km at ca. 58 Ma (Fig. 4). If we assume that fast ~120 mm/yr convergence rates were accommodated across this orogen width at the onset of collision, we derive a fast bulk strain rate of $\sim 1.2 \times 10^{-15} \text{ s}^{-1}$ (Fig. 4), which is almost an order of magnitude higher than the present-day strain rate (i.e., $5.5\text{--}6.6 \times 10^{-16} \text{ s}^{-1}$). This model requires strain rates to decelerate sometime during the Cenozoic to match the observed present-day values and not yield orogenic shortening that exceeds constraints for greatest total (i.e., restored) orogen width. Specifically, if the bulk strain rate from Model #2 was applied for 58 myr, it would yield >15,000 km of orogenic shortening, which is greater than any plate reconstructions allow (e.g., Huang et al., 2017; van Hinsbergen et al., 2019).

Model #3 is based on the decoupling of plate convergence and crustal shortening rates. Although India may have converged with Asia at fast rates (>100 mm/yr) in the early Cenozoic, part of the convergence may have involved relatively “clean” underthrusting or continental subduction (e.g., Ingalls et al., 2016; Cowgill et al., 2016), such

that the budget of Himalaya and intra-Asian shortening rates was significantly slower than the total India-Asia plate convergence rates. As an example, we show in our model that the Himalaya and Tibet accommodated $\sim 70 \pm 10$ mm/yr of the India-Asia plate convergence rate of the total budget of ~120 mm/yr at ca. 58 Ma. This shortening rate applied over the restored width of the southern Asian margin (~3,075 km, derived above) plus the restored Himalayan width of ~1,125 km (i.e., present-day 325 km width plus ~800 km shortening), based on minimum shortening estimates from field-based studies (e.g., Coward and Butler, 1985; Long et al., 2011; Webb et al., 2011), yields strain rates of $5.3 \times 10^{-16} \text{ s}^{-1}$ (Fig. 4). This value is comparable to that of Model #1 and present-day bulk strain rates.

Model #4 is based on the Xigaze backarc basin hypothesis (Kapp and DeCelles, 2019), which involves the collision of the composite India-Xigaze arc continent with Asia at ca. 45 Ma following the closure of the Xigaze backarc. This model implies that India-Asia convergence from 58 Ma to 45 Ma was primarily partitioned into oceanic subduction of the Xigaze backarc beneath southern Asia, and thus the fastest phases of convergence did not drive continental shortening related to major collisional orogeny. For this kinematic framework, we apply ~80 mm/yr of India-Asia convergence starting at ca. 45 Ma, with deformation distributed across the restored width of southern Asia (~3,075 km) and the Himalaya (~1,125 km). This framework yields strain rates of $6.0 \times 10^{-16} \text{ s}^{-1}$ (Fig. 4), similar to Model #1 and present-day bulk strain rates.

It is worth noting that Model #1 is based on paleomagnetic convergence magnitudes, whereas Models #2–4 are based on reconstructions of Himalayan and intra-Asian Cenozoic shortening. The discrepancies between the two restorations (i.e., ~6,050 km vs ~4,200 km) led to the proposal of the Greater Indian Basin hypothesis (e.g., van Hinsbergen et al., 2012). The “missing” ~1,850 km of plate convergence may have been cleanly subducted as oceanic or continental crust (cf. Ingalls et al., 2016) or may not be required due to efficient backarc basin closure and later India-Asia collision as in the Xigaze backarc model (Kapp and DeCelles, 2019). This may also result from imprecise field-based minimum shortening tabulations in the Himalaya due to erosion, underthrusting, intra-unit deformation, and/or poor constraints on early-middle Cenozoic shortening records, such that the size of the pre-collisional Greater India (i.e., Himalaya and Tethyan Himalaya) was greater than the 1,125 km width assumed here.

From the predictions of these models, it is clear there is systematic tradeoff between India-Asia convergence rate and orogen width, which affects estimates of bulk strain rates. Fig. 5 shows this relationship and how Models #1, #3, and #4 yield apparently similar strain rates, even if the convergence rates and orogen widths vary, whereas Model #2 results in significantly faster strain rates. Another implication drawn from the plot in Fig. 5 is that if convergence/shortening rates and orogen width show comparable decreases, the resulting strain rate will remain roughly constant. This was part of the basis for the conclusions reached by Clark (2012).

3. Time-varying HTO shortening rates

The apparent similarity of derived bulk strain rates today and at ca. 58 Ma, with either Model #1 or Model #3, suggests that strain rate may have remained constant across the HTO throughout the Cenozoic. Assuming this, one can predict time-varying bulk orogenic shortening during the Cenozoic. However, 1D orogenic shortening may be accommodated by a variety of processes, most simply via contractional deformation (e.g., folding and reverse faulting) and associated crustal thickening, or lateral out-of-plane motion. Here we argue that HTO shortening was mostly accommodated via convergence-parallel crustal shortening, and therefore our model of time-varying orogenic shortening provides constraints for crustal shortening across the orogen. Although extrusion-type tectonics have long been postulated for the HTO (e.g., Tapponnier et al., 1982; Peltzer and Tapponnier, 1988),

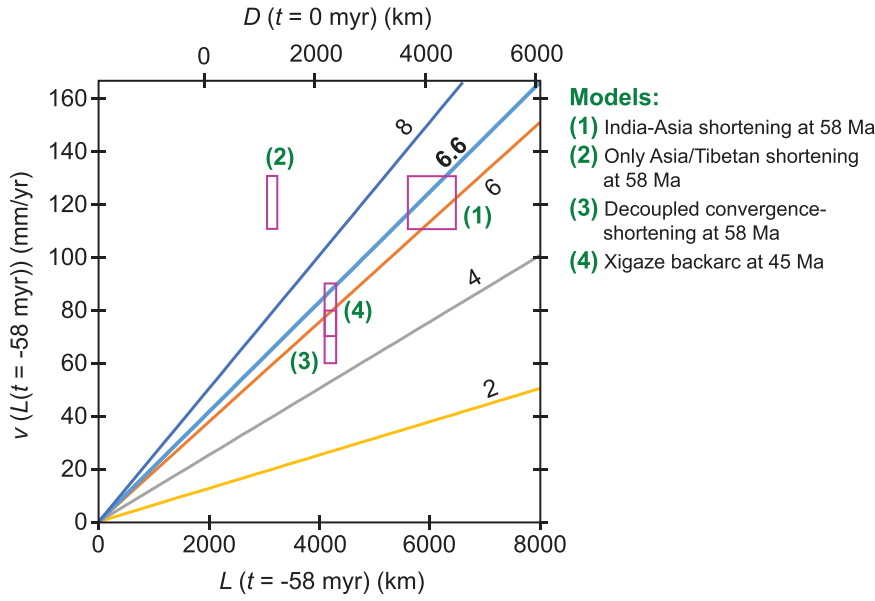


Fig. 5. Convergence (v) and orogen width (L) at 58 Ma ($t = -58$ myr), except Model #4 is at 45 Ma ($t = -45$ myr), plotted with contoured strain rates $\times 10^{-16} \text{ s}^{-1}$. Also shown are the ranges of estimates of v and L (or total shortening D , which is $L(t = -58 \text{ myr}) - L$) for the four kinematic models discussed in the text. For strain rates to have remained roughly constant during the Cenozoic, decreases in v must have been matched by decreases in L , such that v/L would follow a single strain rate contours.

geologic investigations of the major strike-slip faults in central and northern Tibet generally refute wholesale extrusion of the Tibetan Plateau region. Specifically, (1) strike-slip displacements and slip rates are too low (i.e., <100 km and <10 mm/yr, respectively) to fit these extrusion models, and (2) slip rates and displacements approach zero at the eastern and western tips of these faults (e.g., Zhang et al., 1991; Kirby et al., 2007; Duvall and Clark, 2010; Zuza and Yin, 2016; Zuza et al., 2018; Li et al., 2020).

Available age constraints suggest that strike-slip and normal faults within the HTO that accommodate west-east stretching orthogonal to plate convergence initiated at ca. 15–10 Ma (Fig. 1C; e.g., Duvall et al., 2013; Styron et al., 2013, 2015; Wu et al., 2019; Li et al., 2019). Therefore, these Cenozoic structures formed relatively late in the history of the HTO and were not predominant structures in the development of the orogen. In southern Tibet, extensional strain is rather low ($<10\%$; e.g., Bian et al., 2020). The V-shaped conjugate strike-slip faults in central Tibet associated with eastward lithospheric flow were generated via either gravitational spreading or lower crust-upper mantle shear (Yin and Taylor, 2011) and thus are not explicitly accommodating India-Asia convergence. Integrating ca. 15–10 myr of west-east stretching at the present-day strain rate of $\sim 2 \times 10^{-16} \text{ s}^{-1}$ (Fig. 3) would only lengthen the orogen by ≤ 250 km, compared to 1000+ km of shortening during the life span of the HTO. Accordingly, we argue that for most of the Cenozoic, 1D shortening across the HTO was accommodated primarily by crustal shortening and related crustal thickening and erosion. By neglecting out-of-plane motion (i.e., east-directed orthogonal stretching), crustal shortening derived from bulk orogen contraction represent maximum estimates as some convergence—we argue here very little—may have been accommodated by lateral fault slip and crustal motion.

Earthquake focal mechanisms across the plateau classically show reverse faulting across most of the plateau's margins, with normal and strike-slip faulting concentrated at the thickest crustal regions of central Tibet (e.g., Molnar et al., 1993; Copley et al., 2011). Therefore, plate convergence along the southern and northern plateau margins is primarily accommodated by crustal shortening. In a subsequent section, we use bulk orogenic strain rates to model crustal shortening rates across two plateau-bounding fold-thrust belts (i.e., the Himalaya and Qilian Shan-Nan Shan), where we contend that out-of-plane extrusion is minimal.

With these potential limitations noted, the inferred nearly constant strain rate across the HTO can be used to approximate long-term crustal

shortening rates across the orogen (Fig. 6) (Zuza et al., 2019). The present-day velocity field, which shows a linear decrease in the N23E-parallel GPS velocities from the Himalaya to northern Tibet (Fig. 2), implies that orogenic contraction is evenly distributed across the orogen. We infer that the constant v/L ratio should roughly scale from across the entire Himalayan-Tibetan orogen to smaller intra-orogen deformation belts (Fig. 6). These relationships suggest that bulk strain across the orogen should scale to individual thrust belts—i.e., they have the same slope at varying scales, as in Fig. 6—which allows for reconstructed estimates of the pre-deformed widths of orogens or discrete thrust belts, L , at all scales. We recognize that this method is a simplification, and strain may be partitioned locally within specific domains during the evolution of the orogen, but this approach provides meaningful first-order estimates of time-varying crustal shortening.

The convergence-parallel width of the HTO can be quantified as

$$L(t) = L_t + s(t) \quad (1)$$

where $s(t)$ is the length of the orogen that has been progressively shortened during collision to nothing today, such that $s(t = 0 \text{ myr}) = 0$ km. Given that L_t is a constant and $\dot{\epsilon} = v(x = L(t))/L(t)$, it follows that

$$dL/dt = ds/dt = v(x = L(t)) \quad (2)$$

and therefore

$$ds/dt = \dot{\epsilon} \times L(t) = \dot{\epsilon}(L_t + s(t)) \quad (3)$$

Solving equation (3), with the condition $s(t = 0 \text{ myr}) = 0$ km, yields the following relationship:

$$s(t) = L_t(e^{-\dot{\epsilon}t} - 1) \quad (4)$$

Combining equations (1) and (4) yields the width of the HTO as a function of time:

$$L(t) = L_t e^{-\dot{\epsilon}t} \quad (5)$$

and the orogen-scale shortening rate is

$$v(x = L(t)) = \dot{\epsilon} L_t e^{-\dot{\epsilon}t} \quad (6)$$

With this simple kinematic framework, we aim to use a constant bulk strain rate to model shortening rates and total shortening magnitude as a function of time. Of the four models discussed above (Fig. 4), we focus primarily on Model #1 ($\sim 6,000$ km wide orogen including Greater India and Tibet at 58 Ma) and Model #4 (Xigaze backarc basin hypothesis).

Implications of constant strain rate for bulk shortening and orogen width

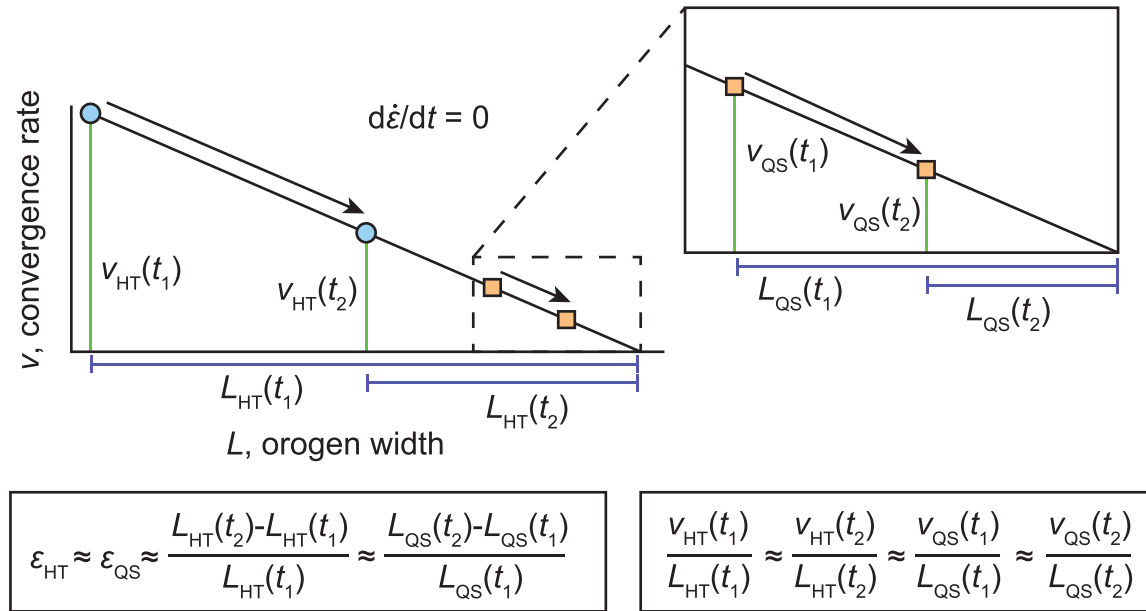


Fig. 6. Schematic plots showing that the consequence of an orogen with a roughly constant strain rate and a linearly decreasing across-strike velocity profile (e.g., Zhang et al., 2004) predicts the strain and strain rate should scale from across the entire Himalayan-Tibetan orogen (blue circles) to individual intra-orogen thrust belts (orange squares). Thus, strain (ϵ) across the entire orogen should approximately equal strain across smaller fault systems. For illustrative purposes, these plots are not to scale; t_1 is some time in the past, and t_2 is today.

