A global open-source database of flood-protection levees on river deltas (openDELvE)

~~Joey O'Dell[1],~~ Jaap H. Nienhuis[1], Jana R. Cox[1], Joey O'Dell[1], Douglas A. Edmonds[2], and Paolo Scussolini[3]

[1] Department of Physical Geography, Universiteit Utrecht, Postbus 80.115, 3508 TC Utrecht, Netherlands
[2] Department of Earth and Atmospheric Sciences, Indiana University Bloomington, 1001 East 10th Street, Bloomington, IN 47405-1405, United States of America
[3] Institute for Environmental Studies, Vrije Universiteit Amsterdam, De Boelelaan 1111, 1081 HV Amsterdam, Netherlands

Correspondence to: ~~Joey O'Dell~~Jaap H. Nienhuis (~~data~~j.h.nienhuis@uu.nl~~opendelve.eu~~)

Abstract. Flood-protection levees have been built along rivers and coastlines globally. Current datasets, however, are generally confined to territorial boundaries (national datasets) and are not always easily accessible, posing limitations for hydrologic models and assessments of flood hazard. Here, we bridge this knowledge gap by collecting and standardising global flood-protection levee data for river deltas ~~present our work to develop a single,~~into the **open**-source global river **del**ta **lev**ee data **e**nvironment, ~~(~~openDELvE~~) which aims to bridge a data deficiency by collecting and standardising global flood-protection levee data for river deltas~~. In openDELvE, we ~~have~~ aggregated levee data from national databases, ~~as well as data stored in~~ reports, maps, and satellite imagery. The database identifies the river delta land areas that the levees have been designed to protect~~, and w~~. Where ~~additional relevant~~ data ~~is~~are available, we record the extent and design specifications of the levees themselves (e.g., levee height, crest width, construction material) in a harmonised format. The 1,657~~01~~ polygons of openDELvE ~~currently~~ contain~~s~~ 19,248 ~~5,089~~ km of levees~~ on deltas,~~ and 44,733.505 km^2 of leveed area~~ in 1,601 polygons~~. For the 153~~2~~ deltas included in openDELvE, ~~on average 19~~17~~8~~% of ~~their habitable~~their land area is confined by ~~verifiable~~ flood-protection levees. ~~Globally, we estimate that between 56% and 5417% of all delta land is confined by flood-protection levees.~~ They are densely populated. Around 26~~0~~% of delta population lives within the 17~~8~~% of delta area that is protected. openDELvE data can help improve flood exposure assessments, which currently do not account for flood-protection levees. ~~We estimate that e~~We find that current flood hazard assessments ~~that do not consider on levees~~may exaggerate ~~flood exposed~~ the delta flood exposure~~exposed to coastal floods~~ by 33~~60~~% on average, but up to 100% for some deltas. ~~Inclusion of levee data in hazard modelling can be important in risk prediction and management (including population patterns, flood protection and land loss).~~ ~~openDELvEThe data is aligned to the recent standards of Findability, Accessibility, Interoperability and Reuse of scientific data (FAIR) and is open-source to maximize use.~~ openDELvE is made public on an interactive platform (www.opendelve.eu), which includes a community-driven revision tool to encourage inclusion of new levee data and continuous improvement and refinement of open-source levee data.

1 Introduction

1.1 What are levees and what do they do?

Levees are banks of sediment or artificial material that prevent water from entering areas where it is not desirable. They are common in delta plains and protect their populations and ~~urban areas~~assets ~~, including in floodplains,~~ from water level fluctuations of rivers and the sea. Levees have been constructed to mitigate flood risk and direct water flows throughout human civilisation. Recorded building of levees along the River Nile in ~~(~~Egypt~~)~~ began around 4600 BP (Westermann, 1919) which indicates the innate link between the settlement of coastal populations and the development of levees. Modern materials and engineering concepts have altered the overall appearance and effectiveness of levees, but the basic principle has remained~~s~~ the same for millennia.

Levees can also also have negative environmental consequences. They alter sediment transport and sedimentation patterns, as sediment deposition behind levees is usually reduced. Areas protected by levees can subside relative to the surrounding (Middelkoop et al., 2010), resulting in increased risk of coastal and river flooding in the longer term (Pinter et al., 2008; Criss and Shock, 2001; Pinter, 2005; Munoz et al., 2018). In particular, dDeltasleveed deltas in particular are at risk to be locked-in (Santos and Dekker, 2020), as areas become sediment- -starved and cease to keep up with sea level rise (Pinter et al., 2016). Another example of the negative effect of levees is in Australia, where undocumented private levees, intending to protect land, resulted in degradation of the floodplain ecosystem, and contributed to flash flood risk by disconnecting the floodplain and channel (Steinfeld et al., 2013).

Because of the negative consequences, contemporary fluvial river and flood management measures/projects scenarios therefore often prioritise nNature-based sSolutions that limit the need for levees (Esteves, 2014; Cohen-Shacham et al., 2016; Van Wesenbeeck et al., 2014). In deltas in particular, levees are sometimes removed and sS, but this is not always possible. edimentation--eEnhancing sStrategies (SES) (Cox et al., 2022) are pursued to restore natural delta functions, but this may not always be possible.

1.2 Why (data of) levees matter

Data on levees areis important, especially for river deltas. People living in river deltas face mounting threats: they are disproportionately affected by coastal flooding and relative sea level rise (Edmonds et al., 2020) and rely on a river sediment supply that is ever diminishing in many places (Dunn et al., 2019). Data on levees can help to assess these threats.

Levees themselves are not new creations, and so the majority of data that references their locations and standards is historical and locked away in paper form (maps, plans etc.). However their effect on the human population can be easily seen. Modern urbanized deltas tend to be heavily embanked by levees because high population densities have demanded protection against river and coastal flooding. Despite these levees, people living in coastal deltas face mounting threats:; they are disproportionately affected by coastal flooding and relative sea level rise (Edmonds et al., 2020) and rely on a river sediment supply that is ever diminishing river sediment supply (Dunn et al., 2019). Data on levees can help to assess these threats.

This is then further compounded by the "levee effect", defined by Gilbert White in 1947 whereby increased levee building creates a false sense of security which can often come at the effect of land use decisions (Hutton et al., 2019) and is particularly well documented in New Orelans both in the lead up to Hurricane Katrina and in the multiple waves of reconstruction post-disaster (Kates et al., 2006). Therefore the demand for levees receives a certain public sentiment through the apparent afforded protection and not the actual standard of protection offered in the area.(Hersher, 2018).

Mapping the presence of levees presence is useful for hydrologic and hydrodynamic modelling. Such models often aimare used to predict inundation during high water levels in rivers or in the sea, discharge events and help active management of risk and hazard to life. Models are also used to design levees by simulating a specific flood return period or flood scenario. B, but large-scale modelling without accurate levee locations results in modelling inaccuracies (Fleischmann et al., 2019). Levees can be incorporated in detailed models (e.g., HEC-RAS, US Army Corps of Engineers, 2020; or Delft3D, Lesser et al., 2004) as a geometric feature within an initial surface topography. For models on larger scales, levees are too small to be included directly and are sometimes presented as a sub-grid feature or through a flood-attenuation proxy (Sampson et al., 2015).

OpposinglyConversely, models are also used to design levees by matching a specific return period or flood scenario, and so modelling data accuracy might also have has a direct impact on the suitability of the height of the levee.

Data on levees can help to better understand human-landscape interactions (Werner and McNamara, 2007). One of these interactions is the so-called “levee effect”, defined by Gilbert White in 1947, whereby ~~increased~~ levee building creates an ~~false~~excessive sense of security which leads to increased development and increased flood exposure (Hutton et al., 2019). This effect of levees is thought to contribute to ~~the vulnerability~~larger exposure to low-probability floods in~~of~~ delta cities. New Orleans after Hurricane Katrina is an example (Kates et al., 2006). Data on levees can help to assess ~~Levees are deployed~~the co-evolution of levees and development prior to, and ~~, (re-)engineered, and altered~~ in response to ~~threats~~disasters, ~~—~~ both present and future, and better understand the levee effect ~~— so their dynamics are challenging to simulate, but remain of interest to the scientific and public communities~~ (Di Baldassarre et al., 2018).

Levee data can also help studies on levee failures, which are a globally significant source of flood risk. Özer et al. (2019) have developed the International Levee Performance Database (ILPD), ~~focussed on assimilating~~presenting data on levee testing and failure events in an ~~whilst also producing an~~ interactive and ~~queryable~~ searchable interface. Levee data for hazard assessment purposes is additionally useful outside ~~of~~ the realm of ~~pure~~ geophysical modelling, and is core to civil engineering and emergency response management for levee performance, such as during the safety and risk calculation of hurricanes (Mitchell et al., 2013). Data can also be relevant for large-scale studies into the effects and costs of levees, and in their comparison ~~to~~with alternative flood-risk reduction strategies in these areas (Ibáñez et al., 2014; Scussolini et al., 2017; Vuik et al., 2019; Cox et al., 2022). The insurance industry, local residents, and homeowners are additional users of levee data and modelling outputs ~~(National Research Council, 2013, p.68 Box 5-1) for~~that may help with their hazard and risk assessments (National Research Council, 2013). ~~, where this data is usually commercially sensitive and not shared, whereas in the case of the US National Levee Database, open data being made available to all prevents this from being a hidden factor.~~

1.3 A (data) gap in levees

Despite the potential use of levee data, locations and characteristics of levees are often poorly documented (Scussolini et al., 2016; Özer et al., 2019), resulting in inaccuracies and challenges for flood risk modelling (Sampson et al., 2015; Trigg et al., 2016; Winsemius et al., 2016; Dullaart et al., 2021), hazard modelling (Di Baldassarre et al., 2009), and ~~projections of delta land loss from sea-level rise~~sea-level rise impact modelling (Nienhuis and van de Wal, 2021). Accurate models require data input about levees including their spatial extent, protected area, and basic attributes, which currently does not exist in a coherent and harmonised single geospatial data format.

~~However their effect on the human population can be easily seen. Despite these levees,~~

Levees themselves are not new creations, and so ~~the majority of~~most data that references their locations and standards is historical and recorded in paper form (maps, plans etc.).

It is typically governments and municipal organisations who plan and construct levees~~, as safe and dry places to live are key to maintaining public health and wellbeing~~. These institutions (e.g., USACE) also maintain them as part of their daily operations and produce maps and datasets about their design, operation, and failure. This gives a plethora of data such as reports and design specifications, which allows for accurate data gathering and collection processes without the need for in-person observation (e.g., USACE National Levee Database, levees.sec.usace.army.mil). Generally, this results in good quality central national databases, sometimes complemented by ~~with~~ higher resolution ~~localised~~ regional variants (e.g., New South

Wales' Distinctive Land Surface Dataset, Australia), but that do not extend past administrative borders. Data availability can also be publicly restricted.

Poor data on levee existence and levee properties have made it such that their presence is often disregarded in global flood modelling (Trigg et al., 2016) and global delta modelling (Nienhuis et al., 2020). The lack of levee data results in suboptimal modelling results (Fleischmann et al., 2019). The WRI AQUEDUCT Global Flood Analyzer is an example (https://www.wri.org/data/aqueduct-global-flood-analyzer). It provides exceptional global-level flood hazard data but does not include levees and results in overpredicted flood exposure for heavily leveed areas such as the Netherlands.

While specific aspects of levee failure have been documented and aggregated globally (i.e., Özer et al., 2019), we are not aware of any open-source approaches that collect, harmonise and attribute information on levee extent. The lack of global registration of levees complicates flood management efforts. As an alternative, FLOPROS (Scussolini et al., 2016) presents a global dataset on existing and policy-level flood protection standards. This implicitly includes the flood protection offered by levees, but does not include data on levees. Other approaches exist that use (semi-)automated algorithms to locate and specify levees from LIDAR data (e.g. Steinfeld et al., 2013; Wing et al., 2019) but this is restricted by data availability and is not yet possible globally. A levee database can help inform those algorithms and provide validation and calibration data. Besides the registration of their existence of levees, communication and awareness of this information is important, to enable the above-listed uses of levee information.

The lack of global registration, and therefore data, of these levees further complicates flood management efforts., Besides the however just registration of their existence of levees,cannot alone be considered a solution to the problem, and indeed we understand that communication and awarenessthis information can be just as is important, to enable the above-listed uses of levee information in the whole-system oversight and evaluation of change. We recognise that attempts exist to document and aggregate sWhile specific aspects of levee failure have been documented and aggregated globally (i.e., Özer et al., 2019), but we are not aware of any competing open-source approaches that collect, and harmonise and attribute information on levee extent and attribute data.

