Dear Editors,

Below you will find our newly revised manuscript that has taken into account the reviewers suggestions. We feel the comments made by the reviewers were helpful and allowed us to construct a better manuscript. To make sure our contribution is impactful as possible, we did two new analyses suggested by the reviewers. Reviewer 1 suggested we provide a clearer connection to predicting delta existence in the stratigraphic record. To do this, we reran our logistic regression with only water and sediment discharge (since these are generally interpretable from the stratigraphic record) and show then how that equation correctly predict delta existence for the Ferron Formation. Reviewer 2 and 3 both brought up the issue of using modern environmental variable data to predict delta formation that may have occurred 1000s of years ago. Based on our review of scaling analyses for delta adaptation time, we assume that deltas are adapted to modern conditions. We now articulate that assumption in the paper more clearly. But, we also think the reviewers idea is intriguing so we reran our logistic regression with sea level change data averaged over the last 26,000 years from the ICE_6G_C model output. This did not affect our results that sea level change is seemingly unimportant for delta formation. We discuss this now in the manuscript.

Sincerely,

Doug Edmonds (on behalf of the authors)

Reviewer comments are provided in bold-faced type. All line and page numbers refer to manuscript version with changes accepted.

Author responses are italicized

REVIEWER 1: Janok Bhattacharya

A delta is defined where there is either a shoreline bulge or distinctive distributary channels, such that estuaries would be largely excluded from the analysis. Single-channel deltas are thus largely excluded as well as lacustrine deltas.

The reviewer is generally correct here in his characterization of our definition of a delta. But our definition is not so restrictive because many of the deltas in our dataset are single-channeled or in estuaries. Though for these cases single-channeled deltas must have a protrusion and estuarine systems must show evidence of sediment distribution. To make our contribution clearer we have added “along marine coastlines” to the title to make it explicit that freshwater lakes are excluded.

The paper shows that deltas are more likely along coastlines with high discharge and low wave power and to a lesser extent low tidal range. It would be nice to compare storm-dominated versus fair-weather waves.

We are not entirely sure what the reviewer is suggesting. But, we suspect he is referring to comparing the storm-dominated against fair-weather waves for a given location to create a nondimensional ratio. But in that case, we are not sure what insight such a ratio would provide given that our simple metric of
the highest 1/3 of the significant wave heights (this is probably more representative of storm wave height) provides a suitable reference for understanding delta formation.

The paper mentioned ancient systems. But made no attempt to test the logistic regression against an ancient example. We have published S2S analysis of a number of Cretaceous systems. I think it is probably fairly robust to estimate slope and sediment and water discharge, wave and tidal power may also be tractable. Conversely, the presence of a delta could be used to predict wave and tidal regime of Qw, Qs, and S are known. It might be good to expand this part of the discussion to include some of these ancient systems.

This is exactly what we have in mind and why we mention ancient systems in our paper. To make this point more concrete, we present a new version of equation 1 that does not include waves or tides and then we show that the for the cases of the Ferron Fm. presented in Bhattacharya et al., (2016) delta likelihood is high. We have now this section starting on P15L21: “Ancient deltaic deposits comprise significant hydrocarbon reservoirs, and provided our analysis holds through geologic time we could predict the presence of deltaic deposits in the rock record if Qw or Qs can be estimated via other geologic methods. If we use a logistic regression that does not include the limiting effects of waves and tide then equation (1) becomes:

\[
\ln(\pi_{\text{delta}}/(1 - \pi_{\text{delta}})) = 0.0016 + 0.0175Q_w + 0.0345Q_s
\]  

(2)

Using this simplified equation, we can predict the likelihood of delta formation for paleoenvironments where sediment and water discharge are more easily constrained. For example, water and sediment flux estimates for rivers of the Ferron Sandstone in the Cretaceous Western Interior Seaway of the United States (Bhattacharya et al., 2016), suggests the likelihood of delta formation is 99%. Even though this is somewhat obvious since we know the Ferron contains deltaic deposits, it highlights how our results could be used to predict the presence of deltaic deposits in the absence of outcrop or seismic evidence.”

I was a bit confused as to the meaning of “the main river mouth”, especially as it applies to systems that may have hundreds of terminal distributary channels like the Lena Delta?

The definition of main river mouth is already provided in the text on P512 and we state that “for rivers with deltas, this is the location of the widest river mouth in the distributary network.”

When calculating slope, the authors state that the search far away from the shoreline, but I wasn’t clear how far away? Is it some scaled distance from the river mouth (e.g., 10-100 x river mouth width)?

In this case, we calculate slopes between the river mouth and ALL bathymetric points within a 20-km radius and take the slope as the 75th percentile of all these slopes. This is stated on P8L15-20.

REVIWER 2: Jinyu Zhang

General comments:

My main concern is about the role of sea-level change. We know it is important for the formation of deltas (Stanley and Warne, 1994, which was cited in this work). I think (1) the range of sea-level change rate of past 26 years might be too limited to influence the delta likelihood; (2) the mapped
deltas initiated way before 26 years – most of them formed around 7 ky. So it is worth being more careful when discussing the sea-level change, even though I notice the authors emphasized ‘RECENT’ sea-level change in abstract and conclusion.

We certainly agree with this point, and as the reviewer points out, we made sure to refer to ‘recent’ sea level change so that this point is always in the reader’s mind. The sea level change estimates are far from what we would prefer but because it is such a crucial control on delta formation we wanted to use the best data we have.

Section 2.1 provides good criteria to recognize delta. However, I’m not sure how authors map them. Are they mapped manually or processed by some programs? If it is processed by programs, how reliable is it?

We map the deltas manually because defining the presence of a depositional body is not straightforward. This leads to some subjectivity, but we are willing to accept some ambiguity to complete the global task.

It might be beyond the scope of this study but some sensitivity analysis should be useful when discussing the relative importance of each parameter. The effect in equation 1 (0.000524, 4.77, -0.952,-0.175) shows the Qw is the least important parameter.

The reviewer is correct that a sensitivity analysis is important. We have already provided something akin to that in section 4.1 Either way, it is important to note that the variables in equation 1 are dimensional so that the size of the coefficients cannot be compared without taking into account the variables of each unit. So while the coefficient on Qw is the smallest, Qw typically is the largest numerical value.

I believe before Galloway ternary classification, people used constructive system and destructive systems, at least for ancient deltas (See William Fisher, 1969, GCAGS). When talking about constructive and destructive forces, it might be good to acknowledge them.

We thank the reviewer for this suggestion and we added this reference, along with others by reviewer 3, into a more complete discussion of this point that duly notes this “constructive vs destructive” idea has been around for a while. Please see P14L10-25 for the new discussion.

Detailed comments:

Page 8 Line 12 Yes, the range here is fairly limited. The rate of Cenozoic eustatic sea-level change went over three order of magnitude. It might be worth comparing the range of eustatic sea level change rate of past 7000 years (when most of the deltas formed) and the past 26 years data here.

Indeed this is a good idea that was also mentioned by another reviewer. We conducted an additional analysis using sea level change rates from 22 kyr ago to present from Ice6G_C model results. We now mention this brief effort in the methods on P9L4-7 and comment on how it does not increase the success of predicting delta formation.

Page 8 Line 20 – Page 9 Line 3 Using numbers instead of ‘high’ and ‘low’ when describing delta density

This is a fair point, but we opted for the qualitative descriptions here because it is easier to read and digest quickly. The quantitative data are in the figure.
Page 12 Line 21 A typo ‘.4.77’ before Qs

Thank you for catching this.

Table 1 There is a mistake for the unit of sea-level change rate, should be L/T (mm/s?), instead of ‘m’
Also please check the unit for Qs and sediment concentration. Sediment discharge refers to sediment transport mass per unit time. Unit is commonly Mt/yr?? Unit for sediment concentration is commonly mg/l?? I guess these units will influence the log odds relationship?

Thank you for catching these issues. Sediment discharge can be reported a number of ways. The dimensions will affect the logistic regression only in the size of the coefficient, but not in the effect of Qs on delta formation.

REVIEWER 3

Specific comments:

Major:

My first and largest concern is with the presentation of the constructive vs. destructive argument. The current manuscript implies that this is a new idea and a new way of thinking about delta formation, when in fact this idea has been widely discussed in the deltaic literature (both sedimentology- and geomorphology-related) since at least the 1960s. Yes, it is slightly in contrast to the Galloway diagram, but it is not in contrast to an abundance of delta literature. It needs to be made clear through (1) an enhanced discussion section (and possibly some more background as well) and (2) increased referencing throughout the paper that this idea is not new. What IS new and exciting is that you can start to quantify this with the data presented in the current manuscript, showing that you have indeed significantly advanced the science while still giving sufficient credit to the vast quantity of existing literature that presents or at least references this framework. I have listed some references for you below, compiled from only a very quick search.