We exclude Model #2 (strain rate of $\sim 10^{-15} \text{ s}^{-1}$) from this analysis because it would result in unrealistic shortening strain values if such a fast strain rate remained constant throughout the Cenozoic. The strain rate in Model #2 must have decreased at some point during the Cenozoic to $\sim 10^{-16} \text{ s}^{-1}$, but the timing of this decrease is unresolvable with current datasets. Thus, we prefer to model the simpler constant-strain rate models. Model #3 (i.e., decoupled convergence and intra-Asian shortening) overlaps the strain rate in Model #1 (Fig. 4) but involves an admittedly arbitrary assumed ratio between total India-Asia convergence and intra-Asian shortening (i.e., we chose $\sim 70 \text{ mm/yr}$ shortening for 120 mm/yr convergence). Therefore, we do not specifically focus on this model, but note that a first-order shortening curve from Model #3's kinematics would overlap with the shortening estimates from the Model #1 model.

Total convergence and shortening rates across the HTO as a function of time are plotted in Fig. 7 following the kinematics of Model #1. Applying a constant strain rate yields shortening/convergence rates that

match plate-circuit reconstructions (Fig. 7), including a decrease in convergence velocities of $\sim 120 \text{ mm/yr}$ at 58 Ma to $\sim 40 \text{ mm/yr}$ today (Molnar and Stock, 2009; van Hinsbergen et al., 2011a). Furthermore, total convergence since initial collision at ca. 58 Ma is estimated to be $\sim 3,000\text{--}4,000 \text{ km}$ (Fig. 4), which is comparable to published estimates cited in Molnar and Stock (2009) and Huang et al. (2017). Model #4 kinematics (black lines in Fig. 7) provide overlapping estimates.

Application of a constant strain rate across the HTO since collision matches external kinematic constraints from plate reconstructions, as previously discussed by Clark (2012). Below, we assume a constant strain, combined with observations that convergence-parallel GPS velocities decrease linearly from the modern plate boundary in the southern Himalaya to northern Tibet, to estimate shortening rates and magnitudes across the orogen through time (Fig. 6).

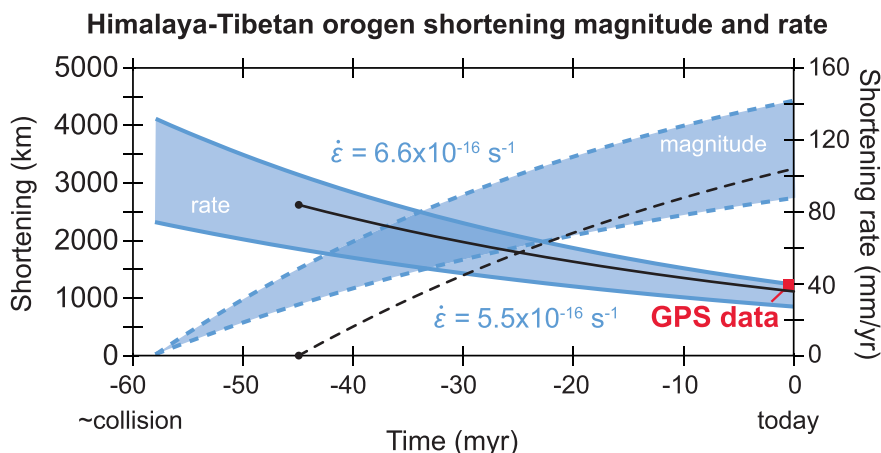


Fig. 7. Modeled time-varying shortening rate (solid lines) and magnitude (dashed lines) across the Himalayan-Tibetan orogen (HTO), with envelop bounds resulting from using two strain rates: $6.6 \times 10^{-16} \text{ s}^{-1}$ (derived from Model #1 in Fig. 4 and Clark, 2012) and $5.5 \times 10^{-16} \text{ s}^{-1}$ (derived from GPS data) (final orogen width, L_t , is $1,900 \text{ km}$) in blue. Black lines show calculations assuming Model #4 kinematics, with deformation starting at ca. 45 Ma and a strain rate of $6.0 \times 10^{-16} \text{ s}^{-1}$. The GPS orogenic shortening rate is from Ge et al. (2015).

4. Modeled shortening across Tibet's northern and southern margins

We integrate our derived bulk strain rate estimates to predict shortening rates across the southernmost and northernmost deformation belts of the HTO: the Himalayan fold-thrust belt in the south and Qilian Shan-Nan Shan thrust belt in the north (Fig. 1). The Himalayan fold-thrust belt analysis is intended to broadly represent deformation south of the India-Asia suture zone, including the Himalayan and Tethyan Himalaya fold-thrust belts (e.g., Ratschbacher et al., 1994; Murphy and Yin, 2003; Kapp et al., 2007; Laskowski et al., 2017). The Qilian Shan-Nan Shan thrust belt analysis consists of the thrust system between the Qaidam Basin and Hexi Corridor (e.g., Meyer et al., 1998; Zuza et al., 2016). We focus our analyses on these bounding thrust belts for three reasons. First, they have relatively robust constraints on the magnitudes and rates of Cenozoic shortening from both surface geology and sub-surface geophysical data. Published estimates of crustal shortening can be compared directly against kinematic predictions of our strain-rate analysis. Shortening magnitudes and rates across the HTO are presently limited in spatial and temporal resolution (see van Hinsbergen et al., 2011a; Yakovlev and Clark, 2014). Second, these thrust-belt domains bound the HTO and are relatively discrete in that crustal deformation becomes negligible south of the Himalaya and north of the Qilian Shan-Nan Shan. The Hexi Corridor at the northern edge of the Tibetan Plateau (Fig. 1C) is assumed as a relatively fixed position as India and Greater India have progressively converged northward since the onset of collision (Fig. 1). Therefore, the change in HTO width can broadly be scaled to a change in the width of these two bounding thrust belts due to progressive shortening. Lastly, as already discussed, these regions are dominated by contractional structures. Therefore, bulk 1D contraction, as modeled in this work, was most likely accommodated dominantly via crustal shortening that can be compared with geological investigations of NNE-SSW directed crustal shortening magnitude and rate.

A similar analysis for northern Tibet was provided in Zuza et al. (2019) using constant strain rates of 7 to $5 \times 10^{-16} \text{ s}^{-1}$. For consistency with this study, we produced new models using strain rates of 6.6×10^{-16}

s^{-1} and $5.5 \times 10^{-16} \text{ s}^{-1}$. Deformation in northern Tibet initiated shortly after initial India-Asia collision, starting in the Eocene or possibly Paleocene (Jolivet et al., 2001; Yin et al., 2008a, 2008b; Craddock et al., 2011; Li et al., 2020). Available quantitative estimates suggest that rapid bedrock cooling and/or local foreland-basin deposition were underway by ca. 40 Ma (e.g., Jolivet et al., 2001; Craddock et al., 2011; Yin et al., 2008a, 2008b; Duvall et al., 2011; Li et al., 2020). We model the shortening rate and magnitude for northern Tibet across the Qilian Shan-Nan Shan thrust belt (Fig. 1) using equations (3) and (4), assuming that shortening either initiated in northern Tibet immediately at the time of collision at ca. 58 Ma or there was some lag time until ca. 40 Ma while strain propagated northward from the India-Asia collisional front. The choice of ca. 40 Ma for delayed thrust-related deformation is based on a broad peak population of cooling ages in the Qilian Shan as recently synthesized in Li et al. (2020). We assume strain rates remained constant through the Cenozoic. Based on long-term geologic shortening estimates from cross-section restorations, Zuza et al. (2018) suggested that the central Qilian Shan experienced geologic strain rates (10 's myr rates) of $2\text{--}7 \times 10^{-16} \text{ s}^{-1}$, which overlaps the present-day strain rate values derived from GPS velocities. This similarity at different timescales supports our assertion that Qilian Shan strain rates may have remained constant during the Cenozoic.

Our calculated time-varying shortening rate and magnitude use a present-day length of the Qilian Shan-Nan Shan thrust belt (L_f) of 350 km. For an early collision initiation, equation (5) suggests that length of the restored thrust belt $L(t = -58 \text{ myr})$ was $\sim 950\text{--}1,150$ km wide and for a delayed initiation, $L(t = -40 \text{ myr})$ was $\sim 700\text{--}800$ km (Fig. 8). These predicted restored thrust belt widths since ca. 58 Ma and 40 Ma suggest 55–85% shortening strain ($\sim 650\text{--}800$ km) and 35–65% shortening strain ($\sim 300\text{--}450$ km), respectively, since the onset of deformation. Fig. 8 predicts that shortening rates across the Qilian Shan have decreased from $\sim 17\text{--}24$ mm/yr at ca. 58 Ma to $17\text{--}12$ mm/yr at ca. 40 Ma to $\sim 6\text{--}7$ mm/yr today. Geodesy suggests contemporary shortening rates of $\sim 6 \pm 1$ mm/yr (Zhang et al., 2004) (Fig. 8).

For the Himalaya (including the Tethyan Himalaya fold-thrust belt), we assumed that collision-related crustal shortening initiated at the time

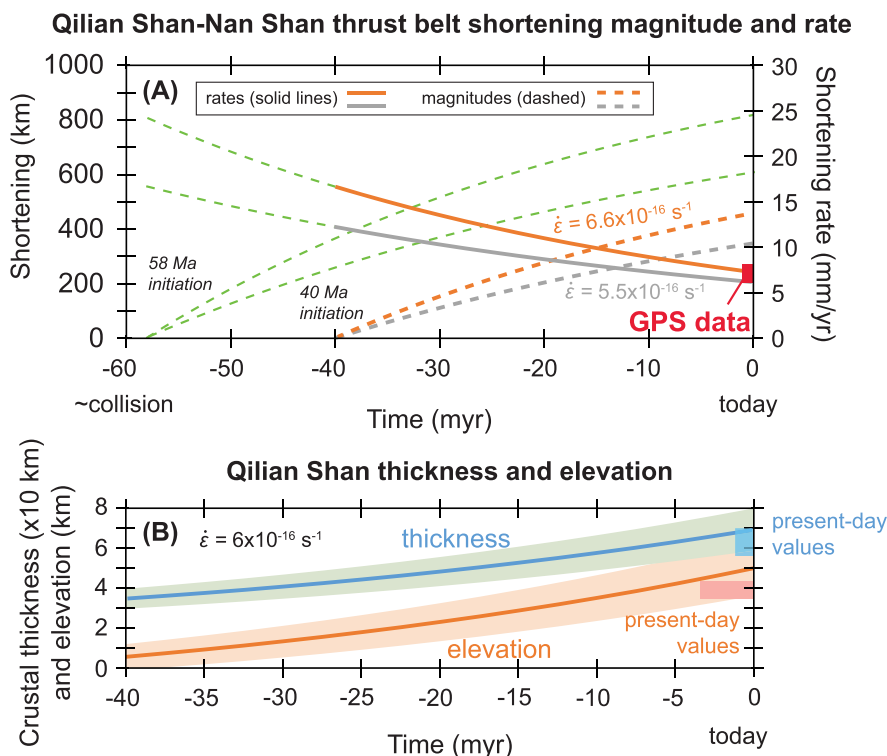


Fig. 8. (A) Modeled time-varying shortening rate (solid line) and magnitude (dashed line) using constant bulk strain rates across the Qilian Shan-Nan Shan thrust belt. Green dashed lines are models invoking ca. 58 myr of deformation (i.e., shortening since the onset India-Asia collision), whereas the grey, blue, and orange curves show modeling of shortening initiating at ca. 40 Ma, consistent with geological observations. GPS data is from Ge et al. (2015). (B) Modeled Qilian Shan crustal thickness starting with an initially 35-km-thick crust subjected to the shortening modeled in Fig. 8A and predicted elevation following this crustal thickening, given isostasy. The model applies ~ 5 km of erosion averaged over the time of deformation. An initial uncertainty of ~ 5 km on crustal thickness is used to add uncertainty bounds to both curves. Present-day crustal thickness is from Yue et al. (2012).

of India-Asia collision at ca. 58 Ma. Constraints for Himalaya-Tethyan deformation in the early Cenozoic are relatively sparse and mostly consist of prograde metamorphic ages (e.g., Smit et al., 2014; see Kapp and DeCelles, 2019 review). From equations (3) and (4), shortening rates and magnitudes as a function of time across the Himalaya were calculated, using a modern thrust belt width (L_f) of 325 km (Fig. 9A). The model predicts a total of ~550–750 km shortening (~61–68% strain) (Fig. 9A). Modeled shortening rates yield a decrease from ~15–22 mm/yr at 58 Ma to 6–7 mm/yr today (Fig. 9A). In the central Himalaya, which approximately represents our selected transect (Fig. 1), our strain-rate modeled present-day shortening rates are significantly lower than observed geodetic shortening rates of 25–20 mm/yr (e.g., Banerjee et al., 2008), long-term geological shortening rates of ~20 mm/yr (DeCelles et al., 2001), or Pleistocene-Holocene shortening rates of 21–18 mm/yr (Mugnier et al., 1999; Lavé and Avouac, 2000) (Fig. 9A). However, several authors have published estimates of pre-Miocene Himalaya-Tethyan fold-thrust belt shortening (249–133 km Ratschbacher et al., 1994; DeCelles et al., 2001) that yield low shortening rates (<10 mm/yr), which have subsequently increased to ~20 mm/yr throughout the Neogene-Quaternary (Lavé and Avouac, 2000; Mugnier et al., 2004).

