Global Study of Tailings Dam Failures by Statistical and Remote Sensing Methods

by

Nahyan Muhammad Rana

A thesis

presented to the University of Waterloo

in fulfillment of the

thesis requirement for the degree of

Doctor of Philosophy

in

Earth Sciences

Waterloo, Ontario, Canada, 2023

Examining Committee Membership

The following served on the Examining Committee for this thesis. The decision of the Examining Committee is by majority vote.

Supervisor:	Stephen G. Evans Professor of Geohazards Department of Earth and Environmental Sciences / Geological Engineering University of Waterloo
Internal Member:	Keith B. Delaney Continuing Lecturer Department of Earth and Environmental Sciences University of Waterloo
Internal Member:	Martin Ross Associate Professor of Applied Quaternary Science Department of Earth and Environmental Sciences University of Waterloo
Internal Member:	David Blowes Professor of Groundwater Geochemistry Department of Earth and Environmental Sciences University of Waterloo
Internal External:	Giovanni Cascante Professor of Geotechnical Engineering Department of Civil and Environmental Engineering University of Waterloo
External Examiner:	Giovanni Crosta Professor of Engineering Geology Universita degli Studi di Milano Bicocca, Italy

Author's Declaration

This thesis consists of material all of which I authored or co-authored: see Statement of Contributions included in the thesis. This is a true copy of the thesis, and any required final revisions, as accepted by my examiners.

I understand that my thesis may be made electronically available to the public.

Statement of Contributions

This thesis consists in part of three manuscripts written for publication as part of the multi-institution CanBreach project (see Acknowledgements section).

Research presented in Chapter 2:

This research was led by Nahyan Rana at the University of Waterloo under the supervision of Dr. Stephen Evans. Nahyan Rana conceived and conceptualized the study design and scope. Nahyan Rana and Negar Ghahramani (The University of British Columbia) collected case history data. Nahyan Rana performed the statistical analyses and wrote the draft manuscript. Andy Small (Klohn Crippen Berger), Nigel Skermer (Independent Consultant), Dr. Stephen Evans, Dr. Scott McDougall (The University of British Columbia), and Dr. W. Andy Take provided supervisory support and intellectual input.

Citation:

Rana, N.M., Ghahramani, N., Evans, S.G., Small, A., Skermer, N., McDougall, S., Take, W.A., 2022a. Global magnitude-frequency statistics of the failures and impacts of large water-retention dams and mine tailings impoundments. Earth-Sci. Rev. 104144. https://doi.org/10.1016/j.earscirev.2022.104144.

Research presented in Chapter 3:

This research was led by Nahyan Rana at the University of Waterloo under the supervision of Dr. Stephen Evans. Nahyan Rana and Dr. Stephen Evans conceived and conceptualized the study design and scope. Nahyan Rana and Negar Ghahramani collected case history data. Nahyan Rana performed the empirical-statistical analyses and wrote the draft manuscript. Andy Small, Dr. Stephen Evans, Dr. Scott McDougall, and Dr. W. Andy Take provided supervisory support and intellectual guidance.

Citation:

Rana, N.M., Ghahramani, N., Evans, S.G., McDougall, S., Small, A., Take, W.A., 2021a. Catastrophic mass flows resulting from tailings impoundment failures. Eng. Geol. 292, 106262. https://doi.org/10.1016/j.enggeo.2021.106262.

Research presented in Chapter 4:

This research was led by Nahyan Rana at the University of Waterloo under the supervision of Dr. Stephen Evans. The research is to be submitted as a manuscript to the journal venue in May 2023. Nahyan Rana conceived and conceptualized the study design and scope. Nahyan Rana performed the remote sensing (InSAR) investigations and wrote the draft manuscript. Evan Deane (BGC Engineering) connected Nahyan Rana to the anonymized mine owner to help obtain ground-truth monitoring prism data. Daniel Adria (Knight Piesold) and Negar Ghahramani contributed to summarizing case history knowledge. Keith Delaney (University of Waterloo), Andy Small, Dr. Stephen Evans, Dr. Scott McDougall, and Dr. W. Andy Take provided supervisory support and intellectual input.

Citation:

Rana, N.M., Delaney, K., Evans, S.G., Deane, E., Small, A., Adria, D., McDougall, S., Ghahramani, N., Take, W.A. Analysis of Sentinel-1 PS-InSAR data using an automated commercial software to monitor tailings dams and predict instability. To be submitted to Remote Sensing of Environment in May 2023.

As lead author of these three main-body chapters, I was responsible for conceptualizing the study design and scope, carrying out data collection and analysis, and drafting and submitting manuscripts. My co-authors provided guidance during each step of the research and delivered feedback on draft manuscripts.

Abstract

The growth of the mining industry, mainly driven by increased resource demands, has accelerated the production and storage of mine waste. Tailings comprise the fine-grained fraction of this waste. Due to their intrinsically hazardous characteristics from a geochemical and geotechnical perspective, tailings are conventionally discharged in slurry form behind constructed dams for public and environmental safety purposes. However, tailings dams have been under significant scrutiny following a number of high-profile, catastrophic breach events in, for example, Canada (2014), Brazil (2015 and 2019), India (2017, 2019, and 2021), China (2017 and 2021), and South Africa (2022). This thesis aims to study such phenomena using advanced statistical and remote sensing methods. Specifically, this thesis focuses on quantifying the failure rate statistics via a magnitude-frequency approach, mapping breach-runout processes by GIS analysis of historical satellite/aerial imagery, and analyzing pre-failure deformation patterns using satellite InSAR. By this, the thesis enhances knowledge on how often tailings dam failures occur, the common causal and failure mechanisms for these events, the potential downstream consequences, and the capabilities of satellite technology to monitor, and predict instability in, tailings dams.

Chapter 2 notes that global magnitude-frequency statistics of dam failures in general remain poorly quantified, out-of-date and/or limited in scope in scientific literature. Addressing this gap would give insight into how the hazard-risk of tailings dams has evolved over time in comparison to water-retention dams, and would provide quantitative benchmarks for estimating likelihoods of dam failures and their societal impacts to support risk assessments. The chapter introduces and analyzes new datasets and estimates of the construction and failures of large reservoir facilities (LRFs) and tailings storage facilities (TSFs) worldwide in the period 1965–2020. Long-standing data gaps on LRF failures in China, and subsequently worldwide, and on constructed TSFs worldwide are addressed by new estimation/extrapolation techniques to illustrate the range of uncertainty in the results.

The total number of LRF failures is estimated to have been between 394 and 608. The annual numbers of newly constructed and failed LRFs declined near-proportionally, thus the cumulative failure rate of LRFs stayed fairly constant, falling in the range of 1.2% to 1.8% as of end-of-2020. The rate drops to at least 0.7% when excluding China. The cumulative fatality rate of LRFs reduced over time to 1.2 deaths per constructed facility, and falls in the range of 64 to 98 deaths per failure, as of end-of-

2020. Failures of LRFs with very high storage capacities (>200 M m^3) have continued to occur since 2016.

In comparison, the annual number of TSF failures stayed relatively constant, whereas the annual construction rate of TSFs is estimated to have increased by ~3x, thus the cumulative failure rate of TSFs declined over time. When assuming the lower-estimate of the number of constructed TSFs (6810), the cumulative failure rate is ~4.4% as of end-of-2020. When adopting the upper-estimate (20,230 TSFs), a rate of ~1.5% is obtained, which falls in the same order as the corresponding rate of LRFs. A review of published estimates of existing TSFs worldwide indicates that the "true" rate is much lower than 4.4% and closer to 1.5%. The cumulative fatality rates of TSF failures reduced over time to 0.1–0.3 death per constructed facility and 6 deaths per failure as of end-of-2020, which are lower than those of LRFs. However, the size and the environmental impact of TSF failures have increased on average worldwide, especially since 2014. The rising global rate of failed tailings volumes has been approximately proportional to the rising global rate of tailings production since the 1990s.

Heavy rainfall events and intensifying precipitation patterns are concluded to be statistically important causative variables for the failures of both LRFs and TSFs. This has implications for the design and management of storage capacity, freeboard, facility drainage and spillways under climate change conditions. The results are applicable broadly on a global scale and are conditioned by uncertainties in the data and the methods used to address data gaps. To improve the robustness of future statistical analyses, a more comprehensive public disclosure effort is necessary, particularly with respect to reservoir facility failures in China and constructed TSFs worldwide.

Chapter 3 focuses on the downstream mass flows resulting from TSF failures (i.e. tailings flows). The chapter is supported by a new global database of 63 tailings flow cases that have been remotely analyzed through a compilation of satellite imagery, digital elevation models and literature. The synthesis provides insight into the influence of impoundment conditions, preconditioning and trigger variables, failure mechanisms and the downstream environment on tailings flow behavior. The database also sheds light on the limitations of data quality and availability in the public domain. Magnitude-frequency statistics indicate that TSF failures that have produced catastrophic tailings flows with total outflow volumes of ≥1 M m^3 have occurred at a mean recurrence interval of 2–3 years over the period 1965–2020. Weather hazards and impoundment drainage issues are identified as major causative variables. The occurrence of liquefaction and/or the incorporation of free water are sufficient conditions

to trigger extremely rapid, highly mobile behavior. Travel path confinement and steeper bed slopes enhance flow velocities (peak of 25–30 m/s) and kinetic energy, whereas flow mobility appears to be exacerbated along major rivers.

Although general trends may be observed in empirical observations, it is concluded that such efforts are prone to substantial uncertainty due to the complexity and variability of site conditions (that are typically unaccounted for in broad statistical approaches) as well as poor data availability and/or quality for many of the selected cases. This highlights the importance of performing site-specific investigations through numerical models, laboratory tests and field observations to better predict post-breach behavior (ideally within a probabilistic framework) when undertaking site assessments.

Chapter 4 focuses on satellite InSAR technology, which has grown in popularity due to its ability to detect millimeter-scale displacements in infrastructure along the line-of-sight (LOS). The literature on InSAR applications on tailings dams is relatively limited compared to other engineering sectors. This has led to limited understanding of (i) whether InSAR can be as accurate or representative as on-the-ground instrumentation, (ii) whether failures of tailings dams can be predicted in advance using InSAR, and (iii) what conditions and variables the quality of InSAR results and monitoring/prediction efforts depend on. To help fill this research gap, open-source Sentinel-1 data is analyzed to undertake a ground-truth assessment at a test site and a forensic investigation of 7 failure cases (2017-2019). The methodology involves the use of a cost-saving commercial software with an automated Persistent Scatterer (PS) workflow, with comparison to a proprietary software, implemented with a dual PS and Distributed Scatterer (DS) algorithm with advanced noise-filtering proficiency, for the ground-truth test site and one forensic case study.

The chapter delivers key considerations that are of practical value. Firstly, commercial InSAR software with Sentinel-1 data provides reasonable accuracy when monitoring consolidation and lower-scale deformation (< 50 mm/yr) but is ineffective in larger or accelerating deformation regimes (> 50 mm/yr), where advanced proprietary algorithms are more appropriate. Secondly, environmental conditions strongly influence the quality of Sentinel-1 PS-InSAR. The best results are in dry, bare-earth or urban terrains, whereas vegetation, water, and snow/ice cover worsen PS-InSAR data quality. Thirdly, Sentinel-1 InSAR is a useful hazard-screening tool in active mine sites with large TSFs or multiple TSFs, or in legacy mines with abandoned TSFs, as it may help guide where to undertake targeted investigations. Fourth, sub-horizontal movements in the north-south direction are generally

poorly captured due to the polar orbit of SAR satellites. Lastly, not all TSF failures show clear warning signs for weeks in advance, and the revisit interval of SAR satellites prevents detection of instantaneous failure modes. Therefore, long-term monitoring programs should ideally be integrated with a combination of remote sensing methods and field instrumentation. Future avenues of research could involve case-study comparisons between medium-resolution Sentinel-1 and high-resolution TerraSAR-X data and ground-truth comparisons between different proprietary algorithms.

Chapter 5 synthesizes the main contributions of this thesis and delivers a concluding outlook on the progress of the tailings industry in reaching its aspirational goal of zero-harm. This commentary is provided in the context of the September 2022 Jagersfontein TSF disaster in South Africa, after which the Church of England investors issued an urgent call for action on (i) a global registry of all TSFs including privately owned and legacy sites and (ii) a global monitoring system that integrates satellites as well as ground sensors, where available, to monitor the "highest-risk" dams. The focus of this thesis has naturally aligned with these goals and the results have contributed to both initiatives. While the Jagersfontein incident offers a sobering reminder that much work remains to be done, a point of optimism is that thesis is just one of many ongoing research projects on tailings management and safety (e.g. liquefaction susceptibility, dam breach-runout modelling, water management, design criteria) – all of which are necessary to facilitate a safer, more sustainable mining industry.

Acknowledgements

This thesis was carried out within the framework of the Canadian Tailings Dam Breach (CanBreach) project that was commenced in mid-2019. The project combined the resources of three Canadian research institutions (The University of British Columbia, Queen's University, and University of Waterloo) with the support of five industrial partners (Imperial Oil Resources, Suncor Energy, BGC Engineering, Klohn Crippen Berger, and Golder Associates) as part of a Collaborative Research and Development (CRD) Grant issued by the Natural Sciences and Engineering Research Council of Canada (NSERC). The two main goals of the CanBreach project were to: 1) significantly improve understanding of tailings dam breach and flow/runout processes and characteristics, and 2) develop improved tailings dam breach and runout analysis methods that increase the accuracy of inundation maps, consequence classifications, and emergency response and preparedness plans. To achieve this, the CanBreach team integrates three key research methods: 1) field (excluded due to COVID-19), statistical and remote observations of the behavior of historical tailings dam breaches and tailings flows led by Dr. Stephen Evans at Waterloo, 2) laboratory-based physical modelling (flume experiments) of tailings dam breach processes led by Dr. Andy Take at Queen's, and 3) numerical modelling of tailings flows led by Dr. Scott McDougall at UBC. This thesis presents the research led by Waterloo during the first phase of the CanBreach research program.

Additional funding during this PhD program (Sept. 2019 – April 2023) consisted of the following:

- Gary Salmon Memorial PhD Scholarship awarded by the Canadian Dam Association in the final year of the PhD program.
- Queen Elizabeth II Graduate Scholarship in Science and Technology (QEII-GSST) awarded by the Ontario Provincial Government of Canada for the final year of the PhD program.
- President's Scholarship awarded by Waterloo for the final year of the PhD program.
- Science Domestic Graduate Scholarship awarded by Waterloo for the entire duration of the PhD program.
- Science Graduate Award provided by Waterloo for the entire duration of the PhD program.
- Three-time Teaching Assistantships (TA) for the 4th year Engineering Geology course (EARTH 438) at Waterloo.

Dedication

To Mom and Dad for their selflessness.

To Dr. Evans for his wise mentorship and consistent support.

Table of Contents

List of Figures

List of Tables

Chapter 1: Introduction

1.1 Context

Mineral extraction and processing is a valuable sector of the world economy. The rising demand of mineral resources due to increased resource demands has necessitated the growth of the mining industry in recent decades. An adverse consequence of this growth has been the acceleration of waste – specifically tailings – production per mine and in total worldwide. Tailings are fine-grained, wet, environmentally hazardous by-products of ore extraction (Blight et al., 2010). For societal and environmental protection, tailings are conventionally discharged as slurries into dammed, open-air impoundments termed as tailings storage facilities (TSFs). In contrast to conventional water-retention dams, tailings dams are evolving structures that are raised in successive stages over the lifecycle of mine operations (Vick, 1983; Blight et al., 2010; Kossoff et al., 2014; Canadian Dam Association, 2020). TSFs are designed to be permanent geomorphological features in the landscape that require expert care and monitoring for decades until closure, abandonment, and reclamation by state authorities (Vick, 1983; Morgenstern, 2012; Schafer et al., 2018; McKenna and van Zyl, 2020).

However, poorly designed or managed TSFs can be susceptible to failure, which in some cases can lead to catastrophic consequences. There have been over 360 reported TSF failures worldwide since 1928 (Islam and Murakami, 2021) and at least 2,800 related deaths (Santamarina et al., 2019), including 272 fatalities in the Feijao (Brazil) disaster in January 2019 (Robertson et al., 2019; de Lima et al., 2020). In November 2015, the collapse of the Fundao TSF in the same mining region in Brazil caused 15 deaths and triggered the largest downstream release of mine waste in history (Morgenstern et al., 2016). In addition to the human risk, TSF failures can also produce major economic impacts and long-lasting contamination, as evidenced by the 2014 Mt. Polley (Canada) environmental disaster (Morgenstern et al., 2015) that affected the livelihoods of local Indigenous communities (Shandro et al., 2017). High-profile TSF failures have also occurred in Italy (e.g. Chandler and Tosatti, 1985; Takahashi, 1991), Romania (e.g. UNEP and OCHA, 2000; Mossa and James, 2013), Hungary (e.g. Mecsi et al., 2013; Banvolgyi, 2018), South Africa (e.g. Midgley et al., 1979; Wagener, 1997; Fourie et al., 2001a,b; Blight and Fourie, 2005; Church of England, 2022), the USA (e.g. Harder and Stewart, 1996; AECOM, 2009), Russia (e.g. Glotov et al., 2018), Chile (e.g. Dobry and Alvarez, 1967; Villavicencio et al., 2013), Mexico (e.g. Macias et al., 2015), Peru (e.g. ICOLD, 2001; Bowker and

Chambers, 2015; Petley, 2020d), Australia (e.g. Jefferies et al., 2019), China (e.g. Reid and Fourie, 2016; Petley, 2020a), and India (e.g. Petley, 2020b; Zhuang et al., 2021).

The seemingly high occurrence rate of such incidents has generated intense public scrutiny of the tailings management industry. In response, global experts and mining stakeholders from the International Council on Mining & Metals (ICMM), the Church of England (COE), the United Nations Environmental Programme (UNEP), and the International Commission on Large Dams (ICOLD) have commissioned new guidelines on improving design and performance standards and hazard-risk assessment protocols (ICMM et al., 2020; ICOLD, 2023).

One approach to elevating the state-of-the-art in risk assessments would be to broaden case history knowledge. The few existing datasets of TSF failures (ICOLD, 2001; Bowker and Chambers, 2015; Islam and Murakami, 2021, WISE, 2021) are important secondary venues of historical documentation, but due to poorly reported case history information in the literature, these inventories may be sparse in quantitative details (e.g. impounded or released volume) and may lack context, definition, and/or consistency for key variables (e.g. runout distance). The resulting knowledge gaps contribute to considerable uncertainties in empirical (e.g. Rico et al., 2008b; Larrauri and Lall, 2018), semi-empirical (Innis et al., 2022), and numerical (Ghahramani et al., 2022) models – all of which can inform risk-based decision-making. An improvement in the quality and quantity of case history documentation would enhance the global understanding of failure rate statistics and breach-runout processes, which in turn would logically lead to better-informed predictive models and risk assessments.

A novel way to fill case history knowledge gaps and contribute to global risk-reduction initiatives is to leverage satellite technology for geospatial documentation and geo-statistical analysis (i.e. remote sensing). The objective of remote sensing is to derive spatiotemporal information of features or phenomena on Earth. In the field of engineering geoscience, two categories of satellite imagery are especially relevant: optical and radar. Optical satellite sensors capture images of the Earth's surface in the visible or near-visible portion of the electromagnetic spectrum, with popular end-uses being visual monitoring and change detection for land-use mapping and disaster impact assessments (e.g. Evans et al., 2021). The primary drawbacks to optical satellite imagery are night-time coverage and inhibited penetration by cloud cover. These obstacles were overcome by the invention of synthetic aperture radar (SAR) satellites, which emit microwave-band signals to monitor the Earth's surface in day-or-night and

all-weather conditions. A uniquely powerful feature of SAR satellites is the ability to detect millimeter-scale changes by interferometric analysis (InSAR) (e.g. Ferretti et al., 2001, 2011; Crosetto et al., 2016).

Improved accessibility to these advanced technologies makes this a timely opportunity to obtain new quantitative insights on historical TSF failures, which can ultimately improve scholarly understanding of tailings dam breach-runout behavior and feed into enhanced regression models, numerical simulations, and hazard-risk assessments.

1.2 Scope and Objectives

The goal of this doctoral dissertation is to enhance knowledge and documentation on failure statistics, breach-runout processes, and pre-failure deformation patterns of tailings dams. Chapter 2 defines the problem by estimating the number of tailings dams and quantifying their historical failure rate, in comparison to water-retention dams for context. Chapter 3 maps the consequence by providing a state-of-the-art overview of tailings flow behaviour via analysis of optical satellite imagery, compilation of case history data, and development of new empirical relationships. Chapter 4 evaluates the potential of a modern solution by exploring the capabilities and limitations of satellite InSAR technology for monitoring tailings dams. The combination of these individual chapters are aimed at providing a holistic story of tailings dam failures. The implications of this thesis are to be reflected by developments in both academic research and industry practice, as discussed in Chapter 5. An overview of the specific objectives, methods, and outcomes of each main chapter is provided below.

1.2.1 Chapter 2: Global Magnitude-Frequency Statistics of the Failures of Tailings Dams in Comparison to Large Reservoir Dams

Chapter 2 presents the first comprehensive global study of failure and fatality rates, per annum and per constructed facility, of constructed dams in scientific literature. The purpose of this chapter is to set the scene for this thesis by quantifying how the rates for TSFs have evolved over the period 1965-2020 in comparison to “large” water-retention dams. This work involves a detailed literature review of datasets and past statistical studies followed by a new, independent, global-scale analysis.

The comparison between the two categories of dam structures is useful because water-retention dams are the closest geotechnical engineering analogs to TSFs with respect to construction methods and breach modelling applications; thus, a direct comparison contextualizes the global hazard-risk of TSFs.

The specific focus on "large" reservoir dams is due to the recently issued ICOLD Registry on Large Dams that documents the construction frequency and basic design characteristics of "large" water-retention dams with storage capacities of ≥ 3 M m^3 or dam heights ≥ 15 m with statistical completeness (i.e. global data on smaller water-retention dams is sparse). Given that the construction of TSFs is poorly documented worldwide, a novel statistical approach is presented that is based on extrapolation from global mineral production data reported by the British Geological Survey (BGS, 2021). This enables the first representative approximation of the number of existing TSFs worldwide.

With the support of new datasets, statistical estimates, and geospatial insights, this chapter illustrates cumulative magnitude-frequency profiles of TSF failures and large reservoir dam failures (in terms of volume and fatalities) and plots normalized cumulative failure and fatality rates. By addressing these pre-existing statistical gaps, this chapter provides quantitative benchmarks for estimating likelihoods of dam failures and their societal impacts to support future risk assessments.

This chapter was published in the journal *Earth-Science Reviews* in August 2022 (Rana et al., 2022a). The supplementary materials to this chapter, which include multiple databases, calculation sheets, a supplementary article, and a compilation of satellite images, are published in the *Borealis* data repository (Rana et al., 2022b).

1.2.2 Chapter 3: Catastrophic Mass Flows from Tailings Dam Failures

Chapter 3 narrows the research focus to "tailings flows" – downstream mass flows resulting from the failures of TSFs. Previous studies (e.g. Rico et al., 2008b; Larrauri and Lall, 2018) developed empirical relationships between the released volume and downstream runout distance based on values reported in publicly available datasets and news articles. However, these pre-existing values were proposed without validation of the runout extent by remote sensing and without contextualizing the flow characteristics in terms of the tailings properties and the nature of the downstream terrain.

The main objectives of this chapter are as follows: (i) present a case history synthesis of tailings flow characteristics from initiation to breach process to runout behavior, as influenced by the tailings properties, impoundment characteristics, and downstream topography (i.e. channelized vs. unconfined; wet vs. dry); (ii) implement a new GIS-remote sensing methodology (Ghahramani et al., 2020) to measure runout distances, inundation areas, and travel path angles of tailings flows dating back to 1965;

and (iii) develop new, more representative empirical relationships between total impounded volume and total released volume, and between total released volume and the three key runout parameters (inundation area, runout distance, and travel path angle). The chapter also illustrates the first magnitude-frequency profile of catastrophic tailings flows (i.e. those with flow volumes of ≥ 1 M m^3).

A new global database of tailings flows is appended to this chapter, supported by a compilation of GIS-processed satellite/aerial imagery of 37 historical cases. The chapter produces scholarly insights on breach-runout behavior and downstream processes as influenced by flow interactions with diverse, complex terrains. The geo-database and empirical observations also have implications for screening-level assessments, semi-empirical models, and numerical simulations in industry practice.

