LIVING BY THE COAST

with accelerating sea-level rise

Floris Calkoen

LIVING BY THE COAST WITH ACCELERATING SEA-LEVEL RISE

Floris Reinier Calkoen

LIVING BY THE COAST WITH ACCELERATING SEA-LEVEL RISE

-

Dissertation

for the purpose of obtaining the degree of doctor
at Delft University of Technology
by the authority of the Rector Magnificus,
Prof. dr. ir. H. Bijl,
chair of the Board for Doctorates
to be defended publicly on
Tuesday, 7 April 2026, at 17:30 o'clock

by

Floris Reinier CALKOEN

This dissertation has been approved by the (co)promotors.

Composition of the doctoral committee:

Rector Magnificus,	chairperson
Prof. dr. ir. S. G. J. Aarninkhof,	Delft University of Technology, promotor
Prof. dr. R.W.M.R.J.B. Ranasinghe,	IHE Delft, NL & University of Twente, NL, promotor
Dr. ir. A. P. Luijendijk,	Delft University of Technology & Deltares, NL, copromotor

Independent members:

Prof. dr. ir. F.C. Vossepoel,	Delft University of Technology
Prof. dr. P. Ruggiero,	Oregon State University, United States
Dr. E.E. Koks,	VU Amsterdam, The Netherlands
Dr. G.A. Corzo Perez,	IHE Delft, The Netherlands
Prof. dr. B.K. van Wesenbeeck,	Delft University of Technology, *reserve member*

Dr. F. Baart, Delft University of Technology & Rijkswaterstaat, has contributed greatly to this research.

This research was supported by the Horizon 2020 programme Coastal Climate Core Services (CoCliCo) (grant agreement No 101003598) and the Deltares strategic research programs on Flooding (Moonshot 2) and Enabling Technologies.

Living by the coast with accelerating sea-level rise
Doctoral dissertation, Delft University of Technology
Author: Floris Reinier CALKOEN

Keywords	coastal science; sea-level rise; coastal erosion; global analysis; impact assessment; climate adaptation; sandy coasts; shoreline projections; earth observation; deep learning; cloud-native geospatial; pangeo

Cover photo	Floris Calkoen
Cover design	Floris Calkoen & Taras Neumyvaikin
Printed by	Ridderprint

An electronic version of this dissertation is available at http://repository.tudelft.nl/

Contents

Summary

The world's coasts are under increasing pressure from intense human development and climate change. This convergence creates a need to better understand how sea-level rise (SLR) will change hazards that take place in the coastal zone. One of these is coastal erosion, a natural process of shoreline adjustment that has the potential to become a hazard where it coincides with livelihoods, assets, infrastructure or even ecosystems where there is no place to migrate landwards. For this reason, a better understanding of the impacts associated with coastal erosion, especially under accelerating SLR, is necessary to inform climate adaptation policy. This thesis addresses this challenge by developing a framework that integrates Earth-observation satellite data, cloud technology, and artificial intelligence (AI) to map the coast at high (100 m) resolution. The resulting geomorphological classification is then used to produce refined probabilistic future shoreline projections, enabling a worldwide impact assessment of future coastal erosion on individual buildings under different SLR scenarios.

The primary objective was achieved through a three-part approach that progressively builds from a technical foundation to the final impact assessment. First, the research developed a scalable, cloud-native analytical framework capable of processing petabyte-scale satellite data archives. Subsequently, this framework was used to map the physical composition of the world's coast by using deep learning to classify Earth-observation data. This 100 m global coastal typology provided the geomorphological context for the final step, where these components were integrated to conduct a first-order impact assessment of future coastal erosion. For this assessment, probabilistic future shoreline change was projected using an equilibrium-profile model, with the typology constraining its application to sandy sediment plains and dune coasts. By intersecting these projected shoreline positions with global building data, the number of potentially impacted buildings was quantified.

The research yields several findings that characterize the state of the global coastal zone. The new machine-learning-derived coastal typology reveals that over 60% of the world's ice-free coast consists of soft, erodible sediments. The analysis estimates that 40% of the global coast is fronted by sandy, gravel, or shingle beaches, a higher figure than current estimates of 31%, with the difference largely attributed to the new, zonally-unbiased transect system used in this thesis. The analysis also reveals geomorphological sensitivities, showing that 20% of these beaches are backed by cliffs. This cross-shore configuration can limit their natural capacity to migrate landward with SLR, especially where cliffs are composed of resistant lithologies, a critical detail this satellite-based approach cannot yet resolve. Turning to coastal development, the analysis shows that a quarter of the world's coast ($\sim$253 000 km) has at least one building, per alongshore km, within the first kilometer of land, hosting ap-

proximately 76 million buildings in total. From these patterns, the thesis introduces the "Empirical Setback Zone" (ESZ), a data-driven metric that quantifies the de facto setback coastal communities implicitly maintain. This metric shows that for 10% of the total developed coasts, the first building is located less than 37 m from the shoreline. Combining these analyses reveals that coastal development is disproportionately concentrated on the most dynamic coasts; for instance, building density on sandy, gravel, and shingle coasts is 354 buildings per alongshore kilometer, significantly higher than the global average of 299.

This high present-day exposure translates into significant future impacts when combined with projections of future shoreline erosion. By 2100, between 1.6 million (under a low-emissions SSP1-2.6 scenario) and 2.4 million (under a high-emissions SSP5-8.5 scenario) buildings are projected to be impacted by coastal erosion. While substantial, this figure represents a conservative lower-bound estimate, as the model's application was first refined by the machine learning (ML)-based coastal typology to sandy sediment plains and dune coasts (21% of the world's coast), before data limitations (primarily the lack of reliable nearshore slope data) further reduced the study area to 9.3%. It is critical, however, to acknowledge the limitations of this impact assessment. The results are subject to uncertainties from the underlying coastal typology and the structural assumptions of the equilibrium-profile model used for the projections. Furthermore, the assessment quantifies impacts by combining the hazard (future coastal erosion) with coastal exposure (building footprints), but it stops short of a full climate risk assessment. Such an assessment typically considers the full probability distribution and would as well require a complete analysis of vulnerability, which in turn includes not only a system's sensitivity to coastal erosion but also its adaptive capacity; the socioeconomic ability to cope with and respond to this hazard.

The scientific contributions of this thesis are twofold. Methodologically, the research demonstrates a scalable, cloud-native framework for applying modern AI to petabyte-scale satellite archives, bridging the gap between detailed local studies and coarse global analyses. Thematically, this framework was used to produce a set of globally consistent data products. These include a new global coastal transect system (GCTS), a high-resolution coastal typology derived using deep learning, several coastal exposure metrics and probabilistic future shoreline projections. Together, these components enable a first-order impact assessment of future coastal erosion on individual buildings, advancing the field from a hazard-focused view towards direct, asset-level impact quantification.

The findings have direct implications for both coastal management and science. For coastal managers, the globally consistent datasets on coastal geomorphology, shoreline dynamics and exposure provide a first-order assessment for identifying potential erosion impacts, allowing them to prioritize attention and contextualize adaptation strategies. For coastal science, the research provides a scalable framework for applying modern AI to large satellite archives, highlights the need for multi-model approaches to project future shoreline change on other coastal types, and underscores the opportunity for community-driven efforts to create the standardized classification schemas and open ML training datasets that can train the next generations of "CoastalAI". Sustaining this progress depends on continued support for the open data policies, open-source software, and reliable cloud infrastruc-

ture that enable the community to build transparent, accessible, and intelligent analytical tools for better coastal management.

Samenvatting

Wereldwijd staan kusten onder toenemende druk door intensieve menselijke ontwikkelingen in dit gebied en klimaatverandering. Deze samenkomst creëert de noodzaak om beter te begrijpen hoe onder zeespiegelstijging de gevaren in de kustzone zullen veranderen. Eén van deze gevaren is kusterosie, een natuurlijk proces van aanpassing van de kustlijn, dat een gevaar kan worden waar het samenvalt met bestaansmiddelen, bezittingen, infrastructuur of zelfs ecosystemen waar die niet landinwaarts kunnen migreren. Om deze reden is een beter begrip van de gevolgen van kusterosie, met name onder het huidige klimaat van versnellende zeespiegelstijging, noodzakelijk om klimaatadaptatiebeleid te ontwikkelen. Dit proefschrift richt zich bij uitstek op deze uitdaging door een raamwerk te ontwikkelen dat aardobservatiesatellietdata, cloudtechnologie en kunstmatige intelligentie integreert om de kust in hoge (100 m) resolutie in kaart te brengen. De resulterende geomorfologische classificatie wordt vervolgens gebruikt om verfijnde probabilistische toekomstige kustlijnprojecties te produceren, wat een wereldwijde gevolgenanalyse van toekomstige kusterosie op gebouwen onder verschillende zeespiegelstijging-scenario's mogelijk maakt.

De hoofddoelstelling werd bereikt via een driedelige aanpak die stapsgewijs opbouwt van een technische basis naar de uiteindelijke gevolgenanalyse. Ten eerste ontwikkelde het onderzoek een schaalbaar, "cloud-native" analytisch raamwerk dat in staat is om satellietdataarchieven op petabyteschaal te verwerken. Vervolgens werd dit raamwerk gebruikt om de fysieke samenstelling van de wereldkust in kaart te brengen door diep-lerende neurale netwerken te gebruiken om aardobservatiedata te classificeren. Deze wereldwijde kusttypologie op 100 m verschafte de geomorfologische context voor de laatste stap, waarin deze componenten werden geïntegreerd om een eerst-orde gevolgenanalyse van toekomstige kusterosie uit te voeren. Voor deze analyse werd probabilistische toekomstige kustlijnverandering geprojecteerd met een evenwichtsprofielmodel, waarbij de typologie de toepassing ervan beperkte tot zandige sedimentvlaktes en duinkusten. Door deze geprojecteerde kustlijnposities te kruisen met wereldwijde gebouwdata, werd het aantal potentieel getroffen gebouwen gekwantificeerd.

Het onderzoek levert verschillende bevindingen op die de toestand van de wereldwijde kustzone karakteriseren. De nieuwe, door machine learning afgeleide kusttypologie onthult dat meer dan 60% van de ijsvrije kust in de wereld uit zachte, erodeerbare sedimenten bestaat. De analyse schat dat 40% van de wereldkust wordt geflankeerd door zand-, grind- of kiezelstranden, een hoger cijfer dan de huidige schattingen van 31%, waarbij het verschil grotendeels wordt toegeschreven aan het nieuwe, zonaal-onbevooroordeelde transectsysteem dat in dit proefschrift werd ontwikkelt. De analyse legt ook geomorfologische gevoeligheden bloot en toont aan dat 20% van deze stranden landwaards wordt begrensd door kliffen.

Deze configuratie in het dwarsprofiel kan hun natuurlijke vermogen om met zeespiegelstijging landinwaarts te migreren beperken, vooral waar kliffen bestaan uit resistente lithologieën; een cruciaal detail dat deze op satellieten gebaseerde aanpak nog niet kan oplossen. Wat betreft kustontwikkeling, toont de analyse aan dat een kwart van de wereld kustlijn ($\sim$253 000 km) ten minste één gebouw binnen de eerste kilometer land heeft, met in totaal ongeveer 76 miljoen gebouwen. Op basis van deze patronen introduceert het proefschrift de "Empirische Bebouwingsvrije Zone" (EBZ), een datagestuurde metriek die de feitelijke bebouwingsvrije zone kwantificeert die kustgemeenschappen impliciet aanhouden. Deze metriek toont aan dat voor 10% van de totale bebouwde kusten het eerste gebouw zich op minder dan 37 m van de kustlijn bevindt. Het combineren van deze analyses onthult dat kustontwikkeling disproportioneel geconcentreerd is op de meest dynamische kusten; zo is de bebouwingsdichtheid op zand-, grind- en kiezelstranden 354 gebouwen per kilometer kust, significant hoger dan het wereldwijde gemiddelde van 299.

Deze hoge huidige blootstelling vertaalt zich in significante toekomstige gevolgen wanneer deze wordt gecombineerd met projecties van toekomstige kusterosie. Er wordt geprojecteerd dat tegen 2100 tussen de 1,6 miljoen (in een lage-emissiescenario SSP1-2.6) en 2,4 miljoen (in een hoge-emissiescenario SSP5-8.5) gebouwen getroffen zullen worden door kusterosie. Hoewel dit een aanzienlijk aantal is, vertegenwoordigt het een conservatieve ondergrens, aangezien de toepassing van het model eerst werd beperkt door de op machinaal geleerde kusttypologie tot zandige sedimentvlaktes en duinkusten (21% van de wereld kustlijn), voordat databeperkingen (voornamelijk het gebrek aan betrouwbare data over de helling van de kust) het studiegebied verder reduceerden tot 9,3%. Het is echter cruciaal om de beperkingen van deze gevolgenanalyse te erkennen. De resultaten zijn onderhevig aan onzekerheden die voortkomen uit de onderliggende kusttypologie en de structurele aannames van het gebruikte evenwichtsprofiel model. Bovendien kwantificeert de analyse de gevolgen door het gevaar (toekomstige kusterosie) te combineren met de blootstelling (gebouwdata), maar is het geen volledige klimaatrisicoanalyse. Een dergelijke analyse houdt doorgaans rekening met de volledige waarschijnlijkheidsverdeling en zou tevens een complete analyse van de kwetsbaarheid vereisen, die op haar beurt niet alleen de gevoeligheid van een systeem voor kusterosie omvat, maar ook het adaptief vermogen: het sociaaleconomische vermogen om met dit gevaar om te gaan en erop te reageren.

De wetenschappelijke bijdragen van dit proefschrift zijn tweeledig. Methodologisch demonstreert het onderzoek een schaalbaar, "cloud-native" raamwerk voor de toepassing van moderne kunstmatige intelligentie op satellietarchieven van petabyteschaal, waarmee de kloof tussen gedetailleerde lokale studies en grofmazige mondiale analyses wordt overbrugd. Thematisch werd dit raamwerk gebruikt om een reeks wereldwijd consistente dataproducten te produceren. Deze omvatten een nieuw "Global Coastal Transect System", een kusttypologie op hoge resolutie afgeleid door middel van diep lerende neurale netwerken, verschillende metrieken voor kustblootstelling en probabilistische toekomstige kustlijnprojecties. Gezamenlijk maken deze componenten een eerst-orde gevolgenanalyse van toekomstige kusterosie op individuele gebouwen mogelijk, waarmee het vakgebied wordt vooruit-

geholpen van een op gevaren gerichte benadering naar directe kwantificering van de gevolgen op asset-niveau.

De bevindingen hebben directe implicaties voor zowel het kustbeheer als de wetenschap. Voor kustbeheerders bieden de wereldwijd consistente datasets over kustgeomorfologie, kustlijndynamiek en kustblootstelling een eerst-orde beoordeling voor het identificeren van potentiële erosiegevolgen, waardoor zij aandacht kunnen prioriteren en adaptatiestrategieën kunnen ontwikkelen. Voor de kustwetenschap biedt het onderzoek een schaalbaar raamwerk voor de toepassing van moderne kunstmatige intelligentie op grote satellietarchieven, benadrukt het de noodzaak van multi-model opzet om ook toekomstige kustlijnveranderingen op andere kusttypen te projecteren, en onderstreept het de kans voor gemeenschapsgestuurde inspanningen om gestandaardiseerde classificatieschema's en open dataset voor machine learning te creëren die de volgende generaties van "CoastalAI" kunnen trainen. Het bestendigen van deze vooruitgang is afhankelijk van voortdurende steun voor open data, open-source software en een betrouwbare cloudinfrastructuur, die de gemeenschap in staat stellen de transparante, toegankelijke en intelligente analytische instrumenten te bouwen die nodig zijn voor een beter toekomst bestendig kustbeheer.

Foreword

Life is thought to have emerged along the coast more than three billion years ago. Back then, Earth was a harsh environment, but it was also a place where interactions between land and sea created the dynamic conditions in which oxygen-producing microbial communities evolved. These so-called "stromatolites"—structures formed by layered mats of cyanobacteria—are so resilient that they have survived all major extinction events and still exist today, for example at Shark Bay in Western Australia. This dynamic nature, which was the driving force behind the earliest-known life-form, remains *the* defining characteristic of coastal environments. This is reflected in the commonly used definition of the coast as "the interface between land and sea, where oceanographic, terrestrial and atmospheric processes interact." In other words, it is the part of land influenced by marine processes and the part of the sea influenced by terrestrial processes.

What do those immense timescales have to do with sand grains and sea-level rise? Everything in a way. Adopting a geological perspective provides a temporally extensive understanding of how Earth changes through time. Earth is a enormous recycling system. Rocks form deep within the planet, cool and solidify, are uplifted to the surface, weathered, eroded, and eventually re-deposited to form new rocks on timescales of millions of years. Although Earth has existed for billions of years, the mountains and landscapes we see today are only a snapshot in its geological history. Geologist learn to see and unravel Earth's history by studying rocks and landscapes. It is a profound way of seeing Earth because the geological timescale is far beyond our natural human perception. It is a way of seeing that can be trained, but may, especially at first, require some imagination or a geological excursion in the mountains.

A geologist might look at a world map and notice how Africa and South America fit together like pieces of a puzzle, or how India is colliding with Asia to form the Himalaya—the highest mountains on Earth. Visiting Iceland, a geologist may step from the Eurasian Plate onto the North American one, while imagining how the island has grown over millions of years as the plates drift apart and magma wells up to form new land. The process that drives the movement of entire continents is known as plate tectonics: the theory that Earth's outer shell is divided into several rigid plates that move relative to each other at roughly the speed of fingernails grow (between one and ten centimeter per year). To humans this seems slow, but over millions of years it adds up to thousands of kilometers—the distance that now separates Africa and South America. The interaction of these plates shapes Earth's surface, creating deep ocean trenches that can reach 11 kilometers below sea level and raising the highest mountains on land. A simple fact illustrates how Earth changes over geological time: the summit of Mount Everest is composed of marine limestone; showing that

the highest point on Earth was once part of the ocean floor. Today, those ancient seabed rocks are being eroded by glaciers, broken down into finer sediments that rivers such as the Ganges carry back to the sea, where they eventually accumulate on the ocean floor and gradually lithify into new rocks. Seen in this geological light, studying Earth's changing landscapes reminds us that change is natural, that landforms evolve, and that geology provides the foundation on which all other Earth processes unfold.

Although coastal scientists typically study processes acting on much shorter timescales than geological time, they too have fascinating ways of seeing coasts. A coastal scientist understands sandy beaches as rivers of sand; systems where sediment is constantly in motion. Material eroded from mountains such as the Alps is carried by rivers like the Rhine to the sea, where it accumulates along coasts such as the Holland Coast. Tides regulate the exchange of sediment from rivers to the sea, while waves can travel thousands of kilometers from distant storms, delivering energy to the shoreline. Waves approaching the coast drive sediment alongshore, while tides and currents redistribute it across the nearshore zone. Sediment is retained on the continental shelf, where waves continually resuspend and push it landward, while gravity draws it seaward again—a process sometimes informally described by coastal engineers as the "bulldozing effect of waves." Some sediment, lifted by the wind, returns to land to form dunes, to be eroded again by a storm, in a perpetual cycle of movement that defines coastal environments. So, the coast we observe today is provided by its geological setting but continuously shaped by wind, waves, tides and climate, with as a consequence, much variety in coasts.

Thus, the coast is a naturally dynamic and complex zone where atmospheric, oceanographic, and terrestrial processes interact to shape landforms that have developed over geological timescales. This complexity has long supported life. As established, life itself may have originated in the coastal zone, and today these regions host some of the planet's most diverse ecosystems with very high biological productivity. Coasts have also attracted humans for millennia—first for their abundance of food and natural resources, later for strategic and logistical purposes, and more recently for recreation. When tides recede, coastal flats and estuaries give way to an abundance of shellfish and other food, cultivating early fishing and gathering. The accessibility of the sea likely encouraged innovations such as boats and fishing nets. Sheltered bays provided natural harbors where ships could anchor and trade with distant civilizations. Throughout history, the coast has acted as a gateway between the hinterland and the wider world, facilitating not only the exchange of goods but also of ideas. In Ancient Greece, maritime trade and naval power helped coastal cities such as Athens become centers of wealth, technological advancement, and cultural exchange—developments often associated with the invention of democracy. As long as the Athenian Empire controlled the sea, these civilizations flourished, but when it lost its naval dominance democracy collapsed with it. Today, in a globalized world, about 80% of international trade occurs by sea, with ships navigating between ports globally. The closure of the Suez Canal has been estimated to cost one to ten billion dollars per day. Beyond trade, coasts continue to attract human settlement for their favorable climates: temperate, mild conditions modulated by the thermal capacity of nearby oceans.

Thus, we now see a highly dynamic and complex zone that provides abundant resources and has attracted humans since prehistory. However, it is no longer only natural processes that shape the coastal zone; human activities increasingly modify it as well. Since the Industrial Revolution (or modern capitalism), greenhouse gas emissions have accelerated climate change, causing sea level to rise. Sea levels have always fluctuated, rising by more than 100 m since the last glacial maximum, with the rise coming to a stillstand in past millennia, yet today they are rising faster than at any time in recorded history, with rates continuing to accelerate. Storms are becoming more intense in many regions, and extreme coastal events are occurring more frequently. At the same time, more people live in the coastal zone than ever before. These two compounding trends—rising sea levels and growing human exposure—place the world's coasts under increasing pressure. In effect, life in the coastal zone is becoming harder, not only for humans but also for species and entire ecosystems, as space becomes limited and adaptation to a changing climate is required. Climate change is also exacerbating many natural hazards, including tropical cyclones (such as Katrina), storm surges, saltwater intrusion, and high-tide flooding. In this work, we focus on one of the hazards: coastal erosion. Coastal erosion is the permanent loss of land to the sea. It is a natural process—"Erosion happens," as they say in the Grand Canyon, and beaches exist because of it. However, this natural process becomes a hazard where humans have settled close to the shore and it threatens things of value, such as assets, infrastructure, livelihoods, or ecosystems.

In this thesis, we study where coastal erosion can become problem for people living in coastal areas worldwide. We do so by mapping the geomorphological composition of the coast globally and relating it to how beaches have changed over the past forty years. These historical trends are extrapolated into the future and combined with projections of sea-level-rise-induced retreat to derive probabilistic estimates of future coastal erosion under different sea-level-rise scenarios. These projections of future shoreline positions are associated with present-day building exposure to assess the impacts of future coastal erosion.

Figure 1: The Mount Everest illustrates how Earth changes over geological time. The summit of Mount Everest (third peak from the left) is composed of marine limestone, showing that the highest point on Earth was once part of the ocean floor. Photo: Floris Calkoen.

Part I

Core Chapters

1 Introduction

1.1 The Coastal Zone Under Pressure

The coast is a dynamic transitional zone where terrestrial, marine, and atmospheric processes interact, shaping a diverse geomorphology inherited from a geological past (Wright et al. 1984; Carter 1988, Ch. 1; Woodroffe 2002, Ch. 1; Masselink 2014, Ch. 1) . It supports highly productive ecosystems and has historically attracted human settlement for its rich resources and its central role in trade, transport, and culture (Burke et al. 2001; Martínez et al. 2007; Barbier et al. 2011). Coastal systems are naturally resilient, having adapted to significant environmental changes, including more than 120 m of sea-level rise (SLR) since the last glacial maximum about 20 thousand years ago (Kraft et al. 1985; Carter et al. 1994; Lambeck et al. 2010; Woodroffe et al. 2012). Today, however, they face direct pressure from human development. Coastal zones are often more densely populated than inland areas (Small et al. 2003; McGranahan et al. 2007; Cosby et al. 2024) and are experiencing higher rates of population growth and urbanization—a trend that is projected to continue (Neumann et al. 2015; Jones et al. 2016). This creates a range of local, non-climate pressures, including the disruption of sediment budgets through river damming, sand mining, and land subsidence (Slagel et al. 2008; Nicholls et al. 2021; Rangel-Buitrago et al. 2023), while coastal armoring and land reclamation fix the coastline in place, disrupting natural processes (Dugan et al. 2008; Bulleri et al. 2010; Wright et al. 2019).

The second major pressure, and the one most relevant to this work, is climate change, which compounds these direct impacts by introducing new stressors and intensifying existing coastal hazards (Oppenheimer et al. 2019). While climate-change-induced stressors like ocean acidification and marine heatwaves directly degrade coastal ecosystems (e.g., Smith et al. 2025), the primary driver of physical change is accelerating global mean SLR (Nerem et al. 2018; Dangendorf et al. 2019). Observed rates have increased from 1.3 mm/yr (1901-1971) to 3.7 mm/yr (2006-2018) (Fox-Kemper et al. 2021). This rise exacerbates multiple hazards, including coastal flooding, groundwater inundation, saltwater intrusion, and coastal erosion (Nicholls 2010; Fox-Kemper et al. 2021; Ranasinghe et al. 2021). Moreover, due to thermal expansion and ice sheet dynamics, sea level is committed to rise for centuries, regardless of future emission scenarios (Calvin et al. 2023, Sec. B.7.2).

The convergence of these two pressures—a highly developed coastal zone and a rising, more hazardous ocean—creates a phenomenon known as "coastal narrowing" (or more specifically "coastal squeeze" when affecting ecosystems) (Doody 2004; Pontee 2013; Lansu et al. 2024). In this "narrowing" state, natural habitats and human systems are trapped, with

ecosystems unable to migrate inland and communities having less room to adapt. This degrades ecosystems, reduces the coast's natural defensive capacity, threatens biodiversity, and makes human adaptation more difficult and costly (Defeo et al. 2009; European Environment Agency. 2024). The result is a coastal zone under unprecedented stress, with climate change set to increase this pressure over the coming centuries.

While adaptation to these pressures is best addressed locally, the scale of the drivers also benefits from a global scientific perspective. A global analysis provides a consistent framework to compare coastal regions, identify hotspots that may be invisible at local level, and facilitates cross-learning. This comparative evidence can inform international policy frameworks, such as the Intergovernmental Panel on Climate Change (IPCC) assessments, which help prioritize and allocate adaptation funds (Reisinger et al. 2020). In this way, a global-scale analyses contribute to effective coastal climate adaptation planning.

1.2 Coastal Erosion: From Natural Process to Risk

1.2.1 Defining the Hazard: Chronic and Episodic Erosion

The shoreline is the dynamic interface between land and sea, its position constantly varying across multiple time and space scales due changes in external forcing conditions like waves, tides and wind (Wright et al. 1977; Stive et al. 2002). One of the most studied coastal processes is coastal erosion. As a natural physical process, coastal erosion is the net removal of sediment by hydrodynamic and/or aeloian forces, which distinguishes it from the simple submergence of land due to rising water levels. It is useful to distinguish between two main types: *episodic erosion* (Fig. 1.1), a rapid but often temporary loss of land caused by individual storms from which a beach may naturally recover (e.g., Dodet et al. 2019), and *chronic erosion*, a persistent, long-term trend of land loss (Hinkel et al. 2013). It is this chronic erosion that is often used interchangeably with the term "shoreline retreat", especially when describing the slow, continuous landward movement of the coastline caused by long-term drivers like SLR.

The drivers of chronic erosion directly mirror the two confounding pressures on the coastal zone. The first is the direct human influence on coastal sediment budgets. Activities such as river damming, sand mining, and the construction of coastal infrastructure can starve beaches of their natural sediment supply and directly cause or increase long-term erosion (Slagel et al. 2008; Hapke et al. 2013; Pranzini et al. 2015). As a result of these human influences, chronic erosion has become a widespread global phenomenon, as documented in several regional and global assessments over the past decades (Bird 1985; Salman et al. 2004; Luijendijk et al. 2018; Mentaschi et al. 2018). The second driver is SLR (Bruun 1962). While not a direct erosive force itself, SLR *enables* (Zhang et al. 2004) chronic erosion by allowing the forcing to act farther up the beach profile. The acceleration of SLR is therefore expected to induce future shoreline retreat, which is a primary concern for coastal management.

Figure 1.1: Coastal erosion is a natural process of shoreline adjustment. This photo shows episodic beach erosion (21 December 2003) on the island of Ameland, the Netherlands. Unlike on developed coasts, where this process may threaten things of value, here it is part of the natural dynamics of the coastal system. Photo: © Rijkswaterstaat | Rob Jungcurt.

1.2.2 Equilibrium-profile Shoreline Projections

With SLR increasing pressure on a highly developed coastal zone, shoreline projections are becoming more relevant for climate adaptation planning. This is not a new concern, as the hazardous effects of SLR have been recognized for decades (Mercer 1978; Gornitz 1991; Nicholls et al. 2025). However, making projections on decadal-to-century timescales is challenging; this is precisely the "centennial gap" where predictive understanding is weakest, and there is no scientific consensus on which models to apply (Woodroffe et al. 2012). Current approaches include numerical models (that resolve physical equations), data-driven models (statistical methods that drive on observations), and simplified process-based models (that focus on one particular process, e.g., SLR) (Vitousek et al. 2017), with each approach facing its own significant challenges. Numerical models are computationally intensive and often require detailed local data that are often unavailable globally (Ranasinghe 2020). Conversely, simply extrapolating historical trends (using statistical (machine-learning) methods) is insufficient, as this cannot account for the non-linear effects of accelerating SLR. It is for these reasons that broad-scale assessments (e.g., Vousdoukas et al. 2020b) have combined historical shoreline-change trends with a simple equilibrium-profile model (the Bruun rule) (Bruun 1962).

The Bruun rule provides a direct link between (relative) SLR and the retreat of sandy shores. Its mechanism says that a sandy equilibrium profile moves upward and landward in

response to (relative) SLR. This is achieved by eroding the upper beach and depositing the sediment on the nearshore seafloor, thereby conserving sediment volume across the active profile. The canonical form of the rule is given by:

$$R = S \times \frac{L}{B + h} \tag{1.1}$$

where R is the shoreline retreat, S is the amount of SLR, L is the length of the active profile, B is the berm height, and h is the depth of closure. The fraction represents the inverse of the active profile slope. The model's use is the subject of a long-standing scientific discussion, as its assumptions, such as a closed, two-dimensional sediment system and a fully erodible sandy profile, are rarely all met in nature (e.g., SCOR Working Group 89 1991; Cooper et al. 2004; Zhang et al. 2004). Even so, for broad-scale assessments, the equilibrium-profile concept offers an important conceptual improvement over a simple "bathtub" approach (e.g., Hoozemans et al. 1993) because it incorporates the fundamental physical principle that the entire active profile responds to SLR.

Several studies have applied this equilibrium-profile concept to project future shoreline change. Initial global estimates used a simplified application ($R = 100 \times S$) of the rule to project land loss (Hinkel et al. 2013). Subsequent assessments introduced methodological refinements, such as using spatially varying nearshore slopes and combining the SLR-induced retreat with extrapolated historical trends to account other processes that affect the sediment budgets, including alongshore sediment transport gradients and anthropogenic influences (Athanasiou et al. 2020; Vousdoukas et al. 2020b). A common gap in these assessments, however, is that they do not explicitly incorporate the geomorphological context of the backshore (e.g., erodible dunes vs. hard, resistant cliffs). This is an important factor for further refined application of the Bruun model, which is predicated on the assumption of a fully erodible sandy profile, and addressing this gap is a key focus of this work.

1.2.3 The Role of Coastal Geomorphology

A coastline's response to drivers like SLR is governed by its physical form and composition, which dictates how hydrodynamic energy is dissipated and sediment is mobilized, from resistant rock cliffs to mobile sandy beaches (Woodroffe 2002; Masselink 2014). Understanding this relationship requires a systematic coastal classification. Here, it is useful to distinguish a classification from a taxonomy: whereas a taxonomy is a more general system of organization, a classification is a technical grouping created for a specific purpose (Fairbridge 2004; Finkl 2004). For this thesis, that purpose is to inform an assessment of future coastal erosion impacts and, more broadly, coastal adaptation planning in the face of accelerating SLR.

Global coastal classification has a long history, with numerous schemes proposed to characterize the diversity of coastal landforms (e.g., McGill 1958). Early classifications were often based on large-scale tectonic settings (Inman et al. 1971). Other schemes have focused on hydrodynamic controls, such as the relative influence of waves and tides (Davies 1964), the relationship between climatic zones and sediment type (Hayes 1967), or the sedimentologi-

cal response to these forces (Boyd et al. 1992). Together, these approaches show that a coast's character is a product of its geological inheritance, material composition, and the processes that construct and erode it.

Developing a consistent coastal typology that functions at a broad scale while retaining the high resolution needed for local analysis has remained a challenge. Pioneering projects like EUROSION provided a structured overview for shoreline management at a continental scale but faced challenges with integrating heterogeneous national data (Salman et al. 2004). Similarly, the DIVA project used a systematic segmentation for global SLR assessments, though its classification was necessarily coarse given the data available at the time (Vafeidis et al. 2008; Hinkel et al. 2013). More recent global classifications have leveraged satellite data to map specific sediment types, such as sandy (Luijendijk et al. 2018) or muddy coasts (Hulskamp et al. 2023). While providing a higher-resolution foundation, a further step is needed to differentiate the geomorphological settings in which these sediments occur. Some studies have used machine learning (ML) algorithms to study geomorphology (Dang et al. 2020; Mao et al. 2022), but these efforts have been limited either in thematic detail or geographic scope. A common opportunity for advancement across these earlier efforts is to expand the focus from the shoreline itself to include the backshore geomorphology (e.g., dunes, cliffs, wetlands), which is essential for predicting how the coast will respond to SLR and how much migration space is available.

1.2.4 Assessing the Impacts of Coastal Erosion

To translate physical projections of shoreline change into meaningful information that can inform large-scale adaptation planning, a formal framework is needed. This work is guided by the IPCC climate risk framework, which defines risk as the potential for adverse consequences for human or ecological systems (Reisinger et al. 2020). In this framework, risk arises from the interaction of three components: the *hazard* (the climate-related physical event, such as shoreline retreat), the *exposure* of human or ecological systems to that hazard, and the *vulnerability* of the exposed system. Vulnerability itself is a function of the system's sensitivity or susceptibility to harm and its lack of capacity to cope and adapt. This relationship, often expressed as:

$$\text{Risk} = \text{Hazard} \times \text{Exposure} \times \text{Vulnerability} \tag{1.2}$$

moves the assessment beyond purely physical phenomena to consider where and how those phenomena matter to society and ecosystems.

Applying this conceptual framework to earlier studies, the assessment by Hinkel et al. (2013) provided an early, comprehensive impact analysis by coupling erosion projections with socio-economic exposure, quantifying consequences such as land loss and forced migration. More recently, broad-scale assessments have built upon satellite-derived data to produce hazard projections at a much higher spatial resolution. The global study by Vousdoukas et al. (2020b) is an example of a state-of-the-art, global hazard assessment, while work

by Athanasiou et al. (2020) for Europe also quantified exposure in terms of land area loss. Building on these advances, this thesis conducts a first-order impact assessment, which quantifies the potential consequences by combining the hazard (future coastal erosion) with the exposure of specific assets (e.g., buildings). By linking the physical hazard to specific societal assets, this method moves the assessment beyond a purely physical hazard focus to consider where and how those hazards matter. This focus on impacts is important because coastal erosion, as a natural process, only becomes a direct societal concern where it intersects with something of value (Fig. 1.2).

Figure 1.2: Coastal erosion becomes a direct societal hazard on developed coastlines. This photo from Ecuador, shows how coastal erosion may undermine coastal infrastructure and threaten coastal developments. While erosion is a natural process, it transforms into a coastal hazard where it threatens something of value. Photo: Arjen Luijendijk.

1.3 The Digital Transformation of Coastal Science

Coastal science is undergoing a rapid digital transformation, driven by a set of innovations that have redefined what is possible at a global scale. This shift began with the opening of multi-decadal satellite data archives, which provided an unprecedented, consistent record of Earth's surface (Wulder et al. 2012; Wulder et al. 2022). The sheer volume of this data has made traditional "download-and-analyze" workflows impractical. The solution emerged from new data processing paradigms (Dean et al. 2008) that were incorporated into user-friendly, data-proximate cloud computing platforms that collocate massive datasets with analysis tools (Gorelick et al. 2017). Accelerated by a culture of open-source software de-

velopment (Raymond 1999), this new ecosystem revolutionized geoscience, enabling the creation of the first high-resolution global maps of phenomena like forest cover change (Hansen et al. 2013). These innovations soon made their way into other domains like surface water monitoring (Donchyts et al. 2016a; Pekel et al. 2016), until the potential of satellite data for coastal science was demonstrated with global applications in shoreline monitoring and tidal flats mapping (Luijendijk et al. 2018; Murray et al. 2018), establishing earth observation (EO) as standard practice among coastal practitioners and scientists (Vos et al. 2023a).

However, it is useful to characterize the methods that were central to this first wave of transformation. While these early studies commonly leveraged machine learning, the algorithms used were primarily traditional methods developed from the 1960s to 1980s, such as decision trees (Breiman et al. 1984), later widely adopted in implementations like Random Forests and Support Vector Machines (Cortes et al. 1995; Breiman 2001). It is helpful to distinguish these methods from the deep learning (DL) models that define the current era of artificial intelligence (AI) (Mitchell 1997; Goodfellow et al. 2016). This modern approach, where foundational work on neural network architectures—from early work on backpropagation and recurrent networks (Rumelhart et al. 1986; Hochreiter et al. 1997) to the development of transformer models (Vaswani et al. 2023)—combined with massive datasets and parallel computing (Krizhevsky et al. 2012; Le 2013), has led to significant advances in fields from protein structure prediction (Jumper et al. 2021) to weather forecasting (e.g., Lam et al. 2023). Recognizing this distinction helps to reveal that the digital revolution in Earth System science since the 2010s was primarily facilitated by new modes of data access and computation, and *not* by the widespread adoption of modern AI. This raises a question: given the demonstrated potential of modern AI in other scientific domains, why have its applications in coastal science been comparatively limited?

Despite its immense potential (e.g., Brown et al. 2022; Jakubik et al. 2023), the broad adoption of deep learning in coastal science has been slower than might be expected, presenting a clear frontier for research. The reasons are multifaceted and rooted in the unique nature of geospatial data and analysis (Zhu et al. 2017; Reichstein et al. 2019; Tuia et al. 2025). Mainstream AI success was built on standardized data, like RGB images, whereas satellite data is fundamentally different: it is multi-spectral, georeferenced to a sphere, and stored in petabyte-scale archives that are too large to download. This creates profound challenges for data handling, spatial sampling, and the creation of large, high-quality training datasets, for which convenient tools are often lacking. Furthermore, coastal AI requires a rare combination of expertise: a deep understanding of the natural science problem, familiarity with big data and cloud technology, and specialized knowledge of AI. While platforms like Google Earth Engine (GEE) provided a revolutionary toolbox for Earth scientists to handle immense raster datasets, their architecture is not optimized for the flexible, iterative workflows required for deep learning (Abernathey et al. 2021; Tuia et al. 2025). Consequently, to leverage latest innovations from AI, a more flexible ecosystem is helpful. This is probably one of the reasons that has led to the growth of open-source communities like Pangeo (Hamman et al. 2018), which are building the interoperable software ecosystems needed for scalable, geo-analytics. Assembling this infrastructure—from data access and scalable compute to

integrated deep learning libraries for both raster and vector data—is a central challenge addressed in this thesis.

1.4 Research Objective

The combined pressures of accelerating climate change and expanding coastal development are placing coastal zones under increasing stress. Coastal erosion, though a natural process, becomes a societal concern where it intersects with something of value. Addressing this challenge at macro-policy level requires a global perspective, which has only recently become feasible through the convergence of open, multi-decadal satellite data archives, cloud technology, and artificial intelligence.

> The primary objective of this thesis is to conduct a global assessment of the impacts of future coastal erosion under different sea-level rise projections by analyzing the vast observational datasets now available from satellites, using the power of cloud technology and the capabilities of artificial intelligence.

To achieve this objective, the following research questions are formulated:

RQ 1: **How can the combination of satellite data, cloud technology, and artificial intelligence enable the study of local coastal phenomena at a global scale?**
This research component lays the technical and methodological foundation for the thesis. It focuses on developing scalable, interoperable workflows that integrate Earth Observation data, cloud technology, and artificial intelligence to support the global analysis of complex, local-scale coastal phenomena.

RQ 2: **What is the global distribution of coastal geomorphological types?**
This research component focuses on mapping the geomorphological composition of the world's coast by applying deep learning, a form of artificial intelligence, to satellite imagery.

RQ 3: **What are the impacts of future coastal erosion under different sea-level rise scenarios?**
This final research component provides an illustrative application of the thesis framework by conducting an impact-focused assessment of future coastal erosion. The approach synthesizes the preceding work by producing probabilistic shoreline projections (the hazard) refined by the coastal typology with global building footprint data (the exposure) to assess potential impacts on the built environment.

1.5 Structure of the Thesis

This thesis is organized into two main parts. Part I presents the six core chapters (Fig. 1.3) that progressively build towards answering the primary research objective. Following this

introduction, Chapter 2 establishes the technical foundation for the thesis by developing a scalable, cloud-native framework for global coastal analysis. Chapter 3 applies this framework to create a novel, global coastal typology, providing the essential geomorphological context for the impact assessment. Chapter 4 then provides an illustrative application of these components to conduct a global assessment of the future impacts of coastal erosion under various sea-level rise scenarios. Subsequently, Chapter 5 synthesizes the key scientific contributions and offers a perspective on the future of global coastal science. This first part concludes with Chapter 6, which presents the principal conclusions of the thesis by directly answering the research questions. Part II contains the appendices, which provide supplementary material, including more detailed descriptions of the data and software products, additional results, and validation analyses that support the findings in the core chapters.

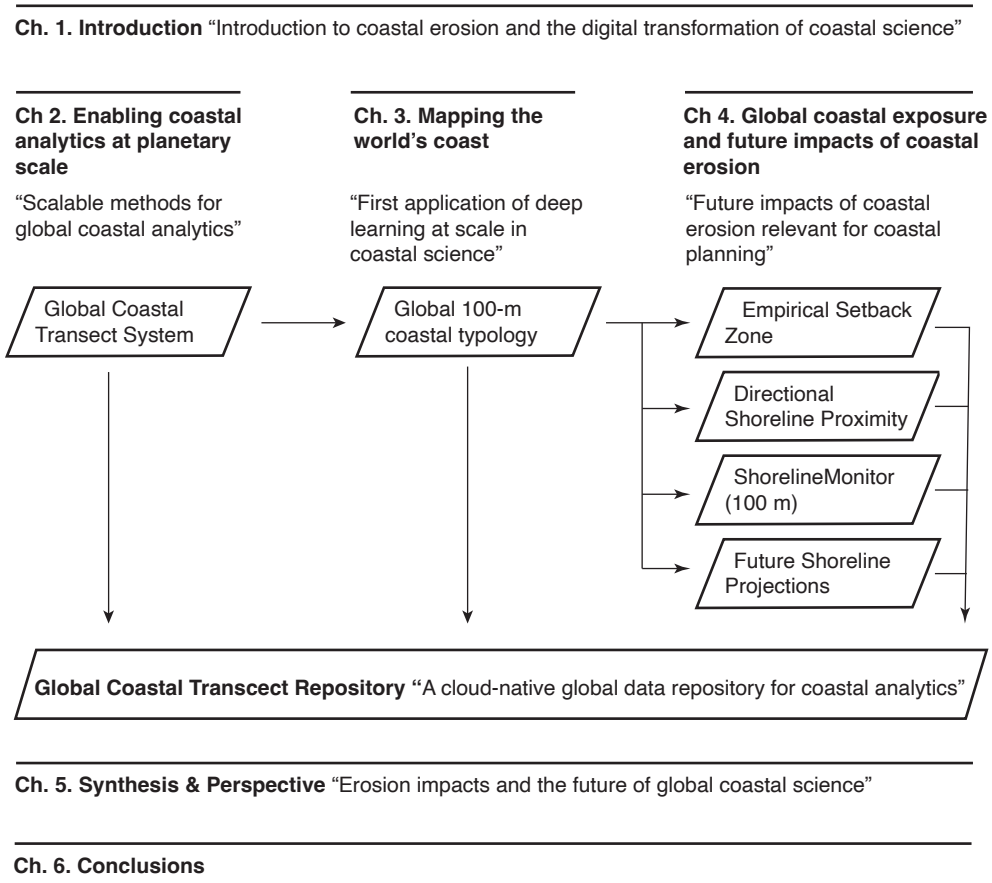

Figure 1.3: Structure of the thesis. This flowchart illustrates the relationship between the core research chapters and the key datasets they produce (parallelograms). The research progresses from establishing a scalable analytical framework (Ch. 2) to creating a global coastal typology (Ch. 3), which provides the foundation for the worldwide erosion impact assessment (Ch. 4). The thesis concludes with a synthesis of the findings and a perspective on the future of global coastal science (Ch. 5), followed by the final conclusions (Ch. 6).

2 Enabling Coastal Analytics at Planetary Scale

Abstract Coastal science has entered a new era of data-driven research, facilitated by satellite data and cloud computing. Despite its potential, the coastal community has yet to fully capitalize on these advancements due to a lack of tailored data, tools, and models. This paper demonstrates how cloud technology can advance coastal analytics at scale. We introduce GCTS, a novel foundational dataset comprising over 11 million coastal transects at 100-m resolution. Our experiments highlight the importance of cloud-optimized data formats, geospatial sorting, and metadata-driven data retrieval. By leveraging cloud technology, we achieve up to 700 times faster performance for tasks like coastal waterline mapping. A case study reveals that 33% of the world's first kilometer of coast is below 5 m, with the entire analysis completed in a few hours. Our findings make a compelling case for the coastal community to start producing data, tools, and models suitable for scalable coastal analytics.

This chapter is based on the following article:

Calkoen, F. R., Luijendijk, A. P., Vos, K., Kras, E., & Baart, F. (2025). Enabling coastal analytics at planetary scale. *Environmental Modelling & Software*, *183*, 106257. https://doi.org/10.1016/j.envsoft.2024.106257

2.1 Introduction

Coastal science has entered a new era of data-driven research (Vitousek et al. 2023a), facilitated by the opening up of historical satellite data catalogs (Wulder et al. 2022) and advances in data processing (Dean et al. 2008), which have been integrated into cloud-based geospatial data analysis platforms (e.g., Gorelick et al. 2017). In coastal science, the potential of Earth-observing satellite data was showed by global analyses to coastal change (e.g., Luijendijk et al. 2018; Murray et al. 2018) and since then the coast is studied at increasing detail using satellite-derived data products (e.g., Warrick et al. 2023).

Currently we can distinguish between two distinct analysis strategies: one aims for global coverage, typically compromising accuracy for spatial extent ("everywhere"), while the other prioritizes accuracy ("anywhere"). Although not all coastal analyses have to be run at broad spatial scales (e.g., Mikkelsen et al. 2024), there are several reasons for why we need coastal analyses at extensive spatial scales. Particularly since the advent of geospatial information system (GIS) in coastal science, it has been acknowledged that analyses at scale facilitate integrated and systematic approaches to coastal classification (Finkl 2004). It also supports development of diverse coastal management strategies at varying spatial scales (Cooper et al. 1998). Moreover, while details and planning of coastal management often happen at local or regional levels, they are supported by legislation at national or international levels (Wong et al. 2014). Finally, analyses at scale facilitate intercomparisons between different regions, enabling peer-to-peer learning, where local coastal management practices that have been adopted successfully can be shared.

The two distinct analysis strategies ("everywhere" vs "anywhere") can be illustrated by contrasting some coastal monitoring tools in more detail. On one hand, analyses at scale ("everywhere") typically use a cloud platform to process petabyte-scale satellite data catalogs by condensing stacks of individual imagery into composites (e.g., Luijendijk et al. 2018; Mao et al. 2021) or cloud-free mosaics (e.g., Hulskamp et al. 2023). These approaches rely on methods that are available on the cloud platform, so that they can be incorporated in server-side compute. While such strategies manage large volumes of data by efficiently processing data in close proximity of where it is stored, it inherently limits the temporal depth and/or restricts the analysis methods to those available on the cloud platform. On the other hand, approaches that process each image in the historical catalog using more sophisticated processing routines or algorithms (e.g., Buscombe et al. 2018; Vos et al. 2019; Al Najar et al. 2023; Muir et al. 2024), have traditionally been confined to local studies or demanded substantial effort (e.g., Vitousek et al. 2023b; Vos et al. 2023b; Castelle et al. 2024). Such approaches typically involve downloading all data from the server to a machine with all required software to run the analysis installed. In summary, analyses focusing on larger geographic areas often sacrifice either temporal depth, algorithmic flexibility, or both, whereas those prioritizing more detailed analyses would face problems to scale up to larger areas if that is desired.