Other localities along strike of the Himalaya that appear to deviate

from the high shortening rates in Nepal are consistent with our modeled modern shortening rates. In Bhutan, long-term geologic shortening rates on Himalayan faults are between 7 mm/yr and 10 mm/yr since late Cenozoic time (10–0 Ma) derived from cross-section restoration (Long et al., 2012; McQuarrie and Ehlers, 2015). These rates contrast to Holocene shortening rates (Burgess et al., 2012) and geodetic shortening rates (Banerjee et al., 2008) of 25–20 mm/yr in the neighboring easternmost Himalaya. In the northwest Himalaya, deformation is partitioned on multiple faults and late Cenozoic convergence rates are 7–14 mm/yr (Powers et al., 1998; Dey et al., 2016). Similarly, these rates are lower than geodetic convergence rates of 18–20 m/yr in the same area (Jade et al., 2004; Banerjee et al., 2008). In the Kashmir Himalaya, estimates of long-term geologic (ca. 10–0 Ma), Pleistocene-Holocene, and modern convergence rates range between conservative estimates of 5 mm/yr and 10–13 mm/yr (Schiffman et al., 2013; Kundu et al., 2014; Gavillot, 2014; Gavillot et al., 2016, 2018), which broadly overlap with our modeled shortening rates.

Published estimates of crustal shortening across the Himalaya since the onset of collision vary between 500 km and 900 km (e.g., Coward and Butler, 1985; DeCelles et al., 2001, 2002; Yin et al., 2010; Long et al., 2011; Webb et al., 2011). These values are consistent with our predicted ~550–750 km of total shortening derived from our strain rate

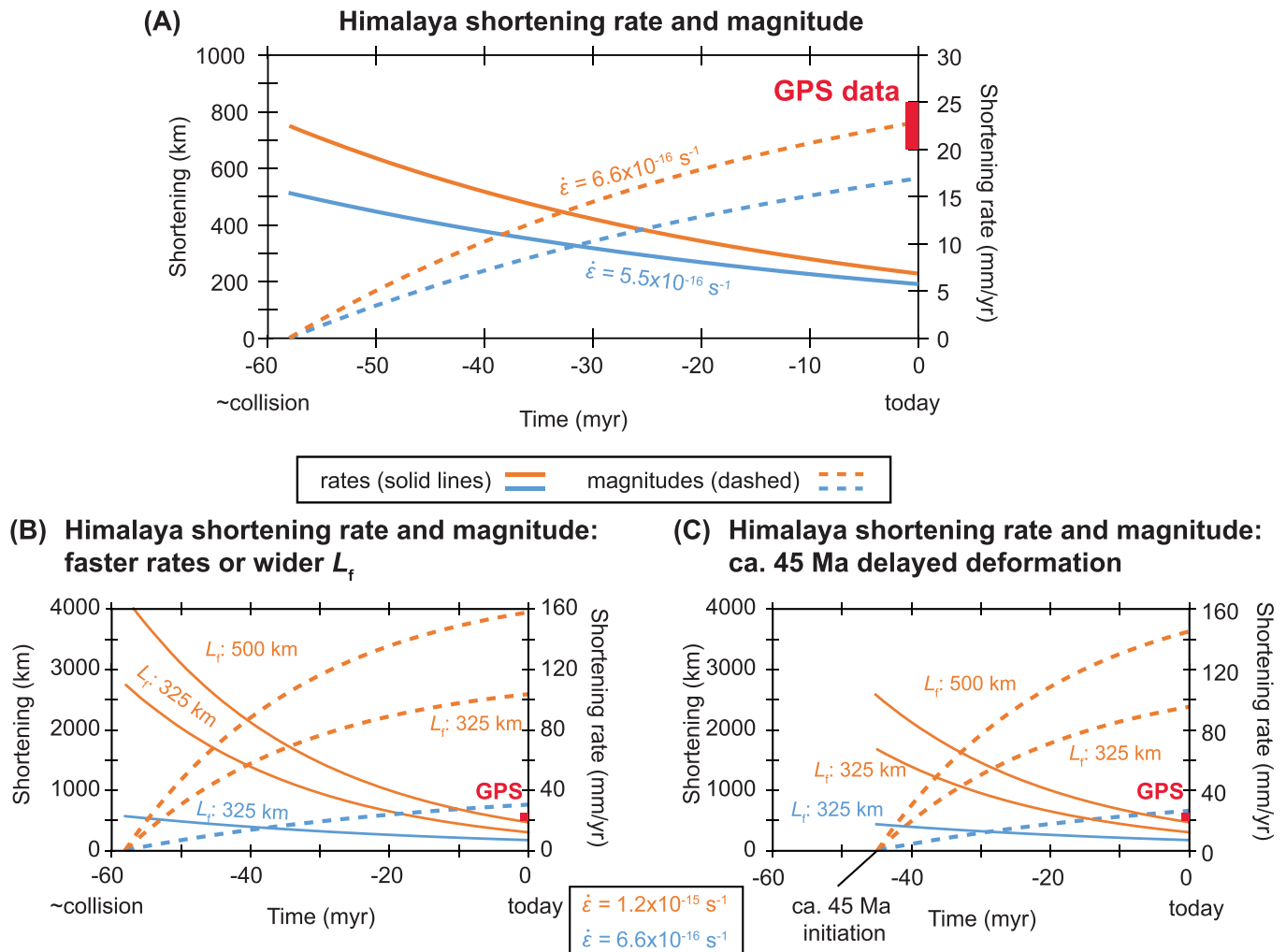


Fig. 9. (A) Modeled time-varying shortening rate (solid lines) and magnitude (dashed lines) with constant bulk strain rates across the Himalayan thrust belt. GPS data is from Banerjee et al. (2008). (B) Modeled time-varying shortening rate and magnitude using a faster bulk strain rate across the Himalayan thrust belt and a wider final orogen length (L_f). Note that these models result in early Cenozoic shortening rates that are faster than the India-Asia convergence rates and shortening magnitudes that are higher than those from palinspastic reconstructions. (C) Similar plot to Fig. 9B, but assuming that deformation started at ca. 45 Ma, as in Model #4 of this study.

model for strain rates between $5.5 \times 10^{-16} \text{ s}^{-1}$ and $6.6 \times 10^{-16} \text{ s}^{-1}$ (Fig. 9A).

Modeled Himalayan shortening rates and magnitudes (Fig. 9A) are generally consistent with observations from the northwestern and eastern Himalaya, but deviate from estimates from the central Himalaya. Because our modeled shortening rates yield lower predicted present-day shortening rate values compared with modern shortening-rate estimates across the central Himalaya, we explored the impacts of other parameters such as a faster bulk strain rate of $1.2 \times 10^{-15} \text{ s}^{-1}$, as derived in Fig. 2C from the cubic and piece-wise fits, and a wider final (modern length) thrust belt width ($L_f=500\text{km}$) (Fig. 9B). Some geologic cross-section reconstruction from the eastern Himalaya and central Nepal imply strain rates of $\sim 10^{-15} \text{ s}^{-1}$ (Khanal and Robinson, 2013; Yin et al., 2010; Happroff et al., 2018). Therefore, 10^{-15} s^{-1} may be a valid approximation of the middle-late Cenozoic strain rate affecting the Himalaya. Using such a higher strain rates results in a faster modeled shortening rate approaching 20 mm/yr at 0 Ma. However, application of these parameters over the entire Cenozoic evolution of the Himalaya imply >2,500 km of net shortening and early Cenozoic Himalayan shortening rates that are equal to, or faster than, India-Asia convergence rates at the onset of collision (i.e., $\geq 100 \text{ mm/yr}$) (Fig. 9B). These results are not well supported by available geologic data but illustrate the effects of the range of values used to constrain strain rate and orogen width on our 1D model.

Delayed onset of collision until ca. 45 Ma (e.g., Model #4) results in lower net shortening, but still predicts >2,000 km for strain rates of $1.2 \times 10^{-15} \text{ s}^{-1}$ and $\sim 700 \text{ km}$ for strain rates of $6.6 \times 10^{-16} \text{ s}^{-1}$ (Fig. 9C). A fast strain rate (10^{-15} s^{-1}) for the Himalaya applied to Model #4's kinematics (Fig. 9C) appears more plausible, based on comparison to geologic observations, than the 10^{-15} s^{-1} strain rate applied to Model #1 (Fig. 9B).

5. Discussion

5.1. Modeled shortening compared against observations

Some of our modeled thrust-belt-scale shortening rates and magnitudes agree with geologic observations, whereas others do not. Modeled total shortening across the Himalaya ($\sim 550\text{--}750 \text{ km}$) (Fig. 9A) overlaps with minimum crustal shortening estimates of $\sim 700 \pm 200 \text{ km}$ observed from geologic mapping and cross-section restorations (e.g., DeCelles et al., 2002; Yin et al., 2010; Long et al., 2011; Webb et al., 2011). However, our modeled shortening rates differ from short-term geologic and geodetic constraints across the central and eastern Himalaya. Present-day and Pleistocene-Holocene shortening rates are $\sim 25\text{--}20 \text{ mm/yr}$ (Lavé and Avouac, 2000; Banerjee et al., 2008; Burgess et al., 2012), which depart from our modeled rates of $\sim 6\text{--}7 \text{ mm/yr}$ (Fig. 9A). If we apply faster constant strain rates or wider deforming regions, our models can predict higher shortening rates closer to observed GPS velocities (Fig. 9B), but this leads to geologically unreasonable conditions, such as extremely high Cenozoic shortening magnitudes, as mentioned above. Conversely, it is possible that field studies may severely underestimate shortening magnitudes based on the limitations of structural reconstructions and the balanced cross section method.

Cenozoic to Pleistocene-Holocene geologic shortening rates across the Himalaya are spatially and temporally variable along strike. Rates determined for regions including the Bhutan, northwest India, and Kashmir Himalaya overlap with our modeled shortening rates ($5\text{--}14 \text{ mm/yr}$; Powers et al., 1998; Long et al., 2012; McQuarrie and Ehlers, 2015; Gavillot, 2014; Gavillot et al., 2016, 2018). Several published long-term geologic shortening rates since onset of collision imply deceleration of Himalaya shortening through time that parallels our model results. For example, a comparable deceleration has been reported from cross-section models across the Bhutan Himalaya since the early Miocene (e.g., Long et al., 2012; McQuarrie et al., 2014). To compare deceleration in shortening across Bhutan with our modeled shortening rates, we use the geologic shortening rates from Long et al.

(2012). The older shortening rates estimated in Long et al. (2012) (i.e., 23 to ca. 10 Ma rates) involve the assumption that shortening initiated at ca. 23 Ma. However, crystallization ages of Bhutan leucogranites are as old as ca. 34 Ma (Hopkinson et al., 2020), which implies that crustal thickening and underthrusting of hydrated sediments (e.g., Harrison and Wielicki, 2016) was active in the Bhutan Himalaya by at least ca. 30 Ma or earlier. Accordingly, we reinterpreted the Long et al. (2012) rates considering that shortening initiated at ca. 30 Ma. McQuarrie et al. (2019) provided updated thermochronology and thermal modeling to estimate temporal variations in shortening rate in the Bhutan Himalayan. In their study, shortening rates estimated for two time segments are faster than the corresponding India-Asia convergence rates, which we are skeptical of given that intra-Asian deformation was also occurring at this time and integrated crustal shortening rates should be equivalent to overall plate convergence rates. All of these aforementioned rates are plotted over our modeled shortening rates in Fig. 10A using both a fast and slow strain rate value (i.e., $1.2 \times 10^{-15} \text{ s}^{-1}$ and $6.6 \times 10^{-16} \text{ s}^{-1}$, respectively). Although this provides fairly coarse resolution, we find that that deceleration of shortening across the Bhutan Himalaya is broadly similar to our strain-rate model, and both predict present-day shortening rates of $<10 \text{ mm/yr}$ (i.e., $3\text{--}8 \text{ mm/yr}$ in Long et al., 2012 versus $6\text{--}7 \text{ mm/yr}$ for the $6.6 \times 10^{-16} \text{ s}^{-1}$ model; Fig. 10A).

The apparent mismatch between long-term geologic (ca. 10–0 Ma) shortening rates of $<10 \text{ mm/yr}$ derived from field studies and the Pleistocene-Holocene-present shortening rates ($\sim 20\text{--}25 \text{ mm/yr}$; e.g., Banerjee et al., 2008; Burgess et al., 2012) has been used as evidence to suggest that active shortening rates across the Himalaya have only very recently accelerated to $\sim 20\text{--}25 \text{ mm/yr}$ (e.g., Long et al., 2012; McQuarrie et al., 2014; McQuarrie and Ehlers, 2015). Our modeling using a constant strain rate may further support this hypothesis. In this sense, our modeled shortening rates of $5\text{--}10 \text{ mm/yr}$ are comparable to published geologic interpretations over the past 10 myr, and the mismatch between these shortening rates and geodetic rates requires further evaluation.