This chapter was published in the journal *Engineering Geology* in July 2021 (Rana et al., 2021a). The supplementary materials to this chapter, which include the database, a supplementary article, and satellite images, are published in the *Borealis* data repository (Rana et al., 2021b).

1.2.3 Chapter 4: Using Sentinel-1 InSAR to Monitor Tailings Dams and Predict Instability

Chapter 4 studies the utility of satellite InSAR in monitoring tailings dams. Prior to this chapter, there have been only 2 TSF failures that have been forensically analyzed by InSAR (2018 Cadia in Australia and 2019 Feijao in Brazil; Carla et al., 2019; Jefferies et al., 2019; Thomas et al., 2019; Gama et al., 2020; Holden et al., 2020; Rotta et al., 2020; Grebby et al., 2021; Hudson et al., 2021; Bayaraa et al., 2022) and 2 case studies on long-term InSAR monitoring of metastable TSFs (Kennecott, Utah and Zelazny Most, Poland; Hu et al., 2017; Mazzanti et al., 2021). This limited archive is mainly due to the conceptual and workflow complexities and the long processing times for SAR data, which can render such studies impractical for users with limited resources.

This chapter significantly expands the InSAR case study inventory by undertaking a ground-truth assessment at a test site and by forensically examining the precursor displacements in 7 failure cases in the period 2017-2019. For all these cases, Sentinel-1 SAR data (freely downloadable on the Alaska Satellite Facility or Copernicus Hub websites) is processed via the Persistent Scatterer (PS) technique (Ferretti et al., 2001; Crosetto et al., 2016) on a cost-saving, automated, commercial software called ENVI SARScape Analytics. In addition, for two of the case study sites (ground-truth test site and one

failure case), Sentinel-1 data is also processed on the state-of-the-art proprietary SqueeSAR software (TRE Altamira) that integrates both PS and Distributed Scatterer (DS) points and an advanced noise-filtering algorithm which enables it to generate a higher-quality set of displacement results in most terrains. In the ground-truth test site, the Sentinel-1 displacement data from both software are compared to *in-situ* monitoring prism data.

The main objectives of this chapter are as follows: (i) analyze the accuracy of Sentinel-1 PS-InSAR as a proactive monitoring tool on a site-scale; (ii) assess if breach location and breach timing can be predicted by the present approach; (iii) identify errors in the results due to inherent limitations of the selected software/algorithm and/or the lower-resolution Sentinel-1 data; and (iv) comment on the influence of environmental conditions (e.g. vegetation or ponded water), dam-satellite geometry, and failure modes on the quality and interpretation of InSAR results.

This comprehensive study allows for the comparison of the quality of the results depending on the selected software, thus providing insights into practical considerations primarily from an engineer's perspective. This work has global implications for implementing satellite-based monitoring systems across a diverse portfolio of TSFs by consultants and dam owners to complement on-site instrumentation. In light of recent proposals by global mining investors to develop InSAR monitoring systems for poorly monitored or closed/abandoned TSFs in high-risk areas (Church of England, 2022), this study also has implications for the potential scaling up of InSAR as a regional-scale, hazard-screening and/or risk prioritization tool.

This chapter will be submitted to the journal *Remote Sensing of Environment* in May 2023.

1.2.4 Chapter 5: Synthesis and Outlook

This concluding chapter synthesizes the scholarly and practical significance and the outcomes of this thesis. The chapter places into global context the novel insights drawn in this thesis, particularly as it relates to the background hazard-risk of TSFs, the geotechnical and geomorphological processes involved in TSF failures, and the capabilities and limitations of using satellites to monitor TSFs. Lastly, the chapter reflects on the ongoing rapid scale of progress in this field and delivers a forward-looking commentary on further work that needs to be done to ensure tailings management becomes a safer, more sustainable practice.

Chapter 2: Global Magnitude-Frequency Statistics of the Failures of Tailings Dams in Comparison to Large Reservoir Dams

Overview:

The failures of engineered dams that retain freshwater or mining waste (tailings) have led to socioeconomic and environmental consequences. However, the global magnitude-frequency statistics of these occurrences remain poorly quantified, out-of-date and/or limited in scope. Addressing this gap would give insight into how the hazard-risk of water-retention (reservoir) dams and mine tailings impoundments has evolved over time, and would provide quantitative benchmarks for estimating likelihoods of facility failures and their societal impacts to support risk assessments. In this study, we analyze new datasets and estimates of the construction and failures of large reservoir facilities (LRFs) and tailings storage facilities (TSFs) worldwide in the period 1965–2020. We address long-standing data gaps on LRF failures in China, and subsequently worldwide, and on constructed TSFs worldwide by adopting multiple estimation/extrapolation approaches to illustrate the range of uncertainty in our results. The total number of LRF failures is estimated to have been between 394 and 608. The annual numbers of newly constructed and failed LRFs declined near-proportionally, thus the cumulative failure rate of LRFs stayed fairly constant, falling in the range of 1.2% to 1.8% as of end-of-2020. The rate drops to at least 0.7% when excluding China. The cumulative fatality rate of LRFs reduced over time to 1.2 deaths per constructed facility, and falls in the range of 64 to 98 deaths per failure, as of end-of-2020. Failures of LRFs with very high storage capacities (>200 M m^3) have continued to occur since 2016. The annual number of TSF failures stayed relatively constant, whereas the annual construction rate of TSFs is estimated to have increased by ~3×, thus the cumulative failure rate of TSFs declined over time. When assuming our lower-estimate of the number of constructed TSFs (6,810), the cumulative failure rate is ~4.4% as of end-of-2020. When adopting our upper-estimate (20,230 TSFs), we obtain a rate of ~1.5%, which falls in the same order as the corresponding rate of LRFs. Our review of published estimates of existing TSFs worldwide indicates that the "true" rate is much lower than 4.4% and closer to 1.5%. The cumulative fatality rates of TSF failures reduced over time to 0.1–0.3 death per constructed facility and 6 deaths per failure as of end-of-2020, which are lower than those of LRFs. However, the size and the environmental impact of TSF failures have increased on average worldwide, especially since 2014. The rising global rate of failed tailings volumes has been approximately proportional to the rising global rate of tailings production since the 1990s. Heavy rainfall events and intensifying precipitation patterns are statistically important causative variables for the failures of both LRFs and TSFs. This has implications for the design and management of storage capacity, freeboard, facility drainage and spillways under climate change conditions. Our results are applicable broadly on a global scale and are conditioned by uncertainties in the data and the methods used to address data gaps. To improve the robustness of future statistical analyses, a more

comprehensive public disclosure effort is necessary, particularly with respect to reservoir facility failures in China and constructed TSFs worldwide.

2.1 Introduction

2.1.1 Preamble

Engineered impoundments encompass vast spatial footprints in modern landscapes. Globally, these constructed impoundments predominantly include: (i) water-retention (or reservoir) dams built for irrigation, flood control (disaster mitigation), water supply, hydroelectric power generation and other purposes (Baxter, 1977; ICOLD, 2020) and (ii) tailings storage facilities (TSFs) built to contain fine-grained, wet, environmentally hazardous sediments that are waste by-products of mineral extraction operations (Vick, 1983; Blight, 2010; Rana et al., 2021a). The number of human-made reservoir dams worldwide likely exceeds 200,000 (USACE, 2019; Ge et al., 2020; ICOLD, 2020). Over 50,000 of these reservoir dams are reported in the International Commission on Large Dams (ICOLD, 2020) Registry based on their conformance to the ICOLD definition of "large" dams – i.e. those with storage capacities of ≥3 M m^3 or dam heights of ≥15 m. These are referred to as large reservoir facilities (or LRFs) in this study. The historical number of constructed TSFs worldwide is in the order of 30,000 according to the World Mine Tailings Failures organization (WMTF, 2020). Projects requiring the construction of water-retention dams and TSFs have spurred socioeconomic development worldwide since the 1950s, commensurate with the rising demand for energy production, food, mineral resources and other benefits (e.g. Mudd and Boger, 2013; Dutta et al., 2016; Di Baldassare et al., 2018; Boulange et al., 2021; Vaidya et al., 2021).

The principal requirement for engineers when designing and operating water-retention dams and TSFs is to ensure their safety throughout the project duration and lifecycle. The performance of these facilities poses a global engineering challenge (Oliver-Smith, 1996; Santamarina et al., 2019; ICMM et al., 2020), especially since they are important strategic components in the emerging green/renewable economy and continued efforts for sustainability (e.g. Dutta et al., 2016; Herrington, 2021; Vaidya et al., 2021). However, catastrophic impoundment failures have resulted in high costs and long-lasting environmental harm over extensive inundation areas (Kossoff et al., 2014; Bowker and Chambers,

2015, 2017; Liu et al., 2019; Association of Dam Safety Officials, 2020, 2021; de Lima et al., 2020; Latrubesse et al., 2020; Ghahramani et al., 2020; Zhong et al., 2021; Rana et al., 2021a; Yu et al., 2021; Xiao et al., 2022). Satellite images of some recent high-profile failures are shown in Figure 2-1. Weather hazards are among the most frequent failure causes of recent cases (e.g. Robertson et al., 2019; ASDSO, 2020, 2021; Latrubesse et al., 2020; Rana et al., 2021a; Yu et al., 2021; Xiao et al., 2022). Floods from failed reservoir dams have caused over 40,000 deaths worldwide since 1965 (Zhang et al., 2016). Over the same period, at least 2,000 people have died due to TSF failures (Santamarina et al., 2019; Islam and Murakami, 2021; WISE, 2021), including 272 in the 2019 Feijao disaster in Brazil (Figure 2-1d; de Lima et al., 2020). These events have raised concern about the long-term safety of engineered impoundments, especially against the backdrop of a changing climate, heightened public pressure (i.e. lower societal tolerance for dam-related adverse consequences), growth in exposed population and increased resource demands (Oliver-Smith, 1996; Guthrie, 2015; Best, 2019; Santamarina et al., 2019; ICMM et al., 2020; Herrington, 2021).

The global performance and safety of engineered dams can be inferred by examining historical magnitude-frequency statistics of their failures and societal (human and environmental) consequences. However, the last review of failure statistics for large reservoir dams (global excluding China) was published over two decades ago (without an assessment of magnitude) and requires updating (Foster et al., 2000). The spatiotemporal distributions of TSF failures have been explored in histograms (Azam and Li, 2010; Lyu et al., 2019; Piciullo et al., 2022) whereas Rana et al. (2021a) estimated the recurrence rate of tailings flows (i.e. downstream mass flows from failed TSFs), but these studies did not cast a wider focus on TSF failure and fatality rates and the mean return periods of TSF failure consequences. Thus, at present, the failure statistics of both reservoir dams and TSFs remain poorly quantified, out-of-date and/or limited in scope.

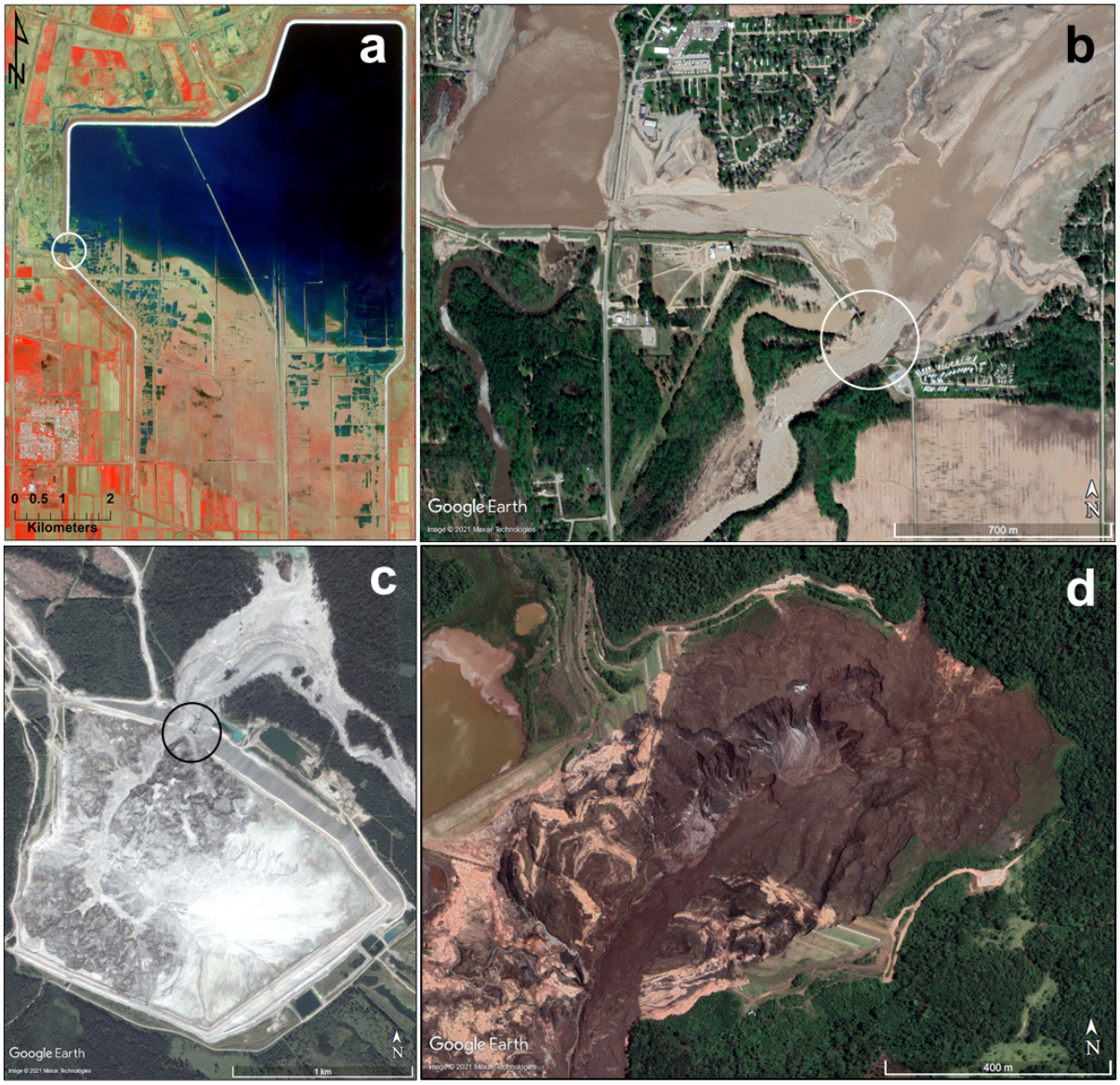

Figure 2-1: Examples of recent high-profile failures of large reservoir facilities (LRFs) and tailings storage facilities (TSFs). [a] 2020 Sardoba LRF failure in Uzbekistan (e.g. Xiao et al., 2022). [b] 2020 Edenville LRF failure in Michigan, USA (e.g. ASDSO, 2021). [c] 2014 Mt. Polley TSF failure in Canada (e.g. Morgenstern et al., 2015). [d] 2019 Feijao TSF failure in Brazil (e.g. Robertson et al., 2019; de Lima et al., 2020). The circles in a-c indicate the dam breach locations. Raw image data courtesy of Planet Inc., Google and Maxar.

The sparsity of comprehensive magnitude-frequency investigations to date is mainly attributed to pre-existing failure datasets that are incomplete or contain poorly constrained data, and in the case of TSFs, the substandard public disclosures of facility construction (Franks et al., 2021). The resulting

lack of up-to-date, normalized statistical insights hinders a clear understanding of how the hazard-risk of water-retention dams and TSFs has evolved over time, which may lead to public misconceptions and inconsistent claims. For example, the assumption of lower-estimates of the number of TSFs worldwide (e.g. ~3,500 in the year 2000; Davies et al., 2000; Davies, 2002) and/or underestimates of the number of water-retention dams worldwide has contributed to interpretations that the failure rate of TSFs is several times (Lyu et al., 2019; Liang et al., 2021) or more than two orders of magnitude (Azam and Li, 2010) greater than that of water-retention dams. However, when assuming the newer, upper-estimate range of 29,000-35,000 TSFs as proposed by WMTF (2020) based on national inventories and published estimates (e.g. ~12,000 TSFs in China alone; Pan et al., 2014; Wang et al., 2017), these interpretations fail to hold true. These inconsistencies underscore the challenges associated with obtaining best-estimates of the "true" global failure rates of both types of engineered structures, as well as the detailed effort necessary to arrive to more representative conclusions that could ultimately clarify existing misconceptions, advance scientific knowledge and benefit future risk assessments.

2.1.2 Scope and objectives

In this study, we undertake a comprehensive review of existing databases and published estimates and use statistical extrapolation and estimation approaches to fill data gaps, in an effort to illustrate the range of uncertainty in our results depending on the underlying assumptions. The objective in performing a detailed review of literature is to evaluate how statistically complete the construction and failure records are, which is an important first step to conducting magnitude-frequency and failure rate assessments. This step enables us to simplify our analysis by focusing on a specific study period for which the required data can be procured or reasonable estimates can be generated. In this case, we select a 56-year period of 1965-2020 due to the insufficient documentation of pre-1965 construction and failures.

We then assess the magnitude thresholds above which the record is either complete over the study period, or for which the degree of incompleteness can be addressed by means of approximation. Here, we find that focusing on LRFs only, rather than all reservoir facilities (for which the construction record is largely incomplete), is the most practical approach due to the recently issued ICOLD Registry. The number and magnitude of LRF failures can be extracted or minimum-estimated from U.S. Bureau of Reclamation (USBR, 2015) and Zhang et al. (2016) for all countries except China, for which

documentation on reservoir facility failures is sparse due to confidentiality. This issue was also encountered by Foster et al. (2000), who consequently excluded China from their statistical analysis. To build on preceding work, we present two "global estimation scenarios" (an upper-estimate and a lower-estimate) based on our analysis of historical frequency numbers of reservoir dam failures in China that are tabulated in Ge et al. (2020).

With regards to the number of constructed TSFs and the volume of produced tailings, the lack of an international registry necessitates us to present two "global extrapolation approaches" based on our analysis of mineral commodity production data (British Geological Survey, 2021; U.S. Geological Survey, 2021): one produces an upper-estimate closer to that proposed by WMTF (2020) and the other produces a lower-estimate that aligns with interpretations in previous studies (e.g. Davies et al., 2000; Davies, 2002; Lyu et al., 2019; Liang et al., 2021). The historical number of TSF failures worldwide is relatively well-documented in published datasets (e.g. ICOLD, 2001; Bowker and Chambers, 2015; Rana et al., 2021b; Islam and Murakami, 2021; Piciullo et al., 2022).

By these advances, we are able to set the central goal of our statistical analysis: to quantify the magnitude-frequency distributions and the cumulative failure and fatality rates of LRFs and TSFs worldwide in the period 1965-2020. The specific research questions that we seek to address by our statistical analysis to gain novel insights are as follows:

1. How has the construction and failure frequency of LRFs and TSFs evolved over time?
2. What are the statistically important causative variables of LRF and TSF failures worldwide?
3. How often have LRF and TSF failures above a specified magnitude (e.g. storage capacity, released volume, inundation area or fatality) occurred?
4. How have the average size of LRF and TSF failures and the environmental consequences of TSF failures evolved over time?
5. What fraction of the globally impounded freshwater and tailings has undergone failure/release over time?
6. What are the global cumulative failure and fatality rates of LRFs and TSFs as of end-of-2020 depending on our different statistical approaches and assumptions?

7. What are the implications of our statistical findings with respect to previous studies and interpretations and to the engineering properties and global safety of LRFs and TSFs?

Our global analysis focuses on LRFs and TSFs as distinct engineered structures, which allows a comparison between their failure statistics. At the outset, it should be emphasized that there are fundamental differences in their engineering characteristics (Section 2.2) and the sources of their data uncertainties (Section 2.3 and Section 2.4), all of which ultimately influence their respective failure statistics and corresponding implications (Section 2.5 and Section 2.6). Our analysis does not consider failure risk from war/terrorism damage and excludes levees and dikes – i.e. engineered structures that prevent coastal and riverine flooding. For clarity and consistency, Table 2-1 provides a list of key terminologies relevant to this study and

Table 2-2 provides the definitions of the various statistical metrics and variables that are used or analyzed in this article.

Our findings are presented with cognizance of the uncertainties that underlie the collated data, the methods used to address existing data gaps, and our statistical estimates. Notwithstanding these limitations, our work is timely and sheds light on the background hazard-risk of engineered impoundments worldwide. Moreover, the data and results presented herein benefit industrial practice by providing benchmark values to support future risk assessments.

Table 2-1: Key terminologies relevant to this study.

Term	**Definition**
Engineered impoundment	A constructed facility that stores waste materials produced by anthropogenic operations and/or materials (including freshwater) that serve societal interests behind an engineered dam. The present study focuses on the two most common types of engineered impoundments worldwide: reservoir facilities and tailings storage facilities.
Engineered dam	A constructed embankment that stores materials (in this study, freshwater or mine tailings) to serve community interests and/or to protect humans, the environment and built infrastructure from potential chemical and physical consequences. The present study distinguishes the engineered dam to be an individual component of an engineered impoundment. Levees and dikes – structures that prevent coastal and riverine flooding – are not analyzed in this study.
Reservoir facility	An artificial impoundment of water (predominantly freshwater) with one or more socio-economic purposes, such as hydroelectric power generation, irrigation, flood control, water supply, fish farming or recreation. The present study considers a reservoir facility to include the constructed dam, the impounded reservoir and appurtenant structures.
Continued from Table 2-1	
Large reservoir facility (LRF)	A large reservoir facility (LRF) is based on ICOLD (2020) that defines large dams as those with storage capacities of ≥3 M m^3 or dam heights of ≥15 m. The storage capacity is defined to be the reservoir capacity during normal operation.
Reservoir	The water (predominantly freshwater) that is artificially impounded by an engineered dam.
Reservoir dam	Also referred to as a “water dam” or “water-retention dam”. A reservoir dam is a constructed embankment that is designed to have sufficient geotechnical integrity to impound water. A reservoir dam may be built across a river channel (to block and control river flow) or across the outlet from a natural lake.
Tailings	Waste by-products of mineral extractive operations that commonly comprise fine-grained, wet, environmentally hazardous sediments. The present study considers tailings to include both tailings solids and interstitial (pore) water.

Tailings storage facility (TSF)	An impoundment of tailings and, in many cases, supernatant ponds (i.e. wastewater). The present study considers a TSF to include the tailings dam, the impounded tailings, the ponded water and appurtenant structures.
Tailings dam	A constructed embankment that is designed to have sufficient geotechnical integrity to impound tailings in perpetuity. Tailings dams are evolving structures that are raised in the upstream, downstream and/or centerline direction in successive stages over several years to decades, depending on the project and site characteristics.
Failure	An accident at an engineered impoundment that suddenly threatens the safety of the exposed environment, infrastructure and/or communities. A failure may include a dam breach, overtopping or seepage without the onset of a breach, or a collapse of an appurtenant structure (e.g. 2017 Oroville LRF spillway collapse in the USA).
Breach	The physical process of a dam losing its geotechnical integrity in the form of a partial or complete collapse resulting in an outflow of impounded materials (e.g. all cases in Figure 2-1).
Failure cause	One or more variables that acts as a trigger for the failure of an engineered impoundment. Examples include heavy rainfall, seismic activity or improper design/construction methods.
Failure mode	The primary mechanism by which the facility failure occurred. Examples include overtopping of the dam, internal erosion (piping) of the dam core or shear failure of foundation soils.

Table 2-2: Definitions of key statistical parameters and variables that are used or analyzed in this study.

Variable	Definition
Study period	The period that the statistical analysis corresponds to. The study period in our analysis is the 56-year period between January 1, 1965 and December 31, 2020. This study period corresponds to the assessment of both construction and failures.
Magnitude-frequency	A plot characterizing the annual frequency of events (i.e. the number of events per annum) with a magnitude X or greater. The magnitude can be represented by storage capacity, released volume, inundation area, fatality or other variables.
Mean return period	The average period of time (expressed in years in this study) between successive events, calculated as the inverse of the estimated annual frequency.
Annual failure rate	The number of failures per the number of constructed facilities worldwide in a given year, expressed as a percentage.
Cumulative failure rate	The cumulative number of failures per the cumulative number of constructed facilities worldwide at the end of a given period, expressed as a percentage. In this study, we investigate the cumulative failure rates over the study period (failures and construction from 1965 onwards) and report them based on end-of-2020 estimates. This means that, in our study, the annual failure rate and the cumulative failure rate are the same only in the year 1965.
Cumulative fatality rate	The cumulative number of fatalities due to facility failures per (i) the cumulative number of constructed facilities or (ii) the cumulative number of facility failures worldwide at the end of a given period, expressed as a percentage. In this study, we investigate the cumulative fatality rates over the study period (fatalities, failures and construction from 1965 onwards) and report them based on end-of-2020 estimates.