Historically, the importance of data-proximate, scalable compute on rasters (Baumann 1993) and the ability to query data onto a uniform grid for comparison (Cornillon et al. 2003) have been well acknowledged. These principles are now typically integrated into geospatial data processing platforms (e.g., Medvedev et al. 2016; Raoult et al. 2017). Later, Gorelick et al. (2017) have arguably revolutionized Earth science by building a geospatial data analytics platform on the public cloud, where users can write and execute code to efficiently perform more complex analyses on vast datasets by co-locating data and compute infrastructure as well as providing a high-level scripting environment. Despite its success, Google Earth Engine (GEE) remains a platform with limited flexibility and modularity (Abernathey et al. 2021) and nowadays there are also more open, flexible, community-driven efforts that aim to facilitate big environmental data analysis. Both Open Data Cube (ODC) (Killough 2018) and Pangeo (Abernathey et al. 2017) emphasize open-source development and collaboration,

with active communities across all continents. ODC focuses on providing a structured framework for managing and analyzing satellite data, while the aim of Pangeo is broader, with its community contributing to a coherent software ecosystem that aims to enable big data geoscience. ODC has particularly been successful in Australia, where Digital Earth Australia (DEA) (Gavin et al. 2018) have also enabled applications in coastal monitoring at scale (Bishop-Taylor et al. 2021). Although ODC is open, flexible and capable of handling large computations, it has primarily been implemented on high-performance clusters at a national level (Killough 2018). Both ODC and Pangeo leverage similar open-source software, but Pangeo, with its larger ecosystem, also supports cloud-based platforms[1], where one can conduct highly specialized analytics. These platforms are built on open, scalable software and can integrate diverse data sources like satellite imagery, climate data and other geospatial features, marking a shift towards more open, flexible and scalable geospatial data analytics.

We contend that if the coastal community wants to gain new insights into urgent coastal challenges at extensive scale, without compromising accuracy or spatio-temporal resolution, it will probably have to start producing tools, models, and data that are suitable for scalable analytics. In this paper, we share our experience in using cloud technology to advance coastal science at scale, effectively bridging the gap from "everywhere" towards "anywhere."

This paper is divided into three parts. The first part discusses data-proximate computing through the implementation of coastal waterline mapping in the cloud. It highlights the advantages of data-proximate computing and explores how a flexible software stack (Pangeo) facilitates scalable coastal analytics. The second part compares different data storage strategies for our novel global coastal transect system (GCTS), a foundational dataset, consisting of cross-shore coastal transect system at 100-m alongshore resolution, that is now made publicly available. The experiments demonstrate how cloud-optimized data formats are significantly more efficient than traditional file formats, making them critical for coastal analytics at scale. In this part we also highlight the importance of standardized metadata specifications. Finally, the third part demonstrates how flexible, scalable compute (part 1), combined with cloud-optimized data exposed through standardized metadata specifications (part 2), enable high-resolution coastal analytics (100-m alongshore) at planetary scale.

As a use case, we compute the percentage of land within the first kilometer of the coastal zone that is below 5 meters above mean sea level. With this work, we aim to show that coastal analytics can be performed at global scale, without compromising accuracy or spatio-temporal resolution, in very reasonable compute times. Can we make a compelling argument for the coastal community to start producing tools, models, and data that are suitable for scalable coastal analytics?

2.2 Methodology

In this section, we present our framework for conducting high-resolution, planetary-scale coastal analytics. We focus on two experiments: data-proximate coastal waterline mapping and strategies for cloud-native data release. Insights gained from these experiments inform a subsequent case study, which extracts elevation data over more than 11 million coastal transects, illustrating how scalable tools, models, and data can advance coastal science. Overall, insights from experiment 1 and experiment 2 enable the case study on coastal elevation mapping. In this section, the methods are described by topic, with some methods used in multiple experiments as well as in the case study. Figure 2.1 provides an overview of the described architecture, encompassing all methods detailed in this section.

[1] Notable examples include Microsoft Planetary Computer (MSPC), Coiled and Earthmover.

2.2.1 Global Coastal Transect System

Cross-shore coastal transects are essential to coastal monitoring, offering a consistent reference line to measure coastal change, while providing a robust foundation to map coastal characteristics and derive coastal statistics thereof. In this work, we introduce the GCTS, a novel foundational dataset compromising more than 11 million cross-shore coastal transects uniformly spaced at 100-m intervals along the shore. In comparison to previous efforts (Luijendijk et al. 2018; Bishop-Taylor et al. 2021; Vos et al. 2023b), this system has several advantages. The dataset has global coverage at 100-m alongshore resolution, for all OpenStreetMap (OSM) coastlines (80° S - 84° N) that are longer than 5 kilometers. We decided to use to define transects at 100-m alongshore resolution because this has shown to be effective for studying coastal dynamics at broad spatial scales (Bishop-Taylor et al. 2021; Vos et al. 2023b); it aligns well with the typical resolution of public satellite imagery (∼10-30 m); and, this resolution is also used in numerical modeling studies (Roelvink et al. 2020). The transect system is derived from a recent (2023-01-23) generalized OSM coastline[2], that was specifically prepared [3], with an optimal balance between smoothing and simplification (Hormann 2014) for a coastal cross-shore transect system at this 100-m alongshore resolution. Expert evaluations and visual comparisons with existing systems, such as the manually digitized CoastSat transects for the Pacific Basin (Vos et al. 2023b), confirm that zoom level 9 of the generalized OSM coastline provides the most accurate transects at this scale. Furthermore, using the most recent coastline data allows us to incorporate the latest crowd-sourced data from the OSM project. Finally, the transects are derived in their (Universal Transverse Mercator (UTM)) projection, a conformal projection that preserves angles locally, maintaining a uniform length of 2000 meters and a spacing of 100 meters apart alongshore, effectively correcting the zonal (latitude) distortions present in earlier global transect systems. GCTS is licensed under CC BY 4.0 licence, which means that you are free to share and adapt the dataset, as long as you give appropriate credit (i.e. cite this paper).

In this first release, we also add administrative boundaries[4] and a north bearing—the angle measured in degrees in a clockwise direction from the north pole. The data is released following best practices, as derived in experiment 2 (See Section 2.3.2) to facilitate convenient, efficient data retrieval, while ensuring that each chunk comfortably fits in the memory of a regular personal computer. The transects are stored in cloud-optimized GeoParquet format within a public cloud container [5] and are available as part of the Coastal Climate Core Services (CoCliCo) Spatio-Temporal Asset Catalog (STAC) catalog [6] under the collection ID "gcts".

2.2.2 Software stack

The methods we use to conduct coastal analytics at extensive spatial scales are deeply integrated with core-packages of the Pangeo stack for big environmental data analysis. The data processing routines employed typically utilize the data models Pandas (McKinney 2010) and Xarray (Hoyer et al. 2017), which, when integrated with geospatial extensions like GeoPandas and Rioxarray, enable advanced geospatial analysis of both raster and vector data, respectively. In coastal analytics, many of the data processing steps, such as computing spectral indices, can be run in parallel. To leverage the efficiency of parallel processing, we utilize Dask, an open-source library for parallel and distributed

[2] OSM data is available under the Open Database License (ODbL) at https://openstreetmap.org.
[3] The generalized coastline was produced by Imagico, DE.
[4] Administrative boundaries are extracted from Overture Maps.
[5] https://coclico.blob.core.windows.net/gcts
[6] https://coclico.blob.core.windows.net/stac/v1/catalog.json

computing (Rocklin 2015). Dask distributes computational tasks across multiple workers, which can also operate on different nodes. This distribution is centrally managed using computational graphs, which provide a structured representation of the tasks. Unlike traditional for loops, which execute sequentially, Dask's graph system allows tasks to be executed concurrently, optimizing performance and reducing computation time. We also use DuckDB, an embeddable analytical relational database management system (Raasveldt et al. 2019) to efficiently filter and retrieve tabular geospatial data from cloud object storage. To enhance the efficiency of distributed computing in coastal monitoring, we have refactored some routines from earlier coastal monitoring efforts (Luijendijk et al. 2018; Vos et al. 2019; Bishop-Taylor et al. 2021) into vectorized functions — operations applied simultaneously across entire arrays of data. This enables integration into the Dask data processing chain for more efficient parallel processing. Furthermore, by executing computations on a server nearby data storage and leveraging on-demand cloud compute infrastructure with tools like Dask Gateway or Dask Jobqueue, we can further reduce processing times by efficiently distributing tasks across a larger network of nodes. For example, when using a cloud compute infrastructure like MSPC, we are able to distribute tasks across more than 100 'workers' who collectively have access to approximately 800 GB of memory. In such distributed network, secondary data movement, such as coastal transects in the third experiment (See section 2.3.3) is minimized by strategically scattering it over the network, allowing it to be referenced through pointers (addresses to the data location) rather than sending the actual data.

2.2.3 Data retrieval

Data is searched for per area of interest, data range or other attributes using STAC, a geospatial data specification that nables efficient localization of spatio-temporal data collections through its standarized application programming interface (API). Data is retrieved from various STAC catalogs, including the MSPC and CoCliCo catalogs. The STAC collections are browsed, and relevant storage locations are parsed by the respective data models (See section 2.2.2). For tabular vector data, such as coastal transects, spatial joins or predicate pushdowns, a technique that filters data at the database level to improve query efficiency, on the bounding box attribute are employed to refine data selection. Raster data, such as satellite imagery, are lazily loaded, meaning that data is only loaded into memory when it is actually needed for computation. The data are read into Xarray using ODC-STAC, a Python library that is part of DEA software ecosystem. To minimize redundant earth observation (EO) data processing, imagery transmitted to multiple downlink stations or covering overlapping areas (Bauer-Marschallinger et al. 2023) on the same orbital ground track is grouped by solar day, selecting only the first occurrence of each day. Precise pixel alignment is achieved by centering pixels on the coordinate grid, ensuring all coordinate axes are anchored at pixel centers.

2.2.4 Data processing

To avoid memory issues and create a scalable, dynamic, parallel data processing strategy, larger areas are divided into a hierarchical grid using so-called quadtiles[7], a geo-data storage and indexing strategy. This strategy operates at specific zoom levels. For example, for processing Sentinel-2 (S-2) tiles, we subdivide them into quadtiles at zoom level 10, which corresponds to an area of approximately 0.35 degrees longitude around the equator. Data processing is performed on a per-quadtile basis, with computational graphs constructed and executed in a single compute call per quadtile. This approach prevents redundant recalculations of intermediate results and ensures a scalable data

[7] https://wiki.openstreetmap.org/wiki/QuadTiles

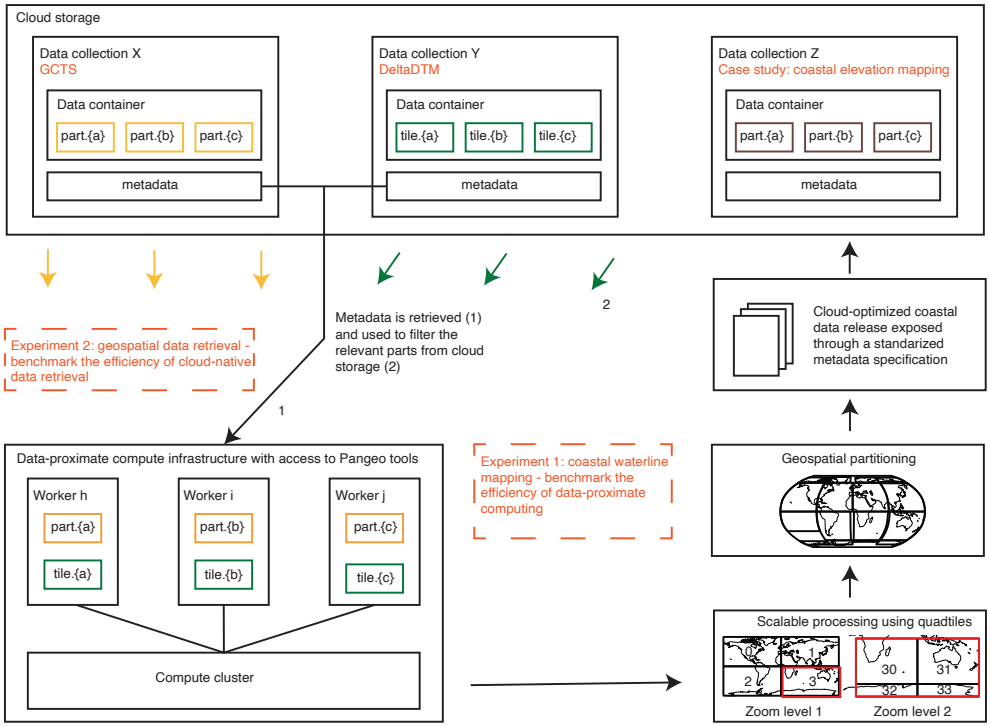

Figure 2.1: The workflow architecture for coastal analytics at extensive spatial scales. Data is stored in cloud object storage and exposed through standardized metadata specifications for efficient retrieval. Compute infrastructure is set up in close-proximity of the data, facilitating scalable data processing via hierarchical quadtiles. The results are stored in geospatial partitions in the cloud and also described following a standardized metadata specification to enable the next iteration of downstream coastal analytics. This study includes two primary experiments: geospatial data retrieval and data-proximate coastal waterline mapping. Insights from these experiments inform a case study demonstrating coastal analytics at a planetary scale by combining coastal transects (GCTS) with elevation data (DeltaDTM).

processing workflow. The chosen zoom level for the quadtiles (zoom level 10) ensures that arrays fit within memory limits, which is essential for complex processing workflows like coastal waterline mapping (See section 2.2.8). We typically also include coastline buffers into the processing pipeline to limit processing solely within our designated coastal region of interest. These buffers, are derived from the OSM coastline as of 2023-01-23, with a radius of 10 km, that is derived per UTM region. While processing, secondary data, such as these coastline buffers, are broadcast across the cluster so that this data is referenced through pointers (addresses to the data location) rather than transferring the actual data, further optimizing the processing efficiency.

2.2.5 Data partitioning

When storing data it is partitioned into manageable chunks, ranging from 100 to 200 MB for raster data and 500 MB for vector data, to enhance data interoperability for downstream coastal analytics. We spatially partition static coastal data using a strategy that compensates for coastline complexity, maintaining uniform partition sizes despite geographical variances in coastal geomorphology. The partitioning strategy begins by estimating the memory usage per geometry and its attributes. It then adds a quadkey (also known as a geohash), a geospatial index that encodes the location of an attribute in a standardized string, suitable for hierarchical binning at a given zoom level (e.g., level 12). Next, the data is sorted by quadkey. Finally, the data is recursively partitioned into chunks that do not exceed a set memory usage threshold. This approach ensures that local data density stays below a predefined partition size threshold. This method is exemplified in appendix B.2, where the GCTS is partitioned, for demonstration purposes, into chunks that do not exceed 100MB. It is evident that areas with higher data density, such as Chile, result in denser partitions. Finally, we typically assign a bounding box attribute with the minx, miny, maxx, and maxy coordinates of the geometry. This attribute is added as a structured datatype to enable query optimization techniques, such as predicate pushdowns, which filter data at the database level to improve query efficiency.

2.2.6 Data release

All data are stored in cloud-optimized formats, to ensure efficient retrieval (Durbin et al. 2020) from cloud object storage services that can serve large volumes of data. Geospatial vector data, such as coastal waterlines and cross-shore transects, are saved as compressed Apache GeoParquet files, a widely-used columnar storage format. Time series data, obtained per station, such as satellite-derived shoreline (SDS) series, are also stored in the GeoParquet format, with additional attributes that enable efficient geospatial data retrieval (See Section 2.2.3). Raster data, such as binary water-occurrence maps without a temporal component, are stored as cloud-optimized GeoTIFF (COG), whereas n-dimensional array's (such as climate data cubes or raster data with temporal component) are stored in Zarr. The data sets are described in STAC collections, that are added to the CoCliCo STAC catalog[8]. Here, each data partition or chunk is referenced as a STAC item. The catalog and its collections are maintained with open-source glsstac-utils[9] software, which utilizes specific extensions like 'proj', 'eo', 'datacube' and 'stac-table' to validate metadata against community standards. To further optimize accessibility, a GeoParquet snapshot of all STAC items is included as a collection asset, eliminating the need for users to index each JSON file individually from the collection.

[8] See https://coclico.blob.core.windows.net/stac/v1/catalog.json
[9] See https://github.com/stac-utils

2.2.7 Usage

All methods and workflows are bundled in CoastPy, a domain-specific Python package that contains tools for scalable, cloud-based coastal analytics, that, especially when combined with cloud computing, can efficiently process large amounts of data. CoastPy is available on the PyPI repository and is actively maintained at https://github.com/TUDelft-CITG/coastpy. It supports Python $\geq$ 3.11 and is compatible with macOS, Windows, and Linux. Detailed installation instructions and usage guidelines are provided in the README file and documentation. The repository includes tutorial notebooks that introduce straightforward applications of coastal analytics at scale. Released under the MIT License, CoastPy promotes open-source collaboration and development. We plan to expand the repository with tools for coastal machine learning (ML) and other advancements, and we welcome contributions from the community to this cloud-based coastal analytics initiative.

2.2.8 Experiments

To study the potential of cloud-native workflows for coastal analytics, we conducted several experiments focusing on data-proximate computing and data storage strategies. Experiment 1 is about coastal waterline mapping near the physical storage of EO data and experiment 2 benchmarks various cloud-optimized data access strategies using a standard geospatial data retrieval pattern. Additionally, in a subsequent case study, we demonstrate the potential of cloud-native workflows for coastal analytics at planetary scale by studying the distribution of elevation in the coastal zone. All experiments were performed using a standard Pangeo container, that a was run on MSPC, which provided access to a Dask Gateway cluster with adaptive scaling. The cluster was configured to provide 8 GB of RAM per Dask worker, with the number of workers dynamically adjusting between 2 and 100 based on processing workload and available compute. Dask acts as a higher-level tool that efficiently manages and distributes Python workflows across available compute resources. We emphasize that this a modular approach, where compute is separated from storage, that is not pertained to MSPC, but can be orchestrated at any cloud provider, (SLURM-based) high performance computing (HPC) cluster or even an ordinary personal computer.

Experiment 1: Coastal waterline mapping

The purpose of this experiment on coastal waterline mapping is to assess the efficiency of data-proximate cloud-computing for coastal monitoring by orchestrating coastal waterline mapping routines next to where the satellite data is stored. Given that shoreline monitoring has de-facto become standard practice in coastal science (Vos et al. 2023a), we expect that this experiment serves as a relatable example for most coastal practitioners.

In this experiment, the S-2 archive of publicly available imagery (level-1A surface reflectance) is efficiently retrieved (See section 2.2.3) from the MSPC STAC catalog. To minimize redundant EO data processing, imagery transmitted to multiple downlink stations or covering overlapping areas (Bauer-Marschallinger et al. 2023) on the same orbital ground track is grouped by solar day, selecting only the first occurrence of each day. The blue, green, red, NIR, SWIR16, and SCL classification bands from the S-2 catalog are retrieved, while precise pixel alignment is achieved by centering pixels on the coordinate grid. Following Vos et al. (2023a), the SWIR16 band is aligned with other spectral bands using bilinear resampling, since this strategy better captures linear features such as waterlines. The S-2 SCL layer, a 10-class land cover classification, is used to mask pixels categorized as "No

Data", "Dark Area Pixels", "Clouds high probability", or "Cirrus". However, the category "Snow and Ice" often represents whitewater in coastal zones, and is therefore deliberately not masked.

In line with well-established approaches(Vos et al. 2023a), coastal waterlines are mapped by applying Otsu-thresholding to mNDWI optical satellite imagery and extracted at sub-pixel resolution using marching-squares. The Otsu-threshold value is computed after applying a pixel classification, to consider exclusively sandy and water pixels in the thresholding process. Additionally, we add a simple quality control filter to exclude imagery with over 95% water pixels in the S2 SCL or CoastSat classification, as such high water content does not yield effective results when using the Otsu'algorithm. Finally, by intelligently designing the computational graph, we simultaneously compute two results: a raster map with the coastal water occurrence probability, as the mean presence of water pixels over time; and a vector layer with waterlines from each image. We emphasize that this experiment does not aim to introduce yet another method for shoreline monitoring, but rather serves as relatable benchmark to the efficiency of data-proximate cloud computing for coastal monitoring.

Experiment 2: Geospatial data retrieval

This experiment benchmarks the efficiency of geospatial data retrieval using eight different data dissemination strategies by retrieving coastal transects for the Basque Country, Spain, from the GCTS. We evaluate the efficiency gains of cloud-optimized data, geospatial sorting, and metadata filtering with STAC across different data models (GeoPandas, Dask GeoPandas, and DuckDB) and retrieval methods (spatial join and predicate pushdown). Geospatial sorting, which involves sorting data based on quadkey (a geohash facilitating efficient spatial indexing, also see 2.2.5), is examined. Metadata filtering, performed on the attributes provided in the STAC collection, allows for selective retrieval of relevant data partitions. We also compare retrieval methods, including spatial join operations, which merge datasets based on their spatial relationship, and predicate pushdown, a query optimization technique that applies filters early in the data retrieval process to enhance performance by reducing the amount of data transferred and processed.

Case study: Coastal elevation mapping

This case study demonstrates the potential of cloud-native workflows for coastal science by combining data-proximate computing (Expt. 1; section 2.2.8) with cloud-optimized data accessed via a standardized metadata specification (Expt. 2; section 2.2.3). Specifically, we integrate the GCTS with DeltaDTM, a novel digital terrain model (Pronk et al. 2024), to determine the percentage of the world's first kilometer of coast that is lower than 5 meters. The GCTS consists of more than 11 million coastal transects, while DeltaDTM includes 7105 tiles of 1 x 1 degree each at a spatial resolution of one arc second (30 m). We migrated DeltaDTM to cloud object storage and described it in a STAC catalog for convenient analysis. Thus, for this analysis both datasets are stored in cloud-optimized formats and accessible using STAC into a cloud-based computational cluster located close to data storage. DeltaDTM tiles are grouped by quadkey at zoom level 4, and relevant transects from the GCTS are retrieved using Dask GeoPandas, with STAC effectively filtering the necessary partitions. The transect data is then broadcast across the client, allowing it to be referenced by pointers without transmitting the data over the network, ensuring efficient processing. This setup establishes an efficient, scalable data processing routine capable of high-resolution extraction on a global scale.

2.3 Results

2.3.1 Where should we run our algorithms?

We analyzed the efficiency and scalability of data-proximate, cloud-based coastal waterline mapping. The results demonstrate that mapping shorelines in close proximity to where the satellite data is stored is 70 times faster for small areas, such as at Narrabeen Beach, Australia (App B.2). This is a significant improvement that highlights the efficiency gains of data-proximate computing, which are specifically relevant for coastal monitoring because this practice requires frequent data processing to track dynamic shoreline changes as satellite imagery comes available. Moreover, a comparison (App. B.1) shows that this cloud-native CoastPy approach produces results that are in good agreement with CoastSat, making it an important step towards instantaneous shoreline mapping at extensive spatial scales. Finally, data-proximate coastal monitoring is more sustainable as it eliminates the need to duplicate or move large datasets and efficiently utilizes on-demand compute resources.

Another key result relates to scalability. We assessed the scalability of our cloud-native approach to coastal waterline mapping by expanding the study area to regional and state levels, first examining the San Francisco Peninsula and then the entire state of California. While scaling up to larger areas (e.g., Ocean Beach) the computation becomes relatively more efficient (App. B.3), with CPU usage almost fully saturated at 100% while operating at scale. While processing areas as large as Sentinel-2 tiles (110 x 110 km) exceeded the memory capacity of our Dask workers (8 GB RAM per worker), the workflow remained efficient at the California state level when tiles were subdivided into zoom-level 10 quadtiles (38 x 38 km) (See section 2.2.5). The results, as shown in Figure 2.2, indicate that coastal areas spanning quadtiles at zoom level 10 can be processed in approximately five minutes on average; that equivalents to coastal waterline mapping at 50 km per seconds, which is several orders of magnitude faster than conventional approaches, such as CoastSat (0.1 km/s; App. B.2). Although we find that occasionally data is spilled to disk during the operation, reporting double wall-clock times (10 minutes), the cluster recovers to a healthy state, demonstrating the robustness of this approach. This improved efficiency at larger scales is because proportionally less time spent configuring the workers and constructing the computational graph (App. B.3). Overall, these findings highlight that specialized coastal monitoring routines can be much more efficiently orchestrated in the cloud, close to where the data is physically stored, with increased efficiency as the study area scales from local to regional levels.

2.3.2 How should we store our data?

We benchmarked eight different geospatial data retrieval strategies. Table 2.1 summarizes the characteristics and performance of each strategy, including mean and standard deviation of execution times over 20 iterations. Detailed distributions of retrieval times are presented in Appendix B.5.

In the initial experiment, as presented in Table 2.1, coastal transects were retrieved from the GCTS, that is stored as a traditional geopackage, which is approximately 5GB in that format. Unlike in other experiments, this retrieval operation had to be conducted on a personal computer due to memory limitations on a standard MSPC Pangeo instance. This setup resulted in data retrieval that is up to 160 times slower (Expt. 1 vs. Expt. 8) compared to those achieved with cloud-optimized data formats. Thus, the occurrence of memory errors and the prolonged execution time for data retrieval in this experiment show the convenience and advantages of cloud-optimized data formats.

Subsequent experiments involve fetching transects from the GCTS stored across several GeoParquet partitions, that are altogether approximately 1GB in this cloud-optimized format. The experiments use various techniques such as spatial joins or predicate pushdown (See section 2.2.8),

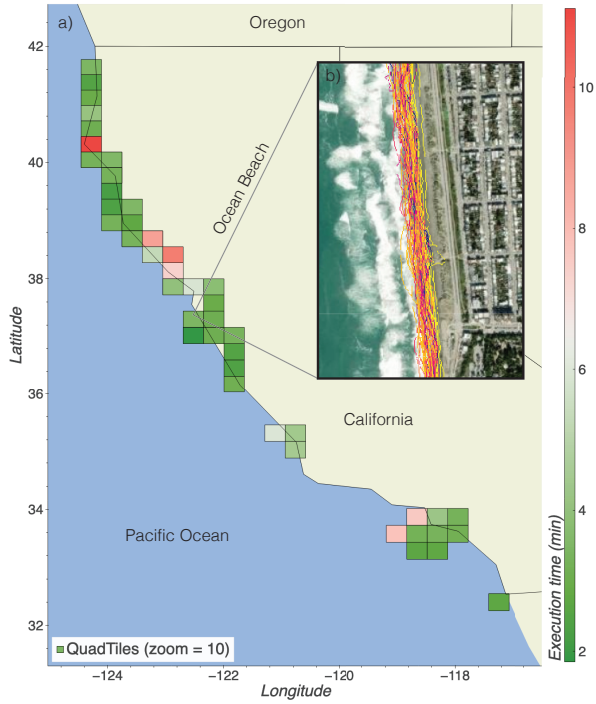

Figure 2.2: Execution time (s) per quadtile for cloud-native coastal waterline mapping across California state, USA. By orchestrating shoreline-monitoring routines in close proximity of the satellite data, coastal waterlines are mapped at approximately 50 km/s.

Table 2.1: Average execution time (s) for various strategies of data retrieval. The crosses (X) indicate whether a certain strategy was applied.

Expt.	Data Model	Query Method	Cloud Opt.	Spatial Sort	STAC Filter	Mean (s)	Std (s)
1	GeoPandas	Spatial join				1071.3	5.0
2	Dask GeoPandas	Spatial join	X	X		41.3	1.1
3	Dask GeoPandas	Spatial join	X			40.5	1.0
4	DuckDB	Spatial join	X	X		25.2	1.1
5	DuckDB	Spatial join	X			24.0	0.8
6	Dask GeoPandas	Spatial join	X	X	X	10.9	0.3
7	DuckDB	Predicate pushdown	X	X		7.4	0.6
8	DuckDB	Predicate pushdown	X	X	X	6.7	0.7

facilitated by a STAC catalog to identify relevant partitions and spatial sorting by quadkey. Surprisingly, the results (Tab. 2.1) indicate that spatial sorting slightly increases data retrieval times. However, the use of STAC metadata to effectively filter the relevant STAC items enhances Dask GeoPandas so that it is capable of retrieving transects four times faster (Expt. 4 vs. Expt 6), likely due to the reduced need to index data objects using HTTP protocol. This efficiency gain demonstrates the advantage of standardized metadata in optimizing data retrieval, while also underscoring the importance of geospatial sorting, as the effectiveness of the STAC metadata filter depends on it.

The results (Tab. 2.1) demonstrate that the DuckDB query engine is more efficient for geospatial data retrieval than Dask-GeoPandas (e.g., Expt 4 vs 2) in this setup, being almost twice as fast. More importantly, the data show that employing predicate pushdown on a bounding box attribute is a three times more efficient strategy for data retrieval than a spatial join operation (Expt 7 vs. Expt 5). For geospatial predicate pushdown to function effectively, the data must be geospatially sorted and include a bounding box column that provides the extent of the geometry. This underscores the importance of adding a bounding box column to geometries and organizing the data geospatially to significantly enhance query performance. Additionally, the data retrieval is slightly faster (10%) when using STAC metadata to selectively filter relevant STAC items (e.g., Expt. 8 vs. Expt. 7), although the efficiency gains from this approach do not match those achieved with Dask GeoPandas (twice as fast). While the differences in execution times among strategies that access cloud-optimized data formats are of a different order than the 160 times improvement reported against conventional data formats, the differences are relevant for analysis at scale. Implementing simple geospatial sorting methods combined with standardized metadata specifications results in approximately 4 times faster performance. This is particularly important when these access patterns are repeated for a large number of similar tasks, demonstrating the effectiveness of these strategies in optimizing geospatial data retrieval at scale.

In summary, these experiments affirm the importance of cloud-optimized data formats for broad-scale coastal analytics, where data interoperability is crucial (Sec. 2.1). Cloud-optimized data formats, when combined with geospatial sorting and query optimization techniques like predicate pushdown, facilitate efficient filtering methods for retrieving data pertinent to specific regions of interest. Additionally, this experiment highlights the value of metadata-driven data access strategies in enhancing the efficiency of data retrieval, which is presumably increasingly important as the number of partitions grows and indexing remote data objects over HTTP protocol becomes a bottleneck. Therefore, adopting cloud-optimized data formats, geospatial sorting, and enabling metadata-driven access methods is essential for facilitating efficient geospatial data retrieval.

2.3.3 How much of the first km of coastal land is below 5 meters?

By using cloud technology as described in earlier experiments (See section 2.3.1 and 2.3.2) we developed a scalable, high-resolution mapping at global scale. Figure 2.3 illustrates the fine spatial resolution at which elevation data is extracted along the landward side of the transects around Barra de la Cruz, Mexico. In total we extract approximately 300 million elevation observations at equally-spaced 100 m alongshore resolution. Compared to the conventional download-and-analyze approach, this method significantly enhances our ability to analyse coastal datasets at global scale. We have developed a scalable method to locate and process DeltaDTM tiles. Similarly, coastal transects for a given tile can be retrieved downloading and processing an extensive 5 GB GeoPackage. A traditional workflow, involving a spatial join for each of the 7,105 tiles, would take approximately 88 days to process just to retrieve the transects. In contrast, using cloud technology, we completed the entire analysis in a few hours, making it 700 times faster.

On average, we find that 33% of the first km in the coastal zone is lower than 5 meter. Figure 2.4 shows the average percentage of very low-lying coastal land (<5 m) in the first km coastal zone, with the 5% (n=25) largest clusters (Anselin 1995) indicated as red dots. The map shows that particularly the gulf of Mexico, the East Coast of the United States and the European Wadden Sea much of the first km of coastal land is lower than 5 meters above mean sea level. In this 1-km coastal zone, the distribution (Fig. 2.4) of areas that are on average lower/higher than 5m is highly bi-modal; which shows that either the 1-km coastal land is mostly below 5m or, land rises fast and elevation is mostly above 5m.

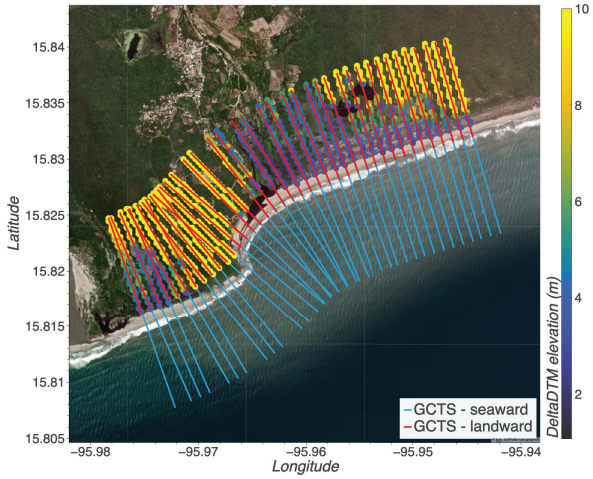

Figure 2.3: Extracting DeltaDTM elevation data over the landward side (1 km) of our coastal transects, around Barra de la Cruz, Mexico.

2.4 Discussion

Cloud technology and the opening up of EO data have started a digital transformation in coastal science. In the introduction, we distinguished between two distinct analysis strategies: one aims for global coverage, often compromising accuracy for spatial extent ("everywhere"), while the other prioritizes accuracy ("anywhere"). This paper demonstrates how leveraging an open, flexible, and scalable geospatial software stack (Pangeo) in combination with cloud-optimized data formats bridges the gap between these approaches, enabling coastal analytics at broad scales without compromising the high spatio-temporal resolution and accuracy that are typically used in more local analyses.

An essential step this approach is decoupling data storage from compute, a strategy that contrasts with the integrated frameworks of existing platforms like GEE. This separation provides more control over both processes, enabling independent management of each.

Compute, is managed on a open, flexible and scalable framework (Pangeo), that, depending on the needs of the analysis, can be deployed on personal computers, HPC, or cloud infrastructure. As shown in the first experiment, this software ecosystem provides a flexible environment that can be scaled up to broader areas while retaining the spatio-temporal resolution used in some specialized coastal monitoring practices.

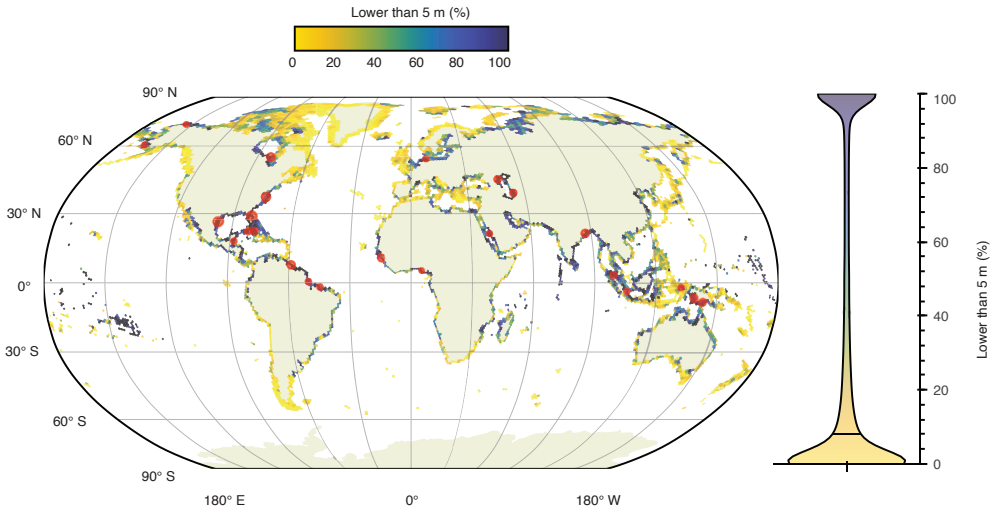

Figure 2.4: The map shows the average percentage of land within the first kilometer of the coastal zone that is lower than 5 meters above mean sea level, based on approximately 300 million observations from DeltaDTM (Pronk et al. 2024) across more than 11 million transects (GCTS). On average, 33% of this coastal zone is below 5 meters. Red dots indicate the 5% largest clusters of predominantly low-lying coastal land. To the right, a violin plot shows the distribution of the percentage of coastal land below 5 meters.

Storage, in turn, is handled using cloud-optimized data formats, while adhering to standardized metadata specifications. As shown in the second experiment, this structured approach enables rapid querying, and ensures that data is optimized for coastal analytics at scale.

Implementing and maintaining such a cloud solution presents some technical challenges. Besides writing scalable software, producing cloud-based data repositories requires expertise in cloud-optimized data formats, data specifications and data partitioning. While such skills may not be commonly expected from coastal scientists, our case study demonstrates that the substantial improvements in efficiency, flexibility, and hence analytical capabilities justify these efforts.

The modularity of cloud-based coastal analytics eases the migration between different computational environments while removing dependency on proprietary platforms (e.g., GEE). Also, by decoupling storage and compute, this system can continuously integrate the latest innovations in both data management and algorithms. Other advantages of cloud-based solutions is that it provides universal access points to data, that are much more accessible than traditional storage solutions, while they typically also reduce the need for data downloading with in effect less data duplication.

Cloud computing has a significant environmental impact (Monserrate 2022), and data-intensive applications—particularly those utilizing artificial intelligence (AI)—-are a major source of its energy consumption (Katal et al. 2023). Coastal research labs that leverage this powerful infrastructure should be mindful of their environmental footprint, especially as their funding is often tied to initiatives promoting sustainability and/or addressing climate change. However, despite its notable carbon footprint, cloud computing can offer more sustainable solutions (Jones 2018). By optimizing resource allocation, reducing idle time, and consolidating large-scale operations, cloud systems can minimize overall energy consumption. Additionally, cloud infrastructure's universal access points reduce the need for data duplication and unnecessary transfers, while cloud-optimized data formats,

through efficient compression, further minimize storage and transmission requirements, enhancing overall sustainability.

Although there are technical barriers and environmental impacts that require careful attention, the importance of cloud technology for large-scale coastal analytics is clear, as also shown in the experimental findings, that are discussed in the next section.

2.4.1 Experimental findings

In this section we discuss several key results from our experiments, beginning with the novel global coastal transect system (GCTS). The transect system benefits from recent OSM data uptake, corrects zonal bias, includes polar latitudes, and offers a finer alongshore resolution (100 m) than existing transect systems. We believe this transect system can serve as a robust foundation for various coastal studies, providing a reliable basis for deriving SDS series, coastal characterizations, classifications, and related statistics.

Secondly, we show that 33% of the first kilometer of coast is below 5 meters. This data, calculated at a 100-m alongshore resolution, is crucial for several kinds of coastal analyses, such as coastal classification and characterizations. This finding is particularly relevant in the context of the low-elevation coastal zone (LECZ), as described by McGranahan et al. (2007), potentially providing a more detailed understanding of coasts that are vulnerable to accelerating climate change. However, the 30-m resolution of the digital terrain model (DTM) is likely insufficient for flood-risk modeling, underscoring the continued relevance of local studies. The workflows used can be adopted to map other coastal variables onto, for example, our transect system (GCTS). With more coastal variables mapped, it will become possible to apply theoretical classification frameworks (e.g., Cooper et al. 1998; Finkl 2004) at extensive scales without compromising spatial resolution and accuracy.

Thirdly, we present open, scalable, flexible methods for efficient coastal analytics at planetary scale, that are up to 700 times faster than traditional download-analyze approaches. This tremendous speed-up is achieved by bringing code to the data rather than moving the data to the code; and, using cloud-optimized data exposed through rich, standardized metadata specifications. By adopting this framework for coastal waterline mapping, we achieved a processing speed of 50 km/s, compared to 0.1 km/s using conventional methods, like CoastSat. Data-proximate computing, increasingly standard in various scientific fields (Gentemann et al. 2021), is very relevant for the coastal community, which increasingly relies on EO data (Vitousek et al. 2023a), that is still typically downloaded and analyzed on institutional premises.

The coastal waterlines mapped during the data-proximate shoreline monitoring experiment are primarily intended to demonstrate the feasibility of efficiently implementing specialized coastal monitoring routines in the cloud, rather than serving as a dataset for studying coastal dynamics. Nevertheless, the methods introduced here have the potential to form the basis for the first-ever global mapping of instantaneous shorelines from the full historical Landsat and/or Sentinel-2 catalog. Crucially, future enhancements must include tidal corrections and improvements to the classification methods to ensure robust generalization across diverse coastal environments (Konstantinou et al. 2023; Vos et al. 2023a), including macro-tidal regions and beaches with unique sand types, like those on volcanic islands.

Fourthly, we show that for coastal analytics at scale, cloud-optimized data exposed through standardized metadata specifications are essential. Our experiments on geospatial data retrieval indicate that data storage format and accompanied metadata are as important as the data itself, especially for large-scale analyses. By eliminating repetitive tasks such as locating data for regions of interest and enhancing data interoperability, we developed data retrieval methods that are up to 160 times faster

than traditional download-and-analyze approaches. Another advantage is that by adopting these principles, the data de facto becomes findable, accessible, interoperable, reusable (FAIR) (Wilkinson et al. 2016).

A crucial step in curating the data is partitioning. In this study, we explored various partitioning methods, with our released GCTS partitioned by quadkey to enhance the efficiency of regional data access and time-series analysis. While partitioning based on administrative boundaries has been suggested (Holmes 2023), it poses challenges in coastal science. Administrative boundaries can enhance user-friendliness, but the coast—being a transitional area between ocean and land—is not always fully encompassed by existing administrative divisions. Additionally, the complexity and variability of coastal regions can result in highly unequal partitions, further complicating this approach.

Our geospatial partitioning strategy, while effective, is not definitive, especially for datasets that grow with time, such as with ongoing satellite missions. Managing spatio-temporal coastal data repositories, such as SDS, presents significant challenges due to conflicting partitioning needs. Spatial partitioning enables efficient regional data access, that is useful for time-series analysis, but requires frequent data rewriting, which is computationally intensive and costly, particularly in cloud-object stores. As a result, repositories may be managed through temporal partitioning, where data is appended as new satellite data becomes available. While this is more logical for continuously updating datasets, it is less efficient for users needing access to time series of certain coastal variables per region.

Fifthly, the cloud-native coastal waterline mapping experiment demonstrates that we can now incorporate advanced ML models into coastal analyses at scale. Although the classification model used in this study is relatively basic (a simple feed-forward network), it serves as proof of concept that advanced coastal deep learning models can be integrated into server-side computing. This capability is critical for implementing advanced coastal ML models (e.g., Buscombe et al. 2018; Al Najar et al. 2023) at extensive spatial scales.

2.4.2 Outlook

We argue that the coastal community should collectively begin constructing publicly available, analysis-ready coastal data repositories, as these will be the critical resources for powering coastal ML and providing a robust foundation for data-driven coastal decision making. The digital revolution in coastal science, triggered by the opening up of historical satellite catalogs and innovations in cloud technology, has provided an unprecedented global perspective of the coastal environment. However, the coastal community is not yet fully capitalizing on recent advances in ML, such as deep learning (DL). Although the ongoing digital revolution in coastal science is most likely going to culminate in the widespread adoption of coastal AI, this advancement requires the availability of high-quality data. Adopting best practices from computer science (Raymond 1999) and making data publicly available (Tenopir et al. 2015) is crucial for achieving this goal, but not enough. Now that coastal science has an appetite for data, its management must be recognized as an integral component of coastal research. Without a focus on data management, we risk accumulating more data than we can effectively analyze. Future coastal data releases should aim to minimize manual, repetitive work in downstream coastal analytics, by providing analysis-ready data, following flexible data schemata that we, as a coastal community, still have to work out ourselves. Yet it is also critical that future coastal analytics integrate strategies to reduce environmental footprints, particularly for the data-intensive applications. To ensure the long-term sustainability and accessibility of the data resources as well as the tools to work with it, governmental support (e.g., Directorate-General for Research and Innovation (European Commission) 2022) to develop and maintain public cloud infrastructure is essential, as reliance solely

on commercial enterprises may compromise the stability and equitable access to potentially critical coastal data. We envision a future where an Earth System data cube (Mahecha et al. 2020), enriched with coastal characteristics, will allow customized views of the coast, possibly even through natural language queries (Zhu et al. 2024) enabled by artificial intelligence.

2.5 Conclusion

This study demonstrates the transformative potential of cloud technology for coastal analytics at extensive spatial scales. We introduced a novel global coastal transect system (GCTS), found that 33% of the first kilometer of coast is below 5 meters, and developed methods that are up to 700 times faster than conventional approaches. The GCTS introduced here can serve as a foundational coastal dataset, offering a robust frame of reference with a global set of coastal stations that can be used for deriving shoreline-change series, coastal characterizations, and related statistics. Our findings highlight that coastal science no longer needs to be constrained by high latency, storage capacity, available compute resources, or specific toolboxes provided by cloud platforms. By leveraging cloud technology and a flexible, scalable software ecosystem, we can perform complex computations close to data storage, drastically reducing analysis time from months to just a few hours. This approach allows us to use all available data without compromising accuracy or resolution, hopefully setting a precedent for future broad-scale coastal studies. We are therefore convinced that if the coastal community aims to address urgent coastal challenges at broad spatial scale without compromising accuracy or spatio-temporal resolution — bridging the gap from "everywhere" towards "anywhere" — it will have to start producing tools, models, and data suitable for scalable coastal analytics.

2.6 Software availability

The authors share their code and instructions to access the data at https://github.com/TUDelft-CITG/coastpy. The Global Coastal Transect System is also available for download at https://zenodo.org/records/14056925.

Software name	CoastPy
Developer	Floris Calkoen
Year first official release	2024
System requirements	Mac, Linux, Windows
Program language	Python
Program size	<1 MB
Availability	https://github.com/TUDelft-CITG/coastpy
License	MIT
Documentation	README, documentation and tutorial notebooks
Citation	F R Calkoen et al. (Jan. 1, 2025f). "Enabling Coastal Analytics at Planetary Scale". *Environmental Modelling & Software* 183, p. 106257. DOI: 10.1016/j.envsoft.2024.106257

3 Mapping the world's coast

A global 100-m coastal typology derived from satellite data using deep learning

Abstract We present a globally consistent, high-resolution (100 m) coastal typology dataset derived from satellite imagery and elevation data using deep learning—the first application of its kind in coastal science. Using a supervised multi-task convolutional neural network, we classified for nearly 10 million coastal transects (one million km of coast) four coastal attributes along the cross-shore profile: (1) sediment type, (2) coastal type, (3) presence/absence of built environment, and (4) presence/absence of human-made coastal defenses. The model, trained on about 1800 globally distributed samples, achieves strong predictive performance with F1 scores ranging from 0.67 to 0.83. Results show that the global coastal sediment distribution consists of 40% sandy, gravel, or shingle; 21% muddy; 13% rocky; and 27% with no sediment. Considering the coastal type, 33% of coasts are cliffed, 22% are sediment plains, 15% are wetlands, and 3% are dune systems (i.e. 26,000 km). Combining sandy, gravel, shingle and muddy sediments, we estimate that 61% of the global coastline consists of soft sediments that are potentially easily erodible. Among sandy, gravel or shingle coasts specifically, 20% are cliff-backed and 16.5% are located on built-up coasts. This global dataset, available in a cloud-optimized format at https://doi.org/10.5281/zenodo.15599096, provides a robust foundation for coastal-change analysis and erosion assessment, and enables new opportunities for broad-scale vulnerability mapping and adaptation planning in the face of accelerating sea-level rise.

This chapter is based on the following article:

Calkoen, F. R., Luijendijk, A. P., Hanson, S., Nicholls, R. J., Moreno-Rodenas, A., De Heer, H., & Baart, F. Mapping the world's coast: a global 100-m coastal typology derived from satellite data using deep learning. *Earth System Science Data Discussions* [preprint], in review, 2025. https://doi.org/10.5194/essd-2025-388

3.1 Introduction

Coastal zones are increasingly under pressure from human development and the cumulative effects of accelerating climate change (Oppenheimer et al. 2019). Urbanization, infrastructure expansion, and land reclamation have transformed coastal landscapes, often leaving limited space for natural processes to unfold (Lansu et al. 2024). As a result, future risks are increasing: erosion and flooding threaten assets, livelihoods and even lives (Barnard et al. 2025), while critical ecosystems face habitat loss and degradation (European Environment Agency. 2024). Anticipating these risks requires reliable climate data and spatial information to support adaptation planning (Le Cozannet et al. 2017). A coastline's response to drivers such as sea-level rise and increased storminess depends largely on its geomorphology (Stive et al. 2002; Woodroffe 2002; Masselink 2014). Yet despite major advances in satellite-based coastal monitoring (Vitousek et al. 2022), no globally consistent, high-resolution coastal typology currently exists to assess vulnerability and ultimately inform sustainable, long-term action.

As early as the 1990s, the potential of geographic information systems to support integrated coastal management was recognized (e.g., Cooper et al. 1998). One of the first broad-scale examples was the EUROSION project (Salman et al. 2004), which highlighted coastal erosion risks across Europe and promoted more coordinated shoreline management among EU member states. Its coastal typology provided a structured spatial overview of European coastlines, compiled from national and regional geological datasets. At the global scale, the DIVA (Dynamic Interactive Vulnerability Assessment) model (Vafeidis et al. 2008) introduced a systematic segmentation of the world's coastlines to support sea-level rise impact assessments (e.g., Hinkel et al. 2013; Schuerch et al. 2018). These initiatives show the growing importance of broad-scale coastal typologies for coastal planning but also illustrate the limitations of their time: they relied on heterogeneous national and regional data sources, which led to inconsistencies in classification and variable spatial coverage and quality.