The modeled present-day shortening rate across the Qilian Shan is $\sim 6\text{--}7 \text{ mm/yr}$ (Fig. 8A), which is consistent with geodetically observed rate of $5\text{--}7 \text{ mm/yr}$ (Zhang et al., 2004). If deformation initiated at ca. 40 Ma in the Qilian Shan-Nan Shan, as opposed to ca. 58 Ma, our modeled shortening rates decreased from $\sim 16 \text{ mm/yr}$ to $\sim 7 \text{ mm/yr}$ and yield crustal shortening of $\sim 400 \text{ km}$ (Fig. 8). These values overlap with available estimates of crustal shortening of $\sim 300 \text{ km}$ and shortening rates across the Qilian Shan-Nan Shan thrust belt through the Cenozoic, with possible deceleration from $>8 \text{ mm/yr}$ prior to ca. 10 Ma and $6.2\text{--}7.4 \text{ mm/yr}$ between 10–0 Ma (Zuza et al., 2018, 2019; Fig. 8A). Alternatively, if the initiation of thrust-related deformation northern Tibet was synchronous with the onset of India-Asia collision at ca. 58 Ma, our modeled crustal shortening magnitude are nearly double what are interpreted from cross section restorations across the Qilian Shan-Nan Shan thrust belt (i.e., minimum shortening magnitudes of $\sim 250\text{--}350 \text{ km}$) (Meyer et al., 1998; Yin et al., 2008a, 2008b; Zuza et al., 2016, 2018, 2019; Allen et al., 2017). Because a delayed initiation of deformation for northern Tibet yielded model results consistent with published geological shortening magnitude and rate observations, we use this 40 Ma initiation age as our preferred model assumption.

Geologic reconstructions of shortening magnitudes and rates are minimum estimates, and erosion and/or cross-cutting and overprinting structures can obscure these values derived from field studies. Furthermore, these fold-thrust systems are often located in remote regions that are difficult to access. Therefore, field-based datasets are relatively incomplete. We argue that these 1D strain-rate-based model results provide a framework to be tested against future kinematic field studies.

5.2. Crustal thickening in northern Tibet

Our modeled shortening across the $\sim 350\text{--}450 \text{ km}$ -wide Qilian Shan-Nan Shan thrust belt is predicted to increase to $\sim 300\text{--}450 \text{ km}$ ($\sim 35\text{--}65\%$

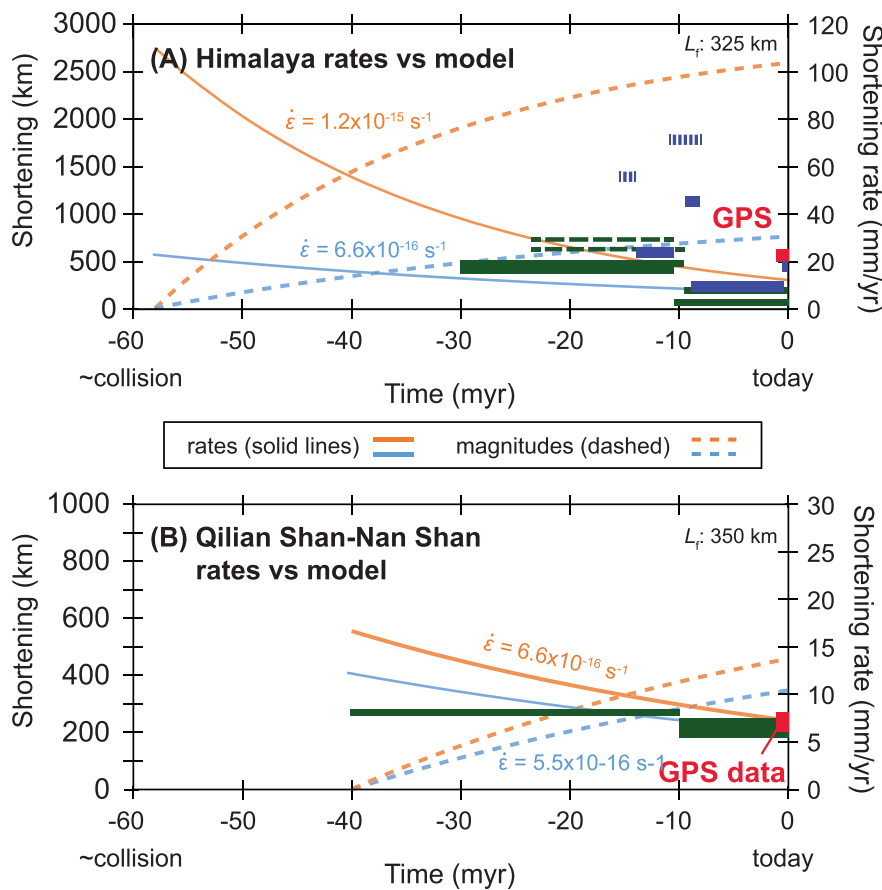


Fig. 10. (A) Modeled time-varying shortening rate and magnitude with constant bulk strain rates ($6.6 \times 10^{-16} \text{ s}^{-1}$ and $1.2 \times 10^{-15} \text{ s}^{-1}$) across the Himalayan thrust belt. Green bars are the Bhutan Himalaya shortening rates from Long et al. (2012); original >10 Ma rates (dashed line) were adjusted to account for earlier deformation and thus slower rates (solid green line). Blue bars are updated Bhutan Himalaya shortening rates from McQuarrie et al. (2019); dashed bars are not preferred shortening rates because they are faster than estimated India-Asia convergence rates. Note the deceleration in shortening rate observed with the geologic shortening rates and predicted by the strain-rate model. GPS data is from Banerjee et al. (2008). (B) Modeled time-varying shortening rate and magnitude applying constant bulk strain rates ($6.6 \times 10^{-16} \text{ s}^{-1}$ and $5.5 \times 10^{-16} \text{ s}^{-1}$) across the Qilian Shan-Nan Shan fold-thrust belt. Green bars are the Qilian Shan-Nan Shan shortening rates synthesized in Zuza et al. (2019). Note the deceleration in shortening rate observed with geologic shortening rates and predicted by the strain-rate model. GPS data is from Ge et al. (2015).

strain) by the present day, over a 40 myr time interval (Fig. 8). To demonstrate the effects of this protracted shortening on crustal thickening, we model pure-shear thickening of a crust, assuming that horizontal shortening strain directly equates to vertical thickening with no volume change. This simplifies potential strain partitioning or the presence of mid-crustal detachments (e.g., Wang et al., 2011b), but allows for first-order comparisons of shortening to vertical thickening (e.g., Yin et al., 1998; Lease et al., 2012). This exercise assumes that the crust was originally ~ 35 km thick prior to deformation (e.g., Zuza et al., 2016, 2019), which is a reasonable estimate for extended pre-Cenozoic crust (Christensen and Mooney, 1995). Erosion and exhumation are minimal in northern Tibet, as evidenced by mostly pre-Cenozoic low-temperature thermochronometric cooling ages across the Qilian Shan-Nan Shan (e.g., George et al., 2001; Jolivet et al., 2001; Baotian et al., 2013; Duvall et al., 2013; Qi et al., 2016; Li et al., 2020). Therefore, we apply ~ 5 km of total erosion evenly through time to our Qilian Shan model. Our derived time-varying shortening magnitudes thicken an initially ~ 35 -km-thick crust to 60–80 km in 40 myr (Fig. 8B). Crustal thickness estimates for northern Tibet are ~ 55 –70 km as constrained by seismic reflection (Gao et al., 2013) and receiver function (Yue et al., 2012; Ye et al., 2015) studies.

Assuming isostasy, this crustal thickness evolution can be related to surface elevation, E , following the relationship of Lee et al. (2015):

$$E = \left(1 - \frac{\rho_c}{\rho_m}\right)H - B \quad (7)$$

where B is a common base level at ~ 4 km below sea level, H is crustal thickness as a function of time, ρ_c is crustal density ($2,900 \text{ kg/m}^3$), and ρ_m is mantle density ($3,330 \text{ kg/m}^3$). Equation (7) adequately calculates elevation versus crustal thickness for most regions on Earth today (Lee et al., 2015). This curve (Fig. 8B) recreates the present-day elevations in

northern Tibet of 4–5 km. The lack of paleoaltimetry data in this region prohibits more robust testing through time (e.g., Deng and Ding, 2015).

5.3. West-east stretching driven by progressive crustal overthickening

At ca. 15–10 Ma there is a marked shift in deformations style across the HTO—with the contemporaneous initiation of strike-slip faulting in northern and central Tibet, and normal faulting in southern Tibet and the Tethyan Himalaya (Fig. 1C)—which may have been related to changes in boundary conditions, crustal thickness, rheology, or a combination of processes (e.g., Yin et al., 1999; Yuan et al., 2013; Sundell et al., 2013; Styron et al., 2015; Li et al., 2019; Bian et al., 2020). We test the impact of pure-shear thickening of the Tibetan crust, using a similar approach to the Qilian Shan example above (section 5.2), starting with a bulk crustal thickness of ~ 35 km. The pre-collisional crustal thickness of the Tibetan crust was probably highly spatially variable (e.g., Murphy et al., 1997; Yakovlev and Clark, 2014; Deng and Ding, 2015; Staisch et al., 2016; Cao et al., 2020; Tang et al., 2020), but we assume that a bulk value of ~ 35 km is reasonable for the averaged extent of the HTO. The model applies constant strain-rate shortening since ca. 58 Ma that thickens the Tibetan Plateau to its approximate present-day thickness of ~ 70 km by ca. 15 Ma (Fig. 11). This result suggests that pure-shear thickening of Tibet may have caused the plateau to obtain its present-day thickness by ca. 15 Ma, when strike-slip and normal faulting initiated across the HTO. These findings imply that protracted deformation under a constant strain rate may have thickened the crust to some critical thickness, where gravitational potential energy (GPE) exceeds applied boundary forces to drive deformation (e.g., Molnar et al., 1993; Jones et al., 1996; England and Molnar, 1997). This could have resulted in a reorientation of the intermediate principal compressive stress orientation from horizontal to vertical by ca. 15 Ma that shifted the

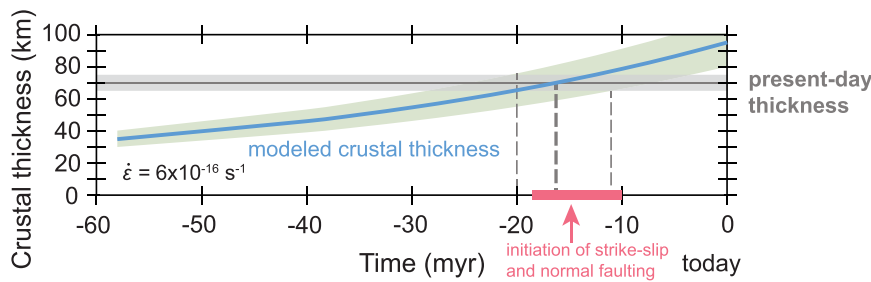


Fig. 11. Modeled crustal thickness of the Tibet region starting with an initially 35-km-thick crust subjected to a shortening strain rate of $6 \times 10^{-16} \text{ s}^{-1}$ since ca. 58 Ma. The model applies ~ 5 km of total erosion averaged over the time of deformation. An initial uncertainty of ~ 5 km on crustal thickness is used to add bounds to the curve. Note that the curve intersects the present-day crustal thickness of much of Tibet (~ 70 km) at ca. 15 Ma, which is when major strike-slip and normal faulting generally initiated across the HTO.

deformation regime (Duvall et al., 2013; Zuza and Yin, 2016; Allen et al., 2017; Li et al., 2019). Accordingly, this supports that the transition to strike-slip and normal-faulting in Tibet was not caused by a punctuated shift in boundary conditions or geodynamic response to a process like dense lithospheric removal (e.g., England and Houseman, 1989; Yuan et al., 2013), but rather the result from protracted deformation and crustal thickening.

5.4. Implications of delayed shortening in northern Tibet

Initially, our bulk-shortening estimates were calculated by assuming that the Hexi Corridor (Fig. 1) remained the northernmost boundary of the Himalayan-Tibetan orogen since initial collision at ca. 58 Ma (Fig. 4). However, our kinematic model for the Qilian Shan and available geologic constraints suggest that a delayed onset of thrust-related deformation at 40 Ma across northern Tibet is more plausible to match the geologic record. If we assume kinematics of Models #1-3 (i.e., models involving deformation of the Tibetan crust since ca. 58 Ma; Fig. 4), this impacts our estimated bulk strain across the HTO for the first ~ 18 myr of orogeny (Fig. 8), whereas delayed deformation in the Qilian Shan is readily compatible with India-Asia collision at ca. 45 Ma in the Xigaze backarc basin model (Model #4 in Fig. 4). With the Xigaze backarc basin model, collision-related deformation across the Tibetan crust would not have started until at ca. 45 Ma (Kapp and DeCelles, 2019). This model has the advantage of explaining widespread deformation initiation across Tibet, including the Qilian Shan, at ca. 45-40 Ma. This scenario implies that the HTO deformed at a constant strain rate of $6.0 \times 10^{-16} \text{ s}^{-1}$.