2.2 Characteristics of engineered impoundments

2.2.1 Reservoir facilities

Reservoir dams are designed to retain water, predominantly freshwater. In most cases, this is achieved by blocking and controlling the flow of rivers. Another common approach is to build a dam across the outlet from a natural lake. The societal functions for reservoir facilities include irrigation, hydroelectric power generation, flood control, water supply, fish farming, recreation and/or navigation (Baxter, 1977; ICOLD, 2020). Large reservoirs occupy volumes in the order of tens to hundreds of M m^3 (e.g. Figure S11 in the Supplementary Article). These projects may result in large-scale modification of natural landscapes, including catchment-wide implications for surface water and groundwater flow patterns, sediment flux and ecological diversity (e.g. Baxter, 1977; Yang et al., 2014;

Fan et al., 2015). Globally, reservoirs are estimated to represent >15% of the annual global runoff, and the surface area of formerly terrestrial habitat now inundated by reservoirs is about the size of France (Nilsson, 2009) or the entire Columbia River drainage basin in North America.

The engineering design of reservoir facilities depends on the terrain, project requirements and economic feasibility. There are two principal kinds of reservoirs: storage reservoirs that sustain flow into the downstream channels and control ordinary fluctuations in discharge, and run-of-the-river impoundments that balance daily variations in the demands for water-to-electricity production by generating hydraulic head (Nilsson, 2009). Foster et al. (2000) identified 12 distinct zoning categories for 11,192 large reservoir dams built worldwide (except China and pre-1930 Japan) up to 1986. Earthfill dams (homogeneous earthfill, zoned earthfill, concrete-face earthfill or earthfill with filter) comprised the majority with about 70%, whereas rockfill dams (zoned earthfill and rockfill dams, concrete-face rockfill dams and central core earth and rockfill dams) made up over 20%.

The potential failure modes for LRFs predominantly include dam overtopping and internal erosion (Foster et al., 2000; Zhang et al., 2016) and in some cases spillway collapse (e.g. Independent Forensic Team Report, 2018) or seismic/static liquefaction of the dam (e.g. Seed et al., 1975; ASDSO, 2021).

2.2.2 Tailings storage facilities (TSFs)

TSFs are designed to store wet, environmentally hazardous sediments (tailings) that comprise the waste by-products at mineral extraction and processing sites, including some industrial and power plant operations (Blight et al., 2010). In some cases, ponded wastewater (supernatant ponds) may overlie the tailings to prevent the generation of surface dusts and acidic drainage and provide reclaim water for use in the milling process. These materials are impounded by a compacted dam that is typically built out of waste earthfill/rockfill, foundation fill and/or the coarse fraction of tailings (Blight, 2010). By convention, the tailings are hydraulically discharged into the facility over the duration of the project. On a much smaller scale, some mines produce and store dry-stack and filtered tailings with lower moisture contents (Williams, 2021). In contrast to dams built to contain reservoirs, tailings dams are progressively raised in stages for several years to decades, commensurate with the rising volume of stored tailings. The dam raise direction may be upstream (i.e. over the impounded tailings), downstream and/or vertically upwards (centerline method) (Blight et al., 2010; Kossoff et al., 2014). Some of the largest TSFs in the world have storage capacities of over 100 M m^3 and exceed 200 m in dam height

(Supplementary Figs. S17 and S18). The construction and permanent presence of large TSFs can have long-term implications for local/regional hydrological and hydrogeological regimes (e.g. Ochieng et al., 2010).

The spatial configuration of TSFs can be diverse depending on the site topography and the project design requirements. Some examples include: (i) fully enclosed impoundments on near-flat terrains (ring-dykes); (ii) partially enclosed impoundments on hillslopes; (iii) impoundments enclosed by a cross-valley embankment; and (iv) tailings storage in open-pit lakes. At any point in time, the status of a TSF is one of the following: (i) active (construction, disposal and other operations are ongoing at the TSF); (ii) inactive (construction, disposal and other operations have temporarily or permanently ceased at the TSF); (iii) closed (the TSF has been officially decommissioned to be reclaimed by responsible authorities and integrated into the landscape); or (iv) abandoned (the entire mine site has ceased all operations and projects, and the site has been abandoned). The monitoring and management of legacy tailings is a topical contemporary issue due to the growing number of abandoned mine sites worldwide (e.g. Courchesne et al., 2021; Gauthier-Manual et al., 2021).

The potential failure modes for TSFs commonly include seismic or static liquefaction of the impounded tailings, dam overtopping by ponded wastewater, internal erosion of the dam core and embankment instability (ICOLD, 2001; Lyu et al., 2019; Rana et al., 2021a; Piciullo et al., 2022).

2.3 Review of datasets and statistical studies

Table 2-3 lists published datasets of the construction and failures of reservoir facilities and TSFs, with reference to the study authors, dataset focus and purpose, temporal coverage and number of entries. Table 2-4 lists recently developed geospatial datasets of reservoir facilities and mine infrastructure (including TSFs).

Table 2-3: A list of published datasets of constructed engineered impoundments and failure incidents that were consulted in this study.

Dataset	Dataset scope	Dataset purpose and application	Number of entries	Temporal coverage
Construction and failures of reservoir facilities				
MacDonald and Langridge-Monopolis (1984)	Reservoir facility failures	Review and empirical modelling of breach	42	1853-1975
Costa (1985)	Reservoir facility failures	Review and empirical modelling of breach	31	1864-1983
Wahl (1998)	Reservoir facility failures	Review and empirical modelling of breach	108	1864-1996
USBR (2015)	Reservoir facility failures	Case summaries and fatality analysis	60	1874-2011
Zhang et al. (2016)	Reservoir facility failures	Review and statistical analysis of failure modes	1,443	1799-2005
USACE (2019)	Reservoir facilities constructed over time in USA	Catalog of U.S. dams with relevant information	91,470	1640-2018
Zhong et al. (2020)	Reservoir facility failures	Empirical modelling of breach-discharge	162	1889-2004
ICOLD (2020)	Reservoir facility construction	Comprehensive register with relevant data	58,713	130 A.D. – 2020
Construction and failures of tailings and mine waste storage sites				
ICOLD (2001)	TSF failures	Comprehensive database	221	1917-1996
Rico et al. (2008b)	Tailings flows	Empirical modelling of outflow-runout	29	1965-2000
Bowker and Chambers (2015)	Mine waste failures	Case collection and documentation	226	1917-2009
Small et al. (2017)	Mine waste flows	Case collection and outflow-runout data	79	1928-2015
Larrauri and Lall (2018)	Tailings flows	Empirical modelling of outflow-runout	35	1965-2015
Ghahramani et al. (2020)	Tailings flows	Empirical modelling of inundation area	33	1965-2019
Franks et al. (2021)	TSF construction	Incomplete disclosure of publicly-owned TSFs	1,743	1817-2020
Rana et al. (2021a)	Tailings flows	Comprehensive database and empirical modelling of outflow-runout	63	1928-2020

Continued from Table 2-3

Rana et al. (2021b)	Mine waste failures	Case collection only	362	1915-2020
Islam and Murakami (2021)	Mine waste failures	Comprehensive database with mapped surface areas	366	1915-2020

Table 2-4: Publications (pre-2022) on geospatial mapping of reservoir facilities and mine infrastructure (including TSFs).

Study	Results
Reservoir facilities	
Mulligan et al. (2020)	Global geospatial database of >38,000 mapped and georeferenced reservoir facilities.
Mine sites and tailings storage facilities (TSFs)	
Maus et al. (2020)	Global geospatial database of mapped surface areas of mine areas (active and abandoned) using automated algorithm on Google Earth Engine, resulting in >21,000 polygons across >6,000 mine sites.
Werner et al. (2020)	Geospatial database of mapped surface areas of 295 mine areas (TSFs, open-cut pits, water storage ponds, mill infrastructure, waste rock piles, and heap leach pads), totaling to >3,700 polygons.
Tang et al. (2020)	Geospatial mapping and place-marking of 5,189 TSFs in China.
Liang et al. (2021)	Update to Maus et al. (2020), resulting in >24,000 mine area polygons across 117 countries.
Islam and Murakami (2021)	Surface areas of 134 historically failed TSFs were mapped on Google Earth.

2.3.1 Reservoir facilities

2.3.1.1 Constructed facilities

The ICOLD Registry is the authoritative source on constructed LRFs worldwide. The most recent update to the Registry was issued in April 2020 (ICOLD, 2020) and includes 58,713 reservoir dams built up to 2020, with the addition of a few under-construction and planned projects. The ICOLD Registry was specifically aimed towards large reservoir dams that impound ≥3 M m^3 or those with heights of ≥15 m for completeness, although the time period for statistical completeness was not stated.

With the exception of some of the largest reservoir facilities in the world, the vast majority of the reservoir facilities listed in the Registry are impounded behind a single dam.

The true number of all reservoir facilities worldwide is unknown, but is likely to exceed 200,000, with ~91,500 catalogued in the USA (USACE, 2019; DeNeale et al., 2019) and ~98,000 present in China (Ge et al., 2020). Aside from reservoir facility name, owner and geographic location (town and country), the ICOLD Registry reported additional data at varying degrees of completeness, including dam type (87% of listed reservoir facilities), year of construction (88%), dam height (99%), dam crest length (60%), storage capacity (97%), societal purpose (67%), surface area (44%) and catchment area (29%). Mulligan et al. (2020) employed an automated algorithm on Google Earth Engine to develop a global geospatial database (with reported coordinates) of over 38,000 reservoir dams from the ICOLD Registry (Table 2-4).

2.3.1.2 Failures

The compilation of reservoir facility failures has mainly been intended for empirical modelling of breach-outflow (MacDonald and Langridge-Monopolis, 1984; Costa, 1985; Wahl, 1998; Zhang et al., 2020). Zhang et al. (2016) provided the largest open-access dataset of reservoir facility failures to date, although with limited quantitative data and with the exclusion of all failures in China. Of the 1,443 events in the period 1799-2005 that were included, 316 cases (22%) reported storage capacity, 100 cases (7%) reported the volume above the breach invert (i.e. volume of released water), 1,105 cases (77%) reported dam height, 570 cases (40%) reported the year of construction of the failed reservoir facility, 171 cases (12%) reported some description of breach geometry (breach height or width), 62 cases (4%) reported peak outflow and 1,217 cases (84%) reported failure causes or failure modes.

The documentation of reservoir facility failures in China remains publicly inaccessible. However, Ge et al. (2020) provided a summary table of the number of reservoir facility failures in China in the period 1954-2014. They reported a total of 3,529 failures, with the annual national average calculated to be 65 in the period 1954-1965, 152 in the period 1966-1976, 43 in the period 1977-1999 and 5 in the period 2000-2014. When adding 1,443 reservoir facility failures worldwide excluding China (Zhang et al., 2016) and 3,529 reservoir facility failures in China (Ge et al., 2020), this equals to almost 5,000 reservoir facility failures that have been reported worldwide in recorded history.

Foster et al. (2000) conducted a detailed statistical assessment of the failures of large reservoir dams constructed up to 1986, excluding those in Japan pre-1930 and in China due to poor data availability. They estimated a failure rate (number of failures / number of constructed facilities, expressed as a percentage) of 1.2% for 11,192 large embankment dams. Foster et al. (2000) and Zhang et al. (2016) observed that the highest statistical likelihood for the failure of a reservoir facility is within the first 5 years of operation, although the exact proportion varies depending on the failure mode and the dam zoning category. The two studies also noted that overtopping and internal erosion (i.e. piping) together comprise 80-90% of reported failure modes. This range is higher than the ~62% proposed by Costa (1985) – the other failure mode being foundation defects – based on an analysis of failure modes of large reservoir dams listed in the ICOLD Registry that was issued at the time.

2.3.2 Tailings storage facilities (TSFs)

2.3.2.1 Constructed facilities

The total number and the complete geographic distribution of TSFs worldwide remains unknown due to the lack of public disclosure requirements. Davies et al. (2000) and Davies (2002) first suggested a "conservative, lower-bound" estimate of ~3,500 TSFs worldwide at the turn of this century. This estimate was cited in Lyu et al. (2019) and Liang et al. (2021). Azam and Li (2010) suggested a much higher figure of 18,401 mine sites worldwide, but the original basis of this estimate was not stated.

WMTF (2020) compiled national inventories and published estimates of TSFs from 22 countries that represent ~60% of the global mineral commodity production in the modern era. They calculated a total of 19,214 TSFs that have been historically constructed in these 22 countries, including active, inactive, closed and abandoned sites. Approximately 12,000 of these TSFs are in China alone (Pan et al., 2014; Wang et al., 2017), ~43% of which have been georeferenced by Tang et al. (2020). By scaling up the summed number (19,214) to include the other countries that make up the remaining 40% of global mineral production, WMTF (2020) obtained a global estimate of ~32,000 TSFs. When assigning ±10% uncertainty boundaries, they proposed that the actual number of existing TSFs worldwide is between 29,000 and 35,000. This estimated range is similar to that proposed in Spencer et al. (2021), who reviewed existing national inventories and reported a minimum-estimate of 13,400-15,400 active and inactive TSFs, and remarked that the true number of all existing TSFs worldwide could be in the order

of ~30,000 when accounting for abandoned TSFs and poorly reported countries. This suggests that the estimate of 18,401 mines by Azam and Li (2010) may not have been an acute underestimate, especially if one assumes a global average of ~1.7 TSFs per mine site.

Franks et al. (2021) presented the first disclosure database of TSFs. Their data is derived from a detailed, but statistically incomplete, record collated through voluntary survey responses from publicly-listed mineral extraction companies. The final database included 1,743 TSFs (~6% of the historical estimate by WMTF, 2020), 725 of which were active as of 2021. Information such as geographic coordinates, primary commodity, status, construction commencement year, dam raise type, storage capacity and hazard classification were all reported. Their database did not include TSFs that are abandoned, state-owned, privately-owned, small-scale and/or artisanal, as well as tailings that are not stored in conventional impoundments (e.g. heap leach pads that can extract copper, nickel, uranium and precious metals by chemical processes via emplaced liners). Moreover, only four TSFs situated in China were reported in the disclosure database. By proportionally extrapolating data in comparison to global mineral commodity production (using calculated mean stored volume per TSF per primary commodity), Franks et al. (2021) estimated that the total number of active TSFs worldwide is ~3,400, and a "conservative, lower-bound estimate" for the total number of active, inactive and closed TSFs (excluding abandoned sites) is ~8,100.

Both WMTF (2020) and Franks et al. (2021) reported an acceleration in the global annual production of mineral commodities, which they interpret to correspond to an accelerating number of TSFs constructed worldwide. They also suggested that TSFs may be increasing in storage volume on average, although temporal box-and-whisker plot distributions of the incomplete disclosure data provided in Franks et al. (2021) does not provide evidence for this suggestion (see Figure S16 in the Supplementary Article).

As listed in Table 2-4, recent global-scale remote sensing studies have mapped over 6,000 mine sites, with the total surface area of remotely visible mines measured to be 57,300 km^2 by Maus et al. (2020) (about the size of Croatia) and 31,400 km^2 by Liang et al. (2021) (about the size of Belgium). The significant difference between the two measurements is attributed to the difference in mapping methodologies and associated uncertainties. Werner et al. (2020) mapped distinct mine infrastructure features in 295 publicly-owned active mines and measured the total surface area of 267 TSFs across

these select mines to be 918 km^2. Tang et al. (2020) mapped 5,189 TSFs in China and obtained a total surface area of 1,884 km^2. By extrapolation from a mine area vs. TSF area relationship using data from Werner et al. (2020) and our independent geospatial mapping, the global surface area of TSFs is estimated to be in the order of 5,400 km^2 (see Section 2.4 and Figure S22 in the Supplementary Article).

2.3.2.2 Failures

Compared to reservoir facility failures, the number of TSF failures over time is relatively well-documented. ICOLD (2001) developed the first global database of 221 TSF failures that occurred in the period 1917-1996. They reported, at varying degrees of completeness, the mine location (state/province and country), dam raise method, dam height, impoundment volume, failure mode, ore type, released volume and runout distance. ICOLD (2001) also included incident summaries that provided context to each event. Bowker and Chambers (2015) built on ICOLD (2001) by documenting 226 events in the period 1917-2009.

Rico et al. (2008b) extracted data from ICOLD (2001) to develop the first empirical relationship between released volume and runout distance, which was later updated by Larrauri and Lall (2018). Ghahramani et al. (2020) used geographic information systems (GIS) methods and satellite imagery to map 33 tailings flows that occurred since 1965. They developed an empirical relationship between released volume and the "primary impact zone", or Zone 1, planimetric inundation area. Rana et al. (2021a,b) subsequently produced a database of 63 tailings flows that occurred in the period 1928-2020 by screening an inventory of 362 mine waste failures since 1915 based on some case selection criteria. They presented empirical relationships between released volume and GIS-mapped runout distance, inundation area and travel path angle in Zone 1 according to the travel path channelization. They also performed a cumulative volume-frequency analysis of tailings flows and estimated that tailings flows with volumes of ≥1 M m^3 have occurred once every 2-3 years worldwide since 1965, on average.

The pre-existing datasets that were consulted by Rana et al. (2021b) included Bowker & Chambers (2015) and Small et al. (2017), which predominantly focused on tailings, but also included some incidents involving other categories of mine waste, such as coarse colliery waste (e.g. 1972 Buffalo Creek, USA; Davies et al., 1972; Wahler & Associates, 1973), spoil tips (e.g. 1966 Aberfan, Wales; Bishop, 1973) and waste rock piles (e.g. 1985 Quintette, Canada; Dawson et al., 1998). This mixing of mine waste storage types was also observed in the supplementary data of Islam and Murakami (2021).

This inconsistency is likely due to the vague definitions of tailings and TSFs in literature preceding Franks et al. (2021) and Rana et al. (2021a,b), and the sparsity of reported information on low-consequence failure events.

Franks et al. (2021) noted that ~10% of disclosed TSFs had experienced a "stability concern" at some point in the operational history (note that a "stability concern" does not necessarily equate to failure). Historical statistics indicate that the number of TSF failures increased dramatically in the 1960s (commensurate with the global increase in mining activity) and that the failure rate has stayed relatively constant since then at around 35-50 events per decade (Lyu et al., 2019; Piciullo et al., 2022). The proportional contribution of weather hazards and drainage mismanagement as failure causes has increased over time, whereas that of seismic activity has reduced (Azam and Li, 2010; Rana et al., 2021a). The failure modes are statistically well-distributed among overtopping, internal erosion and seismic/static liquefaction (Lyu et al., 2019; Piciullo et al., 2022).

Azam and Li (2010) suggested a historical, global failure rate of 1.2% for TSFs (i.e. 220 failures / 18,400 mines, assuming 1 TSF per mine). Halabi et al. (2022) assumed the global estimate of 18,400 existing TSFs and calculated the annual probability of TSF failures to have been in the order of 10^{-4} on average in the period 1950-2019. They also estimated a historical casualty rate of 9 deaths per TSF failure on average for this period.

2.4 Methods

Magnitude-frequency analyses can generate statistical insights into the rates and impacts of recurring hazardous phenomena at regional to global scales. The approach was initially developed to analyze earthquake statistics (Gutenberg and Richter, 1954). Since then, the magnitude-frequency statistics of natural landslides and natural dams have been investigated as well (e.g. Guthrie and Evans, 2004; Evans, 2006, Hungr et al., 2008; Evans et al., 2011; Carrivick and Tweed, 2016). When it comes to anthropogenic hazards such as LRFs and TSFs, the main challenge has been to establish data completeness for specified magnitude and time intervals. Addressing this gap would help generate reasonable estimates of mean return periods and cumulative failure and fatality rates. In turn, this would provide insight into how the hazard-risk of LRFs and TSFs has evolved over time, and would provide quantitative benchmarks for estimating the likelihoods of facility failures and their societal impacts to support future risk assessments.

Figure 2-2 shows a flowchart that summarizes our data collection, estimation and extrapolation approaches and the associated uncertainties that form the basis of our statistical analysis of the failures of LRFs and TSFs. The main challenge in the effort to estimate cumulative failure and fatality rates (Section 2.5.3) is the data sparsity on the number of LRF failures in China (Section 2.4.3.2) and the number of constructed TSFs worldwide (Section 2.4.4). We adopted different assumptions and scaling and extrapolation approaches to obtain upper- and lower-estimates for both variables, as described in the following sub-sections.

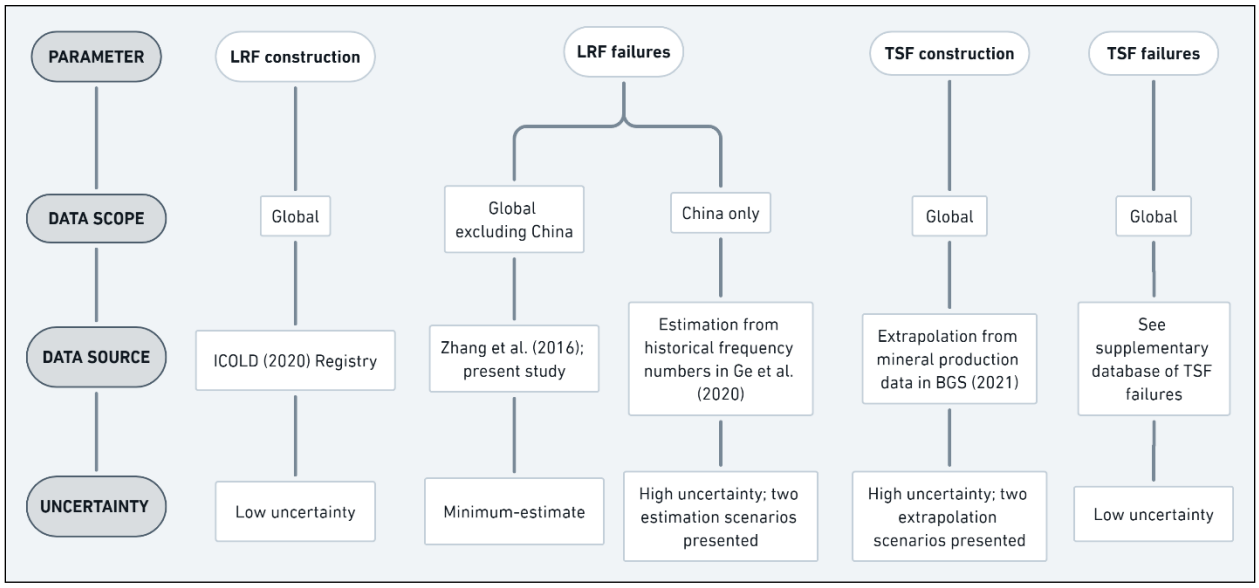

Figure 2-2: Flowchart summarizing the data collection, estimation and extrapolation approach and the associated uncertainties that form the basis of our statistical analysis of the failures of large reservoir facilities (LRFs) and tailings storage facilities (TSFs). The uncertainties in this study mainly stem from the data sparsity on LRF failures in China and constructed TSFs worldwide.

2.4.1 Study period

For our statistical analysis, we set 1965-2020 to be the consistent study period for the construction and failures of both LRFs and TSFs for two reasons. First, the mid-1960s was a critical moment for reservoir facility and TSF engineering following a number of disasters, including the 1963 Vajont LRF failure in Italy (ca. 2,000 deaths; Paronuzzi et al., 2013), the 1963 Liujiatai LRF failure in China (943 deaths; USBR, 2015), the earthquake-triggered failures of 17 TSFs in 1965 in Chile (>200 deaths

caused by the El Cobre tailings flow; Dobry and Alvarez, 1967; Rana et al., 2021a) and the 1966 Aberfan colliery spoil tip failure in Wales (144 deaths; Bishop, 1973). This era marked the beginning of increased public scrutiny and media reporting on reservoir facility and mine waste hazards, concurrent with changes in dam-impoundment design (e.g. the TSF upstream-raise method was banned in Chile). Second, high-resolution, global-coverage satellites were commenced in the mid-1960s with the advent of the Corona satellite, which has allowed for GIS-driven mapping and documentation of failure events and their downstream impacts that have occurred since then (e.g. Ghahramani et al., 2020; Rana et al., 2021a,b).