In the last decade, coastal mapping has been rapidly transformed by two key developments: 1) the opening of satellite Earth observation archives (Wulder et al. 2012), which dramatically increased the availability of consistent, high-resolution data across space and time; and 2) the emergence of user-friendly geospatial cloud platforms, which have enabled convenient broad-scale spatial analysis (Gorelick et al. 2017). Together, these advances have enabled global assessments of coastal change at pixel resolution (e.g., Murray et al. 2018) and have established satellite-based coastal monitoring as standard practice in coastal science. Several studies have used these resources for coastal classification: Luijendijk et al. (2018), to focus its shoreline monitoring on sandy coasts; Mao et al. (2022) for broader geomorphological classification; and Hulskamp et al. (2023) to map muddy coasts globally. However, the resulting typologies often remain narrow in scope and rely on traditional tree-based machine learning methods (e.g., Random Forest), which classify each pixel independently. This limits their ability to capture the spatial structures that define many coastal systems — such as dunes, wetlands, tidal inlets, or engineered coasts.

Deep learning (LeCun et al. 2015) offers a more powerful alternative by directly learning spatial patterns from imagery, enabling end-to-end classification of complex, multi-pixel coastal features. Despite its demonstrated success in other Earth observation domains (e.g., Brown et al. 2022), deep learning remains underutilized in coastal science. A few studies (e.g., Dang et al. 2020; Buscombe et al. 2023; Çelik 2024) have shown its promise, but no global, deep learning-based coastal classification has yet been developed.

Here, we introduce a global coastal typology derived from satellite data using deep learning, developed to support coastal change analysis, erosion assessment, vulnerability mapping, and adaptation planning. The typology adopts a cross-shore perspective, drawing on established classification

frameworks that describe coastal systems through interrelated attributes along the land–sea profile (Sharples et al. 2009; Finkl et al. 2020). The dataset comprises nearly 10 million classified transects (Calkoen et al. 2025f)—covering one million kilometers of coastline between 70°S and 70°N—at 100-meter alongshore resolution, and represents the first global-scale application of deep learning in coastal classification. This paper describes the typology framework, training data, classification model, and resulting dataset, including global summaries and regional comparisons. The typology is distributed in cloud-optimized format, described via a STAC Collection, and archived at: https://doi.org/10.5281/zenodo.15599096.

3.2 Methods

To produce a globally consistent coastal typology, we designed a scalable classification method built around six components: (1) a classification schema defining the coastal attributes of interest; (2) globally available Earth observation datasets serving as model inputs; (3) a consistent transect grid for systematic sampling; (4) labeled training data for supervised learning; (5) a deep learning model capable of capturing spatial patterns across the cross-shore profile; and (6) an inference workflow for global-scale prediction. The following sections describe the methods used in detail.

3.2.1 Coastal typology framework

This study adopts a practical coastal classification (Fairbridge 2004; Finkl 2004) designed to support coastal monitoring, erosion assessment, vulnerability mapping, and adaptation planning. Conceptually, it builds on cross-shore coastal profiling approaches that classify the coast based on (geomorphological) landform characteristics along the land–sea profile (Sharples et al. 2009; Hanson et al. 2010; French et al. 2016; Finkl et al. 2020). Specifically, we follow the distinction made by Sharples et al. (2009) between shore fabric and coastal landform—corresponding to sediment type and coastal type in our framework—and the related separation of sediment stores and coastal landforms proposed by French et al. (2016), which align with these two attributes respectively. In total, the classification distinguishes four coastal attributes along the cross-shore profile: (1) sediment type, describing the sediment composition near the shoreline; (2) coastal type, characterizing the dominant landform; (3) presence/absence of built environment, and (4) presence/absence of human-made coastal defenses (Fig 3.2, panel A). The number of classes is intentionally kept limited to enhance interpretability and facilitate global application. The full classification schema is summarized in Table 3.1.

The sediment type refers to the dominant type in sediment composition of the area just landwards of the immediate shoreline. While drawing on established sediment classification systems—such as Hayes' climatic-sediment framework for the inner continental shelf (Hayes 1967)—our schema integrates shingle, gravel and sandy shores into one class because these are difficult to reliably distinguish from each other in satellite imagery.

The coastal type refers to the geomorphologically-active, dominant landform found landward of the immediate shoreline. It includes natural landforms with an additional class "engineered structures" for areas where the geomorphology is no longer active due to human influences. Sediment plains describe low-lying terrain composed of mostly unconsolidated materials and are commonly found in deltaic environments. In contrast, bedrock plains are also low-lying areas, but then composed of mostly solid rock, often appearing in uplifted regions with resistant lithologies, such as the granitic coasts of the Nordic countries. Cliffs are defined by their elevation and steep gradients, generally exceeding 15 m in height, with typical slopes higher than 30 degrees. Wetland coasts include all three major muddy classes, encompassing salt marshes, mangroves, and sabkha systems.

In addition, the schema includes a transitional class for moderately sloped coasts, where land rises gradually from sea to higher hinterland, a landform that is found in Northern Spain and Ireland. The coastal types are classified taking into account anthropogenic modification. For example, when dunes have been flattened due to urban expansion, the area is categorized as a sediment plain. This applies for example to coastal segments in the Basque region (Spain), where former dune systems, like Zarautz, have been largely urbanized. Similarly, polder landscapes in the Netherlands, which represent areas of reclaimed land, are categorized as sediment plains with defense when they include visible flood protection infrastructure, such as a dike. Areas where natural landform is no longer active due human influences, such as port areas, are categorized under "engineered structures".

The final two attributes in coastal typology act as binary modifiers. The built environment field indicates whether an area is predominantly developed or urbanized. Similarly, the "has defense" attribute specifies the presence of visible hard-engineering structures designed to protect against coastal erosion or flooding.

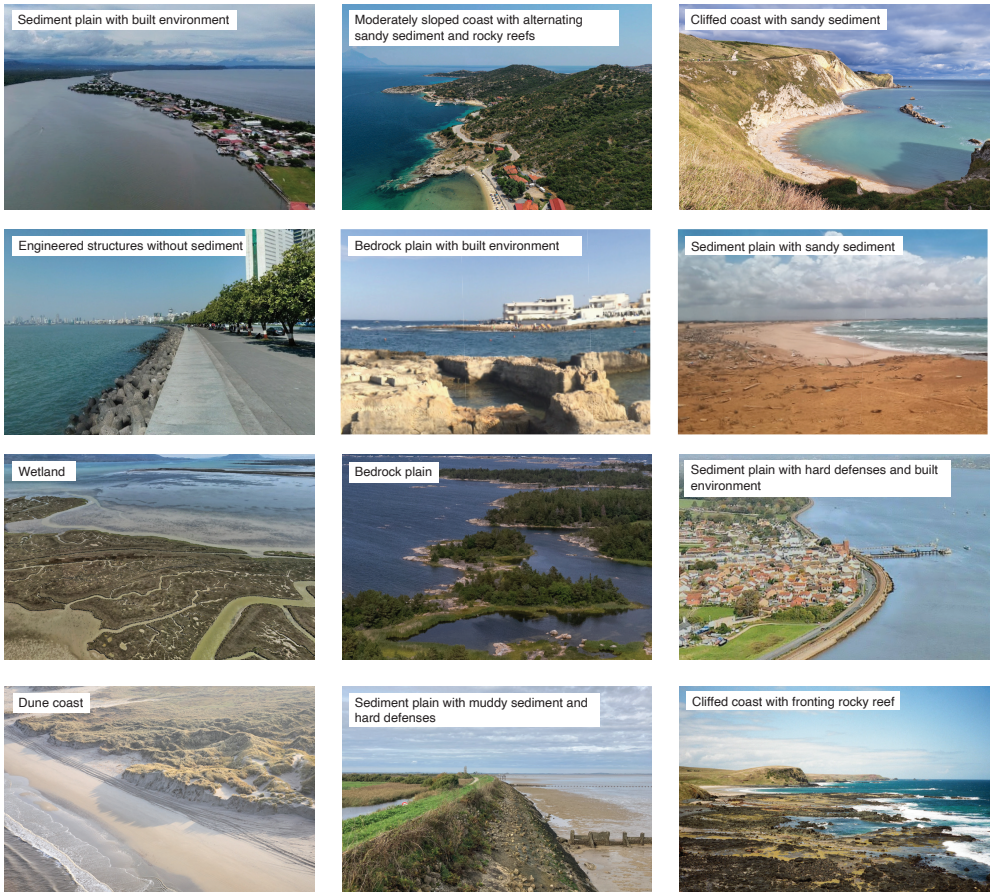

Figure 3.1: Illustrative examples of coastal environments, showing variations in sediment type, coastal type, presence of built environments, and coastal defense structures. The images were contributed by the authors and sourced from publicly available imagery.

Table 3.1: Classification schema used in the coastal typology model, including class definitions.

Attribute	Class name	Description
Sediment type	Sandy, gravel or small boulder sediments	Shorelines composed of unconsolidated materials such as sand, gravel, shingles, and small boulders (0.0652 to 512 mm in diameter).
	Muddy sediments	Shorelines dominated by fine-grained sediments like silt and clay, forming environments such as mudflats and tidal flats.
	Rocky shore platform or large boulders	Shorelines composed of solid rock formations, including shore platforms or large boulders greater than 512 mm in diameter.
	No sediment or shore platform	Shorelines with minimal visible sediment, typically around rocky cliffs, steep faces, or human-made structures such as sea walls.
Coastal type	Cliffed or steep	Coastal areas with cliffs or steep rock faces, generally exceeding 15 m, with slopes of 30 degrees or greater.
	Moderately sloped	Coastal areas with gentle to moderate slopes (<30 degrees), often composed of unconsolidated sediment or soft rock.
	Bedrock plain	Low-lying coastal areas (<15m) primarily formed by consolidated bedrock, including skerries, with minimal variability in elevation.
	Sediment plain	Low-lying coastal areas (<15m) with flat or gently sloping unconsolidated sediment, often featuring beach ridges or washover complexes.
	Dune	Sandy coastal areas characterized by wind-formed dunes, often stabilized by vegetation such as grasses.
	Wetland	Coastal areas periodically flooded, including environments such as tidal flats, salt marshes, mangroves, sabkhas, and peatlands.
	Inlet	A narrow coastal waterway where the sea meets the land creating dynamic systems such as estuaries and lagoons.
	Engineered structures	Coastal areas dominated by engineered structures such as port areas, sea walls, breakwaters and groynes, where the natural coastal landscape is obscured or heavily modified.
Is built environment	True	The coastal area is characterized predominantly by human-made structures, including buildings, industrial complexes, and port facilities.
	False	The coastal area remains largely natural, with minimal or no presence of built structures like buildings, industrial zones, or ports.
Has defense	True	Visible hard engineering structures, designed to protect against coastal erosion and flooding (e.g., sea walls, breakwaters), are present.
	False	No visible hard engineering structure, designed to protect against coastal erosion and flooding (e.g., sea walls, breakwaters), are present.

3.2.2 Collecting training data

The training data for the supervised classification was collected using a custom web application. A total of approximately 1800 training samples were provided by more than 15 experts, with the vast majority (98%) coming from the authors of this paper. As with any supervised learning framework, the quality and character of the classification depend strongly on the underlying training data. In this study, approximately 95% of samples were labeled by the lead author. As a result, the training data, and by extension the resulting typology, reflect the interpretation and domain understanding of that individual. Appendix C.6 shows the global distribution of training samples. While the training dataset covers a broad geographic range, including samples from coastlines around the world, most are concentrated along European coasts.

Each sample was classified based on an area of interest measuring 400 m alongshore by 2 km cross-shore (Fig 3.2, panel B), defined along a transect from the Global Coastal Transect System (GCTS)(Calkoen et al. 2025f). Experts viewed this area on ESRI World Imagery (2.5 m resolution) and assigned labels for each of the four classification tasks: sediment type, coastal type, presence/absence of built environment and presence/absence of human-made coastal defenses. The interface also provides direct links to Google Street View, which users relied on when classes could not be confidently assigned from experience and/or the top-down imagery alone. Labels were assigned to the best of the contributors' ability using these high-resolution sources, even though they were aware that the model would be trained on lower-resolution Sentinel-2 imagery (10 m). Class balance across all four tasks was difficult to maintain, but the overall data collection process aimed for class balance in coastal type, resulting in more even representation for that attribute (Tab. 3.3). We refer to this training dataset as CoastBench (Calkoen et al. 2025g): A global training dataset for coastal classification using satellite imagery and elevation data. It is publicly available, described in a STAC Collection, released under a CC-BY-4.0 license, and accessible at https://doi.org/10.5281/zenodo.15800284, with new samples welcome via the web application.

3.2.3 Satellite data acquisition

This study integrates multi-source Earth observation data to construct a standardized coastal typology datacube with 15 feature maps (Fig. 3.2, panel C). The datacube is constructed in two main stages: (1) generation of an annual cloud-free Sentinel-2 composite, and (2) fusion of this composite with elevation layers and derived spatial features to create the final 15-channel input stack. The input stack combines the cloud-free Sentinel-2 composite—capturing optical and infrared reflectance—with elevation data from the Copernicus DEM (European Space Agency (ESA) 2019) and the coastal DeltaDTM dataset (Pronk et al. 2024). Both Copernicus DEM and DeltaDTM were used because DeltaDTM offers higher accuracy for coastal areas but is limited to elevations below 30 m, whereas Copernicus DEM provides coverage extending beyond this elevation. The flow chart (Panel D) in figure 3.2 shows the data processing steps. All datasets were accessed via STAC APIs: Sentinel-2 and Copernicus DEM through the Microsoft Planetary Computer, and DeltaDTM via the CoCliCo catalog. Inputs were reprojected to a common 10 m UTM grid using bicubic interpolation and normalized to the [0, 1] range to ensure consistent model input scaling.

The classification was performed on standardized image chips of 2.8 km $\times$ 2.8 km (approximately 8 km^2), each centered on a coastal transect from GCTS (Calkoen et al. 2025f). A region of interest, defined as a 400 m alongshore by 2000 m cross-shore rectangle, anchors the target classification zone within each chip, while the surrounding area provides spatial context for the model. This is illustrated in figure 3.2 (Panel C), where the outer red box indicates the complete image chip, and the inner orange box marks the model's target area.

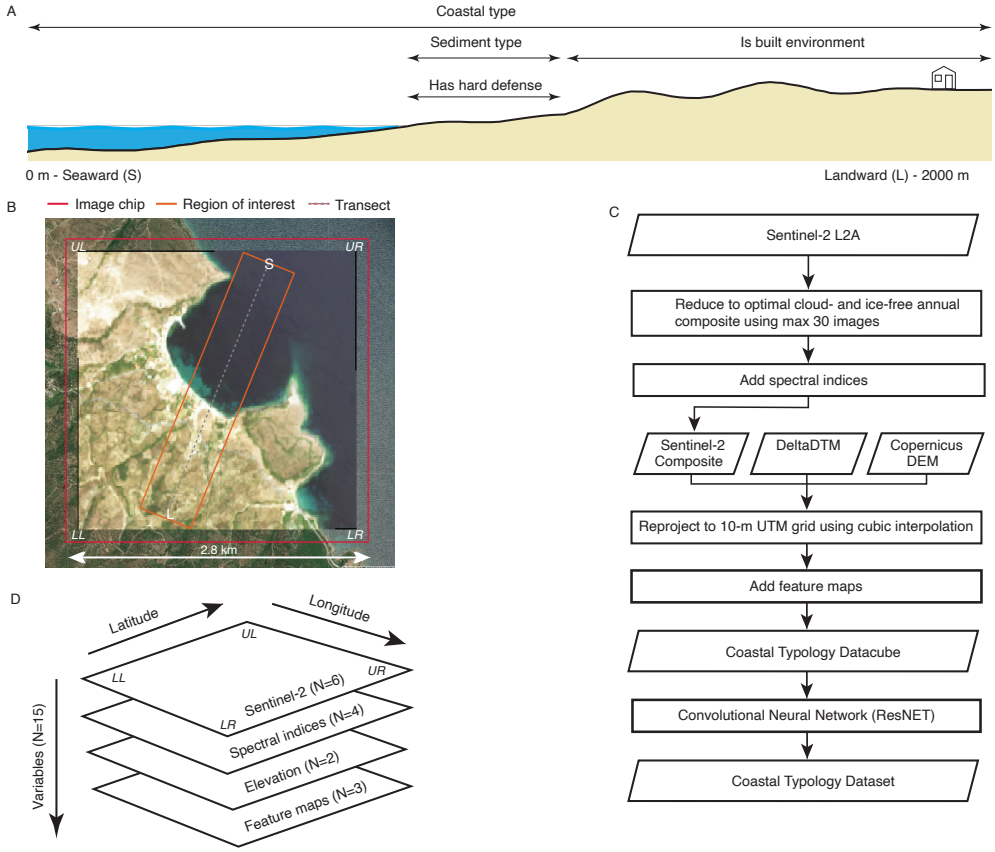

Figure 3.2: Overview of the coastal typology classification workflow. **(A)** Conceptual cross-shore model describing the classification framework, which distinguishes four coastal attributes along the sea–land profile: sediment type, coastal type, whether the area is predominantly built-up, and presence of coastal defenses. **(B)** Illustration of the image sampling process. Each chip (2.8 km × 2.8 km) is centered on a GCTS coastal transect (dashed gray line); the region of interest (orange rectangle) defines the target area for classification, while the surrounding red area provides additional context to the model. The basemap is from Esri World Imagery. **(C)** Flow chart of the data processing and classification pipeline, from raw Sentinel-2 imagery and elevation data to classified coastal attributes. **(D)** Schematic of the Coastal Typology Cube, consisting of 15 variables: 6 Sentinel-2 bands, 4 spectral indices, 2 elevation layers, and 3 spatial feature maps.

Sentinel-2 composites

An annual, cloud-free Sentinel-2 composite was generated for the period 2023-01-01 to 2024-01-01. Imagery was retrieved via the Microsoft Planetary Computer STAC catalog and processed per MGRS tile. To ensure balanced spatial and temporal sampling, scenes were grouped by orbital-track and partitioned into four synthetic groups, while enforcing a five-day minimum interval and selecting the least cloudy scene per period. Cloud and shadow masking used the Sentinel-2 Scene Classification Layer (SCL), excluding pixels labeled as "no data", "dark area pixels", or "clouds high probability". A median composite was computed from remaining pixels, resulting in a, cloud-free, composite image per tile. The final composite includes six reflectance bands—blue, green, red, NIR, SWIR16, and SWIR22—resampled to a unified 10 m UTM grid using bicubic interpolation. The resulting global, cloud- and ice free composite was cataloged as a STAC collection for downstream interoperability.

Coastal typology cube

The coastal typology cube was constructed by combining the annual Sentinel-2 composite with elevation data from the Copernicus DEM and the DeltaDTM product, all resampled to a common 10 m UTM grid. In addition to the six Sentinel-2 reflectance bands and two elevation layers, the cube includes four spectral indices (NDVI, NDWI, MNDWI, NDMI) and a relative elevation layer capturing local terrain contrast. Spatial context is encoded through binary masks for the region of interest (ROI), as well as landward and seaward points of the respective transect. Thus we obtain a coastal typology cube with 15 maps per image chip, normalized to the [0,1] range using robust min-max scaling with the values specified in Table3.2, which also summarizes the resulting input stack used for classification.

Table 3.2: Model input variables used in the coastal typology data cube. Each variable is listed with its source, description, and robust clip range. All variables were scaled using Robust Min-Max scaling with the values provided in the Clip Range column, unless otherwise noted. Variables marked with an asterisk are derived from source data.

Variable	Source	Description	Clip Range
blue	Sentinel-2	Surface reflectance (B2)	[1000, 4000]
green	Sentinel-2	Surface reflectance (B3)	[1100, 4300]
red	Sentinel-2	Surface reflectance (B4)	[1000, 5000]
nir	Sentinel-2	Near-infrared reflectance (B8)	[1000, 6500]
swir16	Sentinel-2	Shortwave infrared reflectance (B11)	[1000, 6000]
swir22	Sentinel-2	Shortwave infrared reflectance (B12)	[1000, 5500]
NDVI	Sentinel-2*	Normalized Difference Vegetation Index	[-0.25, 0.50]
NDWI	Sentinel-2*	Normalized Difference Water Index	[-0.60, 0.30]
MNDWI	Sentinel-2*	Modified NDWI (green–SWIR)	[-0.45, 0.45]
NDMI	Sentinel-2*	Normalized Difference Moisture Index	[-0.25, 0.45]
cop_dem_glo_30	Copernicus DEM	Absolute elevation (m)	[-20, 150]
deltadtm	DeltaDTM	Absolute elevation (m)	[-20, 30]
cop_dem_glo_30_rel	Copernicus DEM*	Relative elevation (local min/max)	[0, 75]
deltadtm_rel	DeltaDTM*	Relative elevation (local min/max)	[0, 30]
region_of_interest_mask	GCTS*	Binary mask for 400 m × 2000 m region of interest	—
landward_mask	GCTS*	Binary mask for landward side of transect	—
seaward_mask	GCTS*	Binary mask for seaward side of transect	—

3.2.4 Deep learning classification model

The classification task is a multi-task, multi-class problem (Ruder 2017), where the model predicts four coastal attributes across the water-land cross-shore profile per image: two multi-class labels (*sediment type* and *coastal type*) and two binary labels (*has defense* and *is built environment*). All tasks are trained jointly using a shared deep learning model based on a ResNet-50 (He et al. 2016) backbone pretrained on ImageNet. The backbone is followed by four parallel classification heads, each producing task-specific predictions from a shared feature representation. Model input consists of 15 channels (the coastal typology cube), including Sentinel-2 reflectance bands, spectral indices, elevation layers, and region-of-interest masks, normalized using robust min-max scaling (see Table 3.2).

To balance task contributions, the total loss was defined as a weighted sum of task-specific loss functions. Multi-class outputs were optimized using categorical cross-entropy, and binary outputs using binary cross-entropy with logits. Loss scaling was set to 1.0 for *sediment type* and *coastal type*, and 0.5 for *has defense* and *is built environment*, prioritizing the more complex typological classes while preserving sensitivity to binary classifications.

The model was implemented in PyTorch Lightning, with experiment tracking and checkpointing managed via Weights & Biases (WandB). A hyperparameter sweep explored learning rate, batch size, and loss scaling factors. Data was split into training, validation, and test sets (70:15:15), while ensuring no spatial overlap between splits. Data augmentation during training included random flips, rotations and affine transformations.

3.2.5 Validation

Model performance is evaluated over 10 independent training runs, each initialized with a different random seed, using a fixed 70:15:15 train–validation–test split. To avoid data leakage, overlapping image chips are assigned exclusively to a single partition. All metrics are computed on the held-out test set and reported (Tab. 3.3) as the mean with standard deviation across the 10 runs. Evaluation focuses on the per-class F1-score as the primary performance metric, due to its robustness to class imbalance compared to accuracy (Christen et al. 2023). Task-level performance is summarized using the macro-F1 metric, defined as the unweighted mean of per-class F1-scores: Macro-F1 $= \frac{1}{N} \sum_{i=1}^{N} \mathrm{F1}_i$, where N is the number of classes and $\mathrm{F1}_i$ is the F1-score for class i.

3.2.6 Inference at scale

For large-scale inference, the final model was retrained on the full dataset, utilizing an 85:15 train-validation split during this phase to maximize training coverage while still having an independent validation partition for early stopping to prevent overfitting. Inference was performed by dynamically constructing coastal typology cubes using the STAC API processing per coastal grid tile (Calkoen et al. 2025f). Tiles were processed using a coastal grid at zoom level 9. Distributed inference was executed on a SLURM cluster using Dask JobQueue, with each worker allocated 32 GB of memory and four threads. Batches were sized at approximately twice the number of active workers to optimize asynchronous requests, throughput and overall resource utilization.

3.2.7 Software

This study used the Pangeo ecosystem (Hamman et al. 2018) and the OpenDataCube (ODC) (Killough 2018) framework, which together provide a scalable foundation for accessing, transforming,

Table 3.3: Per-class F1-scores (mean ± standard deviation) for each classification task. Each score represents the average performance over 10 model runs. The final row reports the macro-F1 score for each task. N refers to the number of samples in the training data. Empty cells indicate that a given class is not defined for that task.

Class	N	Sediment Type	Coastal Type	Is Built Env.	Has Defense
No sediment or shore platform	343	0.70 ± 0.07			
Sandy gravel or small boulder sediments	745	0.84 ± 0.04			
Muddy sediments	242	0.86 ± 0.04			
Rocky shore platform or large boulders	305	0.64 ± 0.09			
Bedrock plain	153		0.62 ± 0.07		
Cliffed or steep	334		0.78 ± 0.06		
Dune	156		0.65 ± 0.08		
Engineered structures	146		0.69 ± 0.11		
Inlet	139		0.61 ± 0.12		
Moderately sloped	209		0.55 ± 0.08		
Sediment plain	348		0.68 ± 0.07		
Wetland	150		0.77 ± 0.08		
False	977/1177			0.86 ± 0.02	0.88 ± 0.02
True	658/458			0.79 ± 0.03	0.66 ± 0.07
Macro F1	–	**0.76 ± 0.04**	**0.67 ± 0.03**	**0.83 ± 0.02**	**0.77 ± 0.04**

and analyzing large-scale geospatial raster data. Raster retrieval is managed via `odc-stac`, while `odc-geo` and its geobox model are used extensively for spatial operations such as reprojection and clipping. Distributed data processing and inference are performed using `Dask 2025.2.0` (Rocklin 2015), enabling parallel, lazy execution across a SLURM cluster. Deep learning routines are implemented in `PyTorch Lightning 2.5`, with `PyTorch 2.2` (Paszke et al. 2019) providing GPU-accelerated training and Weights & Biases (Biewald 2020) used for experiment tracking. All computations were performed using `Python 3.12`, and the code—including data models and deep learning architectures—is available through the open-source `CoastPy` (Calkoen et al. 2025f) package. Geospatial analytics are conducted primarily using `DuckDB 1.2` (Raasveldt et al. 2019), including its spatial extension and H3 bindings for hexagonal hierarchical spatial aggregation.

3.3 Results

We present a globally consistent coastal typology dataset based on the classification of nearly 10 million GCTS transects, covering close to one million kilometers of coastline between 70°S and 70°N, at 100 m alongshore resolution. The following sections describe the spatial distribution of predicted classes at global (Section 3.3.1) and continental scale (Section 3.3.2), provide a local example to illustrate prediction detail (Section 3.3.3) and examine typological relationships across tasks (Section 3.3.4). The section concludes by outlining the structure and accessibility of the released dataset (Section 3.3.5).

3.3.1 A Global Coastal Typology

Figures 3.3 and 3.4 present global maps of predicted *sediment type* and *coastal type*, each aggregated to a level-3 H3 hexagonal grid using the most frequent class (mode) per cell. These coarser sum-

maries, based on nearly 10 million classified coastal transects, reveal broad-scale patterns in sediment type and geomorphology. More regional European maps are included in appendix C.3 and C.3. Similar spatial summaries of the binary attributes *is built environment* and *has defense* are provided in appendix C.1 and C.2.

Sandy coasts dominate along much of Africa, Southeast Asia, and Australia, while muddy shorelines concentrate in tropical deltas and back-barrier systems such as the Gulf of Guinea (Fig. 3.4) and the Wadden Sea (App. C.3). Rocky shore platforms appear in more localized areas (e.g., western Ireland; App. C.3), while high-latitude coasts in Norway, Alaska, and southern Chile are frequently classified as "no sediment or shore platform". Considering *coastal type* map, cliffs dominate in tectonically active or glaciated regions (e.g., Nordic countries; Kamchatka, Russia; and, Southern Chile), wetlands cluster in low-lying tropical zones (e.g., Bangladesh, West Africa), and dune systems are found along the coasts of North Brazil and southwestern France (App. C.4). Engineered structures are concentrated in highly urbanized or industrialized coasts such as Japan, eastern China, and the San Francisco Bay Area.

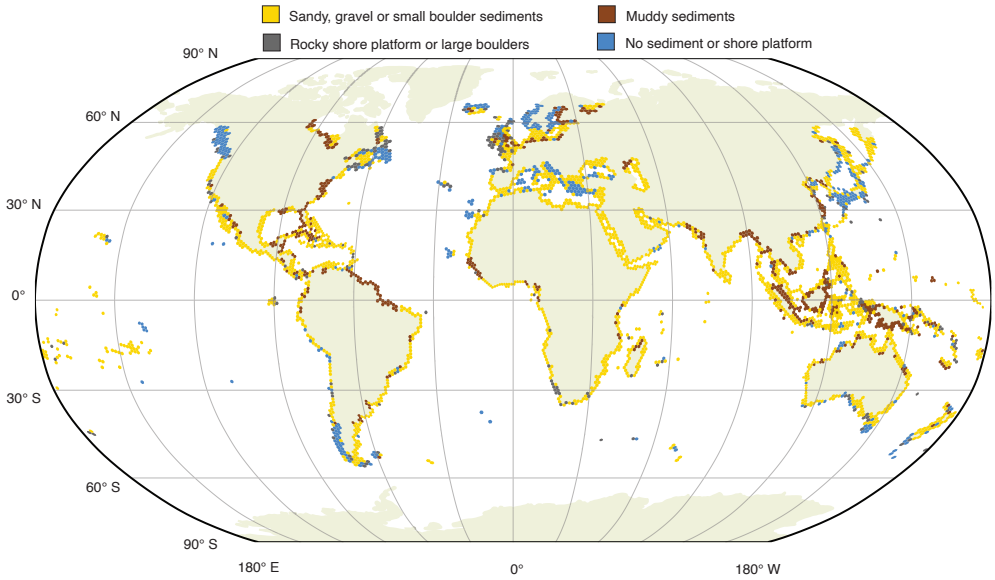

Figure 3.3: Global map of predicted *sediment type*, aggregated on a level-3 H3 hexagonal grid using majority class (mode) per cell. This coarser view summarizes almost 10 million classified coastal transects, highlighting regional patterns. Basemap: Natural Earth.

3.3.2 A quantitative global and continental overview

Table 3.4 provides a quantitative overview of typological coverage by class and continent. Among sediment types, sandy coasts are the most common globally (40%), with particularly high representation in Africa (65%). Muddy shores account for 21% of the global length and are more prevalent in Asia (29%) and Africa (22%), typically reflecting tropical deltas and estuarine systems. Rocky shore platforms—characterized by large boulders or exposed rock—make up 13% and are relatively evenly distributed across continents. The class "No sediment or shore platform" represents 27% of the global coast and is concentrated in high-latitude, high-relief regions such as Europe (36%),

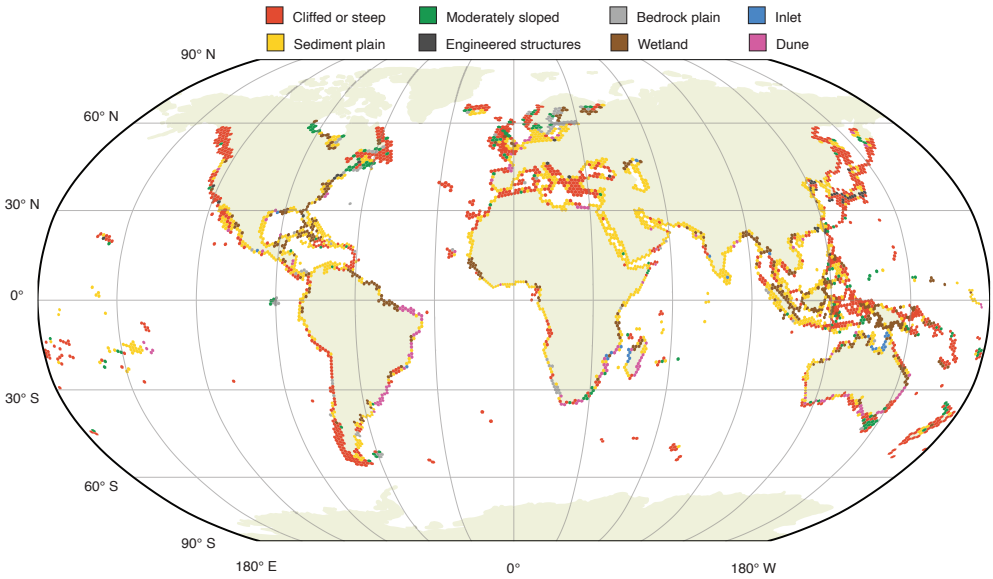

Figure 3.4: Global map of predicted *coastal types*, aggregated on a level-3 H3 grid using the most frequent class (mode) per cell. This coarser view summarizes almost 10 million classified coastal transects, highlighting regional patterns. Basemap: Natural Earth.

North America (37%), and South America (35%), corresponding to the cliff-dominated settings highlighted in Section 3.3.1 (Fig. 3.4).

For coastal types, cliffed or steep coasts are the most frequent globally (33%), most prominent in Europe, North America, and South America. Sediment plains (22%) and wetlands (15%) are more common in Asia and Africa, reflecting the prevalence of low-lying humid coasts. Dune systems are relatively rare (3% globally), but appear more prominently in Africa (10%), often along straight, sediment-rich coasts such as southwestern France (Aquitaine) (App. C.4). Engineered structures make up a modest share (3%), with the highest in Asia (7%) and Europe (5%), consistent with the high density of urbanized and industrialized coastlines in these regions.

The global coastline remains predominantly natural: 87% of segments are not classified as built environment, and 91% show no coastal defense structures, with maps provided in appendix C.1 and C.2. However, regional contrasts are notable. Asia stands out with 21% of its coastline classified as built and 17% showing defense infrastructure, followed by Europe (18% built, 11% defended). In contrast, Africa, Oceania, and South America remain largely undeveloped, with built and defended segments typically below 6%. Overall, the data show that 59% of built-up coastlines also contain coastal defenses. Conversely, and as expected, the vast majority (83%) of defended coasts are situated along built-up areas.

The class distributions are also computed at the country level. Appendix C.5 shows normalized class percentages per country for all four classification tasks using a stacked bar plot, enabling visual comparison of coastal typology across national boundaries. A full tabular summary of these country-level statistics is also provided in the Zenodo archive (Sec. 3.3.5), offering both absolute (km) and relative (%) values for each class and classification task.

Table 3.4: Global and continent-wise summary of coastal typology classes per task, by total coastal length (10^3 km) and relative proportion (%). Class names are abbreviated for clarity (e.g., "Rocky" for "Rocky shore platform or large boulders"). Abbreviations: Global (GL), Europe (EU), Asia (AS), Africa (AF), North America (NA), South America (SA), Oceania (OC), Antarctica (AN).

		Length (10^3 km)								Percentage (%)							
Attribute	**Class**	**GL**	**EU**	**AS**	**AF**	**NA**	**SA**	**OC**	**AN**	**GL**	**EU**	**AS**	**AF**	**NA**	**SA**	**OC**	**AN**
Sediment Type	Sandy/gravel/shingle	374	61	96	35	102	31	46	2	40	35	46	65	34	32	48	49
	Muddy	193	27	62	12	50	19	23	0	21	15	29	22	17	19	25	1
	Rocky	119	25	25	5	37	13	14	1	13	14	12	8	12	13	14	16
	No sediment/platform	252	64	27	3	109	34	12	1	27	36	13	5	37	35	13	33
Coastal Type	Cliffed or steep	306	65	47	8	113	44	26	3	33	36	22	14	38	45	28	60
	Moderately sloped	125	31	21	3	42	14	13	1	13	17	10	6	14	14	14	29
	Bedrock plain	74	20	8	3	32	6	5	0	8	11	4	5	11	6	5	5
	Sediment plain	206	32	67	21	52	13	21	0	22	18	32	40	17	13	22	3
	Dune	26	4	4	5	4	3	6	0	3	2	2	10	1	3	6	1
	Wetland	143	16	45	11	37	16	18	0	15	9	21	20	13	16	19	1
	Inlet	25	2	5	2	10	1	5	0	3	1	2	4	3	1	5	0
	Engineered	33	8	14	1	8	1	1	0	3	5	7	1	3	1	1	2
Is Built Env.	False	816	146	166	51	267	92	90	4	87	82	79	94	89	94	95	96
	True	122	32	44	3	32	6	5	0	13	18	21	6	11	6	5	4
Has Defense	False	852	158	174	52	277	94	92	4	91	89	83	97	93	97	97	98
	True	86	20	36	2	22	3	3	0	9	11	17	3	7	3	3	2

3.3.3 Local-Scale Example: Saunton Sands

Figure 3.5 presents detailed predictions near Saunton Sands, a geomorphologically diverse stretch of coastline in southwest England. This area includes a range of sediment types, landforms, and anthropogenic features, offering a representative setting to illustrate specifically the resolution at which the global coast is mapped.

North of the estuary, the model correctly identifies a well-developed dune system (Fig. 3.5, Panel A). South of the estuary, where dune forms are absent and a low embankment is present (Fig. 3.5, Panel B), the classification as "sediment plain" is appropriate. Within the estuary, sandy tidal flats are mapped as "sandy, gravel or small boulder" sediment type, with adjacent low-relief terrain labeled as "moderately sloped" coastal type (Fig. 3.5, Panel C). Built-up settlements along the estuary are also captured. Further upstream, the model successfully tracks the transition to muddy sediments.

At Croyde (Fig. 3.5, Panel D), the pocket beach is classified as "sandy, gravel or small boulder sediments" without defenses, and its hinterland is labeled "moderately sloped". A "dune" classification could also be argued for this coastal type and this is probably an example that relates to the relatively lower performance scores of these classes (Tab. 3.3). Just north of Croyde, the model detects the shift to a rocky reef and correctly assigns the "cliffed or steep" coastal type.

Further south, near Westward Ho (Fig. 3.5, Panel B), the model identifies the transition to a heavily modified shoreline, assigning both *is built environment* and *has defense*. Beyond this area, predictions capture a return to natural cliffed coastline without defenses. Across the full extent, the outputs show spatial coherence, reflect smooth transitions between class boundaries, and align well with observed landforms and infrastructure.

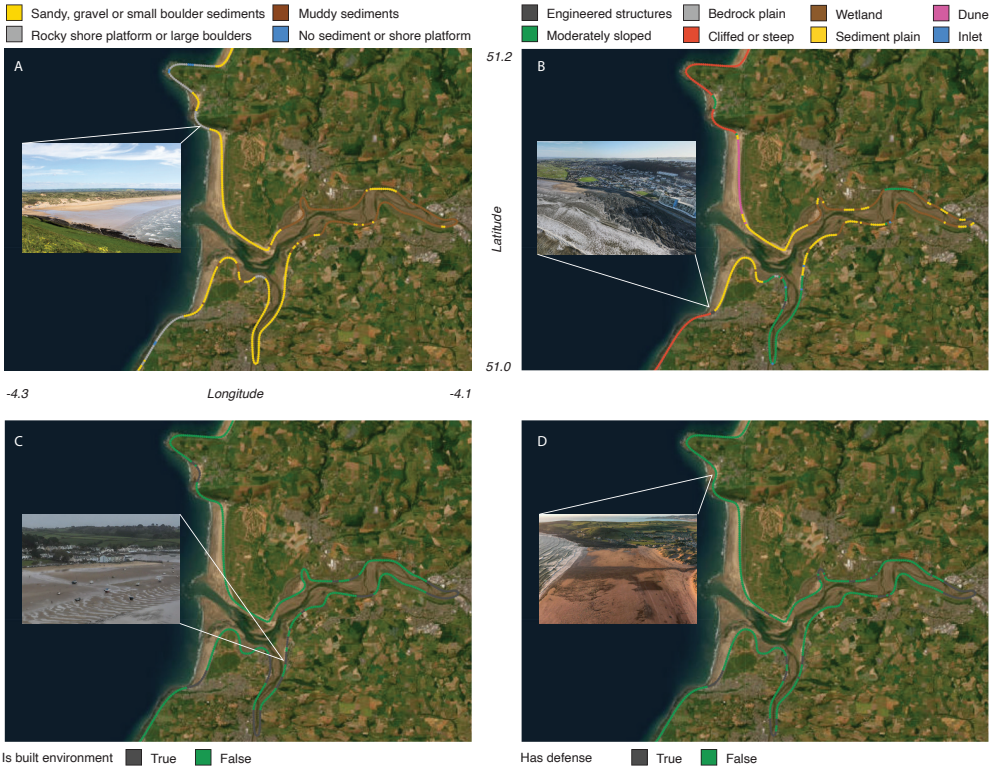

Figure 3.5: The coastal typology classification near Saunton Sands (Devon, UK), shown at its 100 m alongshore resolution for each of the four tasks: (A) *sediment type*, (B) *coastal type*, (C) *is built environment*, and (D) *has defense*. Insets in each panel show supporting photographs sourced from public imagery. The example illustrates how the model distinguishes subtle variations across the four classification tasks, including dunes along Saunton Sands (Inset A), sediment plains and coastal defenses near Westward Ho! (Inset B), built environment at Instow (Inset C), and the absence of defenses around Croyde despite nearby development (Inset D). The basemap is from Esri World Imagery. This example illustrates the resolution and interpretability of the coastal classification at local scale.

3.3.4 Co-occurrence and typological relationships

To explore relationships across the four classification dimensions, we present two complementary visualizations: normalized co-occurrence matrices between *sediment type* and *coastal type* (Figure 3.6), and a multi-level Sankey diagram showing co-distribution patterns across all tasks (Fig. 3.7). The co-occurrence matrices express conditional relationships, with each row representing a single class and columns showing the relative composition of paired classes, normalized to 100%. The Sankey diagram extends this by visualizing how class transitions unfold across all four tasks. Each vertical layer represents one classification dimension, and flow thickness corresponds to the total coastline length (in kilometers) shared by the connected classes.

Several patterns in these visualizations align with well-known coastal geomorphologies. Wetlands co-occur almost exclusively with muddy sediment shorelines (96%; Figure 6b), and dune coasts are paired with sandy shorelines (98%). Cliffed or steep coasts are most frequently associated with the absence of visible sediment (54%), but also co-occur with rocky platforms (22%) and sandy shores (24%), reflecting their geomorphic variability. Inlets are most often matched with sandy shorelines (89%), consistent with their dynamic and nature that is rich in sandy sediments. Engineered structures are strongly linked to sediment-free shorelines (74%), suggesting that hard-structure interventions tend to disrupt natural sediment presence or occur where it is entirely absent. Likewise, coastal defenses are overwhelmingly situated near built environments: 83% of defended coastlines also contain built environment (Fig. 3.7).

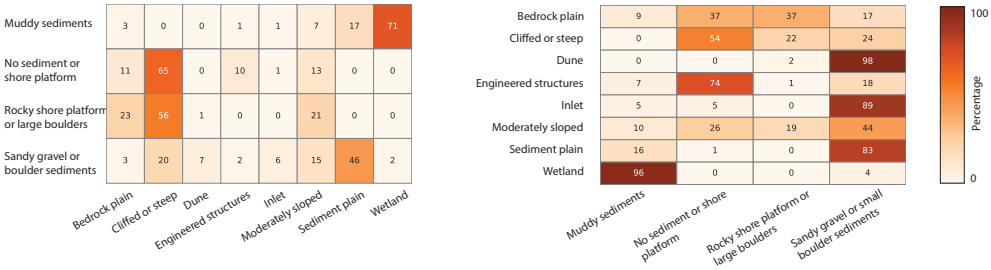

Figure 3.6: Normalized co-occurrence matrices between *shore_type* and *coastal_type*. Each row represents a single class from the source task (either shore or coastal type), and each column indicates the percentage of samples co-labeled with a given class from the other task. Values are normalized such that each row sums to 100%, enabling a direct interpretation of conditional relationships between classification tasks. This visualization highlights systematic co-occurrence patterns and supports interpretation of typological compatibility.

3.3.5 Data Records

The coastal typology dataset (Calkoen et al. 2025c) released in this study contains model predictions for four coastal typology tasks: *sediment type*, *coastal type*, *has defense*, and *is built environment*. Each record corresponds to a unique transect in the GCTS and includes both the predicted class label and a corresponding model confidence score for each task.

For multi-class tasks (*sediment type*, *coastal type*), probabilities are computed using the softmax function and reported per class. For binary tasks, probabilities are derived from the sigmoid function and indicate the likelihood of a positive prediction, using a threshold of 0.5 for classification. Probabilities are stored in columns prefixed with *prob_**, followed by the class name (e.g.,

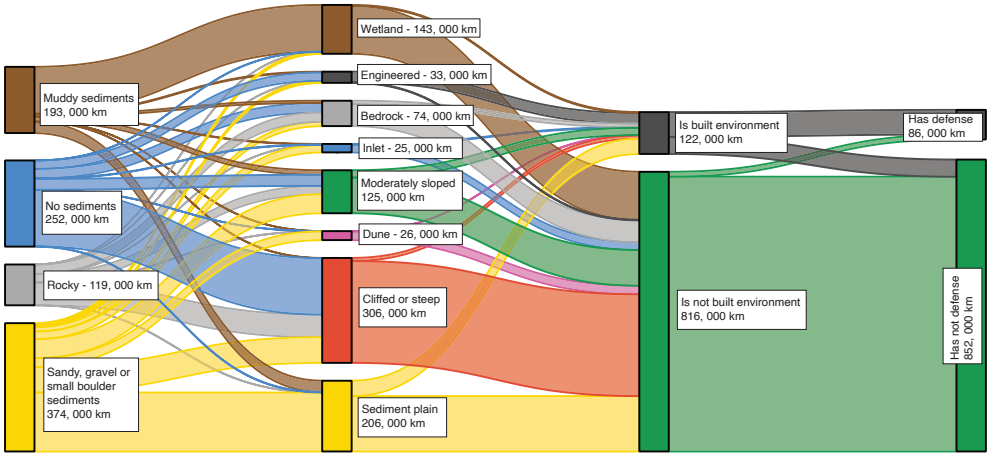

Figure 3.7: Sankey diagram showing the co-distribution of predicted classes across four typology dimensions: *sediment type*, *coastal type*, *is built environment* and *has defense*. Each vertical layer represents one classification task, with all categories in each task summing to the global coastline length in kilometers. The thickness of each flow segment is proportional to the coastline length shared between two connected class labels. This visualization highlights common transitions and conditional relationships between coastal typologies, such the relatively large share of sandy, gravel or small boulder sediments that is backed by cliffed or steep coasts.

prob_muddy_sediments) or task name (e.g., *prob_has_defense*). To enhance downstream coastal analytics, each record includes some key metadata (copied over from GCTS), including the unique transect ID, geographic coordinates, bounding box, administrative region, and tile identifier.

Table 3.5 summarizes all structured fields, including data types, value ranges, and source attribution. The dataset is stored as partitioned Parquet files, organized by coastal grid tile, and available under a CC-BY-4.0 license. It spans the global coastline between $-70°$ and $70°$ latitude while including Iceland, uses EPSG:4326 as its spatial reference system, and reflects Sentinel-2 satellite imagery of 2023. Data can be accessed via a the CoCliCo STAC catalog, with instructions and an example notebook available in the `CoastPy` GitHub repository (Calkoen et al. 2025f). The data is also archived at Zenodo https://doi.org/10.5281/zenodo.15599096.

3.4 Discussion

This study makes two distinct contributions: first, it provides better and previously unknown global estimates of the distribution and composition of coastal systems; second, it represents a methodological advance by introducing deep learning into coastal science for broad-scale coastal analysis of satellite data. The following subsections first discuss the results of our coastal classification by comparing the estimates to existing coastal typologies, then state the inherent assumptions made in the methodology, evaluate the model's performance, and finally highlight methodological innovations.

Table 3.5: Structured output fields in the geospatial classification dataset. Each field is defined by its name, description, data type, value range or format, and source.

Field Name	Description	Data Type	Range/Format	Source
transect_id	Unique GCTS ID; foundational reference	string	–	GCTS
geometry	Centroid point (longitude, latitude)	geometry	EPSG:4326	GCTS
bbox	Bounding box coordinates	array[float]	[minx, miny, maxx, maxy]	GCTS
quadkey	Microsoft tile identifier	string	QuadKey	GCTS
utm_epsg	UTM projection code	int32	e.g., 32633	GCTS
continent	Continent name	string	–	GCTS
country	Country code (ISO2)	string	ISO2	GCTS
common_country_name	Country name (long form)	string	–	GCTS
common_region_name	Administrative region name	string	–	GCTS
shore_type	Predicted sediment type class	categorical	4 classes	Model
coastal_type	Predicted coastal type class	categorical	8 classes	Model
is_built_environment	Predicted built environment presence	boolean	True / False	Model
has_defense	Predicted coastal defense presence	boolean	True / False	Model
prob_* (N=14)	Class probabilities (per task)	float32	[0–1]	Model

3.4.1 The global distribution and composition of coastal systems

Quantifying the global distribution and composition of coastal systems has a long history in coastal science, with estimates varying widely depending on methodology and data availability. An early expert-based inventory (Bird 1985) estimated that approximately 20% of the world's coastlines are sandy. In contrast, more recent large-scale satellite-based approaches have reported higher figures; Luijendijk et al. (2018) find that 31% of global ice-free coastlines are sandy. In our study, we estimate that 40% of the world's coastlines are sandy, gravel or shingle, amounting to 375,000 km between 70°N and 70°S. This higher figure primarily reflects improvements in spatial sampling. Unlike earlier studies based on Web Mercator-derived transects, we use the Global Coastal Transect System (GCTS), which distributes transects evenly (every 100 m) in latitude and thereby corrects zonal biases (Calkoen et al. 2025f). Attributing the difference to the transect system used is supported by a continental comparison: In Africa, where zonal distortion is minimal, our estimate of 65% sandy coastlines closely matches the 67% reported by Luijendijk et al. (2018). This agreement occurs despite differences in classification criteria, including a minimum beach width of 20 m and the exclusion of small boulder beaches in the earlier study, both of which are included in the present classification.