1.4 Levees in hydrologic and hydrodynamic models

Levees, by design, affect the flow of water. They can be incorporated in detailed models (e.g., HEC-RAS [US Army Corps of Engineers, 2020], or Delft3D [https://oss.deltares.nl/web/delft3d]) as a geometric feature against initial surface topography. For models on larger scales, levees are too small to be included directly and are sometimes presented as a sub-grid feature or through a flood-attenuation proxy (Sampson et al., 2015). In both cases, poor data on levee existence and levee properties have made it such that their presence is often disregarded in global flood modelling (Trigg et al., 2016) and global delta modelling (Nienhuis et al., 2020). The lack of levee data (which change and control water and sediment discharge) results in suboptimal modelling scenarios, such as the WRI AQUEDUCT Global Fflood-risk tool Analyzer (https://www.wri.org/applicaitons/aqueduct/floods/), which provides exceptional global-level flood hazard data but does not include levees and results in abstract scenarios for heavily leveed areas such as the Netherlands.

As an alternative to global levee data, FLOPROS (Scussolini et al., 2016) presents a global dataset on existing and policy-level flood protection standards, which implicitly includes also the flood protection offered by levees, but does not include data on levees. FLOPROS provides uniform, global coverage, however individual feature level data is omitted. Other approaches exist that use (semi-)automated algorithms to locate and specify levees from LIDAR data (e.g. Steinfeld et al.,

~~2013; Wing et al., 2019) but these are generally focussed on specific problem definitions and lack global applicability. A global levee database can help inform those algorithms and provide validation and calibration data.~~

1.4~~5~~ Objective

The objective of openDEL~~l~~vE is to provide an ~~attestable~~ source of delta levee protection ~~delta~~data, for both primary use in flood and hazard modelling, as well as secondary community use through increased data availability by publishing the data on a public website (http://www.opendelve.eu) following standard data types. openDELvE includes links to original data sources, as well as~~, and~~ a user-led amendment reporting function. Examples are also given ~~as to how~~of openDELvE use for~~can be applied in~~ hazard modelling and delta modelling improvements.

2 Methods

2.1 Overview

openDELvE is a collection of existing data on levees and protection features on deltas. We have collected data from vector, raster, and documentary sources. This resulted~~s~~ in two geospatial layers – one for levees, and one for leveed areas~~, and one for leveed lines~~ ~~—~~ and a supporting index dataset, linked to the respective delta by a unique identifier and cross-mapped to the river delta dataset of Edmonds et al. (~~,~~2020). Our methods allow for replicable tracing, processing, assimilation, and display of the data. By storing individual level references and data quality, we aim to provide data that is open and transparent. Our work is underpinned by the principles of FAIR science to support reuse by producing data that is Findable, Accessible, Interoperable, and Reusable (Wilkinson et al., 2016). openDELvE development followed these steps: data definition (Sect. 2.2), data collection (2.3), data processing (~~Sect.~~ 2.4), data attribution (Sect. 2.5), data management (Sect. 2.6), and data assurance (Sect. 2.7).

2.2 Data definition

We followed our definition of levees from Sect. 1.1. Levees exist along coasts and rivers globally, but the scope of openDELvE is limited to river deltas (Sect. 2.4.1). We made use of a database of deltaic locations and deltaic area extent by Caldwell et al. (~~Caldwell et al.,~~ 2019) and Edmonds et al. (~~Edmonds et al.,~~ 2020). We further limited ourselves to only storing information on defences that are permanent features, and not temporary/reactive measures. ~~Temporary measures, such as s~~Sandbags and hoardings deployed for flash flooding or imminent but irregular flood issues ~~are not temporally constant~~are temporary, and so are usually not mapped, nor were considered for inclusion in this database.

openDELvE is designed to represent levees as geospatially explicit vector data: lines and polygons. For source data that exists in reports on maps and technical drawings levee presence is often reduced to a raster map element, and so needed to be sufficiently georeferenced and assessed for quality. However, this is still a valid data source and is included in our process. We consider the age, source document, and data quality as we recognise that data may be reworked and requoted a number of times in its lifespan.

openDELvE consists of three data elements: an index table and two vector layers (Table 1), each with a set of standardised attributes (Table 2).

Table 1: Data entities in the live viewing environment and their exported file types as in the research data store

Data Entity	Type	Exported Elements	Purpose
Delta Index	Table	CSV	Contains data decision logs and linking ~~characteristics~~ levees at delta level
Leveed Area	Polygon	SHP, KML	A vector layer containing polygons of the areas protected by levees
Levee Lines	Line	SHP, KML	A vector layer containing lines of the levees and including standardised attributes

~~**Table 1: Data entities in the live viewing environment and their exported file types as in the research data store**~~

Table 2: openDELvE attributes for the three data elements (as per Table 1). Conversion factors and mapping of fields are~~is~~ given in Supplementary Table S2.

Data Entity	Attribute	Purpose
Delta Index	FriendlyName	Name of the delta, if known
	Status	Processed, No Result, Pending, or Not Processed (as per Sect. 2.3)
	PolygonID	Delta ID following Edmonds et al. ~~,~~(2020)
	ISO_2	2-digit code identifying the country where the majority of the delta lies, following ISO 3166-1:2020 alpha-2
	Journal	A timestamped text log of activity at a delta level
	MainRefAPA7	Literature reference for the overall source material for the delta, formatted in APA 7th Edition, if available
	MainRefDOI	Digital Object Identifier for the source material, if available
	NeedsReview	Boolean indicator of requirement for later review of delta
	LastChkDate	Date field signalling last check date of the delta
	LastChkBy	Two-character identifier of the last user who updated the dataset
Leveed Area	NAME	Name of the leveed area feature from the source dataset, if available
	REFERENCE	The identifier for the feature from the source dataset, if available
	DOI	Digital Object Identifier for the source material, if available
	URL	Uniform Resource Locator (web link) for the source material, if available
	LITREF	Literature reference for the source material, formatted in APA 7th Edition
	PolygonID	Delta ID following Edmonds et al. ~~,~~(2020)
	DataQuality	Data quality classification (following Table 3)
Levee Lines	NAME	The name for the feature from the source dataset, if available
	REFERENCE	The identifier for the feature from the source dataset, if available
	DOI	Digital Object Identifier for the source material, if available
	URL	Uniform Resource Locator (web link) for the source material, if available
	LITREF	Literature reference for the source material, formatted in APA 7th Edition
	DefenceLength	The length of the levee feature, as provided in the source dataset, if available (metres)
	DefenceHeight	The height of the levee feature, as provided in the source dataset, if available (metres)
	DefenceWidth	The width of the levee feature, as provided in the source dataset, if available (metres)
	FoundationWidth	The width of the levee foundation, as provided in the source dataset, if available (metres)
	Construction	The primary material that the levee is composed of
	ClassType	Construction or formation type of the feature
	CutoffMaterial	The material that the levee cutoff is composed of
	DesignStandard	Design storm rating of the feature (1/n, decimal)

	DataQuality	Data quality classification (following Table 3)
	PolygonID	Delta ID following Edmonds et al. ~~,~~ (2020)

~~**Table 2: openDELvE attributes for the three data elements (as per Table 1). Conversion factors and mapping of fields is given in Supplementary Table S2.**~~

Levee data in openDELvE include a data quality class and a direct link to the source dataset. We devised the data quality criteria included in Table 3:

Table 3: Data quality definition for levees based upon data provenance, both for use in initial data classification and ongoing maintenance. Criteria are exclusively applied: ~~, and~~ all categories must be met to meet ~~that~~ a certain classification.

Class	**Criteria**
A (Excellent)	Vector data First-order data source (i.e., scientific papers, governmental geospatial data, original publication) Spatially complete[a] (with respect to geopolitical boundaries) Existence verifiable with satellite imagery
B (Good)	Raster data (suitably georeferenced, little to no variance) First-order or re-cited/modified (original accessible) but published within a scientific or government publication Existence verifiable with satellite imagery
C (Acceptable)	Raster data (loosely georeferenced, variance due to old base map or similar) Conjectural or non-scientific source (ex: newspaper) Source >20 years of age, regardless of type Existence (partially[b]) verifiable with satellite imagery
X[c] (Invalid)	Data inaccessible (blocked, hidden, unpublished) Irrecoverable issues with data quality Could not confirm existence of data from other sources using satellite imagery with resolution ≤25m Temporary or reactive measures only (ex: sandbags)

~~**Table 3: Data quality definition for levees based upon data provenance, both for use in initial data classification and ongoing maintenance. Criteria are exclusively applied, and all categories must be met to meet that classification.**~~

[a]Data that were attributable to class X have not been included in the published dataset but are documented in the delta index

[b]We included 'partially verifiable' due to incident patchy local coverage of openly accessible satellite data, as there are instances where sufficient high-resolution imagery was not accessible, but standard-resolution imagery indicated the presence of the feature that was elsewhere published.

[c]Spatially complete was defined as being of the entire levee run, which may be comprised of several subsection maps.

2.3 Data collection

We conducted extensive literature searches using a variety of web searching platforms (i.e., Clarivate Web of Science, Google Search, Google Scholar, OCLC WorldCat) as well as data aggregation platforms (e.g., re3data.org, DataCite, data.gov.uk, data.gov, data.gov.au). Data was collected in a search process that is documented as a log with diary-style entries in the Delta Index table (see Table 1) and recorded at a delta level. Sources for each individual levee are stored at the feature level. This allowed us to record rationale and decision-making process so that both viewers and onward developers of the dataset are aware of the steps taken and explanations for decisions taken in data hand.

With an international scope, searching often required country or location-specific terms (e.g., '*tanggul*' meaning levee or embankment in Indonesian) to aid data discovery, and these were regionally supplemented along with a vocabulary of common delta and levee terms when using academic paper and internet indexing services.

Funding reports from the World Bank projects on flood defence activities has also contributed to the database. Financing documents often contain maps and so we include data from the World Bank where it was discovered in our searches, released publicly, had been reviewed, and contained levee feature level data.

When it was not possible to find data in areas where levees were expected, the place was identified by name using the address search (*gazetteer*) function in ArcGIS and then basic internet searching was performed to find reports of floods or sea-~~-~~level rise related damage. Finally, we made use of the world satellite imagery layer within ArcGIS to review areas where levee source data was inaccessible, and assess by visual means whether it was likely levees were present. We verified areas that we believe may be uninhabited ~~areas~~ using this imagery and classified them accordingly, where satellite imagery confirmed no visible levees, the delta was set to No Result. If levees were visible but we could not verify them with alternative data sources, we set the delta to 'Pending' where external enquiries were taking place and the relevant note was entered in the ~~Journal~~ArcGIS journal (see Table 2). We identify deltas as 'Not Processed' if we have yet to manually review available sources, and no national vector dataset was discoverable for processing via our automated tool.

Many deltas in the delta dataset may be small and uninhabited (Edmonds et al., 2020), have inaccessible data, or have data that we were unable to convert into a format that we could add to the database. We collectively group these deltas as having "No result" in terms of data collection. Note that this does not always mean there is no data. For example, data from the Database nazionale della AgriNature in TErra (DANTE, *formerly known as: ItaliaN LEvee Database [INLED]*) (Barbetta et al., 2015) was not suitable for processing because it only contains a levee start and end point coordinate. We classified~~y~~ these deltas under "No result" because it requires access to a detailed regional-level watercourses database and high-resolution DEM so that an interpretational algorithm could be trained to infer the levee course.

Where available, we included levee attributes (e.g., design storm, wall height, levee material, Table 2). This can inform modelling and therefore work as a stand-alone spatial tool for investigating river delta dynamics. Additionally, the data layers can be used for verification of deductive models for the detection of levees by other means, including LIDAR and remotely sensed data as well as corroborating other data sources, such as OpenStreetMap. As we intend for the database to be globally comparable, we set up a cross matching list (Supp. Table S2) within the project documentation to ensure that the attributes of the levee lines layer were consistent between sources and languages. This was then used for both manual and automated input so that different units of measure, classifications of levee and construction type, and key engineering data were harmonious.