However, for Models #1-3, delayed deformation in the Qilian Shan until ca. 40 Ma would impact that bulk strain-rate calculation. If the Qilian Shan was not deforming at ca. 58 Ma, then the effective northern boundary of the HTO may have been a thrust belt further to the south, such as the Fenghuoshan thrust belt (Fig. 1C) (e.g., Leeder et al., 1988; Staisch et al., 2016), which experienced contractional deformation in the Eocene to late Oligocene (ca. 51-27 Ma; Staisch et al., 2014, 2016). If the Fenghuoshan thrust belt (Fig. 1C) was the northernmost locus of deformation prior to ca. 40 Ma, there must have been a narrower width of the HTO at this time. We envision two possible scenarios to consider the effective orogen width in northern Tibet: (1) variable partitioned deformation within the Qilian Shan-Nan Shan and other thrust belts in the HTO acted in concert to yield a constant bulk strain rate over the life of the orogen (i.e., very low strain in the Qilian Shan-Nan Shan was balanced by higher strain elsewhere) (Fig. 12A), or (2) the bulk HTO orogen width was variable due to delayed thrust-related deformation in the Qilian Shan-Nan Shan until ca. 40 Ma, which implies bulk-strain rate variability throughout the Cenozoic (Fig. 12A).

In one model, the effective northernmost boundary of the HTO may have been adjacent to the Hexi Corridor (Fig. 1C), but local variations in deformation and convergence-parallel velocity allowed the Qilian Shan region to remain relatively undeformed while other regions experienced focused shortening. Such a deformation pattern with a constant bulk strain rate ($\sim 5.6 \times 10^{-16} \text{ s}^{-1}$) may have been partitioned within the Himalaya, Tethyan Himalaya, Gangdese retroarc, Qiangtang, and

Fenghuoshan thrust systems (Ratschbacher et al., 1994; Kapp et al., 2007; Yin et al., 2010; Staisch et al., 2016; Laskowski et al., 2017; Cao et al., 2020), with negligible northward convergence in the Qilian Shan-Nan Shan relative to stable Eurasia (Fig. 12A).

An alternative model suggests that the effective width of the HTO, L , was narrower from ca. 58 Ma to 40 Ma before the Qilian Shan-Nan Shan started deforming (Fig. 12B). For example, if the original pre-shortened width of the present-day ~ 350 -km-wide Qilian Shan-Nan Shan thrust belt was ~ 715 km (i.e., 365 km shortening or 50% strain), and collision-related deformation involved the entirety of the Greater India-Tibet since 58 Ma, then the restored effective width of the pre-collisional HTO was 6,050 km minus 750 km, $L(t = -58 \text{ myr}) \approx 5,300$ km. Using India-Asia convergence rates as in Model #1 of Fig. 4, this yields a marginally higher effective bulk strain rate of $7.2 \times 10^{-16} \text{ s}^{-1}$ from ca. 58 Ma to 40 Ma. Once northern Tibet started deforming at ca. 40 Ma, the orogen width should have increased by ~ 715 km, thus reducing the bulk strain rate of the entire orogen. We note that a strain rate of $5.3 \times 10^{-16} \text{ s}^{-1}$ across the HTO is permissible based on predicted shortening magnitudes and the orogen width at ca. 40 Ma. Potentially decoupled plate convergence and shortening rates (i.e., Model #3 in Fig. 4) complicates this calculation, but one could envision analogous configurations yielding similar strain rates. In summary, despite potential complications to the bulk orogen width through time, the above discussed scenarios can result in a bulk HTO strain rates that broadly overlap $5.6 \times 10^{-16} \text{ s}^{-1}$ strain rates.

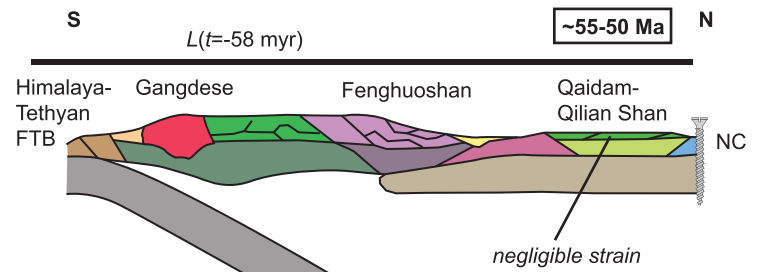
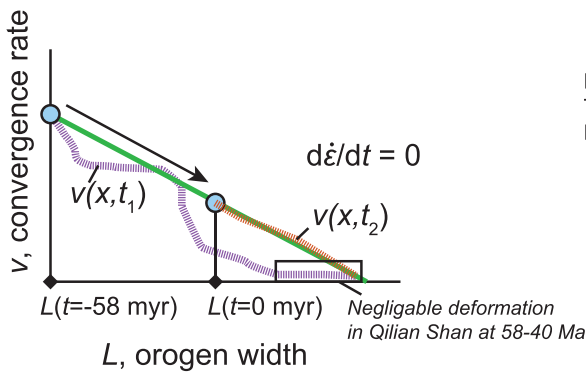
5.5. Can a constant bulk strain rate describe HTO kinematics?

Four kinematic models were discussed in this study to explore the implications of bulk HTO strain rates (Fig. 4). Models #1, #3, and #4 appear to suggest nearly constant $5.6 \times 10^{-16} \text{ s}^{-1}$ bulk strain rates across the orogen. Model #2 implies a faster strain rate ($\sim 10^{-15} \text{ s}^{-1}$) in the early Cenozoic that must have slowed at some point to match the present-day bulk strain rate ($5.6 \times 10^{-16} \text{ s}^{-1}$) and reduce predicted crustal shortening to values less than or equal to plate convergence. Although the models presented here generally assume constant strain rates, we acknowledge that strain rates could have varied through time. Here we offer several ways that bulk orogenic strain rates may have varied through time based on orogen width and convergence rate (Fig. 12C). Apparent strain rate variations across the orogen could be due to a change in boundary conditions, such as an external change in plate-convergence velocity or the width of the actively deformation region. Widening of the orogen during progressive contractional deformation could occur via foreland propagating faulting and footwall accretion along the northern or southern boundaries of the Tibetan Plateau. Southward incorporation of Greater India crust into the Himalaya fold-thrust belt might be an example of this. Alternatively, the orogen-scale bulk strain rate may simply vary during the evolution of the orogen (Fig. 12C) due to a dynamic response driven by increased gravitational potential energy and/or the removal of dense mantle lithosphere (e.g., England and Molnar, 1997; Flesch et al., 2001). This predicts a nonlinear change in convergence rate compared to orogenic shortening (Fig. 12C; cf. Fig. 5).

Himalayan studies suggest temporal variations in Cenozoic

(A) Deformation partitioning:

Bulk strain rate remains constant with variability in velocity field

**(B) Bulk strain rate variations:**

Bulk strain rate changes because northern plateau deformation is delayed

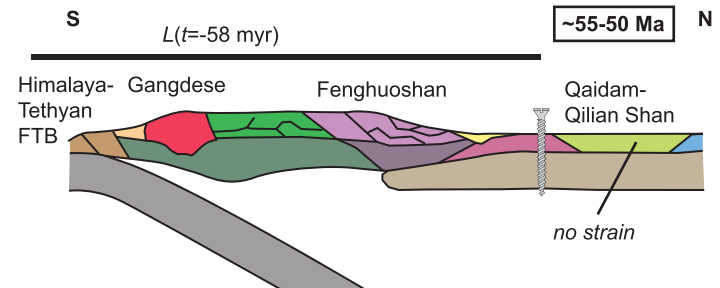
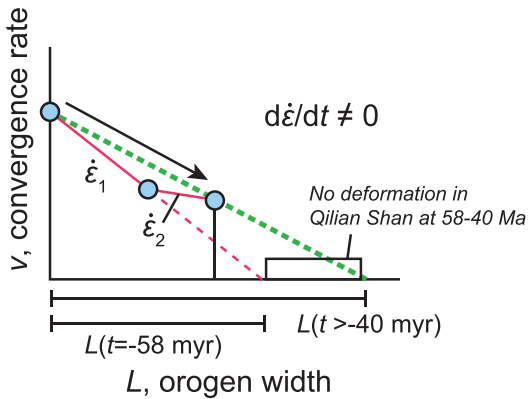
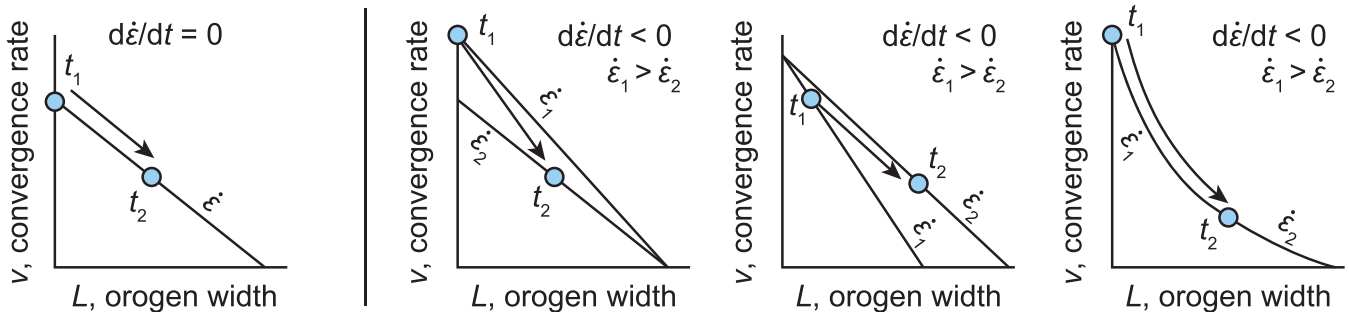
**(C) Variations in strain rate through time**

Fig. 12. Models to reconcile delayed deformation initiation in northern Tibet; schematic cross sections show where the effective northernmost HTO boundary is located during the early stages of deformation (modified from [Staich et al., 2016](#)). (A) The entire HTO may have deformed under a constant bulk strain rate (green line), with deformation variably partitioned to discrete thrust belts. The thick purple line is a hypothetical convergence-parallel velocity curve, which effectively shows negligible deformation in the Qilian Shan-Nan Shan. (B) The lack of deformation in northern Tibet may have reduced the effective orogen width, thus increasing its bulk strain rate (i.e., $\dot{\epsilon}_1$ is predicted by a narrower orogen). Once deformation in the Qilian Shan-Nan Shan initiates, the strain rate slows, $\dot{\epsilon}_2$, as a result of a wider orogen. (C) Schematic plots of how orogen-scale strain rate (v/L) could have universally varied with time due to variations in convergence rate (v) and orogen width (L). If the strain rate has remained constant, v and L will decrease at a constant ratio. However, variation in the strain rate could be caused by a punctuated change in convergence rate, orogen width, and/or intrinsic property of continental deformation, where strain rate varies with time given constant boundary conditions.

shortening rate (e.g., [Long et al., 2012](#); [McQuarrie and Ehlers, 2015](#); [McQuarrie et al., 2019](#)), which must be balanced elsewhere in the orogen. Because convergence rate exponentially decreases since the onset of initial collision (e.g., [Molnar and Stock, 2009](#); [Copley et al., 2010](#)), any

other patterns or shifts in deformation rate must be accounted for. It is presently unknown whether this balance is manifested via enhanced or diminished deformation to the north—that is, reduced shortening in the Himalaya could be balanced by higher shortening rates in central and

northern Tibet and vice versa (e.g., Long et al., 2012) (i.e., an effective strain partitioning scenario, Fig. 12A)—or a shift in the deformational response (e.g., lateral extrusion or more efficient continental subduction). For example, reduced crustal shortening rates in either the Himalaya or Qilian Shan-Nan Shan could be a product of more efficient convergence versus continental shortening decoupling or continental subduction that erases traces of upper crustal deformation (see Cowgill et al., 2016 for discussion of the Caucasus; e.g., Ingalls et al., 2016; Rowley and Ingalls, 2017; cf. van Hinsbergen et al., 2017; Haproff et al., 2019, 2020). Accordingly, the convergence rate across a thrust belt would be severely underestimated by the observed and interpreted upper-crustal shortening rate. Alternatively, assumed subsurface geometries of Himalayan deformation (e.g., duplexing or tectonic wedging) could affect our ability to fully capture geologic deformation rates. Other models that envision either large-scale continental subduction (Ingalls et al., 2016) or oceanic crust subduction as in the Greater India Basin model (van Hinsbergen et al., 2011a, 2012) may be supported by the disconnect between geodetic convergence rate and observed geologic shortening rate in the some localities of the Himalaya, such as in Bhutan (i.e., ~ 7 mm/yr geologic shortening rate since the Miocene versus ~ 20 mm/yr geodetic convergence rate; Banerjee et al., 2008; Long et al., 2012; McQuarrie and Ehlers, 2015).

To first order, bulk orogen-scale HTO strain rates may have remained constant through time due to efficient partitioning of India-Asia convergence versus intra-Tibetan shortening, and this condition permits using a constant strain rate to constrain shortening rate and magnitude through time. In this sense, efficient strain partitioning was critical to keep $\dot{\epsilon}$ approximately constant. We acknowledge that the discussed (1) 18 myr possible delay of the onset of collision-related deformation in the Qilian Shan-Nan Shan compared to India-Asia collision at ca. 58 Ma, and (2) variable shortening rate histories observed within thrust belts may imply some strain rate variability during the evolution of the HTO. However, these effects appear to be secondary controlling factors, and most of our envisioned kinematic models support a relatively constant bulk orogen-scale HTO strain rate. This study may provide indirect support of the Xigaze backarc basin hypothesis, given that the model predicts a similar bulk strain rate ($\sim 6.0 \times 10^{-16} \text{ s}^{-1}$) that may have remained constant to the present ($\sim 6.6 \times 10^{-16} \text{ s}^{-1}$) and is compatible with deformation in northern Tibet starting at ca. 40 Ma (e.g., Li et al., 2020).