For muddy coasts, Hulskamp et al. (2023) reported a global figure of 14%, noting their prevalence in equatorial and deltaic regions. Our estimate is substantially higher at 21% (193,000 km), a difference that can again be attributed to usage of a latitudinally-consistent transect system (GCTS), which better captures the footprint of equatorial muddy coastlines such as those in Southeast Asia, West Africa, and the Amazon delta. An additional factor is the classification schema: whereas Hulskamp et al. (2023) include a separate class for "vegetated" coasts, our model does not distinguish vegetated systems explicitly because we focus on the geomorphology. Many such coasts—especially in estuarine deltas—are both vegetated and muddy, suggesting overlap between their vegetated class and our muddy class.

These comparisons with Luijendijk et al. (2018) and Hulskamp et al. (2023) underline the foundational role of GCTS in global coastal analytics. By using a geographically uniform transect system,

we reduce latitudinal sampling bias and accurately capture equatorial coastlines. Also, GCTS is derived from a more recent OpenStreetMap coastline, hence, has broader and better spatial coverage.

Despite their geomorphological significance and hazard mitigation, dune coasts are relatively rare on a global scale. Our analysis indicates that they comprise 3% of the world's coastline (26,000 km), a relatively short length, especially when contrasted to sandy shorelines more broadly, that account for 40% (375,000 km). This disparity probably reflects the specific formation requirements for dune systems (Moore et al. 2025)—including large sand supply, persistent onshore winds, and sufficient space for inland accumulation—which are not consistently met across most coastal regions. Human impacts may further contribute to their scarcity, as many dune systems have been flattened or modified for coastal development (Lansu et al. 2024). Given the relatively modest classification performance for dune coasts (F1 = 0.65), these findings should be interpreted with caution. Some dune systems may be misclassified into adjacent categories; most notably *moderately sloped* coasts, as shown in a local example (Fig. 3.5, panel B), where dunes are labeled as moderately slopedis.

Young et al. (2019) showed that cliffed coastlines are widespread globally, with 93% of the world's coastal regions containing some cliffed segments. However, their estimate was based on regional presence across 213 coastal units, rather than on proportional shoreline length. In this study, we provide the first globally consistent, length-based estimate of cliff coasts: approximately 33% of the global coastline, or 305,000 km between 70°N and 70°S, is cliffed or steep.

Our classification also enables cross-shore compositional analysis. For example, we find that approximately 20% of sandy coastlines are backed by cliffs or steep slopes (74,000 km). These cliff-backed beaches are of particular concern under rising sea levels (Vitousek et al. 2017), as they typically combine limited sediment supply with minimal room for inland migration, making them among the systems most vulnerable to beach loss and morphological collapse in the coming centuries.

Recent global assessments have quantified "coastal squeeze" (Lansu et al. 2024) and "coastal hardening" (Nawarat et al. 2024), providing insight into the extent of human modification along sandy shorelines. Lansu et al. (2024) report that 33% of sandy coasts have less than 100 m of infrastructure-free space inland, while Nawarat et al. (2024) estimate that 33% are "hardened." In our analysis, we apply more conservative guidelines based on visual interpretation of satellite imagery, supported by self-supervised machine learning. Using this approach, we estimate that 16.5% of sandy, gravel or small boulder shorelines are embedded within built environments, such as settlements or ports and 9.3% have visible coastal defenses. The dataset shows that 11% of the continental European coastline features human-made coastal defenses, substantially higher than the 5% previously reported by EUROSION (Salman et al. 2004). This discrepancy may reflect expanded detection through our automated methods, and/or an actual increase in coastal defenses since that study, and/or potential overclassification due to false positives in the model. Only 3% of sandy shores are classified as engineered coasts—locations where the natural geomorphology is visibly inactive due to heavy human intervention (e.g., port development). Notably, this engineered category is more prevalent in Asia, consistent with widespread recent coastal development in the region. Overall, these findings align with earlier work showing that coastal populations are concentrated in specific regions (McGranahan et al. 2007; Kulp et al. 2019)—particularly in South, Southeast, and East Asia, as well as mid- and low-latitude deltas and cities—while large lengths of coast have little or no direct population pressure (Small et al. 2003; Lincke et al. 2018).

3.4.2 Methodological assumptions

While this study is not constrained by data availability—satellite Earth observations are global and relatively consistent—our methodology relies on several assumptions. First, the classification pre-

sented here is not a fixed taxonomy but a functional typology, intended to support consistent, large-scale analyses of coastal change and erosion risk. Like any typology, it simplifies a naturally continuous landscape: transitional forms may fall between classes, and certain lithological nuances (e.g., distinctions between hard and soft rock) remain beyond the resolution of satellite-based observation.

Second, the classifier uses a supervised learning approach that necessarily reflects the characteristics of its training dataset. In this study, the majority of training samples (approximately 95%) were provided by the lead author. As such, the resulting typology inevitably reflects the judgments and interpretation of that individual. This is a common limitation of broad-scale remote sensing studies that rely on learning-based methods (including tree-based or deep learning approaches), where expert-driven labeling is fundamental to model training. Uncertainty also varies across the classified attributes: sediment type builds on well-established categories and is comparatively robust, whereas coastal type, built environment, and coastal defense are subject to greater interpretative ambiguity due to less standardized definitions.

3.4.3 Model performance

The model demonstrates strong and consistent performance across all four typology tasks, with accuracy scores varying according to task complexity, class prevalence and geomorphological ambiguity in the coastal typology.

Binary tasks—*is built environment* and *has defense*—achieved the highest macro-F1 scores of 0.83 and 0.77, respectively, with low variability across runs (standard deviations ≤ 0.04). As expected, the multi-class tasks were more challenging due to their finer morphological distinctions: *sediment type* reached a macro-F1 of 0.76, while the eight-class *coastal type* scored 0.67.

At the class level, performance also varied. Classes with distinct spectral or morphological features, such as *muddy sediments* (F1 = 0.86), *sandy, gravel or small boulder sediments* (F1 = 0.84) and *wetlands* (F1 = 0.77), were classified with high accuracy, also suggesting that clear visual cues can offset limited training data (e.g., *muddy sediments* and *wetlands*). Conversely, classes with more ambiguous or transitional characteristics—such as *moderately sloped* (F1 = 0.55)—showed lower performance, likely due to their overlap with adjacent classes like *cliffed*, *dune*, and *sediment plain*. Similarly, *dune* coasts yielded a modest F1-score (0.65), reflecting spectral similarity with nearby categories, particularly *sediment plains*, which explicitly include beach ridge systems that can resemble dunes in satellite imagery. For *inlet* (F1 = 0.61), performance varied more across runs, possibly indicating inconsistencies in the training data set for this class.

Despite a relatively small training set (N = 1800), the model exhibited consistent behavior across 10 independent training runs, with low variance in most metrics. While there is a broad correlation between class frequency and performance, exceptions suggest that feature distinctiveness and annotation clarity are also influential. To support future improvements, the annotation tool used for labeling is made publicly available, and we actively invite new contributions from the coastal science community to extend coverage in underrepresented areas or class types.

Model confidence scores during inference (Fig. C.7) indicate high model certainty outside Europe, but this should be interpreted cautiously, as confidence does not necessarily imply true generalization quality. However, confidence levels are broadly consistent across regions and vary more strongly by class, implying that uncertainty arises from typological ambiguity rather than geographic region, supporting the model's global applicability.

3.4.4 Methodological advances

A notable strength of our approach is that it relies exclusively on satellite imagery and elevation data, without incorporating ancillary geospatial variables such as climatic conditions (e.g., wave regime, temperature) or other geospatial data products (e.g., CORINE Land Cover). This design choice minimizes the risk of data leakage from environmental drivers into the classification process, thereby preserving the integrity of downstream analyses. As a result, correlations between coastal typology and external forcing factors—such as climate or socioeconomic variables—can be investigated without introducing circularity or confounding effects.

Another strength of our method is the region-of-interest (400 m x 2000 m) encoding within each image chip (2.8 km x 2.8 km). Each prediction is made for a 400 m alongshore by 2 km cross-shore region (region of interest, that is spatially encoded), while the model sees the broader a 2.8 km $\times$ 2.8 km area, providing additional context. In addition, the landward and seaward direction of the transect is encoded, so that model knows which coast to consider when classifying areas with narrow bays. Overall this setup enables the model to integrate local detail with the surrounding area, as seen in the Saunton Sands example (Section 3.3.3), where it accurately resolves transitions between dunes, sediment plains, rocky shores, urbanized segments, and pocket beaches like Croyde.

Recent efforts by Hanson et al. (2025) and Nyberg et al. (2025) continue the tradition of rule-based coastal classification, building on frameworks such as EUROSION and DIVA. These approaches are well-established and integrate geomorphological expertise with ancillary data to characterize the coast at broad scale. In contrast, our deep learning-based method is fully automated, relying solely on satellite imagery and elevation data, and can be efficiently updated and/or expanded. While promising for global-scale analysis, it represents an emerging approach; the typology definitions may require further refinement, and the training dataset would benefit from contributions by a broader expert community.

Many typological classes in this study span multiple pixels and are characterized by contextual spatial patterns (e.g., dunes, wetlands, and defense structures, among others). Deep learning models naturally accommodate such patterns by learning spatial dependencies, in contrast to traditional pixel-wise or feature-based classifiers (e.g., Random Forest) that often require manually engineered features or hybrid models. Adopting deep learning, reduces the need for complex rule-based post-processing and enables end-to-end learning from raw data. Overall, the accuracy scores for sandy and muddy sediment types are comparable to those of Hulskamp et al. (2023), who report F1-scores of 75% (sandy) and 87% (muddy) using a hybrid tree-based approach incorporating, that includes ancillary geospatial datasets. So, while using a smaller training dataset, without providing ancillary geospatial datasets as inputs, we report better accuracy scores for the deep learning model than has been previously reported for tree-based classifiers.

In the introduction, we highlighted two major developments that have transformed coastal science: the opening-up of Earth observation data and the emergence of user-friendly cloud platforms. This study represents a third step in this evolution, demonstrating how integrating deep learning (or broader AI) with these capabilities provides a natural and powerful solution for capturing the complexity of coastal environments globally. Uniting these three advancements (open satellite data, cloud technology, and artificial intelligence) arguably shows the full potential of modern, data-driven coastal analytics. Crucially, individually, each component has limitations: without open data, research will be unequal, disadvantaging those with less means and lack reproducibility; without cloud technology, scaling from experimental case studies to consistent global applications remains challenging or simply not possible; and without AI, analysis is constrained to traditional

machine learning methods unable to handle the complexity of real-world coastal systems. Only by combining all three can we can advance high-resolution, data-driven coastal science globally.

3.5 Conclusions

The combination of open satellite archives and cloud technology has fundamentally transformed how we study the Earth's coastlines. This study marks a third step in that evolution, demonstrating how integrating deep learning with these capabilities provides a more natural and especially more powerful solution for studying complex coastal systems globally.

The multi-task deep learning model shows strong performance across all four classifications tasks, with performance depending on task complexity, class prevalence and geomorphological ambiguity in the coastal typology. Binary classifications—identifying built environments (F1: 0.83) and coastal defenses (F1: 0.77)—achieve the highest accuracy due to their task simplicity. Performance remains strong for sediment type classification, particularly for clearly defined categories such as sandy (F1: 0.84) and muddy sediments (F1: 0.86). However, classification accuracy decreases for the more complex coastal type classification, especially for the transitional coastal types such as moderately sloped (F1: 0.55) and inlets (F1: 0.61) that have high morphological ambiguity, but are for some classes also negatively affected by the limited amount of training data (e.g., dune coasts) as well as labeling consistency within the training samples (inlets). A notable strength of the method is its reliance solely on satellite data, ensuring unbiased downstream analyses of relationships between coastal types and external environmental conditions.

The dataset offers comprehensive and accurate global estimates of coastal sediment composition, indicating that approximately 40% of global coastlines consist of sandy, gravel, or shingle sediments, 21% are muddy, 13% rocky, and 27% are sediment-free. This shows that over 60% of the world's coastlines are composed of soft sediments, which are easily erodible. Geomorphologically, about 33% of coasts are cliffed, 22% sediment plains, 15% wetlands, and only 3% are dune systems (adding up to 26,000 km), the rarity of which likely reflects specific formation requirements and widespread human alteration.

The cross-shore composition of coastal systems further emphasizes coastal vulnerability: 20% of sandy coastlines are cliff-backed, posing significant risks under accelerating sea-level rise due to limited space for inland retreat. Additionally, 16.5% of sandy coasts are embedded within built environments, which are regions that are particularly susceptible to "coastal squeeze" and human-induced modification. A final notable finding is that 9.3% of the world's coastline features visible coastal defenses, with 59% of built-up coasts also containing coastal defenses and, conversely, 83% of defended coasts situated along built-up areas.

As an automated, data-driven method, the typology can be efficiently refined over time—through improved class definitions, targeted corrections, and expanded training data to reduce expert bias and geographic imbalance. Overall, the dataset may support numerous applications, including coastal change monitoring, erosion risk assessments, and broad-scale vulnerability mapping, thereby providing a critical foundation for coastal adaptation planning in response to accelerating climate change.

3.6 Availability Statements

Code and data availability All data, code, and models used in this study are openly available. The coastal typology dataset is published as partitioned, cloud-optimized Parquet files and can be

accessed through the CoCliCo STAC catalog. A static archive is also available via Zenodo: https://doi.org/10.5281/zenodo.15599096 (Calkoen et al. 2025b). The deep learning model code used to produce the typology is available through the open-source `CoastPy` package at https://github.com/COCLICO/coastpy (Calkoen et al. 2025f).

Sample availability The training samples used to develop the coastal typology model are released as the CoastBench dataset (Calkoen et al. 2025g), available under a CC-BY-4.0 license. The dataset includes annotated labels for sediment type, coastal type, and the presence of built environment and coastal defenses. It is available at https://doi.org/10.5281/zenodo.15800285. The dataset is accessible through a STAC collection and can be expanded via a custom web application. The source code for the application is available at https://github.com/florisCalkoen/coastapp.

4 Global coastal exposure and future erosion impacts on sandy shores

Abstract Coastal zones host productive ecosystems and are critical hubs for human activity, but they face compounding pressure from climate change and human development. While future impacts from coastal flooding are increasingly well-quantified, those from chronic erosion (i.e., long-term shoreline retreat) remain a gap in global climate assessments. Recent broad-scale studies typically focus on coastal erosion as a physical hazard, but do not quantify the impacts to ecosystems, livelihoods, infrastructure, and assets. This study provides the first global, asset-based impact assessment of future coastal erosion under different sea-level rise (SLR) scenarios by integrating probabilistic shoreline change projections with a high-resolution building footprint dataset. Our analysis first quantifies present-day coastal exposure, revealing that approximately 76 million buildings are within 1 km of the shoreline. Along this 253,000 km of "developed" coast, 10% has buildings within just 37 m of the shoreline. Moreover, our analysis reveals a human preference for developing the coasts most sensitive to erosion; development is disproportionately concentrated on sandy, gravel, and shingle coasts, where building density is 18% higher and buildings are sited over 30% closer to the shoreline than on other coastal types. Our analysis then focuses on sandy sediment plains and dune coasts (21% of the world's total coast), although data limitations reduced our final analysis to a 9.3% subset. For these locations, we define a future hazard zone that combines extrapolated historical shoreline-change trends, SLR-induced retreat, and an Empirically-derived Setback Zone (ESZ), defined at the 10th percentile of the shoreline-to-first-building distances. Under a high-emissions scenario, we project a median shoreline retreat of 135 m by 2100. This retreat, when considered together with the ESZ, places up to 2.4 million buildings at risk within our study area—i.e., more than 20% of the analyzed coast faces impacts. Finally, the results show that by 2050, impacts from future coastal erosion are largely a legacy of existing coastal development, with the ESZ and ambient change accounting for over 80% of buildings at risk, regardless of the emissions scenario. By 2100, however, SLR becomes the principal driver, showing the potential impact on coastal communities worldwide and highlighting the need for forward-looking coastal adaptation strategies that respond to climate change.

This chapter is based on the following article:

Calkoen, F. R., Luijendijk, A. P., Barli, P., Hemmes, J., Ranasinghe, R., & Nicholls, R. J. Global coastal exposure and future erosion impacts on sandy shores. *Manuscript in preparation*, 2025.

4.1 Introduction

The coastal zone, the dynamic interface between sea and land, hosts some of the planet's most productive ecosystems and has long been a hub for human settlement and economic activity (Small et al. 2003; Barbier et al. 2011). These zones now face intensifying pressure from two compounding trends. First, climate change is accelerating global mean sea-level rise (SLR), which is projected to reach approximately one meter by 2100 under a high emission scenario (SSP5-8.5), with ongoing rise thereafter, intensifying pressure on the coastal zone (Oppenheimer et al. 2019). Second, expanding human development concentrates assets and livelihoods in these vulnerable areas, often creating a fixed landward boundary (Cosby et al. 2024). This prevents the natural inland migration of the shoreline in response to SLR or other drivers, leading to a phenomenon known as "coastal narrowing" where both human and natural systems have progressively less space to adapt (Pontee 2013; Lansu et al. 2024). Rising sea levels will exacerbate coastal hazards including coastal flooding, salt water intrusion, and coastal erosion (Ranasinghe et al. 2021)—with the latter being a natural morphodynamic process of shoreline adjustment that has the potential to threaten infrastructure, human assets and also ecosystems where inland migration is not possible.

To understand and manage such climate-related hazards, the Intergovernmental Panel on Climate Change (IPCC) provides a framework that defines risk as a function of hazard, exposure, and vulnerability (Reisinger et al. 2020). Applying this conceptual framework to earlier studies, the assessment by Hinkel et al. (2013) provided an early, comprehensive impact analysis by coupling erosion projections with socio-economic exposure, quantifying consequences such as land loss and forced migration. More recently, broad-scale assessments have built upon satellite-derived data to produce hazard projections at a much higher spatial resolution. The global study by Vousdoukas et al. (2020b) exemplifies these advanced hazard assessments. Other related work has also quantified exposure, for instance at the national scale for Japan (Udo et al. 2017) and the continental scale for Europe (Athanasiou et al. 2020), by estimating metrics such as future beach and land area loss.

Building on these advances we here provide the first global impact assessment of buildings at risk from coastal erosion by integrating probabilistic shoreline projections with a high-resolution global building dataset. Our analysis proceeds in three stages. First, we quantify present-day coastal exposure and from these development patterns, we derive a geomorphology-specific "Empirical Setback Zone"—the setback distance that coastal communities implicitly maintain; a proxy for the dynamic zone that a coastal system purportedly needs to provide its defensive capacity or other (natural) services. Subsequently, we project future shoreline change for the sandy sediment plains and dune coasts (∼21% of the world's coast), before data limitations (primarily nearshore slope data) further reduced the study area to 9.3% (∼ 93,000 km). Finally, by integrating these probabilistic shoreline change projections with the ESZ and building footprints, we quantify the number of buildings potentially impacted by future coastal erosion and disaggregate this impact into its constituent components to understand the drivers of risk over the century.

4.2 Results

4.2.1 Global patterns of coastal exposure

Our global coastal exposure analysis reveals that a quarter of the world's coast ($\sim$ 253, 000 km) contains at least one building, per alongshore km, within one kilometer of the shoreline, summing up to approximately 76 million buildings in total. Further analysis shows that this development is often located close to the shoreline: for 10% of these developed coasts, the nearest building is located within just 37 meters of the shoreline (Table 4.1). This broad exposure, however, is not evenly distributed among the different types of coasts.

More detailed analysis shows an apparent human preference for settling on the more dynamic coastal systems. By using a machine-learning-derived coastal typology, that characterizes the sediment type as well as the geomorphological landform along the cross-shore coastal profile, we study coastal development patterns over different types of coasts, such as sandy sediment plains or dune coasts (Calkoen et al. 2025c). In the used coastal typology, sandy, gravel, and shingle beaches make up 40% of the global total (Calkoen et al. 2025c), yet they host more than half of all buildings found within 1 km of the shoreline (42.5 million). Consequently, building density on these soft coasts is 18% higher than the global average, and buildings are sited more than 30% closer to the shoreline, with a median distance of 140 m compared to the global median of 205 m (Fig. 4.1a). This concentration of coastal developments is the highest on sandy low-lying sediment plains, which are both the most densely developed (427 buildings/alongshore km) and have buildings nearest to the shoreline (median distance of 128 m) (Table 4.1, Fig. 4.1b).

From these global development patterns, this study introduces a novel, data-driven metric to quantify the de facto setback zones that coastal communities implicitly maintain: the "Empirical Setback Zone" (ESZ). This metric is defined as the 10th percentile of the shoreline-to-first-building distances for a given coastal type (e.g., sandy sediment plain or dune coast), computed across the 253,000 km of "developed" coast. This value provides an empirical proxy for the dynamic zone a coastal system purportedly requires to provide its (natural) defensive function or other services. However, it is likely that this represents a conservative estimate, as many perceive the benefits of living close to the beach outweigh the risks of erosion (Buck 2025). The ESZ varies significantly by geomorphology, reflecting different physical dynamics and development patterns; for example, it is 64 m for dune coasts but only 34 m for sandy sediment plains (Table 4.1)—a more detailed overview is provided in the Supplementary Information (Tables D.2, D.3).

Mapping historical shoreline change onto this developed coastal zone ($\sim$253,000 km) reveals that while a majority of soft coasts have been historically stable, a significant portion is already experiencing erosion (Fig. 4.2). Of the developed sandy, gravel and shingle and muddy shores with available data, approximately 50% have remained stable (i.e., changing at rates between -0.5 and +0.5 m/year). At the same time, at least 16,000 km of developed sandy coast and 7,800 km of developed muddy coast are eroding at rates faster than 0.5 m/year. While substantial, these historical trends represent a conservative baseline of the hazard. Relying on them alone would fail to capture the anticipated accelerating retreat driven by future SLR, which is paramount for an assessment of the future impacts of coastal erosion.

4.2.2 Future shoreline change

To quantify the impacts of future coastal erosion on buildings, our analysis first projects future shoreline change under various SLR scenarios. We projected this change using an equilibrium-profile approach (expressed via the Bruun rule) for sandy sediment plains and dune coasts (Bruun

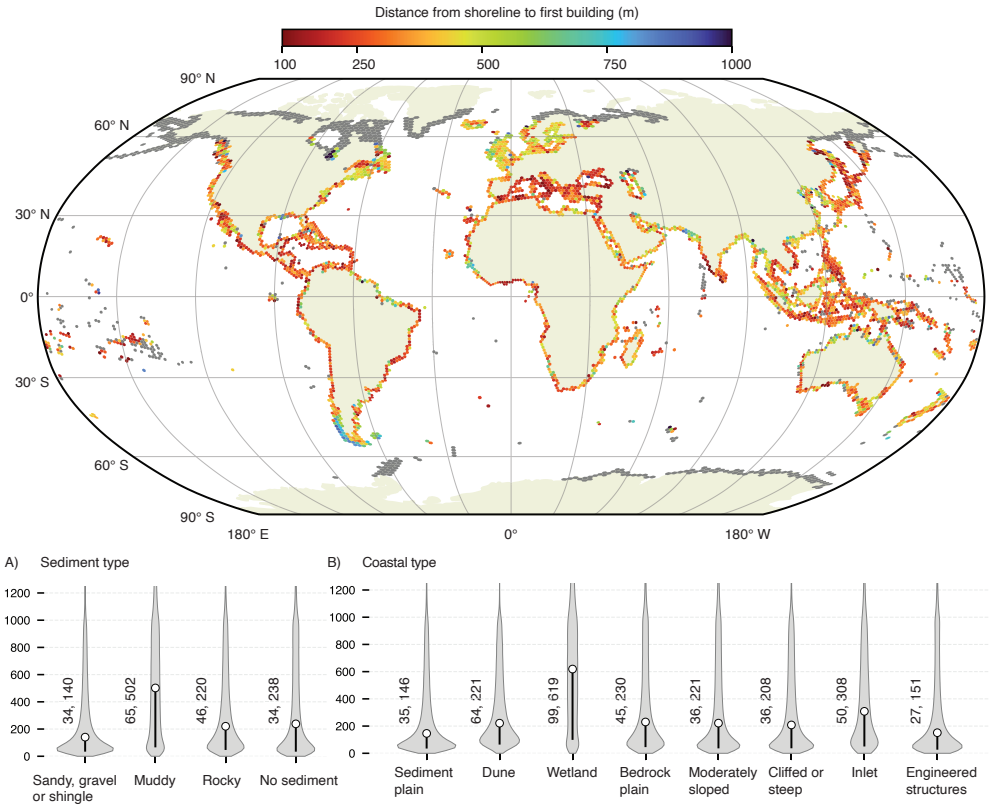

Figure 4.1: Global proximity of coastal development to the shoreline. The main panel maps the distance from the shoreline to the nearest building across all developed coastlines (∼253,000 km), aggregated by a hexagonal H3 grid for visualization. The color scale highlights variation in shoreline-to-building distance (100–1,000 m); grey areas indicate no buildings detected within 2,000 m of the shoreline. Inset panels show the distributions of these distances as violin plots, grouped by **a**, sediment type and **b**, coastal type. For each distribution, the width of the grey shape indicates the probability density. Overlaid on each violin is a box plot where the inner box indicates the 10th–50th percentile range and the white dot marks the median (50th percentile). The annotated numbers explicitly state these 10th and 50th percentile values. Basemap: Natural Earth.

Table 4.1: Global coastal exposure by sediment and coastal type. The table quantifies coastal exposure, first presenting global summaries and then providing a breakdown by sediment type. Metrics include the total length of developed coast (segments with at least one building within 1 km), the number and density of buildings within this zone, and the 10th and 50th percentile distances from the shoreline to the nearest building. For sandy, gravel and shingle coasts, the data are further detailed by two coastal types (dune and sediment plain), as these are used in the subsequent impact analysis. A detailed overview is provided in Supplementary Tables D.2, D.3.

	Developed Coast	Buildings (within 1 km of shoreline)		Distance to Nearest Building (from shoreline, m)	
	Length (km)	Total Count (#)	Density (/km)	10th percentile	Median
Global	**253,102**	**75.7M**	**299**	**37**	**205**
Sandy, gravel or shingle	**119,996**	**42.5M**	**354**	**34**	**140**
Dune	*7,612*	*2.2M*	*288*	*64*	*221*
Sediment Plain	*63,290*	*27.0M*	*427*	*34*	*128*
Muddy	39,995	8.0M	201	65	502
Rocky	35,790	8.3M	232	46	220
No Sediment	57,321	16.9M	295	34	238

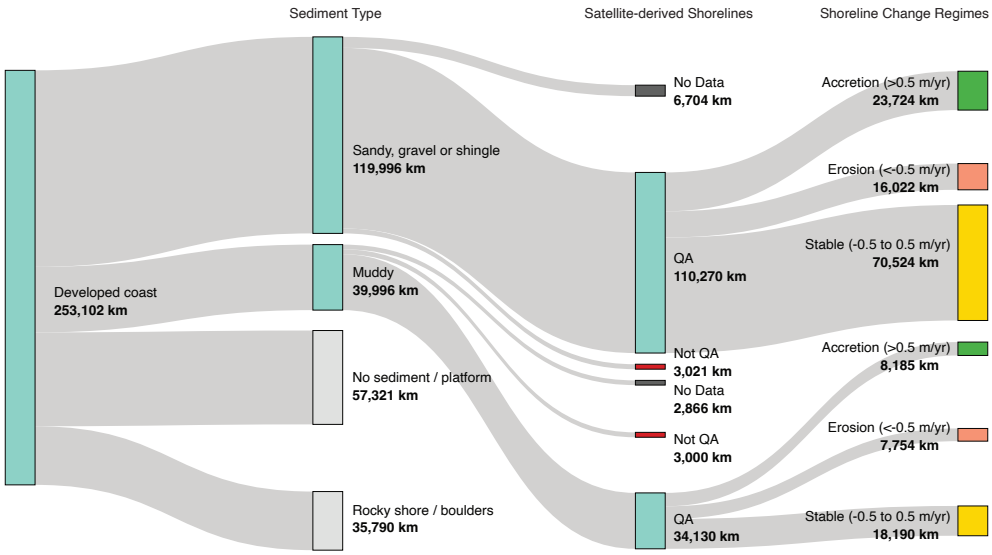

Figure 4.2: Distribution of developed coastlines by historical shoreline change. The diagram shows the classification of global developed coastlines (segments with buildings within 1 km of the shoreline) by their historical shoreline change trend. The flow begins by categorizing coastlines by sediment type (sandy, gravel or shingle; muddy; rocky; or no sediment). Soft-sediment transects (sandy and muddy) are then filtered based on the availability and quality of historical shoreline data. For transects with quality-assured data, shoreline behavior is classified as Erosion (change rate < -0.5 m/year), Stable (-0.5 to +0.5 m/year), or Accretion (> +0.5 m/year). Flow widths are proportional to the total coastline length (in kilometers) at each stage.

1962). This subset, representing ~ 21% of the world's coast, was further filtered based on the availability of high-quality historical shoreline data and nearshore slopes (see Methods). Our final analysis therefore covers 93,000 km, or ~9.3% of the global total (Fig. 4.6). For this study, we projected shoreline positions for 2030, 2050, and 2100 under three integrated scenarios (SSP1-2.6, SSP2-4.5, and SSP5-8.5), with 2010 as baseline. Following earlier probabilistic approaches for projecting shoreline change (Athanasiou et al. 2020; Vousdoukas et al. 2020b), our method combines two components: the continuation of historical ambient change (AC) and additional retreat induced by SLR, with uncertainties propagated using Monte Carlo sampling.

The projections indicate a global retreat of these sandy sediment plains and dune coasts by the end of the century, with the magnitude strongly dependent on the emissions pathway. While local historical ambient change (AC) varies widely, its contribution to the global average change is negligible—a median of only 1 m of progradation by 2100 (Table 4.2). Consequently, while the large variability in local outcomes is strongly influenced by AC, the global median retreat is almost entirely driven by SLR. By 2100, this median projected retreat ranges from 79 m (90% CI: 12 to 166 m) under the low-emissions SSP1-2.6 scenario to 135 m (90% CI: 60 to 250 m) under the high-emissions SSP5-8.5 scenario. While these projections are broadly consistent with previous global assessments (Vousdoukas et al. 2020b), a direct comparison requires acknowledging methodological differences. Our study uses a new ML-derived coastal typology that, most importantly, refines the analysis to sandy sediment plains and dune coasts. This geomorphologically-refined approach likely explains why our projected retreat is more severe at the 5th percentile (a retreat of 60 m) than in previous work. By excluding steep, cliffed coasts our analysis avoids the low retreat estimates that the equilibrium-profile model will project for these steep areas, which likely affected the statistics of shoreline retreat projections in earlier broad-scale assessments.

Table 4.2: Projected shoreline change for the global ~93,000 km of sandy sediment plains and dune coasts. The table presents projected shoreline change (in meters) for dune and sediment plain landforms for 2030, 2050, and 2100 under three emissions scenarios. Projections are decomposed into a scenario-independent ambient change (AC) component and a scenario-dependent sea-level rise (SLR) component. Values represent the median change with the 5th–95th percentile uncertainty range in parentheses. The Total column is derived from the full distribution of summed AC and SLR projections; its median is therefore not the direct sum of the component medians. Negative values indicate erosion.

		SSP1-2.6		SSP2-4.5		SSP5-8.5	
Year	**AC**	**SLR**	**Total**	**SLR**	**Total**	**SLR**	**Total**
2030	0 (-1, 1)	-17 (-33, -3)	**-17 (-35, -2)**	-17 (-33, -3)	**-17 (-34, -2)**	-18 (-34, -4)	**-18 (-35, -3)**
2050	0 (-4, 5)	-34 (-63, -8)	**-35 (-69, -7)**	-36 (-65, -10)	**-37 (-71, -9)**	-41 (-72, -13)	**-42 (-78, -12)**
2100	1 (-12, 15)	-77 (-150, -17)	**-79 (-166, -12)**	-97 (-176, -32)	**-98 (-192, -30)**	-134 (-233, -64)	**-135 (-250, -60)**

4.2.3 Buildings at risk from future coastal erosion

By combining our shoreline projections with global building data, we quantify the impacts of future coastal erosion on coastal buildings. This analysis is conducted for the approximately 93,000 km of sandy sediment plains and dune coasts (9.3% of the global total; Fig. 4.6). Following the IPCC framework, we identify first-order impacts by determining which buildings fall within a potential

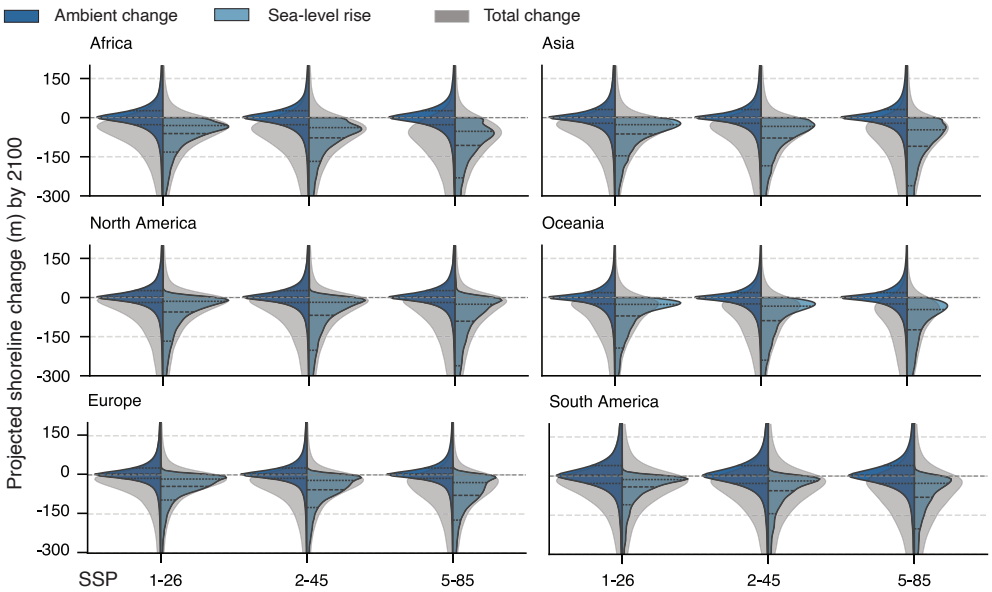

Figure 4.3: Projected shoreline change by continent in 2100. The figure shows probability distributions of projected shoreline change for the year 2100, grouped by continent and under three integrated emission scenarios (SSP1-2.6, SSP2-4.5, and SSP5-8.5). Each distribution is decomposed into contributions from ambient change (AC; dark blue), sea-level rise-induced retreat (SLR; light blue), and the resulting total change (grey). Negative values indicate shoreline retreat.

future hazard zone. This zone is defined as the sum of three components (Fig. 4.5): (1) shoreline retreat from extrapolated historical ambient change (AC), (2) additional retreat from SLR, and (3) the geomorphology-specific ESZ, to account for the fact that buildings are not placed right upon the shoreline. The ESZ represents the lower end of the de facto setback that coastal communities implicitly maintain (e.g., 64 m for dune coasts, 34 m for sandy sediment plains; Tab. 4.1) and serves as an empirical proxy for the dynamic zone a coastal system purportedly requires to maintain its defensive capacity or provide its other (natural) services.

The analysis reveals that the impacts of future coastal erosion are substantial and accelerate throughout the 21st century. Within the analyzed area, the number of buildings potentially impacted is projected to reach about one million by 2050 under all scenarios, rising to between 1.6 million (SSP1-2.6) and 2.4 million (SSP5-8.5) by 2100 (Figure 4.4c). A shift in the drivers of this impact is projected to occur over the century. By 2050, impacts are largely a legacy of present-day coastal development patterns, with the ESZ and AC accounting for over 80% of buildings at risk, regardless of the emissions scenario. This large contribution from AC stems from the aggregation of impacts at all locations with historical erosion; an effect that is masked in the near-zero global median by the counterbalancing effect of accretion elsewhere (Tab. 4.2). By 2100, however, SLR becomes the principal driver. Under the high-emissions scenario, SLR alone accounts for nearly 1.2 million buildings potentially impacted—almost 50% of the total—highlighting the growing impact of climate change on coastal communities throughout the 21st century.

When aggregated by continent, based on country-level assignments, the distribution and nature of these impacts are highly uneven. Asia faces the largest absolute impact, with 1.17 million buildings at risk of future coastal erosion by 2100 under SSP5-8.5, representing 49% of the global total. In contrast, Europe faces the highest proportional threat, with nearly 30% of its analyzed coastline potentially impacted (Supplementary Figs. D.1, D.2). This distinction is important for macro scale prioritization of adaptation planning; for example, while previous studies identified Australia as having widespread shoreline retreat (Vousdoukas et al. 2020b), our analysis finds the potential impacts on buildings there are relatively low, as much of this retreating coastline is undeveloped (Supplementary Fig. D.5). Finally, the results show that impacts intensifies over time as erosion encroaches on more densely developed inland areas. Globally, while the length of "potentially-impacted" coastline increases by 40% between 2050 and 2100 (SSP5-8.5), the number of buildings at risk more than doubles (a 132% increase), a trend seen across all continents (Supplementary Figs. D.1, D.3).

4.2.4 Outlook

By integrating shoreline projections with a global building dataset, this study provides the first worldwide impact assessment of future coastal erosion on buildings throughout the 21st century, taking a step forward towards climate risk assessments of coastal erosion. By 2100, this places between 1.6 million (under a low-emissions SSP1-2.6 scenario) and 2.4 million (under a high-emissions SSP5-8.5 scenario) buildings at risk of being impacted by erosion. This figure represents a conservative lower-bound estimate, as the model's application was refined by the ML-derived coastal typology, which constrained its use to sandy sediment plains and dune coasts (21% of the world's coast) before data limitations (primarily nearshore slope data) further reduced the study area to 9.3% ($\sim$ 93, 000km). This asset-level assessment is analogous to a direct-impact inundation map in flood modeling (Hauer et al. 2019); its purpose is to identify potential risk hotspots at a broad scale, thereby guiding where more detailed local studies are needed, rather than to inform specific local interventions. The finding that SLR becomes the dominant driver of risk after 2050 has an important implication for climate adaptation planning: relying on historical trends alone is no longer

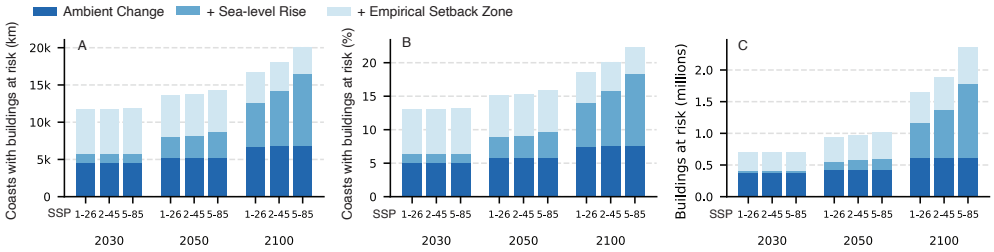

Figure 4.4: Buildings at risk from projected future shoreline retreat. The figure presents buildings at risk for sandy sediment plains and dune coasts, quantified by **a**, the total length of coast with buildings impacted (km); **b**, this length as a percentage of the analyzed coastline; and **c**, the total number of buildings at risk (in millions). Estimates are shown for 2030, 2050, and 2100 under three emissions scenarios (SSP1-2.6, SSP2-4.5, and SSP5-8.5). The stacked bars disaggregate the total statistic into its constituent components: the Empirical Setback Zone (ESZ), sea-level rise (SLR), and ambient change (AC), where the AC component represents the cumulative impact from all coastlines with a historical erosion trend.

sufficient, and proactive coastal management strategies that account for accelerating climate change are required to protect coastal communities worldwide.

An important consequence of this refined approach is that the model's application was first constrained by the coastal typology to sandy sediment plains and dune coasts (21% of the world's coast), before data limitations (primarily nearshore slope data) further reduced the study area to 9.3%. This highlights the need to develop multi-model approaches for projecting shoreline change in other coastal environments. Besides equilibrium-profile models that simulate the *direct* impact of SLR on sandy coasts, more comprehensive global assessment to the impacts of future shoreline retreat, future work would benefit to also integrate the *indirect* impacts of SLR, which are primarily driven by changes in coastal sediment budgets. This requires incorporating models that simulate the increased sediment demand of tidal inlets and estuaries (e.g., Stive et al. 2003; Ranasinghe et al. 2013; Bamunawala et al. 2021; Ranasinghe et al. 2025), the reduced sediment supply from dammed rivers (Dunn et al. 2019), the retreat of soft cliffs (Walkden et al. 2008), and the unique bio-geomorphic dynamics of muddy coasts. For other environments, such as stable, low-lying bedrock plains, the primary hazard from SLR is inundation rather than erosion, meaning the path forward for these coasts is better integration with global coastal flood risk assessments. A hierarchical classification that situates local geomorphologies within broader landform complexes (e.g., estuaries, barrier islands) offers a promising path to structure such a multi-model approach (French et al. 2016).

4.3 Methods

Our analysis integrates several global datasets to quantify buildings impacted by future coastal erosion under different SLR scenarios. This assessment proceeds in a three-stage process. First, we quantify present-day coastal exposure by combining a global building dataset with the global coastal transect system and a high-resolution coastal typology. Second, for sandy sediment plains and dune coasts, we generate probabilistic shoreline projections for the years 2030, 2050, and 2100 under multiple integrated climate scenarios (SSP1-2.6, SSP2-4.5, and SSP5-8.5). Third, we identify buildings potentially impacted by future coastal erosion by determining which of the present-day coastal

buildings fall within a future hazard zone. The methods used for each of these steps are specified in more detail in the following subsections.

4.3.1 Data

This study integrated several global datasets quantify the buildings impacted by future coastal erosion a three-stage process (Table 4.3). The spatial foundation was the Global Coastal Transect System (GCTS) (Calkoen et al. 2025f), which provides a consistent grid of over 11 million cross-shore transects at 100-m alongshore resolution. To quantify exposure, we used building footprints from the Overture Maps Foundation (Buildings theme; release 2025-04-23.0) (Overture Maps Foundation 2025). This dataset was chosen for its completeness, as it combines community-contributed data (e.g., OpenStreetMap) with multiple machine-learning (ML) derived products to provide a comprehensive global building stock. The physical characteristics of the coast at each transect, such as sediment type and geomorphological landform, were obtained from a global ML-derived coastal typology (Calkoen et al. 2025c).

To model shoreline change, we used datasets describing historical change and future SLR projections. Historical shorelines from 1984-2024 were obtained from the ShorelineMonitor dataset (Luijendijk et al. 2018). Future SLR projections were taken from the IPCC Sixth Assessment Report (AR6) dataset (Fox-Kemper et al. 2021). Specifically, we used the geocentric (i.e., not corrected for Vertical Land Motion (VLM)) AR6 distributions, as our framework accounts for VLM through the extrapolated historical AC term, thereby avoiding double-counting. The nearshore slope, a parameter required for modeling SLR-induced shoreline retreat with the Bruun rule, was sourced from the Global Coastal Characteristics (GCC) dataset (Athanasiou et al. 2023). All datasets were mapped onto the GCTS to create a consistent analytical framework.

Table 4.3: Datasets used in the global assessment of buildings at risk from shoreline retreat. The table lists the global datasets used for the analysis, including their provider, purpose, and version.

Dataset Name	Purpose	Data Provider	Version
Global Coastal Transect System	Foundational dataset	Deltares	2024-08-12
Global Coastal Typology	Sediment type & coastal geomorphology	Deltares	2025-07-08
Overture Buildings	Quantifying coastal exposure	Overture Maps	2025-04-23.0
ShorelineMonitor	Historical shoreline change (AC)	Deltares	1984-2024
AR6 Sea-Level Rise Projections	Bruun rule application (SLR)	IPCC / CoCliCo	AR6
GCC Nearshore Slopes	Bruun rule application (SLR)	JRC / Deltares	v1.1
Global Coastal Grid	Scalable data processing	Deltares	2025-01-01

Quantifying present-day coastal exposure

To quantify present-day coastal exposure, we analyzed building data within a 1-km inland zone from the shoreline for each transect in the GCTS. A coastline was defined as "developed" if at least one building was present within this zone. For each developed transect, we calculated several exposure metrics, including the distance from the shoreline to the nearest building, the total number of buildings in this zone, and the total building footprint area. All spatial calculations were performed using the local UTM projection specified for each transect to ensure metric accuracy.

From these exposure data, the ESZ was derived. This metric serves as a data-derived proxy setback zone that coastal communities implicitly maintain, reflecting the minimum zone a coastal sys-

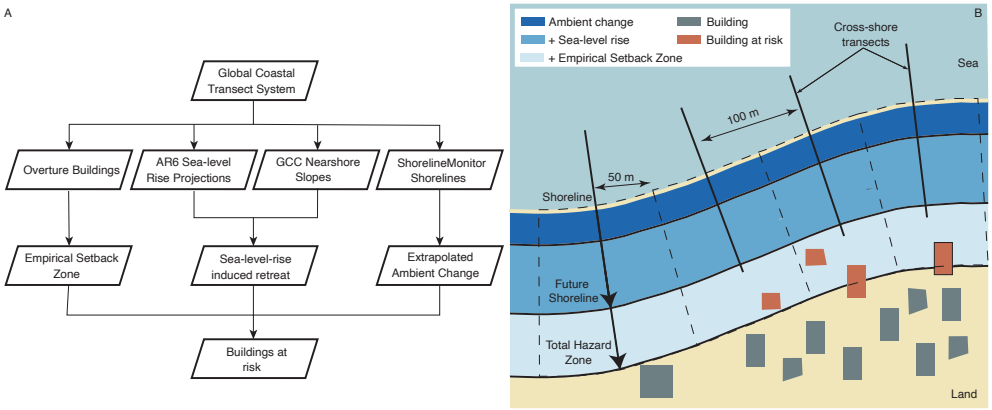

Figure 4.5: Schematic for assessing buildings at risk from future shoreline retreat. The total hazard zone is defined by combining three components: projected shoreline change from ambient trends (AC), additional retreat from sea-level rise (SLR), and the geomorphology-specific Empirical Setback Zone (ESZ). Buildings at risk (shown in red) are quantified by identifying those footprints that intersect with this hazard zone (green-grey).

tem purportedly requires to maintain its defensive capacity or provide its other (natural) services. To calculate the ESZ, all developed transects were grouped by their coastal class (i.e., sediment and geomorphological type), and probability distributions of the shoreline-to-building distances were generated for each class. The ESZ for each class was defined as the 10th percentile of its respective distribution. This geomorphology-specific ESZ was subsequently used as a component in the future impact assessment.

Projecting shoreline change

We projected future shoreline positions probabilistically, following earlier methods (Athanasiou et al. 2020; Vousdoukas et al. 2020b). This approach combines two primary components of shoreline change: (1) the continuation of historical ambient change (AC) and (2) additional retreat induced by SLR. To robustly propagate and combine uncertainties from these components, we performed a multi-stage Monte Carlo simulation (n=1,000) for each transect.

The SLR component was modeled using an equilibrium-profile model (the Bruun rule), which is suitable for coasts composed of non-cohesive, unconsolidated sediments that can freely adjust their profile (Bruun 1962). To identify these areas, we used a global ML-derived coastal typology to select all transects classified as having a "sandy, gravel, or shingle" sediment type on either "sediment plains" or "dune coasts" (Calkoen et al. 2025c). These two geomorphological classes were selected as they are most likely to consist of erodible material, unlike other classes from the typology such as cliffed coasts or the lithologically-ambiguous "moderately sloped" class, which could consist of either unconsolidated sediment or resistant rock. The "sandy, gravel, or shingle" category is treated as a single class because current satellite-derived classifications cannot reliably distinguish between these sediment sizes on the resolution ($\sim$ 10 m) of open satellite imagery (Calkoen et al. 2025c). However, the geomorphological selection of only sediment plain and dune coasts effectively removes most shingle beaches because they typically front steeper, more rocky environments. Nevertheless

we acknowledge that due to classification errors in the underlying ML-derived coastal typology, the Bruun rule may have been applied to some coasts where it is not strictly valid, and vice-versa.