2.4 Data processing

2.4.1 Vector data processing

Where data ~~was~~ were sourced in vector format, we defined a data processing algorithm in the ArcGIS® Model Builder (Supp. Fig. S1) to clip the imported data to the extent of river deltas from Edmonds et al. (2020) with a 100 km 'buffer zone'. This buffer zone is included to maximize OpenDELvE data ~~usability~~usability, but it does not affect reported statistics on delta coverage:~~.~~ ~~A~~all reported data statistics in this paper are for levees strictly within delta boundaries (Figure~~.~~ 1), although these can include shallow marine portions of the delta front as well as upland area (Figure 2, Edmonds et al., 2020). The buffer zone is included to allow extended use of the dataset for upstream fluvial and sediment transport modelling and additionally, should the dataset of Edmonds et al. (2020) be updated, reduces the likelihood that levees are missed from the layer.

The ArcGIS® Model Builder automated import process ~~created~~ is distributed with the dataset so that data can be repeatedly processed and added to the database both now and in the future. We supplemented this by the creation of conversion tables *(Supplementary Table S2)* so that levee attributes, where available, are comparable at a global scale.

2.4.2 Non-vector data processing

We performed georeferencing of levee maps~~/~~ ~~and documentary data~~ where the location was visible using a second, ~~contemporary~~ georeferenced, map and the map could be referenced in ~~less~~ fewer than 5 reference points. This ensured that we were not extensively distorting the source map and therefore it was possible for us to trace in the features as accurately as possible. Where no georeferencing within 5 reference points was possible, or where the map had too few defining features to be georeferenced at all (e.g. map created with too few topographical features, substantial engineered or geological change resulted in difference between map and modern day situation) then the appropriate data quality class (X) was assigned. T~~, and where the map was impossible to suitably georeference, t~~he data source was set aside and the process was documented in the log. Furthermore, where aerial photography was analysed, we defined a set protocol for the inference of leveed area (Supp. Fig. S3).

Data in the "Levee Lines" layer is currently limited to vector levee data sources and does not exist for raster data sources. Ongoing work includes manual review and development of (semi)automated processing steps to retrieve levee lines from raster sources.

2.4.3 Extraction of leveed-areas from levee line information ~~Line data interpretation~~

~~A number of~~Several data sources were processed where only levee lines ~~line data (lines) was made~~ are available, and not levee-protected areas (polygons). In these cases, we estimated levee-protected areas from levees by: (1) manually selecting levees that are not separated by water bodies, and (2) constructing an area confined by these levees (Figure 1). We manually reviewed this process using~~This process was initially developed by manually reviewing~~ datasets that have both ~~existing~~ levee lines and leveed areas ~~datasets~~ (e.g., USACE National Levee Database) and did not result in a large over- or under estimation of the leveed area (Fig. S3)~~ and as a result we developed the following process developed conceptually and then added into the data handling workflow.~~ ~~, and as such this required a level of intepretation using a set logic to produce replicable results. This process was initially developed by manually reviewing existing levee line and leveed area datasets (e.g. USACE National Levee Database) and as a result we developed the following process developed conceptually and then added into the data handling workflow.~~

~~Our processing of this type of data was entirely manual however used geospatial processing tools (Concave Hull) built into the ArcGIS Pro platform. We consider the opportunity for these steps to be coded into an automated workflow.~~. Leveed area generated from this process instead of original data is indicated in the data quality label.

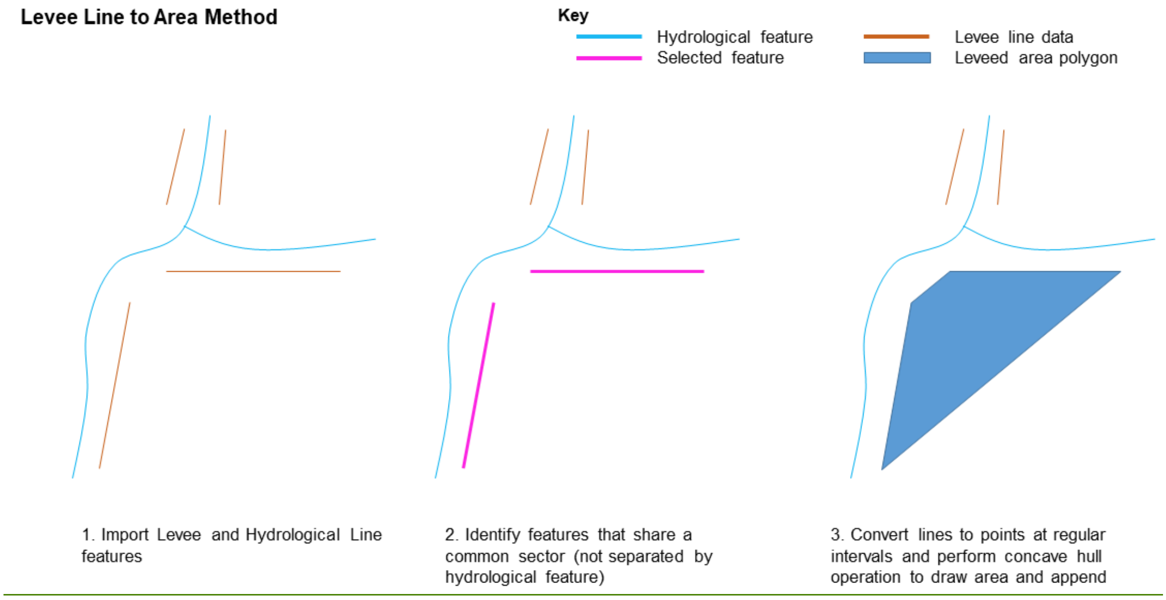

Figure 1: Extraction of leveed-areas from levee line information ~~Interpretation method of levee line data to leveed area polygon~~ (visual representation of process outline in Supplementary Figure S3)

2.5 Data attribution

Every task performed was~~in the journal is~~ recorded in the openDELvE metadata (ArcGIS "journal") for audit purposes, and each entry ~~into the layers~~ is attributed to the data source, including a full literature reference, the source URL, and a DOI (where available). This ensures that we can display this data interactively and that the original source remains permanently available. We also included any digital identifiers from vector datasets so that the individual feature can be tracked and mapped over subsequent data revisions.

We ~~timestamp~~ linked each entry into openDELvE to a delta using the PolygonID from Edmonds et al., (2020), ~~the delta index~~ and additionally flagged deltas that need manual review in the future. ~~This has no effect on data quality, however it~~ It ensures that there is a robust process in the future to signal amendments needed or entries for which ~~where it is apparent that there~~ sources are undocumented or inaccessible~~ data sources available~~. This ~~not only~~ supports ~~local~~ maintenance~~, but also~~ and prevents repetition of previous search activities.

2.6 Data management

The resulting data layers for levee area and levee line feature were created in ArcGIS Pro and hosted on an ArcGIS Online data hub (http://www.opendelve.eu, Figure 3)~~)~~. Additionally, we maintained ongoing research data exports in the DataverseNL environment as the database develops, which also assigns permanent identifiers (DOIs) to the research dataset. Data is stored in three defined entities as per Table 1, and we stored each layer within their own container in the public ArcGIS Online® environment. These layers weare then ~~publicly~~ published to be used as part of the ArcGIS Online Directory and through modern GIS clients via a Web Feature Service (WFS).

The openDELvE platform facilitates an interactive and community driven maintenance of the dataset through an amendment form and additional messages in all metadata files. Suitable new data will be added to openDELvE ~~The project remains actively maintained~~ by the authors at Utrecht University, and made public on the openDELvE webpage and the DataverseNL environment.~~ by assigning permanent identifiers (DOIs) to the research dataset, as well as developing the website alongside, there project remains actively maintained.~~

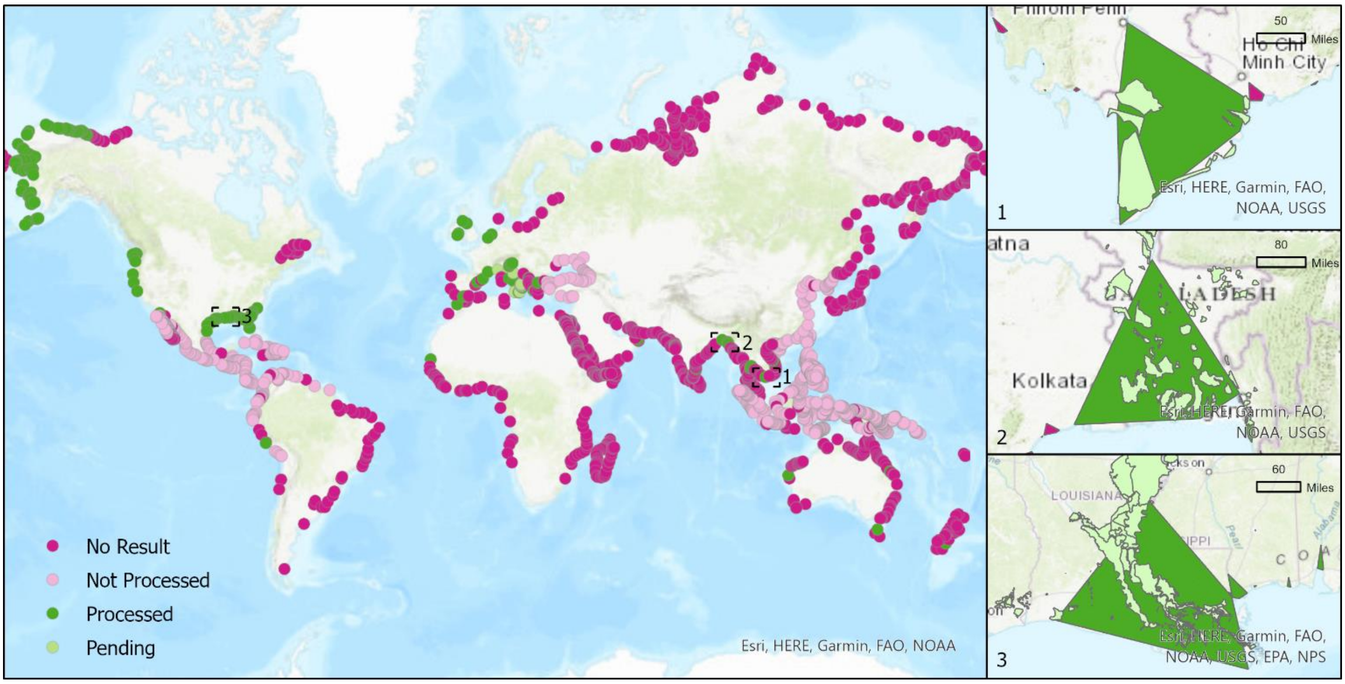

Figure 2~~1~~: Distribution of delta levee dataset completeness and data availability in release of openDELvE (v1.0). Polygons encompass the four-point deltaic extent as defined by Edmonds et al. (2020).