We are grateful to the reviewer for taking the time to collect these references and present them to us. We never intended to miss this literature, but it unfortunately happened. We have now revised the discussion to reflect our contribution more clearly. Please see P14L10-25.

Second, you chose not to include basin depth as an environmental variable in your analysis, and you state that is because the basin depth at the time of delta formation cannot be known from modern bathymetry. You have a similar statement regarding the last 26 years of sea level data. Unfortunately, this is also true for all of the variables you include. Many of these deltas are thousands of years old, at least. How can you take modern river discharge and relate it to delta formation without knowing if the modern river discharge is responsible for modern delta existence? I have two suggestions to remedy this because I still believe your analysis to be a useful one, despite the obvious problem of time. (1) Be abundantly clear in your language throughout the paper that you can map delta existence (not formation) in the modern state and you have some modern environmental variables, but provide the caveat that deltas were formed by past environmental conditions that are largely impossible to know. This is an easy partial fix, although somewhat unsatisfying, but absolutely needs to be stated up front. It is absent from the current manuscript in regard to all other variables.
This is a good suggestion and we thank the reviewer for pointing this out because it helps us clarify our effort to readers. There is a point here that deserves more clarification. What our analysis effectively assumes is not that the modern discharge is the same as at the time of delta formation, but rather that the present deltaic configuration along the shoreline has adapted to modern conditions. Indeed, we now refer in the text to others papers that show delta response time scales are fairly short (order 100-1000 yrs) such that our assumption that river mouths are roughly in equilibrium with current conditions is reasonable. Afterall, work by one of our co-authors shows that waves can easily rework deltas in 100s of yrs (Nienhuis, et al., 2013). We now state this important caveat on P5L18-25: “We use modern data collected for each of these environmental variables, even though some deltas may have initially formed under different conditions 6000 to 8500 years ago as sea-level rise slowed after deglaciation (Stanley and Warne, 1994). We assume that the current river delta (or lack thereof) is adapted to the modern environmental variables because scaling analyses suggest the diffusive response time of river delta deposition and wave reworking is on the order of 100-1000 years (Jerolmack, 2009, Nienhuis et al., 2013). Of course, the diffusive response time depends nonlinearly on delta size, so larger deltas may still be adapting to changing environmental variables.” We also added a line in the introduction at P3L12 that says “We use modern values of these environmental variables under the assumption that present day delta formation is adapted to current conditions.”

It is also worth mentioning that this issue is why we use predam values of sediment and water flux from Milliman and Farnsworth, because it insures that we are using the environmental variable that is most likely responsible for the current delta configuration. Regrettably we did not specify our use of predam values before, but we have now corrected that omission.

In regard to the suggestion of using ‘existence’ in place of ‘formation’ we opt to stick with our original language. The word existence obviously implies formation of a delta and because we do not know when the current deltas achieved their present configuration, we don’t think changing the wording affects our messaging. In this sense, it is possible that many of the deltas in our study have undergone cycles of formation and destruction as water and sediment fluxes and wave fields have changed over time. There are also quite a few documented examples of this presented in Anthony (2015). This makes it difficult to pinpoint when the current deltas experienced their most recent formation. In this way, we view ‘delta formation’ as a nearly continuous process that is always occurring (or not occurring), rather than as a discrete process that happened once 7000 years ago. To help clarify this we refer to ‘modern delta formation’ now.

(2) For a subset of river systems, compile any and all regional historic climate data (may also include paleo reconstructions) to evaluate how the regional conditions were different in the past or perhaps even at the time of delta formation. This will at least give you a distribution of how conditions may have changed over longer timescales and will allow you to determine if some systems might be moving more towards constructive or destructive phases.

This is an intriguing idea. We followed through on this and found that longer term sea-level rates do not actually increase success of predicting delta formation. We used model results from the well-known ICE_6G_C model. Please see P9L2-7 for this new discussion.

With regards to the points made above, part of the difficulty is that it is not straightforward to know when any given delta ‘formed’, which makes it difficult to estimate sea level change over the lifetime of the delta. Even when these longer term RSL rates are used, it still did not have a statistical affect on delta
likelihood and was thrown out of the final logistic regression. There are likely many different reasons for this, but we come back to the point mentioned above: when did the current array of deltas defined in this paper start forming? Perhaps the 26kyr to present time frame is too long to average over, but it at least represents a longer time scale average than the 25 yr duration from AVISO. It is tempting to assume that a duration from 7 kyr to present is more appropriate following Stanley and Warne (1994). But it is worth noting that they record the progradation of deltas following decelerating post glacial sea level. This does not mean the delta did not exist prior to deceleration. In fact, it is possible a smaller delta existed in a more landward portion during the transgression. So even for those deltas the timing of delta formation is not entirely clear.

We think that it is plausible, given the relatively rapid adaptation timescales of deltas we quote above, that deltas may have gone through multiple cycle of construction and destruction. In this sense, it is difficult to cleanly link a ‘time’ of delta formation to historical conditions, which would make it difficult to perform an analysis like this. While this kind of analysis of historic climate data could yield interesting results it does not bear on the outcome of the present study. As we state above, we effectively assume the deltas have adapted to modern conditions, which is reasonable given their relative rapid adaptation timescales. Even if we did find that some deltas have undergone changing conditions, we would still need to show that the morphodynamic adjustment is incomplete. This would necessarily require a modeling study. While we do agree with the reviewer that this may provide some interesting results, we do not think such an analysis will bear strongly on our results or conclusions. Given that our paper is tightly focused we prefer to leave this analysis for a different contribution.


Third, I know that data on upstream and downstream variables is difficult to assemble, especially if you want all apples instead of a mix of apples and oranges. However, I find it surprising that you chose to use only one dataset of river discharge from 2011. Is there not any additional discharge data that can help you use more of your river/delta dataset in your statistical analyses? Have you checked the Global Runoff Data Centre or maybe even the Global Forest Information Service? Increasing your data usage can only make your own dataset(s) more valuable to the community.

Thank you for pointing out these additional datasets. We agree that having more data is strongly preferable, but we wanted to use the best possible data and as the reviewer points out we wanted to avoid comparing different types of data. That is why we used the Milliman and Farnsworth data because they represent measured observations, not models. Moreover, the MF dataset uses available data from the GRDC.

Finally, the paper needs a bit more information on the datasets involved in this study. How did you map the rivers and deltas? Entirely by hand? Over what timeframes are your discharge data? I can go pull the MF2011 dataset myself, but this should be included in your manuscript, as it is for the marine data. Your supplemental table is helpful, but please also include in it your mapped river widths and the values of Qw, Qs, and Ab.

We now clarify that the rivers and deltas are mapped by hand and that the annual fluxes of water and sediment are calculated from many years of data, though the number of years varies for each record.
The Milliman and Farnsworth data are from a book that is available for purchase and we felt it was not appropriate to publish their data here to avoid copyright infringement. We included in our dataset only the data we created, i.e. the marine data. We have now clarified that all of our data are modern in timescale and how we mapped the deltas (we did it by hand). While we did measure the width to make sure the river was > 50 m, it was not recorded and not included in any analyses in this paper and therefore not relevant to reproducing our results. We now state in the methods that the environmental variables were collected over many years.

Minor:

The introduction would benefit from elaborating on what we DO know about the conditions that lead to delta formation. You state that they are not completely known, but you fail to really provide much information about what we do already know, either from a fluid dynamics perspective, sediment supply, or even things like effects of basin depth on delta formation.

We have now added some references and brief discussions in our introduction. Please see P2L9-16

Page 3 line 21: How do you determine (quantitatively) if the protrusion has a ‘relatively smooth depositional shoreline?

We don’t make this determination quantitatively because it would require that the shoreline is defined from high resolution images as current shoreline datasets are spatially too coarse to distinguish between depositional and rocky shorelines.

Why did you choose the 75th percentile of bathymetric slopes?

We choose the 75th percentile because it best characterized the offshore bathymetric slope when we checked our method against select field results.

Page 10 Line 15-16: There is no apparent linear decrease.

We agree that the linear decrease is not strong, and the language we use to describe this reflects that. But, we still maintain that when the tail of the distribution is eliminated there is a linear decrease. We reworded this sentence to more clearly reflect that.

Page 10 Line 21-22: Reserve this observation for the discussion, as it is elaborated on sufficiently there.

Fixed.

Figure 7: Is there a reason this can’t be shown with all the data rather than binned data? At least include as supplemental, if not in the paper itself.

