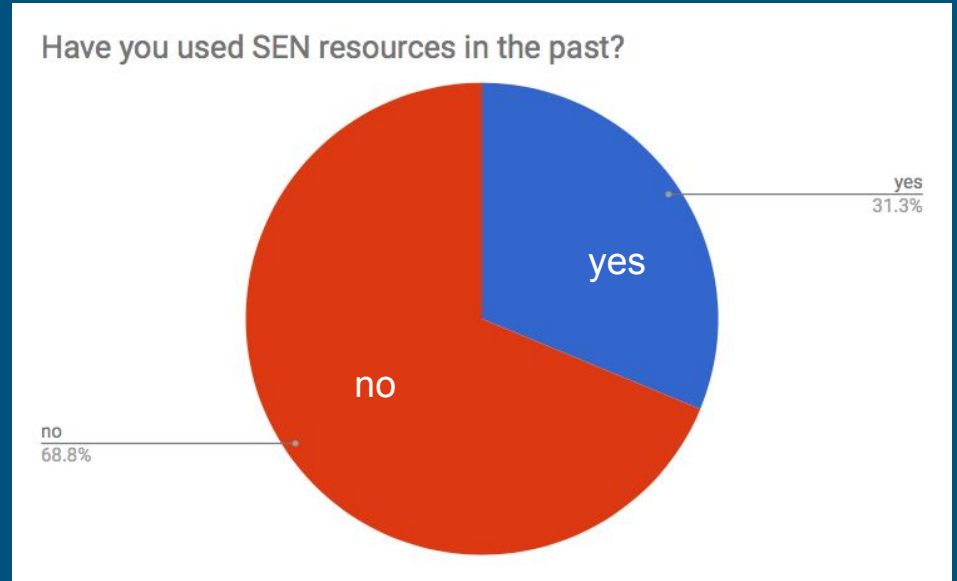# Wrangling your research data

Part 2
SEN Knowledge Base and Data Repos

Leslie Hsu and
Sediment Experimentalist Network Team
May 2018

Goals for this presentation

- Introduce tools to help you meet data management and publication requirements.

- Address issues brought up in pre-workshop survey.

Have you used SEN resources in the past?

yes
31.3%

yes

no
68.8%

no

Does this apply to you?

You are publishing a paper and the journal asks you to provide a link to the supporting data.

You are writing a data management plan and need to describe plans for managing the project's data.

You are trying to set up a lab environment where new members can quickly learn your lab's group knowledge.

You would like to optimize your time when writing papers and proposals, including collaborations.
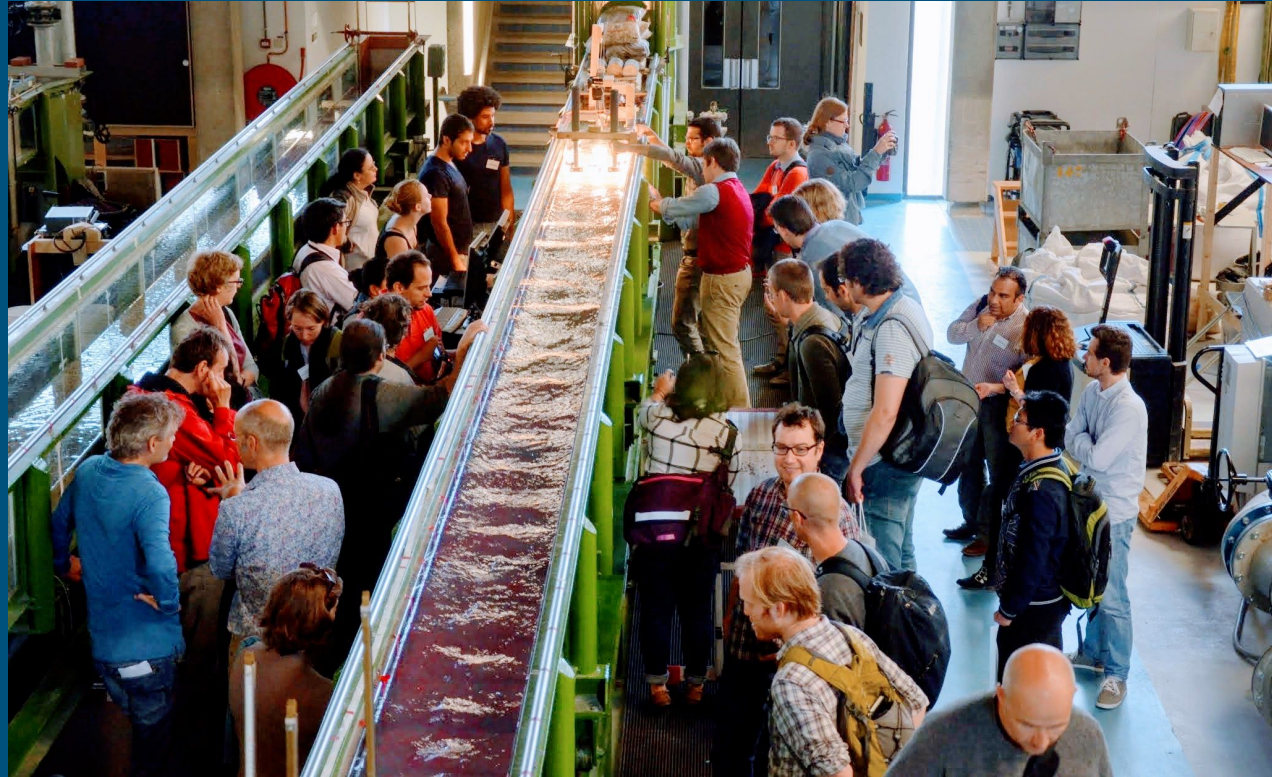
Topics we will review today

