Where and Why Do Submarine Canyons Remain Connected to the Shore During Sea-level Rise? - Insights from Global Topographic Analysis and Bayesian Regression

Anne Bernhardt[1] and Wolfgang Schwanghart[2]

[1]Freie Universität Berlin
[2]University of Potsdam

November 23, 2022

Abstract

The efficiency of sediment routing from land to the ocean depends on the position of submarine canyon heads with regard to terrestrial sediment sources. We aim to identify the main controls on whether a submarine canyon head remains connected to terrestrial sediment input during Holocene sea-level rise. Globally, we identified 821 canyon heads that are currently located at the -120m depth contour (the Last Glacial Maximum shoreline) and 188 canyon heads that remained shore-connected during present-day highstand. Regional hotspots of present-day shore connected canyons (SCCs) are the Mediterranean active margin and the Pacific coast of Central and South America. We used 34 terrestrial and marine predictor variables to predict SCC occurrence using Bayesian regression. Our analysis suggests that narrow and steep shelves and steep continental slopes precondition the maintenance of canyon-head connectivity to the shore. Moreover, SCCs occur preferentially along active margins characterized by low-erodibility bedrock and high water discharge.

Where and Why Do Submarine Canyons Remain Connected to the Shore During Sea-level Rise?
–
Insights from Global Topographic Analysis and Bayesian Regression

Anne Bernhardt[1], Wolfgang Schwanghart[2]

[1]Institute of Geological Sciences, Freie Universität Berlin, Malteserstrasse 74-100, 12249 Berlin, Germany

[2] Institute of Environmental Sciences and Geography, Potsdam University, Karl-Liebknecht-Str. 24-25, 14476 Potsdam-Golm, Germany

Corresponding author: Anne Bernhardt (anne.bernhardt@fu-berlin.de)

Key Points:

- Presently, 188 submarine canyons are connected to the shoreline along the world's major continents
- Narrow shelves and high shelf and continental slope gradients precondition the maintenance of canyon-head connectivity to the shore
- Canyon heads preferentially remain connected to the shore offshore river catchment with low-erodibility bedrock and high water discharge

Abstract

The efficiency of sediment routing from land to the ocean depends on the position of submarine canyon heads with regard to terrestrial sediment sources. We aim to identify the main controls on whether a submarine canyon head remains connected to terrestrial sediment input during Holocene sea-level rise. Globally, we identified 821 canyon heads that are currently located at the -120m depth contour (the Last Glacial Maximum shoreline) and 188 canyon heads that remained shore-connected during present-day highstand. Regional hotspots of present-day shore connected canyons (SCCs) are the Mediterranean active margin and the Pacific coast of Central and South America. We used 34 terrestrial and marine predictor variables to predict SCC occurrence using Bayesian regression. Our analysis suggests that narrow and steep shelves and steep continental slopes precondition the maintenance of canyon-head connectivity to the shore. Moreover, SCCs occur preferentially along active margins characterized by low-erodibility bedrock and high water discharge.

Plain Language Summary

Since the last glaciation about 20,000 years ago, sea level has risen by about 120 meters. As a consequence, most coastlines have migrated landward, inundating large shelf areas. Some of these areas are now dissected by submarine canyons. However, with only 4% of the world's submarine canyons reaching todays coastline, these canyons remain the exception. Here, we aim to identify the environmental factors and processes that control whether rates of headward canyon incision can keep pace with landward migration of the coastline during the Holocene. We determine 34 variables that potentially predict whether a canyon remains connected to the coastline. We find that shore-connected canyons preferentially occur along continental margins with narrow and steep shelves and steep continental slopes. Such seafloor morphology favors the occurrence of erosive sediment flows. Moreover, our analysis supports the occurrence of such canyons offshore river basins, that are characterized by durable bedrock and high water discharge. Such rivers deliver coarse-grained sediment to submarine canyons, which can erode the canyon head and floor. To this end, our analysis offers new insights into the formation and maintenance of submarine canyons that are required to efficiently transport sediments, pollutants and organic carbon from rivers to the ocean floor.

1 Introduction

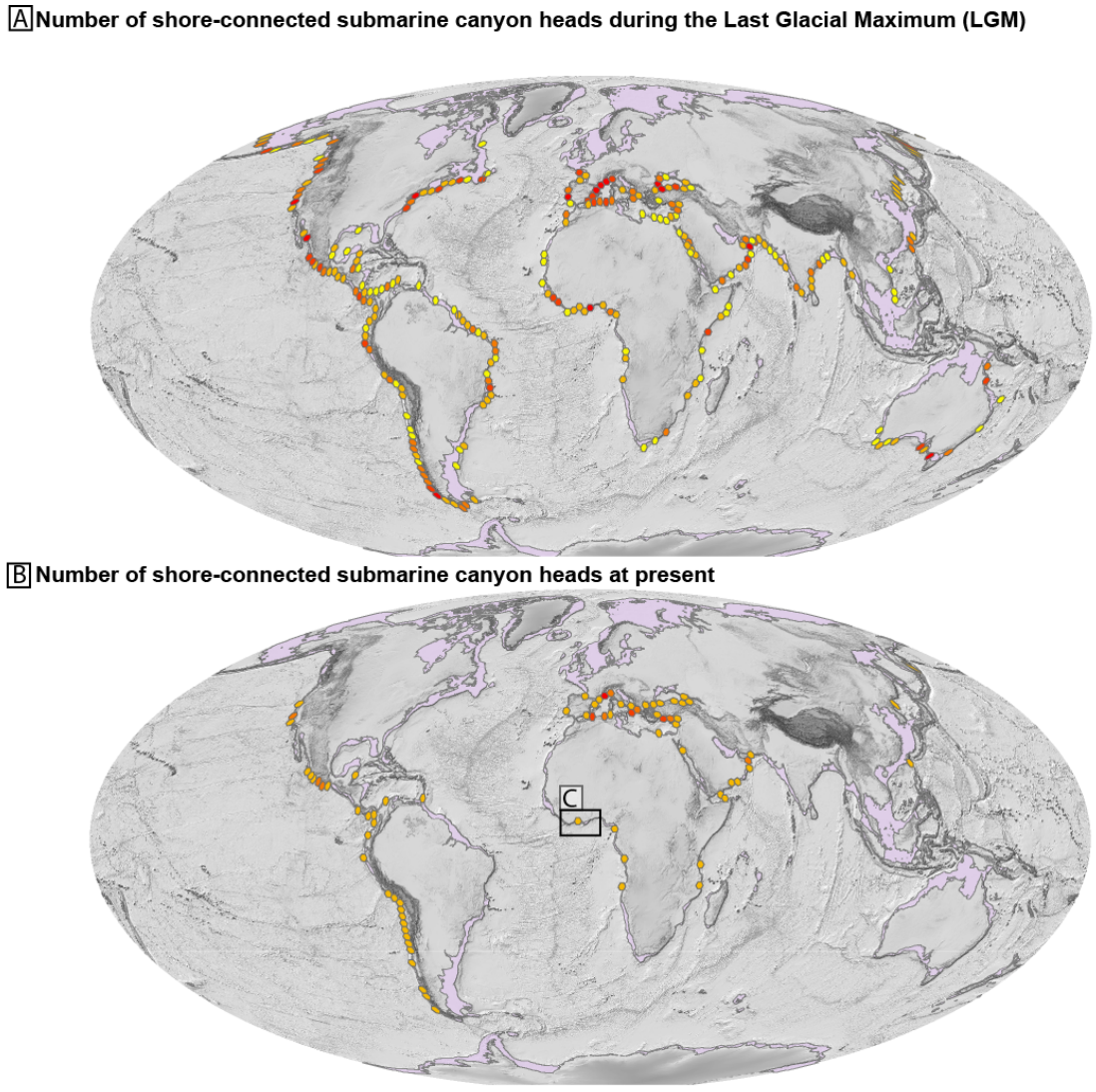
Submarine canyons are prime conduits for sediment-laden flows that link terrestrial sediment sources with deep-marine depocenters. The efficiency at which canyons route sediments is at least in part controlled by how far they extend into the shelf. If the distance between the canyon head and the shore is short, terrestrial sediment, associated pollutants and organic carbon is efficiently delivered to the deep ocean (Covault et al., 2007; Dominguez-Carrió et al., 2020; Galy et al., 2007; Kane and Clare, 2019). We refer to such canyons as shore-connected canyons (SCCs) hereafter.

Whether and where SCCs occur relates to the canyons' ability to erode headward at a pace that keeps up with rates and magnitudes of millennial-scale sea-level rise (e.g., after the Last Glacial Maximum (LGM)) (e.g., Mauffrey et al., 2017). It is remarkable that ~30% of submarine canyons are incised into the shelf but only few canyons are connected to the present-day shoreline (Harris and Whiteway, 2011). However, the detailed controls on why a submarine canyon occurs at a specific site, why it incised into the shelf, or why it remained connected to terrestrial sediment supply during rising sea level are poorly understood (Harris and Whiteway, 2011; Smith et al., 2017; 2018). Several, possibly related factors (Shepard, 1981), were proposed to control submarine canyon occurrence and shelf incision, including narrow continental shelves along active margins (Normark et al., 2009), high shelf gradient (Sweet and Blum, 2016), high sediment flux from onshore catchments (Harris and Whiteway, 2011; Pratson et al., 2007), mass wasting along steep continental slopes (Pratson and Coakley, 1996; Pratson et al., 1994), and submarine groundwater seepage (Pratson et al., 2007).

Most recent work suggests that submarine canyon heads preferentially remain shore-connected if neighboring onshore catchments experience high uplift rates and expose durable bedrock (Smith et al., 2017). Based on a global compilation of submarine canyons classified into 'shelf-incising' and 'blind' (slope-confined), Harris and Whiteway (2011) showed that shelf-incising canyons

prevail along the western, tectonically active margins of the Americas that are characterized by high sediment supply. Smith et al. (2017) focused on the West Coast of the United States and found no correlation between canyon occurrence and shelf gradient and width. Instead, their analysis underscored the role of coarse sediment from durable terrestrial bedrock that controls offshore canyon-head incision, possibly modulated by wave focusing of canyon bathymetry (Smith et al., 2018). A key finding of these studies is that onshore processes and lithological composition are crucial to understanding rates of headward erosion of submarine canyons. Yet, whether these findings pertain to the global patterns of canyons has not been investigated so far.

Here, we study global patterns of SCCs with the aim to identify the main controls on their occurrence. Our analysis is driven by two hypotheses: First, we hypothesize that submarine canyon heads remain connected to the shore upon postglacial sea-level rise if the shelf is narrow and steep and the horizontal distance that the canyon heads need to migrate is low. Second, we test the hypothesis of Smith et al. (2017, 2018) that submarine canyon heads remain preferentially connected to the shoreline when located offshore tectonically uplifting regions which are characterized by durable terrestrial bedrock. We test these hypotheses by globally categorizing canyon heads as present-day SCC heads and canyon heads close to the -120m contour. We assume that the latter canyons were connected to the shore during the LGM, but did not incise further into the shelf during Holocene sea-level rise. We then predict present-day SCCs using Bayesian penalized regression.


A Number of shore-connected submarine canyon heads during the Last Glacial Maximum (LGM)
B Number of shore-connected submarine canyon heads at present
C
Number of canyon-head type per hexagon (size: 50,000 km²)
1
2 - 3
4 - 6
7-9
10-20
shelf (Harris et al., 2014)

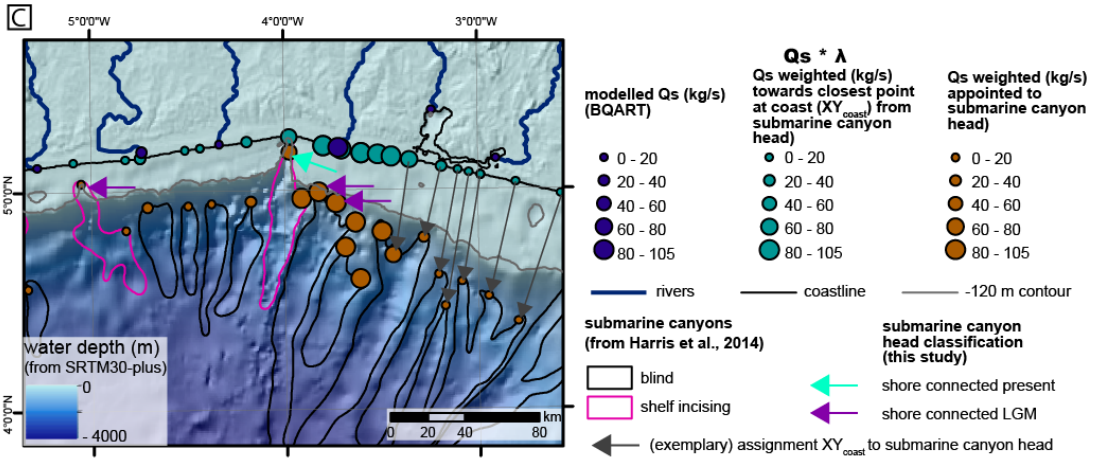

C
5°0'0"W
4°0'0"W
3°0'0"W
5°0'0"N
4°0'0"N
water depth (m)
(from SRTM30-plus)
0
- 4000
0 20 40 80
km
modelled Qs (kg/s) (BQART)
Qs * λ
Qs weighted (kg/s) towards closest point at coast (XY$_{coast}$) from submarine canyon head)
Qs weighted (kg/s) appointed to submarine canyon head)
0 - 20
20 - 40
40 - 60
60 - 80
80 - 105
rivers
coastline
-120 m contour
submarine canyons (from Harris et al., 2014)
blind
shelf incising
submarine canyon head classification (this study)
shore connected present
shore connected LGM
(exemplary) assignment XY$_{coast}$ to submarine canyon head

Fig. 1. Global overview of a) the number of canyon heads close to the -120 contour per hexagon (50,000 km^2) and b) present day SCCs. c) Illustration of the weighting calculation performed using suspended sediment load (Qs) (BQART; Kettner and Syvitski, 2011) offshore the Ivory Coast. Most canyons are ' blind' canyons (Harris et al., 2014) and two canyons are shelf-incising. Note that one canyon shapefile can have several canyon heads. The canyon heads here are classified as shore-connected during the LGM and only the head of the 'Trou sans Fond' submarine canyon (Dietz and Knebel, 1971) presently connects to the shore.