The data must be binned because the data on the y-axis (before binning and averaging) are binary. That is, they represent the observation of a delta or not. So to create a continuous Ld that can be compared to predictions form equation 1 we have to bin the data. We added a sentence in the caption to reflect this.

Page 15 Lines 3-5: But there is a concern with preservation in the rock record as well that makes this difficult.

This concern is now addressed starting at P15L14.
Technical corrections:

Figure 1: Please provide the locations of the coastal environments shown in the caption.

Fixed.

Page 4 Line 15: ‘representative of the river, devoid of significant downstream widening.’

Fixed.

Figure 2: Add to caption what the bounding boxes are or where they’re discussed in the manuscript.

Fixed.

Page 7 Line 19 (and elsewhere basin bathy/depth is mentioned): Add appropriate references. Eg. Carlson et al., 2018, Wang et al., 2019

Added.

Page 10 Line 11: ‘between the mean or median Qs/Qw values’

fixed

Page 11 Line 6: Your independent variables were not ‘collected on all rivers,’ as you state in the next sentence. Please reword.

fixed

Page 13 Line 4: If downstream variables are secondary, reword/reorder to reflect that.

We thank the reviewer for this suggestion but prefer to keep the paragraph organized as it is.
A global delta dataset and the environmental variables that predict delta formation on marine coastlines

Rebecca L. Caldwell¹,², Douglas A. Edmonds¹, Sarah Baumgardner²,³, Chris Paola³, Samapriya Roy⁴, Jaap H. Nienhuis⁵

¹ Department of Earth and Atmospheric Sciences, Indiana University, Bloomington, IN, 47401, USA
² Chevron Energy Technology Company, Chevron Corporation, Houston, TX, 77005, USA
³ Department of Earth Sciences, University of Minnesota, Minneapolis, MN, 55455, USA
⁴ Department of Geography, Indiana University, Bloomington, IN, 47401, USA
⁵ Department of Physical Geography, Utrecht University, Utrecht, NL

Correspondence to: Rebecca Caldwell rcaldwell93@gmail.com

Abstract. River deltas are sites of sediment accumulation along the coastline that form critical biological habitats, host megacities, and contain significant quantities of hydrocarbons. Despite their importance, we do not know which factors most significantly promote sediment accumulation and dominate delta formation. To investigate this issue, we present a global dataset of 5,399 coastal rivers and data on eight environmental variables. Of these rivers, 40% (n = 2,174) have geomorphic deltas, defined either by a protrusion from the regional shoreline, a distributary channel network, or both. Globally, coastlines average one delta for every ~300 km of shoreline, but there are hotspots of delta formation, for example in Southeast Asia where there is one delta per 100 km of shoreline. Our analysis shows that the likelihood of a river to form a delta increases with increasing water discharge, sediment discharge, and drainage basin area. On the other hand, delta likelihood decreases with increasing wave height and tidal range. Delta likelihood has a non-monotonic relationship with receiving basin slope: it decreases with steeper slopes but for slopes > 0.006 delta likelihood increases. This reflects different controls on delta formation on active versus passive margins. Sediment concentration and recent sea-level change do not affect delta likelihood. A logistic regression shows that water discharge, sediment discharge, wave height, and tidal range are most important for delta formation. The logistic regression correctly predicts delta formation 75% of the time. Our global analysis illustrates that delta formation and morphology represent a balance between constructive and destructive forces, and this framework may help predict tipping points where deltas rapidly shift morphologies.

1. Introduction

Deltas provide a variety of ecosystem services, such as carbon sequestration and nitrate removal (Rovai et al., 2018; Twilley et al., 2018), and they provide home to close to half a billion people (Syvitski and Saito, 2007) living within large
agricultural and urban centers (Woodroffe et al., 2006). Deltas form at river mouths where fluvial sediment accumulates nearshore long enough for the deposit to become subaerial. This simple view of delta formation is a statement of sediment mass balance; and understanding where deltas form requires knowing how and why sediment accumulates. Sediment accumulates provided it is supplied and deposited at the coast faster than it is removed. Sediment supply and removal are chiefly determined by the river, waves, tides, rate of relative sea-level change, and offshore bathymetry. To complicate matters, most of these variables can be both sources and sinks of sediment, and their exact roles in the deltaic sediment mass balance remains uncertain. Previous research suggests that rivers are almost always sources (Bates, 1953; Coleman, 1976; Wright, 1977; Syvitski et al., 2005; Syvitski and Saito, 2007), whereas the roles of waves and tides are largely ambiguous (Nienhuis et al., 2015; Hoitink et al., 2017; Lentsch et al., 2018), though there is some evidence to suggest waves are mainly sediment sinks in the delta formation process (Fisher, 1969; Anthony, 2015). The bathymetric characteristics of the offshore basin determine the nearshore hydrodynamics, wave power, and structure of the turbulent jet, which in turn influences sediment deposition patterns and delta formation (Fagherazzi et al., 2015; Jiménez-Robles et al., 2016). Sea-level is also an important part of delta formation, and we know that slower rates of sea-level rise promote delta formation (Stanley and Warne, 1994; Porebski and Steel, 2006; Paola et al., 2011).

Despite these efforts, we do not fully understand what how these different controls combine to create river deltas. The conditions that lead to delta formation are not completely known, but we know the conditions for delta formation are not easily met—pick nearly any oceanic marine shoreline on earth and there will be several of the river mouths that intersect the coast, but only some of these rivers will have a delta. Previous studies on delta formation (Wright et al., 1974; Audley-Charles et al., 1977; Milliman and Syvitski, 1992; Syvitski and Saito, 2007; Nyberg and Howell, 2016) focused on large-scale patterns and concluded that major modern delta locations are influenced largely by tectonic margin type and drainage patterns. While useful, these datasets were biased toward the largest and most populated deltas. Expanding the prediction effort to deltas of all sizes is a logical next step, especially because smaller deltas are thought to be more resilient to rising sea levels (Giosan et al., 2014).

In addition to expanding the range of delta sizes, to understand the controls on delta formation, we need to consider cases where delta formation is suppressed. In this paper we present a global delta dataset and use it to investigate why some
rivers form deltas and others do not. Understanding conditions for modern delta formation should also help exploration for ancient deltaic deposits, which requires predicting where deltas might form under past environmental conditions (Nyberg and Howell, 2016). Similarly, as research moves towards delta risk assessment due to global environmental change (Tessler et al., 2015) and improving efforts to build new deltaic land (Kim et al., 2009), we must understand how different environmental variables govern delta formation. For example, understanding the conditions for delta formation would help restoration efforts that seek to build new deltaic land in places like the Mississippi River Delta (Paola et al., 2011; Edmonds, 2012; Twilley et al., 2016).

To address these issues, we developed a global dataset that includes the locations of 5,399 coastal rivers, information on whether they form deltas or not, and the environmental variables important for that could influence delta formation. We use global datasets of coastlines (Dürr et al., 2011; Nyberg and Howell, 2016), sediment and water (Syvitski and Milliman, 2007; Milliman and Farnsworth, 2011), wave climate hindcasts (Tolman, 2009; Chawla et al., 2013), a tidal inversion model (Egbert and Erofeeva, 2002), ocean bathymetry data (Amante and Eakins, 2009), and rate of sea-level change (https://www.aviso.altimetry.fr). We use modern values of these environmental variables under the assumption that present day delta formation has adapted to current conditions. Of the 5,399 included rivers, 2,174 form geomorphic deltas that are visible in aerial imagery, defined either by a protrusion from the regional shoreline, a distributary channel network, or both. We use statistical relationships between independent environmental variables and the presence or absence of a delta to determine what controls the likelihood of a river to form a delta.

2. A global coastal and river delta dataset

2.1 Identifying river deltas

River deltas are fundamentally systems of sediment accumulation and distribution at the coastline. Accordingly, we identify coastal deltas by distinguishing geomorphic expressions of sediment accumulation and distribution at locations where rivers meet the coast. We consider a river to have formed a delta at the coastline if the river-mouth area contains an active or relict distributary network (Fig. 1c), ends in a subaerial depositional protrusion from the lateral shoreline (Fig. 1d), or does both (Fig. 1c). Distributary networks are an expression of sediment deposition and distribution (Edmonds et al., 2011) and we
identify them by the presence of one or more channels that bifurcate and intersect the coast at different locations. We include relict channels, where they are clearly visible in imagery and connect to the main channel, because they are evidence of sediment distribution and accumulation through avulsion (Slingerland and Smith, 2004). We do not include channels that bifurcate solely around non-deltaic topographic highs. Our second criterion is oceanward-directed shoreline protrusions. We classify a protrusion as deltaic if it has a relatively smooth depositional shoreline, as opposed to rough shorelines associated with rocky coasts (Limber et al., 2014), and if it extends more than ~5 channel widths oceanward relative to the position of the regional shoreline. We map only map protrusions that are associated with the river, ignoring protrusions that may exist near the channel mouth that we judge to be pre-existing undulations in the shoreline. Examples of this include promontories associated with pre-existing geology or depositional protrusions created by other processes, such as wave-driven sediment transport (Ashton et al., 2001).