- Sediment Experimentalist Network (SEN)
- EarthCube
- SEN Knowledge Base (SEN KB)
- Data repository options

Sediment Experimentalist Network (SEN)

A community of people working to bring information and tools for working with data to laboratory geomorphologists and related researchers.

Sediment Experimentalist Network (SEN)

EarthCube

A community of scientists across all geoscience domains, as well as geoinformatics researchers and data scientists. It is a joint effort between the NSF Directorate for Geosciences and the Division of Advanced Cyberinfrastructure.

Has funded 63 projects since 2013. (Has solicitations like NSF-GLD)

Supports both technical development projects and community activities like SEN.

SEN Knowledge Base (SEN-KB)

sedexp.net

SEN's effort to gather our community knowledge in a centralized place.
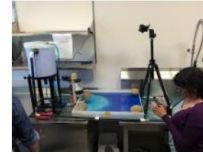
SEN KB categories

- Lab facility
- Equipment
- Experimental set-up
- Method
- Data catalog

SEN-KB - showcase lab capabilities

Lab facility

Experimental Setups:
St. Anthony Falls drip box
Bedform development in recirculating flume
Subaqueous and subaerial currents with net deposition
One-dimensional rice pile avalanche

Equipment:
Data Acquisition Carriage
SAFL Main Channel
SAFL density current basin
20" Flume
24" Flume

Methods:
2d topographic scan with sonar
Mixing salt-water solution for density current experiments
Designing and running the XES-15 experiment in the Jurassic tank
Still Image Correction Set Up

Data:
Bedform adjustment experiments
Rice Pile Experiments

SEN-KB - meet funder requests
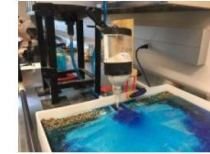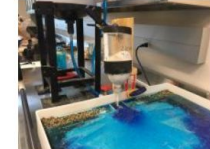
A system for automation of an experimental sedimentology basin at Tulane University
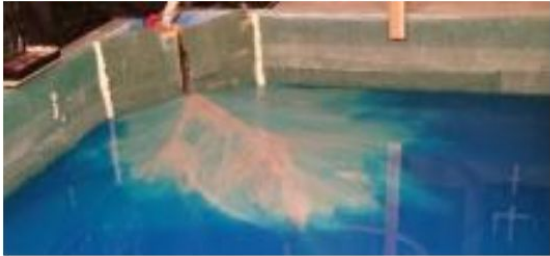
Setup