Deformation across the HTO is inherently a 3D process and our models simplify the effects of lateral or oblique-slip deformation. Differences between the modeled and observed shortening magnitudes and rates may be due to vertical decoupling of deformation between the upper and lower crusts, as discussed above with continental subduction, in addition to along-strike strain variations. For example, perturbations in shortening rate across the Himalaya may be caused by focused deformation in specific regions along the thrust belt or oblique slip and strain partitioning. In the northwest Himalaya, plate convergence is effectively partitioned between strike-slip and thrust structures as observed in field studies (e.g., Styron et al., 2011; Murphy et al., 2014) and geodesy (e.g., Kundu et al., 2014; Schiffman et al., 2013). The same is true within the Tibetan Plateau, where plate convergence can be accommodated, for example, in either the Qilian Shan-Nan Shan to the north or the Longmen Shan to the east (Fig. 1C). However, the relatively late initiation of lateral tectonics at 15–10 Ma (e.g., Yuan et al., 2013; Li et al., 2019) and dominance of primarily NNE-directed contractional deformation parallel to India-Asia plate convergence permit the use of a 1D convergence/shortening model to explore time-varying shortening across the orogen.

Despite these uncertainties, we contend that this work represents a framework to tabulate and predict deformation rates across the HTO within the bounds of known plate convergence. Data from both past and future geological investigations can be used to test these models.

6. Conclusions

In this study, we explored the 1D shortening strain rate across the Himalayan-Tibetan orogen both today and at the time of initial India-Asia collision, assuming a ca 58 Ma collisional age and a ca. 45 Ma collision based on the Xigaze backarc model. The bulk strain rate appears to have been approximately constant during the Cenozoic at $5\text{--}6 \times 10^{-16} \text{ s}^{-1}$. This apparent constant bulk strain rate across the HTO suggests effective strain partitioning as the orogen evolved and different structures initiated. Assuming that this strain rate has remained constant through time and space, we present a simple model of shortening magnitude and rate for the entire orogen. Because of the linear deceleration of present-day GPS velocities across the HTO, we assume that the Himalaya and intra-Tibet thrust belts operate at similar bulk strain rates as the entire orogen and estimate shortening rate and magnitude across individual fold-thrust belts, such as the Himalayan and Qilian Shan-Nan Shan that presently define the southern and northern margins of the orogen, respectively. These belts are dominated by contractional structures, as opposed to lateral strike-slip faulting, and therefore we have confidence that our estimates of shortening in response to plate convergence are not impacted by models of lateral extrusion tectonics.

Modeled shortening magnitudes across the Himalaya match geologic observations (i.e., $\sim 550\text{--}750$ km shortening), but shortening rates estimated from field observations are more variable than our model predicts. Faster applied strain rates ($\sim 10^{-15} \text{ s}^{-1}$) better predict the present-day shortening rate across the Himalaya, but these rates are too fast to have been operating across the Himalaya since initial collision. Estimates of geologic shortening rates from Miocene to present suggest a deceleration of shortening that broadly matches our modeled curves. Both geologic observations and our preferred models predict shortening rates today of <10 mm/yr, which is lower than GPS velocities. The similarity of these geologic interpretations versus our modeled values is compelling, but there is no satisfactory explanation for the discrepancy. Modeled shortening in northern Tibet agrees with geologic observations if thrust-related deformation started at ca. 40 Ma, which is approximately 18 myr after the traditional ca. 58 Ma India-Asia collision age or only shortly after the ca. 45 Ma India-Asia collision predicted by the Xigaze backarc model. Growing datasets of deformation timing in northern Tibet are consistent with a middle Eocene initiation of thrusting and exhumation.

Application of the strain rate model to the HTO can predict the timescales of crustal thickening that can be tested against geologic, geochemical, and geophysical observations. In particular, the shift in deformation at ca. 15 Ma with the initiation of strike-slip and normal faulting across Tibet corresponds to the time when crustal thickening would have reached present-day values according to our model. This implies the progressive shortening and crustal thickening may drive the change deformational style, rather than punctuated shift in boundary conditions or geodynamic impact, such as lithospheric removal.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

Funding for this study was provided by the Office of the Vice President of Research and Innovation at the University of Nevada, Reno and the Tectonics Program of the National Science Foundation (EAR 1914501 to Zuza and Haproff). We appreciate earlier comments by Oliver Jagoutz and an anonymous reviewer on a previous version of this manuscript. Zhong-Hai Li and Andrew Laskowski are thanked for constructive reviews that clarified our ideas and improved the manuscript.

References

- Aitchison, J.C., Davis, A.M., 2004. Evidence for the multiphase nature of the India-Asia collision from the Yarlung Tsangpo suture zone, Tibet. *Geol. Soc. Lond., Spec. Publ.* 226 (1), 217–233.
- Allen, M.B., Walters, R.J., Song, S., Saville, C., De Paola, N., Ford, J., Sun, W., 2017. Partitioning of oblique convergence coupled to the fault locking behavior of fold-and-thrust belts: evidence from the Qilian Shan, northeastern Tibetan Plateau. *Tectonics* 36 (9), 1679–1698.
- Allmendinger, R.W., Reilinger, R., Loveless, J., 2007. Strain and rotation rate from GPS in Tibet, Anatolia, and the Altiplano. *Tectonics* 26 (3).
- An, K., Lin, X., Wu, L., Yang, R., Chen, H., Cheng, X., Xia, Q., et al., 2020. An immediate response to the Indian-Eurasian collision along the northeastern Tibetan Plateau: evidence from apatite fission track analysis in the Kuantan Shan-Hei Shan. *Tectonophysics* 774.
- Banerjee, P., Bürgmann, R., Nagarajan, B., Apel, E., 2008. Intraplate deformation of the Indian subcontinent. *Geophys. Res. Lett.* 35 (18).
- Baotian, P., Qingyang, L., Xiaofei, H., Haopeng, G., Zibian, L., Shaofei, J., Wanming, Y., 2013. Cretaceous and Cenozoic cooling history of the eastern Qilian Shan, northeastern margin of the Tibetan Plateau: evidence from apatite fission-track analysis. *Terra Nova* 25 (6), 431–438.
- Besse, J., Courtillot, V., Pozzi, J.P., Westphal, M., Zhou, Y.X., 1984. Palaeomagnetic estimates of crustal shortening in the Himalayan thrusts and Zangbo suture. *Nature* 311 (5987), 621.
- Bian, S., Gong, J., Zuza, A.V., Yang, R., Tian, Y., Ji, J., Cheng, X., 2020. Late Pliocene onset of the Cona rift, eastern Himalaya, confirms eastward propagation of extension in Himalayan-Tibetan orogen. *Earth Planet. Sci. Lett.* 544, 116383.
- Blisniuk, Peter M., Hacker, Bradley R., Glodny, Johannes, Ratschbacher, Lothar, Bi, Siwen, Wu, Zhenhan, McWilliams, Michael O., Calvert, Andy, 2001. Normal faulting in central Tibet since at least 13.5 Myr ago. *Nature* 412 (6847), 628.
- Burgess, W.P., Yin, A., Dubey, C.S., Shen, Z.K., Kelty, T.K., 2012. Holocene shortening across the Main Frontal Thrust zone in the eastern Himalaya. *Earth Planet. Sci. Lett.* 357, 152–167.
- Cande, S.C., Patriat, P., Dymant, J., 2010. Motion between the Indian, Antarctic and African plates in the early Cenozoic. *Geophys. J. Int.* 183 (1), 127–149.
- Cao, W., Yang, J., Zuza, A.V., Ji, W.Q., Ma, X.X., Chu, X., Burgess, Q.P., 2020. Crustal tilting and differential exhumation of Gangdese Batholith in southern Tibet revealed by bedrock pressures. *Earth Planet. Sci. Lett.* 543.
- Christensen, N.I., Mooney, W.D., 1995. Seismic velocity structure and composition of the continental crust: a global view. *J. Geophys. Res. Solid Earth* 100 (B6), 9761–9788.
- Clark, M.K., 2012. Continental collision slowing due to viscous mantle lithosphere rather than topography. *Nature* 483 (7387), 74–77.
- Clark, M.K., Farley, K.A., Zheng, D., Wang, Z., Duvall, A.R., 2010. Early Cenozoic faulting of the northern Tibetan Plateau margin from apatite (U–Th)/He ages. *Earth Planet. Sci. Lett.* 296 (1–2), 78–88.
- Copley, A., Avouac, J.P., Royer, J.Y., 2010. India-Asia collision and the Cenozoic slowdown of the Indian plate: Implications for the forces driving plate motions. *J. Geophys. Res. Solid Earth* 115 (B3).
- Copley, A., Avouac, J.P., Wernicke, B.P., 2011. Evidence for mechanical coupling and strong Indian lower crust beneath southern Tibet. *Nature* 472 (7341), 79–81.
- Coward, M.P., Butler, R.W.H., 1985. Thrust tectonics and the deep structure of the Pakistan Himalaya. *Geology* 13 (6), 417–420.
- Cowgill, E., Forte, A.M., Niemi, N., Avdeev, B., Tye, A., Trexler, C., Javakishvili, Z., Elashvili, M., Godoladze, T., 2016. Relict basin closure and crustal shortening budgets during continental collision: an example from Caucasus sediment provenance. *Tectonics* 35 (12), 2918–2947.
- Craddock, W., Kirby, E., Zhang, H., 2011. Late Miocene–Pliocene range growth in the interior of the northeastern Tibetan Plateau. *Lithosphere* 3 (6), 420–438.
- DeCelles, P.G., 2004. Late Jurassic to Eocene evolution of the Cordilleran thrust belt and foreland basin system. *western USA. Am. J. Sci.* 304 (2), 105–168.
- DeCelles, P.G., Robinson, D.M., Quade, J., Ojha, T.P., Garzione, C.N., Copeland, P., Upreti, B.N., 2001. Stratigraphy, structure, and tectonic evolution of the Himalayan fold-thrust belt in western Nepal. *Tectonics* 20 (4), 487–509.
- DeCelles, P.G., Robinson, D.M., Zandt, G., 2002. Implications of shortening in the Himalayan fold-thrust belt for uplift of the Tibetan Plateau. *Tectonics* 21 (6).
- DeCelles, P.G., Kapp, P., Gehrels, G.E., Ding, L., 2014. Paleocene-Eocene foreland basin evolution in the Himalaya of southern Tibet and Nepal: Implications for the age of initial India-Asia collision. *Tectonics* 33 (5), 824–849.
- Deng, T., Ding, L., 2015. Palealtimetry reconstructions of the Tibetan Plateau: progress and contradictions. *National Science Review* 2 (4), 417–437.
- Dewey, J.F., Bird, J.M., 1970. Mountain belts and the new global tectonics. *J. Geophys. Res.* 75 (14), 2625–2647.
- Dewey, J.F., Burke, K.C., 1973. Tibetan, Variscan, and Precambrian basement reactivation: products of continental collision. *J. Geol.* 81 (6), 683–692.
- Dey, S., Thiede, R.C., Schildgen, T.F., Wittmann, H., Bookhagen, B., Scherler, D., Strecker, M.R., 2016. Holocene internal shortening within the northwest Sub-Himalaya: Out-of-sequence faulting of the Jwalamukhi Thrust. *India, Tectonics* 35, 2677–2697.
- Dupont-Nivet, G., Horton, B.K., Butler, R.F., Wang, J., Zhou, J., Waanders, G.L., 2004. Paleogene clockwise tectonic rotation of the Xining-Lanzhou region, northeastern Tibetan Plateau. *J. Geophys. Res. Solid Earth* 109 (B4).
- Dupont-Nivet, G., Lippert, P.C., Van Hinsbergen, D.J., Meijers, M.J., Kapp, P., 2010. Palaeolatitude and age of the Indo-Asia collision: Palaeomagnetic constraints. *Geophys. J. Int.* 182 (3), 1189–1198.
- Duvall, A.R., Clark, M.K., 2010. Dissipation of fast strike-slip faulting within and beyond northeastern Tibet. *Geol.* 38 (3), 223–226.
- Duvall, A.R., Clark, M.K., van der Pluijm, B.A., Li, C., 2011. Direct dating of Eocene reverse faulting in northeastern Tibet using Ar-dating of fault clays and low-temperature thermochronometry. *Earth Planet. Sci. Lett.* 304 (3), 520–526.
- Duvall, A.R., Clark, M.K., Kirby, E., Farley, K.A., Craddock, W.H., Li, C., Yuan, D.Y., 2013. Low-temperature thermochronometry along the Kunlun and Haiyuan Faults, NE Tibetan Plateau: Evidence for kinematic change during late-stage orogenesis. *Tectonics* 32 (5), 1190–1211.
- England, P., Houseman, G., 1986. Finite strain calculations of continental deformation: 2. Comparison with the India-Asia collision zone. *J. Geophys. Res. Solid Earth* 91 (B3), 3664–3676.
- England, P., Houseman, G., 1989. Extension during continental convergence, with application to the Tibetan Plateau. *J. Geophys. Res.: Solid Earth* 94 (B12), 17561–17579.
- England, P., Molnar, P., 1997. Active deformation of Asia: from kinematics to dynamics. *Science* 278 (5338), 647–650.
- Flesch, L.M., Haines, A.J., Holt, W.E., 2001. Dynamics of the India-Eurasia collision zone. *J. Geophys. Res. Solid Earth* 106 (B8), 16435–16460.
- Gan, W., Zhang, P., Shen, Z.K., Niu, Z., Wang, M., Wan, Y., Zhou, D., Cheng, J., 2007. Present-day crustal motion within the Tibetan Plateau inferred from GPS measurements. *J. Geophys. Res. Solid Earth* 112 (B8).
- Gao, R., Wang, H., Yin, A., Dong, S., Kuang, Z., Zuza, A.V., Li, W., Xiong, X., 2013. Tectonic development of the northeastern Tibetan Plateau as constrained by high-resolution deep seismic-reflection data. *Lithosphere* 5 (6), 555–574.
- Gaudemer, Y., Tapponnier, P., Meyer, B., Peltzer, G., Shunmin, G., Zhitai, C., Cifuentes, I., 1995. Partitioning of crustal slip between linked, active faults in the eastern Qilian Shan, and evidence for a major seismic gap, the ‘Tianzhu gap’, on the western Haiyuan Fault, Gansu (China). *Geophys. J. Int.* 120 (3), 599–645.
- Gavillot, Y., 2014. Active tectonics of the Kashmir Himalaya (NW India) and earthquake potential on folds, out-of-sequence thrusts, and duplexes. In: PhD thesis.
- Gavillot, Y., Meigs, A., Yule, D., Heermance, R., Rittenour, T., Madugo, C., Malik, M., 2016. Shortening rate and Holocene surface rupture on the Riasi fault system in the Kashmir Himalaya: Active thrusting within the Northwest Himalayan orogenic wedge. *Geol. Soc. Am. Bull.* 128 (7–8), 1070–1094.
- Gavillot, Y., Meigs, A.J., Sousa, F.J., Stockli, D., Yule, D., Malik, M., 2018. Late Cenozoic foreland-to-hinterland low-temperature exhumation history of the Kashmir Himalaya. *Tectonics* 37.
- Ge, W.P., Molnar, P., Shen, Z.K., Li, Q., 2015. Present-day crustal thinning in the southern and northern Tibetan plateau revealed by GPS measurements. *Geophys. Res. Lett.* 42 (13), 5227–5235.
- George, A.D., Marshall, S.J., Wyrwoll, K.H., Jie, C., Yanchou, L., 2001. Miocene cooling in the northern Qilian Shan, northeastern margin of the Tibetan Plateau, revealed by apatite fission-track and vitrinite-reflectance analysis. *Geology* 29 (10), 939–942.
- Haproff, P.J., Zuza, A.V., Yin, A., 2018. West-directed thrusting south of the eastern Himalayan syntaxis indicates clockwise crustal flow at the indenter corner during the India-Asia collision. *Tectonophysics* 722, 277–285.
- Haproff, P.J., Zuza, A.V., Yin, A., Harrison, T.M., Manning, C.E., Dubey, C.S., Chen, J., 2019. Geologic framework of the northern Indo-Burma Ranges and lateral correlation of Himalayan-Tibetan lithologic units across the eastern Himalayan syntaxis. *Geosphere* 15 (3), 856–881.
- Haproff, P.J., Odum, M.L., Zuza, A.V., Yin, A., Stockli, D.F., 2020. Structural and Thermochronologic Constraints on the Cenozoic Tectonic Development of the Northern Indo-Burma Ranges. *Tectonics* 39 (9) e2020TC006231.
- Harrison, T.M., Wielicki, M.M., 2016. From the Hadean to the Himalaya: 4.4 Ga of felsic terrestrial magmatism. *Am. Mineral.* 101 (6), 1348–1359.
- Hopkinson, T., Harris, N., Roberts, N.M., Warren, C.J., Hammond, S., Spencer, C.J., Parrish, R.R., 2020. Evolution of the melt source during protracted crustal anatexis: An example from the Bhutan Himalaya. *Geol.* 48 (1), 87–91.
- Hu, X., Garzanti, E., Moore, T., Raffi, I., 2015. Direct stratigraphic dating of India-Asia collision onset at the Selandian (middle Paleocene, 59±1 Ma). *Geology* 43 (10), 859–862.
- Hu, X., Wang, J., BouDagher-Fadel, M., Garzanti, E., An, W., 2016. New insights into the timing of the India-Asia collision from the Paleogene Quixia and Jialazi formations of the Xigaze forearc basin, South Tibet. *Gondwana Res.* 32, 76–92.
- Huang, W., Lippert, P.C., Jackson, M.J., Dekkers, M.J., Zhang, Y., Li, J., Guo, Z., Kapp, P., van Hinsbergen, D.J., 2017. Remagnetization of the Paleogene Tibetan Himalayan carbonate rocks in the Gamba area: implications for reconstructing the lower plate in the India-Asia collision. *J. Geophys. Res. Solid Earth* 122 (2), 808–825.
- Ingalls, M., Rowley, D.B., Currie, B., Colman, A.S., 2016. Large-scale subduction of continental crust implied by India-Asia mass-balance calculation. *Nat. Geosci.* 9 (11), 848–853.
- Jade, S., Bhatt, B.C., Yang, Z., Bendick, R., Gaur, V.K., Molnar, P., Anand, M.B., Kumar, D., 2004. GPS measurements from the Ladakh Himalaya, India: Preliminary tests of plate-like or continuous deformation in Tibet. *Geol. Soc. Am. Bull.* 116 (11–12), 1385–1391.
- Jagoutz, O., Royden, L., Holt, A.F., Becker, T.W., 2015. Anomalously fast convergence of India and Eurasia caused by double subduction. *Nat. Geosci.* 8 (6), 475–478.
- Jolivet, M., Brunel, M., Seward, D., Xu, Z., Yang, J., Roger, F., Tapponnier, P., Malavieille, J., Arnaud, N., Wu, C., 2001. Mesozoic and Cenozoic tectonics of the northern edge of the Tibetan plateau: fission-track constraints. *Tectonophysics* 343 (1), 111–134.
- Jones, C.H., Unruh, J.R., Sonder, L.J., 1996. The role of gravitational potential energy in active deformation in the southwestern United States. *Nature* 381 (6577), 37–41.
- Kapp, P., DeCelles, P.G., 2019. Mesozoic–Cenozoic geological evolution of the Himalayan-Tibetan orogen and working tectonic hypotheses. *Am. J. Sci.* 319 (3), 159–254.