Our delta identification method does not account for deltaic deposition with no geomorphic signature, such as a single-channel delta infilling a drowned valley that produces no protrusion from the regional shoreline. Although such features may be considered deltaic, we cannot unambiguously identify them as deltas based on aerial imagery alone and we do not include them in the dataset.

We applied the preceding criteria to a scan of oceanic marine coastlines, including most open ocean coasts and the Black sea, using Google Earth. First, we identified all all rivers with width > 50 m reaching the coast that are connected to an upstream catchment (Figs. 1a, 1c–1e). Channels not clearly connected to an upstream catchment, such as tidal channels, were not included in the dataset (Fig. 1b). This was done to restrict the study to coastal depositional landforms that represent the interaction of upstream and downstream environmental variables. We selected rivers at least 50 m in width because they have corresponding data, such as basin area, that can be reliably determined on coarser resolution elevation models. This width designation was applied to the rivers’ bankfull widths, and thus includes any visible mid-channel bars. Channel widths on rivers without a delta were measured at the shoreline or upstream from visible marine influence, such as significant tidal widening (Nienhuis et al., 2018). If a river empties into a gradually widening estuary or embayment, we measured the channel width where it is devoid of a significant downstream widening and thus representative of the river devoid a significant downstream widening. Channel widths on rivers that have deltas were measured immediately upstream of the delta node,
which we define as the location of the most upstream bifurcation. If no bifurcation exists, we use the intersection of the main channel with the regional shoreline (e.g., Fig. 1c and 1d, blue dot). In all cases, channel widths were not measured in areas of clear human influence. This includes, for example, man-made levees that can cause artificial widening or narrowing of channels.

We mapped rivers and deltas on the coastlines of Earth’s continents and large islands (Fig. 2). We exclude small islands where rivers large enough for inclusion are rare and it is difficult to obtain environmental data. Thus, large islands, such as Papua New Guinea and Fiji, were included but not all the associated smaller islands. Coastlines dominated by fjords (as determined using Dürr et al. (2011)) were not included because offshore glacial over deepening and protection from coastal waves and tides make their comparison to most of the world’s coastal deltas difficult. Ephemeral rivers in arid regions were included in the dataset, though the rivers in these regions are often difficult to identify due to poor imagery and difficulty distinguishing the channel banks when they are dry. If a clear distinction was not possible, the river was not included in the dataset. Thus, the total count of rivers and deltas in arid regions should be considered a minimum. Finally, we did not include river channels that do not clearly reach the coast to avoid conflating alluvial fans with deltas.

For each river we marked the latitude and longitude of the main river mouth (Figure 1, RM) (Supplemental Table 1). For rivers without a delta, this is the location where the river meets the coastline (Fig. 1a), and for rivers with deltas, this is the location of the widest river mouth in the distributary network (Fig. 1c–1e). For rivers sheltered by barrier islands or rocky islands, we mark the river mouth landward of those obstructions.

2.2. Environmental variables

To determine controls on delta formation we also compiled data on eight environmental variables (Table 1). We classify the environmental variables into two groups: (1) upstream variables include water and sediment supply from the river, sediment concentration, and the drainage basin area; and (2) downstream variables include wave heights, tidal ranges, bathymetric slopes immediately offshore of the river mouth, and the rate of sea-level change. We use modern data collected for each of these environmental variables, even though some deltas may have initially formed under different conditions 6000 to 8500 years ago as sea-level rise slowed after deglaciation (Stanley and Warne, 1994). We assume that the current river delta...
(or lack thereof) is adapted to the modern environmental variables because scaling analyses suggest the diffusive response time of river delta deposition and wave reworking is on the order of 100-1000 years (Jerolmack, 2009, Nienhuis et al., 2013). Of course, the diffusive response time depends nonlinearly on delta size, so larger deltas may still be adapting to changing environmental variables.

Notably absent in the collected environmental variables are tectonic data. At present, there are no globally available measurements of tectonic activity (e.g., uplift). However, we consider some of the variables to be reasonable proxies for tectonics. For instance, models predicting sediment flux to the ocean represent tectonics in the form of basin area (Syvitski and Morehead, 1999; Syvitski and Milliman, 2007). We also include bathymetric slope, which is a rough proxy for tectonics because, on average, tectonically active margins have steeper slopes than passive margins (Pratson et al., 2007).

2.2.1. Upstream variables

We compiled the four upstream variables from the global river dataset of Milliman and Farnsworth (2011) (hereafter referred to as MF2011). We matched rivers in this dataset with entries in MF2011 based on geographic proximity or by the river name. If neither matching method yielded a confident result, the MF2011 data were not included in this study. If two or more rivers in the MF2011 dataset combine to make one river in this study’s dataset, the data from all relevant MF2011 rivers are included. In cases where matches were found, we included the river ID(s) from MF2011 in our dataset (Supplemental Table 1). Our dataset includes 1,217 MF2011 rivers, representing 1,158 entries in our dataset, (54 entries delta entries contain are made from 2 or more MF2011 rivers), and in those cases we added together the MF2011 values together to form one value for the river mouth or delta. There are 314 MF2011 rivers not included in this dataset because they are too small (< 50 m wide), exist on coastlines not included in our dataset, or could not be matched.

Water discharge \(Q_w\), expressed as mean annual volumetric flux, \(\text{m}^3 \text{s}^{-1}\) data in MF2011 come from the MF2011 dataset. The \(Q_w\) measurements are compiled from various sources of reported gauging station measurements, where the downstream-most gauging station data is used. \(Q_w\) is computed from many years of data, though the number of records for each value is different. Where available, we used the pre-dam \(Q_w\). As MF2011 note, water discharge values may be over- or under-estimated due to distance upstream of the river mouth. In many regions, additional water input downstream of the gauging station increases the true \(Q_w\) value reaching the river mouth. However, in arid regions, water volume may be lost due
to evapotranspiration, groundwater recharge, or irrigation water removal. In total, 17% of rivers (n = 943) in this dataset have $Q_w$ data.

Sediment discharge ($Q_s$, expressed as mean annual volumetric flux, $\text{m}^3\text{s}^{-1}$) data come from the MF2011 dataset of annual sediment load measurements and are converted to $\text{m}^3\text{s}^{-1}$ assuming a density of 2650 kg $\text{m}^{-3}$. $Q_s$ is computed from many years of data, though the number of records for each value is different. Where available, we used the pre-dam $Q_s$. The $Q_s$ data are compiled from various sources of reported loads and most often represent suspended load measurements rather than total load. Bedload is assumed to represent only 10% of total load (Milliman and Meade, 1983), but this estimation may be less valid for small mountainous rivers where relative proportion of bedload can be greater (Amante and Eakins, 2009). Like the $Q_w$ data, many of these measurements may have been made upstream of the actual river mouth, and thus actual $Q_s$ values that reach the river mouth likely vary (e.g., due to fluvial plain deposition downstream of measurement location). Finally, extrapolation of measurements taken over varying lengths of time to represent annual sediment loads is potentially risky (e.g., when considering the significance of event-driven discharge events). In total, 11% (n = 600) of all rivers in this dataset have $Q_s$ data. Sediment concentration ($Q_s/Q_w$) is calculated from the sediment and water discharge data and 11% (n = 571) of all rivers have $Q_s/Q_w$ data.

We also include upstream drainage basin area ($A_b$, km$^2$) in our dataset because it partly sets the magnitude of $Q_w$ and $Q_s$ (Syvitski et al., 2003; Syvitski and Milliman, 2007) and compensates for the relatively small number of rivers with water and sediment data. $A_b$ data come from the MF2011 dataset. Although these values are often well documented for larger river systems, they may sometimes represent the total drainage area upstream of a hydrologic station, which would be a smaller value than total drainage area upstream of the river mouth. Given the potential error, $A_b$ values should be considered a minimum.