These sandy sediment plains and dune coasts were further filtered to include only transects with high-quality historical shoreline data and valid nearshore slope information (i.e., slopes between 1/300 and 1/5). This filtering, detailed in Figure 4.6, resulted in a final analysis extent of ~93,000 km. For this final set of transects, the total shoreline change was derived through a numerical convolution of the AC and SLR components. For each of the 1,000 Monte Carlo realizations, a value for AC was drawn from a normal distribution defined by the observed historical rate (i.e., the regression slope) and its corresponding standard error. This was summed with a SLR driven retreat value, which was drawn with replacement (bootstrapped) from the full distribution of potential shoreline retreat values generated by applying the Bruun rule to the AR6 SLR ensemble. This process yields a final empirical probability distribution of total shoreline change for each transect, from which we extracted the median (p50) and the 90% confidence interval (CI; p5, p95).

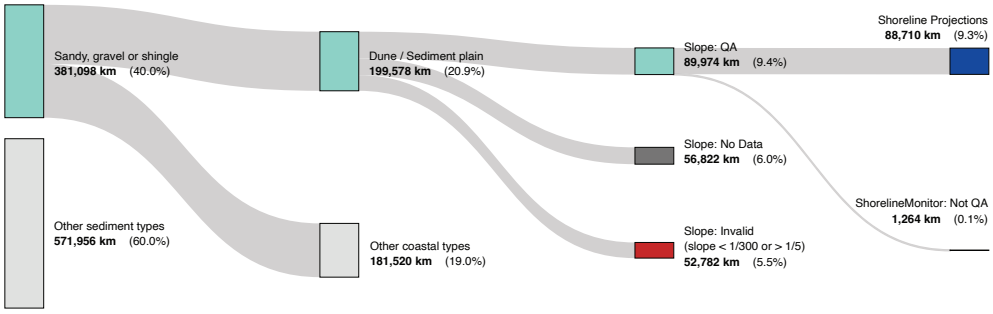

Figure 4.6: Filtering cascade for future shoreline projection analysis. The diagram shows the progressive filtering used to select the portion of the global coast for shoreline projections. The flow begins with all sandy, gravel, or shingle coasts, which are first constrained by geomorphology to "Dune / Sediment plain". This subset is then filtered based on the availability of valid nearshore slope data (i.e., 1/300 < slope < 1/5). A final filter is applied to retain only transects with a high-confidence historical shoreline change record, defined as those with a sufficient number of data points to have a statistically robust linear trend. The final category, "Shoreline Projections", represents the coastline passing all filters. Flow widths are proportional to the total coastline length (in kilometers) at each stage.

4.3.2 Historical Shoreline Change

The historical Ambient Change (AC) component was derived from the ShorelineMonitor dataset, which provides an annual satellite-derived shoreline record from 1984 to 2024 (Luijendijk et al. 2018). To extract a time-series of annual shoreline positions, the continuous shorelines were mapped onto the GCTS (Calkoen et al. 2025f). The primary shoreline series was then extracted from all intersections by tracking back in time and selecting the furthest offshore shoreline observation, unless this observation was identified as part of another shoreline using directional coastline proximity data (e.g., to distinguish the opposite side of an inlet) (Calkoen et al. 2025d) or as an outlier using Median Absolute Deviation (MAD) outlier detection. This process yields a quality-assured time series of annual shoreline positions for each transect, from which a long-term rate of change was calculated using ordinary least-squares linear regression.

Quantifying buildings impacted by future erosion

To quantify the number of buildings impacted from future coastal erosion, we identified which present-day buildings fall within a future hazard zone. This analysis was conducted for the sandy sediment plains and dune coasts where shoreline projections were produced, i.e, $\sim$ 93,000 km (Fig. 4.6). For each transect, the total hazard zone was defined as the cumulative landward extent of three components (Fig. 4.5): (1) projected shoreline retreat from extrapolated historical AC, (2) additional retreat from SLR, and (3) the ESZ (Table 4.1). A building was considered to be "impacted" if its footprint intersected this total hazard zone. To attribute the total impacts to their underlying drivers, we then employed a sequential analysis: for each transect, we first identified buildings impacted solely by the AC component, then the additional buildings captured when including the SLR component, and finally, the further buildings captured by adding the ESZ. This attribution method allowed for the disaggregation of the total impact presented in the results.

4.3.3 Methodological Limitations

The worldwide assessment is subject to a set of (cascading) uncertainties and assumptions, from the core physical model to the final risk quantification. The shoreline projections are based on an equilibrium-profile model (expressed through the Bruun rule), which has been discussed extensively and has well-documented assumptions (SCOR Working Group 89 1991; Cooper et al. 2004; Zhang et al. 2004; Woodroffe et al. 2012). The two most limiting assumptions are that the coast is composed of unconsolidated non-cohesive sediments and that it functions as a closed system with no net longshore sediment transport gradients. Our approach is explicitly designed to address these assumptions, but the methods and data used introduce their own uncertainties.

To address the "erodible profile" assumption, we used a novel, ML-derived coastal typology to define the study area. This foundational dataset has known classification errors that can propagate through the analysis (Calkoen et al. 2025c). For instance, the classification groups sandy, gravel, and shingle beaches together, as these are difficult to resolve in open satellite imagery ($\sim$10 m resolution). While shingle beaches may respond to SLR as equilibrium profiles, the Bruun rule is specifically advocated for sandy coasts. Although our geomorphological selection of sediment plains and dune coasts effectively removes most shingle beaches, some may remain, and the equilibrium-profile model may overestimate retreat here. Furthermore, the nearshore slope is derived from a static global dataset. This introduces uncertainty from both measurement error and the dataset's inability to capture natural and temporal variability, which cascades directly into the shoreline retreat projections.

In line with earlier studies, the "closed system" assumption was addressed by combining the Bruun rule with an extrapolated AC term (Athanasiou et al. 2020; Vousdoukas et al. 2020b). This AC component represents a composite historical signal that implicitly includes factors not modeled by the equilibrium-profile model, such as longshort transport gradients and human interventions. By extrapolating this term, we assume that these historical drivers will continue unchanged. A further limitation is the potential for "double counting," as the historical AC term likely contains some signal from past SLR. Isolating this relatively small signal from a 40-year satellite record that includes large fluctuations from storms (Galgano et al. 1998) and climate oscillations (Vos et al. 2023b) is an ongoing scientific challenge. Our approach, which treats AC and future SLR as independent, follows currently standard practice in broad-scale assessments but may lead to projections on the higher or lower side of the spectrum. Finally, our approach assumes that the historical drivers embedded in the AC term, including VLM, will continue linearly, which may not hold true in areas with accelerating or decelerating subsidence or new large-scale human interventions.

The analysis provides a first-order impact assessment, but stops short of a full climate risk assessment because it does not quantify vulnerability. A complete vulnerability analysis would require evaluating both a system's sensitivity to harm (e.g., structural resilience based on building codes) and its adaptive capacity (the socioeconomic ability of a community to respond). By not including these components, our framework implicitly treats vulnerability as uniform globally. Furthermore, the scope is deliberately focused on buildings as a high-resolution proxy for coastal exposure, which excludes other critical infrastructure and ecosystems. The building count itself is also an imperfect proxy for population or asset value, as a single footprint can represent anything from a beach house to a multi-story apartment.

These cascading limitations—from model assumptions to input data uncertainties and the scope of the impact assessment itself—mean the true uncertainty in the analysis is likely wider than the reported confidence intervals. Therefore, the results should not be used to inform local design but should be interpreted as a transparent baseline of the number of buildings potentially at impacted, upon which more detailed studies can be based.

4.3.4 Software and code availability

Data processing and analysis were conducted using open-source, scalable, geospatial and mostly Pythonic software ecosystem (Pangeo). Geospatial analytics were performed primarily using DuckDB (v1.2) with its spatial and H3 extensions (Raasveldt et al. 2019). The analysis of large data volumes was enabled by the adoption of cloud-optimized data formats (e.g., Parquet), with the several data partitions managed through Spatio-Temporal Asset Catalogs (STAC). The core analytical components developed for this study are integrated into the open-source `coastpy` package (Calkoen et al. 2025f).

4.3.5 Data availability

The datasets generated during this study are available in two public Zenodo repositories. The first, the Global Coastal Transect Repository (GCTR), contains the present-day coastal exposure metrics (e.g., shoreline-to-building distances) and the directional coastline proximity data (DOI: 10.5281/zenodo.16928268) (Calkoen et al. 2025d). The second repository contains the probabilistic future shoreline projections for 2030, 2050, and 2100 under the SSP1-2.6, SSP2-4.5, and SSP5-8.5 scenarios (DOI: 10.5281/zenodo.16928880) (Calkoen et al. 2025e). To enhance downstream coastal analytics, all generated data are linked to the Global Coastal Transect System.

5 Synthesis and Perspective

5.1 Synthesis: A Global Assessment of Future Coastal Erosion Impacts

This thesis set out to to develop a global framework for assessing the future impacts of coastal erosion by combining the vast observational datasets nowadays available from satellites with the power of cloud technology and the capabilities of artificial intelligence (AI). The final chapter (Ch. 6) summarizes the principal conclusions from this research. This synthesis now integrates the findings together to provide a more nuanced view. It discusses how the research advances the field, critically reflects on the assumptions and limitations of the methodology, and considers the broader implications of the results for both coastal science and management.

5.1.1 From a Hazard-Centric to Impact Assessment

Coastal erosion is a natural process that becomes a societal risk where it potentially impacts something of value, i.e., livelihoods, assets, infrastructure and ecosystems. This thesis contributes to a growing body of research aiming to quantify the impacts of coastal erosion at broad spatial scale. An earlier global study by Hinkel et al. (2013), for instance, provided an assessment of coastal erosion impacts in the form of land loss and forced migration, but its analysis was limited by the coarse, segment-averaged data available at the time. More recently, broad-scale assessments (Athanasiou et al. 2020; Vousdoukas et al. 2020b) have leveraged satellite-derived data products to project future shoreline change at higher spatial resolution. While these studies represent an advance in coastal erosion hazard mapping they do not quantify direct impacts on individual assets. The research in this thesis aims to further advance the field, by moving from broad-scale hazard-focused studies towards worldwide asset-level quantification of coastal erosion impacts. This is achieved by generating georeferenced, probabilistic shoreline projections that allow for a direct intersection with global building footprint data. Compared to earlier studies that estimated impacts by multiplying land loss area by average population densities (cf. Hinkel et al. 2013), this asset-level approach offers a more granular understanding of the impacts of future coastal erosion. Overall, the final first-order assessment of the impacts of future coastal erosion on buildings (Ch. 4) is the culmination of a three-part solution developed in this thesis. It was made possible by first refining application of the used equilibrium-profile model to sandy sediment plains and dune coasts by using a novel, machine learning (ML)-derived geomorphological coastal typology (Ch. 3), which in turn was produced by the novel, cloud-native analytical framework that is capable of applying applying modern AI (i.e., deep learning (DL)) at broad spatial scales (Ch. 2).

The coastal exposure analysis (Ch. 4) reveals that a quarter of the world's coast ($\sim$253 000 km) contains buildings within the first kilometer of land, hosting approximately 76 million buildings in total. This global picture of a heavily developed coastal zone is reinforced by the Empirical Setback Zone (ESZ), a concept introduced in this thesis to quantify the de facto setback that coastal communities implicitly maintain. The analysis shows that for 10% of the developed global coast,

the first building is located less than 37 m from the shoreline, leaving very limited space for natural coastal processes to unfold. Furthermore, this development is disproportionately concentrated on soft, erodible coasts. Sandy, gravel, and shingle coasts, for instance, make up only 40% of the global coastline but host over half all coastal buildings, with a building density 18% higher than the global average. The concentration is particularly high on sandy sediment plains, which represent just 21% of the global coast yet host about 30 million buildings (38%).

This high degree of coastal development (coastal building exposure) directly translates into significant future impacts when combined with projections of future shoreline erosion under different sea-level rise (SLR) scenarios. By 2100, between 1.6 million (under a low-emissions SSP1-2.6 scenario) and 2.4 million (under a high-emissions SSP5-8.5 scenario) buildings are projected to be impacted by erosion. While substantial, this figure represents a conservative lower-bound estimate, as the model's application was refined to specific coasts; the ML-based coastal typology first constrained its use to sandy sediment plains and dune coasts (21% of the world's coast), before data limitations (primarily nearshore slope data) further reduced the study area to 9.3% ($\sim 93,000$ km). Moreover, it is important to acknowledge the limitations of this impact assessment. The analysis is subject to cascading uncertainties from the input data (ML-based coastal typology and nearshore slope data) and the structural assumptions of the equilibrium-profile model itself, as discussed in the following sections. Furthermore, the assessment quantifies impacts by combining the hazard with exposure but stops short of a full climate risk assessment. Such an assessment typically considers the full range of retreat probabilities and would as well require a complete analysis of vulnerability, which includes not only a system's sensitivity to harm but also its adaptive capacity—the socioeconomic ability to cope with and respond to the hazard. By treating all areas as equally vulnerable, the framework provides first-order estimates of potential future impact f coastal erosion under different SLR scenarios but it does not capture the vast differences in how various communities can manage that risk. Finally, the hazard definition does not explicitly model episodic storm-driven erosion, which is often the direct cause of damage and may intensify under climate change.

5.1.2 The Synergy of Open Data, Cloud Technology, and AI

The findings presented in this thesis, from the development of a global coastal typology to the final impact assessment, were enabled by a methodological shift that helps bridge the gap between broad but coarse ("everywhere") and detailed but local ("anywhere") analyses. This framework operationalizes the synergy between open data, cloud technology, and AI by adopting a strategy that decouples data storage from compute, centered on three principles. First, bringing code to the data through data-proximate computing is essential, as petabyte-scale satellite archives are too large to move. Second, for the large datasets produced and used, adopting cloud-optimized formats (e.g., GeoParquet and Cloud-Optimized Geotiffs) described in standardized metadata catalogs (e.g., Spatio-Temporal Asset Catalog (STAC)) is critical for performance and interoperability. Finally, orchestrating these workflows within a flexible, open-source software ecosystem (Pangeo) provides the modularity and flexibility that helps analyze complex coastal phenomena at a global scale. The efficiency gains from this approach are substantial; an analysis that is 700 times faster than traditional methods is the difference between a theoretically possible study and a practically feasible one. For example, this open, cloud-native architecture enabled the analysis of over a billion coastal observations from heterogeneous datasets in less than a day, a scale of computation that is difficult to replicate using traditional methods.

In this framework, the global coastal transect system (GCTS) is the foundational spatial grid for all analyses. By using a consistent 100 m alongshore spacing, it corrects the latitudinal (zonal) bias

present in previous global transect systems, resulting in more accurate global statistics. For example, this work estimates that 40% of the world's coasts are sandy, gravel, or shingle beaches, a higher figure than the 31% previously reported (Luijendijk et al. 2018). This difference is consistent with the system's design; estimates for Africa, for instance, are approximately the same between both studies, corroborating the idea that the zonal bias affects global statistics but has less impact on continents near the equator. Moreover, its derivation from a more recent (2023) OpenStreetMap coastline provides better global coverage than previous transect systems, including many more islands. Besides providing more robust statistics and better global coverage, the system has several metadata attributes, such as its bounding box and local coordinate reference system, that enhance scalable geospatial analytics. The adoption of this grid is part of a broader methodological choice, that is a shift to a cross-shore perspective departing from earlier alongshore segmentation approaches (e.g., Salman et al. 2004; Vafeidis et al. 2008). This perspective, now more widely adopted (Luijendijk et al. 2018; Finkl et al. 2020; Vousdoukas et al. 2020b; Athanasiou et al. 2023), is better suited for integrating geomorphological variability across the land-sea profile and for historical shoreline change mapping, which simplifies data-driven projections.

Complementing this data structure, the adoption of an open-source software ecosystem (Pangeo) offers higher flexibility than monolithic, (i.e., all-in-one) cloud platforms like Google Earth Engine (GEE), although it is also more demanding for the user, as many high-level computations must be implemented themselves. The open, flexible, and scalable nature of the Pangeo ecosystem enables the integration of specialized, modern AI methods at a broad spatial scale. The creation of the coastal typology in this thesis is a primary example: a custom convolutional neural network (CNN) was trained and globally deployed, a task not readily supported by all-in-one platforms. This ability to control the entire analytical chain is often crucial for innovative science, especially when experimenting with non-standard tools. Such an approach is also more transparent, as the underlying components can be inspected, and more reproducible, because open architectures are more accessible. However, this places the responsibility for maintaining the architecture on the coastal scientist rather than on a commercial provider. Realizing this potential therefore depends on a stable foundation of open-source software, reliable cloud infrastructure, and open data policies. As recent events have shown, none of these can be taken for granted. These three components, which form the basis of the ongoing digital transformation in coastal science, are discussed more extensively in the "Perspective" section of this thesis (Sec. 5.2).

5.1.3 Deep Learning for Coastal Mapping

The global coastal typology (Ch. 3) provides new, globally consistent estimates of the world's coastal composition. The analysis revises earlier statistics on sediment distribution, estimating that 40% of the global coast is fronted by sandy, gravel, or shingle beaches—a higher figure than the 31% previously reported, with the difference largely attributed to the new, zonally-unbiased transect system used in this thesis. Combined with the 21% of coasts found to be muddy, the typology reveals that over 60% of the world's ice-free coast consists of soft, erodible sediments. Moving backshore, the data show that cliffed coasts are the most common landform (33%), while dune systems are globally rare (3%). Methodologically, this work demonstrates that by leveraging DL, more complex coastal landforms can be mapped than with previously used tree-based learning methods. However, the classification has limitations; it is more robust for sediment type (F1-score of 0.76) than for the more complex coastal type classification (F1-score of 0.67). Nevertheless, the typology provides a unique, global cross-shore perspective on coastal geomorphology. For instance, it reveals that 20% of the world's sandy, gravel, and shingle beaches are backed by cliffs. This configuration may physi-

cally limit the natural capacity of these beaches to migrate inland in response to accelerating SLR, a finding that also highlights a key limitation of this satellite-based approach, as the classification does not account for cliff lithology and therefore cannot distinguish between erodible "soft" cliffs and resistant "hard" cliffs.

The 'sediment type' classification is conceptually robust, building on established frameworks (Hayes 1967), and performs particularly well for sandy, gravel, or shingle beaches (F1-score of 0.84) and muddy sediments (F1-score of 0.86). In contrast, performance is lower for coasts fronted by rocky platforms (F1=0.64) and those without sediment (F1=0.70), highlighting an area for future improvement. The high scores for sandy and muddy sediments are comparable to those reported in other broad-scale studies (Hulskamp et al. 2023); however, a direct comparison is challenging as each study uses its own training data. This underscores the need for a shared, community-curated benchmark dataset to properly evaluate and compare different classification methods. A second area for future work is separating shingle from sandy and gravel beaches. While crowd-sourced datasets like "Coastwards" offer a valuable resource for this task, both these further refinements to the sediment type will be challenging due to the relatively coarse 10 m spatial resolution of open satellite imagery.

In contrast, the 'coastal type' classification is a functional schema developed for this thesis that draws on several cross-shore coastal typologies (e.g., Sharples et al. 2009; French et al. 2016; Finkl et al. 2020). It represents a pragmatic first step that requires further alignment with other emerging coastal classifications (e.g., Hanson et al. 2025; Nyberg et al. 2025). Due to the higher complexity of classifying geomorphological landforms, its accuracy scores are lower than for the sediment type classification; some classes remain ambiguous, such as "moderately sloped" (F1-score of 0.55), and the schema does not yet distinguish critical sub-types like erodible "soft" versus resistant "hard" rock cliffs.

Another key limitation is that the supervised ML classification is a direct product of its training data, "CoastBench". This dataset was labeled predominantly ($\sim 95\%$) by a single expert—the author of this thesis—and therefore the coastal typology necessarily reflects the views and biases of its creator. These potential classification errors propagate through the entire analytical chain, contributing to unquantified uncertainties in the final results. For instance, the findings of the coastal exposure analysis, when grouped by coastal type, may differ from reality. Also, errors in classifying sandy sediment plains and dune coasts mean the erosion model may have been applied where its assumptions are not met (false positives) or omitted from areas where it is applicable (false negatives). While a detailed analysis of the model's precision and recall for these key classes could offer qualitative insight into the error's direction, a direct quantitative correction is difficult given the potential bias in the training data.

Addressing these limitations points toward the need for a community-driven effort centered on two complementary components. The first is the development of a more robust and standardized classification schema, created with input from a diverse group of coastal engineers, morphologists, and geologists. Such an effort would help refine ambiguous classes like "moderately sloped" and expand others, such as the classification of coastal defenses, into a more sophisticated scheme suitable for detailed risk assessments. The second component is the creation of a diverse, global training data repository, a task that could be accelerated with contributions from the broader community, including citizen scientists (Wehn et al. 2021). The automated nature of the DL method ensures that the typology can be iteratively improved as the community converges on these shared standards. While creating such a resource is a significant undertaking, it represents a foundational investment. As algorithms advance and more data comes available, a community-owned benchmark dataset remains an enduring scientific asset capable of guiding future generations of CoastalAI models, a topic explored further in the "Perspective" section of this thesis (Sec. 5.2).

5.1.4 A More Refined Application of a Pragmatic Model

The global assessment of future coastal erosion impacts presented in this thesis is built upon a chain of methods, each with its own assumptions and limitations. This section provides a critical reflection on this methodological chain, beginning with the equilibrium-profile model (the Bruun rule) applied for shoreline projections, and extending to the cascading uncertainties that propagate through the analysis as a whole.

The Bruun rule (Bruun 1962), a simple equilibrium-profile model that predicts shoreline retreat due to SLR alone, is an idealized model that rests on a series of strict assumptions: a closed, two-dimensional (cross-shore) sediment system; a profile composed entirely of erodible sand; an instantaneous equilibrium response to SLR; and no sediment loss to other processes such as overwash or aeolian transport. Because these assumptions are rarely met in nature, the model's validity and application have been the subject of a long-standing debate in coastal science (e.g., Bruun 1962; Bruun 1988; SCOR Working Group 89 1991; Cooper et al. 2004; Zhang et al. 2004; Davidson-Arnott 2005; Ranasinghe et al. 2009; Woodroffe et al. 2012; Cooper et al. 2020; Vousdoukas et al. 2020a; D'Anna et al. 2021), with some arguing it should not be used at all (Cooper et al. 2004).

However, even critics of the model agree on the fundamental principle that the entire active profile of an open, sandy coast will respond to SLR, typically resulting in landward retreat (Cooper et al. 2004; Cooper et al. 2020). For these specific coastal environments, the equilibrium-profile approach is a significant conceptual improvement over a simple "bathtub" model because it incorporates this essential physical principle. Adopting it is also a pragmatic choice, born from the need to make shoreline projections at broad scales and on centennial timescales for adaptation policy—precisely the "centennial gap" where predictive understanding is weakest (Woodroffe et al. 2012). While all models are simplifications of a complex reality, their value lies in providing understanding where perfect prediction is not yet possible (Box 1976; Cartwright 1983). This thesis therefore adopts this pragmatic position and, as the following section details, takes the next step of addressing the model's assumptions more rigorously than in previous broad-scale studies (e.g., Hinkel et al. 2013; Athanasiou et al. 2020; Vousdoukas et al. 2020b) by taking into account the backshore geomorphology.

Given this context, the scientific value of this work lies not in "proving" the Bruun Rule, but in demonstrating *a more refined application*. This was achieved by addressing two core assumptions of the equilibrium-profile model. First, in line with recent assessments, the "closed system" assumption is addressed by combining the SLR-induced retreat with an extrapolated ambient change term, which implicitly accounts for factors like net alongshore sediment-transport gradients and vertical land motion (VLM) (Vousdoukas et al. 2020b). The primary contribution of this thesis, however, is a novel solution to the "erodible profile" assumption. By using the coastal typology, the model's application was constrained to those geomorphological settings—sandy sediment plains and dune coasts—where its use is most physically plausible. Furthermore, georeferencing these constrained projections to a high-resolution (100 m) transect grid elevates the analysis from a coarser hazard assessment (e.g., 5 km) to a more precise, asset-level impact assessment. An important consequence of this refined approach is that the model's application was first constrained by the coastal typology to sandy sediment plains and dune coasts (21% of the world's coast), before data limitations (primarily nearshore slope data) further reduced the study area to 9.3%. This reduction was driven in approximately equal parts by the valid filtering criteria for the nearshore slope ($1/300 < \text{slope} < 1/5$) and by differences in the underlying transect systems. This highlights the need to develop multi-model approaches for projecting shoreline change in other coastal environments, a topic detailed in the next section.

Also with this refined framework, several uncertainties remain and cascade through the analysis. The shoreline projections are highly sensitive to the nearshore slope and SLR (Thiéblemont et al. 2021). The use of a static global dataset for the nearshore slope parameter introduces measurement uncertainty due to its coarse resolution and fails to capture the natural temporal variability (aleatoric uncertainty) of the profile. Conceptually, the structural uncertainty of the equilibrium-profile model itself is significant; it does not account for the rate of SLR, which may induce different coastal responses like barrier overstepping or drowning rather than the gradual retreat the model assumes (Woodroffe et al. 2012; Green et al. 2014). These uncertainties are compounded by classification errors from the coastal typology. For instance, the analysis selected "sandy, gravel, or shingle beaches" (sediment type) on "sediment plains" or "dune coasts" (coastal type). While this geomorphological filtering effectively removes most shingle beaches, some may remain, resulting in higher retreat rates at such locations than would likely occur in reality. More broadly, classification errors mean the Bruun rule may have been applied where it should not be (false positives) or omitted where it is applicable (false negatives). Finally, the analysis provides a first-order assessment of the impacts of future coastal erosion on buildings but stops short of a full climate risk assessment, which would require analyzing vulnerability, including sensitivity and adaptive capacity. These cascading limitations mean the results should be interpreted as a transparent, first-order assessment for identifying potential erosion impacts, rather direct input for local design. A more systematic analysis of how these uncertainties propagate through the assessment was beyond the scope of this thesis but remains important to address in future research.

5.1.5 Implications for Science

After decades of research (1962 - present), an unequivocal answer to how coasts will respond to SLR on centennial timescales cannot be given (Woodroffe et al. 2012; Le Cozannet et al. 2014; Ranasinghe 2020). To better understand, model, and predict the impacts of SLR, this section discusses two complementary research paths. The first path involves a systematic evaluation of modeling approaches—process-based, numerical, and data-driven (Vitousek et al. 2017)—to describe a way towards global shoreline projections. The second path suggests a renewed focus on the geological past to address fundamental assumptions about how coastal systems respond to the high rates of accelerated SLR projected for this century. Finally, the section briefly reflects on the application of modern AI as demonstrated in this thesis, a topic that is explored in greater detail in the final "Perspective" (Section 5.2).

Towards Global Shoreline Retreat Projections A direct path to improving global shoreline projections is to further refine the equilibrium-profile approach used in this thesis. In this study, the model's application was carefully constrained. Based on the coastal typology, its use was first limited to sandy sediment plains and dune coasts, which are estimated to comprise 21% of the world's coast, although this figure is itself subject to the classification errors discussed previously. Within this subset, further data limitations—primarily the lack of reliable nearshore slope data—reduced the model's final application to just 9.3% of the global coast. Addressing the limitations of nearshore bathymetry data therefore presents a promising path forward. Bruun-based shoreline change projections are highly sensitive to the active profile slope (Thiéblemont et al. 2021), a parameter that, from approximately 2050, together with SLR becomes the dominant factor in future erosion impacts. Replacing coarse global datasets like GEBCO with higher-resolution data from navigational charts or satellite-derived bathymetry would primarily address "measurement uncertainty", yielding more realistic confidence intervals and increasing the model's spatial coverage. However, it

is reasonable to assume that even with a perfect coastal classification and complete bathymetry data, this equilibrium-profile approach would remain inapplicable to the vast majority of the world's coasts, a finding consistent with previous broad-scale assessments and review studies (SCOR Working Group 89 1991; Zhang et al. 2004).

Alongside simple equilibrium-profile models, more advanced numerical models that resolve a range of physical processes offer a more detailed view of shoreline evolution. These models move beyond simpler equilibrium concepts by linking shoreline change to wave and storm-driven processes, such as alongshore transport gradients (Roelvink et al. 2020) or cross-shore profile response (Ranasinghe et al. 2012). While these provide a more detailed view of long-term shoreline change, they are computationally more demanding and have not yet been implemented at a global scale. This focus on the local response to hydrodynamic forcing, however, raises the question of whether centennial-scale shoreline change is dominated by these local effects or by system-scale changes in the overall sediment budget. The progression of hybrid frameworks like CoSMoS (Vitousek et al. 2023b) illustrates one path forward, demonstrating how physical models are evolving to integrate observational data (i.e., satellite-derived shoreline (SDS)) with separate components for wave-driven change and long-term SLR-induced retreat.

Data-driven approaches are also unlikely to provide centennial insights for future shoreline change (e.g., Calkoen et al. 2021; Gomez-de la Peña et al. 2023; Adusumilli et al. 2024). These models depend on long observational records, which for shoreline modeling primarily means the satellite archives of the Landsat program. While this 40-year record is excellent for studying episodic events and decadal climate oscillations (e.g., Vos et al. 2023b; Warrick et al. 2023), it is likely too short to resolve the slow, long-term trend of SLR-induced erosion. The core challenge is the low signal-to-noise ratio in the observational record; it can take 80 years or more of data to reliably distinguish the underlying erosion trend from the much larger variability caused by storms and recovery cycles (Galgano et al. 1998). Consequently, with the SLR signal so deeply buried in this natural variability, reliably capturing the long-term trend with models trained on the current satellite record remains a challenge.

Developing a broader, more comprehensive framework for centennial shoreline projections would involve moving beyond the direct impact of profile adjustment to capture the indirect, system-scale impacts of future drivers like relative sea-level rise (RSLR) and human activities on coastal sediment budgets (Cowell et al. 2003). Such an approach suggests adding another hierarchical level of classification to the one used in this thesis, moving from local geomorphologies to broader coastal landform complexes, such as barrier-inlet systems, deltas, and estuaries (French et al. 2016). This would situate each transect within its broader system context, creating a nested understanding from sediment type, through coastal geomorphology, to landform complex. With a system-level classification in place, a multi-model approach becomes possible. For coastal systems dominated by sediment sharing, the framework could focus on quantifying future changes to the sediment budget. For example, a model like ASMITA (Stive et al. 2003) or G-SMIC (Bamunawala et al. 2021) could estimate the increased volumetric demand of sinks like tidal inlets as their accommodation space grows with RSLR, while other components could account for human-induced changes in sediment supply from sources like dammed rivers (Dunn et al. 2019). For other major environments, different models could be applied; for instance, the large proportion of cliffed coasts could be addressed using specific cliff erosion models (Walkden et al. 2008), while for stable, low-lying hard coasts, the primary hazard is inundation. A central research challenge then becomes coupling these components by modeling the sediment pathways that connect sources to sinks, and translating the net volumetric changes into shoreline retreat. Overall, such a multi-model approach is becoming increasingly feasible as earth observation (EO) provides not only the basis for large-scale classification, but also first-order estimates for several input parameters, such as intertidal extent for estuary models (Mur-

ray et al. 2018), turbidity to estimate sediment fluxes (Grandjean et al. 2024), and coastal slopes from satellite lidar or optical imagery (Vos et al. 2021).

Learning from the Geological Past Understanding the centennial-scale impact of SLR can be advanced by looking beyond the instrumental record to coastal geoscience. While historical records, such as 18th-century paintings, can extend our observational record beyond the satellite era (Baart et al. 2011), they primarily document a period of relative sea-level stillstand. They offer little insight into how coastal systems respond to the high rates of RSLR projected for this century—expected to reach 7 to 12 mm/yr by 2100 (Fox-Kemper et al. 2021)—which have no precedent in recent millennia. The only empirical evidence of coastal behavior under such conditions lies in the geological record. As Woodroffe et al. (2012) argue, echoing words commonly attributed to the founder of modern geology James Hutton (1788): "If the present is the key to the past, then the past, seen from the context of the present, can be a guide to the future." This approach is becoming increasingly promising as extensive field campaigns and improved dating techniques yield sea-level index points with millimeter-per-year precision (Vacchi et al. 2021; Hijma et al. 2025). Such paleo-analogues have already identified quantitative thresholds for ecosystem collapse, showing that mangroves cannot keep pace with RSLR rates exceeding ∼6 mm/yr (Saintilan et al. 2020). Although dating past sandy coastal facies is more challenging than dating biogenic sequences, it remains a worthwhile avenue to explore. Reconstructing past shoreline responses from geological analogues offers a dual benefit: it can provide a deeper understanding of system behavior to guide coastal adaptation, while also generating validation data to test the fundamental assumptions within the next generation of centennial shoreline models.

The Adoption of CoastalAI The application of DL and, more broadly, modern AI to large, heterogeneous spatiotemporal data, as demonstrated in this thesis, helps to establish a new capability for coastal science. However, the axiom that any ML model is only as good as the data it was trained on is particularly relevant in this context. The coastal typology presented in Chapter 3 was trained on "CoastBench" (Calkoen et al. 2025g), a ML training dataset that is (so-far) predominantly ($\sim 95\%$) collected by a single expert, i.e., the author of this thesis. Consequently, the resulting classification necessarily reflects the views of its main creator. This highlights a fundamental challenge and opportunity for the field: to move beyond single-expert datasets towards a community-driven approach for creating ML training datasets. This is reflected upon more extensively in the final "Perspective" section (Sec. 5.2).

An important component of such an approach would be a robust and standardized classification schema, developed with input from a diverse group of coastal engineers, morphologists and geologists, potentially with help from citizen scientists. Such an effort could build upon the schema used in this thesis by refining the current classes—for instance, by better articulating degrees of coastal development, replacing ambiguous classes like "moderately sloped" with more descriptive categories, and incorporating lithology to capture resistance to erosion. Another path worth exploring is to add a higher hierarchical level to situate local geomorphologies within broader coastal landform complexes, such as barrier-inlet systems and deltas, a step that would help enable the system-scale modeling approaches discussed previously. Composing such a schema and curating the corresponding training data is a considerable undertaking, yet this is important work, that will have long-lasting scientific value. Even with large, pre-trained geospatial foundation models (Jakubik et al. 2023), which reduce the amount of labeled data needed for fine-tuning, high-quality datasets remain important for guiding their domain applications like coastal classification. Ultimately, while algorithms and computational power will advance, a community-owned classification schema and its correspond-

ing training data would remain a worthy scientific asset that can be used to train generations of CoastalAI.

5.1.6 Implications for Coastal Management

This thesis has several implications and recommendations for coastal management, emerging not only from the final assessment of buildings at risk but also from the datasets developed to support it. The research contributes to coastal management by offering several globally consistent data products: a new Global Coastal Transect System (GCTS); a compatible, high-resolution (100 m) coastal typology; several metrics on coastal exposure; and a first-order impact assessment of future coastal erosion. This impact assessment provides probabilistic shoreline projections under different SLR scenarios, although its scope is carefully constrained; the ML-derived coastal typology first limited the model's application to sandy sediment plains and dune coasts (21% of the world's coast), before data limitations (primarily nearshore slope data) further reduced the study area to 9.3%. By correcting previous bias present in earlier transect systems, this work also offers more accurate statistics, resulting in different estimates of global sediment distribution. In this work the world's coasts that are sandy, gravel, or shingle is estimated at 40%, a higher figure than the 31% of sandy beaches previously reported (Luijendijk et al. 2018), with the difference largely attributed to improvements made to the used underlying transect system. The results also show that a quarter of global coastlines host approximately 80 million buildings within the first kilometer of land (Calkoen et al. 2025a). This development is disproportionately concentrated on the most dynamic coasts; beaches composed of soft, easily erodible sandy, gravel, shingle or muddy sediments make up 60% of the global total but contain 50 million buildings in the first kilometer of land. An even higher concentration occurs on sediment plains and dunes, which represent only 21% of the global coast yet host about 30 million buildings. This intrinsic high exposure is reinforced by the narrow setback societies implicitly maintain. For 10% of the developed global coast, the first building is located less than 37 m from the shoreline, which is a very limited space for natural coastal processes to take place, while these often act as the coast's natural defensive capacity for the societies that live on them.

This high present-day exposure directly translates into significant future impacts when combined with projections of future shoreline erosion. By 2100, between 1.6 million (under a low-emissions SSP1-2.6 scenario) and 2.4 million (under a high-emissions SSP5-8.5 scenario) buildings are projected to be impacted by future erosion. While substantial, this figure represents a conservative lower-bound estimate, as the model's application was first refined by the ML-based coastal typology to sandy sediment plains and dune coasts (21% of the world's coast), before data limitations (primarily the lack of reliable nearshore slope data) further reduced the study area to 9.3%. These findings highlight the importance for coastal adaptation, providing a framework for managers to contextualize the strategic responses of protect, accommodate, retreat or advance (Nicholls et al. 2025). The analysis also reveals a shift in the drivers of future impacts; while near-future impacts are largely a legacy of present-day coastal development, SLR is projected to become the dominant driver of erosion impacts throughout the 21st century.

The primary application of this work for management is as a first-order assessment for identifying potential erosion impacts. Analogous to risk assessment in flood modeling, this asset-level analysis is comparable to a direct-impact inundation map, which is more targeted and conservative than broad exposure metrics like the low-elevation coastal zone (LECZ) (Hauer et al. 2019). Responsible use of this assessment, however, requires understanding its key limitations. First, the underlying coastal typology has lower accuracy for some classes and is not yet a community standard. Second, the impact assessment rests on the well-documented assumptions of the simplified equilibrium-profile

model used for the projections. Finally, the analysis quantifies the physical impacts on buildings but stops short of a full climate risk assessment, which typically considers the full range of retreat probabilities and would as well require a complete analysis of vulnerability, including a community's socioeconomic capacity to adapt to and cope with the hazard.

The insights from this thesis were enabled by a methodological approach that demonstrates the potential of integrating modern AI with open satellite data and cloud technology. The identified limitations, particularly the reliance on a single-expert training dataset, highlight an opportunity for the management and scientific communities to collaborate on developing shared ML training datasets. Sustaining this progress also depends on continued support for open data policies, open-source software, and reliable cloud infrastructure. The challenges and potential solutions regarding these final points are discussed in detail in the final "Perspective" section of this thesis (Section 5.2).

5.2 The Future of Global Coastal Science

A Synergy Between Data, Cloud and AI

Abstract

The digital transformation of coastal science is entering a new phase, defined by the challenge of integrating artificial intelligence with petabyte-scale satellite archives. Progress, however, is now limited by the absence of a shared framework for integrating the field's increasingly diverse data, tools, and models. This perspective argues that overcoming this challenge requires a shift toward an open, community-governed "bazaar" ecosystem that realizes the synergy between open data, cloud technology, and AI. It proposes three principles to guide this transition: a focus on interoperable architectures, a common coastal data strategy, and a community-wide commitment to creating open training data. Adopting these principles is an important step toward developing the robust, transparent and more intelligent analytical tools needed to support climate adaptation and coastal management.

5.2.1 The Digital Transformation of Coastal Science

Coastal science is in the midst of a data revolution. This shift began with the opening of multi-decadal satellite data archives, which provided an unprecedented amount of historical observations on the Earth's surface (Wulder et al. 2012). The volume (petabytes) of this data makes traditional "download-and-analyze" workflows impractical. The solution emerged from new data processing paradigms (Dean et al. 2008) that, when incorporated into user-friendly platforms, collocate massive datasets with analysis tools (Gorelick et al. 2017). This new ecosystem revolutionized geoscience, enabling the creation of the first high-resolution global maps of phenomena like forest cover change (Hansen et al. 2013). These innovations soon made their way into other domains like surface water monitoring (Donchyts et al. 2016b; Pekel et al. 2016), until the potential of satellite data for coastal science was demonstrated by global applications to historical shoreline monitoring and tidal flats mapping (Luijendijk et al. 2018; Murray et al. 2018).

However, it is useful to characterize the methods that were central to this first wave of transformation. While these early studies commonly leveraged machine learning, the algorithms used were primarily traditional methods developed from the 1960s to 1980s, such as decision trees (Breiman et al. 1984), later widely adopted in implementations like Random Forests and Support Vector Machines (Cortes et al. 1995; Breiman 2001). It is helpful to distinguish these methods from the DL models that define the current era of AI (Mitchell 1997; Goodfellow et al. 2016). This modern approach, where foundational work on neural network architectures—from early work on backpropagation and recurrent networks (Rumelhart et al. 1986; Hochreiter et al. 1997) to the development of transformer models (Vaswani et al. 2023)—combined with massive datasets and parallel computing (Krizhevsky et al. 2012; Le 2013), has led to significant advances in fields from protein structure prediction (Jumper et al. 2021) to weather forecasting (e.g., Lam et al. 2023). Recognizing this distinction helps to reveal that the digital revolution in Earth System science since the 2010s was primarily facilitated by new modes of data access and computation, and *not* by the widespread adoption of modern AI. This raises a question: given the demonstrated potential of modern AI in other scientific domains, why have its applications in coastal science been comparatively limited?

The application of deep learning at scale in coastal science remains limited, not from a lack of potential, but due to the unique challenges of geospatial data (Zhu et al. 2017; Reichstein et al. 2019; Tuia et al. 2025). Unlike the standardized images used in many mainstream AI applications, satellite data are heterogeneous, multi-sensor spatiotemporal data stored in petabyte-scale archives, with atmospheric effects and temporal gaps that complicate analysis. Beyond the nature of the data itself, further challenges arise in learning from it. High-quality ground truth is not only scarce but often imperfectly aligned and semantically inconsistent. Furthermore, models trained on one sensor or region may not generalize well to others, a problem known as domain shift. These barriers help explain why CoastalAI—here defined as the scientific sub-discipline that integrates methods from artificial intelligence with the domain knowledge of coastal science to understand, model, and predict the dynamics of the coastal system—remains a frontier. This perspective argues that overcoming these barriers can be enabled by a shift towards a community-governed coastal analytics ecosystem and proposes three guiding principles for it: an interoperable architecture, a common data strategy, and a commitment to open training data.

5.2.2 The Synergy of Open Data, Cloud Technology, and AI

The next phase of the digital transformation in coastal science is likely to be characterized by the widespread adoption of AI to better understand, model, and predict the dynamics of the coastal system. This perspective argues that realizing this potential depends on a community-driven effort

to integrate three interdependent pillars: open data, cloud technology, and AI. While each component is powerful, their synergy within a shared framework is what enables the development of robust, interoperable, and transparent analytical coastal AI systems. This section details these three pillars and the open ecosystem that integrates them, before the next section elaborates on three guiding principles for building and sustaining such an ecosystem.

I. Open Data Publicly funded satellite archives, such as those from the Copernicus and Landsat programs, are foundational public goods for equitable, reproducible EO. Today their use is ubiquitous, with applications from monitoring to socioeconomic risk assessments, disaster response and defense. History shows that open access is not guaranteed; the United States Geological Survey's 2008 policy to make Landsat data free marked a turning point, enabling the broad-scale analyses that define modern EO (Wulder et al. 2012). This precedent highlights the importance of safeguarding open data policies against proposals to commercialize satellite data, which would threaten all applications in their current form.

II. Cloud Technology Cloud technology is paramount for managing the petabyte-scale archives that open data policies have created. Beyond storage, it has facilitated a paradigm shift away from "download-and-analyze" workflows towards approaches that bring computation directly to the data, supported by elastic cloud compute and containerized environments. These workflows are enabled by cloud-optimized data formats (e.g., COG, Zarr), standardized metadata catalogs (e.g., STAC), and interoperable application programming interfaces (APIs) that expose data in a consistent way. Together, these contribute to the creation of analysis-ready data (ARD), a concept operationalized in frameworks like the Australian Geoscience Data Cube (Lewis et al. 2017). By enabling tools to request only required data portions and to scale computation elastically, cloud-native geospatial makes continental- and global-scale analyses more practically feasible and reproducible.

III. Artificial Intelligence As established, the initial digital revolution in Earth System science was driven by traditional machine learning, whereas the current era of AI is characterized by the success of deep learning models (Tuia et al. 2025). These models are particularly suited to the scale and complexity of modern Earth observation data because they can learn hierarchical representations directly from large volumes of heterogeneous spatiotemporal data. This capability reduces the need for extensive, data-specific preprocessing and supports an "end-to-end" learning approach, where the model itself discovers the most relevant predictive patterns from raw data. For coastal science, this means evolving from simple traditional machine-learning methods towards answering complex questions about the dynamics of a coastal system (e.g., de Heer et al. 2025).

The Integrating Ecosystem While each component is powerful, their synergy within a shared framework is what enables the development of robust, interoperable, and transparent analytical coastal AI systems that works at scale. Combining these three pillars into a functional whole depends on the software ecosystem chosen for their integration. The available options can be framed by the classic open-source analogy of the "Cathedral and the Bazaar" (Raymond 1999). Monolithic platforms like GEE are analogous to the cathedral: a centrally-planned, all-in-one structure that is functionally rich but can be rigid, may lead to vendor lock-in, and offers a constrained analytical environment. In contrast, frameworks like the "Open Data Cube" and open-source communities such as "Pangeo" and "Radiant Earth" are analogous to the bazaar: a decentralized ecosystem of interoperable tools built on open standards (Lewis et al. 2017; Hamman et al. 2018). This modular

approach presents two distinct advantages. First, it allows researchers to control the full analytical chain, a flexibility typically advantageous for integrating the custom (AI) models used for novel coastal applications or scientific discovery. Second, as data archives become increasingly scattered across multiple cloud providers, the bazaar model is well-suited to support the distributed or federated analytics required to operate across these independent backends (Abernathey et al. 2021; Traub et al. 2021; Backeberg et al. 2022; Mohr et al. 2025). Such systems depend not only on technical interoperability, for instance through standardized processing APIs like openEO (Schramm et al. 2021), but also on community-driven semantic harmonization to align on common data strategies (Mohr et al. 2025). For these reasons, an open, community-driven model is a more promising path for data-driven coastal science.

5.2.3 Guiding Principles for a Community-Governed Ecosystem

The previous section discussed how an open, community-driven "bazaar"—that realizes the synergy between open data, cloud technology, and AI—is a more promising path for data-driven coastal science than a monolithic "cathedral." Realizing the potential of this open model, however, requires more than just assembling these components; it depends on a shared idea of how to build and sustain the digital ecosystem. This section explores three guiding principles that can help the community govern this ecosystem: an interoperable architecture, a common data strategy, and a commitment to open training data.

An Interoperable Architecture A decentralized "bazaar" model functions best when its data, tools, and models are interoperable, allowing them to work across different platforms and providers. Adopting an open, interoperable architecture is a response to the reality that Earth observation archives are scattered across different cloud providers, with no single platform hosting all data (Mohr et al. 2025). An interoperable architecture facilitates the integration of the custom models and diverse datasets that modern geospatial AI depends on (Tuia et al. 2025). Reliance on single, non-interoperable platforms creates barriers to this goal or has associated risks, as illustrated by three recent examples. The first is the volatility of cloud platforms, exemplified by the abrupt discontinuation of Microsoft's Planetary Computer Hub in June 2024. The platform was unique, collocating massive EO datasets with the flexible, open-source Pangeo software ecosystem. While it had a higher entry barrier than some alternatives (e.g., GEE), it offered full control over the analytical chain. However, with only a few weeks' notice, the compute hub was retired due to internal corporate security policy changes. This decision, while not affecting the underlying data archives, disrupted research projects that relied on its integrated compute environment. The migration of ongoing work (e.g., Calkoen et al. 2025c) to an institutional high performance computing (HPC) cluster revealed a new set of challenges. While the open-source Dask software (Rocklin 2015) could be reconfigured to run on an institutional HPC system, the physical infrastructure was not designed for cloud-native workflows. The network capacity of the university was insufficient for the data-intensive communication with nearby cloud servers, causing network congestion on the university's supercomputer. This illustrates that also modern HPC systems may not be well-equipped for the cloud-native workflows that process petabyte-scale data repositories in modern coastal EO applications. An additional disadvantage was that this workflow became specific to that institution and therefore was not easily shared or reproduced elsewhere. Finally, with proprietary platforms there is the risk of a gradual shift in the terms of service, as seen with GEE. Launched in 2010 with a mission to support forest monitoring, GEE has progressively transitioned towards a commercial product, culminating in its integration into the Google Cloud Platform in 2022 and the formalization of

paid tiers for non-research use. Besides the risk of future discontinuation or costs, "one-stop-shop" platforms have a scientific constraint: their server-side infrastructure is often proprietary, and users are limited to the methods available in its client libraries. Implementing novel or specialized algorithms not provided out-of-the-box is significantly more complex than in an open-source ecosystem, creating a "walled playground." These examples highlight that for research programs and operational climate services, reliance on single commercial or institutional platforms presents considerable risks. Open frameworks are preferable, particularly as they are now available, as demonstrated by the Open Data Cube (ODC) initiative. The ODC is not a centralized service but a framework for nations and organizations to build their own data cubes, empowering them to manage and analyze Earth observation data on their infrastructure of choice. One of its implementations, Digital Earth Australia, has effectively demonstrated how this platform-independent approach can deliver operational, continental-scale analyses (Bishop-Taylor et al. 2021) or tools (Bishop-Taylor et al. 2025) without locking users into a single provider.