Tulane University

TSDs
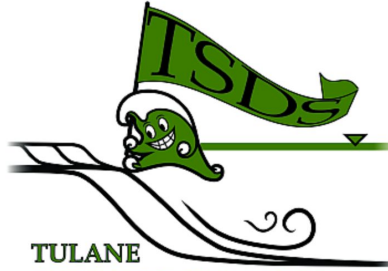
TULANE
SEDIMENT DYNAMICS
& STRATIGRAPHY

Lab facility

TDB_15_1, Tulane Delta Basin

TDB-15-1
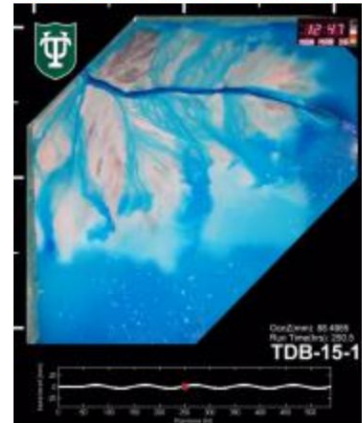
Data catalog

SEN-KB - make unpublished data available

Method

Lab facility

Data catalog

Data & Papers:

Internet resource:
Access Data at figshare: Debris-flow flume force measurements (from Hsu et al., 2014)
Access Data at SEAD: Synthetic Bedrock Strength Measurements for Debris Flow Erosion Experiments
Access Data at SEAD: Debris Flow Flume Video (mp4)
Access Data at SEAD: Debris Flow Flume: Bedrock Erosion Data

Storage capability of some common repositories (as of Sept 25, 2016)
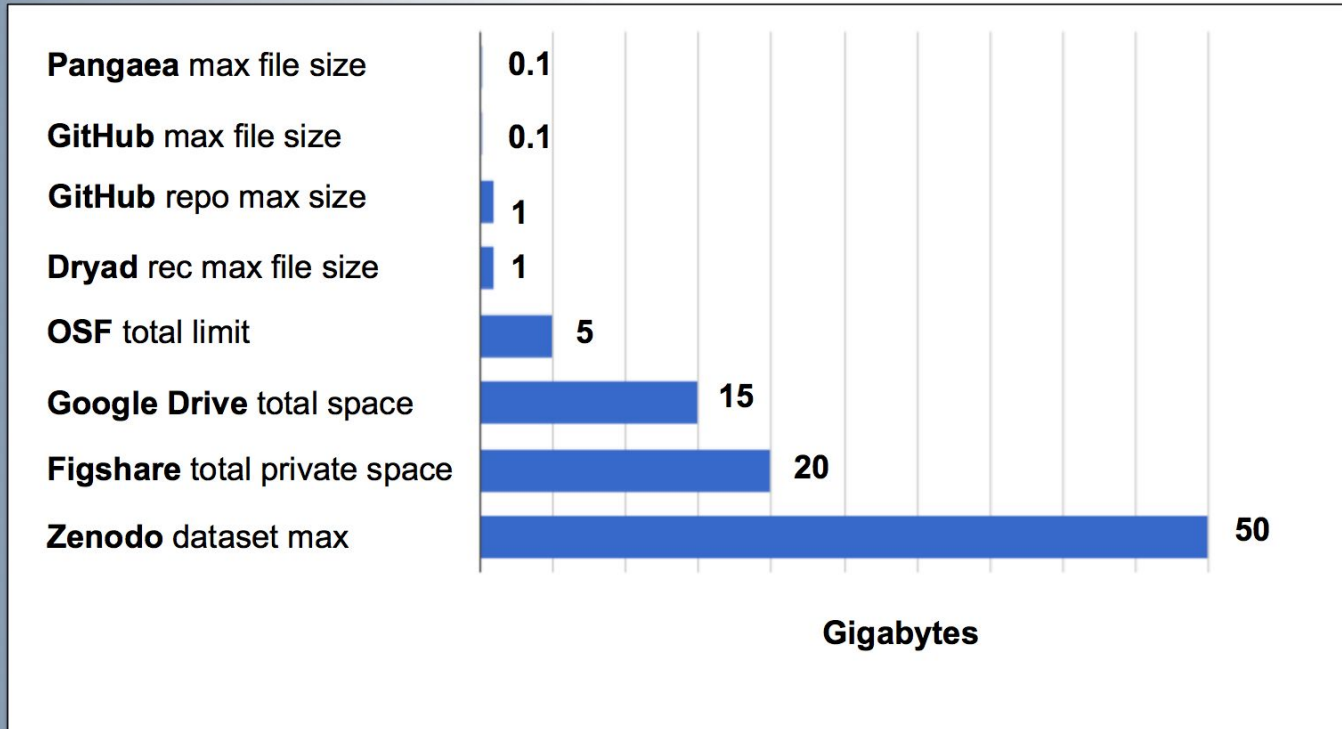
Bar chart titled "Gigabytes":