- Kapp, P., DeCelles, P.G., Gehrels, G.E., Heizler, M., Ding, L., 2007. Geological records of the Lhasa-Qiangtang and Indo-Asian collisions in the Nima area of central Tibet. *Geol. Soc. Am. Bull.* 119 (7–8), 917–933.
- Khanal, S., Robinson, D.M., 2013. Upper crustal shortening and forward modeling of the Himalayan thrust belt along the Budhi-Gandaki River, central Nepal. *Int. J. Earth Sci.* 102 (7), 1871–1891.
- Kirby, E., Harkins, N., Wang, E., Shi, X., Fan, C., Burbank, D., 2007. Slip rate gradients along the eastern Kunlun fault. *Tectonics* 26 (2).
- Kundu, B., Yadav, R.K., Bali, B.S., Chowdhury, S., Gahalaut, V.K., 2014. Oblique convergence and slip partitioning in the NW Himalaya: implications from GPS measurements. *Tectonics* 33 (10), 2013–2024.
- Laskowski, A.K., Kapp, P., Ding, L., Campbell, C., Liu, X., 2017. Tectonic evolution of the Yarlung suture zone, Lopa Range region, southern Tibet. *Tectonics* 36 (1), 108–136.
- Lavé, J., Avouac, J.P., 2000. Active folding of fluvial terraces across the Siwaliks Hills, Himalayas of central Nepal. *J. Geophys. Res. Solid Earth* 105 (B3), 5735–5770.
- Lease, R.O., Burbank, D.W., Zhang, H., Liu, J., Yuan, D., 2012. Cenozoic shortening budget for the northeastern edge of the Tibetan Plateau: Is lower crustal flow necessary? *Tectonics* 31 (3).
- Lee, C.T.A., Thurner, S., Paterson, S., Cao, W., 2015. The rise and fall of continental arcs: interplays between magmatism, uplift, weathering, and climate. *Earth Planet. Sci. Lett.* 425, 105–119.
- Leech, M.L., Singh, S., Jain, A.K., Klemperer, S.L., Manickavasagam, R.M., 2005. The onset of India-Asia continental collision: early, steep subduction required by the timing of UHP metamorphism in the western Himalaya. *Earth Planet. Sci. Lett.* 234 (1–2), 83–97.
- Leeder, M.R., Smith, A.B., Yin, J., 1988. Sedimentology, paleoecology and paleoenvironmental evolution of the 1985 Lhasa to Golmund Geotraverse. *R. Soc. London Philos. Trans.* 327, 107–143.
- Li, B., Chen, X., Zuza, A.V., Hu, D., Ding, W., Huang, P., Xu, S., 2019. Cenozoic cooling history of the North Qilian Shan, northern Tibetan Plateau, and the initiation of the Haiyuan fault: Constraints from apatite- and zircon-fission track thermochronology. *Tectonophysics* 751, 109–124.
- Li, B., Zuza, A.V., Chen, X., Hu, D., Shao, Z., Qi, B., Xiong, X., 2020. Cenozoic multi-phase deformation in the Qilian Shan and out-of-sequence development of the northern Tibetan Plateau. *Tectonophysics* 228423.
- Liu, M., Yang, Y., 2003. Extensional collapse of the Tibetan Plateau: results of three-dimensional finite element modeling. *J. Geophys. Res. Solid Earth* 108 (B8).
- Long, S., McQuarrie, N., Tobgay, T., Grujic, D., 2011. Geometry and crustal shortening of the Himalayan fold-thrust belt, eastern and central Bhutan. *Geol. Soc. Am. Bull.* 123 (7–8), 1427–1447.
- Long, S.P., McQuarrie, N., Tobgay, T., Coutand, I., Cooper, F.J., Reiners, P.W., Hodges, K.V., 2012. Variable shortening rates in the eastern Himalayan thrust belt, Bhutan: Insights from multiple thermochronologic and geochronologic data sets tied to kinematic reconstructions. *Tectonics* 31 (5).
- McKenzie, D.P., 1969. Speculations on the consequences and causes of plate motions. *Geophys. J. Int.* 18 (1), 1–32.
- McQuarrie, N., Ehlers, T.A., 2015. Influence of thrust belt geometry and shortening rate on thermochronometer cooling ages: insights from thermokinematic and erosion modeling of the Bhutan Himalaya. *Tectonics* 34 (6), 1055–1079.
- McQuarrie, N., Tobgay, T., Long, S.P., Reiners, P.W., Cosca, M.A., 2014. Variable exhumation rates and variable displacement rates: Documenting recent slowing of Himalayan shortening in western Bhutan. *Earth Planet. Sci. Lett.* 386, 161–174.
- McQuarrie, N., Eizenhöfer, P.R., Long, S.P., Tobgay, T., Ehlers, T.A., Blythe, A.E., Dering, G.M., 2019. The influence of foreland structures on hinterland cooling: evaluating the drivers of exhumation in the eastern Bhutan Himalaya. *Tectonics* 38 (9), 3282–3310.
- Meigs, A.J., Burbank, D.W., 1997. Growth of the south Pyrenean orogenic wedge. *Tectonics* 16, 239–258.
- Meng, J., Coe, R.S., Wang, C., Gilder, S.A., Zhao, X., Liu, H., Li, Y., Ma, P., Shi, K., Li, S., 2017. Reduced convergence within the Tibetan plateau by 26 Ma? *Geophys. Res. Lett.* 44 (1316), 6624–6632.
- Métivier, F., Gaudemer, Y., Tapponnier, P., Meyer, B., 1998. Northeastward growth of the Tibet plateau deduced from balanced reconstruction of two depositional areas: The Qaidam and Hexi Corridor basins, China. *Tectonics* 17 (6), 823–842.
- Meyer, B., Tapponnier, P., Bourjot, L., Metivier, F., Gaudemer, Y., Peltzer, G., Shunmin, G., Zhitai, C., 1998. Crustal thickening in Gansu-Qinghai, lithospheric mantle subduction, and oblique, strike-slip controlled growth of the Tibet plateau. *Geophys. J. Int.* 135 (1), 1–47.
- Molnar, P., Stock, J.M., 2009. Slowing of India's convergence with Eurasia since 20 Ma and its implications for Tibetan mantle dynamics. *Tectonics* 28 (3).
- Molnar, P., Tapponnier, P., 1975. Cenozoic tectonics of Asia: effects of a continental collision. *Science* 189 (4201), 419–426.
- Molnar, P., England, P., Martinod, J., 1993. Mantle dynamics, uplift of the Tibetan Plateau, and the Indian monsoon. *Rev. Geophys.* 31 (4), 357–396.
- Mugnier, J.L., Leturmy, P., Mascle, G., Huyghe, P., Chalaron, E., Vidal, G., Delcaillau, B., 1999. The Siwaliks of western Nepal: I. Geometry and kinematics. *J. Asian Earth Sci.* 17 (5–6), 629–642.
- Mugnier, J.L., Huyghe, P., Leturmy, P., Jouanne, F., 2004. Episodicity and rates of thrust-sheet motion in the Himalayas (western Nepal).
- Murphy, M.A., Yin, A., 2003. Structural evolution and sequence of thrusting in the Tethyan fold-thrust belt and Indus-Yalu suture zone, southwest Tibet. *Geol. Soc. Am. Bull.* 115 (1), 21–34.
- Murphy, M.A., Yin, A., Harrison, T.M., Durr, S.B., Chen, Z., Ryerson, F.J., Zhou, X., 1997. Significant crustal shortening in south-central Tibet prior to the Indo-Asian collision. *Geology* 25, 719–722.
- Murphy, M.A., Taylor, M.H., Gosse, J., Silver, C.R.P., Whipp, D.M., Beaumont, C., 2014. Limit of strain partitioning in the Himalaya marked by large earthquakes in western Nepal. *Nat. Geosci.* 7 (1), 38–42.
- Najman, Y., Appel, E., Boudagher-Fadel, M., Bown, P., Carter, A., Garzanti, E., Parrish, R., 2010. Timing of India-Asia collision: Geological, biostratigraphic, and palaeomagnetic constraints. *J. Geophys. Res. Solid Earth* 115 (B12).
- Patriat, P., Achache, J., 1984. India-Eurasia collision chronology has implications for crustal shortening and driving mechanism of plates. *Nature* 311 (5987), 615–621.
- Peltzer, G., Tapponnier, P., 1988. Formation and evolution of strike-slip faults, rifts, and basins during the India-Asia collision: An experimental approach. *J. Geophys. Res. Solid Earth* 93 (B12), 15085–15117.
- Powers, P.M., Lillie, R.J., Yeats, R.S., 1998. Structure and shortening of the Kangra and Dehra Dun reentrants, sub-Himalaya, India. *Geol. Soc. Am. Bull.* 110 (8), 1010–1027.
- Qi, B., Hu, D., Yang, X., Zhang, Y., Tan, C., Zhang, P., Feng, C., 2016. Apatite fission track evidence for the Cretaceous–Cenozoic cooling history of the Qilian Shan (NW China) and for stepwise northeastward growth of the northeastern Tibetan Plateau since early Eocene. *J. Asian Earth Sci.* 124, 28–41.
- Ratschbacher, L., Frisch, W., Liu, G., Chen, C., 1994. Distributed deformation in southern and western Tibet during and after the India-Asia collision. *J. Geophys. Res. Solid Earth* 99 (B10), 19917–19945.
- Rowley, D.B., 2019. Comparing paleomagnetic study means with apparent wander paths: a case study and paleomagnetic test of the Greater India versus Greater Indian Basin hypotheses. *Tectonics* 38 (2), 722–740.
- Rowley, D., Ingalls, M., 2017. Reply to ‘Unfeasible subduction?’. *Nat. Geosci.* 10 (12), 879.
- Roy, S., 1976. A possible Himalayan microcontinent. *Nature* 263, 117–120.
- Royden, L.H., Burchfiel, B.C., van der Hilst, R.D., 2008. The geological evolution of the Tibetan Plateau. *Science* 321 (5892), 1054–1058.
- Ryan, W.B.F., Carbotte, S.M., Coplan, J.O., O'Hara, S., Melkonian, A., Arko, R., Weissel, R.A., Ferrini, V., Goodwillie, A., Nitsche, F., Bonczkowski, J., Zensky, R., 2009. Global Multi-Resolution Topography synthesis. *Geochem. Geophys. Geosyst.* 10.
- Schiffman, C., Bali, B.S., Szeliga, W., Bilham, R., 2013. Seismic slip deficit in the Kashmir Himalaya from GPS observations. *Geophys. Res. Lett.* 40 (21), 5642–5645.
- Smit, M.A., Hacker, B.R., Lee, J., 2014. Tibetan garnet records early Eocene initiation of thickening in the Himalaya. *Geology* 42 (7), 591–594.
- Staisch, L.M., Niemi, N.A., Hong, C., Clark, M.K., Rowley, D.B., Currie, B., 2014. A Cretaceous-Eocene depositional age for the Fenghuoshan Group, Hoh Xil Basin: Implications for the tectonic evolution of the northern Tibet Plateau. *Tectonics* 33 (3), 281–301.
- Staisch, L.M., Niemi, N.A., Clark, M.K., Chang, H., 2016. Eocene to late Oligocene history of crustal shortening within the Hoh Xil Basin and implications for the uplift history of the northern Tibetan Plateau. *Tectonics* 35 (4), 862–895.
- Styron, R.H., Taylor, M.H., Murphy, M.A., 2011. Oblique convergence, arc-parallel extension, and the role of strike-slip faulting in the High Himalaya. *Geosphere* 7 (2), 582–596.
- Styron, R.H., Taylor, M.H., Sundell, K.E., Stockli, D.F., Oalman, J.A., Möller, A., Ding, L., 2013. Miocene initiation and acceleration of extension in the South Lunggar rift, western Tibet: Evolution of an active detachment system from structural mapping and (U-Th)/He thermochronology. *Tectonics* 32 (4), 880–907.
- Styron, R., Taylor, M., Sundell, K., 2015. Accelerated extension of Tibet linked to the northward underthrusting of Indian crust. *Nat. Geosci.* 8 (2), 131–134.
- Sundell, K., Taylor, M., Styron, R., Stockli, D., Kapp, P., Hager, C., Liu, D., Ding, L., 2013. Evidence for constriction and Pliocene acceleration of east-west extension in the North Lunggar rift region of west central Tibet. *Tectonics* 32 (5), 1454–1479.
- Tang, M., Ji, W.Q., Chu, X., Wu, A., Chen, C., 2020. Reconstructing crustal thickness evolution from europium anomalies in detrital zircons. *Geology*.
- Tapponnier, P., Peltzer, G.L.D.A.Y., Le Dain, A.Y., Armijo, R., Cobbold, P., 1982. Propagating extrusion tectonics in Asia: New insights from simple experiments with plasticine. *Geol.* 10 (12), 611–616.
- Tapponnier, P., Zhiqin, X., Roger, F., Meyer, B., Arnaud, N., Wittlinger, G., Jingsui, Y., 2001. Oblique stepwise rise and growth of the Tibet Plateau. *Science* 294 (5547), 1671–1677.
- Taylor, M., Yin, A., 2009. Active structures of the Himalayan-Tibetan orogen and their relationships to earthquake distribution, contemporary strain field, and Cenozoic volcanism. *Geosphere* 5 (3), 199–214.
- van Hinsbergen, D.J., Kapp, P., Dupont-Nivet, G., Lippert, P.C., DeCelles, P.G., Torsvik, T.H., 2011a. Restoration of Cenozoic deformation in Asia and the size of Greater India. *Tectonics* 30 (5).
- van Hinsbergen, D.J., Steinberger, B., Doubrovine, P.V., Gassmüller, R., 2011b. Acceleration and deceleration of India-Asia convergence since the Cretaceous: Roles of mantle plumes and continental collision. *J. Geophys. Res. Solid Earth* 116 (B6).
- van Hinsbergen, D.J., Lippert, P.C., Dupont-Nivet, G., McQuarrie, N., Doubrovine, P.V., Spakman, W., Torsvik, T.H., 2012. Greater India Basin hypothesis and a two-stage Cenozoic collision between India and Asia. *Proc. Natl. Acad. Sci.* 109 (20), 7659–7664.
- van Hinsbergen, D.J., Lippert, P.C., Huang, W., 2017. Unfeasible subduction? *Nat. Geosci.* 10 (12), 878.
- van Hinsbergen, D.J., Lippert, P.C., Li, S., Huang, W., Advokaat, E.L., Spakman, W., 5 June 2019. Reconstructing greater India: Paleogeographic, kinematic, and geodynamic perspectives. *Tectonophysics* 760, 69–94.
- Wang, Q.M., Coward, M.P., 1993. The Jiuxi basin, Hexi corridor, NW China: Foreland structural features and hydrocarbon potential. *J. Pet. Geol.* 16 (2), 169–182.