### 2.2.2. Downstream variables

Four downstream variables are included in this dataset. Annual significant wave heights ($H_s$, m) were calculated using the NOAA WAVEWATCH III 30-year Hindcast Phase 2 for 1979–2009 (Tolman, 2009; Chawla et al., 2013). The model outputs 30 years of hourly significant wave height data on five different ocean grids with varying resolution, and the final product is interpolated to a global 0.5-decimal degree grid. We ran a nearest-neighbor search from each $RM$ location to
the nearest grid cell with wave data that is within one grid cell diagonally, which is equivalent to 0.7071 decimal degrees, or 
~80 km at the equator. Because some coasts are missing wave data not all 5,399 rivers have corresponding wave data. For 
each calendar year, we calculate the annual mean of the top 1/3 largest wave heights. The resulting 30 years of annual 
significant (top mean of the largest 1/3 largest) wave height data are representative of the strongest wave action that occurs at 
each location within a year, or representative of a stormy season for areas with strongly seasonal wave climates. The mean of 
these 30 annual values is the mean annual significant wave height ($H_w$).

Median tidal ranges ($H_t$, m) were calculated using the previously published Oregon State University TOPEX/Poseidon 
Global Inverse Solution TPXO model results (Egbert and Erofeeva, 2002). The model outputs tidal harmonics component data 
on a 0.25-decimal degree resolution grid derived from a barotropic inverse solution. Following Baumgardner (2015), we use 
the main tidal components, the lunar semidiurnal and the lunar diurnal, to calculate mean tidal range by building a composite 
tidal sine wave and calculating the average range. We ran a nearest neighbor search from each RM location to all grid cells 
with tidal data that are within the same distance used for the wave search. The median of the tidal range values within this 
search radius is used to represent each river mouth’s tidal range.

Receiving-basin bathymetry is an important attribute of delta formation because it sets the size and shape of the 
volume to be filled from a mass balance perspective, and influences the hydrodynamics (Jiménez-Robles et al., 2016; Carlson 
et al., 2018). The size of the basin could be characterized by the average depth whereas the shape is most simply characterized 
by the bathymetric slope. In most cases, we do not know basin depth prior to delta formation, and current depths offshore 
deltaic river mouths will be deeper than the initial depths if the basin has offshore-dipping bathymetric slopes. Thus, instead 
of using depth, we characterize the receiving basin with bathymetric slopes. Bathymetric slopes ($S_b$) are calculated from 
ETOPO1 bathymetric data (Amante and Eakins, 2009) and RM locations. ETOPO1 is a global surface elevation model with 1 
arc-minute resolution (1/60 decimal degree, or ~1.800 m at the equator). For each river, we collect all bathymetric elevations 
within a 20-km radius from the RM location. We calculate linear slopes between each point and the RM (assumed elevation = 
0 m), and take $S_b$ as the 75th percentile of all slopes. We chose the 75th percentile because it best captures the bathymetric 
slope when we compared our $S_b$ values to spot measurements. We purposefully search far away from the shoreline because 
we want to characterize the offshore depths not affected by sediment deposition from the river.
Rate of sea-level change is calculated from AVISO (Archiving, Validation and Interpretation of Satellite Oceanographic data, https://www.aviso.altimetry.fr). The AVISO dataset combines sea-level change from different satellite altimetry missions from 1992-2018 using the delayed time Ssalto/Duacs multi-mission altimeter data processing system, which corrects biases among instruments and applies inter-calibration to the record. Rates of sea-level change are calculated for every 0.25° x 0.25° cell by finding the best fit to the data over 26 years. The data we use are not corrected for glacial isostatic adjustments. We used modern rates that are measured (not modelled) to be consistent with the other environmental variables. The rates we use are decidedly modern. An argument could be made that although it would be ideal we should compare delta formation to sea-level change data averaged over a their lifespan, not the lifespan of the delta, or these data do not exist. But, we do not know when each delta in this dataset started forming. Nonetheless, if we assume that all deltas started growing since the most recent deglaciation, we can compare modeled relative sea-level change rates since 26 kyr ago to delta formation (Argus et al., 2014). When we use these longer term sea-level change rates, it does not increase the success of predicting delta formation using the method in section 3.2, and because of this we opt to use the modern sea-level change data for consistency. But, those data do not exist. That makes it difficult to compare with deltas, many of which formed 1000s of years ago as sea-level rise started slowing following deglaciation (Stanley and Warne, 1994). It would be ideal to compare delta formation to sea-level change data averaged over their lifespan, but those data do not exist.

3. Results

Our mapping reveals there are 5,399 coastal rivers with widths greater than 50 m, and 2,174 of those rivers (~40%) have a geomorphic delta. Herein, we refer to all 5,399 coastal rivers as “rivers”, the 3,225 that do not have deltas as “river mouths,” and the 2,174 with deltas as “deltas.” These terms are not completely accurate because, for example, an individual “river” that is considered a “delta” rather than a “river mouth” still has at least one main river mouth (RM) and may have additional river mouths for each distributary channel.

3.1. Global distribution of rivers and deltas

River deltas are not distributed evenly on coastlines and there are locations on the world’s coastlines where deltas are unusually common (Fig. 2). These “delta hotspots” occur primarily in Southeast Asia (dashed box Fig. 2b). Notably, these
areas are also densely populated with rivers (Fig. 2a), though river abundance does not always equate to delta abundance. For example, East Asia has high river density but low delta density (black box, Fig. 2b). Similarly, along the west coasts of Central and southern North America (from 5°N to 45°N) the coast is densely populated with rivers, but the northern portion is delta-poor compared to the southern portion. There are also a surprising number of deltas in arid environments. For instance, there is high delta density in the Red Sea and on Baja California. This largely arises because the many alluvial fans coming off the nearby mountains reach the coastline and satisfy our definition of a delta.

Binning these data by latitude reveals preferential locations of rivers and deltas (Fig. 3). The largest numbers of rivers and deltas occur roughly from -12° to 45°, and 66° to 72° (Fig. 3a). This unequal distribution is partly explained by the unequal latitudinal distribution of global shoreline length (Wessel and Smith, 1996) (Fig. 3b). River density, or rivers per shoreline kilometer, shows that globally there is one river for every 230 km of coastline and one delta for every 333 km of coastline. Coastlines within the -6° to -3° bin have the highest density of deltas with roughly one delta per 100 km of shoreline (Figure 3c, solid black line). River density is above average from -45° to 45° (Fig. 3c, white bars). Delta density, however, is above average over a smaller range from -21° to 30° (Fig. 3c, solid black line).

To determine which environments promote delta formation, it is perhaps most instructive to observe locations where the likelihood for rivers to create deltas is highest. Delta likelihood \( L_d \) is defined as the number of deltas relative to the total number of rivers for a given set of samples (Fig. 3d, solid black line). For the entire dataset, 40% of rivers form deltas, and thus the global \( L_d \) is 0.40 (Fig. 3d, dashed black line). Regions where \( L_d \) is higher than the global mean exist from -27° to 30° and 60° to 72°, whereas rivers located from -57° to -27° and 30° to 60° are least likely to form a delta (Fig. 3d).

These latitudinal zones, where rivers are more likely to create deltas, coincide with peaks in environmental variables that influence delta formation. Both \( Q_w \) and \( Q_s \) have notable peaks from roughly -9° to 30° and 60° to 75° (Fig 4a, b), which are similar in location to \( L_d \) peaks. \( S_b \) has a similar high latitude peak, but is missing the equatorial peak (Fig. 4d), probably reflecting the importance of small mountainous rivers in those locations (Milliman and Syvitski, 1992). On the other hand, delta formation is infrequent where \( H_w \) and \( H_s \) are high, namely -57° to -27° and 42° to 60° (Fig. 4e, f). There are no latitudinal changes in \( Q_s/Q_w \), \( S_b \), or \( H_s \) that are easily relatable to delta formation (Fig. 4c, g, h).
3.2. Relationships between environmental variables and delta formation

We explore controls on delta formation by analyzing how the likelihood of a river creating a delta varies with each environmental variable. River mouths and deltas have statistically different population distributions for seven of the eight environmental variables (all but \( Q_s/Q_w \)) (Table 2), suggesting that deltas form under certain ranges of environmental variables.

To determine this, we used the Kolmogorov-Smirnov test, which is a non-parametric, distribution-free test that uses the cumulative distribution functions of the two populations to estimate statistical difference. Although a few variable pairs show some correlation, such as \( Q_w \) and \( A_b \), none have a strong statistical correlation (Pearson correlation coefficient > 0.9), suggesting they exert largely independent controls on delta formation.