- **Pangaea** max file size — 0.1
- **GitHub** max file size — 0.1
- **GitHub** repo max size — 1
- **Dryad** rec max file size — 1
- **OSF** total limit — 5
- **Google Drive** total space — 15
- **Figshare** total private space — 20
- **Zenodo** dataset max — 50

Storage capability of some common repositories (as of Sept 25, 2016)

Repository	Gigabytes
Pangaea max file size	0.1
GitHub max file size	0.1
GitHub repo max size	1
Dryad rec max file size	1
OSF total limit	5
Google Drive total space	15
Figshare total private space	20
Zenodo dataset max	50
Collins dataset	109
Tulane dataset	842

Gigabytes
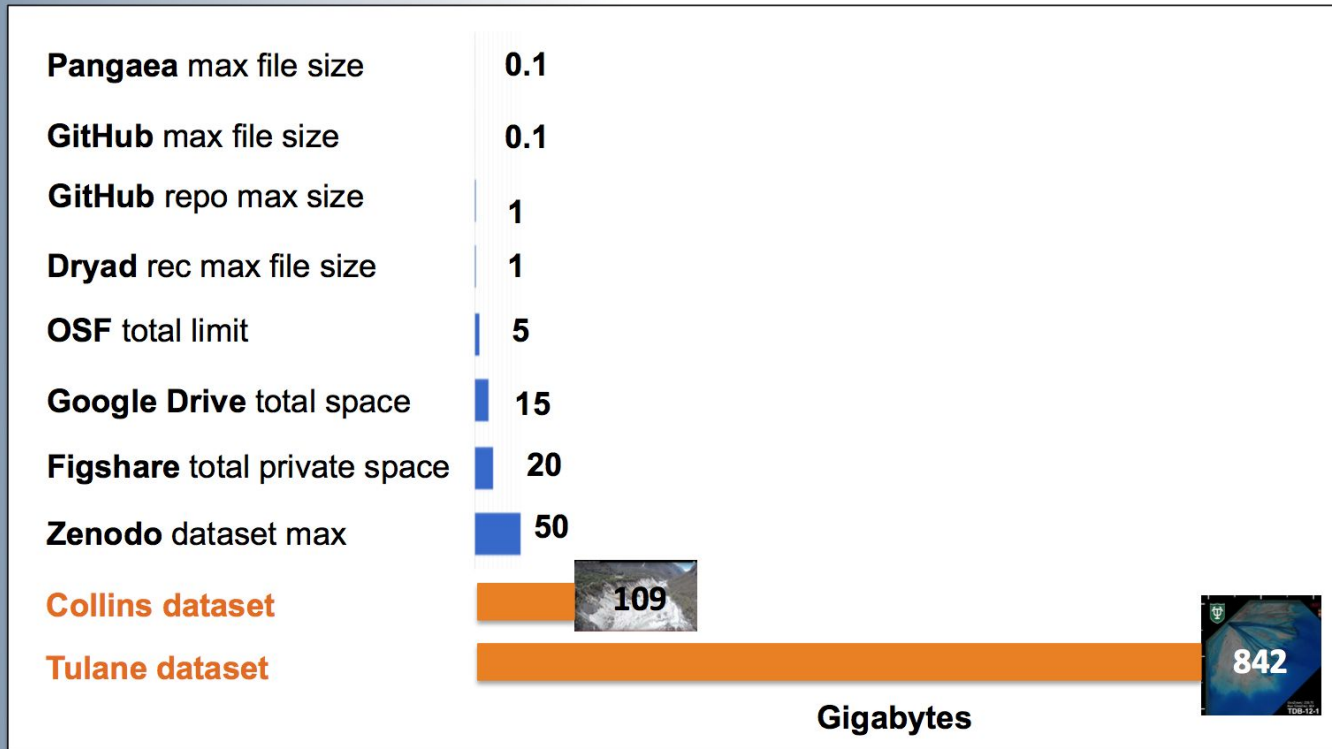
Disciplinary repository example - SEAD