2 Methods and Data

Our study relies on the Shuttle Radar Topography Mission (SRTM30_PLUS) 30-arc second database (~1km resolution) (Becker et al., 2009). We excluded islands and oceanic plateaus and limited our analysis to 50°N and 50°S where most predictor variables (Table 1) are completely available.

2.1 Submarine canyon variables

We used the dataset from Harris et al. (2014) who delineated canyons that extend over a depth range of at least 1000m and are incised at least 100m into the slope. We manually mapped canyon heads and assigned them to the submarine canyon polygon of Harris et al. (2014). We computed the shortest Euclidean distance from each canyon head to the present-day shoreline and to the -120m contour (LGM shoreline). Canyon heads that are located <6km away from either shoreline were classified as 'shore-connected LGM' or 'shore-connected present' (Fig. 1a,b). We chose this distance to be two cells larger than the offshore limit of longshore sediment transport (up to 5 km; Sweet and Blum, 2016), due to the low resolution of the DEM and uncertainties in the mapping process. We manually corrected canyon-head misclassifications during visual inspection or refined interpretation of canyon-head locations from Harris et al. (2014). Although some canyons (Swatch of No Ground, 150 km, Rogers et al., 2015; Indus canyon, 17km, Li et al., 2018) are known to presently receive terrestrial sediment through clinoform progradation, we did not classify them as SCC canyons.

2.2 Terrestrial variables

Topographic analysis was conducted with ArcGIS, MATLAB and TopoToolbox (Schwanghart and Scherler, 2014). Terrestrial drainage-basin statistics (elevation, gradient, area, and river steepness index (ksn)), and river-outlet locations were determined using the 500m-grid and flow directions of the HydroSHEDS compilation (Lehner et al., 2008) (Table 1).

We obtained estimates of water discharge (Qw) at each river outlet by integrating annual runoff from the Global Runoff Data Centre (Fekete et al., 2000). We used the BQART model for pre-human suspended sediment flux (Qs) (Syvitski and Kettner; 2011). Bedload was estimated by two empirical equations from Qs (Table 1). We extracted a mean, area-weighted erodibility index for each drainage basin using the global erodibility index (GEroID) of Moosdorf et al. (2018) which ranges from low erodibility (=0.8) for metamorphic rocks to high erodibility (=3.2) for unconsolidated sediments. Area-weighted means of annual rainfall and its distribution were calculated using Tropical Rainfall Measurement Mission data (TRMM, courtesy of B. Bookhagen and Boers et al. (2014)) and assigned to each river outlet. Finally, we assigned peak-ground accelerations (PGA) to each river outlet using the Global Seismic Hazard Map (Shedlock et al., 2003).

2.1.1 Weighting of terrestrial variables to each submarine canyon head

To assign terrestrial variables to submarine canyon heads, we first determined the closest point on the shoreline (XY_{coast}) for each canyon head (Fig. 1c), and then computed the distance between XY_{coast} and river outlets on the adjacent continent. These distances d_i together with the catchment areas A_i of each river outlet i subsequently served as weights in a distance-weighted averaging approach. Specifically, we calculated the weights λ_i as:

$$\lambda_i = \frac{A_i}{d_i^3} \Big/ \left(\sum_i \frac{A_i}{d_i^3} \right) \qquad \text{Eq. 1,}$$

and assigned the weighted averaged variable to the corresponding canyon head. Our weighting scheme reflects that a one-to-one assignment of river outlets to canyon heads is often infeasible and accounts for the increased importance of outlets that are nearby and have large contributing catchments. Moreover, our approach exonerates us from choosing an arbitrary number of river-drainage basins that may shed sediment into a particular submarine canyon. We weighted the variables towards XY_{coast} and not to the canyon head itself to avoid incorporating the canyon-head to shore distance into the predictor variables, as this is implicitly what we are aiming to predict.

Predictor variable	Unit	Description & Computation	Data Source	Abbreviation in table	Weighting factor λ
terrestrial predictors					
distance to the nearest river mouth - present day	km	distance to the nearest river mouth - present day	(SRTM30_PLUS) (Becker et al., 2009) HydroSHEDS (Lehner et al., 2008)	d_weighted	$\lambda_i = \frac{A_i}{d_i^3} / (\Sigma_i \frac{A_i}{d_i^3})$
max. & mean elevation, area of onshore catchment (present day)	m and km2	upslopestats fuction of TopoToolbox (Matlab)	see above	elev_max_weighted elev_mean_weighted area_weighted	see above
mean river steepness index in onshore catchment (present day)		ksn and upstreammean function of TopoToolbox,	see above	ksn_weighted	see above
gradient in onshore catchment (present day)	unitless	gradient function of TopoToolbox	see above	grad_weighted	see above
discharge of onshore catchment (present day)		flowacc function of TopoToolbox		Qw_weighted	see above
modeled suspended sediment flux (pre-human) Qs	kg/s	Qs at the output of river mouths was assigned to theh HydroSHEDS stream network using the knnsearchlatlon & matchpairs function of TopoToolbox/ Matlab	Qs values of Syvitski and Kettner, 2011 (BQART)	Qs_weighted	see above
bedload estimated from empirical equation of Bagnold (pre-human)	kg/s	$Q_{bed} = 0.8 \times Q_s^{0.5926}$	Bagnold, 1966, Syvitski & Saito, 2007, Qs values of Syvitski and	QbBagnold_weighted	see above
bedload estimated from empirical equation of Metivier (pre-human)	kg/s	$Q_{bed} = aQ_s^b$ for $Q_s \leq (\frac{a}{c})^{1/(d-b)}$ $Q_{bed} = cQ_s^d$ a=0.833^(1-b); b=1.3240;c=0.437^(1-d); d=0.647;	Metivier et al., 2004; Meunier et al., 2006; Qs values of Syvitski and Kettner, 2011	QbMetivier_weighted	see above
peak ground acceleration	m/s2	peak ground acceleration from the Global Seismic Hazard Map of the Global Seismic Hazard Assessment Program (GSHAP) which depicts PGA with a 10% probability of exceedance in 50 years- was assigned to each river outlet using the flowacc function of TopoToolbox	Shedlock et al., 2003	GSHAP_weighted	see above
weighted global erodibility index (GEroID)	unitless	GEroID ranges from 1.0 (low erodibility) to 3.2 (high erodibility). Area weighted mean of GEroID for each drainage basin was assigned to each river outlet using the upslopestats function of the TopoToolbox.	global erodibility index dataset (GEroID) of Moosdorf et al., (2018)	ero_index	see above
mean annual rainfall	mm/day	Area weighted mean of annual rainfall for each drainage basin was assigned to each river outlet using the upslopestats function of the TopoToolbox.	courtesy of B. Bookhagen	'TRMM_mean_weighted'	see above
yearly standard deviation of the daily rainfall -	mm/day	Area weighted mean of the standard deviation of the mean annual rainfall for each drainage basin was assigned to each river outlet using the upslopestats function of the TopoToolbox.	courtesy of B. Bookhagen	'TRMM_STDV_weighted'	see above
The ratio of the 90th to 50th percentile of the mean annual rainfall.	unitless	The ratio of the 90th to 50th percentile of the mean annual rainfall. This is a measure for how extreme the rainfall is. High values indicate extreme rainfall events. Area weighted mean of the 90th/50th percentile of the mean annual rainfall for each drainage basin was assigned to each river outlet using the upslopestats function of TopoToolbox.	Boers et al., 2014	'TRMM_90_50_weighted'	see above

marine predictors					
gradient of the adjacent shelf	unitless	computed outline of shelf shape file (Harris et al., 2014), Laplace interpolation between slope boundaries to create a smoothed shelf without submarine canyons. Computed mean gradient of the smoothed continental slope surrounding each canyon head in a circular shaped polygon with a radius of 10 km	SRTM30_PLUS Becker et al., 2009 & Harris et al., 2014	shelf_gradient'	no weighting
gradient of the adjacent continental slope	unitless	analogous to shelf_gradient but using a circular shaped polygon with a radius of 80 km	see above	'csgradient'	no weighting
max. & mean shelf width (present day)	km	determined the 100 nearest-neighbor DEM pixels (86 km) along the outer shelf boundary from each submarine canyon head (using the *bwdist* and *knnsearch* function in Matlab) and calculated the mean shelf width from these . Where shelf-incising canyons reduced shelf width, results were corrected after visual inspection.	Harris et al., 2014	dshelf_mean' 'dshelf_max'	no weighting
max. & mean shelf width (LGM)	km	identical calculation to the present day shelf width, but based on the SRTM30_PLUS, where 120 m were added to the elevation data.	see above	dshelf_mean_LGM' 'dshelf_max_LGM'	no weighting
max. & mean depth of the shelf edge	m	water depth of shelf edge was calculated using mapped shelf outline and SRTM_30plus data. Shelf edge depth was assigned to each canyon head using from the 20 nearest neighbor DEM pixels (17 km) using the the knnsearch function in a 20 cell window	SRTM30_PLUS Becker et al., 2009 & Harris et al., 2014	shelf_edge_z'	no weighting
max. & mean storm surge height	m	determined the 20 nearest-neighbor DEM pixels from the GTSR dataset along the coast from each submarine canyon head (using the *knnsearch* function in Matlab) and calculated the mean storm surge height from these 20 pixels.	Muis et al., 2016	mean_storm_surge	no weighting
wave height & wave period	m	wave height & period from Sentinel 2 data	Bergsma and Almar, 2020	wave_height_m' 'wave_period_m'	no weighting
depth of closure	m	depth of closure along the adjacent coastline	Bergsma and Almar, 2020	depth_of_closure_m'	no weighting
submarine groundwater discharge predictors					
For all parameters from the modeled submarine groundwater discharge (SGD) parameters from Luijendijk et al. (2020): To assign a weighted submarine groundwater discharge value SGD, we chose the centroid of each coastal watershed polygon provided by the authors and performed the weighting to the nearest point on the coast with regard to the canyon head					
modeled fresh submarine groundwater discharge (SGD)	m2/yr		Luijendijk et al. 2020	fsgd_best_weighted'	$\lambda_i = \frac{1}{d_i^3} / (\Sigma_i \frac{1}{d_i^3})$
modeled near-shore terrestrial discharge (NGD)	m2/yr		Luijendijk et al. 2020	'ngd_best_weighted'	see above
modeled total coastal groundwater discharge (CGD)	m2/yr	CGD=NGD+SGD	Luijendijk et al. 2020	cgd_best_weighted'	see above

* weighting has been performed as follows: we calculated the nearest point (Euclidean distance) on the shore for each submarine canyon head.

We weighted all weighted parameters by a weigthing factor lambda.

Table 1. Summary of predictor variables.

2.3 Submarine groundwater discharge

Luijendijk et al. (2020) simulated fresh submarine groundwater discharge (SGD) in coastal systems using a spatially resolved numerical model. To assign weighted SGDs to canyon heads, we computed the centroid of each watershed, assigned the SGD to that centroid and weighted the value by its distance (d) to each XY_{coast} of each canyon head using the weighting factor λ:

$$\lambda_j = \frac{1}{d_j^3} \Big/ \left(\sum_i \frac{1}{d_j^3}\right) \qquad \text{Eq. 2.}$$

2.4 Marine variables

To acquire the mean gradient of the continental slope in the vicinity of each canyon head, we extracted the outlines of the slope shapefile of Harris et al. (2014). We set DEM values within the extent of canyons to NoData and used a Laplacian interpolation to smoothly interpolate inward from these outlines. The technique is referred to as image inpainting (Stolle et al., 2019) and reconstructs a continental slope devoid of canyons. Analogously, we calculated the shelf gradient adjacent to each canyon head using the shelf shapefile of Harris et al. (2014). To calculate shelf width at each canyon head, we extracted the DEM cells at the oceanward shelf boundary and calculated the shortest Euclidian distance to the shoreline. We chose the 100 nearest-neighbor pixels (~86km) along the oceanward shelf boundary and calculated the mean shelf width from these pixels. We used a large number of nearest-neighbors to minimize the impact of canyon-head indentation into the shelf (Fig. 1c). Results were inspected visually and corrected where wide canyons artificially decreased values of shelf width. For the LGM shelf width, the same calculation

was based on a DEM where 120m of elevation where added to each cell to simulate the LGM landscape. Shelf-edge depth was calculated by extracting the water depth from the 20 nearest-neighbor cells (~17km) of the outer shelf boundary. We assigned storm-surge heights of 1-in-100-year extreme sea levels of Muis et al. (2016) to each canyon head (mean of 20 nearest-neighbor cells). Bergsma and Almar (2020) extracted global wave heights and periods from ESA's Sentinel2 constellation and calculated the depth of closure (maximum water depth of littoral sediment transport of Hallermeier, 1980). We extracted wave height, period and depths of closure for each XY_{coast} and assigned these to adjacent canyon heads.