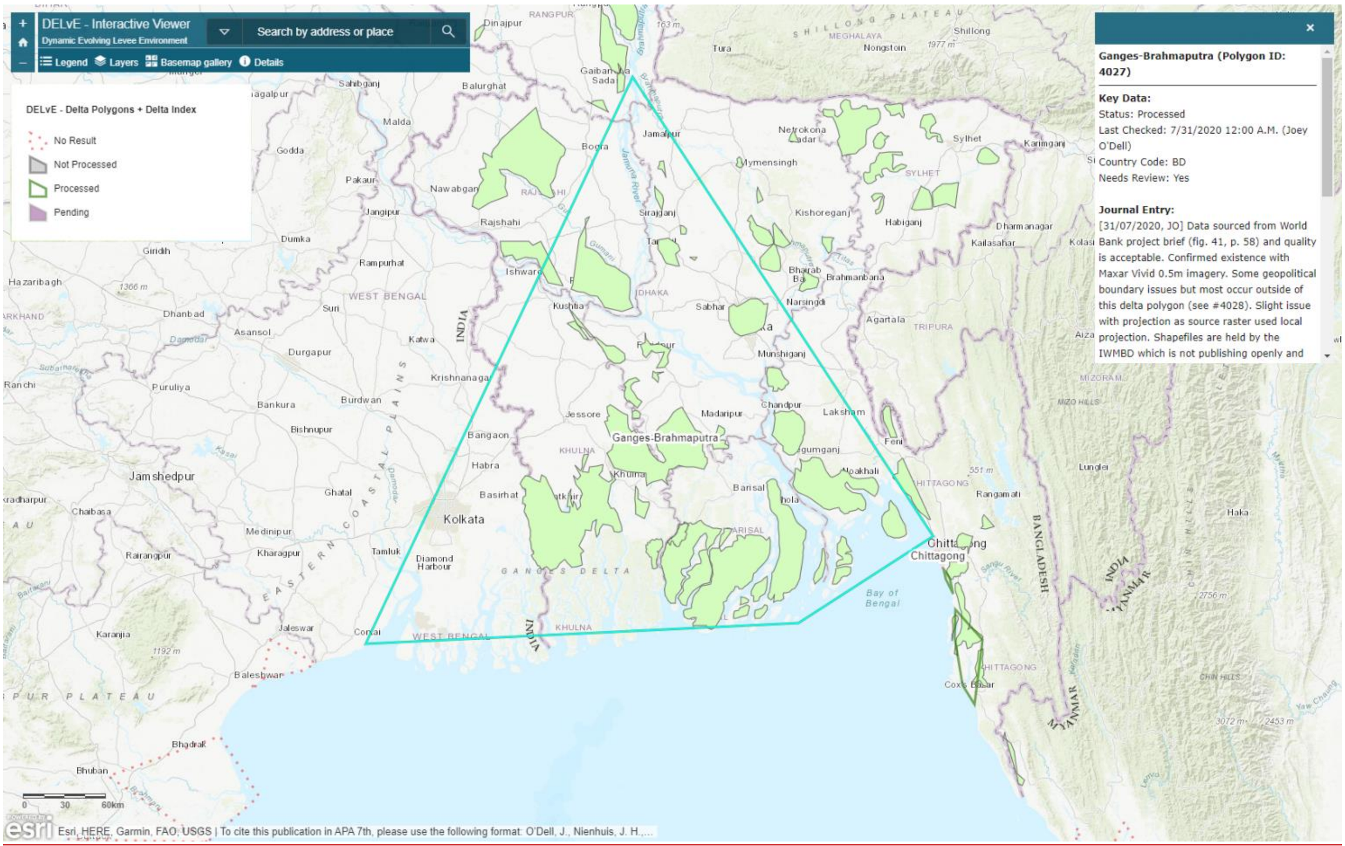

Figure 3: Interactive browsing interface to openDELvE built using the ArcGIS Online platform. Area of focus is the Ganges-Brahmaputra delta, Bangladesh. Data content as per openDELvE version 1.0. Available publicly at: http://opendelve.eu

2.7 Data assurance

Before releasing the dataset, we performed several checks on the data and metadata (Table 4). We then generated metadata compliant with the EU INSPIRE geospatial metadata standard (European Parliament, 2007) using the built~~-~~in ArcGIS® Pro wizard for each data element (Table 2), and for the dataset in its entirety. This included interactive help-text for the model builder~~ GUI~~. We ~~self-validated~~checked ~~-~~the m~~M~~etadata files for completeness using the metadata wizard in the ArcGIS® Pro system.

Table 4: Categories and criteria for the data validation performed on the dataset

Type	**Criteria**
Duplicate Check	There are no duplicate delta polygon IDs (*PolygonID)* in the index
Orphan Check	All linked delta polygon IDs matched a delta polygon in the dataset There were no unsuccessful joins between the data layers
Null Check	Where there was no match to a delta polygon, this returned -1 Where it was not (yet) possible to match the polygon to a delta, this returned null
Visual Check	Visually verify if data appears ~~as should be reasonable to expect~~reasonable (i.e., within 100 km of delta polygon border, within proximity of water feature, of a shape that is coincident to ~~fluvial~~ delta morphology)
Metadata Check	All fields in the ArcGIS® Pro metadata wizard completed

~~**Table 4: Categories and criteria for the data validation performed on the dataset**~~

2.8 Applications of openDELvE

2.8.1 Land-use assessment with Copernicus Global Land Cover Layers~~MODIS IGBP~~

We used th~~NASA~~~~Th~~e ~~-Moderate Resolution Imaging Spectroradiometer (MODIS) International Geosphere-Biosphere Programme (IGBP) land use/land cover data of NASA~~Copernicus global land cover dataset (Buchhorn et al., 2020) ~~(MCD12Q1 2019) (Friedl and Sulla-Menashe, 2019)~~ ~~was used~~ to identify land use types and patterns within ~~the~~ deltas and within leveed areas on deltas ~~processed in openDELvE (~ 150 deltas)~~. Copernicus land cover data separates 16 natural vegetation classes, 4 non vegetated classes, and 2 human-influenced land cover classes, ~~at a~~ on a global 100-m grid ~~with~~. We selected the land cover data from 2019 and calculated for each land use ~~how much delta area~~and for each delta the area that is either ~~differences in land use occurrence between~~ protected from ~~and~~ or exposed to ~~from~~ flooding ~~delta land~~. ~~The MODIS raster data was loaded into ARCGIS Pro for anaylsis~~. ~~Using Zonal Statistics, the majority land-use within and outside the leveed areas was identified within each processed delta polygon. The percentage of all land use types was calculated using Tabulate Area.~~

2.8.2. Population ~~patterns~~density with LandScan™

The Oak Ridge National Laboratory's LandScan™ population data~~distribution~~ ~~raster~~ (https://landscan.ornl.gov/) was used to ~~identify~~ calculate population density~~dynamics~~~~patterns~~ ~~in~~within levees~~s~~ ~~vs~~ and ~~population~~ outside of ~~un~~leveesd ~~delta areas~~. LandScan provides ~~an~~ globally yearly gridded data ~~ambient population~~ at a ~~value~~ 30-arc sec (~1 km) resolution, counting ~~averaged over 24 hours, therefore not just counting~~ resident ~~population but also~~and transitory population ~~(e.g. commuters, students)~~. We used data from 2020 and ~~sum~~calculate delta population within and outside leveed areas ~~and to this end sufficiently de-aggregates census data to provide a truer value of risk and impact. Total population within and outside leveed areas was identified for~~ for each processed delta polygon ~~of openDELvE (~ 150 deltas) using Zonal Statistics~~.

2.8.3 Coastal flooding analysis with COAST-RP

The COAST-RP dataset (COastal dAtaset of Storm Tide Return Periods) of Dullaart et al. (2021) provides spatial extent of coastal floods from storms at 30-arcsec resolution for storm return periods from 1 to 1000 years. This is based on a global hydrodynamic model of the ocean that provides coastal water levels. COAST-RP then propagates these water levels in land using a static inundation model on top of a state-of-the-art global elevation dataset and assuming a water level attenuation factor based on distance. COAST-RP does not include levee data. Here we intersected openDELvE data with COAST-RP to estimate simulated flood extents that might, in reality, be protected from coastal flooding by levees. ~~This dataset is combined with openDELvE~~ We assessed storm return periods of 10, 100, and a 1000 years, but, because of limited levee height data and~~or~~ levee quality data, we have not assessed actual protection but rather potential protection.

~~to indicate how the presence of levees can alter the prediction of coastal flooding. The COAST-RP dataset (COastal dAtaset of Storm Tide Return Periods) of Dullaart et al. (2021) provides spatial extent of floods This of the ocean, and on static inundation model, which includes a state-of-the-art global elevation dataset and which accounts for water level attenuation. and on providesandlevees for of various return periods in deltas. This dataset is combined with openDELvE to indicate how the presence of levees can alter the prediction of coastal flooding. The spatial extent of floods for 10, 100 and 1000-year return periods was calculated for the processed delta polygons of openDELvE (~ 150 deltas) and the leveed areas using Tabulate Area.~~

3. Results

3.1 openDELvE extent & summary

The current release of openDELvE contains 11,188 levees with a combined length of 19,248 km. These levees protect 1,657 separate areas that collectively span 44,734 km^2-, of which 41,399 km^2 isare on a delta (following definition in Sect. 2.4) (Table 5). Most of the data in openDELvE (97% of the leveed area) is derived from vector or high-resolution raster sources and is of good quality (Figure 4). Our leveed area data layer contains 44,734 km^2 of identified delta area protected by levees (following definition in Sect. 2.4) (Table 5). Levee line data that we have processed contains 5,089 km of levees. We have processed levee information for 15223 deltas of the 2,174 deltas identified by Caldwell et al. (2019), representing 287528%% of minimumthe global delta area (24239,0444266,885663 km^2 of 847,936 564874,142771 km^2) (Figure 2), although we appreciate that the total habitable delta area (defined as the area within a delta polygon that is not water) is calculated by Edmonds et. al. (2020) as at least 847,936 km^2. AnotherWe find 1,0987 deltas (59% of global delta area) that are pristine and/or small., where lLevees are unlikely and ; or, where we did not find data information on levee presence could be found, and nor we could wenot identify levees visually either (No Result category) (5922% of global minimum delta area). This includes deltas like the Amazon and Lena. A further 9244 deltas remain unprocessed, largely because data is unavailable — or await data input from external sources, however these are also small and collectivelyis represent only 1212% of global geomorphic minimal delta extent area,. and as such wWe have processed the largest deltas and the remaining deltas to process are small in surface areaare less likely to have levees.

Levees protect 17% (41,399 km^2) of delta area for the 153 deltas included in openDELvE, but of the delta area?

Comparing flood-protected delta area against delta extent, we find thatConsidering only processed deltas, 1910% (44,73443,875 km^2 of 239,044426,663 km^2) of the processed delta area is contained within levees. protection varies regionally. It is 2% in Asia-Pacific but 39% in Europe and C. Asia, and this broadly reflects levee presence but also data availability and data publishing policies between different regions (Figure 4). Protected delta area also varies per delta, from fully unprotected deltas such as the Colville (0%, USA) to mostly protected deltas such as the Rhine-Meuse (70%, NL). Our delta areas also include (coastal) surface water, which is 20% of the Rhine-Meuse land area, therefore the protection percentages will be higher if only land is considered.

deltaof limited levee data availability in openDELvE in Asia in particular.

Total global minimum delta area extent is 847,936564,771 km^2 (Edmonds et al., 2020), which means that, at least, 58% of global delta area is within verifiable levees. This number should be considered a minimum as many deltas remain unprocessed and we suspect that many levees exist that are not (yet) in the openDELvE dataset as even when discovering data, there existed data sources that were incompatible with the licensing of the dataset, for which we document in the delta index.

Percentage coverage of delta area by levees between continentals zones (using the UN Region from Edmonds et. al. (2020) ranges from 13% (Africa)6% (Americas) to 5417% (Europe and C. Asia & Asia-Pacific) and this broadly reflects the different data publishing policies in these regions. As can be seen in Fig. 12, we recognise that the global distribution of data is sub-optimal, and we investigate the imbalance further in the discussion (Sect. 4.2). As discussed in Sect. 2.4.1, the data from Edmonds et al. (2020) consists of polygons drawn from four maximal extent points to create a four sided polygon which represents maxima and not the absolute extent of the delta, so we have calculated these statistics based upon the geomorphic area data provided by Edmonds et al. (2020).

Of the leveed area data in openDELvE, 90% of the leveed area dataset by area (1,641 unique features) is considered of excellent or good quality (Fig. 3). This indicates that most of our sources are vector and high-quality raster data, which we believe supports high-quality onward data propagation to, and consumption by, the hydrological and risk modelling communities.

Continental Zone (UN Region)	**Number of deltas with levee data present in openDELvE[a]**	**Total number of unique leveed areas represented in dataset**	**Total geomorphic minimum deltaic area[a] [km²]**	**Area protected by levees within the delta[a] [km²]**	**Coverage of delta area by levees as % of deltaic area (computed)**
Africa	3	9	4,3524,3584,766	56957069	1323 %
Americas	100	301	105,76699,26225 8,991	13,00015,282344	1215 6 %%
Asia-Pacific	19	83	130,000128,9701 48,121	2,11025,396402	220 17 %
Europe & C. Asia	310	6886	6,4926,45414,755	2,5603,4862,560	54 1397 %
Processed Total	**1532**	**461479**	**246239,044426,6 33885**	**44,73441,3993,87 5**	**19 107 %**
No Result	1,0978	-	519,039502,9281 24,736	-	-
Unprocessed (Pending & Not Processed)	924	-	105,96413,40210 8,218	-	-
Global Total	**2,174**	**47961[a]**	**874,142847,9365 64,771[b]**	**41,39944,73443,8 75**	**5 58 %**

Table 5: Summary of processed features and deltaic area at openDELvE release v1.0(current release) per geographic region and area totalled, figures are rounded to nearest whole integer.