Journal of Sedimentary Research

SEPM — Society for Sedimentary Geology

Archive   Current Issue   About The Journal ▾   About the Society ▾   SEPM Member Sign In   This Journal ▾   Advanced Search

Acknowledgments

Figures & Tables   Contents   GeoRef   References   Related

Fig. 1.—
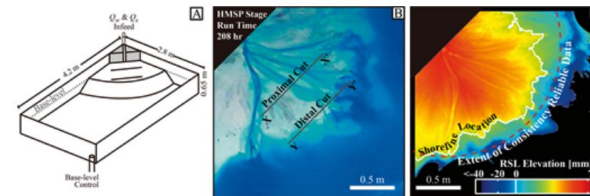
View large   Download slide

Disciplinary repository

SEAD Internal Repository

Title:	**TDB_12_1**
Persistent Identifier:	http://doi.org/10.5967/M03N21GX
Creation Date:	2017-07-26
Creator(s):	Li, Qi
	Straub, Kyle
	Qi Li; Kyle Straub
Contact(s):	qli1@gatech.edu;kmstraub@tulane.edu
	Straub, Kyle M.

Abstract: TDB-12-1: Fan-delta experiment performed in Tulane University Delta Basin. Experiment evolved under constant forcings of water (0.17 l/s), sediment (0.00017 l/s), and sea-level rise rate 0.25 (mm/hr). Experiment run time was 1285 hr. Experiment used a strongly cohesive sediment that had a wide grain size distribution with a median diameter of 65 microns. Experiment performed to explore autogenic sediment transport and stratigraphy with topography monitored every 1 hour of run time.

Keywords: experiment, delta, flume

License: PUBLIC DOMAIN

Contents: Total Size (unzipped): 35.08 GB
Number of Files: 2997
Largest File: 2.29 GB
Number of Folders: 1
Longest Folder Path: 5
Data Mimetypes: text/plain, video/quicktime, application/x-rar-compressed, application/zip, multi/files-zipped, application/x-zip-compressed, application/vnd.openxmlformats-officedocument.wordprocessingml.document, application/octet-stream

Individual files can be downloaded through the links below. The buttons below allow retrieval of the entire publication package or the complete metadata document, respectively.

Name	Size
General Experiment Meta data_TDB_12.docx	14.64 kB
▶ 🗀 TDB_12_1 Results and Processed Data	--
▶ 🗀 TDB_12_1 Logs	--
▶ 🗀 TDB_12_1 Design, Preparation, and Evaluation	--
▶ 🗀 TDB_12_1 Original Data	--
▶ 🗀 TDB_12_1 Data Processing, Programs, Code, Settings	--

Collections

Viewing most recent collections

Tulane Sediment Dynamics and Quantitative Stratigraphy Group

Datasets associated with the Tulane Sediment Dynamics and Quantitative Stratigraphy Group.

💼 10 ▦ 0

Institutional repository

4TU.Centre for Research Data

| Contact | L

Dataset **Data presented in the paper "Zooming in and out: scale-dependence of extrinsic and intrinsic factors affecting salt marsh erosion"**

▶ ▶ ▶ ▶ Link as https://doi.org/10.4121/uuid:214b7b5f-84cc-4b16-8026-3b76463e128d | How to cite this data

▼go to DATA section▼

<< more info...