2.5 Predictive Modeling – Bayesian penalized regression

The aim of this study is to identify the controls on continued shelf incision and maintenance of canyon-head-to-shore connectivity from the LGM to present. We computed a hexagonal net (50,000km^2/hexagon) and computed the number of present-day SCC-heads in each tile (Fig. 1b). This is the number we predict (the 'response'), using a 34 predictor variables ('predictors', Table 1). To extract the weighted predictors for each tile, we computed the hexagon midpoints and their corresponding nearest location at the coast ($XY_{coast\_hexgrid}$) and weighted the XY_{coast}-weighted predictors for individual canyon heads onto this coastal location using inverse-distance weighting (Eq. 2, Table 2).

To identify the most important predictors and to globally predict the number of present-day SCC heads, we employed Bayesian penalized regression. Bayesian statistics apply probabilities to statistical problems offering a way to learn from new data to update prior beliefs while accounting for uncertainties (e.g., Efron, 2013; Korup, 2020). A frequentist approach to penalization is lasso regression which uses a penalty term to shrink small regression coefficients to zero (hence reducing or eliminating the predictor from the model) (Tibshirani, 2011). In Bayesian penalized

regression, penalization is incorporated through the choice of prior distribution of the parameters (e.g., van Erp et al., 2019). We used *bayesreg*, a MATLAB toolbox for fitting Bayesian penalized regression models (Makalic and Schmidt, 2016). All predictors were centered and scaled. As we predict counts of present-day SCCs per hexagon, we chose a Poisson distribution for the response. Our choice of shrinkage prior followed the procedure of van Erp (2020) (see Supporting Information for prior sensitivity analyses, Figs. S3-S6). All priors result in similar prediction root-mean-square errors (RMSE=0.93-0.95), Watanabe–Akaike information criteria of 220-231, pseudo R^2 of 0.5, and posterior distributions of the regression coefficients (Figs. S4-S5). Based on the lowest RMSE and the stable sampling performance (Table S1-5, Fig. S3), we show the results of the lasso-shrinkage prior. Finally, we quantified the importance of each predictor adopting the Bayesian feature-ranking algorithm of Makalic and Schmidt (2011). The rank corresponds to the strength of the association between the predictors and the response where lower ranks denote more important predictors (Fig. 3d, Table S1-S5). The final rank of the predictor is based on the 75th percentile of the complete set of rankings for each posterior sample. Complete model parameters are listed in the Supporting Information.

3 Results

3.1 Present-day, shore-connected canyon occurrence

Our data set comprises 4717 canyon heads, of which 2988 are classified as blind canyons and 1729 as shelf-incising. From the latter type, 821 were classified as LGM SCCs and 188 as present-day SCCs. During the LGM, SCCs are distributed globally along passive and active margins (Fig. 1a). In contrast, during today's sea-level highstand, most SCCs straddle along active margins (n=114, Fig. 2a) with spatial hotspots along the Mediterranean active margin, and the Pacific coast of central South America and Central America (Fig. 1b). Moreover, present-day

SCCs occur frequently along the Californian coast, the Indian-Ocean coast of the Arabic Peninsula and the East coast of the Black Sea. Isolated SCCs occur along the coasts of Africa (Fig. 1b).

Figure 2b shows the number of present-day SCCs per hexagon plotted against the 14 highest-rank predictors (Fig. 3d). LGM SCCs occur at shelf widths from <2km to 400km (Fig. S2), whereas present-day SCCs occupy narrow shelf widths from <2km to 31km (Fig. 2b). Only eight present-day SCCs occur at shelf widths between 20km to 31km, 33 at shelf widths between 10km to 20km, and the majority occurs at shelves <10km wide. SCCs occur preferentially where the difference between the present-day and LGM shelf width was minimal (<27km) and become most abundant at differences <9km. One exception is the Congo canyon which occurs at a 52km-wide shelf (Fig. S2). Present-day SCCs occur along a wide range of erodibility indices but are absent at highly erodible catchment lithologies (GEroID>2). The maximum number of present-day SCCs (n=20) is located offshore southern France.

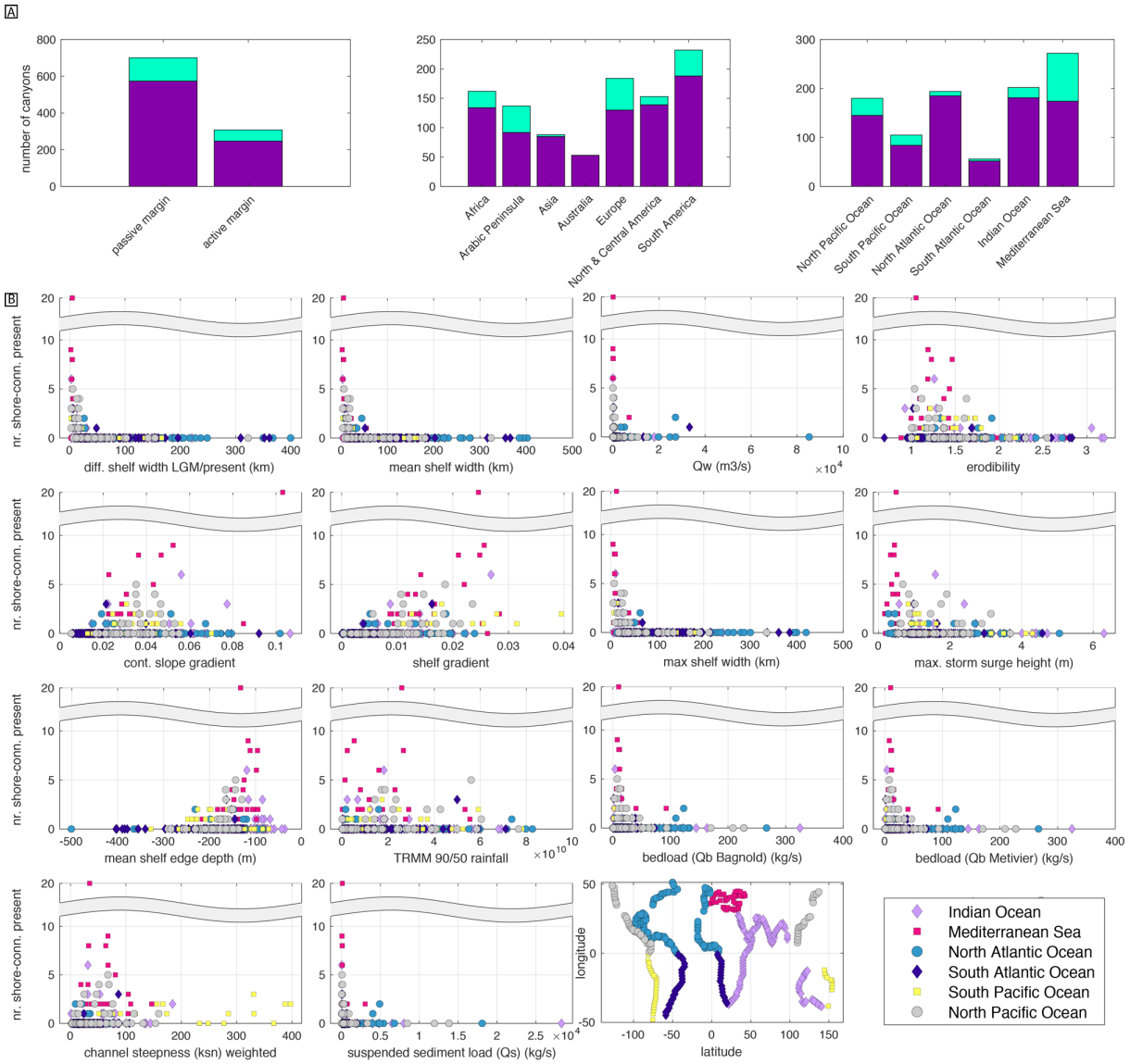


Figure 2. a) Distribution of canyon-head types along margin types, continents, and oceans. b) Predictors plotted against the number of present-day SCCs per hexagon. Data points with fraction=0 represent hexagons that contain canyon heads close to the -120m contour, but no present-day SCCs.

3.2 Prediction of present-day shore-connected canyons (SCCs)

Predictions from Bayesian penalized regression are largely consistent with the regional hotspots of present-day SCCs along the Pacific coast of central South America and Central America, and the Mediterranean active margin (Fig. 3a, b). SCCs along the eastern Black Sea and some along the Indian-Ocean coast of the Arabic Peninsula are also predicted. However, the frequency of SCCs along the Californian coast and the African passive margin (e.g., Congo

canyon) are underestimated (Fig. 3b, c). Individual numbers are predicted with a RMSE of 0.93 and maximum and minimum raw residuals of 7.4 and -4.9, respectively (Figs. 3c, S6). Hence, while the model reliably predicts the spatial patterns of present-day SCC hotspots, predictions of SCC counts are subject to high uncertainties.

Figure 3d shows the posterior distributions of the regression coefficients of the 14 top-ranked predictors (Table S1). Variable ranking reveals the LGM to present-day difference in shelf width and present-day shelf width as the top two predictors (Table 1). The erodibility index of the onshore catchments, the continental slope gradient, and the shelf gradient (rank 4-6) are the only predictors whose posterior distributions do not include zero in their 95% credible interval (regression coefficients of zero indicate no predictive value) (Fig. 3d, Table S1). In summary, shelf and slope morphologic predictors dominate the top ranks and the 14 most important predictors. Two climatic parameters (Qw and TRMM 90/50 percentile) are ranked 3rd and 10th and river sediment bedload (QbBagnold) ranks eleventh. The only onshore topographic parameter amongst the 14 most important predictors is river channel steepness (ksn) on rank 13.

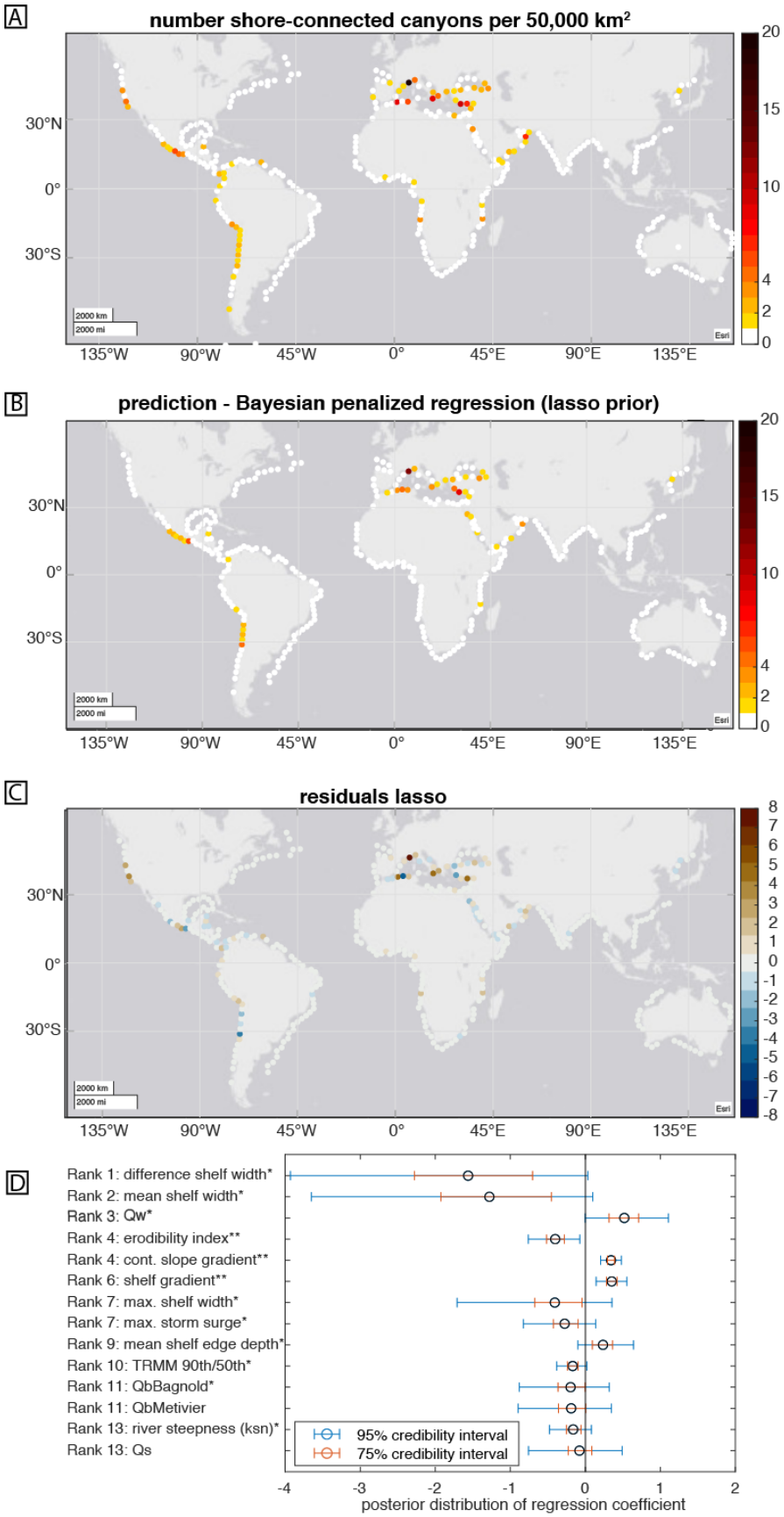

A
number shore-connected canyons per 50,000 km²
30°N
0°
30°S
135°W
90°W
45°W
0°
45°E
90°E
135°E
20
15
10
4
2
0
B
prediction - Bayesian penalized regression (lasso prior)
C
residuals lasso
8
7
6
5
4
3
2
1
0
-1
-2
-3
-4
-5
-6
-7
-8
D
Rank 1: difference shelf width*
Rank 2: mean shelf width*
Rank 3: Qw*
Rank 4: erodibility index**
Rank 4: cont. slope gradient**
Rank 6: shelf gradient**
Rank 7: max. shelf width*
Rank 7: max. storm surge*
Rank 9: mean shelf edge depth*
Rank 10: TRMM 90th/50th*
Rank 11: QbBagnold*
Rank 11: QbMetivier
Rank 13: river steepness (ksn)*
Rank 13: Qs
95% credibility interval
75% credibility interval
-4
-3
-2
-1
0
1
2
posterior distribution of regression coefficient

Figure 3. a) Global maps showing the numbers of present-day SCC heads per tile, b) the predicted of numbers of SCC heads using Bayesian lasso penalized regression, c) the residuals, d) shows the 95% (75%) credibility intervals of the posterior samples of the regression coefficients and the ranking of the 14 most important predictors. The first (second) asterisk is shown when the 75% (95%) credible interval for the corresponding predictor does not include 0.