By restricting our statistical analysis to a specific study period, our approach slightly differs from that of Foster et al. (2000), who analyzed the failure rates and probabilities of large reservoir dams constructed worldwide (excluding China) up to a specific year (1986). A limitation of our approach is that some LRF and TSF failures over our study period of 1965-2020 occurred in facilities that were constructed before 1965; thus, the selection of the same study period for construction and failures yields some influence on our cumulative failure rate estimates (Section 5.3). However, our approach is the most practical option within the comprehensive scope of the present study (which considers LRF failures in China as well as a global analysis of TSFs) to circumvent issues related to the availability and quality of data preceding 1965. An important precondition to facilitating and simplifying our statistical analyses was to ensure a reasonable degree of dataset completeness and temporal consistency for both LRFs and TSFs. Setting a strict study period of 1965-2020 satisfied this objective.

2.4.2 Large reservoir facility (LRF) construction

Given that the ICOLD Registry is complete for constructed large reservoir dams, we filtered the dataset for LRFs built in the period 1965-2020, leading to a final number of 33,424 sites. We assumed that the number of large reservoir dams in the Registry (33,424) for 1965-2020 is roughly equal to the number of LRFs. In this study, it was not considered necessary to filter LRFs according to their construction methodology, as our aim was to inform how the two engineering sectors have evolved in their safety over time from a high-level statistical perspective.

As a check for uncertainty and error, we compared the number of LRFs constructed in the USA in the period 1965-2017 as listed in the ICOLD Registry to that in the U.S. National Inventory of Dams

(USACE, 2019), after filtering the latter by excluding dikes and levees. The number was only about 1% higher in USACE (2019), which indicates low uncertainty.

2.4.3 Large reservoir facility (LRF) failures

2.4.3.1 Global excluding China

We filtered the reservoir facility failure dataset in Zhang et al. (2016) for LRF failures. There were 316 listed reservoir facility failures in the period 1965-2005 and, based on available data on dam height and/or storage capacity, we verified that 113 of these events were LRF failures (69% in the USA). This is a global minimum-estimate as all reservoir facility failures in China are missing in this dataset and a majority of the listed cases (especially those in the USA) did not report dam heights and/or storage capacities (see Section 3.1.2). Only 16 of the 113 events (14%) reported the volume above breach invert (i.e. volume of released water), whereas storage capacity was reported for 65 of the 113 cases (58%). As a result, we chose storage capacity (rather than released volume) as a measure of magnitude for LRF failures in our analysis, but the extracted cumulative storage capacity data still represents a minimum-estimate.

We accessed USACE (2019) with the aim of filling in gaps on dam height and storage capacity to check if any of the other reservoir facility failures in USA with missing data can be classified as LRF failures. However, this effort was unsuccessful either because the failure site names listed in Zhang et al. (2016) did not match the constructed dam site names in USACE (2019) or there were multiple dam sites in USACE (2019) that shared duplicate/similar names.

2.4.3.2 Incorporating China

To incorporate LRF failures in China in our frequency analysis, we consulted Ge et al. (2020), who summarized the number of reservoir facility failures in China in the period 1954-2014. The authors did not supply the full database due to confidentiality. Instead, they tabulated the number of failures over the periods 1954-1965, 1966-1976, 1977-1999 and 2000-2014 based on the following storage capacity classes set by the Chinese Ministry of Water Resources: 0.1 to 1 M m^3 (small reservoir [2]); 1 to 10 M m^3 (small reservoir [1]); 10 to 100 M m^3 (medium reservoir); 100 to 1,000 M m^3 (large reservoir [2]); and ≥1,000 M m^3 (large reservoir [1]). This means that the classification of large dams by the Chinese Ministry of Water Resources is fundamentally different from that of ICOLD.

To estimate the annual numbers of LRF failures in China and subsequently worldwide (Section 2.5.1.1 and Section 2.5.3), we adopted two different approaches: Estimation Scenario 1 (a scaling-based approach that produces a lower estimate) and Estimation Scenario 2 (an assumption-based approach that leads to a higher estimate). The details for the two scenarios are as follows:

- **Estimation Scenario 1 (scaling-based approach that represents a lower-estimate)**: Using the given summary table in Ge et al. (2020), we obtained a sum of ~475 reservoir facilities with storage capacities of ≥1 M m^3 that failed in China in the period 1965-2014. The number of failures in the year 1965 was estimated by annual averaging over the given period of 1954-1965. To scale this failure number (475) to LRFs, we first calculated the global cumulative failure rate in the period 1965-2020 for LRFs with storage capacities of ≥10 M m^3, which we deduce to be a statistically complete interval. The obtained rate was 1.2% (101 failed / 8,313 constructed worldwide), which we assumed to approximately correspond to the global failure rate of all LRFs in the period 1965-2020. By trial-and-error, we scaled down 475 by a factor of 45% to match the failure rate of 1.2%. By this approach, we estimate that there were 261 LRF failures in China in the period 1965-2014 and 394 LRF failures worldwide in the period 1965-2020. To estimate the annual numbers of LRF failures in China, we applied the scaling factor of -45% to the annual averages derived from Ge et al. (2020) to estimate the following for LRF failures in China: 16.9 in 1965; 12.7 per annum in the period 1966-1976; 4.1 per annum in the period 1977-1999; and 0.8 per annum in the period 2000-2014.

- **Estimation Scenario 2 (assumption-based approach that represents an upper-estimate)**: We assumed that the sum of ~475 reservoir facilities with storage capacities of ≥1 M m^3 that failed in China in the period 1965-2014 (estimated from Ge et al., 2020) is approximately equal to the number of LRF failures in China in the same period. By this approach, we estimate that there were 608 LRF failures worldwide in the period 1965-2020, and that the annual average numbers of LRF failures in China were as follows: 30.8 in 1965; 23.0 per annum in the period 1966-1976; 7.4 per annum in the period 1977-1999; and 1.4 per annum in the period 2000-2014.

In our temporal and normalized rate analyses (Section 2.5.1.1 and Section 2.5.3 respectively), we presented the results as derived from both estimation scenarios to illustrate the range of uncertainty involved in the number of LRF failures in China and subsequently worldwide.

To explain how we incorporated storage capacities of failed reservoir facilities in China in our temporal and cumulative rate analyses, we provide the following example. According to Ge et al. (2020), there were 231 reservoir facility failures with storage capacities between 1 and 10 M m^3 in China in the period 1966-1976. Since the real failure data is not given in Ge et al. (2020) and the average is unknown, to provide a minimum-approximation, we assumed that all of the 231 reservoir facility failures had storage capacities of 1 M m^3 (the minimum of a given range), which would yield a total of >231 M m^3 in the period 1966-1976. For annual estimates, we averaged this total number over the corresponding time period (i.e. >21 M m^3 per annum in the period 1966-1976). The same approach was applied to the higher storage capacity ranges (10-100, 100-1,000 and ≥1,000 M m^3) and the other time intervals (1965, 1977-1999 and 2000-2014) given in Ge et al. (2020). This approach provides a minimum-approximation for a temporal analysis but is not appropriate for a cumulative volume-frequency plot, which requires real failure data.

By reviewing several publications (e.g. Graham, 1999; He et al., 2008; USBR, 2015; Mahmoud et al., 2020) and news articles, we extended our record of LRF failures to include an additional 18 events worldwide (excluding China) in the period 1965-2020 that are not found in existing datasets. For fatality-frequency analysis, we extracted life loss data from pre-existing datasets (Graham, 1999; He et al., 2008; USBR, 2015; Zhang et al., 2016; Mahmoud et al., 2020). For consistency in our analysis, we filtered the data for fatalities caused by the failures of LRFs. Our review suggests that this dataset is statistically complete for LRF failures worldwide that caused ≥10 deaths in the period 1965-2020. The failures of some smaller reservoir facilities have also caused human impacts, but these events were not considered in our fatality assessment.

As supplementary material, we provide a database of 36 LRF failures in the period 1965-2020 that were not included in previous global databases or had some gaps that we filled by our review or analysis. We identified the locations of 30 failure sites and supply GIS-processed, pre- and post-event satellite imagery for 19 cases, including the 1975 Banqiao and Shimantan dam breaches in China that collectively caused over 25,000 deaths (e.g. Xu et al., 2008; USBR, 2015). The database file also includes a summary calculation sheet for the two estimation scenarios.

2.4.4 Tailings storage facility (TSF) construction

Given the incomplete global documentation of existing TSFs, described in Section 2.3.2, a key objective and contribution of this study was to independently estimate the volume of produced (stored) tailings and the number of constructed TSFs worldwide in the period 1965-2020. This would help us estimate a more representative cumulative failure rate of TSFs. The most practical approach was to analyze global mineral commodity production data and use published commodity-to-tailings conversion factors, bulk densities, and mean and median storage volumes per TSF per primary commodity. This approach is a more detailed and modified extension to that of Franks et al. (2021) (see Section 2.3.2.1).

Global mineral production data are available as open-access annual reports issued by the British Geological Survey (BGS, 2021) and the U.S. Geological Survey (USGS, 2021). The reports date back to the 1910s, when it became necessary to document mineral resources for major nations partaking in wartime efforts. We reviewed the disclosure database in Franks et al. (2021) and the mine waste failure database in Islam and Murakami (2021) and selected 13 important commodities for statistical extrapolation: coal/fly ash, gold, silver, iron, phosphate, zinc, lead, copper, bauxite, aluminum, nickel, diamond and uranium. For our study period of 1965-2020, BGS (2021) contained data on all of these commodities, whereas USGS (2021) was missing data on coal/fly ash. We therefore relied on the BGS (2021) data for our analysis. As a check of uncertainty, we compared annual mineral production values in BGS (2021) versus USGS (2021) and found that the differences were consistently between <1% and 8%.

For each year in the period 1965-2020, we converted the metric tonnes of produced minerals to tonnes of produced tailings by applying multiplication factors provided in the supplementary information of Bullock et al. (2021). They developed these conversion factors based on the ore type in order to estimate and project the CO_2 removal potential from stored tailings worldwide. The conversion factors were not available for coal/fly ash and phosphate, for which we relied on Joshi and Lothia (1997), Ahmaruzzaman (2010) and Yang et al. (2017). We converted tailings mass to tailings volume by assuming mean bulk density values of specific types of tailings based on our review of numerous publications (e.g. Torres and Brito, 1966; Okusa et al., 1980; Matyas et al., 1984; McGregor and Blowes, 2002; Mecsi, 2013; Morgenstern et al., 2016; Robertson et al., 2019; Boshoff et al., 2018;

Jefferies et al., 2019; Behera et al., 2020; Zmemla et al., 2020; Feng et al., 2021; Rana et al., 2021b; Zubkova et al., 2021). By this approach, we estimated that the total volume of produced (and stored) tailings in the period 1965-2020 is just over 2.5×10^{11} m^3.

To estimate the number of TSFs constructed worldwide in the period 1965-2020, we adopted two approaches: the Mean Extrapolation method (which produces a lower-estimate that aligns with interpretations in previous studies) and the Median Extrapolation (which produces a higher and, by our deduction, a more representative estimate) method. The details are as follows:

Mean Extrapolation (lower-estimate): We divided the total volume of tailings produced each year by the "mean" volume per TSF of the corresponding primary commodity. These mean values were derived from the incomplete disclosure data in Franks et al. (2021). This approach led to a cumulative estimate of 6,810 TSFs constructed in the period 1965-2020. In comparison, by a similar mean extrapolation method, Franks et al. (2021) estimated 8,100 constructed TSFs throughout history and assigned this number as a "conservative lower-bound" for active, inactive and closed TSFs, but excluding abandoned TSFs.

Median Extrapolation (upper-estimate and potentially representative estimate): We divided the total volume of tailings produced each year by the "median" volume per TSF of the corresponding primary commodity. The median values were derived from the disclosure database in Franks et al. (2021). This approach led to a cumulative estimate of 20,230 TSFs constructed in the period 1965-2020.

There are two major reasons why we find that the median number is more representative. First, in terms of global proportions, the database in Franks et al. (2021) is over-represented by larger-volume TSFs that are commonly built and operated by publicly listed companies. In other words, TSFs that are of smaller sizes, artisanal, privately owned, situated in poorly reported regions (e.g. China, India, South America and Eastern Europe) or abandoned are not represented in Franks et al. (2021), yet their failures are included in global failure datasets (see supplementary data). This means that the calculated mean per-TSF-per-commodity storage volumes are more statistically influenced by the large TSFs than the calculated median in the disclosure data. For example, the mean storage volumes for copper, iron and phosphate TSFs were calculated to be unrealistically high at 91, 57 and 96 M m^3 respectively. In comparison, the median-derived values were 21, 13, and 29 M m^3 respectively. As such, the obtained

median values are concluded to be closer to the "true" global mean of per-TSF storage volumes for all commodities.

Second, as described in Section 2.3.2.1, WMTF (2020) estimated that 29,000-35,000 TSFs have been constructed historically worldwide (~12,000 in China alone; Pan et al., 2014; Wang et al., 2017) based on government inventories, published figures and extrapolation estimates. Similarly, Spencer et al. (2021) reviewed existing national inventories and reported a minimum-estimate of 13,400-15,400 active and inactive TSFs, and remarked that the true number of all existing TSFs worldwide could be in the order of ~30,000 when accounting for abandoned TSFs and poorly reported countries.

As such, we consider the median-derived number of 20,230 to be more representative, as it is closer to the WMTF (2020) estimate and is not likely a severe underestimate of the true number of TSFs worldwide. However, previous interpretations on the failure rates of TSFs have been based on lower-estimates of the number of TSFs worldwide (e.g. ~3,500 at the turn of the century; Davies et al., 2000; Davies, 2002; Lyu et al., 2019; Liang et al., 2021). In our temporal and cumulative rate analyses (Section 2.5.2.1 and Section 2.5.3), we showed the results based on both our Median Extrapolation and Mean Extrapolation methods. This allowed us to illustrate the differences in the calculated TSF failure and fatality rates when adopting our lower-estimate of the number of constructed TSFs (Mean Extrapolation), as done in previous studies, versus when adopting our upper-estimate and potentially representative estimate of the "true" number of constructed TSFs (Median Extrapolation).

We note that, for a few commodities (gold, silver, zinc and diamond), even the median storage volumes were higher than anticipated, which resulted in underestimated numbers of TSFs for these commodities. Given that some commodities were not considered in our extrapolation (e.g. cobalt, tin and platinum), our obtained estimate of 20,230 could still be a global underestimate, but it remains difficult to confirm this. We also cannot confirm whether the numbers of TSFs corresponding to the other commodities (coal/fly ash, iron, phosphate, lead, copper, bauxite, aluminum, nickel and uranium) are overestimates. There is also inevitable uncertainty in the multiplication factors provided in Bullock et al. (2021), especially for the four commodities that correspond to underestimated numbers of TSFs.

A summary worksheet of our calculations and a comparison between our extrapolation estimates and the incomplete disclosure data in Franks et al. (2021) are provided in supplementary material.

2.4.5 Tailings storage facility (TSF) failures

To develop an updated database of TSF failures, we first assembled all TSF and mine waste failures and relevant information that have been listed in preceding inventories (

Table 2-3), amounting to almost 370 cases between 1915 and 2020. We shortened the inventory to 328 cases in the period 1965-2020. As stated in Section 2.3.2.2, a limitation in previous datasets is that some incidents involving other categories of mine waste (spoil tips, coarse colliery waste and waste rock dumps) were categorized as TSF failures. The present study adopts a strict definition of tailings and TSFs (Table 2-1) that we applied when screening the 328 cases in the period 1965-2020. The incident summaries and remarks given in ICOLD (2001) and Islam and Murakami (2021) supported this effort by providing case-specific context. In addition, we geo-located ~85% of the TSF failure sites, thus building on the work of Rana et al (2021a,b) and Islam and Murakami (2021). This allowed us to verify the type of mine waste site and the failure extent by examining historical remote sensing imagery, such as Corona, Landsat, USGS aerial photographs, Sentinel, SPOT (on Google Earth), RapidEye, PlanetScope and Worldview (on Google Earth) (similar to the approach of Rana et al., 2021a,b).

Based on our definition and case selection/screening criteria, we identified 303 TSF failures that occurred in the period 1965-2020. This updated database is provided as supplementary data. The 303 TSF failures occurred across 249 distinct mine sites. Of the 303 TSF failures, we reported the geographic coordinates for 259 cases and the elevation/altitude for 215 cases. We assigned a qualitative measure of uncertainty to the reported location of each TSF: low uncertainty (high certainty that the failed TSF has been geo-located), medium uncertainty (the failed TSF is/was in the general vicinity of the place-marked mine site but its exact location cannot be identified) and high uncertainty (the failed TSF or the corresponding mine site could not be geo-located).

With the support of previous datasets and case history literature, we reported the dam height (158 cases), primary ore type (298 cases), total impounded volume (92 cases), total released volume (142 cases) and number of deaths (all cases). We also listed the causative variable(s) for 229 cases based on the definitions provided in Rana et al. (2021a) and the incident descriptions provided in ICOLD (2001), Bowker and Chambers (2015), Islam and Murakami (2021) and Rana et al. (2021a,b).

Table 2-5 introduces a new three-tiered qualitative index termed Consequence Level for TSF failures. The Consequence Level helps distinguish case histories according to the inundation extent (within vs. outside the mine perimeter) and the geochemical vs. physical nature of their impacts. The Consequence Levels were assigned to 285 (out of 303) cases in the period 1965-2020 with sufficient

information in scientific literature, news articles and satellite imagery. The process of assigning the qualitative index required some subjective interpretation and judgment.

Table 2-5: A new three-tiered qualitative index that groups tailings storage facility (TSF) failures according to the main consequences related to socioeconomic and environmental features.

Consequence Level (CL)	**Definition**	**Case example**	**No. of cases in the period 1965-2020**
CL1 (Localized Impact)	The consequences of the TSF failure were localized within the immediate vicinity of the failed TSF without any notable physical or chemical impacts to environmental (e.g. water bodies and vegetation), social and infrastructure features outside the mine perimeter. These events have commonly caused minor to moderate, temporary disruption to the mine operation, damages to mine equipment and, in worst-case scenarios, injuries and fatalities among on-site workers.	2019 Cadia, Australia	89
CL2 (Off-Site Geochemical Impact)	The TSF failure resulted in mainly geochemical consequences to socio-environmental features outside the mine perimeter. In news coverage, Level 2 events are typically reported as tailings "spills" or "leakages", and the main cause of concern is the geochemical impact. These events typically resulted in significant contamination of surface waters, groundwater, and soils (potentially for several to dozens of km when spilled into a river), and potential human injuries and fatalities by consumption of polluted waters and foods.	1995 Omai, Guyana	110
CL3 (Off-Site Physical and Geochemical Impact)	The TSF failure caused major geomorphological and socioeconomic consequences by extremely rapid (>5 m/s) runout along the downstream travel path that may extend several to dozens of km beyond the mine perimeter. The TSF failure triggered a mass flow capable of destroying engineered structures (houses, buildings and bridges) and causing mass casualties and environmental disasters, typically triggering a large-scale emergency response and widespread news coverage. Level 3 events also result in long-lasting geochemical (Level 2) consequences.	2019 Feijao, Brazil	86

It is difficult to confirm whether the 303 TSF failures in the period 1965-2020 is a statistically complete figure or whether more TSF failures have occurred in poorly reported regions. To

qualitatively account for this uncertainty, we refer to this number (303) as the "reported" number of TSF failures worldwide over the study period.

2.5 Statistical analysis

2.5.1 Large reservoir facilities (LRFs)

2.5.1.1 Temporal characteristics

Figure 2-3 displays the annual and cumulative numbers and the cumulative storage capacity of newly constructed LRFs worldwide in the period 1965-2020. There have been 33,424 constructed LRFs in this period. The number of constructed LRFs per annum has declined by over 90% since the peak during the 1970s. The total storage capacity of LRFs built in this period is about 1.1 x 10^{13} m^3 (almost the volume of Lake Superior, the world's third largest freshwater lake), with the cumulative number rising linearly since the early 1980s. Additional graphical insights on the geometric and geographic characteristics of constructed LRFs are provided in the Supplementary Article.

The number of LRF failures worldwide per annum, based on our estimates in Figure 2-4, has declined at a rate almost proportional to the LRF construction rate. This decline can be explained by previous statistical observations that the greatest likelihood for a newly constructed reservoir facility to fail is within the first 5 years of operation (Foster et al., 2000; Zhang et al., 2016). Therefore, one can expect that as the rate of LRF construction declines, so will the number of LRF failures. Our Estimation Scenario 1 indicates that 394 LRF failures occurred worldwide in the period 1965-2020, whereas our Scenario 2 produces a higher estimate of 608 LRF failures worldwide over the same period (Figure 2-4b). China accounts for 66% (Scenario 1) to 78% (Scenario 2) of LRF failures but just 42% of constructed LRFs worldwide. We observe four distinct failure periods in Figure 2-4a: the peak era of LRF failures appears to end in 1965 (23 cases in Scenario 1 and 37 cases in Scenario 2); 1966 to the late 1970s (13 to 22 cases per annum in Scenario 1 and 24 to 32 cases per annum in Scenario 2); late 1970s to the turn of the century (4 to 14 cases in Scenario 1 and 7 to 17 cases in Scenario 2); and ≤5 events per year since then (with 0 LRF failures in 2015 and 2016). Based on our minimum-estimate data, the cumulative storage capacity of failed LRFs since 1965 is >2.3 x 10^{10} m^3, which makes up >0.2% of the cumulative storage capacity of constructed LRFs.

The cumulative number of fatalities caused by LRF failures over the study period is ~40,100 (Figure 2-4d). Due to incomplete data, we considered only those LRF failures with a minimum death toll of 10 people. The cumulative total is a representative estimate based on our assumption of ~10,000 deaths for the 1979 Machhu disaster in India (the unofficial death toll ranges from 2,500 to 15,000 in literature and news articles).

The histograms in Figure 2-5 show that the annual number of LRF failures worldwide over the study period was predominantly ≤10 in both Estimation Scenarios. The histograms in Figure 2-5b to Figure 2-5d divide the study period into four successive 14-year intervals, illustrating how the frequency of LRF failures worldwide has reduced considerably over time in both Estimation Scenarios.

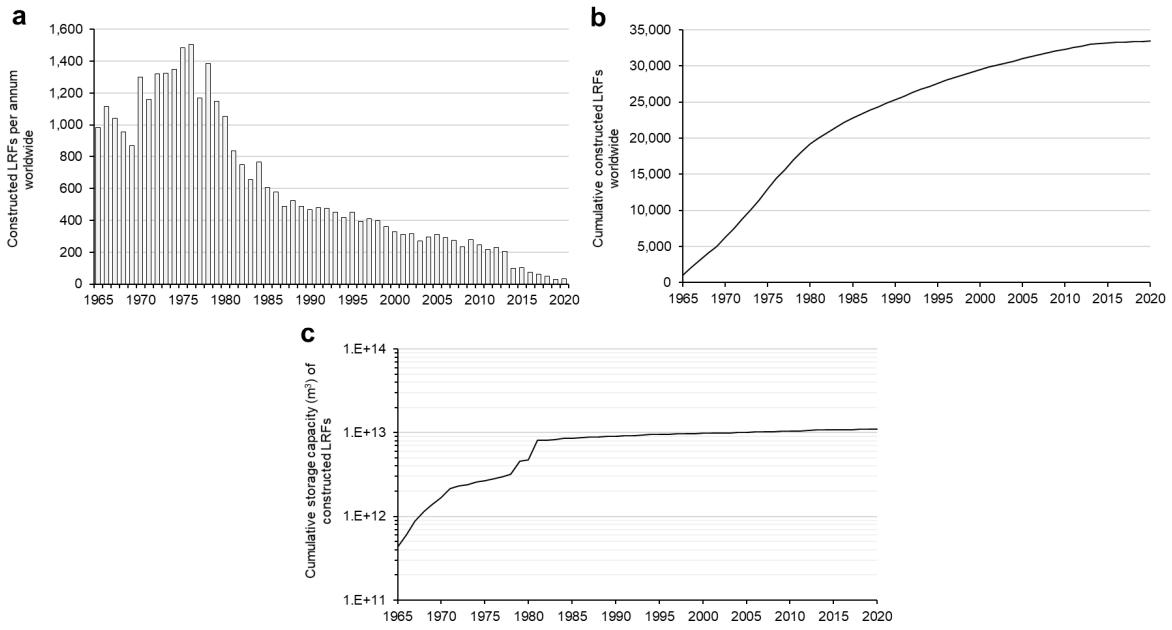

Figure 2-3: Constructed LRFs worldwide in the period 1965-2020. [a] Number of newly constructed LRFs per annum worldwide. The construction rate has declined over time. [b] Cumulative number of newly constructed LRFs worldwide over the study period, equaling to 33,424. [c] Cumulative sums of storage capacities of newly constructed LRFs, presented in semi-log scale, based on a complete record of 33,424 LRFs. As of end-of-2020, the cumulative total is 1.1 x 10^{13} m^3.