Home

Upload datasets

Personal page

» Search in Data
» Search in "info"

title		Data presented in the paper "Zooming in and out: scale-dependence of extrinsic and intrinsic factors affecting salt marsh erosion"
creator		orcid *Wang, H. (Heng)*
creator		orcid *van der Wal, D. (Daphne)*
creator		orcid *Li, X. (Xiangyu)*
creator		orcid *van Belzen, J. (Jim)*
creator		orcid *Herman, P. (Peter)*
creator		orcid *Hu, Z. (Zhan)*
creator		orcid *Zhenming, G. (Ge)*
creator		orcid *Zhang, L. (Liquan)*
creator		orcid *Bouma, T. (Tjeerd)*
contributor		GELIFES, Groningen University
contributor		Institute of Estuarine and Coastal Research, School of Marine Science, Sun Yatsen University
contributor		NIOZ Royal Netherlands Institute for Sea Research, Department of Estuarine and Delta Systems, and Utrecht University
contributor		State Key Laboratory of Estuarine and Coastal Research, East China Normal University
date accepted		2017-07-24
date created		2015-09 through 2017-05
date published		2017
description		This research aims to identify the critical extrinsic (wind exposure and foreshore morphology) and intrinsic factors (soil and vegetation properties) affecting the erosion of salt marsh edges at different spatial scales. The dataset includes the rates of cliff lateral retreat (.txt) quantified using a time series of aerial photographs taken over four salt marsh sites in the Westerschelde estuary, the Netherlands, the wind exposure of marsh edges (.txt), the input data of SWA (Simulating WAve Nearshore) model (.dat and .txt) and the output wave height of Westerschelde estuary (.dat)
language		en
publisher		NIOZ Royal Netherlands Institute for Sea Research
subject		Foreshore morphology ◊ Root biomass ◊ Salt marsh edges determination ◊ Sediment erodibility ◊ Sediment grain size ◊ So properties ◊ SWAN ◊ Vegetation properties ◊ Wind exposure
▲ in collection		General collection of datasets
spatial coverage		Westerschelde (estuary)
map		Map [kml]
related publication		Zooming in and out: scale-dependence of extrinsic and intrinsic factors affecting salt marsh erosion (article, 2017)
licence		General terms of use

TUDelft

TU/e

UNIVERSITEIT TWENTE.

WAGENINGEN
UNIVERSITY & RESEARCH

DATA

? Dataset files (3.0 MiB) >> **download complete dataset (zip)** | **download separate files**

Commercial repository

search on figshare

Browse

ReadMe.rtf (1.77 kB)

ARCHIVE

data.zip (15.82 MB)

Cite   Download all (15.82 MB)   Share   Embed   + Collect (you need to log in first)

Turbulence as the main driver of resuspension in oscillatory flow through vegetation

24.01.2018, 16:27 by Rafael Tinoco, Giovanni Coco

Records of velocity and suspended sediment concentration obtained from multiple ADVs and OBSs as described in Tinoco & Coco 2018, JGR-Earth Surface.

Log in to write your comment here...

Current practices for citing data

2016: 13 of 19 experimental JGR-ES papers in 2016 say that for the supporting data, look in the paper, the references, or "Contact the author."

2017-18: 3 of 10 experimental JGR-ES papers in 2017-18 say that for the supporting data, look in the paper, the references, or "Contact the author."

The ideal way to cite data

Referenced in the text of the paper.

Listed in the references.

With a persistent identifier (usually DOI) and link.

In this format:

Yu, L., and Straub, K.M., 2017, TDB_15_1, SEAD,
http://doi.org/10.5967/M00V89W1.

Another way

In the acknowledgments:

The processed data set is available at
https://doi.org/10.17882/54834.