Delta likelihood \((L_d)\) generally increases as the upstream environmental variables increase (Fig. 5). Increasing \( Q_n, Q_s, \) and \( A_b \) causes a linear increase in semi-log space in \( L_d \) (Figs. 5a-b, d). Deltas have characteristic \( Q_n, Q_s, \) and \( A_b \) values that are an order of magnitude larger than those of river mouths (statistically significant, \( p < 0.05 \)) (Table 2). These data suggest that rivers with small water and sediment discharge and/or that come from small drainage basins rarely form deltas, whereas rivers with larger values of the upstream variables frequently create deltas. Sediment concentration \((Q_s/Q_w)\) exerts no clear control on \( L_d \) (Fig. 5c), and there is no statistical difference between the mean and median \( Q_s/Q_w \) values for rivers mouths versus those for deltas (Table 2).

Rivers are less likely to create deltas where \( H_w \) and \( H_t \) are large. \( L_d \) shows a clear linear decrease as \( H_w \) increases (Fig. 6a). Rivers that experience little wave energy at the coast \(( H_w < 1 \text{ m})\) create a delta more than half of the time \((L_d \approx 0.5-0.6)\), but delta formation becomes nearly impossible for larger wave heights. \( L_d \) also seems to show a linear decrease with \( H_t \) (Fig. 6b), if the long tail of the distribution is eliminated where the sample size is small \((H_t > 7 \text{ m})\), but this relationship shows significantly more scatter than that with \( H_w \). If the long tail of the distribution is eliminated where the sample size is small \((H_t > 8 \text{ m})\), the relationship is clearer. The population of river mouths has higher mean and median \( H_w \) and \( H_t \) than rivers with deltas (statistically significant, \( p\)-value < 0.05) (Table 2).

\( S_b \) displays a non-monotonic relationship where \( L_d \) decreases then increases across the range (Fig. 6c). \( S_b \) data are bimodally distributed for the rivers in our dataset, suggesting rivers empty into two types of receiving basins (separated by the dashed gray line in Fig. 6c). Based on visual observation, the shallowly dipping basin types reflect passive margins, and the
...steeply-dipping basins active margin, though we did not pursue a more robust confirmation. If these basin types are separated, there is a clearer relationship between $S_b$ and $L_d$. For shallowly-dipping basins ($S_b < 0.006$), there is a negative relationship between $L_d$ and $S_b$ (Fig. 6c, left of dashed gray line), and delta likelihood increases as slope decreases. In steeply-dipping basins ($S_b > 0.006$), $L_d$ is approximately constant to slightly increasing as slopes steepen (Fig. 6c). There is no clear relationship between sea-level change ($H_s$) and $L_d$ (Fig. 6d), which is somewhat surprising given that river mouths and deltas have statistically different mean and median $H_s$ values (Table 2).

To quantify the relative importance of the environmental variables for delta formation, we develop an empirically-derived logistic regression. The result of a logistic regression is a statistical model that predicts a dichotomous outcome (in this case, a river creates a delta, or it does not) based on multiple independent variables. This dataset contains eight total independent variables collected on all most rivers, where four are upstream variables ($Q_u$, $Q_s$, $Q_v$, $A_b$), and four are downstream variables ($H_w$, $H_t$, $H_d$, $S_b$). Of the 5,399 rivers in this dataset, 490 of them (9.1%) have data available for all eight independent variables.

The data meet the assumptions of binary logistic regression because the dependent variable has two mutually exclusive outcomes, and the sample size is large (45 samples or more per independent variable). Additional assumptions that the data must meet include having little to no multicollinearity and no outliers. We tested for multicollinearity by calculating the Pearson correlation coefficients ($R$) between all continuous independent variables and no variables exhibited $R > 0.9$. We also remove 143 rivers that have outliers in any of the independent variables based on a modified $z$-score, where an absolute value modified $z$-score > 3.5 is considered an outlier (Iglewicz and Hoaglin, 1993). The final subset of data used for the regression has $n = 4776$ rivers (2498 rivers without deltas, 228 rivers with deltas). The samples were randomly separated into training (2/3 of the samples) and validation (1/3 of the samples) subsets, each of which represented similar distributions of independent and dependent variables. We do this to see how well the logistic regression can predict delta formation on river mouths not used in the development of original equation.

The binary logistic regression between the probability that a river will create a delta and the eight environmental variables yields the following log odds relationship:

$$
\ln(\frac{\pi_{delta}}{1 - \pi_{delta}}) = 1.3845 + 0.000524Q_u - 0.00589Q_s + 0.24562Q_v - 0.97452H_w - 0.17875H_t - 0.12875H_d
$$

(1)
where $\pi_{\text{delta}}$ is the probability that a river will form a delta, and ranges from 0 (river is unlikely to form a delta) to 1 (river is most likely to form a delta). This is different from the $L_d$ values presented earlier only because it is predicted whereas $L_d$ was measured. Environmental variables with $p > 0.05$ ($Q_s/Q_w, A_b, S_b, \text{and} H_s$) are not included in the final empirical relationship, because any controls these variables exert on delta formation is minimal (e.g., variations in $Q_s/Q_w$ have no clear relationship with $L_d$, Fig. 5d) or are related to variations in the other important variables (e.g., $A_b$ influences $Q_w$ and $Q_s$).

Thus, the combination of environmental variables that comprises the right side of equation (1) predicts the log odds that a river will form a delta. When tested using the validation subset, equation (1) has a 75.4% success rate at predicting delta presence (Fig. 7), where $\pi_{\text{delta}} > 0.5$ is considered a prediction that a delta exists, and $\pi_{\text{delta}} < 0.5$ is considered a prediction that no delta exists.

This empirically-derived relationship can be used to calculate the probability that a certain combination of the most important environmental variables will form a delta. For example, using environmental variable values for the Godavari River in the right-hand side ($\text{RHS}$) of equation (1) results in $\text{RHS} = 3.93$. The probability that the Godavari River should form a delta is $\pi_{\text{delta}} = \frac{e^{\text{RHS}}}{1+e^{\text{RHS}}} = 0.98$. Thus, the environmental variables that conspire to form the Godavari River are very likely to form a delta, which is not surprising given the existence of the large Godavari River delta.

4. Discussion

4.1. Which environmental variables most strongly control delta formation?

We have considered the relationships between eight environmental variables and delta formation. However, and determining which ones matter the most for delta formation is not straightforward. After all, most combinations of environmental variables that exist globally completely suppress delta formation (60% of the rivers included in this dataset do not have a delta). Our likelihood analysis shows that deltas are more likely to form at river mouths with large water discharge $Q_w$ (Fig. 5a), sediment discharge $Q_s$ (Fig. 5b), and drainage basin area (Fig. 5c), and with small significant wave heights $H_s$ (Fig. 6a), and tidal ranges $H_t$ (Fig. 6b). Results suggest that upstream variables exert a primary control. Increasing upstream variables ($Q_s, Q_w, A_b$) across their value range accounts for the full range of delta likelihood—that is, the smallest $Q_s, Q_w$, and $A_b$ values have $L_d \approx 0$, and largest $Q_s, Q_w$, and $A_b$ have $L_d \approx 1$ (Figure 5). Downstream
variables seem to be of secondary importance for forming deltas. In contrast, increasing the downstream variables \((H_w, H_t)\) decreases the likelihood that a river forms a delta, but does not produce the full range of possible \(L_d\) values. At the lowest values of \(H_w\) and \(H_t\), delta likelihood is still 0.5. The relationship with \(H_w\) is more significant, it has a steeper slope and less scatter compared to \(H_t\). In fact, downstream variables seem to be of secondary importance for forming deltas. Furthermore, when we remove \(H_w\) and \(H_t\) from equation (1) the prediction success rate decreases by only 3%, from 75.4% to 71.2%.

These controls on delta formation explain first-order latitudinal variations observed in Figures 3 and 4. For example, the peaks in water and sediment discharge values from -9° to 30° and 60° to 75° (Fig. 4) likely explain the similarly located peaks in delta formation (Fig. 3). The suppressing effects of waves and tides can also be seen at a global scale. Low delta formation rates from -57° to -27° and 30° to 60° are likely due to large \(H_w\) and \(H_t\) values in these regions, where \(Q_w\) and \(Q_s\) are low (Figs. 3, 4). Moreover, the zone from 60° to 75° that has increased \(Q_w\) and \(Q_s\) values (Fig. 3) also has some of the lowest \(H_w\) and \(H_t\) values (Fig. 4). Thus, while high \(Q_w\) and \(Q_s\) values in this region promote delta formation, the decreased \(H_w\) and \(H_t\) values also allow delta formation to occur.