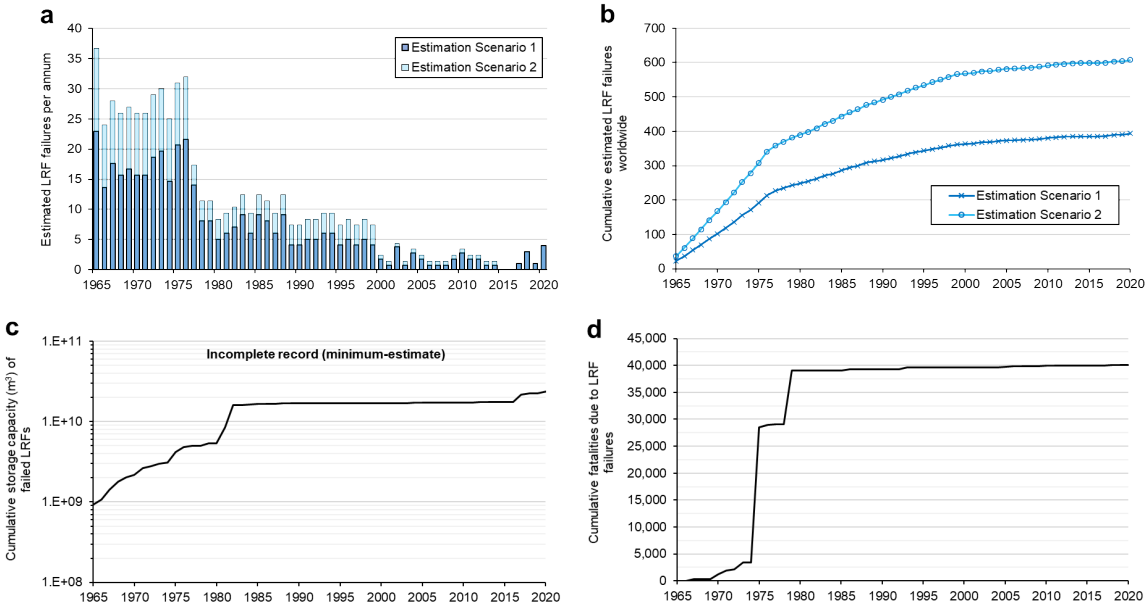

Figure 2-4: Estimated LRF failures worldwide in the period 1965-2020. [a] Estimated annual number of LRF failures and [b] cumulative estimated number of LRF failures over time based on two estimation approaches. Scenario 1 leads to a lower-estimate of 394 LRF failures whereas Scenario 2 produces an upper-estimate of 608 LRF failures worldwide (see Section 2.4.3.2). Both Estimation Scenarios indicate that the annual number of LRF failures worldwide has declined over time. [c] Cumulative sums of storage capacities of LRF failures, presented in semi-log scale, based on an incomplete record (minimum-estimate data) from Zhang et al. (2016), Ge et al. (2020) and our supplementary database of LRF failures. As of end-of-2020, the cumulative total is >2.3 x 10^{10} m^3, which makes up >0.2% of the cumulative total for constructed LRFs. [d] Cumulative number of fatalities caused by LRF failures over the study period, totaling to ~40,100 deaths. We assumed a representative estimate of ~10,000 deaths for the 1979 Machhu disaster in India (unofficial death toll ranges from 2,500 to 15,000). Due to incomplete data, only LRF failures resulting in a minimum death toll of 10 people are considered here. The fatality data was extracted from USBR (2015), Zhang et al. (2016) and other case history sources listed in our supplementary database of LRF failures.

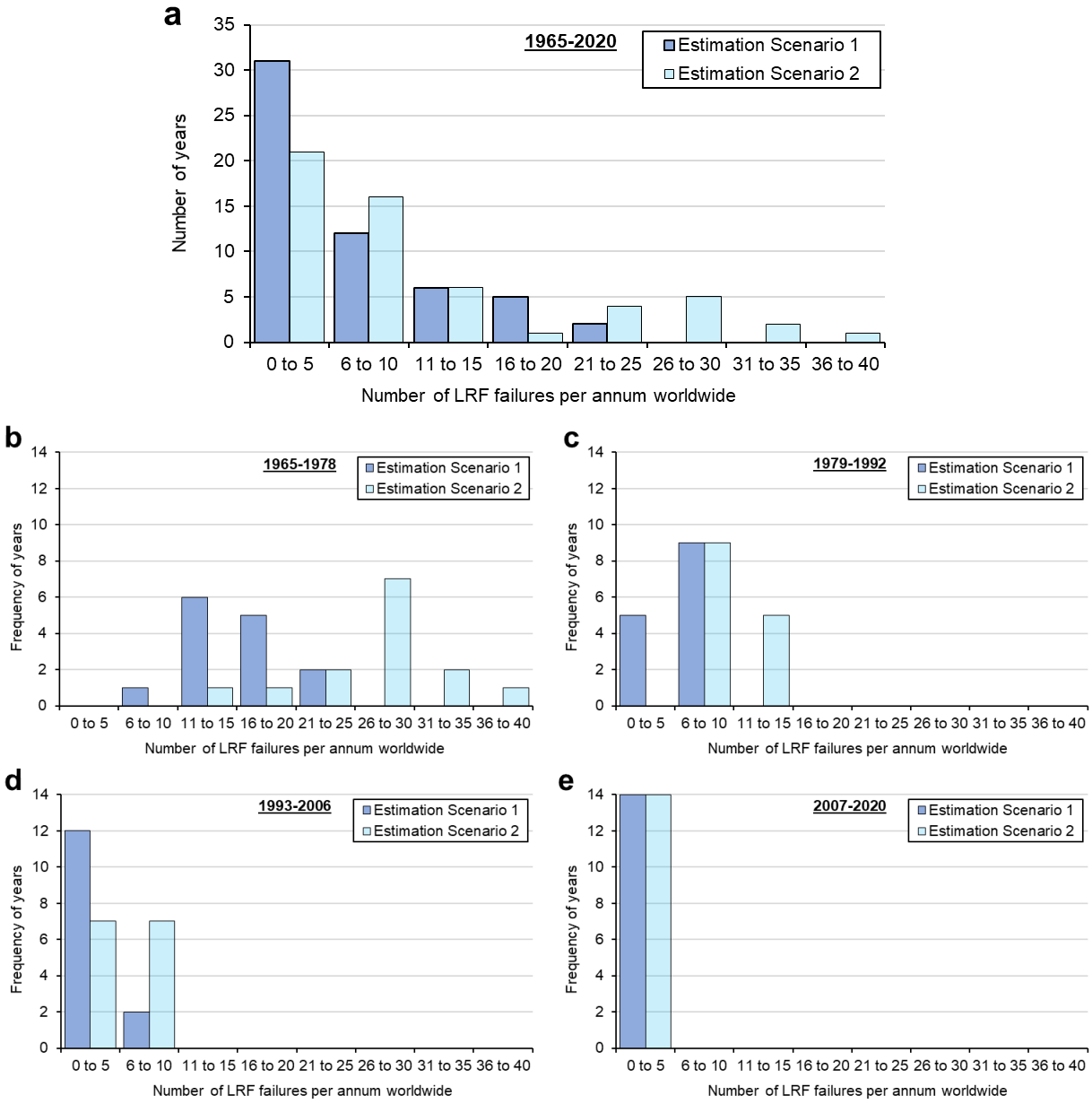

Figure 2-5: Histogram distribution of the estimated annual number of LRF failures worldwide in the period [a] 1965-2020, which is then separated into four successive 14-year intervals: [b] 1965-1978, [c] 1979-1992, [d] 1993-2006 and [e] 2007-2020. The estimates are presented based on our two different approaches: Estimation Scenario 1 and Estimation Scenario 2. See Section 2.4.3.2 for further details. Both Estimation Scenarios indicate that the annual number of LRF failures worldwide has declined over time. The frequency data is derived or estimated from Zhang et al. (2016), Ge et al. (2020) and our supplementary database of LRF failures.

2.5.1.2 Magnitude-frequency distributions

Figure 2-6 illustrates the magnitude-frequency profiles of LRF failures in terms of storage capacity and fatalities. In Figure 2-6a, we show the cumulative storage capacity-frequency curve for only LRF failures with storage capacities of ≥10 M m^3 in the period 1965-2020. The reason for this is to reduce the influence and degree of data incompleteness for LRF failures. In other words, the storage capacity data is more complete (though still not fully complete) for high-magnitude LRF failures, such as those with minimum storage capacities of 10 M m^3, which allows us to generate a more representative annual frequency profile with a sufficient number of data points.

A robust power-law relationship is observed for the 54 cases. We calculated a mean annual frequency of >0.8 for failed LRFs with storage capacities of ≥20 M m^3. This implies a mean return period (i.e. inverse of annual frequency) of <1.3 years. For failed LRFs with storage capacities of ≥50 M m^3, the mean return period is estimated to have been <2 years. In comparison, the average frequency of LRF construction worldwide in the period 1965-2020 was 110 per annum (one every 3 days) and 67 per annum (one every 5 days) for those with storage capacities of ≥20 M m^3 and ≥50 M m^3, respectively. We also draw attention to the 78 "mega" LRF projects with storage capacities of ≥50,000 M m^3 that form high-end outliers in Figure 2-6a.

Given that the number of LRF failures markedly declined over time, we divided the cumulative frequency curves into two successive 28-year intervals in Figure 2-6b. Using the power-law function corresponding to the period 1965-1992, the mean return periods are calculated to have been <1 year for failed LRFs with storage capacities of ≥20 M m^3 and <1.5 years for those with storage capacities of ≥50 M m^3. In the period 1993-2020, the mean return periods for LRF failures decreased to <2 years (for ≥20 M m^3) and <3 years (for ≥50 M m^3). This indicates that the recurrence rate of high-magnitude LRF failures has reduced by two-fold over time.

Figure 2-6c plots the cumulative fatality-frequency profile for 27 LRF failures in the period 1965-2020 with a minimum death toll of 10 people. The mean return period for LRF failures with ≥20 deaths and ≥100 deaths was ~2.5 years and ~5 years, respectively. In Figure 2-6d, we divided the study period into two time intervals (as done in Figure 2-6b). In the period 1965-1992, the mean return period was calculated to have been ~2 years for LRF failures with ≥20 deaths and ~3 years for events with ≥100 deaths. In the period 1993-2020, the corresponding mean return periods reduced to ~3.5 years and ~14

years. This implies that the frequency of fatal LRF failures has declined over time, with the reduction factor calculated to be ~2x for events with ≥20 deaths and ~4.5x for events with ≥100 deaths.

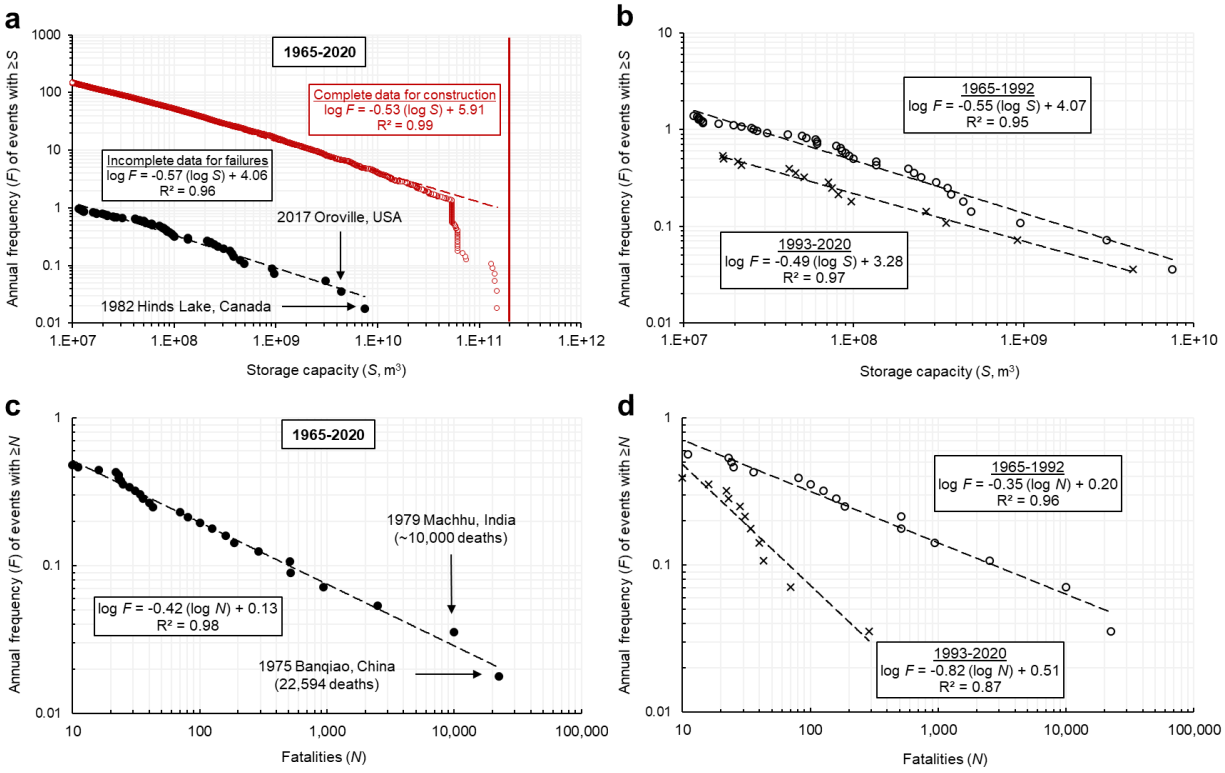

Figure 2-6: Magnitude-frequency profiles of LRF failures in the period 1965-2020. [a] Cumulative storage capacity-frequency curves of LRFs with minimum storage capacities of 10 M m^3 that have failed (black; n=54) and been constructed (red; n=8,313). The red vertical line is manually inserted to indicate the maximum credible storage capacity of a constructed LRF (2 x 10^{11} m^3) based on historical data in the ICOLD Registry. [b] The failure data from [a] is divided into two successive 28-year intervals: 1965-1992 (open circles) and 1993-2020 (crosses). [c] Cumulative fatality-frequency curves of failed LRFs. A total of 27 events with a minimum death toll of 10 people are considered. Note that the unofficial death toll for 1979 Machhu (India) ranges from 2,500 to 15,000 in literature and news articles, and we assumed a representative estimate of ~10,000. [d] The data from [c] is divided into two successive 28-year intervals: 1965-1992 (open circles) and 1993-2020 (crosses).

2.5.2 Tailings storage facilities (TSFs)

2.5.2.1 Temporal characteristics

Figure 2-7 shows the results from our Median Extrapolation and Mean Extrapolation approaches to estimate the number of constructed TSFs worldwide in the period 1965-2020. Our Median Extrapolation produces an upper-estimate that, based on our review, is closer to the representative estimate, whereas our Mean Extrapolation produces a lower-estimate that aligns with interpretations in previous studies (Section 2.3.2.1 and Section 2.4.4). For comparison, we indicate by a red circle in Figure 2-7b the conservative estimate of ~3,500 TSFs in the year 2000 by Davies et al. (2000) and Davies (2002), which is equal to our Mean Extrapolation estimate for the same year.

Our Median Extrapolation approach indicates that the number of TSFs constructed (or commenced) per annum has increased over time from ~150 in 1965 to ~650 in 2020, representing an increase of ~330%. In comparison, our Mean Extrapolation approach indicates a similar global rate of increase from ~50 to ~220 new TSFs per annum. The cumulative estimated number of constructed TSFs is 20,230 based on our Median Extrapolation, which is three times the Mean Extrapolation estimate of 6,810. The incomplete disclosure database in Franks et al. (2021) identifies only 1,336 constructed TSFs in the same period. The global volume of produced (and stored) tailings per annum has accelerated over time (Figure 2-7c), with the cumulative volume estimated to be just over 2.5×10^{11} m^3 (almost the volume of Lake Winnipeg, Canada's sixth largest freshwater lake). In comparison, the cumulative volume reported in the disclosure database in Franks et al. (2021) is 3.9×10^{10} m^3, which equals to 16% of our estimated total.

Figure 2-8a shows that the global rate of reported TSF failures has remained fairly constant since 1966. The histogram distributions of reported TSF failures in Figure 2-9 illustratively confirm this; the annual number of TSF failures worldwide is predominantly between 2 and 6 events over the study period. The peak of 21 failures in the year 1965 was largely due to the M_w 7.6 earthquake in western Chile that caused the simultaneous failures of 17 TSFs (Dobry and Alvarez, 1967; Villavicencio et al., 2013; Rana et al., 2021a,b; Islam and Murakami, 2021), leading to a national ban on the upstream dam-raise method. In total, there were 303 reported TSF failures in the period 1965-2020 (Figure 2-8c). The cumulative number of fatalities resulting from these events was 2,023 (Figure 2-8e). We categorized 285 of the 303 cases according to their Consequence Levels (Table 2-5). Of these 285 cases, 89 were

of Consequence Level 1 (CL1: localized impact), 110 were of CL2 (off-site geochemical impact) and 86 were of CL3 (off-site physical and geochemical impact). The rate of CL1 incidents declined over time, commensurate with the marginally rising rate of CL2 cases. The rate of CL3 events remained relatively constant.

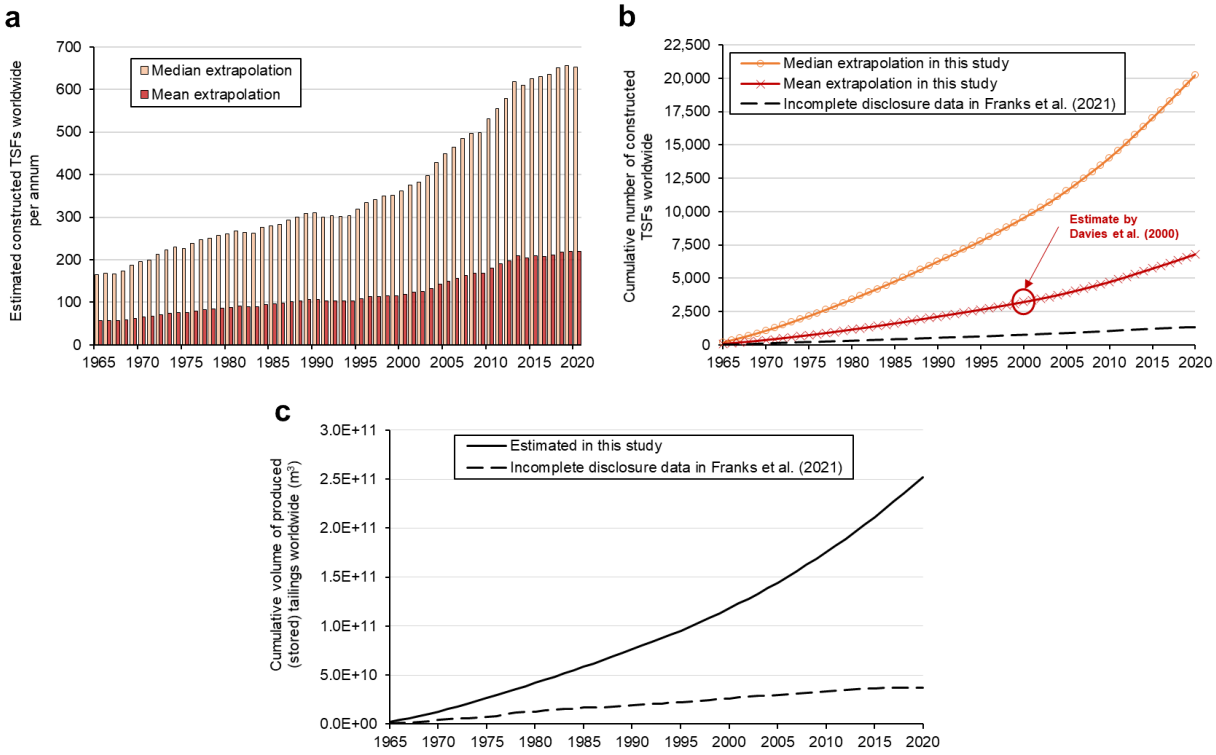

Figure 2-7: Estimated construction rate of TSFs worldwide in the period 1965-2020. [a] Annual numbers of constructed TSFs estimated from mineral production data provided in BGS (2021) based on two approaches: the Median Extrapolation (our upper-estimate and potentially representative estimate) and the Mean Extrapolation (our lower-estimate that aligns with interpretations in previous studies). See Section 2.4.4 for details. Both approaches indicate that the annual construction rate of TSFs has increased over time. [b] Cumulative number of constructed TSFs, totaling to 20,230 based on the Median Extrapolation and 6,810 based on the Mean Extrapolation. The inserted red circle indicates the conservative estimate of ~3,500 TSFs in the year 2000 by Davies et al. (2000) and Davies (2002), which matches our estimate for the same year when using the Mean Extrapolation approach. In comparison, there were 1,336 constructed TSFs reported in the disclosure database in Franks et al. (2021). [c] Estimated cumulative volume of produced (stored) tailings worldwide in comparison to 5-year planned storage volumes for 1,336 TSFs in Franks et al. (2021). About 16% of our estimated volume is accounted for in Franks et al. (2021).

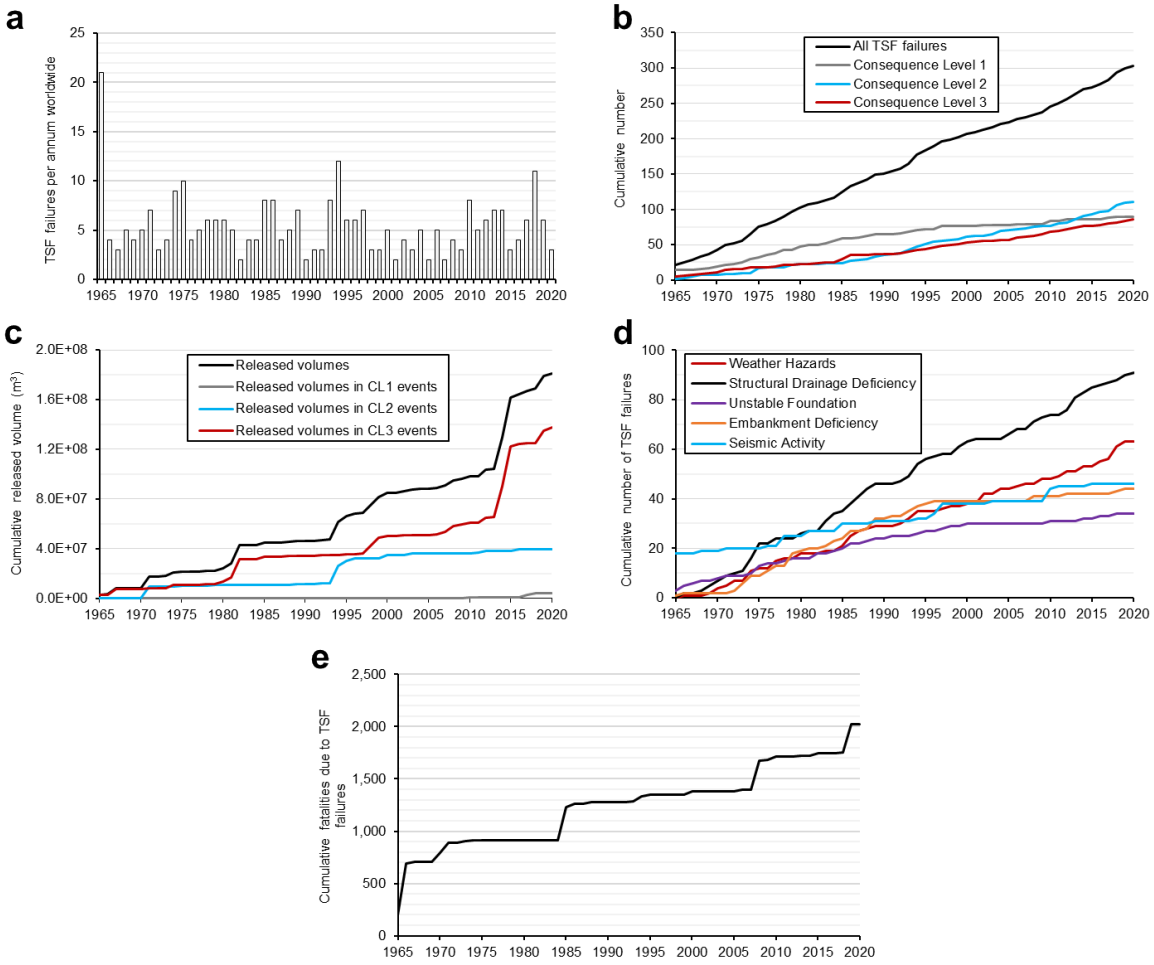

Figure 2-8: Reported TSF failures worldwide in the period 1965-2020, as listed in our supplementary database. [a] Annual numbers of reported TSF failures, which sum to a cumulative total of 303. [b] Cumulative number of reported TSF failures according to corresponding Consequence Levels (n=285 out of 303 cases). Refer to Table 2-5 for definitions of Consequence Level 1 (CL1: localized impact), CL2 (off-site geochemical impact) and CL3 (off-site physical and geochemical impact). [c] Released volumes from failed TSFs (n=153 out of 303 cases), including by categorization of CLs 1-3. The average released volume from failed TSFs has risen and the released volumes in CL3 events has increased at a greater rate compared to CL2 events, especially since 2014. [d] Cumulative number of reported TSF failures according to causative variables (n=229 out of 303 cases). Note that many TSF failures have been caused by multiple variables. The proportional contributions of structural drainage deficiency and weather hazards as causative variables has increased, whereas that of seismic activity, embankment deficiency and unstable foundation has decreased. [e] Cumulative number of fatalities due to TSF failures over the study period, totaling to 2,024 deaths. The fatality data was extracted from Bowker and Chambers (2015), WISE (2021), Rana et al. (2021a,b) and Islam and Murakami (2021).