[a]openDELvE contains 1,601 leveed area polygons but they are partially overlapping mainly due to the structure of administrative units in the USACE NLD. Overlapping sections are therefore are only counted once and for the purpose of this table article we 'dissolved' the layer to ensure that area was not double counted.

[b]Calculated by dissolving leveed area dataset to a single layer and compared to supplementary dataminimum delta geometry from Edmonds et al. (2020).

Figure 32: Interactive browsing interface to openDELvE built using the ArcGIS Online platform. Area of focus is the Ganges-Brahmaputra delta, Bangladesh. Data content as per openDELvE version 1.0. Available publicly at: http://opendelve.eu

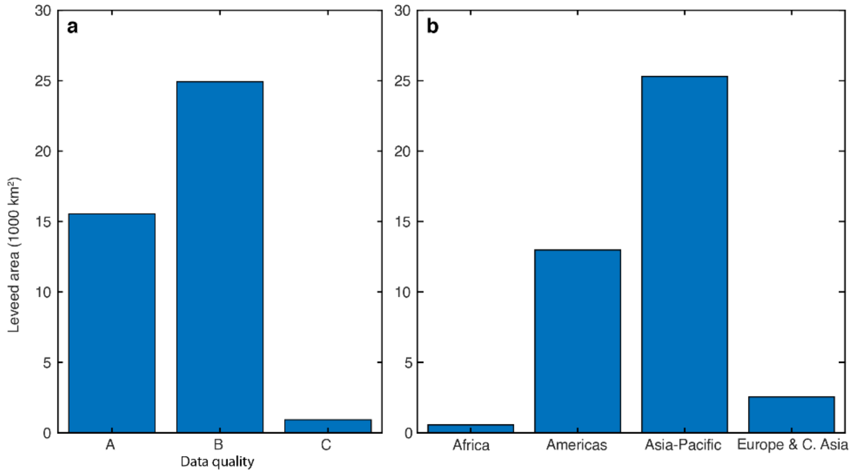

Figure ~~4~~3: (a) Distribution of data quality classification in openDELvE ~~(v1.0)~~ given for each individual leveed area feature, classified according to the data quality matrix (Table 3). (b)

~~Figure 5~~4: Distribution of leveed area data in openDELvE ~~(v1.0)~~ by UN Region. ~~using allocation of deltas to region by Edmonds et al. (2020)~~

3.2 ~~openDELvE potential~~ Demonstrative applications of openDELvE

~~The inclusion ofof the d~~Data on levees ~~inlevee data~~ can bring important insights and more accurate predictions in delta studies. Levees are sometimes included in small-scale studies, but not yet in large-scale or (global) studies (e.g. Dullaart et al., 2021, Nienhuis and van de Wal., 2021). Global studies are becoming more common, in part because of global challenges such as climate change (IPCC, 2021).

Here we showcase uses of openDELvE, including flood-protection of land use (what type of land is protected and what will be at risk), flood-protection of delta population (how many people live in flood-protected vs flood-prone areas), and potential improvements of flood hazard models in deltas (Figure 5).

~~Tosed~~

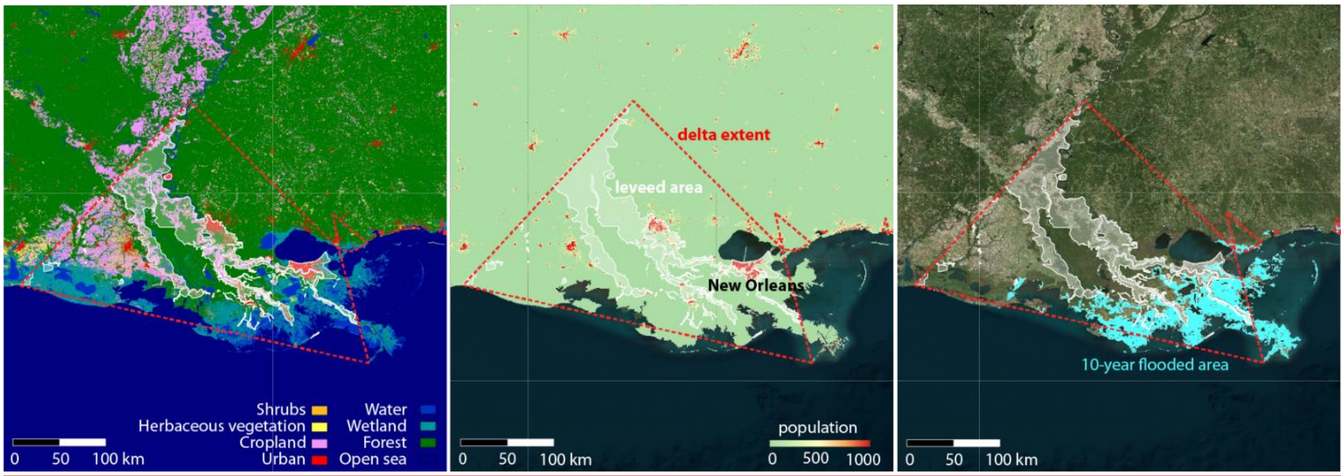

Figure ~~6~~5: Examples of land use (Copernicus Global Land Service, Buchhorn et al., 2020), population (LandScan, https://landscan.ornl.gov/), and flooded area (Dullaart et al., 2021) within and outside levees ~~Side-by-side comparison of MODIS and LandScan™ data showing land use and urban zones~~ in the Mississippi Delta

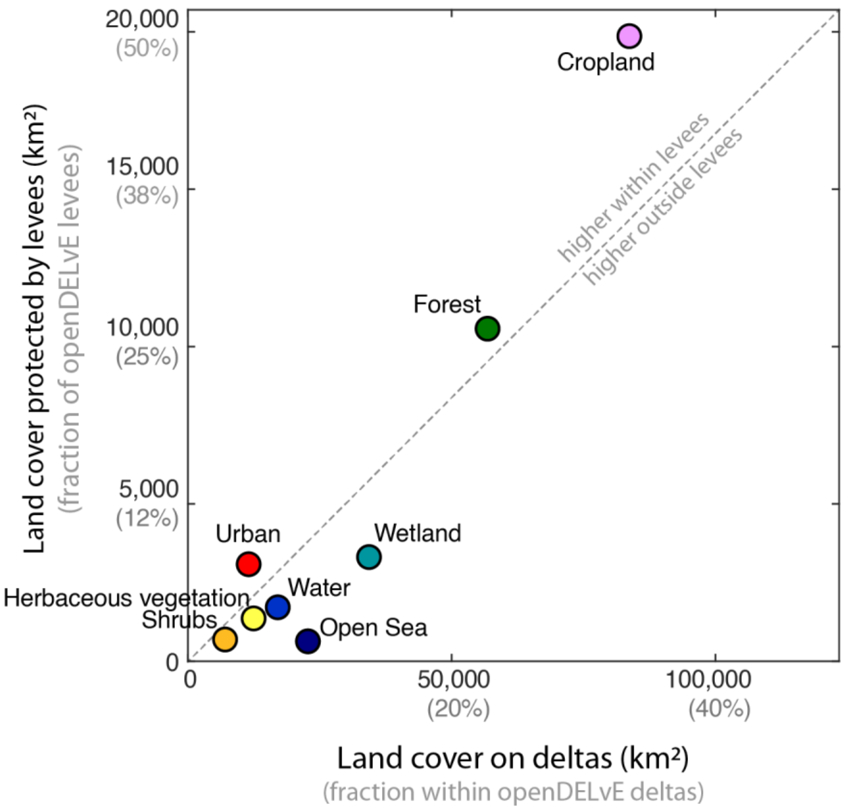

Figure 6: Average land cover for 153 deltas within flood-protection levees compared to deltas as a whole.

First, an intersection

between openDELvE and land use data shows that ~~that~~ land-use patterns ~~in general~~ differ significantly inside leveed areas compared to the rest of the delta (~~see~~ Figure 5, Figure ~~7~~6). Urban and built-up land are concentrated within leveed areas, whereas wetlands and water bodies are more likely to be found outside levees. For example, 48% of flood-protected delta area is used as cropland, compared to ~~only Cropland, for example, covers 25% of leveed delta area, but only~~ 31~~10~~% of the non-flood-protected ~~remaining (flood-prone)~~ delta area ~~of all delta area , resulting in a +15% difference~~ (Figure 6). Over- and under representation of different land use classes is likely ~~due to the fact that~~because ~~because L~~leveesd are constructed to protect land with ~~areas tend to~~ higher valued ~~land~~, such as ~~protect~~ urban, ~~and~~ built-up areas and croplands. Levees are therefore important for food availability and access (Islam and Al Mamun, 2020), protection of urban centres and urban infrastructure (Jongman et al., 2012), and for reducing exposure to flooding (Lumbroso et al., 2017). There is a second effect that can also play a role. The existence of levees could lead to greater investment and development of urban and agricultural land compared to areas outside levees (Hutton et al., 2019), the so-called "levee effect". openDELvE does not include ~~a~~ year of construction for levees, so that it is not possible~~that is needed~~ to separate these two effects.~~Meanwhile, water bodies and permanent wetlands which are not typically protected by levees.openDELvE was also combined with other existing datasets to indicate how it can be used in hazard modelling. Our analysis of population data within our subset of deltas (the openDELvE processed polygons) in section 3 supports the concept of the levee effect, where buildings and heavily urbanised zones are concentrated within leveed areas (see Figure 6) and this can result in the proliferation of buildings in such zones (see Figure 7).~~

~~We do however acknowledge thatWhilst our dataset is limited by the existence of data, however despite the limited number of deltas currently available, the concept of higher urbanisation being reflected in the leveed areas remains valid. Historically, in agricultural-rich zones such as the Mississippi Delta (Figure X)6) levee breaches have resulted in mass outward migration and an unexpected link into modernisation of historic process, as in the case of the 1927 Mississippi flood. (Hornbeck and Naidu, 2014)However with increasing percentages of urban settlement, this leads to a changed outlook on the future management of~~

levee and deltaic systems, as with the changing modern trend in intentional breaches of levees to manage ecological, hydrological, and geomorphological challenges that are resultant from urbanisation. (Nichols and Viers, 2017)

We have found that land-use patterns in general differ significantly inside leveed areas compared to the rest of the delta (see Figure 7). Leveed areas tend to protect urban and built-up areas and croplands. Meanwhile, water bodies and permanent wetlands which are not typically protected by levees.

Figure 7: Percentage difference between the modal land use for leveed areas versus total delta area for deltas processed in openDELvE (~ 150 deltas). Colour of the bar matches the MODIS scheme (see Figure 6).

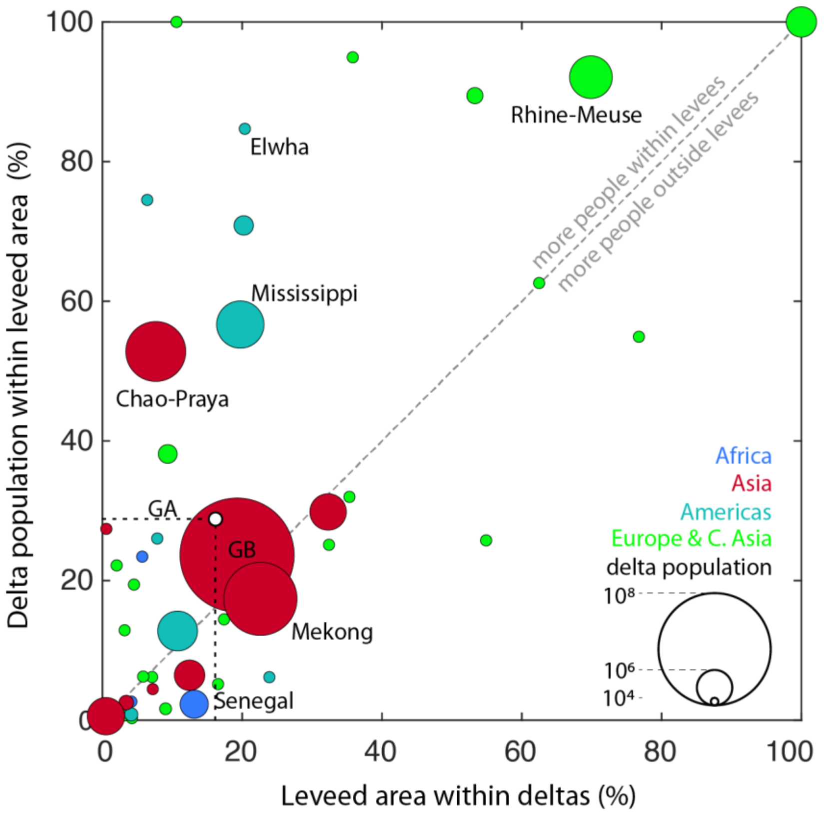

Figure 7: Flood-protected delta area vs flood-protected delta population, both as a fraction of the delta total. GB is the Ganges-Brahmaputra. The dotted line indicates the global average (GA).