4 Discussion

4. 1 Limitations of the data set

Based on Bayesian reasoning and guarded against overfitting, we have identified a number of credible predictors for SSC occurrence along the world's coasts. However, we have not explored the fact that our predictors themselves – and the response – are prone to uncertainties. That our results are robust against the choice of prior distributions (and thus different degrees of uncertainty that we assign to each parameter) suggests that our inferences are not strongly affected by these uncertainties. Notwithstanding, there are other uncertainties that are hardly quantified and included in the error analysis. For example, the onset of canyon incision has rarely been dated but can date back to several million years with a complicated evolution of repeated episodes of erosion and infilling (Maier et al., 2018; Mauffrey et al., 2017). Here, we implicitly assume that the present-day SCCs eroded back to the shore during the Holocene sea-level rise. However, older canyons may have more complicated histories. In addition, our predictors largely represent modern conditions, some of which may not represent active phases of canyon-head incision (e.g., Qs, rainfall, groundwater discharge, wave height). Other variables such as bedload are difficult to determine, in particular on longer timescales (e.g., Nitsche et al., 2011). The degree to which these uncertainties may overshadow the importance of bedload transport to the shelf remains shrouded.

Moreover, canyons may incise along tectonic faults, reoccupy fluvial valleys on the shelf (Maier et al., 2018; Mauffrey et al., 2017), or preferentially incise along shelves built by erodible stratigraphy. Additionally, erosive sediment-gravity currents can be triggered by seasonal

downward flow of dense shelfwater (Canals et al., 2006), possibly fostering shoreward canyon-head migration. These predictors were not included as they are unavailable for the spatial scale and extent of this study.

Headward canyon erosion may be additionally enhanced by wave-driven scouring (Smith et al., 2018). Regionally variable details such as the orientation of wave crests and the direction of shore-parallel (littoral) sediment transport with regard to the canyon orientation could not be considered here. While we included the (water) depth of closure in the analysis, we note that the width of the ocean-littoral cell is more appropriate to characterize along-coast wave-induced sediment transport. We would need to combine the closure depth with high-resolution coastal bathymetry which is unavailable today (Bergsma et al., 2019). We included fresh SGD as a predictor, however, submarine canyon formation may also be related to seepage of recirculated seawater (Pratson et al., 2007), which is not quantified at a global scale. Lastly, our analysis assumes that canyons are distributed independently from each other. However, canyon presence can influence the fluid escape and thus the hydrology of the neighboring canyon (Orange et al., 1994). Such coupling mechanisms induce spatial autocorrelation (e.g. clustering) in canyons occurrence which ultimately may bias our modelling results.

Our Bayesian approach to modelling SCC occurrence embraces the idea that uncertainties about model parameters depend on the availability of data, and that their posterior distributions can be updated once new data becomes available.

4.2 Controls on SCC occurrence

Our analysis supports our first hypothesis that submarine canyons remain connected to the shoreline where the shelf is narrow and steep (Figs. 2b, 3d). Shelf-morphologic parameters are

consistently the most 'important' predictors in the Bayesian regression models (Fig. 3d) along with the continental slope gradient. Shoreward canyon incision during sea-level rise occurs by mass wasting or erosive sediment-gravity flows. Steep continental slopes and shelves facilitate shoreward erosion of the canyon head as sediment-gravity flows achieve increased velocities and bottom shear stresses (e.g., Middleton, 1966), and hence, greater degrees of erosivity and canyon-floor downcutting. This, in turn, can cause oversteepening of canyon walls that results in mass wasting and upslope migration of such failure (Densmore et al., 1997; Pratson and Coakley, 1996) or of canyon-thalweg knickpoints supporting backward canyon-head erosion (Guiastrennec-Faugas et al., 2020). Shelf-sediment failure in the canyon head can initiate turbidity currents, which further erode the canyon head and thalweg (Pratson and Coakley, 1996).

Our global analysis also supports the hypothesis of Smith et al. (2018) that present-day SCCs occur preferentially offshore high-relief or tectonically active regions underlain by durable bedrock. The erodibility index holds rank 4 and virtually no present-day SCCs occur at erodibility indices<2 (Figs. 2b, 3). Low-erodibility lithologies in catchments with high water discharge (rank 3) that form bedload in rivers (rank 11) and that are delivered to canyon heads may promote incision in the underlying shelfal bedrock (Cook et al., 2013; Smith et al., 2017), fostering canyon-wall oversteepening and retrogressive failure.

The three SCC hotspots are located along tectonically active margins. However, topographic predictors indicating high onshore relief (channel steepness (ksn), catchment elevation and gradient) are of minor importance (ranks 13, 18, 31, Table S1).

5 Conclusions

Canyon heads close to the -120m contour, the shoreline during the Last Glacial Maximum, are globally abundant (n=821, between 50°N-50°S, along major continents, excluding islands).

Presently, there are only 188 shore-connected canyons (SCCs), most of which belong to three spatial hotspots: the Mediterranean active margin and the Pacific coasts of central South and Central America. SCCs prevail along margins were the shelf is narrow and steep and the difference between the LGM and the present-day shelf width is low. Using Bayesian penalized regression, we predict the spatial patterns of these hotspots using a subset of predictors. The list of the most important predictors is dominated by shelf morphological parameters followed by the continental slope gradient. Hence, on a global scale, low shelf width and high shelf and continental slope gradients precondition the maintenance of canyon-head connectivity to the shore. Narrow and steep shelves minimize the distance that a canyon head has to erode towards the shore over a single sea-level cycle. Moreover, our analysis supports the findings of Smith et al. (2017; 2018) in which the presence of durable bedrock in onshore catchments favors submarine canyon incision. In combination with high water discharge these catchments deliver coarse-grained bedload, which erodes the canyon head and floor.

Large, isolated SCCs (e.g., Congo canyon) are not predicted by our global regression model. These canyons may occur due to conditions that could not be considered in our global assessment, such as the reoccupation of shelf-incised fluvial channels and/or underlying faults. Although we are able to predict regional patterns of SCC occurrence, we are not able to reliably predict individual numbers of SCCs. Hence, we here identify the major global controls on the maintained shore connection of canyon heads but detailed causes of individual SCC occurrence seem to be highly variable from canyon to canyon.

Acknowledgments

We thank Albert Kettner and Jaia Syvitski for providing BQART predictions; Erwin Bergsma for closure depths, wave heights and periods; Elco Luijendijk for submarine groundwater

discharge data and Nils Moosdorf for erodibility indices. Canyon and predictor data are available in the Supporting Information and will be uploaded to the GFZ Data Services.

References

Becker, J.J., Sandwell, D.T., Smith, W.H.F., Braud, J., Binder, B., Depner, J., Fabre, D., Factor, J., Ingalls, S., Kim, S.H., Ladner, R., Marks, K., Nelson, S., Pharaoh, A., Trimmer, R., von Rosenberg, J., Wallace, G., Weatherall, P., 2009. Global Bathymetry and Elevation Data at 30 Arc Seconds Resolution: SRTM30_PLUS. Mar. Geod. 32, 355–371.

Bergsma, E.W.J., Almar, R., 2020. Coastal coverage of ESA' Sentinel 2 mission. Adv. Sp. Res. 65, 2636–2644.

Bergsma, E.W.J., Almar, R., Maisongrande, P., 2019. Radon-augmented Sentinel-2 satellite imagery to derive wave-patterns and regional bathymetry. Remote Sens. 11, 1–16.

Canals, M., Puig, P., de Madron, X.D., Heussner, S., Palanques, A., Fabres, J., 2006. Flushing submarine canyons. Nature 444, 354–7.

Cook, K.L., Turowski, J.M., Hovius, N., 2013. A demonstration of the importance of bedload transport for fluvial bedrock erosion and knickpoint propagation. Earth Surf. Process. Landforms 38, 683–695.

Covault, J.A., Normark, W.R., Romans, B.W., Graham, S.A., 2007. Highstand fans in the California borderland: The overlooked deep-water depositional systems. Geology 35, 783–786.

Densmore, A., Anderson, R., McAdoo, B., Ellis, M., 1997. Hillslope evolution by bedrock landslides. Science (80-.). 275, 369–372.

Dietz, R.S., Knebel, Ha.J., 1971. Trou sans Fond submarine canyon. Deep Sea Res. 18, 441–447.

Dominguez-Carrió, C., Sanchez-Vidal, A., Estournel, C., Corbera, G., Riera, J.L., Orejas, C., Canals, M., Gili, J.M., 2020. Seafloor litter sorting in different domains of Cap de Creus continental shelf and submarine canyon (NW Mediterranean Sea). Mar. Pollut. Bull. 161.

Efron, B., 2013. Bayes' Theorem in the 21st Century. Science (80-.). 340, 1223–1226.

Fekete, B.M., Vörösmarty, C.J., Grabs, W., 2000. Global, composite runoff fields based on observed river discharge and simulated water balances (UNH/GRDC composite runoff fields V1.0), [online]. http://www.grdc.sr.unh.edu/, 2000.

Galy, V., France-Lanord, C., Beyssac, O., Faure, P., Kudrass, H., Palhol, F., 2007. Efficient organic carbon burial in the Bengal fan sustained by the Himalayan erosional system. Nature 450, 407–10.

Guiastrennec-Faugas, L., Gillet, H., Silva Jacinto, R., Dennielou, B., Hanquiez, V., Schmidt, S., Simplet, L., Rousset, A., 2020. Upstream migrating knickpoints and related sedimentary processes in a submarine canyon from a rare 20-year morphobathymetric time-lapse (Capbreton submarine canyon, Bay of Biscay, France). Mar. Geol.

Hallermeier, R.J., 1980. A profile zonation for seasonal sand beaches from wave climate. Coast. Eng. 4, 253–277.

Harris, P.T., Macmillan-lawler, M., Rupp, J., Baker, E.K., 2014. Geomorphology of the oceans. Mar. Geol. 352, 4–24.

Harris, P.T., Whiteway, T., 2011. Global distribution of large submarine canyons: Geomorphic differences between active and passive continental margins. Mar. Geol. 285, 69–86.

Kane, I.A., Clare, M.A., 2019. Dispersion, accumulation, and the ultimate fate of microplastics in deep-marine environments: A review and future directions. Front. Earth Sci. 7.

Korup, O., 2020. Bayesian geomorphology. Earth Surf. Process. Landforms.

Lehner, B., Verdin, K., Jarvis, A., 2008. New global hydrography derived from spaceborne elevation data. Eos (Washington. DC). 89, 93–94.

Li, Y., Clift, P.D., Böning, P., Blusztajn, J., Murray, R.W., Ireland, T., Pahnke, K., Helm, N.C., Giosan, L., 2018. Continuous Holocene input of river sediment to the Indus Submarine Canyon. Mar. Geol. 406, 159–176.

Luijendijk, E., Gleeson, T., Moosdorf, N., 2020. Fresh groundwater discharge insignificant for the world's oceans but important for coastal ecosystems. Nat. Commun. 11, 1260.

Maier, K.L., Johnson, S.Y., Hart, P., 2018. Controls on submarine canyon head evolution: Monterey Canyon, offshore central California. Mar. Geol. 404, 24–40.

Makalic, E., Schmidt, D.F., 2011. A simple Bayesian algorithm for feature ranking in high dimensional regression problems. Lect. Notes Comput. Sci. (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics) 7106 LNAI, 223–230.

Makalic, E., Schmidt, D.F., 2016. High-Dimensional Bayesian Regularised Regression with the BayesReg Package 1–18.

Mauffrey, M.A., Urgeles, R., Berné, S., Canning, J., 2017. Development of submarine canyons after the Mid-Pleistocene Transition on the Ebro margin, NW Mediterranean: The role of fluvial connections. Quat. Sci. Rev. 158, 77–93.

Middleton, G. V., 1966. Experiments on Density and Turbidity Currents: II. Uniform Flow of Density Currents. Can. J. Earth Sci. 3, 627–637.

Moosdorf, N., Cohen, S., von Hagke, C., 2018. A global erodibility index to represent sediment production potential of different rock types. Appl. Geogr. 101, 36–44.

Muis, S., Verlaan, M., Winsemius, H.C., Aerts, J.C.J.H., Ward, P.J., 2016. A global reanalysis of storm surges and extreme sea levels. Nat. Commun. 7.