The cumulative volume released by reported TSF failures in the period 1965-2020, based on minimum-estimate data (due to missing data for many CL1 cases and some CL2 cases), was >181 M m^3 (Figure 2-8c). CL3 events made up ~76% of this cumulative volume. This proportion rose from 62% before the occurrences of the 2014 Mt. Polley (Canada) and 2015 Fundao (Brazil) TSF failure disasters with a combined released volume of 57 M m^3 (Morgenstern et al., 2015, 2016). CL1 events comprised a minor proportion of released volumes of reported TSF failures. We also find that the released volumes in CL3 events has increased at a greater rate compared to CL2 events, especially since 2014.

In Figure 2-8d, we categorized 229 (out of 303) TSF failures in the period 1965-2020 according to their causative variables as defined in Rana et al. (2021a). We note that a complex TSF failure event may have multiple causative variables. Structural drainage issues contributed to 91 TSF failures (40%), followed by weather hazards (64 cases; 28%), seismic activity (46 cases; 20%), embankment deficiencies (44 cases; 19%) and unstable foundations (34 cases; 15%). The relative proportions of seismic activity, embankment deficiency and unstable foundation have decreased over time, likely reflecting improvements in design and practice. However, weather hazards (primarily heavy rainfall) and TSF drainage deficiencies (insufficient drainage capacity or structural collapse) remain a concern, especially when considering that these two variables can act in conjunction with exacerbating effects (Rana et al., 2021a).

Additional graphical insights on the geometric and geographic characteristics of constructed and failed TSFs based on the disclosure data in Franks et al. (2021) and our supplementary database are provided in our Supplementary Article.

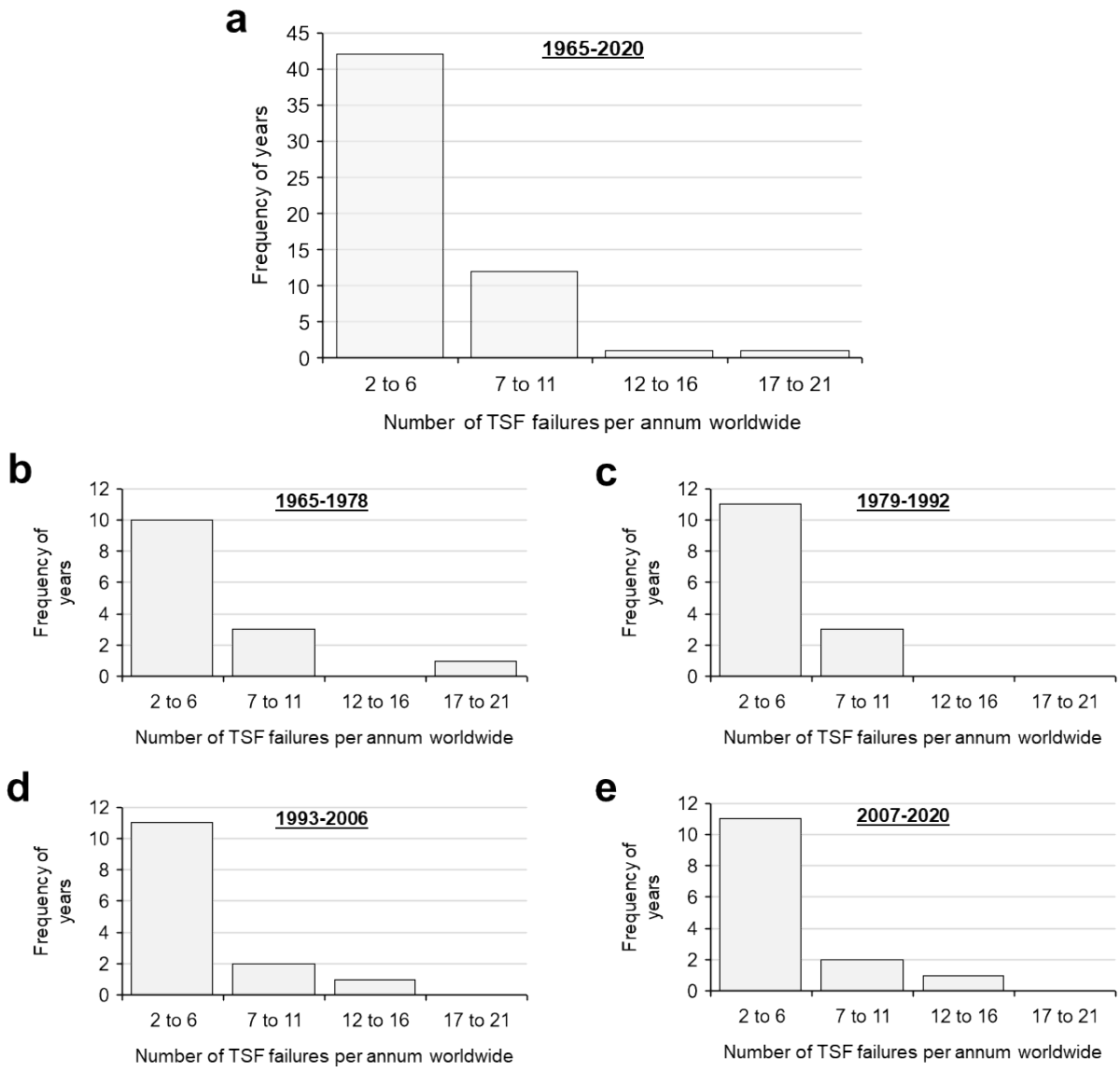

Figure 2-9: Histogram distribution of the annual number of reported TSF failures worldwide in the period [a] 1965-2020, which is then separated into four successive 14-year intervals: [b] 1965-1978, [c] 1979-1992, [d] 1993-2006 and [e] 2007-2020. The rate of TSF failure occurrences has remained fairly constant through the study period.

2.5.2.2 Magnitude-frequency distributions

Figure 2-10 displays the magnitude-frequency profiles of TSF failures in the period 1965-2020 in terms of volume, Zone 1 inundation area (Ghahramani et al., 2020; Rana et al., 2021a,b) and fatalities. We obtain a robust power-law relationship for TSF failures with released volumes of ≥1 M m^3 (Figure 2-10a). The mean return periods for TSF failures with released volumes of ≥1 M m^3 and ≥5 M m^3 are calculated to have been ~1.5 years and ~7 years, respectively. For comparison, we plotted the 5-year storage volume vs. annual frequency data for constructed TSFs from the incomplete disclosure data in Franks et al. (2021). (Note that our extrapolated estimates of constructed TSFs cannot be considered here as magnitude-frequency plots require real-world data points.) Furthermore, we manually inserted a trendline that is parallel to the one for TSF failures and that covers the straight-line portion of the 5-year storage volume data. Using this linear minimum-estimate formula, we calculated that ≥3 TSF construction projects are commenced every year with planned 5-year storage volumes of at least 50 M m^3. We also remark that the high-end outliers represent "mega" TSF projects with 5-year planned storage volumes of ≥500 M m^3, which have been commenced at least once every ~3 years on average worldwide.

Figure 2-10b plots the annual frequency profile for 19 tailings flows with a minimum Zone 1 inundation area of 0.5 km^2. Tailings flows with Zone 1 inundation areas of ≥1 and ≥10 km^2 occurred once every ~4 years and ~18 years, respectively, on average. The annual frequency of the 2015 Fundao case is overestimated by the best-fit trendline because an event of its magnitude (i.e. Zone 1 inundation area) approaches the maximum credible event.

Figure 2-10c shows the annual frequency profile for 19 TSF failures with a minimum death toll of 10 people. The 1966 Sgorigrad case in Bulgaria, reportedly the deadliest TSF failure in history with 488 deaths, is close to the maximum credible fatal event and therefore its annual frequency is overestimated by the best-fit trendline. We calculate that tailings flows that caused ≥10 deaths and ≥100 deaths occurred once every ~3 years and ~12 years, respectively, on average.

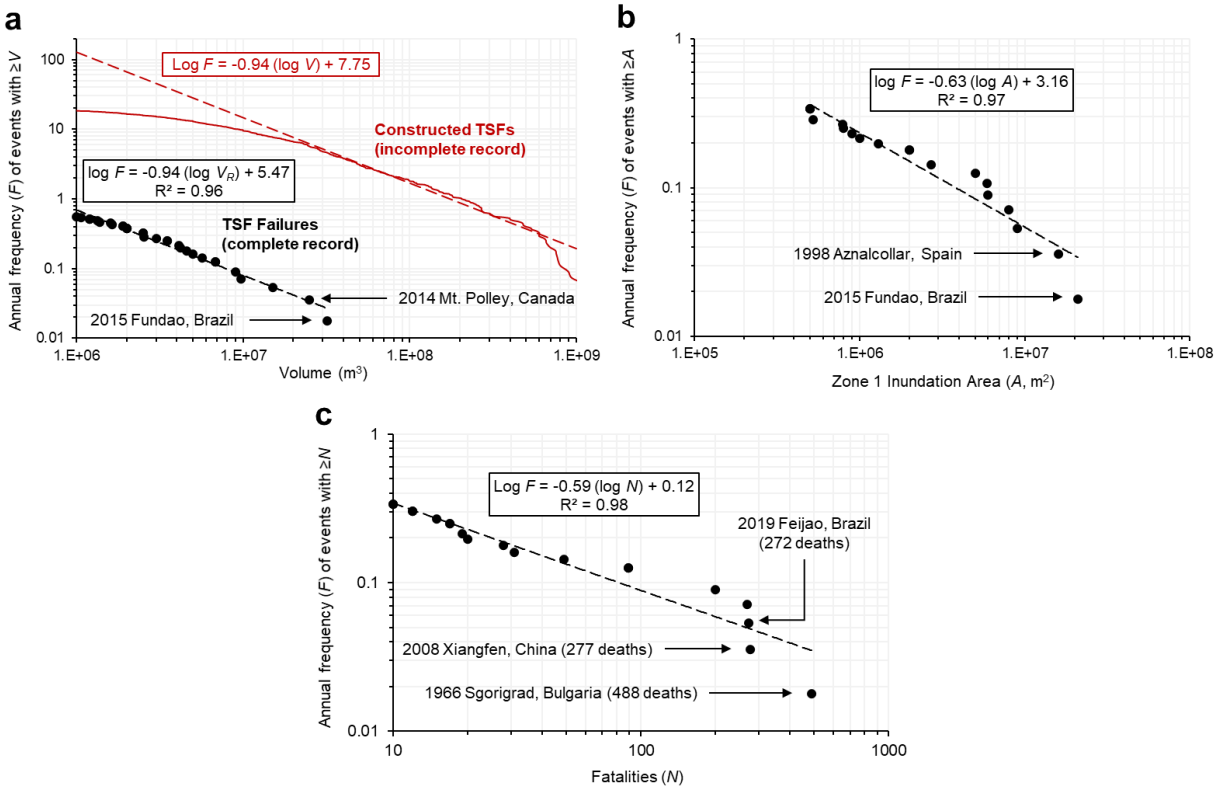

Figure 2-10: Magnitude-frequency profiles of TSF failures in the period 1965-2020. [a] Cumulative frequency curves for released volumes (V_R) of ≥1 M m^3 from 31 failed TSFs (shown in black; data extracted from Bowker and Chambers, 2015, Islam and Murakami, 2021, and Rana et al., 2021a,b) in comparison to the cumulative 5-year planned storage volume vs. frequency profile for 1,260 TSFs constructed worldwide (shown in red; data extracted from incomplete record in Franks et al., 2021). This plot builds on the work of Rana et al. (2021a), who focused only on downstream tailings flows that were captured on satellite imagery. [b] Annual frequency profile of Zone 1 inundation areas (≥0.5 km^2) of 19 tailings flows, with the data extracted from Ghahramani et al. (2020) and Rana et al. (2021a,b). [c] Cumulative fatality-frequency curve for 19 TSF failures with a minimum death toll of 10 people, based on data reported in WISE (2021), Rana et al. (2021a,b), Islam and Murakami (2021) and the present study.

2.5.3 Normalized rates

Figure 2-11 plots the failure rates of LRFs and TSFs normalized by the number of constructed facilities in the period 1965-2020. The annual rates of LRF failures per constructed LRF generally remained below 3% in both Estimation Scenarios until 2018 (Figure 2-11a), when the annual numbers of constructed LRFs dropped to a historic low. Of greater significance is Figure 2-11b where we plot the cumulative number of failures normalized by cumulative number of constructed facilities. The

global cumulative failure rate of LRFs ranges from ~1.2% (1 out of ~83 facilities) in Estimation Scenario 1 to ~1.8% (1 out of ~55 facilities) in Estimation Scenario 2 as of end-of-2020. Both approaches indicate that the rates have remained fairly constant since 1980, which is statistically explained by the near-proportional global decline in the number of constructed LRFs and LRF failures. We plot a separate curve in Figure 2-11b for all countries except China and obtained a lower, nearly constant cumulative failure rate of ≥0.7% (1 out of ≤143 facilities) as of end-of-2020. These estimates indicate that the failure rate of LRFs in China was anomalously high during the study period. We note that Foster et al. (2000) obtained a failure rate of 1.2% for large reservoir dams constructed up to 1986, excluding China. In Figure 2-11b, our corresponding estimate is ≥0.7%. This discrepancy may be due to the differences in statistical approaches (see Section 2.4.1) and underlying uncertainties, and/or may confirm that our rates are indeed minimum-estimates.

The annual rates of TSF failures per constructed TSF vary widely depending on the extrapolation method (Figure 2-11a). In both the Median Extrapolation and Mean Extrapolation approaches, the peak annual rates occur in 1965, which marked a record year of TSF failures. After 1965, the Median Extrapolation approach leads to annual rates that do not exceed 5% whereas the Mean Extrapolation produces rates as high as 14%. Figure 2-11b shows that, in contrast to LRFs, the cumulative failure rate of TSFs declined over time. This is explained by the rising rate of constructed TSFs and the fairly constant rate of TSF failures over time worldwide. The Median Extrapolation (which represents our upper-estimate and potentially representative estimate of the number of constructed TSFs) indicates that the cumulative failure rate declined from over 3% before 1980 to ~1.5% (1 out of ~67 facilities) as of end-of-2020. In comparison, our Mean Extrapolation (which represents our lower-estimate of the number of constructed TSFs) illustrates a decline from over 9% before 1980 to ~4.4% (1 out of ~23 facilities) as of end-of-2020. If one assumes the conservative estimate of ~3,500 TSFs in the year 2000 by Davies et al. (2000), Davies (2002) and our Mean Extrapolation approach, a cumulative failure rate of ~6.5% is obtained as of that year.

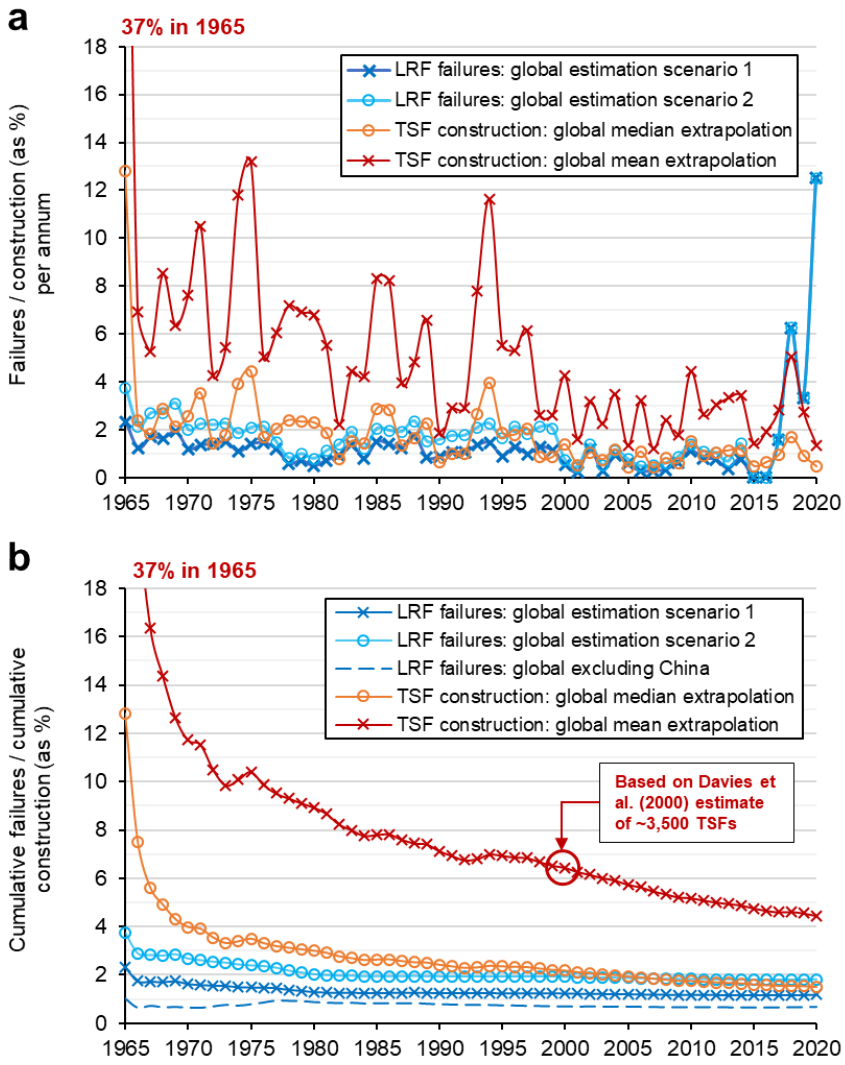

Figure 2-11: Normalized estimated failure rates for LRFs and TSFs worldwide in the period 1965-2020 based on our two estimation scenarios for LRF failures (Section 2.4.3.2 and Figure 2-4) and two extrapolation approaches for constructed TSFs (Section 2.4.4 and Figure 2-7). [a] Annual rates (expressed as %) of the number of failures per the number of constructed facilities for both LRFs and TSFs. [b] Cumulative rates (as %) of the cumulative number of failures per the cumulative number of constructed facilities for both LRFs and TSFs. For LRFs, we estimate a global cumulative failure rate of 1.2% in Scenario 1 (lower-estimate of the number of LRF failures) and 1.8% in Scenario 2 (upper-estimate of the number of LRF failures) as of end-of-2020. We plot a separate global curve excluding China and obtain a cumulative rate of ≥0.7% as of end-of-2020. For TSFs, we estimate a global cumulative failure rate of 1.5% by the Median Extrapolation (upper-estimate and potentially representative estimate of the number of constructed TSFs) and 4.4% by the Mean Extrapolation (lower-estimate of the number of constructed TSFs). Based on the lower-estimate of ~3,500 TSFs in the year 2000 by Davies et al. (2000), Davies (2002) and our Mean Extrapolation approach, the cumulative failure rate was 6.5%.

Figure 2-12a shows the ratios (as %) of cumulative storage capacities of failed vs. constructed LRFs in the period 1965-2020. The record for storage capacities of failed LRFs is incomplete and therefore represents a minimum-estimate. The observed "spikes" in the graph correspond to failures of LRFs with very high storage capacities, such as 1975 Banqiao and Shimantan in China (combined total of 586.4 M m^3), 1981 San Luis in USA (3,083 M m^3), 1982 Hinds Lake in Canada (7,500 M m^3), 2017 Oroville in USA (4,400 M m^3), 2018 Swar in Myanmar (266.8 M m^3), 2018 Xe-Pian in Laos (385 M m^3) and 2020 Sardoba in Uzbekistan (922 M m^3). As of end-of-2020, the cumulative storage capacity ratio of failed versus constructed LRFs is >0.21%, which represents a peak in the study period. This indicates that, since 2016, the cumulative storage capacity of newly constructed LRFs has increased at a slower rate than that of LRF failures with very high storage capacities (>200 M m^3), thus causing the recent spike. Due to lack of data on released volumes, it is difficult to estimate the percentage of the cumulative storage capacity of failed LRFs that has manifested as downstream floods.

Figure 2-12b plots the ratios (as %) of cumulative released volumes from failed TSFs (minimum-estimate due to missing data for many CL1 cases and some CL2 cases) and the cumulative produced volumes of tailings worldwide in the period 1965-2020. As of end-of-2020, the percentage of produced tailings that have failed is ~0.07%. The occurrences of the 2014 Mt. Polley (25 M m^3) and 2015 Fundao (32 M m^3) disasters triggered a notable increase in the cumulative ratio by about 0.03%.

Figure 2-13 illustrates the cumulative fatality rates of LRFs and TSFs normalized by the cumulative number of failures (Figure 2-13a) and the cumulative number of constructed facilities (Figure 2-13b) worldwide in the period 1965-2020. The rates have declined and stabilized over time for both types of engineered impoundments. For LRFs, the rate was at its peak in the 1970s due to the 1975 Banqiao and Shimantan (China) and 1979 Machhu (India) disasters that collectively led to over 35,000 deaths (note that the unofficial death toll for Machhu ranges from 2,500 to 15,000 in literature, and we assumed a representative estimate of ~10,000). As of end-of-2020, the fatality rate for LRFs is 1.2 deaths per constructed LRF (Figure 2-13a). When normalized by cumulative LRF failures, as of end-of-2020, our Estimation Scenario 1 indicates a fatality rate of 98 deaths per failure, whereas in Estimation Scenario 2, the rate drops to 64 deaths per failure (Figure 2-13b). The cumulative fatality rates for TSFs are much lower than those of LRFs. As of end-of-2020, the Median Extrapolation yields a rate of 0.1 death per constructed TSF, whereas the Mean Extrapolation indicates a corresponding rate of 0.3 (Figure 2-13a).

When normalized by cumulative TSF failures, the rate has remained relatively constant at 6 deaths per failure (Figure 2-13b).