Downstream bathymetric slope \((S_b)\) displays a complex relationship with delta likelihood. At slopes < 0.006, delta likelihood decreases with increasing slope (Fig. 6c), because all else being equal, deeper areas should take longer to fill with sediment and they are also less effective at damping incoming waves and tides. But, interestingly, for slopes > 0.006, delta likelihood increases with steeper slopes, which is more difficult to explain. Based on visual observation, the steeply-dipping basins reflect active margins, and the shallowly-dipping basin types reflect passive margins, though we did not pursue a more robust confirmation. If these steeper margins relate to active margins, then larger sediment sizes and higher supply on active margins may explain the difference with delta likelihood than that for the shallowly-dipping basin types (Audley-Charles et al., 1977; Orton and Reading, 1993; Milliman and Farnsworth, 2011). After all, the supply of coarser sediment to the coast is more easily retained nearshore (Caldwell and Edmonds, 2014), thereby increasing the likelihood of delta formation.

4.2. The roles of rivers, waves, and tides in delta formation

Our data suggest that deltas are fundamentally created by water and sediment discharge, whereas waves, and possibly tides, suppress delta formation. This is consistent with the notion that delta formation is the result of
constructive upstream forces set by the river, and destructive downstream marine forces (Fisher, 1969, Boyd et al., 1992, Anthony, 2015). This idea, initially proposed by Fisher (1969), provides a different perspective compared to existing thoughts on delta formation. The Galloway (1975) diagram is to the foundational study of delta morphology and formation from Galloway (1975). Galloway’s diagram implies that deltaic formation and morphology is the result of the interplay of river, waves, and tides. But, Galloway’s diagram remains largely qualitative and it is not clear how the forces of rivers, waves, and tides are quantified, nor it is clear what kinds of predictions the diagram makes. In fact, our data offer a different view of deltaic formation than the one proposed by Galloway. Our data suggest that delta formation is the result of constructive upstream forces set by the river, and destructive downstream marine forces. Consider the case of a purely wave dominated delta, Galloway’s diagram would predict a cuspate delta, but instead, our data clearly show that the most wave-dominated delta is no delta at all, consistent with other works (the work of Nienhuis et al., 2013; Boyd et al., 1992). This suggests to us that the concept of delta formation and morphology might be better cast as a balance between constructive and destructive forces as initially proposed by Fisher (1969).

From this perspective new questions emerge: If we consider the perspective that delta formation is the result of a balance between constructive and destructive forces, then new questions emerge: How do wave and tidal processes influence the ability of fluvial processes to construct deltas? How stable is the balance between a given set of constructive and destructive forces? With regard to the last question, there are examples of rapid changes in delta morphology through time, which suggests that the balance can be precarious. The Rhône River delta clearly shifted in morphology from channel-network dominated in the 16th century to its more familiar wave-smoothed shape today, as floods and sediment loads declined during the Little Ice Age (14th-19th centuries) (Provansal et al., 2015). The Po River delta in Italy showed three morphological transitions each time the balance between river and waves changed over the last 4000 years (Anthony et al., 2014). These examples from the past should direct our attention to how the current configuration of deltas might change in the future. We know that anthropogenic climate change is changing wave conditions (Reguero et al., 2019) and humans are drastically changing water discharge and sediment flux to coastal rivers (Syvitski and Milliman, 2007). It is unclear how the coastal deltas of the world will adapt as these changes in boundary conditions change. Future work would
benefit from linking our empirically derived delta likelihood predictor with metrics of delta morphology to understand when morphological shifts might occur.

5. Implications

River deltas are the final filters of sediment before it is discharged to the global oceans (Sawyer et al., 2015). Although only 40% of rivers in our dataset form deltas, our results show that 5.9 Bt/yr, or 85% of the measured global sediment flux (Milliman and Farnsworth, 2011), enters river deltas before reaching the ocean. This is not entirely surprising because the presence of a delta requires sediment and our data show that sediment rich rivers tend to create deltas. But, we currently do not know what proportion of that sediment is retained in the delta. This retention should be considered when calculating global sediment flux to the oceans (e.g., Milliman and Farnsworth, 2011), because deltas are exceptionally good at impounding sediment since channel networks are optimized to achieve this goal (Edmonds et al., 2011; Tejedor et al., 2016; Tejedor et al., 2017). Limited calculations suggest deltas can retain up to 30% of the sediment supplied to them (Goodbred and Kuehl, 2000; Syvitski and Saito, 2007; Kim et al., 2009). However, as we have shown here, certain environmental variables promote sediment accumulation via delta formation, and these same environmental variables may promote sediment retention by certain deltas. Thus, our results may prove useful for quantifying the full, and presently unaccounted for, deltaic sink in the global sediment cycle. River deltas are the final filters of water and sediment before they are discharged to the global ocean (Sawyer et al., 2015). As we have shown here, certain environmental variables promote sediment accumulation and delta formation. This accumulation results in the storage of sediment, yet all existing efforts to calculate sediment flux to the global ocean ignore sediment deposited in deltas (Milliman and Farnsworth, 2011). In an analogy with blue carbon, we define the volume of sediment deposited on the coastline, in deltas, or just offshore, as “blue sediment.” Our results suggest that the amount of blue sediment stored in river deltas at yearly to millennial timescales could be significant. Based on our results, we find that 5.9 Bt/yr, or 85%, of the measured global sediment flux (Milliman and Farnsworth, 2011), moves through a river delta before being discharged into the ocean. This is important because deltas are exceptionally good at impounding sediment because their extensive channel networks self organize to evenly cover the topset, so that during flood all areas are nourished with sediment (Edmonds et al., 2011; Tejedor et al., 2016; Tejedor et al., 2017). Limited calculations suggest deltas retain 30% of the sediment...
supplied to them (Goodbred and Kuehl, 2000; Syvitski and Saito, 2007; Kim et al., 2009), in which case deltas may be an important, and presently unaccounted for, sink in the global sediment cycle.

We also propose that our data and analyses have important implications for resource exploration and coastal restoration. Although using equation (1) to predict delta formation for modern rivers is somewhat redundant, it may prove useful for predicting past or future delta existence. Ancient deltaic deposits comprise significant hydrocarbon reservoirs, and provided this analysis holds through geologic time, we could predict the presence of deltaic deposits in the rock record if $Q_w$ and $Q_s$, or $A_b$, can be estimated via other paleohydraulic methods. If we use a logistic regression that does not include the less dominant limiting effects of waves and tides, then equation (1) becomes:

$$\ln(\frac{\pi_{\text{delta}}}{1 - \pi_{\text{delta}}}) = 0.0016 + 0.0175Q_w + 0.0345Q_s$$

Using this simplified equation, which shows a 71% success rate when tested using the validation subset, we can predict the likelihood of delta formation for paleoenvironments where sediment and water discharge are more easily constrained. For example, water and sediment flux estimates for rivers of the Ferron Sandstone in the Cretaceous Western Interior Seaway of the United States (Bhattacharya et al., 2016) suggests the likelihood of delta formation is 99%, and indeed, the Ferron contains deltaic deposits. Even though this is somewhat obvious since we know the Ferron contains deltaic deposits, it highlights how our results could be used to predict the presence of deltaic deposits in the rock record the absence of corroborating direct observational evidence.

Looking forward, this relationship can be used to predict future deltaic formation. Global environmental change will continue to put coastal environments at risk, largely by land loss due to accelerated sea-level rise and decreased sediment delivery to the coast. Coastal restoration and hazard-mitigation techniques often involve the creation of new deltaic land via controlled river diversions (e.g., Kim et al. (2009)), though it can be difficult to predict the risk related to such projects. Predictions made using equation (1) can help the decision-making process concerning setting controllable environmental variables, such as water discharge. For example, in a hypothetical environment where a river diversion is being considered, and the current set of environmental variables yields $RHS = -0.2005$ (which suggests the probability of delta formation is $\pi_{\text{delta}}$...
= 0.45), a 600 m$^3$ s$^{-1}$ increase in $Q_w$ alone will increase the probability of delta formation 8% (from 0.45 to $0.54^{\frac{3}{2}}$) (assuming the increased $Q_w$ has no effect on other variables).

6. Conclusions

Based on analysis of a new data-set comprising 5,399 coastal rivers that are at least 50 m wide, along with eight environmental variables, we find that only 40% (2,174) of coastal rivers have deltas, and these are unevenly distributed geographically, with delta formation being more likely in latitudes -27° to 30° and 60° to 72°. Likelihood of delta formation increases with increasing sediment flux, water discharge, and basin area, whereas likelihood decreases with increasing significant wave height and tidal range and significant wave height. Receiving-basin bathymetry has a bimodal effect on likelihood of delta formation. At slopes less than 0.006, delta formation decreases with increasing slope, but the trend is reversed at slopes greater than 0.006. Recent sea-level change and sediment concentration have no clear effect on delta formation. Finally, we derive a logistic regression that predicts probability of delta formation with an accuracy of 74.5%. Together our results suggest that delta formation is a balance between the constructive forces of, such as water and sediment discharge, and the destructive forces, such as of waves and tides.