Second, our analysis with openDELvE and population data suggests that, for the 153 deltas in openDELvE, 74% of delta population lives outside flood-protected areas. Population densities are higher inside flood-protected areas:
We find that Ppopulation dynamics and patternsdensity also differ inside leveed delta regions compared to the whole delta averages. Lleveed areas occupy 178% of delta area, on average, but protect 260% of delta inhabitants. However,, but theseis global averages hides large differences between deltas and regions (Figure 7). We have identified that in Africa and Asia-Pacific the relationship between leveed area and amount of people living within the leveed area mostly follows a 1:1 relation. There is large deviation however in Europe & Central Asia and the AmericasIn Europe (85%) and the Americas (41%) we find -where there can be a high dense population densitys inside therelatively small leveedlarge fraction of the delta population to be protected areas, of the delta e.g., the Ems-Dollard and Rhine-Meuse Scheldt deltas in the Netherlands (92%), and the Mississippi –in the USA (57%, Figure 5). This is not the case acrossin Africa (3%) and Asia-Pacific (24%). Looking at population densities the pattern is different. In Asian Pacific? deltas, 800 people per km^2 live outside flood-protected areas, compared to 15,000 people per km^2 inside. In contrast, iIn Europe and the Americas there is only a 5 and 9 fold increase in population densities within flood-protected areas, respectively. The different patterns could be the result of competing factors

in the co-evolution of levees and cities. There, we report no difference in population density between leveed and non-leveed delta areasthe relationship between leveed area and the amount of people protected by levees mostly follows a 1:1 relation, meaning the population density within levees is approximately equal to the population density outside levees. Meanwhile deltas like the Mekong, and Ganges Brahmaputra Meghnes (GBM) closely followed a 1:1 relation with 15-30% of the delta poulation living within the leveed area (15-30% of the total delta area).

Although levees are constructed to protect people and are therefore expected to protect morethe most populated areas, they are also constructed in vulnerable delta locations, away from likely locations of major cities that were built prior to levee constructionin regions that historically did not build levees (e.g., Bangladesh).

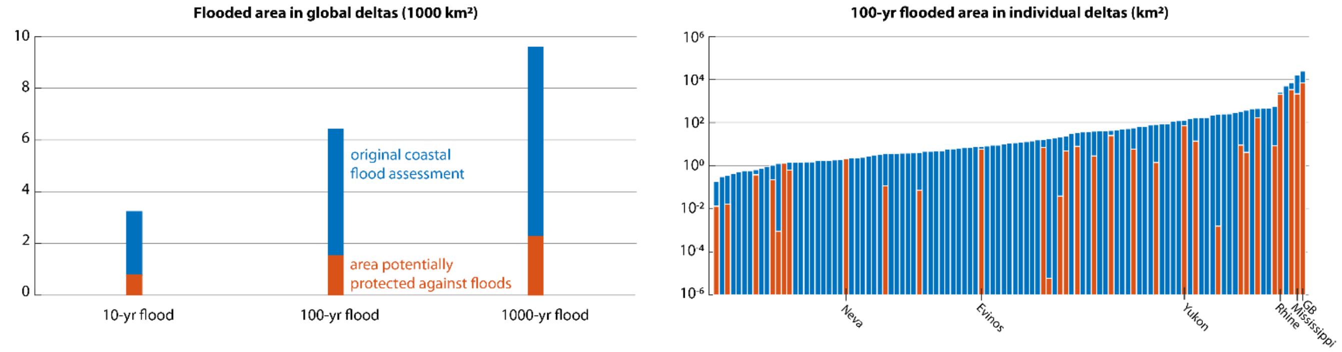

Figure 8: Delta area potentially protected against coastal floods (in red), compared to all exposed delta area (in blue), for all deltas (left) and individual deltas (right). GB in the Ganges-Brahmaputra.

Figure 8: Scatter plot indicating the relationship between percentage of the total delta population living withing leveed area and percentage leveed area per delta. Each delta processed in openDELvE (~ 150 deltas) is represented with colour indicating the UN region and the dashed line is the 1:1 relation.

In a third demonstration of openDELvE use, we assess the intersection of levees with global coastal flood assessments (Figure 5, Dullaart et al., 2021). When neglecting the presence of levees it would seemWe find that 13% (32,261 km^2) of the combined area of the 153 deltas is exposed to coastal floods everywith return period of 10 years (Figure 8). This increases to 26% (63,179 km^2) and 39% (95,879 km^2) for 100-year and 1000-year floods, respectively. However, when accounting for ther levees in some of this flood-exposed area is potentially protected by levees. Using openDELvE, we find that these could reduce flood exposure by 25% and protect between (8,206 km^2-) in the case of (10-year floods), and by XXX24% to (22,744 km^2-) in the case of(1000-yr floods) of delta land against flooding (Figure 8). Protection against floods varies greatly between deltas. Large and densely populated deltas have a mixed lLs vary widely across large and densely populated deltasFor the Rhine delta it is 78%, in the the Ganges-Brahmaputra delta it is (29%, and in)Mississippi (13%) the Mississippi delta it is (13%) the Ganges-Brahmaputra (29%) (Figure 5, Figure 8). Since openDELvE does not include data on levee heights and levee protection standards, we cannot associate each levee with a magnitude of flood; therefore, Tthese numbers are represent an approximation of the best-case protection offered by leveespreliminary and will be made more accurate if more data becomes available in openDELvE about levee heights, or designed levee protection levels (e.g., from FLOPROS from Scussolini et al., 2016).

Moreover, our analysis indicates that, if we assume that levees hold, the amount ea mapsof people at risk in deltas can be significantly lower than hazard modelling without levees indicates. In deltas where levees were present, the flooded area can be decreased on average by 20-65%. In particular, levees significantly decrease predicted flooded area in 1:100 floods. The predicted flooded area is more accurate however, for 1:1000 year floods.

~~**Figure 9: Box and whisker plot indicating the decrease in flooded area due to the presence of levees for 10,100 and 1000 year return period floods for deltas with levees processed in openDELvE**~~

~~4.~~ 4.0 Discussion

4.1 How representative is openDELvE?

As summarised in Table 5, we found that 1~~9~~7% of ~~the geomorphic~~ the delta area ~~(which can include the shallow marine portions of the delta front, Edmonds et al., 2020)~~ processed in openDELvE is protected by ~~a~~ levees. This should be considered a rough estimate. For deltas covered by nationally maintained databases (e.g., Mississippi, Rhine-Meuse) the data quality is good. There, ~~T~~there is rich metadata and ~~there is~~ little chance of false negatives (~~no levee in~~ openDELvE ~~but~~ missing existing levees ~~present nevertheless~~). Data quality and coverage in other deltas (e.g. Ganges-Brahmaputra, Mekong) is poorer, and this appears to be linked to the lack of a nationally or regionally coordinated platform for levee data sharing. ~~While individual levees or leveed areas are represented, t~~There, the ~~is a high~~ chance is higher of false negatives ~~-~~ and therefore undercounting ~~for the delta as a whole~~of leveed area.

Trying to assess global leveed area for all 2,174 global deltas, ~~by~~ including the unprocessed and no result categories, the fraction of delta area that is flood-protected is likely to be lower than 17%. Many of the “No Result” deltas are in sparsely populated areas (the Amazon, the Arctic). We expect those to have fewer levees compared to the 153 deltas within openDELvE. ~~–~~ Global delta levee area is probably higher than 5%, given that this would mean openDELvE currently includes all levees on deltas. The fraction of delta area that is protected can also be somewhat greater than 17% because of limited levee data availability in openDELvE, in Asia in particular.

~~Extrapolating our findings to a global level (Table 5) by including the unprocessed and no result categories, we might conclude that global leveed delta area is likely lower than 19%. Most deltas where we could not find information on levees (no result) are small and uninhabited, those represent 59% of global delta area. We expect those to have fewer levees compared to deltas within openDELvE. Global delta levee area is probably higher than 5%, given that this would mean openDELvE currently includes all levees on deltas. 54% is a likely maximum because this is what we found for Europe and Central Asia, where many deltas are included and data quality is generally good.~~

4.2 Global barriers to data availability

Data sovereignty is an emerging topic within global modelling that revolves around the value, sharing, and ownership of data in a global context. Whilst we acknowledge that breakthroughs have been made in the academic world of data sharing, through the formation of data initiatives (~~i.e~~e.g., FAIR) and for standardised data sharing (e.g., INSPIRE, European Parliament, 2007), data in the private and governmental sectors can still be considered as an internal asset. Tang et al. (2020) define the term ‘data sovereign’ to identify someone with the capabilities, skill set, and hierarchical position to facilitate data sharing across ~~global~~ borders~~, in turn facilitating the share of knowledge and contributing overall to the global economy~~.

~~We~~ In our search for information, we realized that~~identified~~ that countries and governmental organisations which have core values supporting open ~~(governmental)~~ data tend to treat levee information as a ‘product’ and therefore appoint a central data repository or facilitated ordering process to act as a ‘data sovereign’. ~~These~~Some repositories may not themselves hold the actual data but act as centrally maintained indices of ~~‘open government’~~ national data. Examples of national repositories ~~these~~ are the US data.gov platform, ~~(~~which holds record locators for the US Army Corps of Engineers *National Levee Database*~~)~~, the UK data.gov.uk Open Data platform, (which holds record locators for the UK Environment Agency *Asset Information Management System*), the Dutch data.overheid.nl, (which holds record locators for the Rijkswaterstaat (Dutch Ministry of Infrastructure and Water Management) *Dataregister*), and Australian data.gov.au, (which holds record locators for the various state-led systems in place across the country). These repositories may not themselves hold the actual data but act as centrally maintained indices of 'open government' data.

Other countries and institutions treat data differently,

Data ownership and 'sovereignty' canwhich can act as a roadblock to progress towards a harmonised global database. We found fewer data for deltas in Africa, China, South-East Asia, the Southern and Central Americas - as well as those in the Russian Federation and late-accession members to the EU. We have attempted to counter this by using extensive, disparate sources, but we understand that it is beneficial when the creators and maintainers of such levees additionally release information on their locations and engineered properties. We recognise that dData is often locked awaystored in local, national, or offline archives that we were unable to access, and as such this is only a partially complete dataset. Acknowledging this, we have attempted to conduct research using a best effort multilingual approach but recognise that local knowledge will ultimately prevail in the provision of high-quality data.

4.3 Bias in data availability

There is a clear global divide in data availability, with countries such as the UK, USA, Australia, and bloc systems such as the EU gathering vector flood defence data on a centralised platform. We identify that these data environments have defined a clear 'data sovereign' and this is enshrined both in law and local process. In contrast, levee data availability in deltas in Africa, South-East Asia, the Southern and Central Americas - as well as those in the Russian Federation and late-accession members to the EU to the governmental authorities deemed competent??? either simply do not have this data, or it is locked away in archives and not available in globally published formats. Due to this uneven availability of data, the countries that most urgently need accurate modelling of floods and other delta processes for successful management are those for which these data are missing, at least in openDELveThis creates a bias in our database, as there is an unfair co-occurrence of data inavailability and the requirement for data to enhance modelling and management in these areas.

We hope that bBy publishing our data as openly as possible (following FAIR principles) we intend tocan encourage not only external inspection but also suggestion of changes and further data additions. For this reason, wWe explicitly have made our map and submission system as lightweight as possible with the minimal number of questions and responses. We also encourage the furthest and widest possible spectrum for reuse. Additional crowdsourced or "*volunteer geographic information" (VGI,* Young et al., 2020) projects such as 510 - an initiative of the Netherlands Red Cross - and OpenStreetMap.org (amongst other platforms) may be able to further expand data on levees. We acknowledge that for greater global adoption a local presence is needed, and recognise the work of Young et al. (2020) in documenting the deployment of, and challenges associated with, a globally diverse data collection project, however including crowed-sourced data was out of the scope of our research.