Community-Governed Data Strategy A decentralized "bazaar" of tools and data functions effectively when its participants share a common language. A community-governed data strategy can establish this by promoting both technical interoperability and semantic harmonization (Mohr et al. 2025). Technical interoperability allows systems to exchange and process data, while semantic harmonization ensures that the meaning of that data is unambiguous. As coastal science has transitioned from a data-poor to a data-rich field, the importance of such standards grows. The history of the Argus coastal imaging network (Holman et al. 2007) illustrates the challenges of data management without modern standards: decades of imagery were collected, yet today this data resides in fragmented archives that are difficult to access with modern tools. These data management challenges are not new. The climate and ocean modeling communities faced similar issues and addressed them through the Climate Forecast (CF) conventions (Eaton et al. 2003). Originating in the 1990s, CF established a standard way to describe gridded and discrete data and introduced a common vocabulary—the CF Standard Name Table—that provides precise definitions for variables. This shared semantics enabled international efforts like the Climate Model Intercomparison Project (CMIP) (Eyring et al. 2016). For a modern, cloud-native coastal ecosystem, a similar path is possible. The STAC specification provides a standard for data discovery, but it does not describe the internal schema of the data itself. Achieving semantic harmonization therefore requires a domain-specific data specification. The recently launched Field Boundaries for Agriculture ("Fiboa") initiative offers a pragmatic model that the coastal community might follow (fiboa contributors 2025). It defines a simple core schema and uses a flexible extension model to add richer attributes. A similar "simple core, rich extensions" model could be applied to coastal science. For coastal analytics, a specification could define a foundational data unit (e.g., a set of alongshore coastal stations) with a core schema containing only essential information like a unique identifier, geometry, bounding box, and administrative divisions. Community-developed extensions could then attach domain-specific data, such as a 'satellite-derived shoreline' extension with attributes for shoreline chainage, timestamps, and tidal stage, or a 'geomorphology' extension with a coastal typology. When data producers adopt such standards, they simplify the preprocessing steps for the entire community, making datasets interoperable and easier to compare and benchmark. This, in turn, allows for standardized processing workflows, like the "recipes" from the Pangeo Forge initiative, that create ARD by default (Stern et al. 2022), moving from case studies to specialized coastal data services. Adopting and maintaining a standardized, community-governed specification for common tasks like shoreline monitoring is an important step toward making the growing volume of coastal data a cohesive and interoperable resource.

Open Training Data With an interoperable architecture and common data standards in place, the final principle for a coastal "bazaar" is a commitment to creating high-quality, open training data. This remains a primary challenge for adopting modern AI in coastal science (Zha et al. 2025). Modern AI is particularly effective because it can learn complex patterns directly from vast quantities of raw data, bypassing the manual feature engineering of earlier methods (Reichstein et al. 2019). The performance of these models, however, depends entirely on the quality of their training data. While creating public training datasets is becoming standard practice in related fields like flood mapping (e.g., Bonafilia et al. 2020; Notarangelo et al. 2025), coastal science has few such resources, with only fragmented datasets available for tasks like waterline monitoring and coastal classification (Seale et al. 2022; Buscombe et al. 2023; Calkoen et al. 2025c). Such training datasets represent a new form of scientific asset; while algorithms evolve, robust, community-curated benchmark datasets can remain an enduring resource (e.g., Deng et al. 2009). Creating them requires a rigorous approach. The community-governed data specification proposed previously can help by ensuring that datasets have consistently defined labels, a clear ontology, and harmonized geospatial metadata in a machine-readable format. Model generalization can be improved by creating leakage-proof splits for training, validation, and testing. Furthermore, bias can be reduced by using multiple annotators who follow clear guidelines and by ensuring balanced coverage across diverse coastal systems. The "CoastBench" dataset developed in this thesis highlights the limitations of a dataset labeled predominantly by a single expert. This underscores the importance of collaborative efforts, where a diverse community of scientists can collectively build, refine, and curate the training data needed to support future generations of CoastalAI models.

5.2.4 Outlook

The growing volume of Earth observation data offers an opportunity to understand the coastal system in higher detail, but only if this raw data can be converted into useful information, for which AI provides excellent tools. As this perspective has discussed, realizing the potential of CoastalAI can be enhanced by moving beyond siloed platforms toward an open, community-governed ecosystem that effectively integrates open data, cloud technology, and AI. The guiding principles discussed here—an interoperable architecture, a common data strategy, and a commitment to open training data—provide a practical framework for such an ecosystem. Adopting these principles is not merely a technical exercise, but a collaborative investment by the coastal science community to build the shared infrastructure required to better understand, model, and predict the dynamics of the coastal zone, and ultimately, to better manage the world's coasts.

6 Conclusions

This thesis set out to to develop a global framework for assessing the future impacts of coastal erosion by combining the vast observational datasets now available from satellites with the power of cloud technology and the capabilities of artificial intelligence (AI). The research conducted to achieve this objective has yielded the following conclusions, presented in direct response to the research questions posed in Chapter 1.

Regarding RQ1 (Scalable Methods) The combination of satellite data, cloud technology, and AI enables the study of local coastal phenomena, such as coastal erosion, at a global scale by bridging the gap between broad but coarse ("everywhere") and detailed but local ("anywhere") analyses. Satellite data provide global, consistent, high-resolution observations; AI provides methods to analyze large quantities of heterogeneous spatiotemporal data; and, cloud technology provides the digital infrastructure to perform these analyses at scale. This is achieved by adopting a strategy that decouples data storage from compute, centered on three principles. First, bringing code to the data through data-proximate computing is essential, as petabyte-scale satellite archives are too large to move. Second, for the big datasets produced and used, adopting cloud-optimized formats (e.g., GeoParquet) that are described in standardized metadata catalogs (e.g., Spatio-Temporal Asset Catalog (STAC)), is essential for interoperability and performance. Finally, orchestrating these workflows within a flexible, scalable open-source software ecosystem (Pangeo) provides the modularity and flexibility required to analyze complex coastal phenomena at vast spatial scales, without compromising on spatiotemporal resolution. An extra advantage for the latter is that it avoids both dependence on a single commercial provider and the hidden biases that can arise from proprietary platforms where data is often preprocessed in specific ways. By embracing a "cloud-native geospatial" approach, AI can be applied at scale to satellite data. A final important finding, however, is that one of the larger bottlenecks for leveraging modern AI is the need, hence, creation of large, representative, and ideally collaboratively-built training datasets, which is a non-trivial task that would be useful to account for in future research.

Regarding RQ2 (Coastal Geomorphology) The geomorphological composition of the world's coast was mapped at 100 m resolution using deep learning, creating a novel, globally consistent coastal typology. This analysis provides new insights into the global distribution of coastal landforms, revealing that over 60% of the world's ice-free coast consists of soft, erodible sediments. The typology revises earlier statistics on sediment distribution, estimating that 40% of the global coast is fronted by sandy, gravel, or shingle beaches, a higher figure than the 31% previously reported, with the difference largely attributed due to the new, zonally-unbiased transect system used in this thesis. The analysis also reveals geomorphological sensitivities, showing that 20% of these sandy, gravel and shingle beaches are backed by cliffs. This cross-shore configuration limits their natural capacity to migrate landward with sea-level rise where cliffs are composed of resistant lithologies; a critical detail this satellite-based approach cannot resolve yet. Methodologically, this work helps establish deep learning in coastal sci-

ence. However, the classification has limitations. While robust for sediment type (F1-score of 0.76), it is less accurate for the more complex coastal type classification (F1-score of 0.67), with some classes like "moderately sloped" remaining ambiguous due to their overlap with adjacent classes. The automated nature of the method, however, ensures that the typology can be iteratively improved as the community collaborates on standardized class definitions and more diverse, openly available training datasets.

Regarding RQ3 (Future Erosion Impacts) This research provides a global impact assessment of future coastal erosion by first quantifying present-day coastal exposure and then combining this with probabilistic future shoreline projections produced under different sea-level rise (SLR) scenarios. The global exposure analysis reveals that a quarter of the world's coast (∼253 000 km) has at least one building in the first kilometer of land, hosting approximately 76 million buildings within 1 km of the shoreline. From these patterns, the thesis introduces the "Empirical Setback Zone" (ESZ), a data-driven metric that quantifies the de facto setback coastal communities implicitly maintain. For 10% of the developed coast, the first building is located less than 37 m from the shoreline, leaving little space for natural coastal dynamics or human modifications. The analysis also shows that development is disproportionately concentrated on soft, erodible coasts; for instance, building density on sandy, gravel, and shingle coasts is 354 buildings per kilometer, significantly higher than the global average of 299. This high present-day exposure translates into significant future impacts. Through a refined application of an equilibrium-profile model, the analysis shows that while near-future impacts are mostly a legacy of historical development, sea-level rise is projected to become the dominant driver of future erosion impacts over the 21st century. By 2100, this places between 1.6 million (under a low-emissions SSP1-2.6 scenario) and 2.4 million (under a high-emissions SSP5-8.5 scenario) buildings at risk of being impacted by erosion. This figure represents a conservative lower-bound estimate, as the model's application was refined by the machine learning (ML)-based coastal typology, which constrained its use to sandy sediment plains and dune coasts (21% of the world's coast) before data limitations (primarily nearshore slope data) further reduced the study area to 9.3% (∼93 000 km). This refined approach, however, is still subject to uncertainties from the underlying ML-based coastal typology and highlights the need for better nearshore slope data as well as for multi-model approaches to assess future erosion impacts on other coastal types. Finally, this work provides a first-order impact assessment of probabilistic future shoreline change under different SLR scenarios, but stops short of a full climate risk assessment, which would require a complete analysis of vulnerability, including sensitivity and adaptive capacity.

Outlook This thesis has demonstrated how the synergy of open satellite data, cloud technology, and artificial intelligence enables the detailed study of coastal phenomena, making it possible to integrate large, heterogeneous observational datasets to assess erosion impacts at the level of individual buildings. Beyond the specific findings for coastal science, this work serves as a broader demonstration for geoscience of a promising path for analyzing large, heterogeneous datasets. Looking ahead, as the volume of Earth observation data grows, a primary challenge for coastal science will be the integration of the field's diverse data, tools, and models. As argued in the perspective of this thesis (Ch. 5.2), a promising path forward is a collective shift toward an open, community-governed ecosystem. The guiding principles discussed—an interoperable architecture, a common data strategy, and a commitment to open training data—can help realize this effort. This will ultimately enable the community to build the robust, transparent, and intelligent analytical tools required to better manage the world's coasts in an era of climate change.

Part II

Appendices

A A Cloud-Native Data Repository for Coastal Science

This chapter describes the primary scientific outputs of this dissertation: a global coastal data repository and an accompanying open-source software library for scalable coastal analytics. Together, these components form a cloud-native ecosystem designed to enable a wide range of global coastal analyses. The software library contains generic components for applications such as coastal monitoring, coastal mapping, and intertidal bathymetry estimation. The library's design integrates with the Pythonic "cloud-native geospatial" software ecosystem, including tools like DuckDB for SQL-based spatial analytics.

A.1 A Global Data Repository

The data products described in this chapter adhere to the FAIR data principles (Wilkinson et al. 2016). They are stored in cloud-optimized formats and made findable and accessible through a public Spatio-Temporal Asset Catalog (STAC) collection (https://github.com/FlorisCalkoen/stac-phd), with persistent versions archived on Zenodo for long-term access. While this section provides a summary of each dataset, the online STAC catalog serves as the definitive source for detailed, machine- and human readable metadata.

Global Coastal Transect System (GCTS)

Description	A foundational dataset of over 11 million cross-shore transects spaced at 100 m intervals along global coastlines (80°S - 84°N). Derived from a recent (Jan. 2023) OpenStreetMap coastline, an important feature is that transects are generated in their local UTM projection, ensuring uniform length and correcting for latitudinal distortions present in previous systems. GCTS provides a standardized spatial framework for monitoring coastal change, mapping coastal variables, and provides a robust basis for downstream coastal analytics.
STAC Identifier	`gcts`
Version	2024-08-02
Format	GeoParquet
Spatial Framework	Global, 100 m resolution transects
Reference	F R Calkoen et al. (Jan. 1, 2025f). "Enabling Coastal Analytics at Planetary Scale". *Environmental Modelling & Software* 183, p. 106257. DOI: 10.1016/j.envsoft.2024.106257
License	Creative Commons Attribution 4.0 (CC BY 4.0)
Access	A persistent version is archived on Zenodo (DOI: 10.5281/zenodo.14056924)

Global Coastal Grid

Description	A global coastal tiling system designed for scalable geospatial analytics in coastal areas. Derived from a recent (Jan. 2023) OpenStreetMap coastline, it provides structured grids (SlippyMapTiles) at multiple zoom levels (2-10) and buffer sizes (500m-15000m). Each tile contains precomputed intersections with countries, continents, and Sentinel-2 MGRS tiles, making it convenient for broad-scale data processing. Filenames, such as `coastal_grid_z9_10000m.parquet` correspond to the zoom level and buffer size.
STAC Identifier	`coastal-grid`
Version	2025-01-01
Format	GeoParquet
Spatial Framework	SlippyMapTiles (zoom levels 2-10)
Reference	F R Calkoen et al. (Jan. 1, 2025f). "Enabling Coastal Analytics at Planetary Scale". *Environmental Modelling & Software* 183, p. 106257. DOI: 10.1016/j.envsoft.2024.106257
License	Creative Commons Attribution 4.0 (CC BY 4.0)
Access	A persistent version is archived on Zenodo (DOI: 10.5281/zenodo.16925825)

Sentinel-2 L2A Annual Composite (2023)

Description	An annual, cloud-free composite of Sentinel-2 L2A imagery for the year 2023. Generated per MGRS tile, a median composite was computed from scenes sourced from the Microsoft Planetary Computer, using a method that ensures balanced spatial and temporal sampling while masking clouds and shadows. The final product includes six reflectance bands (Blue, Green, Red, NIR, SWIR16, and SWIR22) resampled to a uniform 10 m resolution.
STAC Identifier	`s2-l2a-composite`
Version	2025-04-04
Format	Cloud-Optimized GeoTIFF (COG)
Spatial Framework	Global 15-km coastal zone, 10 m resolution (Global Coastal Grid)
Reference	F R Calkoen et al. (July 24, 2025c). "Mapping the World's Coast: A Global 100-m Coastal Typology Derived from Satellite Data Using Deep Learning". *Earth System Science Data Discussions*, pp. 1–37. DOI: 10.5194/essd-2025-388
License	Creative Commons Attribution 4.0 (CC BY 4.0)
Access	Direct access is recommended via the CoCliCo STAC catalog using the `coastpy` library. Due to its large size (~1 TB over more than 200 thousand tif files), the dataset is not archived in a persistent Zenodo repository, but only available via the cloud (token can be requested from Deltares).

CoastBench

Description	A global training dataset of approximately 1,800 expert-labeled samples for coastal classification. The labels are connected to the GCTS grid, with each sample having labels for four attributes along the cross-shore coastal profile: sediment type, coastal type, built environment presence, and coastal defense presence. This training dataset is the basis for the Global 100 m Coastal Typology product.
STAC Identifier	-
Version	1.0
Format	GeoParquet
Spatial Framework	GCTS
Reference	F R Calkoen et al. (July 3, 2025g). *CoastBench: A Global Training Dataset for Coastal Classification Using Satellite Imagery and Elevation Data*. Version 2025-04-09. Zenodo. DOI: 10.5281/zenodo.15800285
License	Creative Commons Attribution 4.0 (CC BY 4.0)
Access	A persistent version is archived on Zenodo (DOI: 10.5281/zenodo.15800284). For analysis, direct access is recommended via the CoCliCo STAC catalog using the `coastpy` library.

Global 100 m Coastal Typology

Description	Provides global predictions for four coastal attributes along the cross-shore coastal profile: sediment type, coastal type, presence of built environments, and presence of engineered defenses. Derived using a convolutional neural network trained on the CoastBench dataset, the predictions are mapped to the GCTS framework. Each record includes the predicted class label and associated model confidence for all four attributes.
STAC Identifier	`coastal-typology`
Version	2025-04-09
Format	GeoParquet
Spatial Framework	GCTS
Reference	F R Calkoen et al. (July 24, 2025c). "Mapping the World's Coast: A Global 100-m Coastal Typology Derived from Satellite Data Using Deep Learning". *Earth System Science Data Discussions*, pp. 1–37. DOI: 10.5194/essd-2025-388
License	Creative Commons Attribution 4.0 (CC BY 4.0)
Access	A persistent version is archived on Zenodo (DOI: 10.5281/zenodo.15599096). For analysis, direct access is recommended via the CoCliCo STAC catalog using the `coastpy` library.

ShorelineMonitor: Satellite-Derived Shorelines

Description	Annual satellite-derived shorelines (SDS) from Landsat imagery (1984-2024), represented as vector line features. This dataset extends the work of Luijendijk et al. (2018) and includes attributes such as shoreline curvature and expanded metadata.
STAC Identifier	`shorelinemonitor-shorelines`
Version	2025-01-09
Format	GeoParquet
Spatial Framework	Deltares ShorelineMonitor Box System
Reference	In progress
License	Restricted
Access	Available upon reasonable request from Deltares.

ShorelineMonitor: Satellite-Derived Shoreline Series

Description	Shoreline positions from the ShorelineMonitor SDS dataset mapped onto the GCTS framework. This dataset provides a time series of shoreline locations per transect and includes flags to identify a primary shoreline-change series for trend analysis.
STAC Identifier	`shorelinemonitor-series`
Version	2025-01-09
Format	GeoParquet
Spatial Framework	GCTS
Reference	In progress
License	Restricted
Access	Available upon reasonable request from Deltares.

Global Coastal Transect Repository (GCTR)

Description	An enriched version of the GCTS designed for coastal analytics. It extends the foundational transects by adding two key attribute sets: Directional Coastline Proximity, which provides 12 distance measurements to the OSM coastline along fixed bearings (0°, 30°, ..., 330°) to characterize local geomorphology, and Coastal Exposure Metrics, which quantify the distance to, count of, and total area of nearby building footprints.
STAC Identifier	`gctr`
Version	2025-08-22
Format	GeoParquet
Spatial Framework	GCTS
Reference	F R Calkoen et al. (Aug. 22, 2025d). *The Global Coastal Transect Repository (GCTR): An Analyis-Ready, Cloud-Native Data Repository for Coastal Analytics*. Zenodo. DOI: 10.5281/zenodo.16928269
License	Creative Commons Attribution 4.0 (CC BY 4.0)
Access	A persistent version is archived on Zenodo (DOI: 10.5281/zenodo.16928268). For analysis, direct access is recommended via the CoCliCo STAC catalog using the `coastpy` library.

Future Shoreline Projections

Description	Probabilistic projections of future shoreline positions for 2030, 2050, and 2100 under three SSP scenarios (SSP1-2.6, SSP2-4.5, SSP5-8.5). Projections are derived from a model combining historical trends with SLR-induced retreat (Bruun rule) and are provided for sandy sediment plains and dune coasts where the model was applied. The dataset includes the Monte-Carlo distribution (1000 samples) as well summary statistics (median and 90% confidence interval) for the projected change at each GCTS transect.
STAC Identifier	`shoreline-projections`
Version	2025-08-22
Format	GeoParquet
Spatial Framework	GCTS
Reference	F R Calkoen et al. (Aug. 22, 2025e). *Future Shoreline Projections*. Version 2025-08-22. Zenodo. DOI: 10.5281/zenodo.16928881
License	Creative Commons Attribution 4.0 (CC BY 4.0)
Access	A persistent version is archived on Zenodo (DOI: 10.5281/zenodo.16928881). For analysis, direct access is recommended via the CoCliCo STAC catalog using the `coastpy` library.

A.2 The coastpy Analysis Library

`CoastPy` is a domain-specific Python package containing tools for scalable, cloud-based coastal analysis. The library is deeply integrated with the Pangeo and OpenDataCube software ecosystem, leveraging packages like Dask for parallel computation, Xarray for labeled n-dimensional arrays, and GeoPandas for vector data computation. It provides convenient functions to produce, manage and search STAC catalogs, retrieve data from cloud storage, and execute spatial analytics workflows, typically using DuckDB. All data described in this data repository chapter can be accessed programmatically using this library.

Software name	CoastPy
Developer	Floris Calkoen
Year first release	2024
System requirements	Python $\geq$ 3.11; macOS, Linux, Windows
Dependencies	Dask, Xarray, GeoPandas, DuckDB, PySTAC
Availability	github.com/TUDelft-CITG/coastpy; PyPI
License	MIT License
Citation	F R Calkoen et al. (Jan. 1, 2025f). "Enabling Coastal Analytics at Planetary Scale". *Environmental Modelling & Software* 183, p. 106257. DOI: 10.1016/j.envsoft.2024.106257

B Appendix to Chapter 2: Enabling Coastal Analytics at Planetary Scale

This appendix provides supplementary material supporting the findings of the "Enabling Coastal Analytics at Planetary Scale" chapter, structured to align with the supplementary information of the original publication.

B.1 Coastal Waterline Mapping: Comparisons and Performance

B.1.1 Qualitative Comparison: CoastSat vs. CoastPy

Figure B.1 presents time series data for a transect (PF6) at Narrabeen Beach, Australia, comparing field observations, CoastSat, and CoastPy. The figure demonstrates that CoastPy achieves comparable accuracy in mapping coastal waterlines to CoastSat. The minor discrepancies between CoastSat and CoastPy can be attributed to the tidal correction applied in CoastSat but not yet implemented in CoastPy. These results confirm that the methods used in this cloud-native coastal monitoring approach are effectively configured.

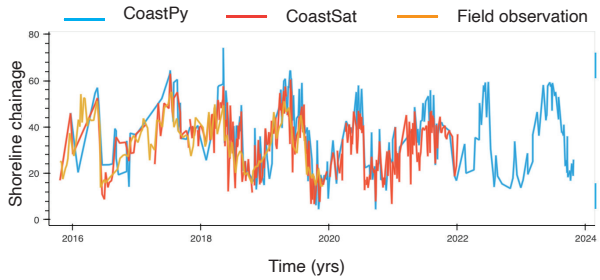

Figure B.1: Comparison of CoastPy, CoastSat, and field-observation satellite-derived shoreline (SDS)-series for transect PF6 at Narrabeen Beach, Australia.

B.1.2 Computational Performance: CoastSat

Figure B.2 shows time-series data of memory and CPU usage during the process of mapping SDS for Narrabeen Beach, Australia, using CoastSat. The mapping of all shorelines from the Sentinel-2

catalog took 3.75 hours. During this period, memory usage was saturated at approximately 50% on average, while CPU usage averaged around 15%. In total, 851 Sentinel-2 Top of Atmosphere (TOA) images were processed, resulting in 450 unique shorelines after accounting for duplicates due to overlapping tiles.

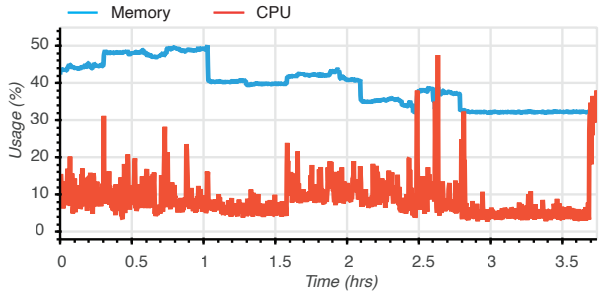

Figure B.2: CPU and memory usage over time for CoastSat while mapping SDS from the full Sentinel-2 catalog for Narrabeen Beach, Australia.

B.1.3 Computational Performance: CoastPy

Figure B.3 presents a performance report with CPU, memory, and bandwidth usage of the Dask compute cluster while mapping coastal waterlines using CoastPy for Ocean Beach, USA. The area of interest covers two quadtiles (zoom level 10), necessitating two iterations of mapping, which is evident from the repetitive pattern in the data over time. Each tile mapping follows a repetitive pattern of scheduling, initial computation with few workers, adaptive scaling, and computation with many workers. During the scheduling phase, CPU, memory, and bandwidth usage are low. As the computation begins with a small number of workers, CPU usage increases until the cluster adaptively scales to more workers to handle the larger computational workload. The adaptive scaling phase shows relatively low CPU usage as workers are configured with the necessary software and data. The compute-at-scale phase occupies approximately one-fourth of the total time, similar to each of the other phases, and is characterized by high bandwidth usage (intensive write operations), showing that the majority of coastal waterlines are mapped during this phase. This cloud-native approach to coastal waterline mapping achieves near-full CPU saturation when computing results.

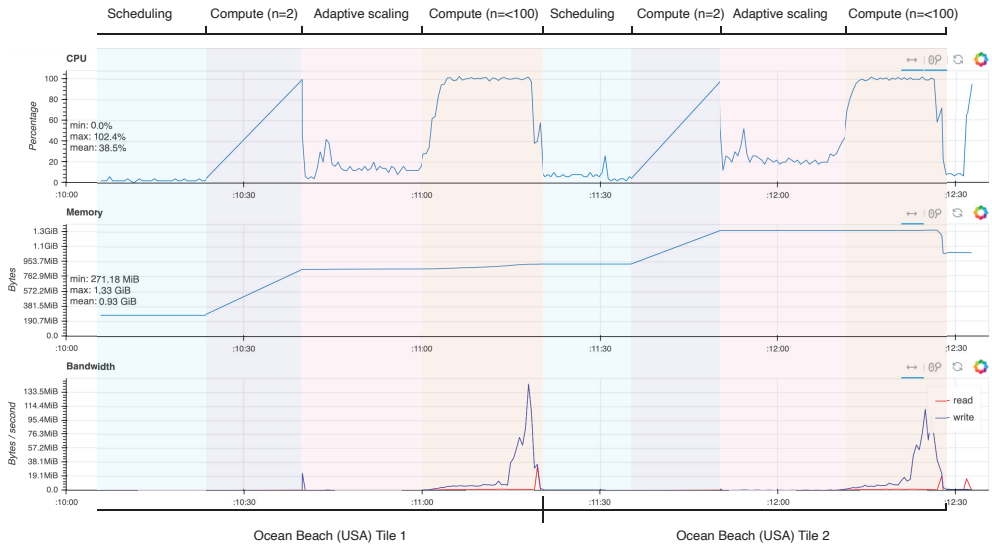

Figure B.3: Performance report of the Dask compute cluster during coastal waterline mapping using CoastPy for Ocean Beach, USA. The report details CPU, memory, and bandwidth usage over time, highlighting the repetitive patterns of scheduling, initial computation, adaptive scaling, and compute-at-scale phases for two quadtiles.

B.2 Geospatial Data Partitioning

Figure B.4 presents a global map of the global coastal transect system (GCTS), spatially partitioned into parts, each containing a maximum size of 100 MB. The data is sorted by quadkey, with entries from Alaska at the beginning and entries from Australia at the end. Bounding boxes representing the different partitions are overlaid on the map, showing that transect records are grouped by spatial area. This partitioning approach facilitates efficient geospatial data retrieval by allowing metadata to discard irrelevant partitions. The figure also shows that areas of higher coastal complexity, such as Indonesia, have relatively high data density.

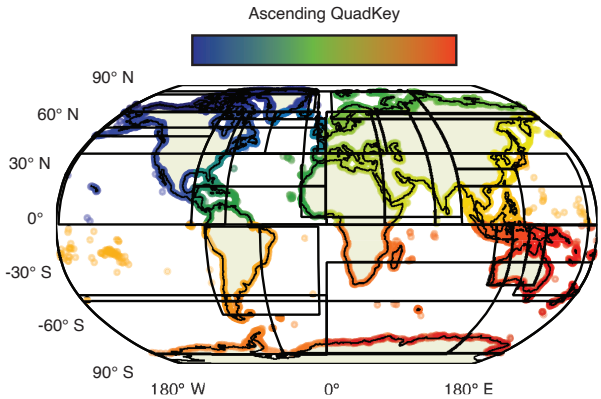

Figure B.4: Global map of the geospatially partitioned GCTS data, sorted by quadkey. The dataset is divided into partitions with a maximum size of 100 MB, enabling efficient geospatial data retrieval.

B.3 Data Retrieval Benchmarks

Figure B.5 presents the distribution of data retrieval times for various data release strategies, as detailed in Table 2.1 in the main chapter. The small variance across all experiments indicates a robust setup, ensuring consistent performance across different retrieval methods.

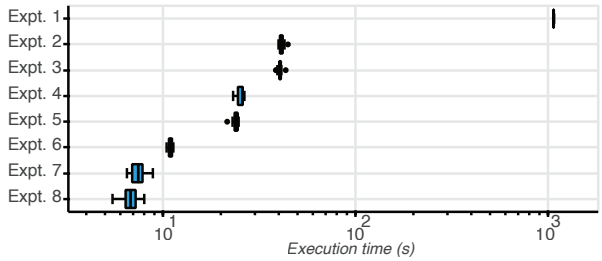

Figure B.5: Distribution of data retrieval times for different data release strategies. The specific operations applied in each experiment are detailed in Table 2.1.

C Appendix to Chapter 3: Mapping the World's Coast

This appendix provides supplementary material for the "Mapping the world's coast: a global 100 m coastal typology derived from satellite data using deep learning" chapter. It is organized into two main parts: supplementary results, including additional global, regional, and national maps; and model performance details, which provide further insight into the training data and classification confidence.

C.1 Supplementary Results

This section contains additional visualizations that complement the results presented in the main chapter, offering a more detailed view of the global coastal typology.

C.1.1 Global Maps of Binary Classifications

Figures C.1 and C.2 provide the global spatial summaries for the two binary classification tasks, complementing the primary sediment and coastal type maps shown in the main text.

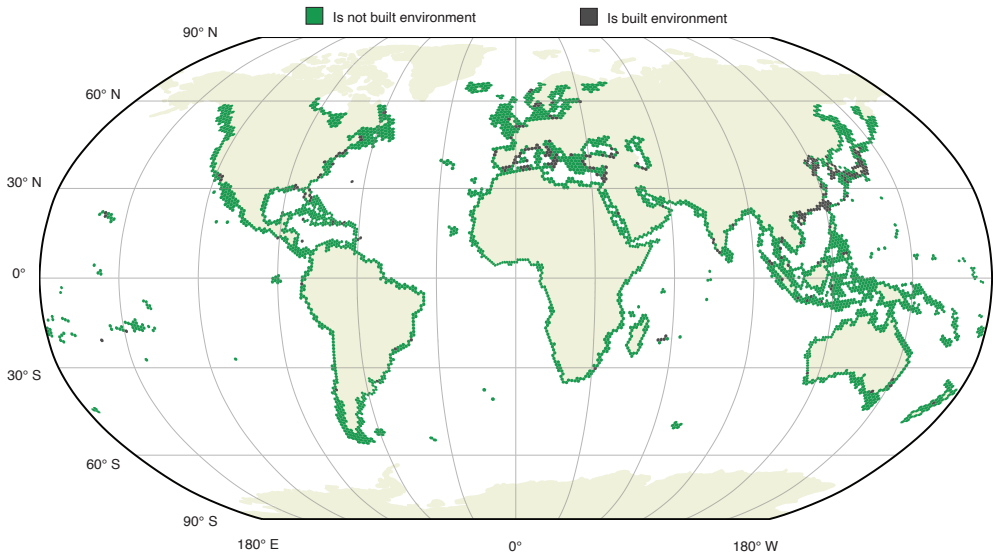

Figure C.1: Global map of predicted *is built environment*, aggregated on a level-3 H3 grid for visualization purposes using the most frequent class (mode) per cell. This coarser view summarizes almost 10 million classified coastal transects, highlighting regional patterns. Basemap: Natural Earth.

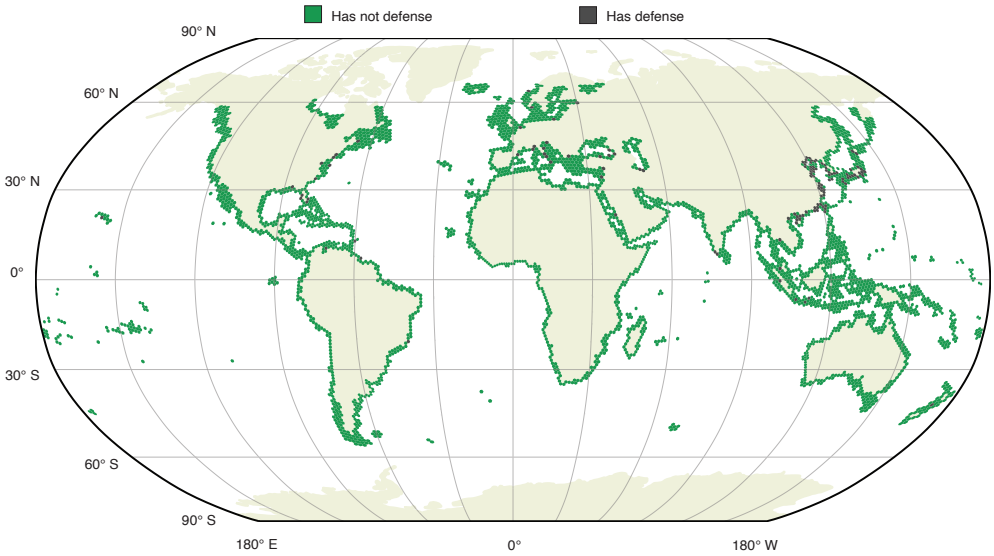

Figure C.2: Global map of predicted *has defense*, aggregated on a level-3 H3 grid for visualization purposes using the most frequent class (mode) per cell. This coarser view summarizes almost 10 million classified coastal transects, highlighting regional patterns. Basemap: Natural Earth.

C.1.2 European Coastal Typology

The following figures provide a more detailed regional view of the predicted coastal typology for continental Europe, as referenced in Section 3.1 of the main chapter.

Figure C.3: Map of continental Europe with predicted *sediment type*, aggregated on a level-5 H3 hexagonal grid for visualization purposes using majority class (mode) per cell. This continental view summarizes around 1.2 million classified coastal transects, highlighting regional patterns. Basemap: Natural Earth.

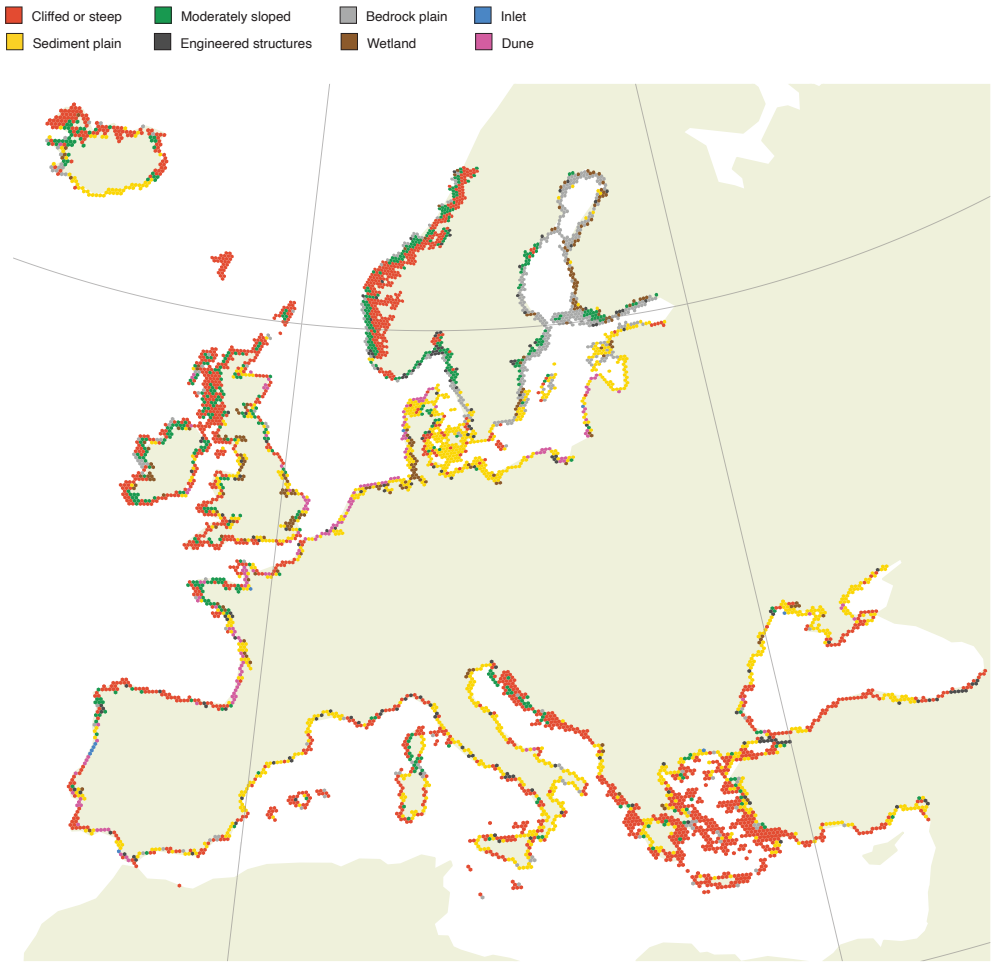

Figure C.4: Map of continental Europe with predicted *coastal type*, aggregated on a level-5 H3 hexagonal grid using majority class (mode) per cell. This continental view summarizes around 1.2 million classified coastal transects, highlighting regional patterns. Basemap: Natural Earth.

C.1.3 Coastal Typology per Country

Figure C.5 provides a quantitative breakdown of the coastal typology at the national level, supporting the summary statistics presented in Table 3.3 of the main chapter.

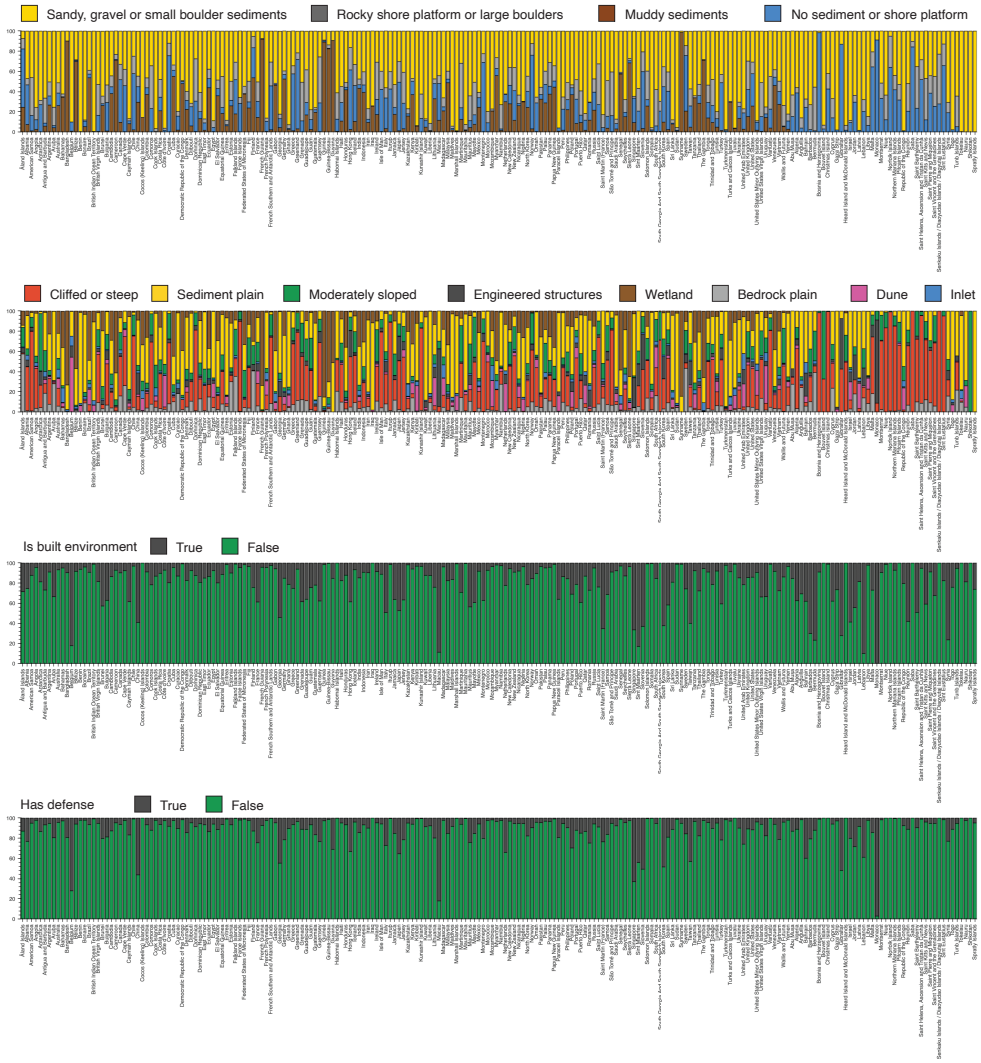

Figure C.5: Normalized class distributions per country for each of the four classification tasks: *sediment type*, *coastal type*, *is_built_environment*, and *has_defense*. For each task, stacked bars represent the percentage of transects assigned to each class within a given country. This allows for comparative analysis of coastal typology across countries.

C.2 Model Performance and Validation

This section provides supplementary information related to the model's training and validation, supporting the discussion of model performance in Section 3.5.3 of the main chapter.

C.2.1 Global Distribution of Training Samples

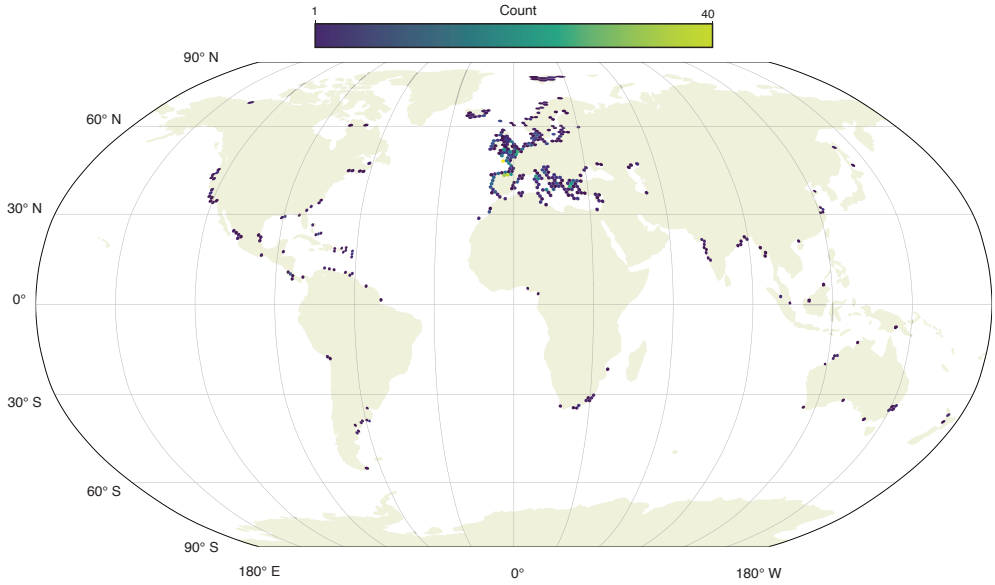

Figure C.6: Global distribution of training samples per H3 cell (zoom level 3). Basemap: Natural Earth.

C.2.2 Classification Confidence and Uncertainty

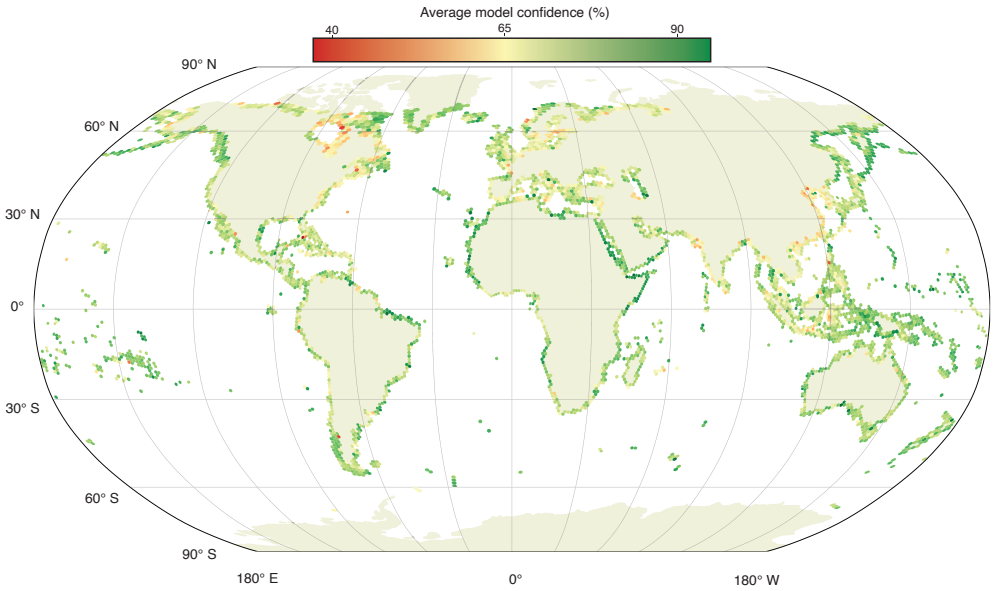

Figure C.7: Model confidence per H3 cell (zoom level 3) averaged across the four typology tasks. Confidence is computed for each transect as follows: for multi-class tasks (*shore_type*, *coastal_type*), it is defined as the maximum predicted class probability; for binary tasks (*has_defense*, *is_built_environment*), it is the distance from the classification threshold (0.5), scaled between 0 and 100. This visualization highlights regions where model predictions are more certain (green) or more ambiguous (red), providing spatial insight into model reliability. Basemap: Natural Earth.

D Appendix to Chapter 4: Global Coastal Erosion

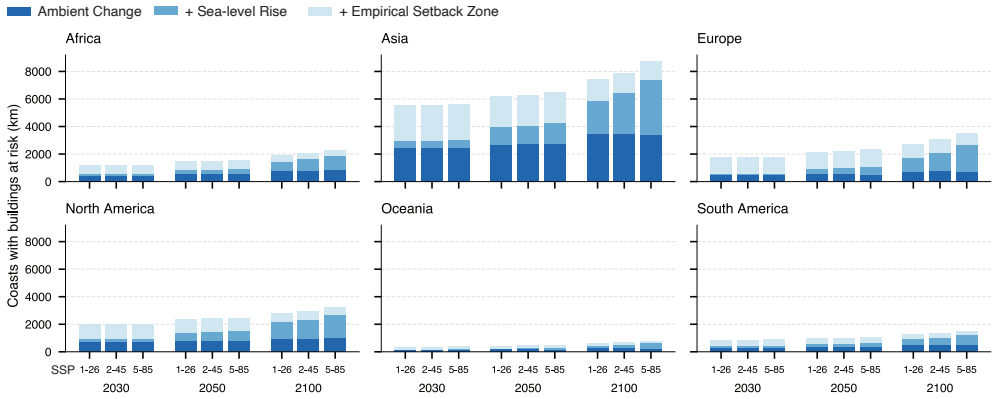

Figure D.1: Projected length of coast at risk, by continent. This figure supplements Figure 4 by disaggregating the projected risk by continent, quantified as the total length of coast (km) where buildings are at risk. Each panel shows results for 2030, 2050, and 2100 under three emissions scenarios (SSP1-2.6, SSP2-4.5, and SSP5-8.5). The stacked bars disaggregate the total risk into contributions from ambient change (AC), sea-level rise (SLR), and the empirically-derived natural buffer.

Table D.1: Shoreline-to-building distances by continent, sediment, and coastal type. This table provides a detailed breakdown of the 10th (p10) and 50th (p50) percentile distances (in meters) from the shoreline to the nearest building. The data is disaggregated by continent and major sediment type, with a further breakdown for key coastal types on sandy, gravel, or shingle coasts.

	Distance by Sediment Type (m)								**Distance on Sandy Coasts by Coastal Type (m)**							
	Sandy		Muddy		Rocky		No Sediment		Dune		Sediment Plain		Engineered		Cliffed/Steep	
Region	**p10**	**p50**	**p10**	**p50**	**p10**	**p50**	**p10**	**p50**	**p10**	**p50**	**p10**	**p50**	**p10**	**p50**	**p10**	**p50**
Global	34	140	65	502	46	220	34	238	64	221	34	128	24	101	31	135
Africa	44	161	74	570	57	207	31	196	67	213	41	140	19	74	39	187
Asia	30	144	60	442	43	233	28	180	62	288	31	145	21	109	27	123
Europe	36	142	73	568	46	227	35	276	77	282	37	119	26	88	30	137
N. America	34	115	62	557	47	188	39	231	54	139	33	101	31	105	35	115
Oceania	45	160	88	557	58	234	42	217	78	335	46	146	29	94	45	140
S. America	36	135	71	556	46	229	39	354	59	164	33	112	18	80	35	218

Table D.2: Coastal exposure metrics by continent and sediment type. This table expands on the exposure metrics in Table 1, quantifying the total length of developed coast (km), and the total count and density of buildings within the 1-km developed zone. Data is disaggregated by continent for each of the four major sediment types.