Experimental clay-laden density flow velocity and deposit data from the Eurotank laboratory

Date	2016
Author(s)	Hermidas Navid[1], Eggenhuisen Joris T.[2], Silva Jacinto Ricardo[3], Luthi Stefan M.[1], Toth Ferenc[4], Pohl Florian[2]
Affiliation(s)	1 : TUDelft, Faculty of Civil Engineering and Geosciences, Delft, Netherlands 2 : Utrecht University, Department of Earth Sciences, Utrecht, Netherlands 3 : IFREMER, Marine Geosciences Unit, France 4 : University of Szeged, Faculty of Science and Informatics, Szeged, Hungary
DOI	10.17882/54834
Publisher	SEANOE
Keyword(s)	flume experiments, clay-laden density flow, Eurotank, deposits, debrite, turbidite
Abstract	This dataset contains flume experiment data from a series of runs that were performed at the Eurotank laboratory of Utrecht university. It includes velocity data as well as photographs that were taken from the deposits emplaced by the density flows inside the flume. The flume was 3.7 m long, 0.22 m wide, and 0.5 m high with glass side walls. For the experiments, the slope was adjusted to 6, 8, and 9.5 degrees. Before the runs, mixtures composed of 1) sand, silt, clay, and water or, 2) sand, clay, and water, were prepared and stirred for approximately 30 minutes to 2 hours in a 0.45 m³ mixing tank and the flume was filled with tap water. The sediment volume concentration was varied between 9%, 15%, and 21% and contained 2/3 sand and 1/3 clay for the runs without silt and 1/3 sand, 1/3 silt, and 1/3 clay, for the runs including silt. The mixtures were then pumped with a discharge rate of 10 or 15 m³/h into the flume. Velocity data was obtained using two Ultrasonic Doppler Velocity Profiler probes (UVP Duo Mx, 1 MHz) which were placed at 0.7 m and 2.7 m distance from the inlet, at a height of 0.11 m above the bed, and with an abgle of 60 degree to the bed, facing up stream. This dataset was used in Hermidas et al., 2018 to classify clay-rich subaqueous density flow structures.
Licence	(cc) BY-NC
Utilisation	Data are published without any warranty, express or implied. The user assumes all risk arising from his/her use of data. Data are intended to be research-quality, but it is possible that these estimates or the data themselves contain errors. It is the sole responsibility of the user to assess if the data are appropriate for his/her use, and to interpret the data accordingly. Authors welcome users to ask questions and report problems.
Acknowledgments	We would like to thank The Netherlands Research Center For Integrated Solid and Earth Science (ISES) for funding this project. We are grateful to J. de Leeuw and J. Trabucho-Alexandre at Utrecht university for their assistance in using the flume and their helpful discussions and inputs.
Sensor metadata	Each Ultrasonic Doppler Velocity Profiler (UVP) probe emitted and received 32 bursts, followed by a dead time of approximately 10 ms before the burst sequence of the second probe. The duration of each burst was 5 µs. The time resolution of a burst sequence was 0.16 ms. The time between successive burst sequences of a single probe was 0.22 s. The thickness of a measurement bin along the z direction was 0.64 mm. The total duration of each flume measurement was approximately 1 min.

Data				
File	Size	Format	Processing	Access
Velocity data and photographs of the emplaced deposits	3 GB	TEXTE	Raw data	Open access

References
Hermidas Navid, Eggenhuisen Joris T., Jacinto Ricardo Silva, Luthi Stefan M., Tóth Ferenc, Pohl Florian (2018). A classification of clay-rich subaqueous density flow structures. Journal of Geophysical Research: Earth Surface, -.

Activity 1: SEN Knowledge Base

Scan the Lab Facilities page on www.sedexp.net - is your lab represented?

  If no, create a page.

  If yes, make sure the content is up to date.

Activity 2:

Browse the SEN Knowledge Base.

If there is content from you or your lab, is it properly linked?

If there isn't content there from you or your lab yet, sketch out a contribution that would help you organize your information.

What other capabilities are needed?

Activity 3: (Try this at home)

You (or a member of your lab) spend 4 hours (half a day) mapping out what information would be useful to contribute to SEN to help make sure there is adequate data description for other lab members or future users of the data.

Guidance for sharing/publishing data

1. Share raw data if possible. If too large, share processed data, such as tables.

2. Share at a trusted digital repository, as opposed to your own server. In order of preference: disciplinary, institutional, and commercial repositories are available.

3. Obtain a persistent identifier (such as a DOI).

4. Keep your own copy of the data.

5. Use proper citation to give and receive credit.

6. Check with SEN for any policy or tool updates. sedimentexp@gmail.com

Quick 2 question feedback - thank you!

bit.ly/sen2018