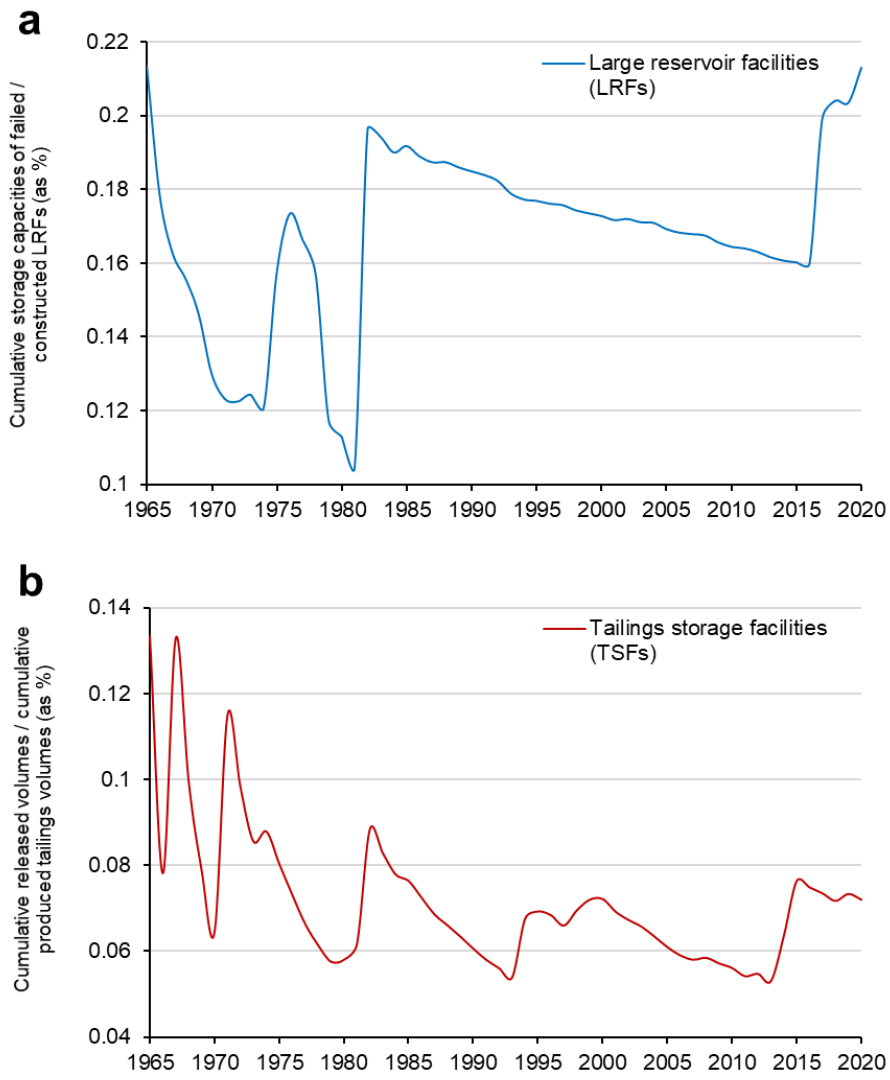

Figure 2-12: [a] Cumulative rates (as %) of the cumulative storage capacities of failed LRFs per the cumulative storage capacities of constructed LRFs based on minimum-estimate data in the period 1965-2020. [b] Cumulative rates (as %) of the cumulative released volumes of failed TSFs per the cumulative produced volumes of tailings in the period 1965-2020. The ratio of cumulative storage capacities of failed LRFs per constructed LRFs was >0.21% as of end-of-2020. This is a peak in the study period, which indicates that the cumulative storage capacity of newly constructed LRFs has increased at a slower rate than that of recently failed LRFs with very high storage capacities. The percentage of produced tailings that have failed is ~0.07% as of end-of-2020 and the cumulative ratio has remained fairly constant since the 1990s.

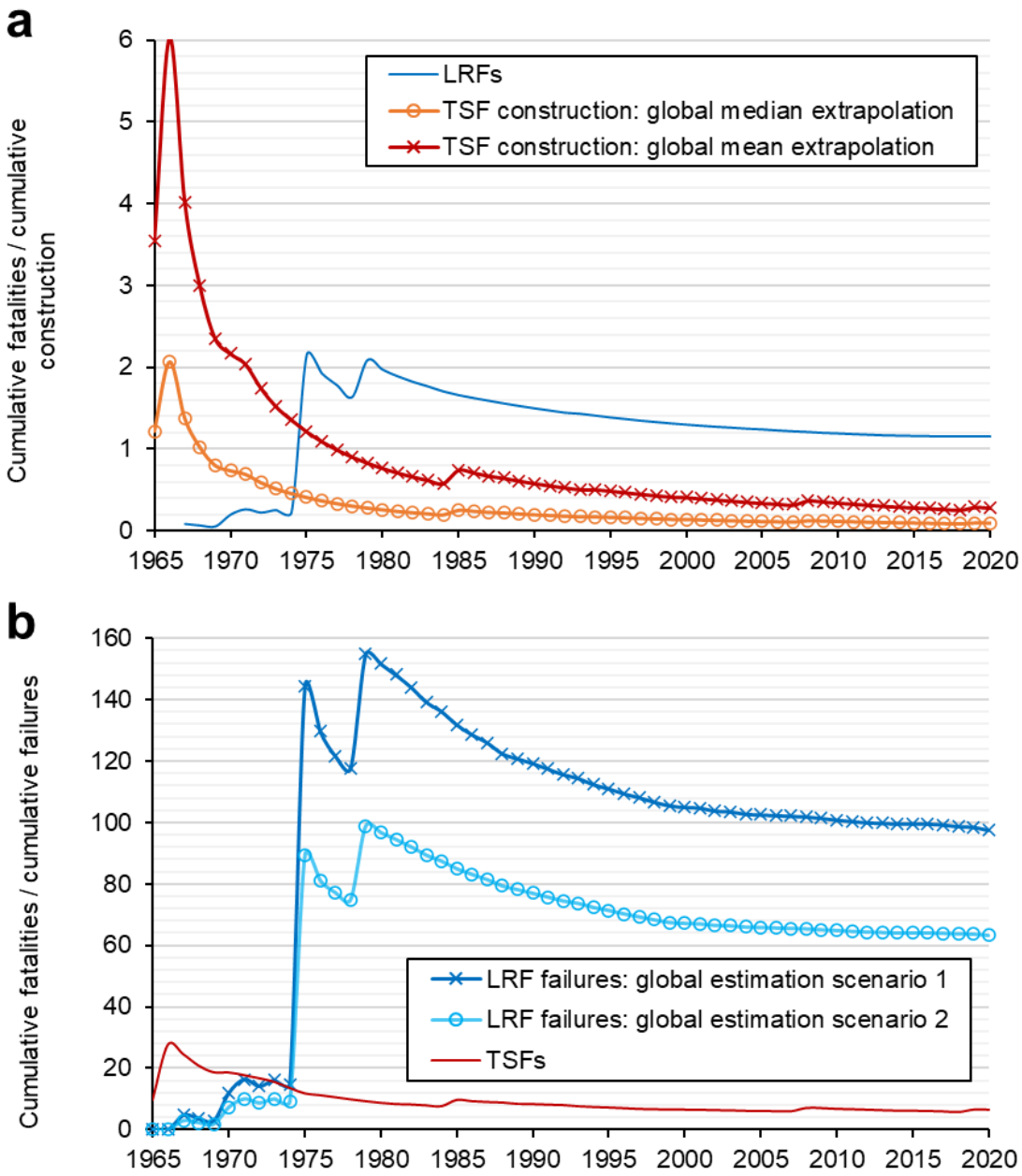

Figure 2-13: Cumulative fatality rates of LRFs and TSFs estimated as annual rates of cumulative fatalities per [a] cumulative constructed facility and [b] cumulative failures in the period 1965-2020. The estimated number of constructed TSFs in [a] is based on our Median Extrapolation (total of 20,230 TSFs; upper-estimate and potentially representative estimate) and Mean Extrapolation (total of 6,810 TSFs; lower-estimate) approaches (see Section 2.4.4 for details). The estimated number of LRF failures in [b] is based on our Scenario 1 (total of 394 LRF failures) and Scenario 2 (total of 608 LRF failures).

2.6 Discussion

2.6.1 Implications

Our comprehensive study produces key insights into the historical hazard-risk of engineered impoundments worldwide. LRFs and TSFs are two distinct engineering systems with different characteristics and design criteria. This corresponds to a notable difference in their statistical properties, particularly as it relates to the temporal changes in their cumulative construction, failure and fatality rates, their relationships with facility age and the geographic distribution of their failure occurrences.

Yet, our study also indicates an important statistical similarity between LRFs and TSFs: their global cumulative failure rates appear to be in the same order if we adopt the Median Extrapolation approach as a representative estimate of the number of constructed TSFs. This observation holds true in both Estimation Scenarios for LRF failures. This contradicts previous interpretations that the failure rate of TSFs is more than two orders of magnitude greater than water-retention dams (Azam and Li, 2010) or that the likelihood of TSF failures is several times higher than water-retention dams (Lyu et al., 2019; Liang et al., 2021). Our review and findings suggest that these previous interpretations were based on lower-estimates of the number of TSFs worldwide (e.g. ~3,500 at the turn of this century; Davies et al., 2000; Davies, 2002; Lyu et al., 2019; Liang et al., 2021), the missing historical data on water-retention dam failures in China, and the unknown number of water-retention dams worldwide. Based on our examination of existing records and estimates (see Section 2.3.2.1), we find that the "true" number of constructed TSFs worldwide in the period 1965-2020 is much closer to 20,230 (Median Extrapolation) than it is to 6,810 (Mean Extrapolation). This implies that the "true" global cumulative failure rate of TSFs is much lower than 4.4% (derived from Mean Extrapolation) and closer to 1.5% (derived from Median Extrapolation). In comparison, our results indicate that the "true" global cumulative failure rate of LRFs likely falls in the range of 1.2% to 1.8%. This finding has important implications in how researchers and practitioners perceive the global risk of TSFs – especially when comparing to that of water-retention dams – and in the underlying assumptions when undertaking hazard-risk assessments.

Moreover, we find that the cumulative failure rate of an engineered structure is not necessarily a constant number, but rather may evolve through time. This observation is particularly relevant to TSFs, which have experienced a steady decline in cumulative failure rate over the study period based on both our Median Extrapolation and Mean Extrapolation estimates.

2.6.1.1 Failures of large reservoir facilities (LRFs)

The frequency of LRF failures has significantly reduced over the study period, which may reflect global-scale improvements in LRF design and operation. We also observe that the global rate of LRF failures is correlated to the global rate of LRF construction. The greatest statistical likelihood for a reservoir facility to fail is between the time of first filling and the 5th year of operation, after which the likelihood generally declines over time (Foster et al., 2000; Zhang et al., 2016). This may partly explain the observed constant cumulative failure rate for LRFs. Despite this, major failures of LRFs with very high storage capacities (>200 M m^3) have continued to occur since 2016. Although data on causative variables is sparse, our supplementary database of LRF failures indicates that the majority of LRF failures since 2002 (at least 11 of 19 events, or ≥58%) have been associated with heavy rainfall. Heavy rainfall-induced failures of LRFs are predominantly manifested through overtopping or internal erosion of the dam. This has important implications for the design of storage capacity, freeboard and spillways under conditions of climate change.

We further note that in the period 1965-2020, China had 66% (Estimation Scenario 1) to 78% (Estimation Scenario 2) of LRF failures but just 42% of constructed LRFs worldwide. The USA had 14% (Estimation Scenario 2) to 21% (Estimation Scenario 1) of LRF failures and 12% of constructed LRFs worldwide. This implies the global hazard-risk of LRF failures was geographically concentrated in these two particular nations over the study period.

2.6.1.2 Failures of tailings storage facilities (TSFs)

In contrast to LRFs, there is no obvious correlation between facility age and failure likelihood in the case of TSFs, whereby even inactive or closed TSFs may possess an appreciable probability of failing several decades after commencement (e.g. Santamarina et al., 2019; Islam and Murakami, 2021). This is due to the inherent properties of TSFs: (i) hydraulically discharged tailings are predisposed to attain a sensitive, metastable state if drainage is not well-controlled in the facility, regardless of the age of the TSF (e.g. Morgenstern et al., 2016; Robertson et al., 2019); and (ii) TSFs are progressively built up over time, thus they continue to gain mass/energy during the operational stage, and they remain in a dynamic state until after closure.

This long-term engineering challenge can be further complicated by regional-scale intensifying rainfall patterns. For example, the sudden static liquefaction failure of the inactive Feijao TSF in 2019 in Brazil was preconditioned by higher-than-average rainfall over the preceding wet season, which led to the loss of suction strength (negative pore pressures) in previously unsaturated tailings (Robertson et al., 2019). On a global scale, structural drainage deficiencies and weather hazards (which can often act in conjunction with exacerbating effects; Rana et al., 2021a) are the two most dominant causative variables for TSF failures (Figure 2-8d). Therefore, similar to LRFs, the design and management of TSF drainage systems with the consideration of long-term, regional-scale precipitation patterns will be an important issue to address under conditions of climate change.

The annual number of TSF failures worldwide has stayed relatively constant and not declined (on average) since 1966. However, our estimates suggest that the number of constructed TSFs worldwide has increased over time, which has translated into a declining cumulative failure rate. Compared to LRF failures, TSF failures are more geographically well-distributed worldwide (e.g. Figure S14 and Figure S15 in the Supplementary Article). Our supplementary data and article (Figure S15) show a marked decrease in the frequency of TSF failures in the USA and Chile. The decrease in Chile is attributed to the national ban on the upstream-raise method in 1970 (Villavicencio et al., 2013). We also showed that the proportional contributions of seismic activity, embankment deficiency and unstable foundation as causative variables for TSF failures has declined over time worldwide.

Yet, we also observed that the global average released volume from failed TSFs has risen, primarily due to 3 catastrophic events since 2014 (2014 Mt. Polley, 2015 Fundao and 2019 Feijao). This increase is nearly proportional to the estimated increase in global tailings production volumes (see Figure 2-12 that shows a fairly constant cumulative ratio of 0.5-0.8% over the past few decades). Furthermore, although CL2 and CL3 events continue to occur at similar rates, the released volumes in CL3 events has been increasing at a greater rate, especially since 2014 (Figure 2-8b and Figure 2-8c). This implies that the environmental impact due to TSF failures has increased on average.

2.6.1.3 Global F-N and environmental risk

Another key insight in this study is that the global cumulative fatality rate has declined for both LRFs and TSFs between 1965 and 2020 – a period during which the global population increased by 130%. This observation is consistent irrespective of whether the fatality data is normalized by the number of constructed facilities or by the number of failures. We attribute this finding to systematic, incremental improvements in dam safety programs and risk assessment and management frameworks for both dam engineering sectors.

The fatality rates for LRFs are much higher than those for TSFs, which we attribute to the following three factors. First, LRFs are typically located closer to populated areas due to societal benefits (e.g. agricultural, hydroelectric and recreation reservoirs), whereas many mine sites are located farther away from residential areas. Second, for a given volume, the inundation extent of water floods from failed LRFs exceeds that of tailings flows which typically have significant sediment content. Lastly, on average, LRFs have larger storage capacities and tend to release a greater proportion of impounded contents relative to TSFs.

To contextualize our cumulative fatality-frequency results, we plot in Figure 2-14 the best-fit regression lines corresponding to LRF failures in 1965-1992 and 1993-2020 (Figure 2-6d) and TSF failures in 1965-2020 (Figure 2-10c) on an *F-N* diagram, which is commonly used as a framework for societal risk assessment for natural and engineering hazards (e.g. Baecher and Christian, 2003; DeNeale et al., 2019; Strouth and McDougall, 2021). *F-N* curves display the annual frequency (or annual probability), *F*, of *N* or more fatalities in log-log space. LRF failures have experienced a drop in their global *F-N* risk between the periods 1965-1992 and 1993-2020, which now appears to be comparable to that of TSF failures. The difference in the best-fit regression slopes of the LRF (1993-2020) and TSF curves implies the following: (i) the annual frequencies of LRF failures that resulted in ≥10 to ≥20 deaths are slightly higher than those of TSF failures that caused ≥10 to ≥20 deaths; (ii) the annual frequencies of LRF failures with ≥30 to ≥80 deaths are comparable to those of TSF failures that caused ≥30 to ≥80 deaths; and (iii) the annual frequencies of TSF failures that caused 100 or more deaths slightly exceed those of LRF failures that resulted in 100 or more deaths.

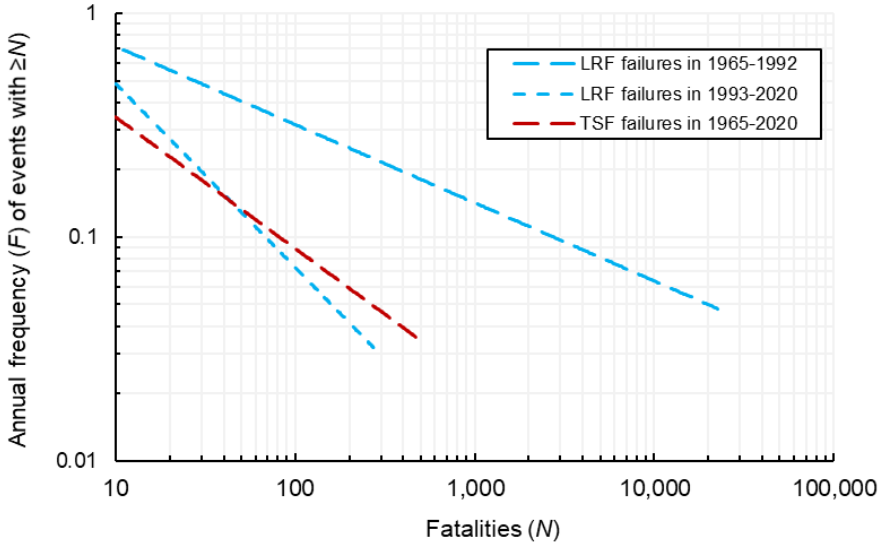

Figure 2-14: An F-N chart illustrating our global cumulative fatality-frequency results, represented by only the best-fit regression lines, for LRFs (1965-1992 and 1993-2020; Figure 2-6) and TSFs (1965-2020; Figure 2-10c).

We choose not to include in Figure 2-14 the societal risk acceptability criteria set by several studies and regulatory agencies for natural and engineering hazards (e.g. Versteeg, 1988; ANCOLD, 1994; GEO, 1998; Baecher and Christian, 2003; CDA, 2013; Briaud, 2020) due to the following issues: (i) each of the "Unacceptable" guidelines reflects the tolerable risk thresholds and policy decisions that are subjectively set by the specific author/regulator; (ii) the *F-N* societal risk acceptability criteria and analyses are more appropriate and representative in site-scale to regional-scale (rather than global-scale) risk studies; and (iii) while site-scale risks can each be acceptable under a guideline, probabilistically combining the site-scale risks may lead to unacceptability in an *F-N* chart, which can lead to inconsistent interpretation (Baecher and Christian, 2003). Due to these reasons, the guidance derived from *F-N* acceptability criteria is typically viewed as qualitative in conventional geotechnical practice (Baecher and Christian, 2003).

With respect to environmental risk, a key difference between LRF and TSF failures is the long-term fate and extreme geochemical consequences associated with downstream tailings flow deposits. Contaminant products from heavy metals and sulfides in released tailings, which are sometimes mixed with highly acidic or alkaline wastewater, can persist in the impacted environment and ecosystem for several years to decades, even after mechanical removal of the waste (e.g. Macklin et al., 2003; Liu et

al., 2005; Kraus et al., 2006; Kossoff et al., 2014; Gabriel et al., 2021). Tailings flows belonging to the categories CL2 and especially CL3 typically pollute surface water bodies (resulting in loss of inclusive biota), groundwater, riparian soils and vegetation, forestry and/or agricultural resources. Hazardous air quality (suspended tailings dust) in the aftermath of CL3 tailings flows is also a source of long-term health concerns to local populations and can impede disaster relief efforts (e.g. Galvão et al., 2020). As a result, post-failure clean-up, remediation and restoration project timelines and costs can be much greater than initially budgeted for, with major implications for the regional or national economy (e.g. Bowker and Chambers, 2015, 2017). Such events can also lead to large-scale displacement of affected communities due to the degraded quality of essential needs (e.g. water, food and air) and the loss of livelihoods (e.g. Amnesty International, 2017; Milanez et al., 2021). Table 2-5 and Figure 2-8b showed that 196 of 285 (~69%) TSF failures in the study period were assigned as either CL2 or CL3. This indicates that 3-4 TSF failures occur every year worldwide on average that result in reported downstream geochemical consequences.

In summary, although both LRF and TSF failures can produce physically destructive mass flows, the geochemical fate of tailings flows is of special concern as it relates to environmental risk. We note that this discussion ignores the local-scale environmental hazards that can result from the mere presence (not just failures) of some large TSFs (e.g. Dong et al., 2020; Okereafor et al., 2020). This has implications for the long-term sustainability of legacy TSFs given our observations of the increased global rate of tailings production and TSF construction over the study period (Figure 2-7), and keeping in mind that a significant proportion of TSFs worldwide have yet to undergo closure or decommissioning.

2.6.1.4 Characterizing the “global safety” of LRFs and TSFs

All of the findings in this study raise an important question: which of our numerous statistical results should define the “global safety” of LRFs and TSFs? Ultimately, the conversation on the “global safety” of either type of facility should incorporate all of these nuances and should not be narrowed down to a singular parameter or number. In this regard, we draw the following summary conclusions:

- The annual numbers of constructed and failed LRFs have declined almost proportionally, thus the global cumulative failure rate of LRFs has stayed nearly constant. The “true” rate likely falls in the range of 1.2% to 1.8% as of end-of-2020 based on our two estimation

approaches. The rate drops to at least 0.7% as of end-of-2020 if China is excluded. The global cumulative fatality rate of LRFs has reduced over time to 1.2 deaths per constructed facility, and falls within the range of 64 to 98 deaths per failure based on our estimates, as of end-of-2020. However, major failures of LRFs with very high storage capacities (>200 M m^3) have continued to occur since 2016.

- The annual number of TSF failures has stayed relatively constant, whereas the annual construction rate of TSFs has increased over the study period, thus the global cumulative failure rate of TSFs has declined over time. When assuming our lower-estimate of the number of constructed TSFs between 1965 and 2020, we obtain a cumulative failure rate of ~4.4% as of end-of-2020. Adopting our higher-estimate yields a rate of ~1.5% as of end-of-2020, which falls in the same order as the corresponding rate of LRFs. Our review of published estimates and figures of the "true" number of existing TSFs worldwide indicates that the "true" cumulative failure rate is much lower than 4.4% and closer to 1.5%. The cumulative fatality rate of TSF failures has reduced over time to 0.1-0.3 death per constructed facility and 6 deaths per failure as of end-of-2020; these rates are lower than the corresponding rates of LRFs. However, the size and the environmental impact of TSF failures have both increased on average, especially since 2014. The increased global rate of failed tailings volumes is roughly proportional to the increased rate of global tailings production since the 1990s.
- Heavy rainfall events and intensifying precipitation patterns are statistically important causative variables for the failures of both LRFs and TSFs. This has implications for the design and management of storage capacity, freeboard, facility drainage and spillways under climate change conditions.

2.6.2 Uncertainties, assumptions and limitations

The results presented in this study are applicable broadly on a global scale, having been conditioned by uncertainties in the data and the methods used to address data gaps. Methodological details of our estimation and extrapolation approaches are provided in Section 2.4 and the supplementary information. To improve the robustness of future statistical analyses, a more comprehensive public

disclosure effort is necessary, particularly with respect to reservoir facility failures in China and constructed TSFs worldwide.

We summarize below the most important sources of uncertainties in our statistical methods and, to avoid repetition, refer to the original sections where the basis and implications of these uncertainties are discussed:

- Focusing our statistical analysis of reservoir facilities to only those that fall within the "large dam" criteria (Section 2.3.1.1);
- Focusing our statistical analysis to a specific study period (Section 2.4.1);
- The missing storage capacity data, and the resulting minimum-estimate, of the number of LRF failures worldwide except China over the study period (Section 2.4.3.1);
- Adopting two statistical estimation scenarios to approximate the number of LRF failures in China and subsequently worldwide over the study period (Section 2.4.3.2); and
- Adopting two statistical extrapolation approaches to estimate the volume of produced (stored) tailings and the number of constructed TSFs worldwide over the study period (Section 2.4.4).

2.7 Summary and conclusions

The failures of engineered dams that retain freshwater or mining waste (tailings) have led to major socioeconomic and environmental consequences. However, the global, historical magnitude-frequency statistics of these occurrences have remained poorly quantified, out-of-date and/or limited in scope. In this study, we presented an analysis of new datasets and estimates of the construction and failures of large reservoir facilities (LRFs) and tailings storage facilities (TSFs) worldwide in the 56-year study period of 1965-2020. We addressed long-standing data gaps on LRF failures in China, and subsequently worldwide, and on constructed TSFs worldwide by adopting multiple estimation approaches for each variable in an effort to illustrate the range of uncertainty in our results. By addressing these gaps, we generated insights into how the hazard-risk of large water-retention dams and mine tailings impoundments has evolved over time. Additional databases, key insights, and methodological details on our estimation methods are provided as supplementary information.

Regarding LRFs, we observed the following global-scale statistics:

1. The total number of LRF failures worldwide in the period 1965-2020 was estimated to be in the range of 394 to 608. Between 66% and 78% of these failures occurred in China and between 14% and 21% occurred in the USA, which implies a geographic concentration of LRF hazard-risk in these two nations.