4.4 Applications of openDeELvE: the importance of levees in hazard modelling

An important hazard for people living in coastal zones is the impact and extent of coastal flooding (Edmonds et al., 2020). The inclusion of levees in hazard modelling can give different patterns and insights and can give more accurate predictions of potential impacts. Levees are rarely included in modelling and hazards studies and yet, as identified in this research their inclusion can provide useful insight into population distribution in deltas (who is living in protected areas), land use (what type of land is protected and what will be at risk) and flood return period analysis (how many people are protected) in deltas.

For example, we have identified that in our subset of data, levees primarily protect croplands and urban land uses, which is important for considerations of: food availability and access (Islam and Al Mamun, 2020) protection of urban centres and urban infrastructure and resulting costs of flood damage (Jongman et al., 2012) and risk of death in flooding scenarios (Lumbroso et al., 2017).

We also identify that in several deltas in the Americas and Europe there are very dense populations living inside leveed portions of the delta. This is particularly relevant in terms of levee maintenance (cost, frequency, urgency) (Özer et al., 2019) and the future of levees as a flood protection measure versus pursual of other flood risk reduction strategies in these areas (Ibáñez et al., 2014; Vuik et al., 2019). The inclusion of levees can also give a more accurate picture of flood risk in deltas. As shown, levees protect large portions of deltas from predicted flood risk, particularly in terms of 1:100 year floods, which these structures are typically built for. However, as 1:1000 year floods may become more likely as the climate changes (Oppenheimer and Glavovic, 2022), it is important that the long-term flood protection that levees truly provide is assessed.

In densely urban and populated zones, this dataset could be used to indicate infrastructure risks (what is protect by levees, what is outside the leveed protection) and to estimate potential economic losses if floods were to occur (e.g. loss of crops, damage to housing etc. outside the leveed area), important consideration in terms of flood risk management.

Our research indicates that wetlands and ecologically valuable areas are not typically protected by levees (as many survive and thrive due to tidal influence). Thus, information on levees should combined with information on ecological species (flora and fauna) to indicate which species, if any, will be protected as sea-levels rise and floods become more common.

Furthermore as nature-based solutions and sedimentation enhancing strategies (Cox et al., 2022) are pursued, several of which involve the deliberate breaching of levees, openDeELvE can provide useful information, particularly if combined with ecological and land-use maps for best possible locations for such strategies to be implemented.

The inclusion of levees can also give a more accurate picture of flood risk in deltas. As shown, levees protect large portions of deltas from predicted flood risk, particularly in terms of 1:100 year floods, which these structures are typically built for. However, as 1:1000 year floods may become more likely as the climate changes (Oppenheimer and Glavovic, 2022), it is important that the long-term flood protection that levees truly provide is assessed.

4.53 Future outlook

By publishing our data as openly as possible (following FAIR principles) we intend to encourage not only external inspection but also suggestion of changes and further data additions. For this reason, we developed a webpage (www.opendelve.eu) and a new-data submission system. We encourage users to refer us to levee data that we missed, and seek partnerships with local experts of countries for additional data inclusion. Additional crowdsourced or "*volunteer geographic information" (VGI,* Young et al., 2020) projects such as 510 - an initiative of the Netherlands Red Cross - and OpenStreetMap.org may be able to further expand data on levees. We recognise the work of Young et al. (2020) in documenting the deployment of, and challenges associated with, a globally diverse data collection project, however including crowed-sourced data was out of the scope of our research.

Levee data can also be expanded using different means. There is the possibility of openDELvE to function as a training dataset for statistical (machine learning) models for levee and flood detection (Wing et al., 2019). By publishing our data with an open licence (Creative Commons Attribution) we encourage its reworking and reuse.~~Our dataset is incomplete, mainly due to the varying nature of data publishing. By publishing data as openly and FAIR as possible, and incorporating a feedback system into our web application, we aim to set a solid foundation for the development of a global levee database. Publishing openDELvE provides a foundation for the development not only of this dataset further, but also for the inclusion of higher-quality levee data in hydrodynamic models in general. Aside from direct use in global models, there is also the possibility of openDELvE to function as a training dataset for statistical (machine learning) models for levee and flood detection (Wing et al., 2019). By publishing our data with an open licence (Creative Commons Attribution) we encourage onward consumption and reworking of the database.~~

~~The data availability landscape is continually improving, and further hope that by showing the encouraging work we have achieved, further data sources will become available, and this will further foster the culture of open data and data sharing in the earth sciences and beyond. Data accessibility remains a factor in the limitation and applicability of any such system. Publishing levee data could be an important step not only in supporting flood-risk modelling, but additionally in the increased visibility and ownership of their existence. We greatly encourage further data submissions and amendments to our database as the global geoscience community identify local sources of data, but furthermore encourage authorities and public works bodies to make their data openly accessible.~~

5. Conclusion

o~~O~~penDELvE is a global delta levee database. We have standardised levee attributes and features from disparate data sources to allow for global comparability and obtained a database of ~~of~~ 11,188 levees with a combined length of 19,248 km~~461 unique leveed areas and 5,089 km of levee extent~~. ~~. We have standardised levee attributes and features to allow global comparability, finding that, f~~For the deltas ~~we processed~~in openDELvE we find that~~,~~ 4~~4~~1,399~~733.505~~ km^2 of their area is contained within levees. This represents 1~~7~~9% of their area and 5% of global delta area. Levees predominantly protect delta cropland, which comprises 48% of protected delta area. Only 26% of delta population is protected by levees, but this varies greatly ~~between~~across deltas, from 3% ~~across~~ in some deltas in Africa to 92% in the Rhine delta. Levees potentially protect up to 8,206 km^2 (10-year floods) or 22,744 km^2 (1000-yr floods) of delta land against flooding.

openDELvE can improve delta flood hazard modelling, global delta hazard assessment, and studies of sustainable delta management in the face of sea-level rise and other anthropogenic pressures. Our database is biased due to data availability, with more data available for Europe, Central Asia, and the Americas than for Africa and Asia-Pacific. openDELvE is FAIR, openly available, and we encourage contributions from other researchers and experts via http://www.opendelve.eu.

~~We have shown that use of this data set can prove important in hazard modelling including population dynamics, land use and flood risk. The inclusion of levees can aid in more sustainable delta management in the face of sea-level rise and urbanisation. more data available for ,for . openDELvE is~~

The database is FAIR, openly available and we encourage contributions from other researchers or leveeand experts. Our database is biased due to data availability, with UN regions of Asia-Pacific and Americas having a higher data availability than Africa and Europe & Central Asia.

Additionally, we acknowledge that the lack of levee data for countries whose political, financial, or administrative structure prevents open data publication, and indeed we identify the great work of many academics such as Wang et al. (2021) in creating calculated attempts as flood protection layers to facilitate disaster management and hydrological research despite this lack of data. We recommend that governments, where possible, seek to appoint a 'data sovereign' to take on this task.

Code availability

The ArcGIS® Model Builder template used to process vector data is published within the research dataset, available on DataverseNL at https://doi.org/10.34894/2WZ0S9.

Data availability

The research dataset is publicly available on DataverseNL at https://doi.org/10.34894/2WZ0S9. The layers and viewing interface are publicly consultable at http://www.opendelve.eu and are additionally hosted in the ArcGIS Online Portal for use with ArcGIS® and other OGC-compatible GIS packages. Additional data used in the applications section (population, land-use, flooded area etc.) are available through original sources, with findings per delta summarized can be found in Supplementary Table 1.

Author contribution

JO curated and maintained the data, performed the investigation, and prepared the draft manuscript with contributions from the co-authors. JHN conceptualised, validated, and supervised the project, visualised the results, and edited the manuscript. JRC supervised the project, visualised the results and edited the manuscript. DAE and PS provided key digital resources and reviewed edited the manuscript.

Competing interests

The authors declare that they have no conflict of interest.

Disclaimer

Figure 1, 2 & Supplementary Figure 1 were created using ArcGIS® software by Esri (ArcGIS® Pro [ver. 2.6.2] and ArcGIS® Online). Figures 4-8 were created using Matlab 2021a. Cartographic data displayed in the images is supplied by Esri under licence from HERE, Garmin, FAO, NOAA, EPA, NPS, and the USGS.

Acknowledgements

The authors wish to thank the many organisations who have actively chosen to embrace open data standards and share such rich datasets as this is a crucial part of the database. We are grateful to Ece Özer (TNO) and Alex Curran (formerly of TU Delft) whose work on the SAFElevee project and the ILPD, and their detailed publications as a result, were instrumental in the discovery of closed datasets. We further thank Silvia Barbetta, Albert Kettner, Yoshiki Saito, and Dhruvesh Patel for their help in gaining access to international data. We would additionally like to thank Fergus Miller Kerins for his help in testing and exploring deployment possibilities for the dataset. We thank ~~XX~~Job Dullaart and Sanne Muis for providing access to the COAST-RP flood return dataset. Finally, we are very appreciative of the support from Maarten Zeylmans van Emmichoven that allowed the data processing to continue remotely during the pandemic. JO is supported by funding from the WCFD at UU. JHN is supported by NWO vi.veni.192.132.

References

Di Baldassarre, G., Castellarin, A., Montanari, A., and Brath, A.: Probability-weighted hazard maps for comparing different flood risk management strategies: a case study, Nat. Hazards, 50, 479–496, https://doi.org/10.1007/s11069-009-9355-6, 2009.

Di Baldassarre, G., Kreibich, H., Vorogushyn, S., Aerts, J., Arnbjerg-Nielsen, K., Barendrecht, M., Bates, P., Borga, M., Botzen, W., Bubeck, P., De Marchi, B., Llasat, C., Mazzoleni, M., Molinari, D., Mondino, E., Mård, J., Petrucci, O., Scolobig, A., Viglione, A., and Ward, P. J.: Hess Opinions: An interdisciplinary research agenda to explore the unintended consequences of structural flood protection, Hydrol. Earth Syst. Sci., 22, 5629–5637, https://doi.org/10.5194/hess-22-5629-2018, 2018.

Barbetta, S., Camici, S., Maccioni, P., and Moramarco, T.: National Levee Database: monitoring, vulnerability assessment and management in Italy, in: Geophysical Research Abstracts, EGU2015-10170, 2015.

Caldwell, R. L., Edmonds, D. A., Baumgardner, S., Paola, C., Roy, S., and Nienhuis, J. H.: A global delta dataset and the environmental variables that predict delta formation on marine coastlines, Earth Surf. Dyn., 7, 773–787, https://doi.org/10.5194/esurf-7-773-2019, 2019.

Cohen-Shacham, E., Walters, G., Janzen, C., and Maginnis, S.: Nature-based solutions to address global societal challenges, edited by: Cohen-Shacham, E., Walters, G., Janzen, C., and Maginnis, S., IUCN International Union for Conservation of Nature, https://doi.org/10.2305/IUCN.CH.2016.13.en, 2016.

Cox, J. R., Paauw, M., Nienhuis, J. H., Dunn, F. E., van der Deijl, E., Esposito, C., Goichot, M., Leuven, J. R. F. W., van Maren, D. S., Middelkoop, H., Naffaa, S., Rahman, M., Schwarz, C., Sieben, E., Triyanti, A., and Yuill, B.: A global synthesis of the effectiveness of sedimentation-enhancing strategies for river deltas and estuaries, Glob. Planet. Change, 103796, https://doi.org/10.1016/j.gloplacha.2022.103796, 2022.

Criss, R. E. and Shock, E. L.: Flood enhancement through flood control, Geology, 29, 875, https://doi.org/10.1130/0091-7613(2001)029<0875:FETFC>2.0.CO;2, 2001.

Dullaart, J. C. M., Muis, S., Bloemendaal, N., Chertova, M. V., Couasnon, A., and Aerts, J. C. J. H.: Accounting for tropical cyclones more than doubles the global population exposed to low-probability coastal flooding, Commun. Earth Environ., 2, 135, https://doi.org/10.1038/s43247-021-00204-9, 2021.

Dunn, F. E., Darby, S. E., Nicholls, R. J., Cohen, S., Zarfl, C., and Fekete, B. M.: Projections of declining fluvial sediment delivery to major deltas worldwide in response to climate change and anthropogenic stress, Environ. Res. Lett., 14, 084034, https://doi.org/10.1088/1748-9326/ab304e, 2019.