- Wang, J., Hu, X., Jansa, L., Huang, Z., 2011a. Provenance of the Upper Cretaceous–Eocene deep-water sandstones in Sangdanlin, southern Tibet: Constraints on the timing of initial India–Asia collision. *J. Geol.* 119 (3), 293–309.
- Wang, C., Gao, R., Yin, A., Wang, H., Zhang, Y., Guo, T., et al., 2011b. A mid-crustal strain-transfer model for continental deformation: A new perspective from high-resolution deep seismic-reflection profiling across NE Tibet. *Earth Planet. Sci. Lett.* 306 (3–4), 279–288.
- Webb, A.A.G., Yin, A., Harrison, T.M., C  lerier, J., Gehrels, G.E., Manning, C.E., Grove, M., 2011. Cenozoic tectonic history of the Himalaya (northwestern India) and its constraints on the formation mechanism of the Himalayan orogen. *Geosphere* 7 (4), 1013–1061.
- Webb, A.A.G., Guo, H., Clift, P.D., Husson, L., M  ller, T., Costantino, D., Yin, A., Xu, Z., Wang, Q., 2017. The Himalaya in 3D: Slab dynamics controlled mountain building and monsoon intensification. *Lithosphere* 9 (4), 637–651.
- Wu, C., Zuza, A.V., Zhou, Z., Yin, A., McRivette, M.W., Chen, X., Geng, J., 2019. Mesozoic–Cenozoic evolution of the Eastern Kunlun Range, central Tibet, and implications for basin evolution during the Indo–Asian collision. *Lithosphere* 11 (4), 524–550.
- Wu, C., Liu, C., Fan, S., Zuza, A.V., Ding, L., Liu, W., Zhou, Z., 2020a. Structural analysis and tectonic evolution of the western domain of the Eastern Kunlun Range, northwest Tibet. *GSA Bull.* 132 (5–6), 1291–1315.
- Wu, C., Li, J., Zuza, A.V., Liu, C., Liu, W., Chen, X., Li, B., 2020b. Cenozoic cooling history and fluvial terrace development of the western domain of the Eastern Kunlun Range, northern Tibet. *Palaeogeography, Palaeoclimatology, Palaeoecology* 560.
- Yakovlev, P.V., Clark, M.K., 2014. Conservation and redistribution of crust during the Indo–Asian collision. *Tectonics* 33 (6), 1016–1027.
- Ye, Z., Gao, R., Li, Q., Zhang, H., Shen, X., Liu, X., Gong, C., 2015. Seismic evidence for the North China plate underthrusting beneath northeastern Tibet and its implications for plateau growth. *Earth Planet. Sci. Lett.* 426, 109–117.
- Yin, A., Taylor, M.H., 2011. Mechanics of V-shaped conjugate strike-slip faults and the corresponding continuum mode of continental deformation. *GSA Bull.* 123 (9–10), 1798–1821.
- Yin, A., Nie, S., Craig, P., Harrison, T.M., Ryerson, F.J., Xianglin, Q., Geng, Y., 1998. Late Cenozoic tectonic evolution of the southern Chinese Tian Shan. *Tectonics* 17 (1), 1–27.
- Yin, A., Kapp, P.A., Murphy, M.A., Manning, C.E., Mark Harrison, T., Grove, M., et al., 1999. Significant late Neogene east–west extension in northern Tibet. *Geology* 27 (9), 787–790.
- Yin, A., Dang, Y.Q., Wang, L.C., Jiang, W.M., Zhou, S.P., Chen, X.H., Gehrels, G.E., McRivette, M.W., 2008a. Cenozoic tectonic evolution of Qaidam basin and its surrounding regions (Part 1): The southern Qilian Shan–Nan Shan thrust belt and northern Qaidam basin. *Geol. Soc. Am. Bull.* 120 (7–8), 813–846.
- Yin, A., Dang, Y.Q., Zhang, M., Chen, X.H., McRivette, M.W., 2008b. Cenozoic tectonic evolution of the Qaidam basin and its surrounding regions (Part 3): Structural geology, sedimentation, and regional tectonic reconstruction. *Geol. Soc. Am. Bull.* 120 (7–8), 847–876.
- Yin, A., Dubey, C.S., Webb, A.A.G., Kelty, T.K., Grove, M., Gehrels, G.E., Burgess, W.P., 2010. Geologic correlation of the Himalayan orogen and Indian craton: Part 1. Structural geology, U–Pb zircon geochronology, and tectonic evolution of the Shillong Plateau and its neighboring regions in NE India. *Geol. Soc. Am. Bull.* 122 (3–4), 336–359.
- Yuan, D.Y., Ge, W.P., Chen, Z.W., Li, C.Y., Wang, Z.C., Zhang, H.P., Dayem, K.E., 2013. The growth of northeastern Tibet and its relevance to large-scale continental geodynamics: a review of recent studies. *Tectonics* 32 (5), 1358–1370.
- Yue, H., Chen, Y.J., Sandvol, E., Ni, J., Hearn, T., Zhou, S., Jin, G., 2012. Lithospheric and upper mantle structure of the northeastern Tibetan Plateau. *J. Geophys. Res. Solid Earth* 117 (B5).
- Zhang, P., Burchfiel, B.C., Molnar, P., Zhang, W., Jiao, D., Deng, Q., Song, F., 1991. Amount and style of late Cenozoic deformation in the Liupan Shan area, Ningxia Autonomous Region, China. *Tectonics* 10 (6), 1111–1129.
- Zhang, P.Z., Shen, Z., Wang, M., Gan, W., B  rgmann, R., Molnar, P., Hanrong, S., 2004. Continuous deformation of the Tibetan Plateau from global positioning system data. *Geology* 32 (9), 809–812.
- Zheng, D., Clark, M.K., Zhang, P., Zheng, W., Farley, K.A., 2010. Erosion, fault initiation and topographic growth of the North Qilian Shan (northern Tibetan Plateau). *Geosphere* 6 (6), 937–941.
- Zuza, A.V., Yin, A., 2016. Continental deformation accommodated by non-rigid passive bookshelf faulting: an example from the Cenozoic tectonic development of northern Tibet. *Tectonophysics* 677, 227–240.
- Zuza, A.V., Cheng, X., Yin, A., 2016. Testing models of Tibetan Plateau formation with Cenozoic shortening estimates across the Qilian Shan–Nan Shan thrust belt. *Geosphere* 12 (2), 501–532.
- Zuza, A.V., Wu, C., Reith, R.C., Yin, A., Li, J., Zhang, J., Liu, W., 2018. Tectonic evolution of the Qilian Shan: An early Paleozoic orogen reactivated in the Cenozoic. *GSA Bulletin* 130 (5–6), 881–925.
- Zuza, A.V., Wu, C., Wang, Z., Levy, D.A., Li, B., Xiong, X., Chen, X., 2019. Underthrusting and duplexing beneath the northern Tibetan Plateau and the evolution of the Himalayan–Tibetan orogen. *Lithosphere* 11 (2), 209–231.