Nitsche, M., Rickenmann, D., Turowski, J.M., Badoux, A., Kirchner, J.W., 2011. Evaluation of bedload transport predictions using flow resistance equations to account for macro-roughness in steep mountain streams. Water Resour. Res. 47.

Normark, W.R., Normark, W., Piper, D., Piper, D., Covault, J. A., Romans, B., Romans, B., Dartnell, P., Covault, Jacob A., Dartnell, Peter, Sliter, R.W., Sliter, R.W., 2009. Submarine canyon and fan systems of the California Continental Borderland. Earth Sci. urban Ocean South. Calif. Cont. Borderl. 454, 141–168.

Orange, D.L., Anderson, R.S., Breen, N.A., 1994. Regular canyon spacing in the submarine environment the link between hydrology and geomorphology. GSA Today.

Pratson, L., Coakley, B., 1996. A model for the headward erosion of submarine canyons induced by downslope-eroding sediment flows. Geol. Soc. Am. Bull. 108, 225–234.

Pratson, L.F., Nittrouer, C.A., Wiberg, P.L., Steckler, M.S., Swenson, J.B., Cacchione, D.A., Karson, J.A., Murray, A.B., Wolinsky, M.A., Gerber, T.P., Mullenbach, B.L., Spinelli, G.A., Fulthorpe, C.S., O'grady, D.B., Parker, G., Driscoll, N.W., Burger, R.L., Paola, C., Orange, D.L., Field, M.E., Friedrichs, C.T., Fedele, J.J., 2007. Seascape Evolution on Clastic Continental Shelves and Slopes, in: Continental Margin Sedimentation. pp. 339–380.

Pratson, L.F., Ryan, W.B.F., Mountain, G.S., Twichell, D.C., 1994. Submarine canyon initiation by downslope-eroding sediment flows : Evidence in late Cenozoic strata on the New Jersey continental slope.

Rogers, K.G., Goodbred, S.L., Khan, S.R., 2015. Shelf-to-canyon connections: Transport-related morphology and mass balance at the shallow-headed, rapidly aggrading Swatch of No Ground (Bay of Bengal). Mar. Geol. 369, 288–299.

Schwanghart, W., Scherler, D., 2014. Short Communication: TopoToolbox 2 - MATLAB-based software for topographic analysis and modeling in Earth surface sciences. Earth Surf. Dyn. 2, 1–7.

Shedlock, K., Giardini, D., Gruenthal, G., Zhang, P., 2003. The GSHAP global seismic hazard map. Seismol. Res. Lett. 71, 679–686.

Shepard, F.P., 1981. Submarine canyons: multiple causes and long-time persistence. Am. Assoc. Pet. Geol. Bull. 65, 1062–1077.

Smith, M.E., Finnegan, N.J., Mueller, E.R., Best, R.J., 2017. Durable Terrestrial Bedrock Predicts Submarine Canyon Formation. Geophys. Res. Lett. 44, 10,332-10,340.

Smith, M.E., Werner, S.H., Buscombe, D., Finnegan, N.J., Sumner, E.J., Mueller, E.R., 2018. Seeking the Shore: Evidence for Active Submarine Canyon Head Incision Due to Coarse

Sediment Supply and Focusing of Wave Energy. Geophys. Res. Lett. 45, 12,403-12,413.

Stolle, A., Schwanghart, W., Andermann, C., Bernhardt, A., Fort, M., Jansen, J.D., Wittmann, H., Merchel, S., Rugel, G., Adhikari, B.R., Korup, O., 2019. Protracted river response to medieval earthquakes. Earth Surf. Process. Landforms 44.

Sweet, M.L., Blum, M.D., 2016. Connections Between Fluvial To Shallow Marine Environments and Submarine Canyons: Implications For Sediment Transfer To Deep Water. J. Sediment. Res. 86, 1147–1162.

Syvitski, J.P.M., Kettner, A., 2011. Sediment flux and the anthropocene. Philos. Trans. R. Soc. A Math. Phys. Eng. Sci. 369, 957–975.

Tibshirani, R., 2011. Regression shrinkage and selection via the lasso: A retrospective. J. R. Stat. Soc. Ser. B Stat. Methodol. 73, 273–282.

van Erp, S., 2020. A tutorial on Bayesian penalized regression with shrinkage priors for small sample sizes, in: van de Schoot, R., Miočevic, M. (Eds.), Small Sample Size Solutions. Routledge, pp. 71–84.

van Erp, S., Oberski, D.L., Mulder, J., 2019. Shrinkage priors for Bayesian penalized regression. J. Math. Psychol. 89, 31–50.

Geophysical Research Letters

Supporting Information for

Where and Why do submarine canyons remain connected to the shore during sea level rise?
–
Insights from global topographic analysis and Bayesian regression

Anne Bernhardt[1], Wolfgang Schwanghart[2]

[1]Institute of Geological Sciences, Freie Universität Berlin, Malteserstrasse 74-100, 12249 Berlin, Germany
[2] Institute of Earth and Environmental Science, Potsdam University, Karl-Liebknecht-Str. 24-25, 14476 Potsdam-Golm, Germany

Contents of this file

Additional Supporting Information (Files uploaded separately)

Introduction

This Supporting Information includes a map that shows the global distribution of individual shore-connected canyon heads during the Last Glacial Maximum (LGM) and the present day (Fig. S1) and various scatter plots of each predictor to the individual canyon head (Fig. S2a, Data Set S1). In the main text, we bin this data into hexagonal polygons of a size of 50,000 km^2 to upscale the data set for Bayesian penalized regression modeling (Fig. S2b, Data Set 2).

Moreover, we include the details of the sensitivity analysis that guided the choice of the shrinkage prior (Figs. S3 to S6). The model parameters for all five shrinkage priors are included in Tables S1 to S5.

Global distribution of shore-connected canyon heads

In the main text of the manuscript, we show the present-day shore-connected canyon (SCC) heads as the fraction SCCs binned in 50,000 km^2 hexagons. For a more complete picture, we show the global distribution of individual SCCs during the present-day and the Last Glacial Maximum (Fig. S1) and the number of present-day SCC heads plotted against every predictor.

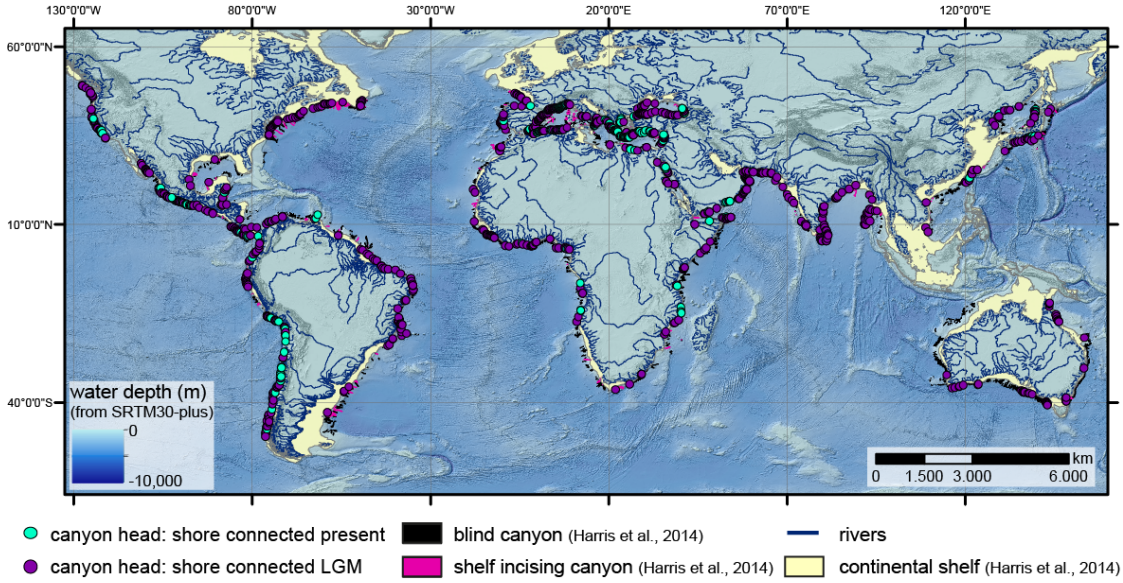
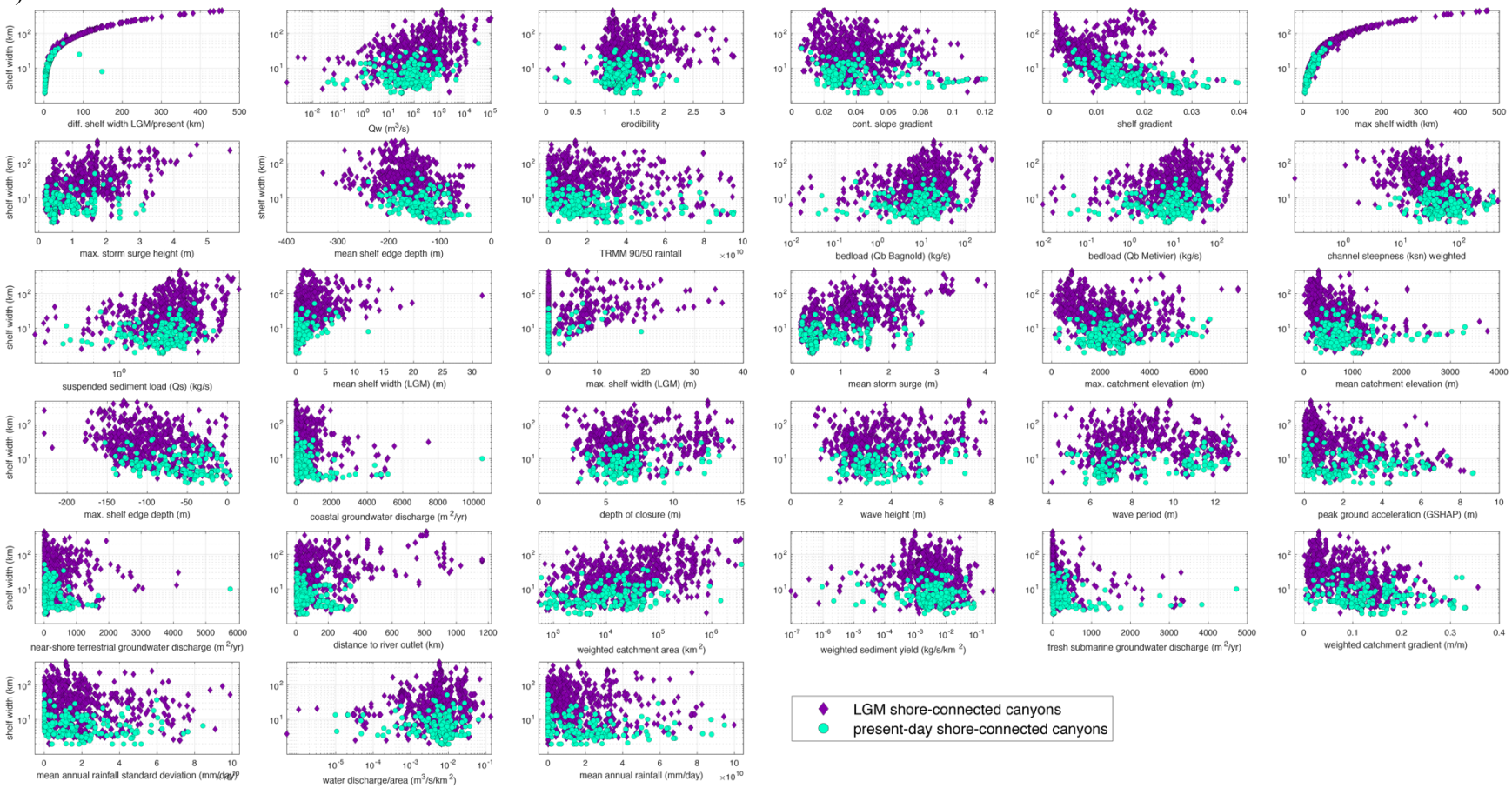


Fig. S1. Global distribution of shore-connected canyon (SCC) heads during the present day and the Last Glacial Maximum.

a)


shelf width (km)
diff. shelf width LGM/present (km)
Qw (m3/s)
erodibility
cont. slope gradient
shelf gradient
max shelf width (km)
max. storm surge height (m)
mean shelf edge depth (m)
TRMM 90/50 rainfall
bedload (Qb Bagnold) (kg/s)
bedload (Qb Metivier) (kg/s)
channel steepness (ksn) weighted
suspended sediment load (Qs) (kg/s)
mean shelf width (LGM) (m)
max. shelf width (LGM) (m)
mean storm surge (m)
max. catchment elevation (m)
mean catchment elevation (m)
max. shelf edge depth (m)
coastal groundwater discharge (m2/yr)
depth of closure (m)
wave height (m)
wave period (m)
peak ground acceleration (GSHAP) (m)
near-shore terrestrial groundwater discharge (m2/yr)
distance to river outlet (km)
weighted catchment area (km2)
weighted sediment yield (kg/s/km2)
fresh submarine groundwater discharge (m2/yr)
weighted catchment gradient (m/m)
mean annual rainfall standard deviation (mm/day)
water discharge/area (m3/s/km2)
mean annual rainfall (mm/day)
LGM shore-connected canyons
present-day shore-connected canyons

b)