Edmonds, D. A., Caldwell, R. L., Brondizio, E. S., Siani, S. M. O., and Siani, E.: Coastal flooding will disproportionately impact people on river deltas, Nat. Commun., 11, 4741, https://doi.org/10.1038/s41467-020-18531-4, 2020.

Esteves, L. S.: What is Managed Realignment?, Springer, Dordrecht, 19–31, https://doi.org/10.1007/978-94-017-9029-1_2, 2014.

European Parliament: Directive 2007/2/EC of the European Parliament and of the Council of 14 March 2007 establishing an Infrastructure for Spatial Information in the European Community (INSPIRE), 2007.

Fleischmann, A., Paiva, R., and Collischonn, W.: Can regional to continental river hydrodynamic models be locally relevant? A cross-scale comparison, J. Hydrol. X, 3, 100027, https://doi.org/10.1016/j.hydroa.2019.100027, 2019.

Friedl, M. and Sulla-Menashe, D.: MCD12Q1 MODIS/Terra+Aqua Land Cover Type Yearly L3 Global 500m SIN Grid, https://doi.org/10.5067/MODIS/MCD12Q1.006, 2019.

Hersher, R.: Levees Make Mississippi River Floods Worse, But We Keep Building Them, All Things Consid. Environ. Energy Collab., 2018.

Hornbeck, R. and Naidu, S.: When the Levee Breaks: Black Migration and Economic Development in the American South, Am. Econ. Rev., 104, 963–990, https://doi.org/10.1257/aer.104.3.963, 2014.

Hutton, N. S., Tobin, G. A., and Montz, B. E.: The levee effect revisited: Processes and policies enabling development in Yuba County, California, J. Flood Risk Manag., 12, e12469, https://doi.org/10.1111/jfr3.12469, 2019.

Ibáñez, C., Day, J. W., and Reyes, E.: The response of deltas to sea-level rise: Natural mechanisms and management options to adapt to high-end scenarios, Ecol. Eng., 65, 122–130, https://doi.org/10.1016/j.ecoleng.2013.08.002, 2014.

Islam, M. M. and Al Mamun, M. A.: Beyond the risks to food availability – linking climatic hazard vulnerability with the food access of delta-dwelling households, Food Secur., 12, 37–58, https://doi.org/10.1007/s12571-019-00995-y, 2020.

Jongman, B., Ward, P. J., and Aerts, J. C. J. H.: Global exposure to river and coastal flooding: Long term trends and changes, Glob. Environ. Chang., 22, 823–835, https://doi.org/10.1016/j.gloenvcha.2012.07.004, 2012.

Kates, R. W., Colten, C. E., Laska, S., and Leatherman, S. P.: Reconstruction of New Orleans after Hurricane Katrina: A research perspective, Proc. Natl. Acad. Sci., 103, 14653–14660, https://doi.org/10.1073/pnas.0605726103, 2006.

Lumbroso, D. M., Suckall, N. R., Nicholls, R. J., and White, K. D.: Enhancing resilience to coastal flooding from severe storms in the USA: international lessons, Nat. Hazards Earth Syst. Sci., 17, 1357–1373, https://doi.org/10.5194/nhess-17-1357-2017, 2017.

Middelkoop, H., Erkens, G., and van der Perk, M.: The Rhine delta-a record of sediment trapping over time scales from millennia to decades, J. Soils Sediments, 10, 1–12, https://doi.org/10.1007/s11368-010-0237-z, 2010.

Mitchell, T., Kluskens, R., Woldringh, B., Van Der Meer, M. T., Kamp, R. G., De Gooijer, C., and Hillen, M. M.: Integrating Levee Performance Assessments into Complex Flood Protection Systems, in: US Society of Dams Annual Conference, 2013.

Munoz, S. E., Giosan, L., Therrell, M. D., Remo, J. W. F., Shen, Z., Sullivan, R. M., Wiman, C., O'Donnell, M., and Donnelly, J. P.: Climatic control of Mississippi River flood hazard amplified by river engineering, Nature, 556, 95–98, https://doi.org/10.1038/nature26145, 2018.

National Research Council: Levees and the National Flood Insurance Program, National Academies Press, Washington, D.C., 63–96 pp., https://doi.org/10.17226/18309, 2013.

Nichols, A. L. and Viers, J. H.: Not all breaks are equal: Variable hydrologic and geomorphic responses to intentional levee breaches along the lower Cosumnes River, California, River Res. Appl., 33, 1143–1155, https://doi.org/10.1002/rra.3159, 2017.

Nienhuis, J. H. and van de Wal, R. S. W.: Projections of Global Delta Land Loss From Sea-Level Rise in the 21st Century, Geophys. Res. Lett., 48, e2021GL093368, https://doi.org/10.1029/2021GL093368, 2021.

Nienhuis, J. H., Ashton, A. D., Edmonds, D. A., Hoitink, A. J. F., Kettner, A. J., Rowland, J. C., and Törnqvist, T. E.: Global-scale human impact on delta morphology has led to net land area gain, Nature, 577, 514–518, https://doi.org/10.1038/s41586-019-1905-9, 2020.

Oppenheimer, M. and Glavovic, B.: Sea Level Rise and Implications for Low Lying Islands, Coasts and Communities, in: Special Report on the Ocean and Cryosphere in a Changing Climate, edited by: Pörtner, H.-O., Roberts, D. C., Masson-Delmotte, V., and Zhai, P., Cambridge University Press, https://doi.org/10.1017/9781009157964, 2022.

Özer, I. E., van Damme, M., and Jonkman, S. N.: Towards an International Levee Performance Database (ILPD) and Its Use for Macro-Scale Analysis of Levee Breaches and Failures, 12, 119, https://doi.org/10.3390/w12010119, 2019.

Pinter, N.: One Step Forward, Two Steps Back on U.S. Floodplains, Science (80-.)., 308, 207–208, https://doi.org/10.1126/science.1108411, 2005.

Pinter, N., Jemberie, A. A., Remo, J. W. F., Heine, R. A., and Ickes, B. S.: Flood trends and river engineering on the Mississippi River system, Geophys. Res. Lett, 35, 23404, https://doi.org/10.1029/2008GL035987, 2008.

Pinter, N., Huthoff, F., Dierauer, J., Remo, J. W. F., and Damptz, A.: Modeling residual flood risk behind levees, Upper Mississippi River, USA, Environ. Sci. Policy, 58, 131–140, https://doi.org/10.1016/j.envsci.2016.01.003, 2016.

Sampson, C. C., Smith, A. M., Bates, P. D., Neal, J. C., Alfieri, L., and Freer, J. E.: A high-resolution global flood hazard model, Water Resour. Res., 51, 7358–7381, https://doi.org/10.1002/2015WR016954, 2015.

Santos, M. J. and Dekker, S. C.: Locked-in and living delta pathways in the Anthropocene, Sci. Rep., 10, 1–10, https://doi.org/10.1038/s41598-020-76304-x, 2020.

Scussolini, P., Aerts, J. C. J. H., Jongman, B., Bouwer, L. M., Winsemius, H. C., De Moel, H., and Ward, P. J.: FLOPROS: an evolving global database of flood protection standards, Nat. Hazards Earth Syst. Sci., https://doi.org/10.5194/nhess-16-1049-2016, 2016.

Scussolini, P., Tran, T. V. T., Koks, E., Diaz-Loaiza, A., Ho, P. L., & Lasage, R.: Adaptation to sea level rise: A multidisciplinary analysis for Ho Chi Minh City, Vietnam. Water Res. Res., 53, 10,841– 10,857. https://doi.org/10.1002/2017WR021344, 2017

Steinfeld, C. M. M. M., Kingsford, R. T., and Laffan, S. W.: Semi-automated GIS techniques for detecting floodplain earthworks, Hydrol. Process., 27, 579–591, https://doi.org/10.1002/hyp.9244, 2013.

Tang, C., Plasek, J. M., Zhu, Y., and Huang, Y.: Data sovereigns for the world economy, Humanit. Soc. Sci. Commun., 7, 184, https://doi.org/10.1057/s41599-020-00664-y, 2020.

Trigg, M. A., Birch, C. E., Neal, J. C., Bates, P. D., Smith, A., Sampson, C. C., Yamazaki, D., Hirabayashi, Y., Pappenberger, F., Dutra, E., Ward, P. J., Winsemius, H. C., Salamon, P., Dottori, F., Rudari, R., Kappes, M. S., Simpson, A. L., Hadzilacos, G., and Fewtrell, T. J.: The credibility challenge for global fluvial flood risk analysis, Environ. Res. Lett., 11, 094014, https://doi.org/10.1088/1748-9326/11/9/094014, 2016.

US Army Corps of Engineers: River Analysis System (HEC-RAS), https://www.hec.usace.army.mil/software/hec-ras/, 2020.

Vuik, V., Borsje, B. W., Willemsen, P. W. J. M., and Jonkman, S. N.: Salt marshes for flood risk reduction: Quantifying long-term effectiveness and life-cycle costs, Ocean Coast. Manag., 171, 96–110, https://doi.org/10.1016/j.ocecoaman.2019.01.010, 2019.

Wang, D., Scussolini, P., and Du, S.: Assessing Chinese flood protection and its social divergence, Nat. Hazards Earth Syst. Sci., 21, 743–755, https://doi.org/10.5194/nhess-21-743-2021, 2021.

Werner, B. T. and McNamara, D. E.: Dynamics of coupled human-landscape systems, 91, 393–407, https://doi.org/10.1016/j.geomorph.2007.04.020, 2007.

Van Wesenbeeck, B. K., Mulder, J. P. M., Marchand, M., Reed, D. J., De Vries, M. B., De Vriend, H. J., and Herman, P. M. J.: Damming deltas: A practice of the past? Towards nature-based flood defenses, Estuar. Coast. Shelf Sci., 140, 1–6, https://doi.org/10.1016/j.ecss.2013.12.031, 2014.

Westermann, W. L.: The Development of the Irrigation System of Egypt, Class. Philol., 14, 158–164, https://doi.org/10.1086/360222, 1919.

Wilkinson, M. D., Dumontier, M., Aalbersberg, Ij. J., Appleton, G., Axton, M., Baak, A., Blomberg, N., Boiten, J. W., da Silva Santos, L. B., Bourne, P. E., Bouwman, J., Brookes, A. J., Clark, T., Crosas, M., Dillo, I., Dumon, O., Edmunds, S., Evelo, C. T., Finkers, R., Gonzalez-Beltran, A., Gray, A. J. G., Groth, P., Goble, C., Grethe, J. S., Heringa, J., t Hoen, P. A. C., Hooft, R., Kuhn, T., Kok, R., Kok, J., Lusher, S. J., Martone, M. E., Mons, A., Packer, A. L., Persson, B., Rocca-Serra, P., Roos, M., van Schaik, R., Sansone, S. A., Schultes, E., Sengstag, T., Slater, T., Strawn, G., Swertz, M. A., Thompson, M., Van Der Lei, J., Van Mulligen, E., Velterop, J., Waagmeester, A., Wittenburg, P., Wolstencroft, K., Zhao, J., and Mons, B.: Comment: The FAIR Guiding Principles for scientific data management and stewardship, Sci. Data, 3, 1–9, https://doi.org/10.1038/sdata.2016.18, 2016.

Wing, O. E. J., Bates, P. D., Neal, J. C., Sampson, C. C., Smith, A. M., Quinn, N., Shustikova, I., Domeneghetti, A., Gilles, D. W., Goska, R., and Krajewski, W. F.: A New Automated Method for Improved Flood Defense Representation in Large-Scale Hydraulic Models, Water Resour. Res., 55, 11007–11034, https://doi.org/10.1029/2019WR025957, 2019.

Winsemius, H. C., Aerts, J. C. J. H., Van Beek, L. P. H., Bierkens, M. F. P., Bouwman, A., Jongman, B., Kwadijk, J. C. J., Ligtvoet, W., Lucas, P. L., Van Vuuren, D. P., and Ward, P. J.: Global drivers of future river flood risk, Nat. Clim. Chang., 6, 381–385, https://doi.org/10.1038/nclimate2893, 2016.

Young, J. C., Lynch, R., Boakye-Achampong, S., Jowaisas, C., Sam, J., and Norlander, B.: Volunteer geographic information in the Global South: barriers to local implementation of mapping projects across Africa, GeoJournal, 1–17, https://doi.org/10.1007/s10708-020-10184-6, 2020.