	Sandy, gravel or shingle			Muddy			Rocky			No Sediment		
	Coast	Bldgs. in 1-km Zone		Coast	Bldgs. in 1-km Zone		Coast	Bldgs. in 1-km Zone		Coast	Bldgs. in 1-km Zone	
Region	**(km)**	**Count (#)**	**Density**	**(km)**	**Count (#)**	**Density**	**(km)**	**Count (#)**	**Density**	**(km)**	**Count (#)**	**Density**
Global	119,996	42.5M	354	39,995	8.0M	201	35,790	8.3M	232	57,321	16.9M	295
Africa	10,598	4.8M	451	1,608	341k	212	1,424	665k	467	1,066	565k	530
Asia	50,141	17.5M	349	18,243	4.0M	221	10,017	2.0M	204	14,138	6.0M	425
Europe	25,323	8.8M	348	9,897	1.7M	177	12,844	3.1M	241	28,990	6.5M	225
N. America	17,902	5.1M	283	6,316	1.0M	159	7,096	1.3M	189	8,737	2.6M	299
Oceania	7,850	1.7M	223	1,734	311k	179	2,528	462k	183	1,461	342k	234
S. America	8,181	4.6M	563	2,197	615k	280	1,882	675k	359	2,930	859k	293

Table D.3: Coastal development metrics for sandy coasts by coastal type. This table provides a detailed breakdown of exposure on sandy, gravel, or shingle coasts. It quantifies the extent of developed coastline (km), and the total count and density of buildings within the 1-km developed zone, disaggregated by continent and key coastal types.

	Dune			Sediment Plain			Engineered			Cliffed or Steep		
	Coast	Bldgs. in 1-km Zone		Coast	Bldgs. in 1-km Zone		Coast	Bldgs. in 1-km Zone		Coast	Bldgs. in 1-km Zone	
Region	**(km)**	**Count (#)**	**Density**	**(km)**	**Count (#)**	**Density**	**(km)**	**Count (#)**	**Density**	**(km)**	**Count (#)**	**Density**
Global	7,612	2.2M	288	63,290	27.0M	427	4,708	3.2M	674	15,585	3.0M	191
Africa	1,705	476k	279	5,864	3.2M	545	106	123k	1163	1,011	317k	314
Asia	1,515	313k	207	29,276	12.1M	412	2,696	1.7M	624	5,881	890k	151
Europe	1,502	556k	370	10,679	4.6M	430	889	732k	823	4,524	961k	212
N. America	1,006	194k	193	9,968	3.2M	326	699	355k	508	1,897	381k	201
Oceania	652	96k	147	3,314	895k	270	178	116k	652	1,249	191k	153
S. America	1,232	560k	454	4,190	3.0M	718	139	164k	1178	1,024	233k	228

Table D.4: Projected shoreline change disaggregated by component and continent. This table supplements Table 2 by disaggregating the projected shoreline change (in meters) into its ambient change (AC) and sea-level rise (SLR) components for each continent. Values represent the median change with the 5th–95th percentile uncertainty range in parentheses. The Total is derived from the full distribution of summed samples and is not the direct sum of the component medians. Region codes: Global (GL), Africa (AF), Asia (AS), Europe (EU), North America (NA), Oceania (OC), South America (SA).

Year	SSP	Component	GL	AF	AS	EU	NA	OC	SA
2030	1-2.6	Ambient Change	0 (-1, 1)	0 (-2, 1)	0 (-1, 2)	0 (-1, 2)	0 (-1, 1)	0 (-1, 1)	0 (-2, 2)
		Sea-Level Rise	-17 (-33, -3)	-15 (-27, -5)	-17 (-37, 0)	-16 (-31, -1)	-19 (-33, -6)	-21 (-40, -5)	-13 (-24, -4)
		Total Change	**-17 (-35, -2)**	**-16 (-29, -5)**	**-17 (-39, 1)**	**-15 (-32, 0)**	**-18 (-34, -6)**	**-23 (-44, -5)**	**-14 (-27, -3)**
	2-4.5	Ambient Change	0 (-1, 1)	0 (-2, 1)	0 (-1, 2)	0 (-1, 2)	0 (-1, 1)	0 (-1, 1)	0 (-2, 2)
		Sea-Level Rise	-17 (-33, -3)	-16 (-28, -5)	-17 (-36, 0)	-16 (-32, -1)	-18 (-32, -6)	-21 (-39, -5)	-14 (-24, -4)
		Total Change	**-17 (-34, -2)**	**-16 (-30, -5)**	**-17 (-37, 1)**	**-15 (-33, 0)**	**-18 (-33, -6)**	**-23 (-43, -5)**	**-14 (-27, -4)**
	5-8.5	Ambient Change	0 (-1, 1)	0 (-2, 1)	0 (-1, 2)	0 (-1, 2)	0 (-1, 1)	0 (-1, 1)	0 (-2, 2)
		Sea-Level Rise	-18 (-34, -4)	-16 (-28, -6)	-18 (-38, 0)	-16 (-32, -2)	-19 (-33, -7)	-23 (-40, -9)	-14 (-25, -5)
		Total Change	**-18 (-35, -3)**	**-17 (-30, -6)**	**-18 (-40, 0)**	**-15 (-32, -1)**	**-19 (-34, -6)**	**-24 (-43, -10)**	**-15 (-28, -4)**
2050	1-2.6	Ambient Change	0 (-4, 5)	0 (-6, 5)	0 (-4, 6)	1 (-2, 6)	0 (-3, 5)	0 (-5, 4)	1 (-6, 9)
		Sea-Level Rise	-34 (-63, -8)	-31 (-53, -12)	-35 (-72, 0)	-30 (-57, -6)	-37 (-64, -15)	-44 (-76, -21)	-28 (-49, -11)
		Total Change	**-35 (-69, -7)**	**-32 (-61, -11)**	**-35 (-77, 0)**	**-28 (-59, -3)**	**-36 (-67, -13)**	**-49 (-88, -21)**	**-29 (-58, -7)**
	2-4.5	Ambient Change	0 (-4, 5)	0 (-6, 5)	0 (-4, 6)	1 (-2, 6)	0 (-3, 5)	0 (-5, 4)	1 (-6, 9)
		Sea-Level Rise	-36 (-65, -10)	-33 (-56, -14)	-37 (-73, 0)	-32 (-59, -8)	-39 (-65, -18)	-48 (-81, -24)	-30 (-52, -13)
		Total Change	**-37 (-71, -9)**	**-35 (-63, -13)**	**-37 (-77, -2)**	**-30 (-61, -6)**	**-38 (-68, -16)**	**-52 (-92, -25)**	**-32 (-61, -10)**
	5-8.5	Ambient Change	0 (-4, 5)	0 (-6, 5)	0 (-4, 6)	1 (-2, 6)	0 (-3, 5)	0 (-5, 4)	1 (-6, 9)
		Sea-Level Rise	-41 (-72, -13)	-37 (-62, -18)	-43 (-82, -5)	-35 (-64, -11)	-42 (-72, -21)	-55 (-89, -31)	-35 (-58, -17)
		Total Change	**-42 (-78, -12)**	**-39 (-69, -17)**	**-43 (-87, -6)**	**-33 (-66, -8)**	**-42 (-75, -18)**	**-59 (-100, -31)**	**-37 (-67, -14)**
2100	1-2.6	Ambient Change	1 (-12, 15)	-1 (-18, 14)	1 (-12, 17)	3 (-7, 17)	1 (-8, 13)	0 (-14, 10)	3 (-16, 26)
		Sea-Level Rise	-77 (-150, -17)	-69 (-127, -24)	-82 (-173, 0)	-63 (-124, -13)	-81 (-144, -34)	-108 (-196, -45)	-65 (-122, -22)
		Total Change	**-79 (-166, -12)**	**-74 (-147, -20)**	**-81 (-186, 2)**	**-57 (-129, -4)**	**-79 (-153, -28)**	**-121 (-228, -48)**	**-68 (-144, -12)**
	2-4.5	Ambient Change	1 (-12, 15)	-1 (-18, 14)	1 (-12, 17)	3 (-7, 17)	1 (-8, 13)	0 (-14, 10)	3 (-16, 26)
		Sea-Level Rise	-97 (-176, -32)	-88 (-153, -41)	-103 (-197, -20)	-80 (-148, -27)	-99 (-175, -49)	-135 (-233, -71)	-85 (-149, -40)
		Total Change	**-98 (-192, -30)**	**-93 (-173, -37)**	**-101 (-209, -18)**	**-75 (-153, -17)**	**-96 (-183, -42)**	**-146 (-267, -73)**	**-89 (-173, -31)**
	5-8.5	Ambient Change	1 (-12, 15)	-1 (-18, 14)	1 (-12, 17)	3 (-7, 17)	1 (-8, 13)	0 (-14, 10)	3 (-16, 26)
		Sea-Level Rise	-134 (-233, -64)	-121 (-200, -69)	-145 (-260, -56)	-110 (-194, -52)	-129 (-234, -73)	-186 (-309, -110)	-118 (-195, -68)
		Total Change	**-135 (-250, -60)**	**-126 (-220, -65)**	**-143 (-274, -52)**	**-104 (-199, -42)**	**-126 (-241, -65)**	**-197 (-344, -111)**	**-124 (-220, -60)**

Table D.5: Projected shoreline change disaggregated by coastal type and continent. This table provides a further disaggregation of the total projected shoreline change (in meters) for sandy coasts, categorized by geomorphic type (Dune or Sediment Plain) and continent. Values represent the median change with the 5th–95th percentile uncertainty range in parentheses. Negative values indicate shoreline retreat. Region codes: Global (GL), Africa (AF), Asia (AS), Europe (EU), North America (NA), Oceania (OC), South America (SA).

Year	SSP	Coastal Type	GL	AF	AS	EU	NA	OC	SA
2030	1-2.6	Dune	-16 (-31, -3)	-12 (-22, -3)	-19 (-45, 1)	-17 (-36, -1)	-24 (-42, -9)	-17 (-33, -4)	-12 (-21, -4)
		Sediment Plain	-17 (-35, -2)	-17 (-32, -5)	-17 (-38, 1)	-15 (-31, 0)	-18 (-33, -6)	-25 (-48, -5)	-15 (-29, -3)
	2-4.5	Dune	-16 (-31, -3)	-12 (-22, -3)	-19 (-45, 1)	-18 (-37, -1)	-23 (-41, -8)	-17 (-33, -4)	-12 (-21, -4)
		Sediment Plain	-17 (-35, -2)	-18 (-32, -6)	-17 (-37, 1)	-15 (-32, 0)	-18 (-32, -6)	-25 (-48, -6)	-15 (-29, -3)
	5-8.5	Dune	-17 (-32, -4)	-12 (-23, -3)	-20 (-46, 1)	-18 (-37, -2)	-24 (-41, -9)	-18 (-33, -7)	-12 (-21, -4)
		Sediment Plain	-18 (-36, -3)	-18 (-33, -6)	-18 (-39, 0)	-15 (-31, -1)	-18 (-33, -6)	-27 (-48, -11)	-16 (-30, -4)
2050	1-2.6	Dune	-31 (-63, -8)	-23 (-45, -6)	-39 (-89, 0)	-31 (-67, -2)	-44 (-80, -17)	-36 (-68, -15)	-25 (-46, -8)
		Sediment Plain	-35 (-70, -6)	-36 (-65, -12)	-35 (-76, 0)	-28 (-57, -3)	-35 (-65, -13)	-54 (-98, -24)	-31 (-62, -7)
	2-4.5	Dune	-34 (-65, -10)	-25 (-47, -7)	-41 (-92, -2)	-34 (-69, -6)	-46 (-81, -20)	-39 (-71, -18)	-27 (-48, -10)
		Sediment Plain	-37 (-72, -9)	-38 (-68, -15)	-36 (-76, -2)	-30 (-59, -6)	-37 (-67, -15)	-58 (-103, -28)	-34 (-65, -10)
	5-8.5	Dune	-38 (-71, -13)	-28 (-51, -10)	-48 (-100, -7)	-37 (-75, -8)	-51 (-88, -23)	-45 (-78, -23)	-31 (-53, -13)
		Sediment Plain	-42 (-79, -12)	-43 (-74, -19)	-43 (-86, -6)	-33 (-64, -8)	-41 (-73, -18)	-66 (-112, -35)	-39 (-72, -14)
2100	1-2.6	Dune	-71 (-151, -14)	-52 (-111, -9)	-91 (-213, 0)	-61 (-147, 3)	-91 (-176, -35)	-90 (-177, -33)	-58 (-116, -14)
		Sediment Plain	-80 (-168, -12)	-81 (-158, -24)	-80 (-183, 2)	-57 (-126, -4)	-77 (-150, -27)	-135 (-255, -54)	-72 (-154, -11)
	2-4.5	Dune	-90 (-178, -30)	-67 (-130, -21)	-115 (-244, -21)	-82 (-174, -17)	-112 (-209, -51)	-111 (-206, -52)	-74 (-137, -29)
		Sediment Plain	-100 (-195, -30)	-101 (-186, -42)	-99 (-206, -18)	-74 (-149, -17)	-94 (-180, -41)	-164 (-298, -82)	-95 (-185, -31)
	5-8.5	Dune	-124 (-229, -57)	-92 (-165, -42)	-164 (-313, -56)	-117 (-230, -42)	-149 (-270, -78)	-150 (-263, -82)	-101 (-174, -52)
		Sediment Plain	-137 (-254, -61)	-138 (-237, -73)	-141 (-271, -51)	-103 (-194, -42)	-123 (-237, -64)	-221 (-386, -124)	-133 (-236, -63)

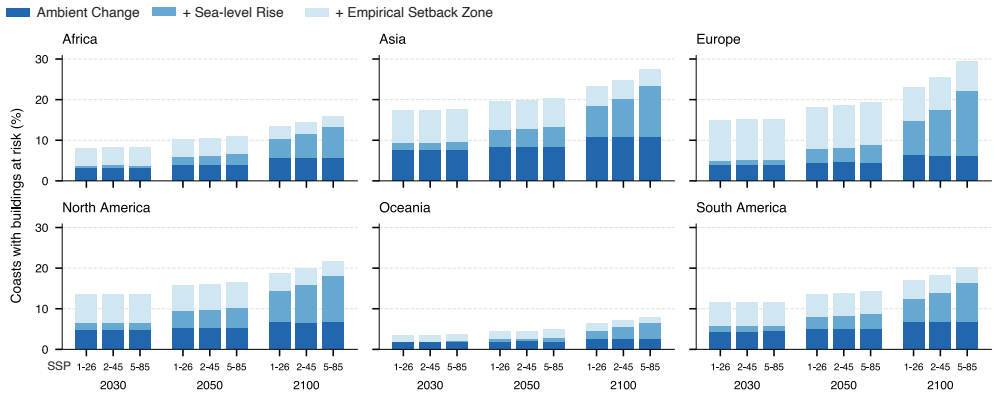

Figure D.2: Projected percentage of coast at risk, by continent. This figure supplements Figure 4 by disaggregating the projected risk by continent, quantified as the percentage of the analyzed coastline where buildings are at risk. Each panel shows results for 2030, 2050, and 2100 under three emissions scenarios (SSP1-2.6, SSP2-4.5, and SSP5-8.5). The stacked bars disaggregate the total risk into contributions from ambient change (AC), sea-level rise (SLR), and the empirically-derived natural buffer.

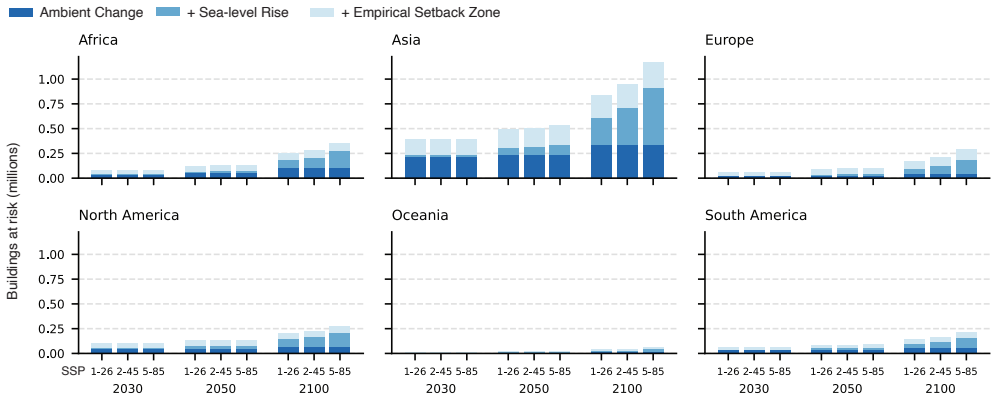

Figure D.3: Projected number of buildings at risk, by continent. This figure supplements Figure 4 by disaggregating the projected risk by continent, quantified as the total number of buildings at risk (in millions). Each panel shows results for 2030, 2050, and 2100 under three emissions scenarios (SSP1-2.6, SSP2-4.5, and SSP5-8.5). The stacked bars disaggregate the total risk into contributions from ambient change (AC), sea-level rise (SLR), and the empirically-derived natural buffer.

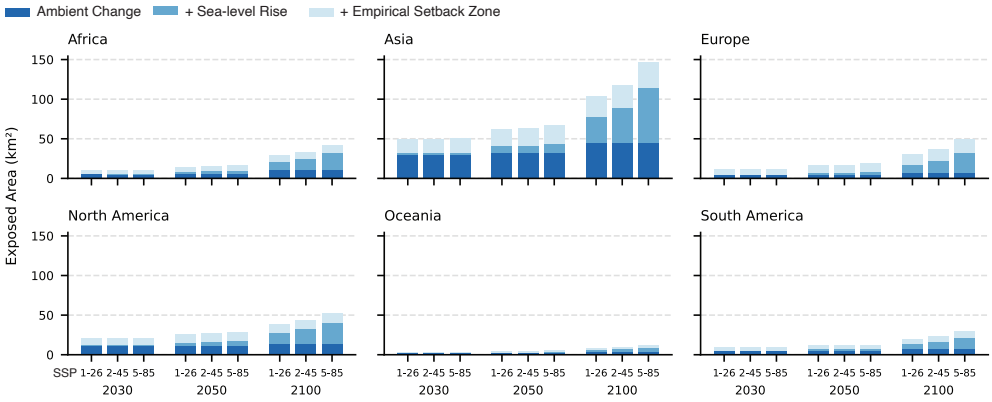

Figure D.4: Projected building area at risk, by continent. This figure supplements Figure 4 by disaggregating the projected risk by continent, quantified as the total building area at risk (in million m^2). Each panel shows results for 2030, 2050, and 2100 under three emissions scenarios (SSP1-2.6, SSP2-4.5, and SSP5-8.5). The stacked bars disaggregate the total risk into contributions from ambient change (AC), sea-level rise (SLR), and the empirically-derived natural buffer.

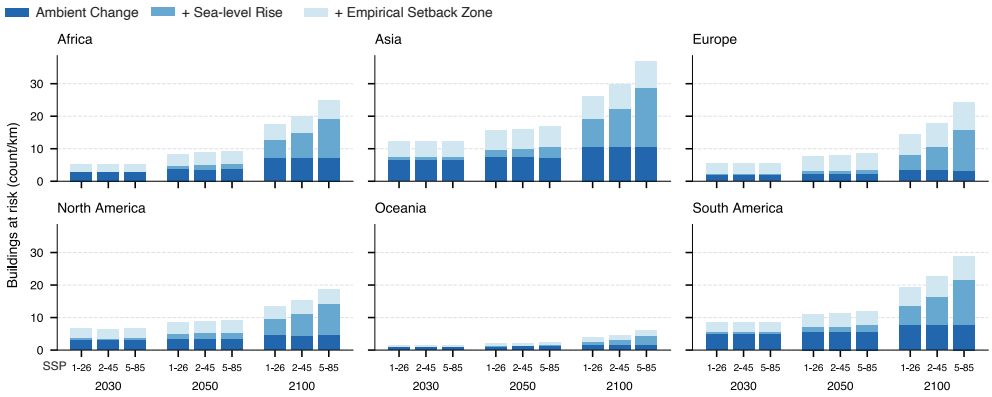

Figure D.5: Projected density of buildings at risk, by continent. This figure supplements Figure 4 by disaggregating the projected risk by continent, quantified as the density of at-risk buildings (count per km of analyzed coastline). Each panel shows results for 2030, 2050, and 2100 under three emissions scenarios (SSP1-2.6, SSP2-4.5, and SSP5-8.5). The stacked bars disaggregate the total risk density into contributions from ambient change (AC), sea-level rise (SLR), and the empirically-derived natural buffer.

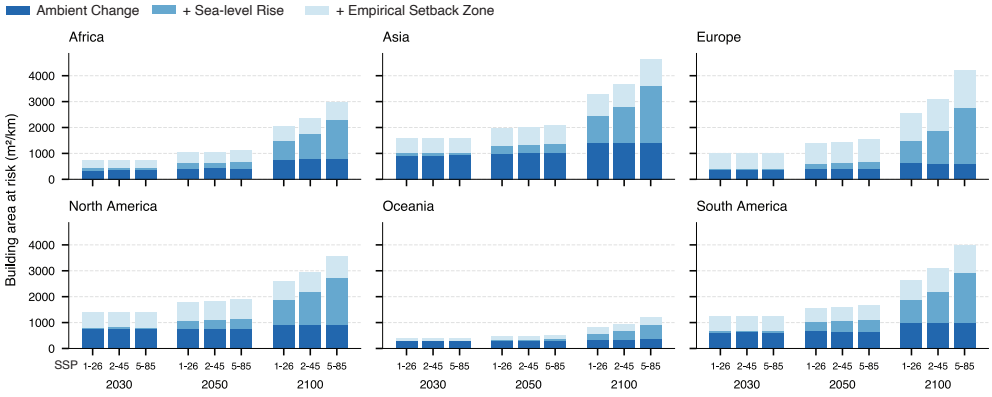

Figure D.6: Projected density of building area at risk, by continent. This figure supplements Figure 4 by disaggregating the projected risk by continent, quantified as the density of at-risk building area (m^2 per km of analyzed coastline). Each panel shows results for 2030, 2050, and 2100 under three emissions scenarios (SSP1-2.6, SSP2-4.5, and SSP5-8.5). The stacked bars disaggregate the total risk density into contributions from ambient change (AC), sea-level rise (SLR), and the empirically-derived natural buffer.

Acknowledgements

Studying beaches is a wonderful thing. Arguably, it is the nicest research topic one can have, and it is probably something I would be happy to do for many more years. Yet, writing this now, after four years of PhD, I cannot wait to go on holidays to the mountains. That doesn't have anything to do with beaches I guess, but rather with the nature of a PhD itself, and a strong overall feeling that it's time to move on. Maybe it was more difficult than I initially expected, and surely I found it challenging to find a good work-life balance; even on holidays, I would still be thinking about coasts, although maybe I should not have taken beach holidays. I also found it stressful at times: before starting this job, I don't think I had ever missed a deadline; in the past four years, I am not sure I made any. Nevertheless, it has been a transformational experience that I do not regret: I will not do another PhD, but I would definitely do this one again. And therefore I am very grateful to have been given this opportunity, for which I would like to thank some people in particular.

First, I am very grateful for all the financial, intellectual, and supervisory support I have had during the past four years from Deltares, Delft University of Technology, and the European project, CoCliCo, that funded my PhD. Some other things may seem small, but made the difference for me. I was allowed to buy the computer that I thought would work best for me (a very nice Mac), while the rest of the organization uses Windows (poor them). Others were transformational. I could travel to conferences, not only in The Netherlands but also to other countries in Europe and even to Australia and the USA. These journeys were very special, and I carry the experiences with me every day. I could plan my own days (very nice, but equally dangerous) and develop my own thoughts for four years, with nothing but support from colleagues, supervisors, and friends—and I realize this is an exceptional privilege.

A PhD is supposed to demonstrate that you can conduct independent research, and that autonomy is one of the things I truly love about it. While I was probably convinced of my independence when I started, looking back I can now see how important supervision is for becoming an independent researcher. Therefore, I can only start by thanking my supervisory team, Arjen, Rosh, Stefan, and Fedor, for their invaluable guidance over the past four years. Each of you made your own meaningful contributions, of which I will here just name a few.

Arjen, thank you for always guiding me towards the main research objective, ensuring the shoreline projections remained in focus. This was a broad research project with a European, but preferably global scope, involving many detailed steps—several a potential research topic in itself—and it was easy to get lost in the details. At those exact moments, you always provided clear, pragmatic, and strategic recommendations on how and why to move forward. At the same time, I appreciate the trust you showed by letting me find my own way in the final year, and more generally the deep and trusting working relationship we have built over the years. We still have a lot of results to wrap up, and beyond that, I look forward to collaborating with you again in the future.

Rosh, thank you for your thorough reviews, with increasingly detailed comments, and the open discussions we had. Over the years, I have come to greatly appreciate your skill as a sharp reader. Especially in the final stage, when things were coming together, your input was invaluable: it made the thesis not only more scientifically sound but also helped me find a more modest tone, which

will probably make the work much better received. I look forward to the day our paths cross again; perhaps on the topic of (coastal) climate risk analysis.

Stefan, although your role was perhaps smaller than initially thought as you became Dean of the faculty, you were always there when it mattered. It was an honour (and maybe an even greater pleasure) to represent you for the CIRN meeting at the FRF in Duck. I also appreciate that one personal morning call about the thesis, which shows that you are there, in a nice supportive way, when it counts. Your optimism, work ethos, leadership, and ability to bring people together are an inspiration, and I am grateful to have seen this firsthand for a few years.

Fedor, first, thank you for being the technical helpdesk. One can ask you anything about computers, statistics or, yes, whatever, preferably by chat without formalities, and you will reply within seconds—usually with an extra, unsolicited ironic remark. I also admire your open mindset and the way you are able to put things into the bigger picture, which was very welcome while we were writing articles together. I am also thankful for the emerging technologies you introduced me to at the start of my PhD; while it was hard to predict back then whether this was a path worth exploring (at least for me), I am now incredibly happy with where it led.

Then I would also like to thank a few people in particular. Robert Nicholls, firstly for the pleasant collaboration within CoCliCo, but perhaps more importantly, for ensuring the shoreline projections did not become an official project deliverable—that would have substantially increased my stress levels.

Kilian Vos, Stefano Conti and Rai Ibaceta for being glorious hosts while I was visiting the Water Research Laboratory in Sydney. Thanks to you, a day at the office has never been better. I remember surfing during lunch breaks and coming to the office without shoes. And believe me Europeans, this is not as weird as it sounds; it is a genuinely appreciated part of "Australia's beach culture," where one is encouraged to walk around barefoot, also in the office.

The year after, travelling to California, Sean Vitousek offered me a place to stay with him and his family, which was not only a pleasant and exceptional experience, but also an inspiration in how to host others travelling from afar. I enjoyed living for a couple of weeks with Vinny, Merigold, Sylvia, and Sean in the Californian redwoods, listening to Vinny's musical aspirations over dinner. I will also never forget the day Sean decided to join me surfing at Natural Bridges (Santa Cruz). Sean may be in the ocean almost every day, but it is rare for him to be on a surfboard without a foil. Yet, on that day, when he saw Natural Bridges was doing its thing, he joined, and it was only a short while before the tall Hawaiian shoreline modeller was getting covered up under a barrel.

My coastal friends in Delft (and Amsterdam). Elias, I think we are already beyond our ten-year anniversary—wasn't it 2015 when we did fieldwork in Fuente Álamo? Since then, our paths have followed a surprisingly similar direction, and I hope they continue to do so now that we are both wrapping up our PhDs. Jakob, we have been in the same boat and that creates a special bond. Beyond that, I am thankful for the friendship we have built and look forward to surfing more often with you soon. And finally a big thanks to my TU Delft roomies (Su, Mario, Paul, Gijs, Marthe, Jakob, Kevin, Bart, Ted and Vincent): while my days there may not have been the most productive, they were totally worth it for the coffee breaks, inspiration, and maybe even for Bart and Mario's music hour.

To my colleagues and students at Deltares: Antonio Moreno-Rodenas for your warm support throughout the PhD, the collaborations, and the numerous slightly-geeky scientific jokes. Etiënne Kras, thank you for the several pleasant collaborations and the fun we had while traveling abroad. I am also grateful to Helena van der Vegt, my department head, who at every personal development meeting focused on offering a helping hand, rather than discussing my personal growth in the organization. Finally, I was incredibly lucky with the students I have worked with. Dani, Prayla, Hugo,

and Juul, it was a pleasure to work with each of you, an experience from which I learned a lot, and I hope our future paths cross again.

Finally, I would like to thank my friends, family, and loved ones for their support, love, and the moments of recreation we shared. To my parents, Heleen and Jan Aarnoud for all the talks, support, encouraging words and advice. And also to my brother Diederick, for finishing our renovation project while I started this job. To my grandparents Piet, Ria, and Liesbeth, who have encouraged scientific curiosity and have always been following closely my studies. As well as to many other family members, for supporting me throughout these years and staying involved with the research. To Nienke and Caroline, who have both been there like no one else—you will always have a special place in my heart. To Marguerite, for her last-minute suggestion to slightly adjust the title, as for the numerous delicious Sunday dinners. To Flavia, for all the coffee breaks and laughs on Science Park; and of course for the fantastic dinners we have made along the way. To Dido, my forever geologist companion. And then obviously also to the "de Bronnenkamer" crew, with whom I spent all those hours in the PC Hoofthuis library (RIP Vrijdag 13 Juni 2025). Well, then this is also the moment to thank the librarians for just letting us be there, year in year out, while we saw generations of students come and go. To Theo and Mathias, for support from France and the Basque Country: for encouraging a PhD in the first place; for showing genuine interest in the topic over the years; and, for all nice waves, pintxos and music we have shared since 2018—"Gorra Gorra, ya nos vemos". And finally to Heartbreak Hotel (Dennis, Valentijn, Luuk, Sterre, Saar & Taras), for all those years of living together. It's open, warm, quirky, unpredictable, sometimes wild, but always close. And although we rather do our own thing, we always end up together. It's happening, it's a vibe and I wouldn't like to change it for anything else.

Amsterdam, October 8, 2025 // Barranco, March 10, 2026

Curriculum Vitae

Full Name Floris Reinier Calkoen
Born August 5, 1992 in Amsterdam, The Netherlands
Location Amsterdam, The Netherlands
Email floris@calkoen.nl
Website https://floris.calkoen.nl

Education

PhD in Coastal Science **2021–2025**
Delft University of Technology & Deltares

MSc in Data Science **2019–2020**
University of Amsterdam

MSc in Marine Environment and Resources (Erasmus Mundus) **2017–2019**
Universities of Bordeaux, Basque Country, Southampton, and Lisbon

BA in Philosophy **2015–2017**
Vrije Universiteit, Amsterdam

BSc in Earth Science (Geology) **2012–2016**
Vrije Universiteit, Amsterdam

VWO, Natuur & Techniek **2004–2011**
Hervormd Lyceum Zuid, Amsterdam

Research & Professional Experience

PhD Researcher **2021–2025**
Deltares; Delft University of Technology, Delft

- Global coastal analytics on erosion impacts using EO, AI, and cloud technology.
- Developed the cloud data infrastructure for the European H2020 CoCliCo project.
- Produced and maintained several open, global coastal datasets.

Research Intern **2019**
University of Lisbon, Lisbon

- Analyzed multi-decadal coastal change in Portugal using satellite imagery, with results presented in an interactive web visualization.

Selected Peer-Reviewed Publications

1. Calkoen, F. R., Luijendijk, A. P., Hanson, S., Nicholls, R. J., Moreno-Rodenas, A., De Heer, H., & Baart, F. (2025). Mapping the world's coast: a global 100-m coastal typology derived from satellite data using deep learning. *Earth System Science Data Discussions* [Preprint, in review]. https://doi.org/10.5194/essd-2025-388
2. Calkoen, F. R., Luijendijk, A. P., Vos, K., Kras, E., & Baart, F. (2025). Enabling coastal analytics at planetary scale. *Environmental Modelling & Software, 183*, 106257. https://doi.org/10.1016/j.envsoft.2024.106257
3. Calkoen, F. R., Luijendijk, A. P., Rivero, C. R., Kras, E., & Baart, F. (2021). Traditional vs. Machine-Learning Methods for Forecasting Sandy Shoreline Evolution Using Historic Satellite-Derived Shorelines. *Remote Sensing, 13*(5), 934. https://doi.org/10.3390/rs13050934

Selected Awards & Grants

- TU Delft Education Innovation Projects Grant 2024
- TU Delft Open Education Stimulation Fund Grant 2023
- Prof. Collins Award for Best EMMER MSc Thesis Presentation 2019
- Erasmus+ Traineeship Grant 2019
- MER Consortium Scholarship 2018 & 2019
- VSBfonds Scholarship 2017

List of Outputs

This chapter provides a comprehensive list of the scholarly outputs generated during the course of this PhD research.

1 Peer-Reviewed Journal Publications

First Author

1. Calkoen, F. R., Luijendijk, A. P., Hanson, S., Nicholls, R. J., Moreno-Rodenas, A., De Heer, H., & Baart, F. (2025). Mapping the world's coast: a global 100-m coastal typology derived from satellite data using deep learning. *Earth System Science Data Discussions*, in review. https://doi.org/10.5194/essd-2025-388. [Chapter 3]

2. Calkoen, F. R., Luijendijk, A. P., Vos, K., Kras, E., & Baart, F. (2025). Enabling coastal analytics at planetary scale. *Environmental Modelling & Software*, 183, 106257. https://doi.org/10.1016/j.envsoft.2024.106257. [Chapter 2]

3. Calkoen, F. R., Luijendijk, A. P., Rivero, C. R., Kras, E., & Baart, F. (2021). Traditional vs. Machine-Learning Methods for Forecasting Sandy Shoreline Evolution Using Historic Satellite-Derived Shorelines. *Remote Sensing*, 13(5), 934. https://doi.org/10.3390/rs13050934.[1]

Co-author

1. de Heer, H. J., Calkoen, F. R., & Moreno-Rodenas, A. Geospatial classification through fine-tuned large language and vision models. *Manuscript in preparation, 2025.*

2. Hanson, S. E., Nicholls, R. J., Calkoen, F. R., Le Cozannet, G., & Luijendijk, A. P. (2025). A geospatial database of coastal characteristics for erosion assessment of Europe's coastal floodplains. *EGUsphere*, in review. https://doi.org/10.5194/egusphere-2025-2371.

3. Thiéblemont, R., le Cozannet, G., Rohmer, J., Privat, A., Guidez, R., Negulescu, C., Philippenko, X., Luijendijk, A. P., Calkoen, F. R., & Nicholls, R. J. (2024). Sea-level rise induced change in exposure of low-lying coastal land: implications for coastal conservation strategies. *Anthropocene Coasts*, 7(1), 8. https://doi.org/10.1007/s44218-024-00041-1.

4. Hulskamp, R., Luijendijk, A. P., van Maren, B., Moreno-Rodenas, A., Calkoen, F. R., Kras, E., Lhermitte, S., & Aarninkhof, S. (2023). Global distribution and dynamics of muddy coasts. *Nature Communications*, 14(1), 8259. https://doi.org/10.1038/s41467-023-43819-6.

[1] This work was conducted prior to the PhD research but is included due to its direct relevance to the dissertation's topic.

5. Vos, K., Splinter, K. D., Palomar-Vázquez, J., Pardo-Pascual, J. E., Almonacid-Caballer, J., Cabezas-Rabadán, C., Kras, E., Luijendijk, A. P., Calkoen, F. R., Almeida, L. P., Pais, D., Klein, A. H. F., Mao, Y., Harris, D., Castelle, B., Buscombe, D., & Vitousek, S. (2023). Benchmarking satellite-derived shoreline mapping algorithms. *Communications Earth & Environment*, 4(1), 1-17. https://doi.org/10.1038/s43247-023-01001-2.

2 Datasets

1. Calkoen, F. R., Luijendijk, A. P., Barli, P., Hemmes, J., Ranasinghe, R., & Nicholls, R. (2025). *Future Shoreline Projections*. Zenodo. https://doi.org/10.5281/zenodo.16928881.
2. Calkoen, F. R., Luijendijk, A.P. (2025). *The Global Coastal Transect Repository (GCTR): An Analyis-Ready, Cloud-Native Data Repository for Coastal Analytics*. Zenodo. https://doi.org/10.5281/zenodo.16928269.
3. Calkoen, F. R., Luijendijk, A. P., Hanson, S., Nicholls, R., Moreno-Rodenas, A., De Heer, H., & Baart, F. (2025). *Mapping the World's Coast: A Global 100 m Coastal Typology Derived from Satellite Data Using Deep Learning*. Zenodo. https://doi.org/10.5281/zenodo.15607678.
4. Calkoen, F. R., Luijendijk, A. P., Hanson, S., Nicholls, R., Moreno-Rodenas, A., De Heer, H., & Baart, F. (2025). *CoastBench: A Global Training Dataset for Coastal Classification Using Satellite Imagery and Elevation Data*. Zenodo. https://doi.org/10.5281/zenodo.15800285.
5. Calkoen, F. R. Luijendijk, A.P. (2025). *Sentinel-2 L2A Annual Composite (2023)*. Deltares. https://github.com/TUDelft-CITG/coastpy.
6. Calkoen, F. R., Luijendijk, A. P. (2025). *Global Coastal Grid*. Zenodo. https://doi.org/10.5281/zenodo.16925825.
7. Calkoen, F. R., Kras, E., & Luijendijk, A. P. (2025). *ShorelineMonitor: Satellite-derived Shorelines (1984-2024)*. Deltares.
8. Calkoen, F. R., Barli, P., Kras, E., & Luijendijk, A. P. (2025). *ShorelineMonitor: Satellite-derived Shoreline Series*. Deltares.
9. Calkoen, F. R., Luijendijk, A. P., Vos, K., Kras, E., Baart, F. (2024). *Global Coastal Transect System (GCTS)*. Zenodo. https://doi.org/10.5281/zenodo.14056925.

3 Software

1. Calkoen, F. R. (2024). `coastpy`: A Python library for scalable coastal analytics. GitHub Repository. https://github.com/TUDelft-CITG/coastpy. The software is described in detail in Calkoen et al. (2025f).
2. Calkoen, F. R., Bosboom, J., Pearson, S., Antolínez, J. A., Christiaanse, J. C., de Bruijn, K., Guerra-Medina, D., Hoogervorst, C. D., Pupić Vurilj, M., & van den Berg, M. (2023). *Coastal Dynamics Open Codebook*. GitHub Repository. https://github.com/FlorisCalkoen/CoastalCodebook. Archived repository, project continued at https://github.com/Coastal-Dynamics.

4 Talks, Workshops & Proceedings

1. Calkoen, F. R., (2025, June). *Future Shoreline Retreat*. Oral presentation at the CoCliCo H2020 Final Event, Brussels, Belgium.
2. Calkoen, F. R. (2024, October). *Enabling Coastal Analytics at Planetary Scale*. Oral presentation at the NCK Themadag Remote Sensing, Leiden, Netherlands.
3. Calkoen, F. R., Luijendijk, A. P., Hanson, S., & Nicholls, R. J. (2023). Mapping the Coast from Space: What Can We Learn from Pixels? In *AGU Fall Meeting Abstracts* (Vol. 2023, EP42A-02).
4. Vitousek, S., Vos, K., Brown, S., & Calkoen, F. R. (2023, December). *Coastal Change Monitoring with Earth-observing Satellites*. Session Convener at the AGU Fall Meeting, San Francisco, CA, USA.
5. Calkoen, F. R., (2023, May). *Coastal Monitoring at Continental Scale*. Oral presentation at the Coastal Imaging Research Network (CIRN) Workshop, North Carolina, USA.
6. Calkoen, F. R., Baart, F., Kras, E., & Luijendijk, A. P. (2023). A Novel Data Ecosystem for Coastal Analyses. In *EGU General Assembly Conference Abstracts* (EGU-15964).
7. Calkoen, F. R., Luijendijk, A. P., Moreno-Rodenas, A., & Aarninkhof, S. (2022). Mapping Coastal Typology Using Publicly Available Earth Observation Data and Deep Neural Networks. In *Coastal Engineering Proceedings* (37, pp. 158-158).
8. Calkoen, F. R. (2022, October). *Mapping Building Footprints Using Modern Cloud Technology*. Workshop at the CoCliCo H2020 General Assembly, Orléans, France.

5 Research Grants

1. Pearson, S., Bosboom, J., & Calkoen, F. R. (2024). *Coastal Dynamics Open Quizbook (CDoq)*. TU Delft Education Innovation Projects Grant.
2. Pearson, S., Bosboom, J., & Calkoen, F. R. (2023). *Coastal Dynamics Open Quizbook*. TU Delft Open Education Stimulation Fund Grant.

Acronyms

AI	Artificial Intelligence
API	Application Programming Interface
ARD	Analysis-Ready Data
CC BY	Creative Commons Attribution License
CF	Climate Forecast
CMIP	Climate Model Intercomparison Project
CNN	Convolutional Neural Network
CoCliCo	Coastal Climate Core Services
COG	Cloud-Optimized GeoTIFF
DEA	Digital Earth Australia
DEM	Digital Elevation Model
DL	Deep Learning
DNN	Deep Neural Network
DTM	Digital Terrain Model
EO	Earth Observation
FAIR	Findable, Accessible, Interoperable, Reusable
GCTR	Global Coastal Transect Repository
GCTS	Global Coastal Transect System
GDAL	Geospatial Data Abstraction Library
GEE	Google Earth Engine
GIS	Geospatial Information System
HPC	High Performance Computing
IPCC	Intergovernmental Panel on Climate Change
LECZ	Low-Elevation Coastal Zone
LSTM	Long Short-Term Memory
ML	Machine Learning
MSPC	Microsoft Planetary Computer
NDWI	Normalized Difference Water Index
NN	Neural Network
ODC	Open Data Cube
OGC	Open Geospatial Consortium
OSM	OpenStreetMap
Parquet	Apache Parquet
RCP	Representative Concentration Pathway
RSLR	Relative Sea-Level Rise
S-2	Sentinel-2
SDS	Satellite-Derived Shoreline
SLR	Sea-Level Rise
SSP	Shared Socioeconomic Pathway

STAC	Spatio-Temporal Asset Catalog
UTM	Universal Transverse Mercator
VLECZ	Very Low-Elevation Coastal Zone
VLM	Vertical Land Motion

Bibliography

Abernathey, R et al. (Aug. 31, 2017). "Pangeo NSF Earthcube Proposal". DOI: 10.6084/m9.figshare.5361094.v1.

Abernathey, R P et al. (Mar. 1, 2021). "Cloud-Native Repositories for Big Scientific Data". *Computing in Science & Engineering* 23:2, pp. 26–35. DOI: 10.1109/MCSE.2021.3059437.

Adusumilli, S et al. (2024). "Predicting Shoreline Changes Along the California Coast Using Deep Learning Applied to Satellite Observations". *Journal of Geophysical Research: Machine Learning and Computation* 1:3, e2024JH000172. DOI: 10.1029/2024JH000172.

Al Najar, M, G Thoumyre, E W J Bergsma, R Almar, R Benshila, and D G Wilson (Apr. 1, 2023). "Satellite Derived Bathymetry Using Deep Learning". *Machine Learning* 112:4, pp. 1107–1130. DOI: 10.1007/s10994-021-05977-w.

Anselin, L (1995). "Local Indicators of Spatial Association—LISA". *Geographical Analysis* 27:2, pp. 93–115. DOI: 10.1111/j.1538-4632.1995.tb00338.x.

Athanasiou, P, A van Dongeren, A Giardino, M I Vousdoukas, R Ranasinghe, and J Kwadijk (July 17, 2020). "Uncertainties in Projections of Sandy Beach Erosion Due to Sea Level Rise: An Analysis at the European Scale". *Scientific Reports* 10:1, p. 11895. DOI: 10.1038/s41598-020-68576-0.

Athanasiou, P, A van Dongeren, M Pronk, A Giardino, M Vousdoukas, and R Ranasinghe (Nov. 28, 2023). "Global Coastal Characteristics (GCC): A Global Dataset of Geophysical, Hydrodynamic, and Socioeconomic Coastal Indicators". *Earth System Science Data Discussions*, pp. 1–32. DOI: 10.5194/essd-2023-313.

Baart, F et al. (Oct. 20, 2011). "Using 18th Century Storm-Surge Data from the Dutch Coast to Improve the Confidence in Flood-Risk Estimates". *Natural Hazards and Earth System Sciences* 11:10, pp. 2791–2801. DOI: 10.5194/nhess-11-2791-2011.

Backeberg, B et al. (July 13, 2022). "An Open Compute and Data Federation as an Alternative to Monolithic Infrastructures for Big Earth Data Analytics". *Big Earth Data* 0:0, pp. 1–19. DOI: 10.1080/20964471.2022.2094953.

Bamunawala, J et al. (July 7, 2021). "Twenty-First-Century Projections of Shoreline Change along Inlet-Interrupted Coastlines". *Scientific Reports* 11:1, p. 14038. DOI: 10.1038/s41598-021-93221-9.

Barbier, E B, S D Hacker, C Kennedy, E W Koch, A C Stier, and B R Silliman (2011). "The Value of Estuarine and Coastal Ecosystem Services". *Ecological Monographs* 81:2, pp. 169–193. DOI: 10.1890/10-1510.1.

Barnard, P L et al. (Jan. 2025). "Projections of Multiple Climate-Related Coastal Hazards for the US Southeast Atlantic". *Nature Climate Change* 15:1, pp. 101–109. DOI: 10.1038/s41558-024-02180-2.

Bauer-Marschallinger, B and K Falkner (Aug. 1, 2023). "Wasting Petabytes: A Survey of the Sentinel-2 UTM Tiling Grid and Its Spatial Overhead". *ISPRS Journal of Photogrammetry and Remote Sensing* 202, pp. 682–690. DOI: 10.1016/j.isprsjprs.2023.07.015.

Baumann, P (1993). "Language Support for Raster Image Manipulation in Databases". In: *Graphics Modeling and Visualization in Science and Technology*. Ed. by M Göbel and J C Teixeira. Springer, Berlin, Heidelberg, pp. 236–245. DOI: 10.1007/978-3-642-77811-7_19.

Biewald, L (2020). *Experiment Tracking with Weights and Biases*. https://wandb.ai/site/articles/experiment-tracking.

Bird, E C F (Jan. 1, 1985). *Coastline Changes. A Global Review*. John Wiley and Sons Inc.,New York, NY.

Bishop-Taylor, R, R Nanson, S Sagar, and L Lymburner (Dec. 15, 2021). "Mapping Australia's Dynamic Coastline at Mean Sea Level Using Three Decades of Landsat Imagery". *Remote Sensing of Environment* 267, p. 112734. DOI: 10.1016/j.rse.2021.112734.

Bishop-Taylor, R, C Phillips, S Sagar, V Newey, and T Sutterley (May 29, 2025). "EoTides: Tide Modelling Tools for Large-Scale Satellite Earth Observation Analysis". *Journal of Open Source Software* 10:109, p. 7786. DOI: 10.21105/joss.07786.

Bonafilia, D, B Tellman, T Anderson, and E Issenberg (June 2020). "Sen1Floods11: A Georeferenced Dataset to Train and Test Deep Learning Flood Algorithms for Sentinel-1". In: *2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)*. 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW). IEEE, Seattle, WA, USA, pp. 835–845. DOI: 10.1109/CVPRW50498.2020.00113.

Box, G E P (Dec. 1976). "Science and Statistics". *Journal of the American Statistical Association* 71:356, pp. 791–799. DOI: 10.1080/01621459.1976.10480949.

Boyd, R, R Dalrymple, and B Zaitlin (Oct. 1992). "Classification of Clastic Coastal Depositional Environments". *Sedimentary Geology* 80:3–4, pp. 139–150. DOI: 10.1016/0037-0738(92)90037-R.

Breiman, L (Oct. 1, 2001). "Random Forests". *Machine Learning* 45:1, pp. 5–32. DOI: 10.1023/A:1010933404324.

Breiman, L, J Friedman, R A Olshen, and C J Stone (1984). *Classification and Regression Trees*. Wadsworth International Group.

Brown, C F et al. (June 9, 2022). "Dynamic World, Near Real-Time Global 10 m Land Use Land Cover Mapping". *Scientific Data* 9:1, 1, p. 251. DOI: 10.1038/s41597-022-01307-4.