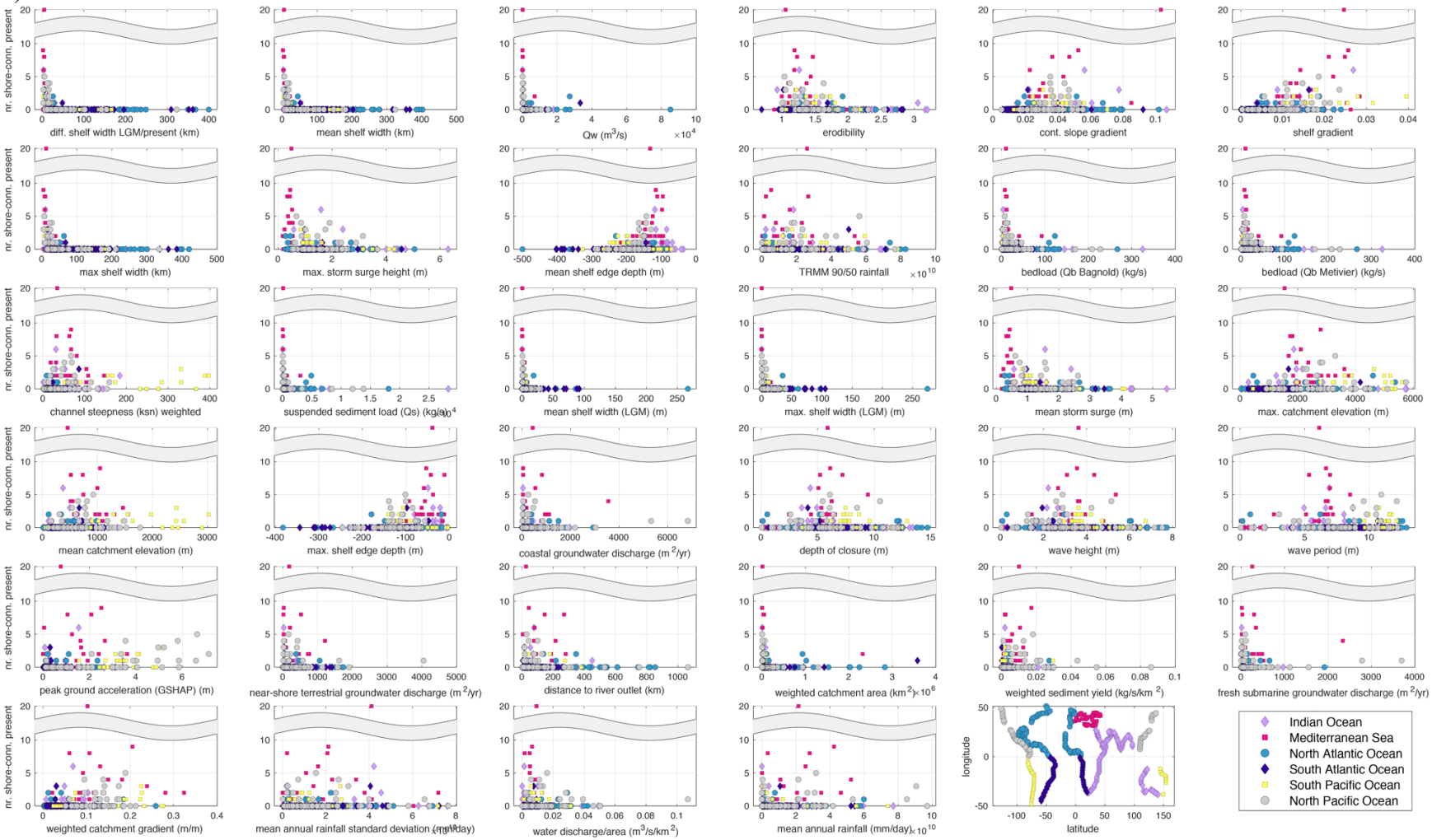


Fig. S2. a) Cross-plots of 33 predictors against mean, present-day shelf width for present-day and LGM shore-connected canyons (SCCs). Note the log-scale of the y-axis and some x-axes. b) Predictors plotted against the total numbers of present-day and LGM shore-connected canyons (SCCs) in each hexagon.

Sensitivity Analyses of shrinkage priors

To identify the most important predictors on the number of present-day shore-connected canyon (SCC) heads and to predict these numbers on a global scale, we employ Bayesian penalized regression. Bayesian statistics applies probabilities to statistical problems offering a statistical way to learn from new data to update prior beliefs (e.g., Efron, 2013) while accounting for uncertainties and Bayesian regression is an emergent state-of-the-art tool in geomorphology (Korup, 2020). In the frequentist approach to penalized regression a penalty term is introduced with the aim to shrink small regression coefficients towards zero and therefore reducing or eliminating the predictor variable from the model, while large coefficients remain large (e.g., Tibshirani, 2011). In Bayesian penalized regression, penalization is incorporated through the prior distribution (e.g., van Erp et al., 2019). Therefore, the shrinkage prior allows us to determine important predictors and to discard these predictors that are unimportant to predict the number of SCC canyon heads. We used used *bayesreg*, a Matlab toolbox for fitting Bayesian penalized regression models with continuous shrinkage prior densities for penalized regression models (Makalic and Schmidt, 2016, version 1.9, 2017-2020). As we are predicting counts of present-day SCC canyons per hexagon, we specified a Poisson distribution for the response variable *in bayesreg*. To choose the shrinkage prior, we followed the suggested procedure of van Erp (2020): We applied all five shrinkage priors available for the Poission problem in *bayesreg* with their default hyperparameters (Figs. S3-S6, Table S1-S5). All priors result in similar prediction root mean square errors (RMSE=0.93-0.95; Table S1-S5, Fig. S6), Watanabe–Akaike information criteria (WAIC=220-231), pseudo R^2 of ~0.5, and posterior distributions of the regression coefficients (Figs. S4). ***Based on the lowest RMSE*** (Table S1) ***and the efficient and stable sampling performance*** (Table S1-5, Fig. S3)***, we show the results of the lasso shrinkage prior in the main text of the manuscript.*** To determine the importance of each predictor, we use the Bayesian feature ranking algorithm of Makalic and Schmidt (2011). The rank corresponds to the strength of the association between the predictors and the response (present-day SCC counts per hexagon) where lower ranks denote more important predictors (Table S1-S5). Because the ranking process is repeated for each posterior sample, the final rank of the predictor is determined from the complete set of rankings based on the 75th percentile.

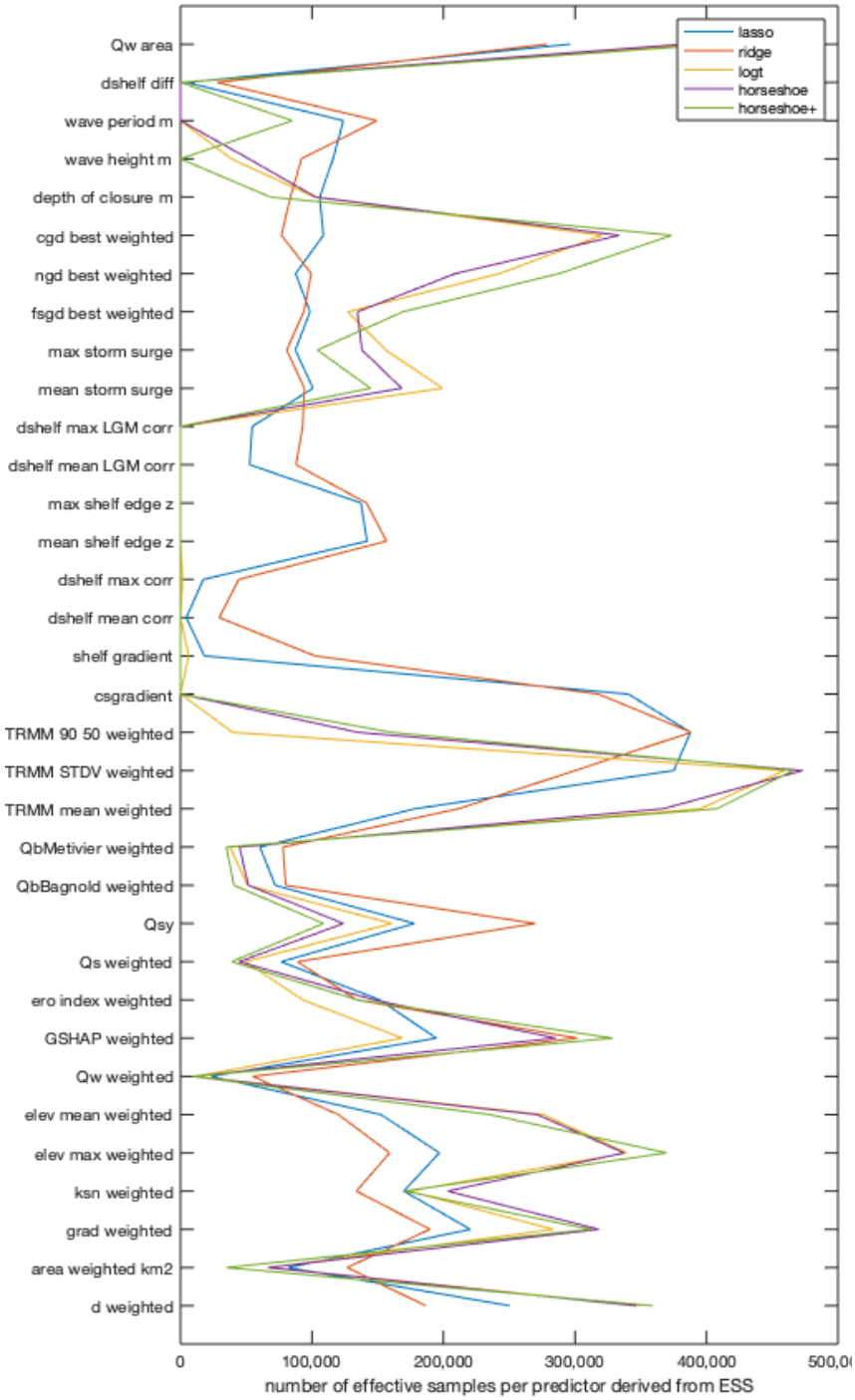


Figure S3. The number of samples used to compute the posterior distribution of all predictors for each shrinkage prior. This number was computed by applying the effective sample size (ESS, Table S1-S5) to the total MCMC sample size of n= 500 000. The logt, the horseshoe and the horseshoe+ shrinkage priors ineffectively sample some predictors and the efficient sample size is zero for e.g., the shelf width. Effective sampling is most stable when using the lasso and the ridge prior with the minimum number of samples used is 4,385 and 27,921 respectively.

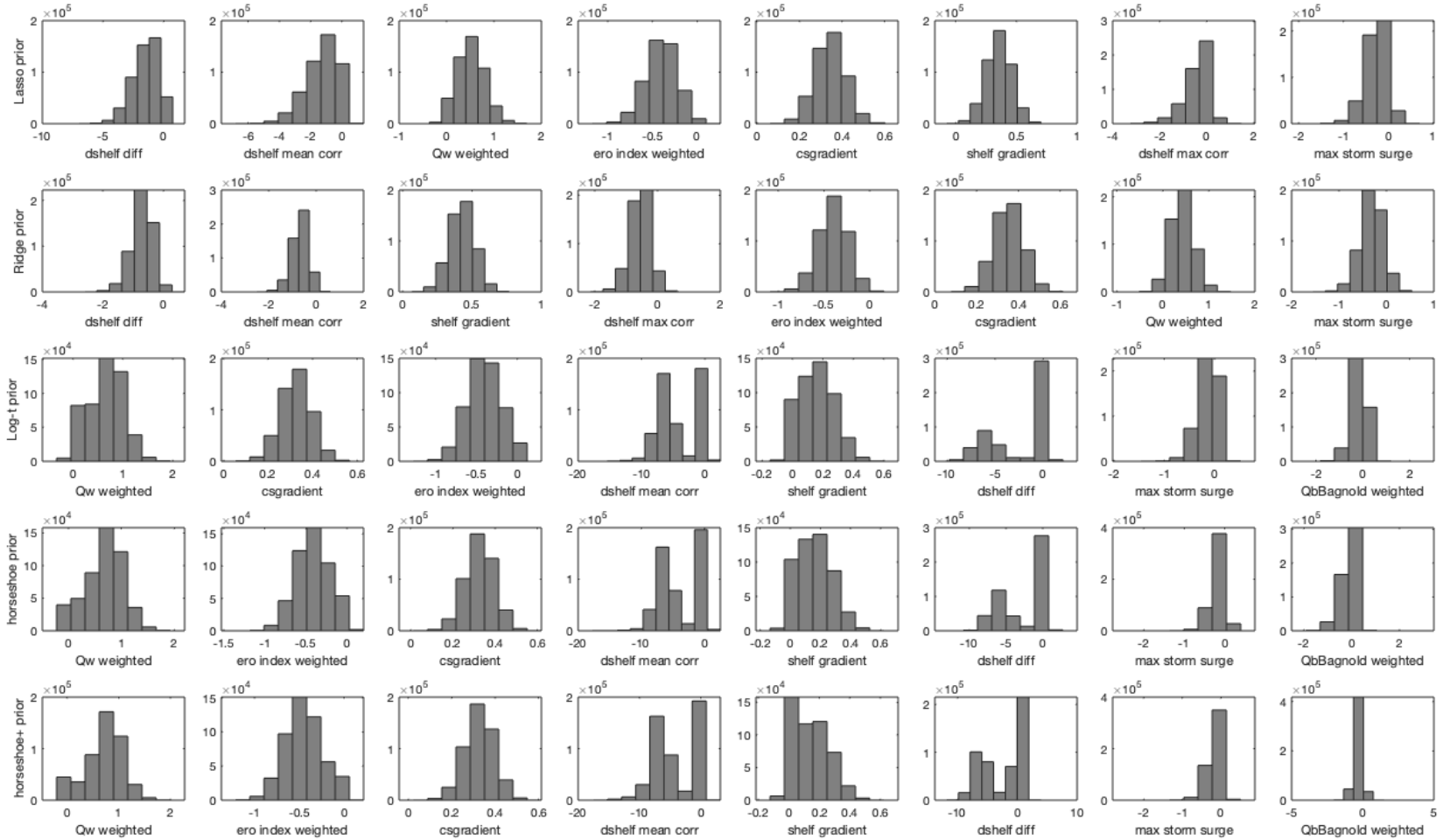


Fig. S4. Posterior distributions of the 8 most highly ranked predictor coefficients for each shrinkage prior ordered by rank. For abbreviations of predictors see Table 1 in the main manuscript. The predictors on the left have the highest ranking (see also Table S1-S5).