7. Acknowledgements

DAE would like to acknowledge funding and support from National Science Foundation grants 1135427, 1426997, and 1812019. Sea-level products were processed by SSALTO/DUACS and distributed by AVISO+ (https://www.aviso.altimetry.fr) with support from CNES. We thank M. Domaracki, I. Thomas, K. Rhodes, A. Whaling, and S. Adams for helping with data collection and organization.

Tables

Table 1. Independent variables: upstream and downstream environmental variables

### UPSTREAM VARIABLES

<table>
<thead>
<tr>
<th>Quantity</th>
<th>Notation (units)</th>
<th># of Rivers with Available Data</th>
<th>Source</th>
</tr>
</thead>
<tbody>
<tr>
<td>Water discharge</td>
<td>$Q_w$ (m$^3$ s$^{-1}$)</td>
<td>943</td>
<td>1</td>
</tr>
<tr>
<td>Sediment discharge</td>
<td>$Q_s$ (m$^3$ s$^{-1}$)</td>
<td>600</td>
<td>1</td>
</tr>
<tr>
<td>Sediment concentration</td>
<td>$Q_s/Q_w$ (–)</td>
<td>571</td>
<td>1</td>
</tr>
<tr>
<td>Drainage basin area</td>
<td>$A_b$ (km$^2$)</td>
<td>1,143</td>
<td>1</td>
</tr>
</tbody>
</table>

### DOWNSTREAM VARIABLES

<table>
<thead>
<tr>
<th>Quantity</th>
<th>Notation</th>
<th># of Rivers with Available Data</th>
<th>Source</th>
</tr>
</thead>
<tbody>
<tr>
<td>Wave height</td>
<td>$H_w$ (m)</td>
<td>5,209</td>
<td>2</td>
</tr>
<tr>
<td>Tidal range</td>
<td>$H_t$ (m)</td>
<td>5,259</td>
<td>3</td>
</tr>
<tr>
<td>Bathymetric slope</td>
<td>$S_{b,r}$ (–)</td>
<td>5,358</td>
<td>4</td>
</tr>
<tr>
<td>Sea level change rate</td>
<td>$H_s$ (mm yr$^{-1}$)</td>
<td>5,172</td>
<td>5</td>
</tr>
</tbody>
</table>

Table 2: Statistical differences between rivers with no deltas and rivers with deltas. Percentages are calculated relative to the total number of rivers with no deltas (3,225) and with deltas (2,174). $D$ is the two-sample Kolmogorov-Smirnov test statistic and is equal to the maximum variance between the cumulative distribution functions of the two populations tested. $p$ is the $p$-value at the 5% significance level. $h$ is the test decision, where 1 rejects the null hypothesis that the distributions are from the same population and 0 accepts the null hypothesis.
<table>
<thead>
<tr>
<th>Variable</th>
<th>All Data</th>
<th>River Mouths (no deltas)</th>
<th>Deltas</th>
<th>KS test results</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>min</td>
<td>max</td>
<td>median</td>
<td>n (%)</td>
</tr>
<tr>
<td>$Q_w$</td>
<td>0.03</td>
<td>$2.1 \times 10^4$</td>
<td>880</td>
<td>95.13</td>
</tr>
<tr>
<td>$Q_i$</td>
<td>$4.8 \times 10^5$</td>
<td>14.37</td>
<td>0.13</td>
<td>0.012</td>
</tr>
<tr>
<td>$Q_i \times C_w$</td>
<td>$9.8 \times 10^6$</td>
<td>0.078</td>
<td>6.7$ \times 10^{-4}$</td>
<td>8.5$ \times 10^{-5}$</td>
</tr>
<tr>
<td>$A_h$</td>
<td>180</td>
<td>$6.3 \times 10^5$</td>
<td>$6.9 \times 10^4$</td>
<td>8011</td>
</tr>
<tr>
<td>$H_w$</td>
<td>0.067</td>
<td>5.18</td>
<td>1.33</td>
<td>1.16</td>
</tr>
<tr>
<td>$H_i$</td>
<td>0.003</td>
<td>15.44</td>
<td>1.74</td>
<td>1.57</td>
</tr>
<tr>
<td>$S_n$</td>
<td>5.0$ \times 10^{-5}$</td>
<td>0.25</td>
<td>0.011</td>
<td>0.0021</td>
</tr>
<tr>
<td>$H_s$</td>
<td>-12.45</td>
<td>11.86</td>
<td>3.41</td>
<td>3.36</td>
</tr>
</tbody>
</table>
Figures

Figure 1. Examples of (a) a river mouth without a delta (Mexico), (b) headless tidal channels not included in this dataset, (c) a delta with land both upstream and downstream of the regional shoreline vector (marked by dashed line), and location of delta node demarcated with blue dot (Godavari River, India), (d) a delta distinguished by a shoreline protrusion only (Red River, Turkey), (e) a delta distinguished by a distributary channel network only. RM locations mark the main river channel mouth (Amazon River, South America).
Figure 2: Global distribution of coastal (a) rivers (includes both river mouths and deltas) and (b) deltas only. Each colored line segment is 3° long. Black (solid and dashed) boxes refer to hotspots of delta formation discussed in the text.
Figure 3. Histograms showing the latitudinal distribution ($3^\circ$ bins) of (a) total number of rivers (white) and number of rivers with deltas (gray), (b) total shoreline length of surveyed coastlines measured from the global shoreline database [Wessel and Smith, 1996], (c) All rivers (including deltas) per shoreline kilometer (white bars), where solid gray line shows rivers with no deltas (river mouths), and solid black line shows rivers with deltas, and (d) Solid black line is the ratio of deltas per river (delta likelihood, $L_d$), and the white bars are total number of rivers (including deltas).
Figure 4. Latitudinal variation of the independent variables used in this study. All panels show the median value for 3° bins. (a) water discharge, $Q_w$; (b) sediment discharge, $Q_s$; (c) sediment concentration, $Q_s/Q_w$; (d) drainage basin area, $A_b$; (e) mean annual significant wave height, $H_w$; (f) median tidal range, $H_t$; (g) bathymetric slope, $S_b$; (h) rate of sea-level change, $H_s$. For a, c, and d the outliers have been cut off for viewing purposes.
Figure 5. Differences in upstream environmental variables for rivers with and without deltas. (top panel) Scatter plots of delta likelihood, defined as number of rivers with a delta relative to total number of rivers in that interval. (bottom panel) Histograms binned into equal log-spaced intervals. Gray boxes outline ranges represented by 1% or less of total sample number.
Figure 6. Differences in downstream environmental variables for rivers with and without deltas. (top panel) Scatter plots of delta likelihood, defined as number of rivers with a delta relative to total number of rivers in that interval. (bottom panel) Histograms binned into equal log-spaced intervals. Gray boxes outline ranges represented by 1% or less of total sample number.

Figure 7. Scatter plot of measured versus predicted delta formation. Equation (1) was used to calculate predicted probability of delta formation, $\pi_{\text{delta}}$, using rivers with necessary data available, $n = 4776$ (2/3 of which was used for training, and 1/3 used for testing/validation). To compare measured to predicted values from equation 1 to our collected data we need to transform the binary observation of delta presence or absence into a continuous variable. To do this we created 20 equal intervals ($L_d = 0.05$ bin widths) and averaged $\pi_{\text{delta}}$ values. $L_d$ is calculated for each bin as the number of rivers with deltas divided by total number of rivers. Dashed line represents a 1:1 relationship.

References


Milliman, J. D., and Syvitski, J. P. M.: Geomorphic/tectonic control of sediment discharge to the ocean; the importance of small mountainous rivers, Journal of Geology, 100, 525-544, 1992.


Twilley, R. R., Bentley, S. J., Chen, Q., Edmonds, D. A., Hagen, S. C., Lam, N. S.-N., Willson, C. S., Xu, K., Braud, D.,
Hampton Peele, R., and McCall, A.: Co-evolution of wetland landscapes, flooding, and human settlement in the Mississippi


Wessel, P., and Smith, W. H.: A global, self-consistent, hierarchical, high-resolution shoreline database, Journal of

Woodroffe, C. D., Nicholls, R. J., Saito, Y., Chen, Z., and Goodbred, S. L.: Landscape variability and the response of Asian

Wright, L. D., Coleman, J. M., and Erickson, M. W.: Analysis of Major River Systems and Their Deltas: Morphologic and
Process Comparisons, Coastal Studies Institute, Louisiana State University, Baton Rouge, LA, 1974.

Wright, L. D.: Sediment transport and deposition at river mouths: A synthesis, Geological Society of America Bulletin, 88,