Bruun, P (Feb. 1, 1962). "Sea-Level Rise as a Cause of Shore Erosion". *Journal of the Waterways and Harbors Division* 88:1, pp. 117–130. DOI: 10.1061/JWHEAU.0000252.

– (1988). "The Bruun Rule of Erosion by Sea-Level Rise: A Discussion on Large-Scale Two- and Three-Dimensional Usages". *Journal of Coastal Research* 4:4, pp. 627–648.

Buck, M (Feb. 23, 2025). "Living on the Edge: How Do Perceptions of Coastal Erosion Risk Affect Residential Mobility Decisions?" *Local Environment*, pp. 1–15. DOI: 10.1080/13549839.2025.2467868.

Bulleri, F and M G Chapman (Feb. 2010). "The Introduction of Coastal Infrastructure as a Driver of Change in Marine Environments". *Journal of Applied Ecology* 47:1, pp. 26–35. DOI: 10.1111/j.1365-2664.2009.01751.x.

Burke, L, Y Kura, K Kassem, C Revenga, M Spalding', and D McAllister (2001). "Coastal Ecosystems". *World Resources Institute Washington, DC*, p. 93.

Buscombe, D and A C Ritchie (July 2018). "Landscape Classification with Deep Neural Networks". *Geosciences* 8:7, 7, p. 244. DOI: 10.3390/geosciences8070244.

Buscombe, D, P Wernette, S Fitzpatrick, J Favela, E B Goldstein, and N M Enwright (Jan. 20, 2023). "A 1.2 Billion Pixel Human-Labeled Dataset for Data-Driven Classification of Coastal Environments". *Scientific Data* 10:1, p. 46. DOI: 10.1038/s41597-023-01929-2.

Calkoen, F, A Luijendijk, C R Rivero, E Kras, and F Baart (Jan. 2021). "Traditional vs. Machine-Learning Methods for Forecasting Sandy Shoreline Evolution Using Historic Satellite-Derived Shorelines". *Remote Sensing* 13:5, p. 934. DOI: 10.3390/rs13050934.

Calkoen, F R, A P Luijendijk, P Barli, J Hemmes, R Ranasinghe, and R J Nicholls (Sept. 2025a). "Global Coastal Erosion: Present-Day Exposure and Future Risk".

Calkoen, F R et al. (Apr. 9, 2025b). *Mapping the World's Coast: A Global 100 m Coastal Typology Derived from Satellite Data Using Deep Learning*. Zenodo. DOI: 10.5281/zenodo.15607678.

Calkoen, F R et al. (July 24, 2025c). "Mapping the World's Coast: A Global 100-m Coastal Typology Derived from Satellite Data Using Deep Learning". *Earth System Science Data Discussions*, pp. 1–37. DOI: 10.5194/essd-2025-388.

Calkoen, F R and A P Luijendijk (Aug. 22, 2025d). *The Global Coastal Transect Repository (GCTR): An Analyis-Ready, Cloud-Native Data Repository for Coastal Analytics*. Zenodo. DOI: 10.5281/zenodo.16928269.

Calkoen, F R, A P Luijendijk, P Barli, J Hemmes, R Ranasinghe, and R Nicholls (Aug. 22, 2025e). *Future Shoreline Projections*. Version 2025-08-22. Zenodo. DOI: 10.5281/zenodo.16928881.

Calkoen, F R, A P Luijendijk, K Vos, E Kras, and F Baart (Jan. 1, 2025f). "Enabling Coastal Analytics at Planetary Scale". *Environmental Modelling & Software* 183, p. 106257. DOI: 10.1016/j.envsoft.2024.106257.

Calkoen, F R et al. (July 3, 2025g). *CoastBench: A Global Training Dataset for Coastal Classification Using Satellite Imagery and Elevation Data*. Version 2025-04-09. Zenodo. DOI: 10.5281/zenodo.15800285.

Calvin, K et al. (July 25, 2023). *IPCC, 2023: Climate Change 2023: Synthesis Report. Contribution of Working Groups I, II and III to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change [Core Writing Team, H. Lee and J. Romero (Eds.)]. IPCC, Geneva, Switzerland.* Intergovernmental Panel on Climate Change (IPCC). DOI: 10.59327/IPCC/AR6-9789291691647.

Carter, R W G (1988). *Coastal Environments: An Introduction to the Physical, Ecological and Cultural Systems of Coastlines*. First paperback edition with corrections, 1989; second printing, 1991; third printing, 1991. Academic Press, London; San Diego; New York; Boston; Sydney; Tokyo; Toronto.

Carter, R W G and C D Woodroffe (1994). *Coastal Evolution: Late Quaternary Shoreline Morphodynamics*. Cambridge University Press.

Cartwright, N (June 9, 1983). *How the Laws of Physics Lie*. Oxford University Press. DOI: 10.1093/0198247044.001.0001.

Castelle, B, E Kras, G Masselink, T Scott, A Konstantinou, and A Luijendijk (June 6, 2024). "Satellite-Derived Sandy Shoreline Trends and Interannual Variability along the Atlantic Coast of Europe". *Scientific Reports* 14:1, p. 13002. DOI: 10.1038/s41598-024-63849-4.

Çelik, O İ (2024). "Leveraging Deep Learning for Coastal Monitoring: A VGG16-based Approach to Spectral and Textural Classification of Coastal Areas with Sentinel-2A Data". *Applied Ocean Research*.

Christen, P, D J Hand, and N Kirielle (Oct. 6, 2023). "A Review of the F-Measure: Its History, Properties, Criticism, and Alternatives". *ACM Comput. Surv.* 56:3, 73:1–73:24. DOI: 10.1145/3606367.

Cooper, J A G and S McLaughlin (1998). "Contemporary Multidisciplinary Approaches to Coastal Classification and Environmental Risk Analysis". *Journal of Coastal Research* 14:2, pp. 512–524.

Cooper, J a G et al. (Nov. 2020). "Sandy Beaches Can Survive Sea-Level Rise". *Nature Climate Change* 10:11, pp. 993–995. DOI: 10.1038/s41558-020-00934-2.

Cooper, J A G and O H Pilkey (Nov. 2004). "Sea-Level Rise and Shoreline Retreat: Time to Abandon the Bruun Rule". *Global and Planetary Change* 43:3–4, pp. 157–171. DOI: 10.1016/j.gloplacha.2004.07.001.

Cornillon, P, J Gallagher, and T Sgouros (2003). "OPeNDAP: Accessing Data in a Distributed, Heterogeneous Environment". *Data Science Journal* 2, pp. 164–174. DOI: 10.2481/dsj.2.164.

Cortes, C and V Vapnik (Sept. 1995). "Support-Vector Networks". *Machine Learning* 20:3, pp. 273–297. DOI: 10.1007/BF00994018.

Cosby, A G et al. (Sept. 28, 2024). "Accelerating Growth of Human Coastal Populations at the Global and Continent Levels: 2000–2018". *Scientific Reports* 14:1, p. 22489. DOI: 10.1038/s41598-024-73287-x.

Cowell, P J et al. (2003). "The Coastal-Tract (Part 1): A Conceptual Approach to Aggregated Modeling of Low-Order Coastal Change". *Journal of Coastal Research* 19:4, pp. 812–827.

D'Anna, M, D Idier, B Castelle, S Vitousek, and G Le Cozannet (Sept. 2021). "Reinterpreting the Bruun Rule in the Context of Equilibrium Shoreline Models". *Journal of Marine Science and Engineering* 9:9, 9, p. 974. DOI: 10.3390/jmse9090974.

Dang, K B, V B Dang, Q T Bui, V V Nguyen, T P N Pham, and V L Ngo (2020). "A Convolutional Neural Network for Coastal Classification Based on ALOS and NOAA Satellite Data". *IEEE Access* 8, pp. 11824–11839. DOI: 10.1109/ACCESS.2020.2965231.

Dangendorf, S et al. (Sept. 2019). "Persistent Acceleration in Global Sea-Level Rise since the 1960s". *Nature Climate Change* 9:9, pp. 705–710. DOI: 10.1038/s41558-019-0531-8.

Davidson-Arnott, R G D (Nov. 2005). "Conceptual Model of the Effects of Sea Level Rise on Sandy Coasts". *Journal of Coastal Research* 2005:216, pp. 1166–1172. DOI: 10.2112/03-0051.1.

Davies, J (Jan. 17, 1964). "A Morphogenic Approach to World Shorelines". *Zeitschrift für Geomorphologie* 8:5, pp. 127–142. DOI: 10.1127/zfg/mortensen/8/1964/127.

De Heer, H J, F R Calkoen, and A Moreno-Rodenas (2025). "Geospatial Classification through Fine-Tuned Large Language and Vision Models".

Dean, J and S Ghemawat (Jan. 1, 2008). "MapReduce: Simplified Data Processing on Large Clusters". *Communications of the ACM* 51:1, pp. 107–113. DOI: 10.1145/1327452.1327492.

Defeo, O et al. (Jan. 1, 2009). "Threats to Sandy Beach Ecosystems: A Review". *Estuarine, Coastal and Shelf Science* 81:1, pp. 1–12. DOI: 10.1016/j.ecss.2008.09.022.

Deng, J, W Dong, R Socher, L J Li, K Li, and L Fei-Fei (June 2009). "ImageNet: A Large-Scale Hierarchical Image Database". In: *2009 IEEE Conference on Computer Vision and Pattern Recognition*. 2009 IEEE Conference on Computer Vision and Pattern Recognition, pp. 248–255. DOI: 10.1109/CVPR.2009.5206848.

Directorate-General for Research and Innovation (European Commission) (2022). *The Digital Twin Ocean: An Interactive Replica of the Ocean for Better Decision Making*. Publications Office of the European Union.

Dodet, G et al. (Jan. 2019). "Beach Recovery from Extreme Storm Activity during the 2013–14 Winter along the Atlantic Coast of Europe". *Earth Surface Processes and Landforms* 44:1, pp. 393–401. DOI: 10.1002/esp.4500.

Donchyts, G, F Baart, H Winsemius, N Gorelick, J Kwadijk, and N van de Giesen (Sept. 2016a). "Earth's Surface Water Change over the Past 30 Years". *Nature Climate Change* 6:9, pp. 810–813. DOI: 10.1038/nclimate3111.

Donchyts, G, J Schellekens, H Winsemius, E Eisemann, and N Van de Giesen (May 2016b). "A 30 m Resolution Surface Water Mask Including Estimation of Positional and Thematic Differences Using Landsat 8, SRTM and OpenStreetMap: A Case Study in the Murray-Darling Basin, Australia". *Remote Sensing* 8:5, p. 386. DOI: 10.3390/rs8050386.

Doody, J P (Jan. 1, 2004). "'Coastal Squeeze'— an Historical Perspective". *Journal of Coastal Conservation* 10:1, pp. 129–138. DOI: 10.1652/1400-0350(2004)010[0129:CSAHP]2.0.CO;2.

Dugan, J E, D M Hubbard, I F Rodil, D L Revell, and S Schroeter (July 2008). "Ecological Effects of Coastal Armoring on Sandy Beaches". *Marine Ecology* 29:s1, pp. 160–170. DOI: 10.1111/j.1439-0485.2008.00231.x.

Dunn, F E, S E Darby, R J Nicholls, S Cohen, C Zarfl, and B M Fekete (Aug. 1, 2019). "Projections of Declining Fluvial Sediment Delivery to Major Deltas Worldwide in Response to Climate Change and Anthropogenic Stress". *Environmental Research Letters* 14:8, p. 084034. DOI: 10.1088/1748-9326/ab304e.

Durbin, C, P Quinn, and D Shum (2020). *Task 51-Cloud-Optimized Format Study*. NASA.

Eaton, B et al. (2003). "NetCDF Climate and Forecast (CF) Metadata Conventions". *URL: http://cfconventions. org/Data/cf-conventions/cf-conventions-1.8/cf-conventions. pdf*.

European Environment Agency. (2024). *European Climate Risk Assessment*. Publications Office, LU.

European Space Agency (ESA) (2019). *Copernicus DEM - Global and European Digital Elevation Model*. DOI: 10.5270/ESA-c5d3d65.

Eyring, V et al. (May 26, 2016). "Overview of the Coupled Model Intercomparison Project Phase 6 (CMIP6) Experimental Design and Organization". *Geoscientific Model Development* 9:5, pp. 1937–1958. DOI: 10.5194/gmd-9-1937-2016.

Fairbridge, R W (Jan. 1, 2004). "Classification of Coasts". *Journal of Coastal Research* 20, 1 (201), pp. 155–165. DOI: 10.2112/1551-5036(2004)20[155:COC]2.0.CO;2.

fiboa contributors (2025). *Field Boundaries for Agriculture (Fiboa) Specification*.

Finkl, C W (2004). "Coastal Classification: Systematic Approaches to Consider in the Development of a Comprehensive Scheme". *Journal of Coastal Research* 20:1, pp. 166–213.

Finkl, C W and C Makowski (June 17, 2020). "Coastal Belt Linked Classification (CBLC): A System for Characterizing the Interface between Land and Sea Based on Large Marine Ecosystems, Coastal Ecological Sequences, and Terrestrial Ecoregions". *Journal of Coastal Research* 36:4, pp. 677–693. DOI: 10.2112/JCOASTRES-D-20A-00001.1.

Fox-Kemper, B et al. (2021). "Climate Change 2021: The Physical Science Basis. Contribution of Working Group I to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change". *Cambridge Univ. Press*.

French, J, H Burningham, G Thornhill, R Whitehouse, and R J Nicholls (Mar. 1, 2016). "Conceptualising and Mapping Coupled Estuary, Coast and Inner Shelf Sediment Systems". *Geomorphology*. Simulating Decadal Coastal Morphodynamics 256, pp. 17–35. DOI: 10.1016/j.geomorph.2015.10.006.

Galgano, F A, B C Douglas, and S P Leatherman (1998). "Trends and Variability of Shoreline Position". *Journal of Coastal Research*, pp. 282–291.

Gavin, D et al. (July 2018). "Digital Earth Australia - from Satellite Data to Better Decisions". In: *IGARSS 2018 - 2018 IEEE International Geoscience and Remote Sensing Symposium*. IGARSS 2018 - 2018 IEEE International Geoscience and Remote Sensing Symposium. IEEE, Valencia, pp. 8633–8635. DOI: 10.1109/IGARSS.2018.8518160.

Gentemann, C L et al. (2021). "Science Storms the Cloud". *AGU Advances* 2:2. DOI: 10.1029/2020AV000354.

Gomez-de la Peña, E, G Coco, C Whittaker, and J Montaño (Nov. 13, 2023). "On the Use of Convolutional Deep Learning to Predict Shoreline Change". *Earth Surface Dynamics* 11:6, pp. 1145–1160. DOI: 10.5194/esurf-11-1145-2023.

Goodfellow, I, Y Bengio, and A Courville (2016). *Deep Learning*. The MIT Press.

Gorelick, N, M Hancher, M Dixon, S Ilyushchenko, D Thau, and R Moore (2017). "Google Earth Engine: Planetary-scale Geospatial Analysis for Everyone". *Remote Sensing of Environment* 202, pp. 18–27. DOI: 10.1016/j.rse.2017.06.031.

Gornitz, V (Mar. 1, 1991). "Global Coastal Hazards from Future Sea Level Rise". *Palaeogeography, Palaeoclimatology, Palaeoecology* 89:4, pp. 379–398. DOI: 10.1016/0031-0182(91)90173-O.

Grandjean, T J et al. (June 2024). "Critical Turbidity Thresholds for Maintenance of Estuarine Tidal Flats Worldwide". *Nature Geoscience* 17:6, pp. 539–544. DOI: 10.1038/s41561-024-01431-3.

Green, A N, J A G Cooper, and L Salzmann (Feb. 2014). "Geomorphic and Stratigraphic Signals of Postglacial Meltwater Pulses on Continental Shelves". *Geology* 42:2, pp. 151–154. DOI: 10.1130/G35052.1.

Hamman, J, M Rocklin, and R Abernathy (Apr. 1, 2018). "Pangeo: A Big-data Ecosystem for Scalable Earth System Science", p. 12146.

Hansen, M C et al. (Nov. 15, 2013). "High-Resolution Global Maps of 21st-Century Forest Cover Change". *Science* 342:6160, pp. 850–853. DOI: 10.1126/science.1244693.

Hanson, S et al. (Sept. 2010). "Capturing Coastal Geomorphological Change within Regional Integrated Assessment: An Outcome-Driven Fuzzy Logic Approach". *Journal of Coastal Research* 265, pp. 831–842. DOI: 10.2112/JCOASTRES-D-09-00078.1.

Hanson, S E, R J Nicholls, F R Calkoen, G Le Cozannet, and A P Luijendijk (July 2, 2025). *A Geospatial Database of Coastal Characteristics for Erosion Assessment of Europe's Coastal Floodplains*. DOI: 10.5194/egusphere-2025-2371. Pre-published.

Hapke, C J, M G Kratzmann, and E A Himmelstoss (Oct. 2013). "Geomorphic and Human Influence on Large-Scale Coastal Change". *Geomorphology* 199, pp. 160–170. DOI: 10.1016/j.geomorph.2012.11.025.

Hauer, M E et al. (Dec. 9, 2019). "Sea-Level Rise and Human Migration". *Nature Reviews Earth & Environment* 1:1, pp. 28–39. DOI: 10.1038/s43017-019-0002-9.

Hayes, M O (Apr. 1, 1967). "Relationship between Coastal Climate and Bottom Sediment Type on the Inner Continental Shelf". *Marine Geology* 5:2, pp. 111–132. DOI: 10.1016/0025-3227(67)90074-6.

He, K, X Zhang, S Ren, and J Sun (June 2016). "Deep Residual Learning for Image Recognition". In: *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). IEEE, Las Vegas, NV, USA, pp. 770–778. DOI: 10.1109/CVPR.2016.90.

Hijma, M P et al. (Mar. 20, 2025). "Global Sea-Level Rise in the Early Holocene Revealed from North Sea Peats". *Nature* 639:8055, pp. 652–657. DOI: 10.1038/s41586-025-08769-7.

Hinkel, J et al. (Dec. 1, 2013). "A Global Analysis of Erosion of Sandy Beaches and Sea-Level Rise: An Application of DIVA". *Global and Planetary Change* 111, pp. 150–158. DOI: 10.1016/j.gloplacha.2013.09.002.

Hochreiter, S and J Schmidhuber (1997). "Long Short-Term Memory". *Neural Comput.* 9:8, pp. 1735–1780. DOI: 10.1162/neco.1997.9.8.1735.

Holman, R and J Stanley (June 2007). "The History and Technical Capabilities of Argus". *Coastal Engineering* 54:6–7, pp. 477–491. DOI: 10.1016/j.coastaleng.2007.01.003.

Holmes, C (Oct. 27, 2023). *The Admin-partitioned GeoParquet Distribution*. Radiant Earth Insights. https://medium.com/radiant-earth-insights/the-admin-partitioned-geoparquet-distribution-59f0ca1c6d96.

Hoozemans, F M J, M Marchand, and H A Pennekamp (1993). *A Global Vulnerability Analysis: Vulnerability Assessment for Population, Coastal Wetlands and Rice Production on a Global Scale*. Delft Hydraulics, the Netherlands.

Hormann, C (Sept. 2014). "Generalisierung im Raster für Karten kleiner Maßstäbe — mit Anwendungsbeispielen aus OpenStreetMap". *KN - Journal of Cartography and Geographic Information* 64:5, pp. 276–280. DOI: 10.1007/BF03544188.

Hoyer, S and J Hamman (Apr. 5, 2017). "Xarray: N-D Labeled Arrays and Datasets in Python". *Journal of Open Research Software* 5:1. DOI: 10.5334/jors.148.

Hulskamp, R et al. (Dec. 13, 2023). "Global Distribution and Dynamics of Muddy Coasts". *Nature Communications* 14:1, p. 8259. DOI: 10.1038/s41467-023-43819-6.

Hutton, J (1788). "X. Theory of the Earth; or an Investigation of the Laws Observable in the Composition, Dissolution, and Restoration of Land upon the Globe." *Earth and Environmental Science Transactions of the Royal Society of Edinburgh* 1:2, pp. 209–304.

Inman, D L and C E Nordstrom (1971). "On the Tectonic and Morphologic Classification of Coasts". *The Journal of Geology* 79:1, pp. 1–21.

Jakubik, J et al. (Nov. 8, 2023). *Foundation Models for Generalist Geospatial Artificial Intelligence*. DOI: 10.48550/arXiv.2310.18660. Pre-published.

Jones, B and B C O'Neill (July 2016). "Spatially Explicit Global Population Scenarios Consistent with the Shared Socioeconomic Pathways". *Environ. Res. Lett.* 11:8, p. 084003. DOI: 10.1088/1748-9326/11/8/084003.

Jones, N (Sept. 12, 2018). "How to Stop Data Centres from Gobbling up the World's Electricity". *Nature* 561:7722, pp. 163–166. DOI: 10.1038/d41586-018-06610-y.

Jumper, J et al. (Aug. 2021). "Highly Accurate Protein Structure Prediction with AlphaFold". *Nature* 596:7873, pp. 583–589. DOI: 10.1038/s41586-021-03819-2.

Katal, A, S Dahiya, and T Choudhury (2023). "Energy Efficiency in Cloud Computing Data Centers: A Survey on Software Technologies". *Cluster Computing* 26:3, pp. 1845–1875. DOI: 10.1007/s10586-022-03713-0.

Killough, B (July 2018). "Overview of the Open Data Cube Initiative". In: *IGARSS 2018 - 2018 IEEE International Geoscience and Remote Sensing Symposium*. IGARSS 2018 - 2018 IEEE International Geoscience and Remote Sensing Symposium. IEEE, Valencia, pp. 8629–8632. DOI: 10.1109/IGARSS.2018.8517694.

Konstantinou, A, T Scott, G Masselink, K Stokes, D Conley, and B Castelle (Aug. 1, 2023). "Satellite-Based Shoreline Detection along High-Energy Macrotidal Coasts and Influence of Beach State". *Marine Geology* 462, p. 107082. DOI: 10.1016/j.margeo.2023.107082.

Kraft, J C and M J Chrzastowski (1985). "Coastal Stratigraphic Sequences". In: *Coastal Sedimentary Environments*. Ed. by R A Davis. Springer New York, New York, NY, pp. 625–663. DOI: 10.1007/978-1-4612-5078-4_9.

Krizhevsky, A, I Sutskever, and G E Hinton (2012). "ImageNet Classification with Deep Convolutional Neural Networks". In: *Advances in Neural Information Processing Systems*. Vol. 25, pp. 1097–1105.

Kulp, S A and B H Strauss (Oct. 29, 2019). "New Elevation Data Triple Estimates of Global Vulnerability to Sea-Level Rise and Coastal Flooding". *Nat Commun* 10:1, p. 4844. DOI: 10.1038/s41467-019-12808-z.

Lam, R et al. (Dec. 22, 2023). "Learning Skillful Medium-Range Global Weather Forecasting". *Science* 382:6677, pp. 1416–1421. DOI: 10.1126/science.adi2336.

Lambeck, K et al. (2010). *Paleoenvironmental Records, Geophysical Modelling and Reconstruction of Sea Level Trends and Variability on Centennial and Longer Time Scales*. Wiley-Blackwell New York.

Lansu, E M et al. (Jan. 10, 2024). "A Global Analysis of How Human Infrastructure Squeezes Sandy Coasts". *Nature Communications* 15:1, p. 432. DOI: 10.1038/s41467-023-44659-0.

Le, Q V (2013). "Building High-Level Features Using Large Scale Unsupervised Learning". In: *2013 IEEE International Conference on Acoustics, Speech and Signal Processing*. IEEE, pp. 8595–8598.

Le Cozannet, G, M Garcin, M Yates, D Idier, and B Meyssignac (Nov. 1, 2014). "Approaches to Evaluate the Recent Impacts of Sea-Level Rise on Shoreline Changes". *Earth-Science Reviews* 138, pp. 47–60. DOI: 10.1016/j.earscirev.2014.08.005.

Le Cozannet, G et al. (Dec. 2017). "Sea Level Change and Coastal Climate Services: The Way Forward". *Journal of Marine Science and Engineering* 5:4, p. 49. DOI: 10.3390/jmse5040049.

LeCun, Y, Y Bengio, and G Hinton (May 28, 2015). "Deep Learning". *Nature* 521:7553, pp. 436–444. DOI: 10.1038/nature14539.

Lewis, A et al. (Dec. 1, 2017). "The Australian Geoscience Data Cube — Foundations and Lessons Learned". *Remote Sensing of Environment*. Big Remotely Sensed Data: Tools, Applications and Experiences 202, pp. 276–292. DOI: 10.1016/j.rse.2017.03.015.

Lincke, D and J Hinkel (July 2018). "Economically Robust Protection against 21st Century Sea-Level Rise". *Global Environmental Change* 51, pp. 67–73. DOI: 10.1016/j.gloenvcha.2018.05.003.

Luijendijk, A, G Hagenaars, R Ranasinghe, F Baart, G Donchyts, and S Aarninkhof (Dec. 2018). "The State of the World's Beaches". *Sci Rep* 8:1, p. 6641. DOI: 10.1038/s41598-018-24630-6.

Mahecha, M D et al. (Feb. 25, 2020). "Earth System Data Cubes Unravel Global Multivariate Dynamics". *Earth System Dynamics* 11:1, pp. 201–234. DOI: 10.5194/esd-11-201-2020.

Mao, Y, D L Harris, Z Xie, and S Phinn (Sept. 2022). "Global Coastal Geomorphology – Integrating Earth Observation and Geospatial Data". *Remote Sensing of Environment* 278, p. 113082. DOI: 10.1016/j.rse.2022.113082.

– (Nov. 2021). "Efficient Measurement of Large-Scale Decadal Shoreline Change with Increased Accuracy in Tide-Dominated Coastal Environments with Google Earth Engine". *ISPRS Journal of Photogrammetry and Remote Sensing* 181, pp. 385–399. DOI: 10.1016/j.isprsjprs.2021.09.021.

Martínez, M L, A Intralawan, G Vázquez, O Pérez-Maqueo, P Sutton, and R Landgrave (Aug. 1, 2007). "The Coasts of Our World: Ecological, Economic and Social Importance". *Ecological Economics*. Ecological Economics of Coastal Disasters 63:2, pp. 254–272. DOI: 10.1016/j.ecolecon.2006.10.022.

Masselink, G (2014). *Introduction to Coastal Processes & Geomorphology*. Second edition. Routledge, Oxon [England]. 1 p. DOI: 10.4324/9780203785461.

McGill, J T (July 1958). "Map of Coastal Landforms of the World". *Geographical Review* 48:3, p. 402. DOI: 10.2307/212259.

McGranahan, G, D Balk, and B Anderson (Apr. 1, 2007). "The Rising Tide: Assessing the Risks of Climate Change and Human Settlements in Low Elevation Coastal Zones". *Environment and Urbanization* 19:1, pp. 17–37. DOI: 10.1177/0956247807076960.

McKinney, W (2010). "Data Structures for Statistical Computing in Python". In: Python in Science Conference. Austin, Texas, pp. 56–61. DOI: 10.25080/Majora-92bf1922-00a.

Medvedev, D, G Lemson, and M Rippin (July 18, 2016). "SciServer Compute: Bringing Analysis Close to the Data". In: *Proceedings of the 28th International Conference on Scientific and Statistical Database Management*. SSDBM '16. Association for Computing Machinery, New York, NY, USA, pp. 1–4. DOI: 10.1145/2949689.2949700.

Mentaschi, L, M I Vousdoukas, JF Pekel, E Voukouvalas, and L Feyen (Dec. 2018). "Global Long-Term Observations of Coastal Erosion and Accretion". *Sci Rep* 8:1, p. 12876. DOI: 10.1038/s41598-018-30904-w.

Mercer, J H (Jan. 1, 1978). "West Antarctic Ice Sheet and CO2 Greenhouse Effect: A Threat of Disaster". *Nature* 271:5643, pp. 321–325. DOI: 10.1038/271321a0.

Mikkelsen, A B, K K McDonald, J Kalksma, Z H Tyrrell, and C H Fletcher (Mar. 29, 2024). "Three Years of Weekly DEMs, Aerial Orthomosaics and Surveyed Shoreline Positions at Waikīkī Beach, Hawai'i". *Scientific Data* 11:1, p. 324. DOI: 10.1038/s41597-024-03160-z.

Mitchell, T M (1997). *Machine Learning*. McGraw-Hill Series in Computer Science. McGraw-Hill, New York. 414 pp.

Mohr, M et al. (Feb. 1, 2025). "Federated and Reusable Processing of Earth Observation Data". *Scientific Data* 12:1, p. 194. DOI: 10.1038/s41597-025-04513-y.

Monserrate, S G (Jan. 27, 2022). "The Cloud Is Material: On the Environmental Impacts of Computation and Data Storage". *MIT Case Studies in Social and Ethical Responsibilities of Computing*, Winter 2022. DOI: 10.21428/2c646de5.031d4553.

Moore, L J et al. (June 2025). "Ecomorphodynamics of Coastal Foredune Evolution". *Nature Reviews Earth & Environment* 6:6, pp. 417–432. DOI: 10.1038/s43017-025-00672-z.

Muir, F M E, M D Hurst, L Richardson-Foulger, A F Rennie, and L A Naylor (May 14, 2024). "VedgeSat: An Automated, Open-Source Toolkit for Coastal Change Monitoring Using Satellite-Derived Vegetation Edges". *Earth Surface Processes and Landforms*. DOI: 10.1002/esp.5835.

Murray, N J et al. (2018). "The Global Distribution and Trajectory of Tidal Flats". *Nature* 565:7738, pp. 222–225. DOI: 10.1038/s41586-018-0805-8.

Nawarat, K, J Reyns, M I Vousdoukas, T M Duong, E Kras, and R Ranasinghe (Dec. 5, 2024). "Coastal Hardening and What It Means for the World's Sandy Beaches". *Nature Communications* 15:1, p. 10626. DOI: 10.1038/s41467-024-54952-1.

Nerem, R S, B D Beckley, J T Fasullo, B D Hamlington, D Masters, and G T Mitchum (Feb. 27, 2018). "Climate-Change–Driven Accelerated Sea-Level Rise Detected in the Altimeter Era". *Proceedings of the National Academy of Sciences* 115:9, pp. 2022–2025. DOI: 10.1073/pnas.1717312115.

Neumann, B, A T Vafeidis, J Zimmermann, and R J Nicholls (2015). "Future Coastal Population Growth and Exposure to Sea-Level Rise and Coastal Flooding - A Global Assessment". *PLOS ONE* 10:3. Ed. by L Kumar, e0118571. DOI: 10.1371/journal.pone.0118571.

Nicholls, R J (July 5, 2010). "Impacts of and Responses to Sea-Level Rise". In: *Understanding Sea-Level Rise and Variability*. Ed. by J A Church, P L Woodworth, T Aarup, and W S Wilson. Wiley-Blackwell, Oxford, UK, pp. 17–51. DOI: 10.1002/9781444323276.ch2.

Nicholls, R J and C Ballesteros (2025). "Sea-Level Change and Its Implications". In: *Reference Module in Earth Systems and Environmental Sciences*. Elsevier, B9780323960267002034. DOI: 10.1016/B978-0-323-96026-7.00203-4.

Nicholls, R J et al. (Apr. 2021). "A Global Analysis of Subsidence, Relative Sea-Level Change and Coastal Flood Exposure". *Nature Climate Change* 11:4, 4, pp. 338–342. DOI: 10.1038/s41558-021-00993-z.

Notarangelo, N, C Wirion, and F Van Winsen (Feb. 6, 2025). "STURM-Flood: A Curated Dataset for Deep Learning-Based Flood Extent Mapping Leveraging Sentinel-1 and Sentinel-2 Imagery". *Big Earth Data*, pp. 1–27. DOI: 10.1080/20964471.2025.2458714.

Nyberg, B, A Gilmullina, W Helland-Hansen, J Nienhuis, and J Storms (Jan. 2025). "Global Coastal Exposure Patterns by Coastal Type from 1950 to 2050". *Cambridge Prisms: Coastal Futures* 3, e12. DOI: 10.1017/cft.2025.10001.

Oppenheimer, M et al. (2019). *Sea Level Rise and Implications for Low Lying Islands, Coasts and Communities*. The Intergovernmental Panel on Climate Change.

Overture Maps Foundation (2025). *Overture Maps Buildings Theme, Release 2025-04-23.0*.

Paszke, A et al. (2019). "PyTorch: An Imperative Style, High-Performance Deep Learning Library". In: *Advances in Neural Information Processing Systems 32*. Ed. by H Wallach,

H Larochelle, A Beygelzimer, F d'Alché-Buc, E Fox, and R Garnett. Curran Associates, Inc., pp. 8024–8035.

Pekel, JF, A Cottam, N Gorelick, and A S Belward (Dec. 2016). "High-Resolution Mapping of Global Surface Water and Its Long-Term Changes". *Nature* 540:7633, 7633, pp. 418–422. DOI: 10.1038/nature20584.

Pontee, N (Nov. 2013). "Defining Coastal Squeeze: A Discussion". *Ocean & Coastal Management* 84, pp. 204–207. DOI: 10.1016/j.ocecoaman.2013.07.010.

Pranzini, E, L Wetzel, and A T Williams (Aug. 2015). "Aspects of Coastal Erosion and Protection in Europe". *Journal of Coastal Conservation* 19:4, pp. 445–459. DOI: 10.1007/s11852-015-0399-3.

Pronk, M et al. (Mar. 6, 2024). "DeltaDTM: A Global Coastal Digital Terrain Model". *Scientific Data* 11:1, p. 273. DOI: 10.1038/s41597-024-03091-9.

Raasveldt, M and H Mühleisen (June 25, 2019). "DuckDB: An Embeddable Analytical Database". In: *Proceedings of the 2019 International Conference on Management of Data*. SIGMOD '19. Association for Computing Machinery, New York, NY, USA, pp. 1981–1984. DOI: 10.1145/3299869.3320212.

Ranasinghe, R (Feb. 6, 2020). "On the Need for a New Generation of Coastal Change Models for the 21st Century". *Scientific Reports* 10:1, 1, p. 2010. DOI: 10.1038/s41598-020-58376-x.

Ranasinghe, R, D Callaghan, and M J F Stive (Feb. 1, 2012). "Estimating Coastal Recession Due to Sea Level Rise: Beyond the Bruun Rule". *Climatic Change* 110:3, pp. 561–574. DOI: 10.1007/s10584-011-0107-8.

Ranasinghe, R, T M Duong, S Uhlenbrook, D Roelvink, and M Stive (Jan. 2013). "Climate-Change Impact Assessment for Inlet-Interrupted Coastlines". *Nature Climate Change* 3:1, pp. 83–87. DOI: 10.1038/nclimate1664.

Ranasinghe, R and M J F Stive (Dec. 2009). "Rising Seas and Retreating Coastlines". *Climatic Change* 97:3, pp. 465–468. DOI: 10.1007/s10584-009-9593-3.

Ranasinghe, R, Z B Wang, J Bamunawala, and T M Duong (Feb. 4, 2025). "On the Time Lag between Sea-Level Rise and Basin Infilling at Tidal Inlets". *Scientific Reports* 15:1, p. 4231. DOI: 10.1038/s41598-025-86699-0.

Ranasinghe, R et al. (2021). "Climate Change Information for Regional Impact and for Risk Assessment".

Rangel-Buitrago, N, W Neal, O Pilkey, and N Longo (Mar. 15, 2023). "The Global Impact of Sand Mining on Beaches and Dunes". *Ocean & Coastal Management* 235, p. 106492. DOI: 10.1016/j.ocecoaman.2023.106492.

Raoult, B, C Bergeron, A López Alós, JN Thépaut, and D Dee (2017). "Climate Service Develops User-Friendly Data Store". *ECMWF newsletter* Meteorology, pp. 22–27. DOI: 10.21957/P3C285.

Raymond, E (1999). "The Cathedral and the Bazaar". *Knowledge, Technology & Policy* 12:3, pp. 23–49.

Reichstein, M et al. (Feb. 2019). "Deep Learning and Process Understanding for Data-Driven Earth System Science". *Nature* 566:7743, pp. 195–204. DOI: 10.1038/s41586-019-0912-1.

Reisinger, A et al. (Sept. 2020). *The Concept of Risk in the IPCC Sixth Assessment Report: A Summary of Cross-Working Group Discussions. Guidance for IPCC Authors*. Intergovernmental Panel on Climate Change (IPCC).

Rocklin, M (2015). "Dask: Parallel Computation with Blocked Algorithms and Task Scheduling". In: *Proceedings of the 14th Python in Science Conference*. Vol. 130. Citeseer, p. 136.

Roelvink, D, B Huisman, A Elghandour, M Ghonim, and J Reyns (2020). "Efficient Modeling of Complex Sandy Coastal Evolution at Monthly to Century Time Scales". *Frontiers in Marine Science* 7.

Ruder, S (June 15, 2017). *An Overview of Multi-Task Learning in Deep Neural Networks*. DOI: 10.48550/arXiv.1706.05098. Pre-published.

Rumelhart, D E, G E Hinton, and R J Williams (1986). "Learning Representations by Back-Propagating Errors". *Nature* 323:6088, pp. 533–536.

Saintilan, N et al. (June 5, 2020). "Thresholds of Mangrove Survival under Rapid Sea Level Rise". *Science* 368:6495, pp. 1118–1121. DOI: 10.1126/science.aba2656.

Salman, A, S Lombardo, and P Doody (2004). "Living with Coastal Erosion in Europe: Sediment and Space for Sustainability". *Eurosion project reports*.

Schramm, M et al. (Jan. 2021). "The openEO API–Harmonising the Use of Earth Observation Cloud Services Using Virtual Data Cube Functionalities". *Remote Sensing* 13:6, p. 1125. DOI: 10.3390/rs13061125.

Schuerch, M et al. (Sept. 2018). "Future Response of Global Coastal Wetlands to Sea-Level Rise". *Nature* 561:7722, 7722, pp. 231–234. DOI: 10.1038/s41586-018-0476-5.

SCOR Working Group 89 (1991). "The Response of Beaches to Sea-Level Changes: A Review of Predictive Models". *Journal of Coastal Research* 7:3, pp. 895–921.

Seale, C, T Redfern, P Chatfield, C Luo, and K Dempsey (Sept. 2022). "Coastline Detection in Satellite Imagery: A Deep Learning Approach on New Benchmark Data". *Remote Sensing of Environment* 278, p. 113044. DOI: 10.1016/j.rse.2022.113044.

Sharples, C et al. (2009). "The Australian Coastal Smartline Geomorphic and Stability Map Version 1: Project Report". *Prepared for Geoscience Australia and the Department of Climate Change by the School of Geography and Environmental Studies, University of Tasmania, Hobart*.

Slagel, M J and G B Griggs (May 2008). "Cumulative Losses of Sand to the California Coast by Dam Impoundment". *Journal of Coastal Research* 243, pp. 571–584. DOI: 10.2112/06-0640.1.

Small, C and R J Nicholls (2003). "A Global Analysis of Human Settlement in Coastal Zones". *Journal of Coastal Research* 19:3, pp. 584–599.

Smith, K E et al. (Mar. 2025). "Ocean Extremes as a Stress Test for Marine Ecosystems and Society". *Nature Climate Change* 15:3, pp. 231–235. DOI: 10.1038/s41558-025-02269-2.

Stern, C et al. (Feb. 10, 2022). "Pangeo Forge: Crowdsourcing Analysis-Ready, Cloud Optimized Data Production". *Frontiers in Climate* 3. DOI: 10.3389/fclim.2021.782909.

Stive, M J F and Z B Wang (Jan. 1, 2003). "Chapter 13 Morphodynamic Modeling of Tidal Basins and Coastal Inlets". In: *Elsevier Oceanography Series*. Ed. by V C Lakhan. Vol. 67. Advances in Coastal Modeling. Elsevier, pp. 367–392. DOI: 10.1016/S0422-9894(03)80130-7.

Stive, M J et al. (Dec. 2002). "Variability of Shore and Shoreline Evolution". *Coastal Engineering* 47:2, pp. 211–235. DOI: 10.1016/S0378-3839(02)00126-6.

Tenopir, C et al. (Aug. 26, 2015). "Changes in Data Sharing and Data Reuse Practices and Perceptions among Scientists Worldwide". *PLOS ONE* 10:8. Ed. by P Van Den Besselaar, e0134826. DOI: 10.1371/journal.pone.0134826.

Thiéblemont, R, G Le Cozannet, J Rohmer, A Toimil, M Álvarez-Cuesta, and I J Losada (July 30, 2021). "Deep Uncertainties in Shoreline Change Projections: An Extra-Probabilistic Approach Applied to Sandy Beaches". *Natural Hazards and Earth System Sciences* 21:7, pp. 2257–2276. DOI: 10.5194/nhess-21-2257-2021.

Traub, J, Z Kaoudi, JA Quiané-Ruiz, and V Markl (Mar. 10, 2021). "Agora: Bringing Together Datasets, Algorithms, Models and More in a Unified Ecosystem [Vision]". *SIGMOD Rec.* 49:4, pp. 6–11. DOI: 10.1145/3456859.3456861.

Tuia, D et al. (2025). "Artificial Intelligence to Advance Earth Observation: : A Review of Models, Recent Trends, and Pathways Forward". *IEEE Geoscience and Remote Sensing Magazine*, pp. 2–25. DOI: 10.1109/MGRS.2024.3425961.

Udo, K and Y Takeda (June 2017). "Projections of Future Beach Loss in Japan Due to Sea-Level Rise and Uncertainties in Projected Beach Loss". *Coastal Engineering Journal* 59:2, pp. 1740006-1-1740006–16. DOI: 10.1142/S057856341740006X.

Vacchi, M, K M Joyse, R E Kopp, N Marriner, D Kaniewski, and A Rovere (June 29, 2021). "Climate Pacing of Millennial Sea-Level Change Variability in the Central and Western Mediterranean". *Nature Communications* 12:1, p. 4013. DOI: 10.1038/s41467-021-24250-1.

Vafeidis, A T et al. (July 2008). "A New Global Coastal Database for Impact and Vulnerability Analysis to Sea-Level Rise". *Journal of Coastal Research* 244, pp. 917–924. DOI: 10.2112/06-0725.1.

Vaswani, A et al. (Aug. 2, 2023). *Attention Is All You Need*. DOI: 10.48550/arXiv.1706.03762. Pre-published.

Vitousek, S, P L Barnard, and P Limber (2017). "Can Beaches Survive Climate Change?" *Journal of Geophysical Research: Earth Surface* 122:4, pp. 1060–1067. DOI: 10.1002/2017JF004308.

Vitousek, S, D Buscombe, K Vos, P L Barnard, A C Ritchie, and J A Warrick (Jan. 2023a). "The Future of Coastal Monitoring through Satellite Remote Sensing". *Cambridge Prisms: Coastal Futures* 1, e10. DOI: 10.1017/cft.2022.4.

Vitousek, S, D Buscombe, K Vos, P L Barnard, A Ritchie, and J Warrick (Nov. 28, 2022). "The Future of Coastal Monitoring through Satellite Remote Sensing". *Cambridge Prisms: Coastal Futures*, pp. 1–43. DOI: 10.1017/cft.2022.4.

Vitousek, S, K Vos, K D Splinter, L Erikson, and P L Barnard (2023b). "A Model Integrating Satellite-Derived Shoreline Observations for Predicting Fine-Scale Shoreline Response to Waves and Sea-Level Rise Across Large Coastal Regions". *Journal of Geophysical Research: Earth Surface* 128:7, e2022JF006936. DOI: 10.1029/2022JF006936.

Vos, K et al. (Sept. 29, 2023a). "Benchmarking Satellite-Derived Shoreline Mapping Algorithms". *Communications Earth & Environment* 4:1, 1, pp. 1–17. DOI: 10.1038/s43247-023-01001-2.

Vos, K, W Deng, M D Harley, I L Turner, and K D Splinter (Nov. 18, 2021). *Beach-Face Slope Dataset for Australia*. preprint. Geosciences – Marine Geology. DOI: 10.5194/essd-2021-388.

Vos, K, M D Harley, I L Turner, and K D Splinter (Feb. 2023b). "Pacific Shoreline Erosion and Accretion Patterns Controlled by El Niño/Southern Oscillation". *Nature Geoscience* 16:2, pp. 140–146. DOI: 10.1038/s41561-022-01117-8.

Vos, K, K D Splinter, M D Harley, J A Simmons, and I L Turner (Dec. 1, 2019). "CoastSat: A Google Earth Engine-enabled Python Toolkit to Extract Shorelines from Publicly Available Satellite Imagery". *Environmental Modelling & Software* 122, p. 104528. DOI: 10.1016/j.envsoft.2019.104528.

Vousdoukas, M I et al. (Nov. 2020a). "Reply to: Sandy Beaches Can Survive Sea-Level Rise". *Nature Climate Change* 10:11, pp. 996–997. DOI: 10.1038/s41558-020-00935-1.

– (Mar. 2020b). "Sandy Coastlines under Threat of Erosion". *Nature Climate Change* 10:3, pp. 260–263. DOI: 10.1038/s41558-020-0697-0.

Walkden, M and M Dickson (May 2008). "Equilibrium Erosion of Soft Rock Shores with a Shallow or Absent Beach under Increased Sea Level Rise". *Marine Geology* 251:1–2, pp. 75–84. DOI: 10.1016/j.margeo.2008.02.003.

Warrick, J A, K Vos, D Buscombe, A C Ritchie, and J A Curtis (2023). "A Large Sediment Accretion Wave Along a Northern California Littoral Cell". *Journal of Geophysical Research: Earth Surface* 128:7, e2023JF007135. DOI: 10.1029/2023JF007135.

Wehn, U et al. (Sept. 2021). "Impact Assessment of Citizen Science: State of the Art and Guiding Principles for a Consolidated Approach". *Sustainability Science* 16:5, pp. 1683–1699. DOI: 10.1007/s11625-021-00959-2.

Wilkinson, M D et al. (Dec. 2016). "The FAIR Guiding Principles for Scientific Data Management and Stewardship". *Sci Data* 3:1, p. 160018. DOI: 10.1038/sdata.2016.18.

Wong, P P et al. (2014). "Coastal Systems and Low-Lying Areas". *Climate change* 2104, pp. 361–409.

Woodroffe, C D (2002). *Coasts: Form, Process and Evolution*. Cambridge University Press.

Woodroffe, C D and C V Murray-Wallace (Oct. 2012). "Sea-Level Rise and Coastal Change: The Past as a Guide to the Future". *Quaternary Science Reviews* 54, pp. 4–11. DOI: 10.1016/j.quascirev.2012.05.009.

Wright, L and B Thom (Oct. 1, 1977). "Coastal Depositional Landforms: A Morphodynamic Approach". *Progress in Physical Geography: Earth and Environment* 1:3, pp. 412–459. DOI: 10.1177/030913337700100302.

Wright, L D and A D Short (1984). "Morphodynamic Variability of Surf Zones and Beaches: A Synthesis". *Marine geology* 56:1–4, pp. 93–118.

Wright, L D, J P M Syvitski, and C R Nichols (2019). "Coastal Systems in the Anthropocene". In: *Tomorrow's Coasts: Complex and Impermanent*. Ed. by L D Wright and C R Nichols. Vol. 27. Springer International Publishing, Cham, pp. 85–99. DOI: 10.1007/978-3-319-75453-6_6.

Wulder, M A, J G Masek, W B Cohen, T R Loveland, and C E Woodcock (July 1, 2012). "Opening the Archive: How Free Data Has Enabled the Science and Monitoring Promise of Landsat". *Remote Sensing of Environment*. Landsat Legacy Special Issue 122, pp. 2–10. DOI: 10.1016/j.rse.2012.01.010.

Wulder, M A et al. (Oct. 1, 2022). "Fifty Years of Landsat Science and Impacts". *Remote Sensing of Environment* 280, p. 113195. DOI: 10.1016/j.rse.2022.113195.

Young, A P and J E Carilli (May 2019). "Global Distribution of Coastal Cliffs". *Earth Surface Processes and Landforms* 44:6, pp. 1309–1316. DOI: 10.1002/esp.4574.

Zha, D et al. (May 31, 2025). "Data-Centric Artificial Intelligence: A Survey". *ACM Computing Surveys* 57:5, pp. 1–42. DOI: 10.1145/3711118.

Zhang, K, B C Douglas, and S P Leatherman (May 1, 2004). "Global Warming and Coastal Erosion". *Climatic Change* 64:1, p. 41. DOI: 10.1023/B:CLIM.0000024690.32682.48.

Zhu, X X et al. (Dec. 2017). "Deep Learning in Remote Sensing: A Comprehensive Review and List of Resources". *IEEE Geoscience and Remote Sensing Magazine* 5:4, pp. 8–36. DOI: 10.1109/MGRS.2017.2762307.

Zhu, Y et al. (Jan. 19, 2024). *Large Language Models for Information Retrieval: A Survey*. DOI: 10.48550/arXiv.2308.07107. Pre-published.