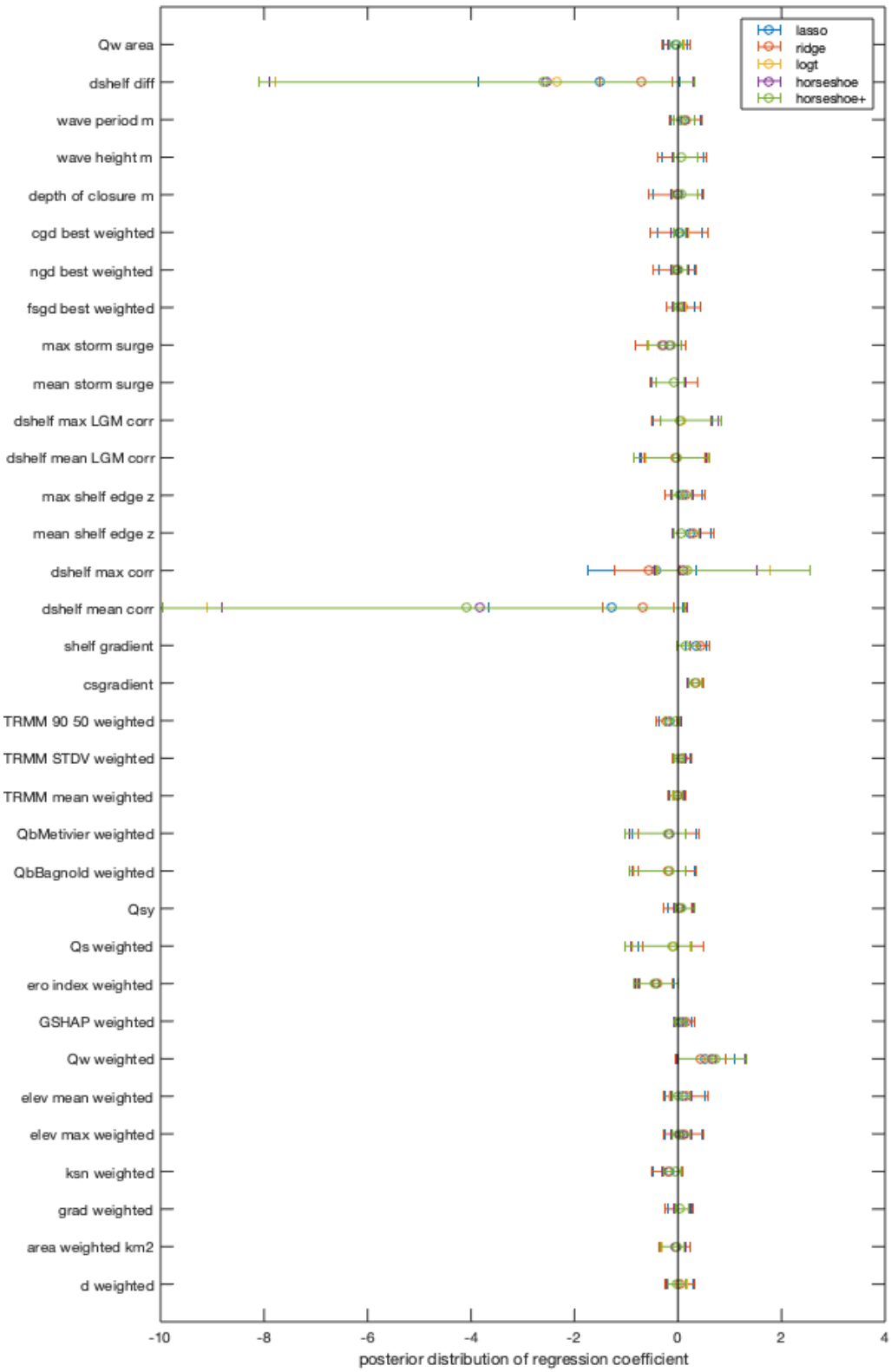


Fig. S5. Comparison of posterior mean estimates and 95% credibility intervals obtained employing the five shrinkage priors. All five shrinkage priors return comparable posterior distributions for all predictors.

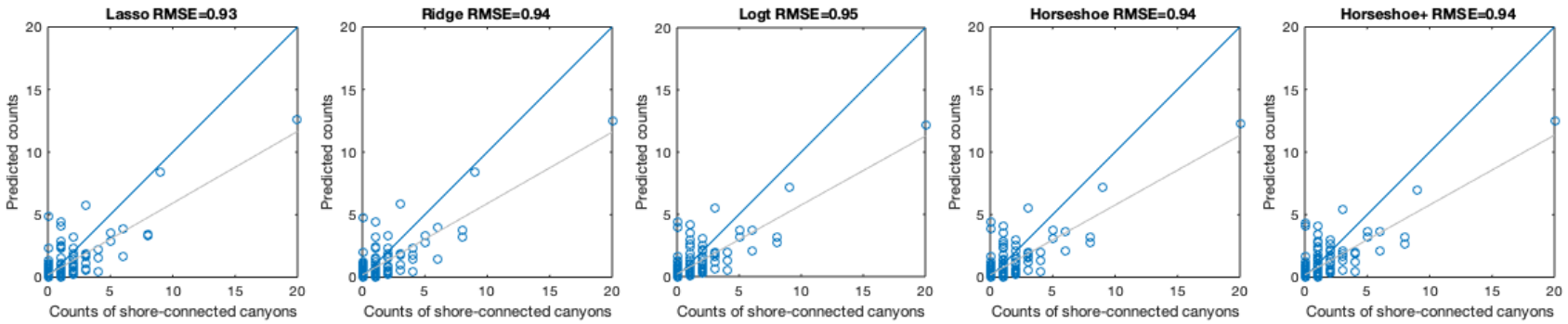


Fig. S6. Comparison of the actual counts of presently shore-connected canyons and the predicted counts from Bayesian regression for all five priors. Blue lines show 1:1 reference lines and grey lines are least-square lines.

Bayesian Poisson lasso regression

MCMC Samples	= 500000	Number of obs	= 347
MCMC Burnin	= 10000	Number of vars	= 34
MCMC Thinning	= 100	Overdispersion	= 0.8113
		Pseudo R2	= 0.5433
		WAIC	= 225.2

Parameter	mean(Coef)	std(Coef)	[95% Cred. Interval]		tStat	Rank	ESS
dshelf_diff	−1.56206	1.08041	−3.92990	0.03486	−1.446	1 *	0.9
dshelf_mean_corr	−1.27846	1.03177	−3.65104	0.10040	−1.239	2 *	1.0
Qw_weighted	0.51996	0.28837	−0.00313	1.10846	1.803	3 *	4.7
ero_index_weighted	−0.40133	0.17550	−0.76035	−0.07394	−2.287	4 **	30.6
csgradient	0.34328	0.07050	0.20376	0.48017	4.869	4 **	64.6
shelf_gradient	0.35196	0.10420	0.14393	0.55243	3.378	6 **	3.5
dshelf_max_corr	−0.40650	0.52613	−1.70896	0.35483	−0.773	7 *	3.4
max_storm_surge	−0.27517	0.24821	−0.82367	0.13795	−1.109	7 *	17.6
mean_shelf_edge_z	0.23568	0.19393	−0.10125	0.64301	1.215	9 *	29.0
TRMM_90_50_weighted	−0.16829	0.10326	−0.37948	0.02046	−1.630	10 *	77.4
QbBagnold_weighted	−0.19589	0.29873	−0.87897	0.31708	−0.656	11 *	13.1
QbMetivier_weighted	−0.18814	0.30983	−0.89673	0.34718	−0.607	11	12.6
ksn_weighted	−0.16178	0.14417	−0.47983	0.08125	−1.122	13 *	32.5
Qs_weighted	−0.07812	0.29859	−0.75618	0.48980	−0.262	13	15.1
dshelf_mean_LGM_corr	−0.04603	0.29824	−0.72504	0.52654	−0.154	15	10.6
dshelf_max_LGM_corr	0.04874	0.27654	−0.49854	0.66027	0.176	15	11.3
mean_storm_surge	−0.08212	0.21863	−0.51612	0.37389	−0.376	15	20.1
elev_max_weighted	0.08788	0.17218	−0.23608	0.45679	0.510	18	37.1
elev_mean_weighted	0.08991	0.19010	−0.25397	0.51385	0.473	18	31.1
max_shelf_edge_z	0.08559	0.17409	−0.25132	0.45013	0.492	18	28.3
cgd_best_weighted	0.02929	0.21100	−0.40450	0.47055	0.139	18	22.7
depth_of_closure_m	0.00675	0.22203	−0.47156	0.45017	0.030	18	20.6
wave_height_m	0.06235	0.19763	−0.31599	0.49351	0.316	18	22.8
wave_period_m	0.12036	0.15004	−0.14830	0.44328	0.802	18 *	23.1
GSHAP_weighted	0.09205	0.08636	−0.07039	0.26624	1.066	25 *	37.3
ngd_best_weighted	−0.01770	0.17018	−0.37254	0.32690	−0.104	25	19.2
d_weighted	0.02468	0.13168	−0.23803	0.29497	0.187	27	46.5
area_weighted_km2	−0.04490	0.14728	−0.36530	0.22925	−0.305	27	16.7
Qsy	0.04341	0.11821	−0.18968	0.28499	0.367	27	34.0
fsgd_best_weighted	0.05517	0.13495	−0.21795	0.33239	0.409	27	17.7
grad_weighted	0.03136	0.11737	−0.20493	0.26851	0.267	31	38.6
TRMM_STDV_weighted	0.05218	0.08909	−0.11779	0.23506	0.586	31	60.2
Qw_area	−0.04156	0.11699	−0.28938	0.18159	−0.355	31	58.8
TRMM_mean_weighted	−0.00986	0.07347	−0.16134	0.13176	−0.134	34	35.1

Table S1. Summary statistics of the Bayesian regression model for a Poisson-distributed target variable using the Lasso shrinkage prior.

Bayesian Poisson ridge regression

MCMC Samples	= 500000	Number of obs	= 347
MCMC Burnin	= 10000	Number of vars	= 34
MCMC Thinning	= 100	Overdispersion	= 0.8922
		Pseudo R2	= 0.5309
		WAIC	= 231.79

Parameter	mean(Coef)	std(Coef)	[95% Cred. Interval]		tStat	Rank		ESS
dshelf_diff	−0.71734	0.35353	−1.50067	−0.10654	−2.029	1	**	5.6
dshelf_mean_corr	−0.68055	0.35157	−1.45794	−0.06795	−1.936	2	**	5.9
shelf_gradient	0.42786	0.09128	0.24564	0.60389	4.688	3	**	20.5
dshelf_max_corr	−0.55698	0.31900	−1.23248	0.03092	−1.746	3	*	8.8
ero_index_weighted	−0.40168	0.15943	−0.72758	−0.10295	−2.520	5	**	26.8
csgradient	0.35233	0.07020	0.21336	0.48851	5.019	5	**	63.6
Qw_weighted	0.42153	0.24502	−0.04049	0.92372	1.720	7	*	11.0
max_storm_surge	−0.31487	0.24278	−0.82230	0.13738	−1.297	8	*	16.2
mean_shelf_edge_z	0.29119	0.19334	−0.08119	0.68069	1.506	9	*	31.3
TRMM_90_50_weighted	−0.21302	0.10348	−0.42017	−0.01452	−2.059	10	**	77.7
ksn_weighted	−0.20467	0.15339	−0.51828	0.08438	−1.334	11	*	26.8
QbBagnold_weighted	−0.19017	0.28576	−0.77199	0.36311	−0.665	11	*	16.0
QbMetivier_weighted	−0.17609	0.29358	−0.77343	0.39486	−0.600	11		15.6
Qs_weighted	−0.06947	0.29506	−0.66857	0.50145	−0.235	14		17.8
dshelf_mean_LGM_corr	−0.05237	0.29857	−0.66333	0.51763	−0.175	14		17.6
elev_mean_weighted	0.13777	0.21622	−0.27814	0.57403	0.637	16		24.0
GSHAP_weighted	0.14339	0.09217	−0.03942	0.32227	1.556	16	*	60.3
dshelf_max_LGM_corr	0.04239	0.29029	−0.51780	0.63160	0.146	16		18.6
cgd_best_weighted	0.01573	0.27458	−0.52818	0.56499	0.057	16		15.3
depth_of_closure_m	−0.03676	0.26611	−0.57623	0.48109	−0.138	16		16.7
elev_max_weighted	0.10844	0.19486	−0.27518	0.49268	0.557	21		31.8
max_shelf_edge_z	0.13913	0.19118	−0.24553	0.50718	0.728	21	*	28.3
mean_storm_surge	−0.08957	0.22930	−0.52873	0.37963	−0.391	21		18.8
wave_height_m	0.06646	0.23411	−0.38871	0.53991	0.284	21		18.4
wave_period_m	0.13856	0.16499	−0.17921	0.46978	0.840	21	*	29.8
fsgd_best_weighted	0.10524	0.16622	−0.23135	0.42790	0.633	26		18.7
ngd_best_weighted	−0.06317	0.20914	−0.47284	0.35509	−0.302	26		19.9
d_weighted	0.03501	0.14541	−0.25398	0.31615	0.241	28		37.3
area_weighted_km2	−0.04064	0.15304	−0.35532	0.24680	−0.266	28		25.3
grad_weighted	0.02332	0.13713	−0.25066	0.28725	0.170	28		37.9
Qsy	0.00373	0.13461	−0.26626	0.26154	0.028	28		54.0
TRMM_STDV_weighted	0.07897	0.09924	−0.11608	0.27284	0.796	28	*	60.0
Qw_area	−0.03192	0.13236	−0.29602	0.22342	−0.241	28		55.8
TRMM_mean_weighted	−0.01776	0.08227	−0.18363	0.13937	−0.216	34		42.2

Table S2. Summary statistics of the Bayesian regression model for a Poisson-distributed target variable using the Ridge shrinkage prior.

Bayesian Poisson log-t regression

MCMC Samples	= 500000	Number of obs	= 347
MCMC Burnin	= 10000	Number of vars	= 34
MCMC Thinning	= 100	Overdispersion	= 0.84204
		Pseudo R2	= 0.5478
		WAIC	= 221.04

Parameter	mean(Coef)	std(Coef)	[95% Cred. Interval]		tStat	Rank	ESS
Qw_weighted	0.64485	0.36312	-0.01116	1.29222	1.776	1 *	2.3
csgradient	0.32521	0.06772	0.18980	0.45577	4.802	2 **	0.0
ero_index_weighted	-0.40354	0.20696	-0.80743	-0.00637	-1.950	3 **	18.6
dshelf_mean_corr	-4.08243	3.29695	-9.09343	0.16254	-1.238	4 *	0.0
shelf_gradient	0.15768	0.11171	-0.01403	0.38290	1.412	5 *	1.3
dshelf_diff	-2.32835	2.97636	-7.77587	0.30061	-0.782	6	0.0
max_storm_surge	-0.15122	0.18088	-0.57593	0.07726	-0.836	7 *	31.2
QbBagnold_weighted	-0.16587	0.27288	-0.86671	0.14828	-0.608	8	10.2
QbMetivier_weighted	-0.17423	0.29312	-0.93889	0.15384	-0.594	8	7.5
mean_storm_surge	-0.07968	0.14491	-0.42390	0.14808	-0.550	10	39.9
Qs_weighted	-0.09063	0.26637	-0.86971	0.24803	-0.340	11	9.9
TRMM_90_50_weighted	-0.06391	0.08319	-0.26498	0.05043	-0.768	11 *	7.9
dshelf_max_corr	0.11055	0.57141	-0.40173	1.76716	0.193	11	0.3
mean_shelf_edge_z	0.07038	0.13077	-0.11098	0.41517	0.538	11	0.0
dshelf_mean_LGM_corr	-0.01343	0.27514	-0.61810	0.56898	-0.049	11	0.0
dshelf_max_LGM_corr	0.04910	0.24847	-0.32567	0.77222	0.198	11	0.0
depth_of_closure_m	0.06005	0.12958	-0.14715	0.37344	0.463	11	20.4
wave_height_m	0.06766	0.12344	-0.12012	0.36625	0.548	11	7.7
wave_period_m	0.06628	0.10550	-0.08435	0.33066	0.628	11	0.0
area_weighted_km2	-0.03463	0.10761	-0.31534	0.14813	-0.322	20	13.6
ksn_weighted	-0.06209	0.09432	-0.30500	0.05942	-0.658	20	34.7
Qsy	0.05823	0.09415	-0.08211	0.29151	0.618	20	32.1
max_shelf_edge_z	0.03241	0.10199	-0.14783	0.29198	0.318	20	0.0
d_weighted	-0.01303	0.08781	-0.22224	0.16561	-0.148	24	70.0
grad_weighted	0.04101	0.07846	-0.08762	0.23334	0.523	24	56.7
elev_max_weighted	0.02593	0.09453	-0.14374	0.26655	0.274	24	67.9
elev_mean_weighted	0.01866	0.09588	-0.15203	0.26438	0.195	24	55.2
ngd_best_weighted	0.01664	0.08124	-0.14189	0.20935	0.205	24	48.5
cgd_best_weighted	0.01457	0.08085	-0.14296	0.19910	0.180	24	64.1
Qw_area	-0.02126	0.07333	-0.20600	0.10880	-0.290	24	78.3
GSHAP_weighted	0.02617	0.05629	-0.07060	0.16229	0.465	31	33.7
TRMM_mean_weighted	0.00702	0.04829	-0.09520	0.11477	0.145	31	79.1
TRMM_STDV_weighted	0.01524	0.05640	-0.09216	0.15063	0.270	31	91.8
fsgd_best_weighted	0.00678	0.06030	-0.12228	0.13647	0.112	31	25.5

Table S3. Summary statistics of the Bayesian regression model for a Poisson-distributed target variable using the Log-t shrinkage prior.

Bayesian Poisson horseshoe regression

MCMC Samples	= 500000	Number of obs	= 347
MCMC Burnin	= 10000	Number of vars	= 34
MCMC Thinning	= 100	Overdispersion	= 0.84202
		Pseudo R2	= 0.5483
		WAIC	= 220.43

Parameter	mean(Coef)	std(Coef)	[95% Cred. Interval]		tStat	Rank		ESS
Qw_weighted	0.66939	0.35528	−0.00834	1.29974	1.884	1	*	1.8
ero_index_weighted	−0.41196	0.20657	−0.81206	−0.00714	−1.994	2	**	29.5
csgradient	0.32641	0.06782	0.19084	0.45685	4.813	2	**	0.0
dshelf_mean_corr	−3.83752	3.29468	−8.81877	0.16504	−1.165	4	*	0.0
shelf_gradient	0.15917	0.11236	−0.01504	0.38657	1.417	5	*	0.0
dshelf_diff	−2.55542	3.03296	−7.89894	0.29174	−0.843	6		0.0
max_storm_surge	−0.15361	0.18559	−0.58608	0.07646	−0.828	7	*	27.6
QbBagnold_weighted	−0.17126	0.27912	−0.88873	0.14860	−0.614	8		10.3
QbMetivier_weighted	−0.17885	0.29640	−0.94937	0.15372	−0.603	8		9.0
mean_storm_surge	−0.07844	0.14765	−0.42946	0.15080	−0.531	10		33.7
Qs_weighted	−0.09408	0.27747	−0.90488	0.26075	−0.339	11		8.9
TRMM_90_50_weighted	−0.06163	0.08172	−0.26123	0.05071	−0.754	11	*	27.0
dshelf_max_corr	0.09902	0.55813	−0.43958	1.51143	0.177	11		0.0
mean_shelf_edge_z	0.07343	0.13379	−0.10694	0.42744	0.549	11		0.0
dshelf_mean_LGM_corr	−0.02548	0.29437	−0.71227	0.55022	−0.087	11		0.0
dshelf_max_LGM_corr	0.04669	0.25256	−0.34908	0.78539	0.185	11		0.0
depth_of_closure_m	0.06118	0.12833	−0.13570	0.37671	0.477	11		20.5
wave_height_m	0.06741	0.12253	−0.11142	0.36718	0.550	11		10.4
wave_period_m	0.06435	0.10454	−0.08374	0.32869	0.616	11		0.0
area_weighted_km2	−0.03769	0.10741	−0.32408	0.13661	−0.351	20		13.3
ksn_weighted	−0.05951	0.09314	−0.30264	0.05923	−0.639	20		40.6
Qsy	0.05951	0.09405	−0.07476	0.29497	0.633	20		24.8
max_shelf_edge_z	0.03170	0.10119	−0.14362	0.29355	0.313	20		0.0
d_weighted	−0.01236	0.08662	−0.22075	0.16358	−0.143	24		69.3
grad_weighted	0.03969	0.07666	−0.08253	0.23084	0.518	24		63.6
elev_max_weighted	0.02540	0.09233	−0.13659	0.26272	0.275	24		67.5
elev_mean_weighted	0.01822	0.09419	−0.14537	0.26263	0.193	24		54.3
ngd_best_weighted	0.01690	0.07775	−0.13082	0.20209	0.217	24		41.8
cgd_best_weighted	0.01511	0.07540	−0.12747	0.18471	0.200	24		66.7
Qw_area	−0.02054	0.07145	−0.20243	0.10465	−0.287	24		76.6
GSHAP_weighted	0.02466	0.05464	−0.06901	0.15747	0.451	31		57.1
TRMM_mean_weighted	0.00714	0.04645	−0.08985	0.11162	0.154	31		73.4
TRMM_STDV_weighted	0.01287	0.05402	−0.09173	0.14163	0.238	31		94.6
fsgd_best_weighted	0.00735	0.05666	−0.11122	0.12846	0.130	31		26.9

Table S4. Summary statistics of the Bayesian regression model for a Poisson-distributed target variable using the Horseshoe shrinkage prior.

Bayesian Poisson horseshoe+ regression

MCMC Samples	= 500000	Number of obs	= 347
MCMC Burnin	= 10000	Number of vars	= 34
MCMC Thinning	= 100	Overdispersion	= 0.86572
		Pseudo R2	= 0.5477
		WAIC	= 219.92

Parameter	mean(Coef)	std(Coef)	[95% Cred. Interval]		tStat	Rank		ESS
Qw_weighted	0.71418	0.35507	−0.00380	1.33041	2.011	1	*	1.8
ero_index_weighted	−0.44661	0.20932	−0.83984	−0.00246	−2.134	2	**	27.0
csgradient	0.32659	0.06687	0.19294	0.45509	4.884	3	**	0.0
dshelf_mean_corr	−4.09754	3.52100	−9.95034	0.13110	−1.164	4	*	0.0
shelf_gradient	0.14127	0.11732	−0.01683	0.37995	1.204	5	*	0.0
dshelf_diff	−2.60921	3.11463	−8.09472	0.31647	−0.838	6		0.0
max_storm_surge	−0.14638	0.19037	−0.58850	0.06083	−0.769	7	*	20.8
QbBagnold_weighted	−0.17864	0.30172	−0.93833	0.13830	−0.592	8		8.1
QbMetivier_weighted	−0.19242	0.32647	−1.01850	0.14108	−0.589	8		7.0
mean_storm_surge	−0.07452	0.14969	−0.43670	0.13178	−0.498	10		28.9
Qs_weighted	−0.10571	0.31087	−1.03834	0.25547	−0.340	11		7.8
Qsy	0.05859	0.09549	−0.05405	0.30958	0.614	11		21.7
dshelf_max_corr	0.16848	0.73127	−0.43150	2.54914	0.230	11		0.0
dshelf_mean_LGM_corr	−0.03261	0.32482	−0.86076	0.59956	−0.100	11		0.0
dshelf_max_LGM_corr	0.04886	0.26281	−0.35061	0.85043	0.186	11		0.0
depth_of_closure_m	0.06361	0.12902	−0.10212	0.39195	0.493	11		13.7
wave_height_m	0.06804	0.12392	−0.08430	0.38085	0.549	11		0.0
wave_period_m	0.05438	0.10045	−0.06950	0.32512	0.541	11		17.0
area_weighted_km2	−0.03499	0.10172	−0.32764	0.10918	−0.344	19		7.0
ksn_weighted	−0.04179	0.08429	−0.27784	0.04996	−0.496	19		34.1
TRMM_90_50_weighted	−0.04470	0.07449	−0.24314	0.04426	−0.600	19		32.0
mean_shelf_edge_z	0.05855	0.12390	−0.09126	0.40812	0.473	19		0.0
d_weighted	−0.00997	0.07659	−0.20842	0.14273	−0.130	23		71.8
grad_weighted	0.03254	0.06904	−0.06258	0.22052	0.471	23		62.4
elev_max_weighted	0.01789	0.07644	−0.11030	0.22582	0.234	23		73.9
elev_mean_weighted	0.01340	0.08094	−0.11708	0.22922	0.166	23		47.1
GSHAP_weighted	0.01732	0.04607	−0.05618	0.14114	0.376	23		65.7
max_shelf_edge_z	0.02437	0.08925	−0.12068	0.27462	0.273	23		0.0
ngd_best_weighted	0.01310	0.06076	−0.09495	0.16625	0.216	23		57.6
cgd_best_weighted	0.01123	0.05625	−0.08799	0.14401	0.200	23		74.7
Qw_area	−0.01202	0.05806	−0.16801	0.09020	−0.207	23		81.6
TRMM_mean_weighted	0.00621	0.03793	−0.07062	0.10031	0.164	32		81.6
TRMM_STDV_weighted	0.00763	0.04376	−0.07834	0.11974	0.174	32		93.1
fsgd_best_weighted	0.00642	0.04319	−0.07884	0.10614	0.149	32		33.8

Table S5. Summary statistics of the Bayesian regression model for a Poisson-distributed target variable using the Horsehoe+ shrinkage prior.

Data Set S1. Information and predictor variable values for each individual canyon head.

Data Set S2. Information and predictor variable values for numbers of submarine canyon heads in each 50 000 km^2 hexagon.