2. The annual rate of construction and failures have decreased almost proportionally over the study period, which translates into a nearly constant cumulative failure rate that likely falls within the range of 1.2% to 1.8% as of end-of-2020. This figure drops to a nearly constant rate of ≥0.7% when China is excluded.

3. The mean return period of failures increased from <1 year in the period 1965-1992 to <2 years in the period 1993-2020 for LRFs with storage capacities of ≥20 M m^3. Similarly, the mean return period increased from <1.5 years (1965-1992) to <3 years (1993-2020) for LRF failures with storage capacities of ≥50 M m^3. This implies that the recurrence rate of high-magnitude failures of LRFs reduced by two-fold over the study period.

4. The mean return period increased from ~2 years to ~3.5 years for LRF failures with ≥20 deaths and from ~3 years to ~14 years for events with ≥100 deaths. This indicates that the reduction factor for the number of fatal failures over time was ~2x for LRF failures with ≥20 deaths and ~4.5x for failures with ≥100 deaths.

5. The ratio of cumulative storage capacities of failed LRFs per constructed LRFs was >0.21% as of end-of-2020. This represents a peak in the study period as, since 2016, the cumulative storage capacity of newly constructed LRFs has increased at a slower rate than that of failed LRFs with very high storage capacities (>200 M m^3).

6. The cumulative fatality rates were 1.2 deaths per constructed LRF, and fell within the range of 64 to 98 deaths per LRF failure based on our estimates, as of end-of-2020.

7. Although data on causative variables is scarce, our supplementary database of LRF failures indicates that at least 11 of 19 cases (≥58%) since 2002 were associated with heavy rainfall.

Regarding TSFs, we observed the following global-scale statistics:

1. By extrapolation from mineral commodity production data, we obtained a lower-estimate of 6,810 TSFs and an upper-estimate of 20,230 TSFs that have been constructed worldwide in the period 1965-2020. Both estimates indicated that the annual number of constructed TSFs increased by ~3 times over this period. However, the annual number of TSF failures worldwide remained fairly constant, thus the cumulative failure rate declined over time. When assuming our lower-estimate of the number of constructed TSFs, the cumulative failure rate is estimated to be ~4.4% as of end-of-2020. When adopting our upper-estimate, the rate is ~1.5% as of end-of-2020, which falls in the same order as the corresponding rate of LRFs. Our review of published estimates indicates that the "true" number of constructed TSFs over the study period is much higher than our lower-estimate (6,810) and closer to our upper-estimate (20,230). This implies that the "true" cumulative failure rate is much lower than 4.4% and closer to 1.5%.

2. TSF failures with released volumes of ≥1 M m^3 and ≥5 M m^3 occurred once every ~1.5 years and ~7 years, respectively, on average.

3. Tailings flows with primary impact zone (Zone 1) inundation areas of ≥1 and ≥10 km^2 occurred once every ~4 years and ~18 years, respectively, on average.

4. TSF failures with ≥10 deaths and ≥100 deaths occurred once every ~3 years and ~12 years, respectively, on average.

5. The proportional contributions of structural drainage deficiency and weather hazards as causative variables of TSF failures has increased, whereas that of seismic activity, embankment deficiency and unstable foundation has decreased.

6. The percentage of produced tailings that have failed is ~0.07% as of end-of-2020 and the cumulative ratio has remained fairly constant since the 1990s. The global average released volume from failed TSFs has risen at a rate roughly proportional to the estimated increase in the global rate of produced tailings. Moreover, the average released volume in Consequence Level 3 events (off-site physical and geochemical impacts) has increased at a greater rate compared to that of Consequence Level 2 events (off-site geochemical impacts), especially since 2014.

7. The cumulative fatality rates are estimated to be 6 deaths per TSF failure and 0.1-0.3 death per constructed TSF as of end-of-2020.

Based on these comprehensive quantitative insights, we draw the following broad conclusions:

1. The annual numbers of LRF construction and failures have declined almost proportionally, thus the cumulative failure rate of LRFs has stayed constant. The cumulative fatality rate of LRF failures has reduced over time. However, major failures of LRFs with very high storage capacities have continued to occur.

2. The annual number of TSF failures has stayed relatively constant whereas the annual construction rate of TSFs has increased, thus the cumulative failure rate of TSFs has continued to decline and, based on our review and analysis, may fall in the same order as the rate of LRFs. The cumulative fatality rate of TSF failures has reduced over time and is lower than the corresponding rate of LRFs. However, the size and the environmental impact of TSF failures have increased on average worldwide at a rate approximately proportional to the increased global rate of tailings production.

3. Heavy rainfall events and intensifying precipitation patterns are statistically important causative variables for the failures of both LRFs and TSFs. This has implications for the design and management of storage capacity, freeboard, facility drainage and spillways under climate change conditions.

Our results serve as quantitative benchmarks for estimating likelihoods of facility failures and their societal impacts to support future risk assessments. These results are applicable broadly on a global scale, having been conditioned by uncertainties in the data and the methods/approaches used to address existing data gaps. To improve the robustness of future statistical analyses, a more comprehensive public disclosure effort is necessary, particularly with respect to reservoir facility failures in China and constructed TSFs worldwide.

2.8 Data availability

Supplementary materials to this article are published in the following Borealis data repository (Rana et al., 2022b): https://doi.org/10.5683/SP3/K7SANT

2.9 References

Given that this thesis is in part article-based, the content of Chapter 2 represents material published in the following citations:

Rana, N.M., Ghahramani, N., Evans, S.G., Small, A., Skermer, N., McDougall, S., Take, W.A., 2022a. Global magnitude-frequency statistics of the failures and impacts of large water-retention dams and mine tailings impoundments. Earth-Sci. Rev. 104144. https://doi.org/10.1016/j.earscirev.2022.104144.

Rana, N.M., Ghahramani, N., Evans, S.G., Small, A., Skermer, N., McDougall, S., Take, W.A., 2022b. Supplementary materials to "Global magnitude-frequency statistics of the failures and impacts of large water-retention dams and mine tailings impoundments". Borealis, v2. https://doi.org/10.5683/SP3/K7SANT.

Rana, N.M., Ghahramani, N., Evans, S.G., McDougall, S., Take, W.A., 2021d. The magnitude-frequency of tailings flows resulting from failures of tailings impoundments. In: Proceedings of the 2021 International Tailings and Mine Waste Conference, Vancouver, Canada.

Chapter 3: Catastrophic Mass Flows from Tailings Dam Failures

Overview:

Tailings dam failures have received significant attention in recent years due to the catastrophic downstream consequences, as evidenced by the 2019 Feijao disaster in Brazil and numerous precedents. This work presents a timely review of tailings flows with the support of a comprehensive global database of 63 cases that have been remotely analyzed through a compilation of satellite imagery, digital elevation models and literature. The synthesis provides insight into the influence of impoundment conditions, preconditioning and trigger variables, failure mechanisms and the downstream environment on tailings flow behavior. The database also sheds light on the limitations of data quality and availability in the public domain. Magnitude-frequency statistics indicate that tailings dam breaches that have produced catastrophic mass flows with total outflow volumes of ≥1 M m^3 have occurred at a mean recurrence interval of 2–3 years over the period 1965–2020. Weather hazards and impoundment drainage issues are identified as major causative variables. The occurrence of liquefaction and/or the incorporation of free water are sufficient conditions to trigger extremely rapid, highly mobile behavior. Travel path confinement and steeper bed slopes enhance flow velocities (peak of 25–30 m/s) and kinetic energy, whereas flow mobility appears to be exacerbated along major rivers. Although general trends may be observed in empirical observations, such efforts are prone to substantial uncertainty due to the complexity and variability of site conditions (that are typically unaccounted for in broad statistical approaches) as well as poor data availability and/or quality for many of the selected cases. This highlights the importance of performing site-specific investigations through numerical models, laboratory tests and field observations to better predict post-breach behavior (ideally within a probabilistic framework) when undertaking site assessments.

3.1 Introduction

3.1.1 Preamble

The production and safe storage of tailings is an integral waste management procedure of standard mining projects as well as some industrial and power plant operations. Tailings are fine-grained waste rocks that are conventionally discharged as slurries into permanent storage facilities (Vick, 1983; Blight, 2010; Kossoff et al., 2014; Jefferies and Been, 2016). Given that tailings disposal is necessary insofar as operations continue, tailings dams are perpetually evolving structures that are raised in stages over the span of several years to decades, with some impoundments being among the largest man-made structures in the world (Vick, 1983; ICOLD, 2001). When unfavorable conditions persist in a storage

facility (i.e. when a large volume of tailings are very wet, contractive and/or overlain by ponded water), a triggering mechanism may initiate a dam breach resulting in a mass movement of failed materials into the downstream environment (Dobry and Alvarez, 1967; Chandler and Tosatti, 1995; Blight, 2010; Morgenstern et al., 2015, 2016; Robertson et al., 2019).

Voluminous tailings flows are capable of catastrophic, highly mobile behavior, with some events rendering considerable infrastructural damage, geomorphological impact, environmental contamination and socio-economic cost (Troncoso et al., 1993; Kossoff et al., 2014; Takahashi, 2014; Macias et al., 2015; Bowker and Chambers, 2015, 2017; Cuervo et al., 2017; de Lima et al., 2020). Some tailings flows have also led to mass casualties when travelling through inhabited areas, with a few examples including 1937 Los Cedros in Mexico (300+ deaths; Macias et al., 2015), 1965 El Cobre in Chile (200+ deaths; Dobry and Alvarez, 1967), 1985 Stava in Italy (268 deaths; Chandler and Tosatti, 1995; Takahashi, 2014) and 2019 Feijão in Brazil (272 deaths; Robertson et al., 2019; de Lima et al., 2020). The occurrence of three high-profile incidents (Mt. Polley in Canada and Fundão and Feijão in Brazil) over a span of just five years (2014–2019) has further raised concerns about large at-risk facilities (Morgenstern et al., 2015, 2016; Bowker and Chambers, 2017; Cuervo et al., 2017; Morgenstern, 2018; Robertson et al., 2019; Santamarina et al., 2019; de Lima et al., 2020; ICMM et al., 2020; Franks et al., 2021; Chovan et al., 2021). Two recent examples of catastrophic tailings flows are presented in Figure 3-1 (Petley, 2020a, 2020b).

A seemingly high frequency of such incidents has generated public scrutiny of the tailings management industry (Azam and Li, 2010; Morgenstern, 2018; Santamarina et al., 2019; ICMM et al., 2020; WISE, 2020; Franks et al., 2021; Chovan et al., 2021). These events continue to add to a growing list of tailings dam failures that includes over 350 reported incidents (ICOLD, 2001; Azam and Li, 2010; Lyu et al., 2019; Ghahramani et al., 2020; WISE, 2020). These failure cases also occur at a time of increasing mine waste production commensurate with declining ore grades and increasing demand for metals (ICOLD, 1996; IIED, 2002; Mudd, 2007; Jones and Boger, 2012; Mudd and Boger, 2013; Adiansyah et al., 2015; Owen et al., 2020), thus raising an urgent need to reduce the frequency and consequences of these incidents.

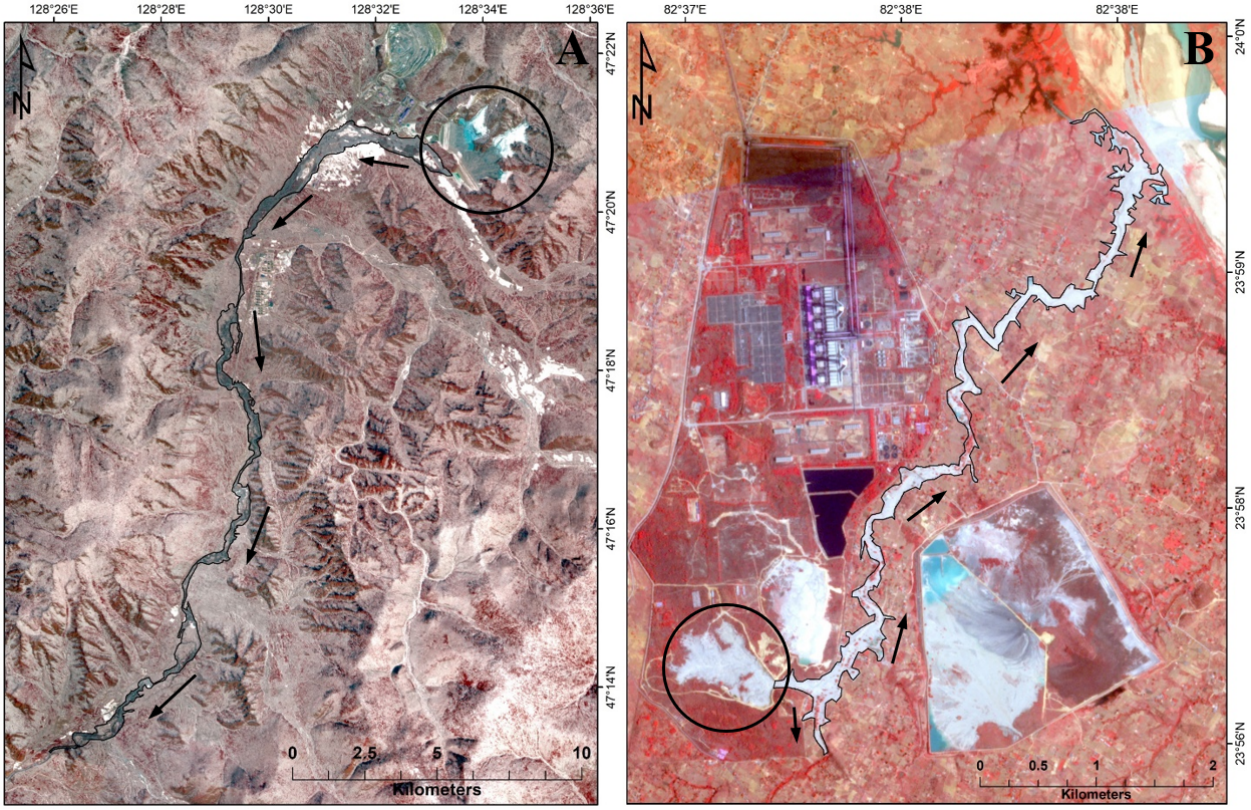

Figure 3-1: PlanetScope (3 m resolution) satellite images (infrared band) of two catastrophic tailings flows reported in the first half of 2020. [A] On March 28, the failure of the decant drainage system under the Luming molybdenum tailings facility in China triggered a catastrophic runout of 2.53 M m^3 of materials into the mining camp and down Yijimi River (Petley, 2020a). [B] On April 10, a corner breach of a fly ash retaining dam in India released a significant volume of materials down a river channel, reportedly killing at least six people across several engulfed villages (Petley, 2020b). Locations of the failed dams are circled while black arrows indicate flow directions. Raw image data courtesy of Planet Labs.

3.1.2 Scope and objectives

Geotechnical assessments have conventionally been carried out to retrospectively examine the root causes of catastrophic tailings dam failures, with particular focus on the stability of the embankment, tailings and foundation (e.g. Fourie et al., 2001; Olalla and Cuellar, 2001; Morgenstern et al., 2015, 2016; Robertson et al., 2019). This work has contributed to the steady development in the conceptual understanding of mining geotechnics over time (Bishop, 1973; Vick, 1983; Blight, 2010; Jefferies and Been, 2016). Comparatively less research attention has been directed towards the post-failure behavior of tailings within the context of high-mobility mass movements (e.g. Takahashi, 1991, 2014; Macias et al., 2015; Cuervo et al., 2017; de Lima et al., 2020). Present literature on tailings flows has mainly been

limited to outflow-runout empirical assessments (e.g. Rico et al., 2008b; Larrauri and Lall, 2018; Ghahramani et al., 2020) with some case studies devoted to characterizing runout behavior (e.g. Blight et al., 1981; Takahashi, 1991, 2014; Macias et al., 2015; de Lima et al., 2020). However, it remains to be holistically examined through a case history perspective how the field behavior (i.e. post-breach outflow and runout behavior) of tailings flows may be influenced by variable predisposing hydro-geotechnical conditions, trigger and breach mechanisms and downstream terrain properties.

To address this knowledge gap, this paper introduces a novel comprehensive database of 63 tailings flows (listed in Table 3-1) that ensued from failed tailings facilities between 1928 and 2020. The dataset – provided as a supplementary Excel file – considers tailings and impoundment properties, predisposal and triggering factors, breach-outflow characteristics and downstream runout behavior (contingent upon data availability). In a supplementary article, we describe the case selection criteria, data procurement methods and sources, data uncertainty and remote sensing-GIS techniques. Using this wealth of information, this paper undertakes a timely global review of the complex conditions and processes involved in the initiation and field behavior of tailings flows. This synthesis also presents some empirical observations while evaluating the limitations associated with broad, deterministic approaches to predicting site-specific post-breach behavior. In this regard, we must note that the database is not intended to provide revised empirical correlations to be unconditionally applied in breach-runout analyses, but rather to provide assistance to practitioners performing site assessments identify case histories with characteristics and pre-failure conditions similar to the dam under consideration. Given the number of terminologies that are presented in this comprehensive work, Table 3-2 lists the descriptions of relevant variables in order to promote consistency and allay potential misinterpretation of key parameters.

Table 3-1: Chronologically sorted list of 63 selected tailings flows from 1928–2020. The complete database and associated descriptions of case selection criteria, data procurement methods, data uncertainty and remote sensing-GIS techniques are provided as supplementary material. The case IDs are used as data point indicators in plots presented in Section 3.7.

ID	Site / Event	Country	Year	ID	Site / Event	Country	Year
1	Barahona	Chile	1928	33	Inez	USA	2000
2	Los Cedros	Mexico	1937	34	Sasa	Macedonia	2003
3	Huogudu	China	1962	35	Comurhex	France	2004
4	Bellavista	Chile	1965	36	Cavendish	UK	2007

5	Cerro Negro	Chile	1965	37	Mineracao	Brazil	2007
6	El Cobre	Chile	1965	38	Kingston	USA	2008
7	La Patagua	Chile	1965	39	Xiangfen	China	2008
8	Los Maquis	Chile	1965	40	Karamken	Russia	2009
9	Sgorigrad	Bulgaria	1966	41	Ajka	Hungary	2010
10	Hokkaido	Japan	1968	42	Caudalosa	Peru	2010
11	Certej	Romania	1971	43	Las Palmas	Chile	2010
12	Cities Service	USA	1971	44	Kayakari	Japan	2011
13	Bafokeng	South Africa	1974	45	Gullbridge	Canada	2012
14	Deneen Mica	USA	1974	46	Philex	Philippines	2012
15	Arcturus	Zimbabwe	1978	47	Talvivaara	Finland	2012
16	Mochikoshi	Japan	1978	48	Obed Mtn.	Canada	2013
17	Church Rock	USA	1979	49	Dan River	USA	2014
18	Tyrone Mine	USA	1980	50	Mt. Polley	Canada	2014
19	Cerro Negro	Chile	1985	51	Fundão	Brazil	2015
20	Niujiaolong	China	1985	52	Luoyang	China	2016
21	Stava	Italy	1985	53	Jharsuguda	India	2017
22	Veta del Agua	Chile	1985	54	Mishor Rotem	Israel	2017
23	Ollinghouse	USA	1985	55	Tonglvshan	China	2017
24	Stancil	USA	1989	56	Cadia	Australia	2018
25	Merriespruit	South Africa	1994	57	Cieneguita	Mexico	2018
26	Tapo Canyon	USA	1994	58	Cobriza	Peru	2019
27	Omai	Guyana	1995	59	Essar	India	2019
28	Marcopper	Philippines	1996	60	Feijão	Brazil	2019
29	Pinto Valley	USA	1997	61	Hindalco	India	2019
30	Aznalcóllar	Spain	1998	62	Luming	China	2020
31	Aitik	Sweden	2000	63	Singrauli	India	2020
32	Baia Mare	Romania	2000				

Table 3-2: A list of important terminologies presented in our collated research on tailings flows. This summary table is aimed to promote consistency and allay potential misinterpretation of key parameters. Further details on these variables, including data procurement methods, sources and uncertainty, are provided throughout the present manuscript and supplementary materials.

Term (Notation)	**Definition / Description**
Tailings	Fine-grained, wet to saturated waste rocks that are conventionally impounded behind a dam. In the present work, we define "tailings" to include both tailings solids and interstitial (pore) water. Although predominantly a product of mining operations, we also consider fine-grained, wet to saturated impounded waste from industrial and power plant operations (e.g. bauxite and fly ash) to be classed as tailings in this work.
Supernatant pond	Ponded free water overlying the tailings within the impoundment. The supernatant pond consists of process water used during mineral extraction and may also include naturally sourced free water such as rainfall and snowmelt.

Continued from Table 3-2

Total impounded volume (V_T)	The total volume of tailings and supernatant pond that are impounded behind a dam prior to a breach.
Impounded volume of supernatant pond (T_F)	The total volume of the supernatant pond that is impounded prior to a dam breach.
Impounded volume of tailings (T_T)	The total volume of tailings that is impounded prior to a dam breach.
Total outflow volume (V_F)	The total volume of tailings and supernatant pond that flows through a dam breach. The total outflow volume equals to the sum of the outflow volume of tailings and the outflow volume of the supernatant pond.
Total outflow ratio (V_F/V_T)	The ratio of total outflow volume to total impounded volume at a breached dam site, typically expressed as a percentage.
Mean dam breach width (B_w)	The average horizontal width of the breach opening as estimated from GIS measurements on satellite imagery or as reported in literature for cases without available satellite imagery.
Zone 1	Primary impact zone: the extent of the main solid tailings deposit characterized by remotely visible or field-confirmed sedimentation over the downstream terrain and above mean bankfull elevations if extending into downstream watercourses. The present work reports measurements of runout distance, inundation area and travel path angle in Zone 1. Refer to Figure 3-8, supplementary materials and Ghahramani et al. (2020) for further details on Zone 1 and Zone 2.
CDA class	A conceptual framework devised by the Canadian Dam Association (CDA) that categorizes breach-flow events into four classes (CDA, 2020): 1A (liquefied tailings with supernatant pond), 1B (non-liquefied tailings with supernatant pond), 2A (liquefied tailings without supernatant pond) and 2B (non-liquefied tailings without supernatant pond). The assignment of CDA classes is a subjective endeavour, particularly for complex breach events. The present work applies this framework primarily as a qualitative indicator variable to observe the influence of tailings liquefaction and supernatant pond on tailings flow behavior.
Travel path confinement	A qualitative attribute of the downstream terrain that indicates whether the travel path was channelized (i.e. laterally confined), terminally confined and/or unconfined. If a tailings flow runout was partially channelized and unconfined, we grouped the case according to the dominant mode of confinement.
Travel path substrate	A qualitative attribute of the downstream environment that describes the travel path substrate based on GIS interpretation of satellite imagery. The present work applies this attribute primarily as an indicator variable to observe the effects of travelling on major river channels on tailings flow mobility.

3.2 Tailings storage facilities

Mining operations involve the development of a geological ore body to extract minerals while producing vast quantities of mine waste. The nature of this waste varies from dry, angular, very coarse (sandy gravel to boulder-sized) waste rocks stripped from the overburden to very wet, fine-grained (sandy to silty) tailings that comprise the residuum of the comminution process (Blight, 2010). Although predominantly produced from mining activities, tailings are also generated from some industrial and power plant operations (e.g. bauxite from alumina and fly ash from coal). These materials are also considered as tailings in this work, whereby we define "tailings" to include both tailings solids and interstitial (pore) water (Table 3-2). The types of tailings may be generally grouped as either hard rock (e.g. lead, zinc, copper, gold, silver, molybdenum, iron, nickel and uranium) or soft rock (e.g. coal, fly ash, phosphate and bauxite) depending on the specific gravity (Vick, 1983; Small et al., 2017; Ghahramani et al., 2020). The use of hard and soft rock tailings as a qualitative indicator variable (as in Section 3.7 of this chapter) has previously been recommended as a simpler way to distinguish the type of ore due to the numerous distinct types of tailings and the data sparsity corresponding to each type (Small et al., 2017; Ghahramani et al., 2020). As reflected in our dataset of tailings flows, the number of data points still remains very low for soft rock tailings (n=6) compared to hard rock tailings (n=33).

To protect against environmental contamination, tailings are conventionally impounded behind a compacted dam that is typically built out of the coarse (sandy) fraction of tailings, foundation materials and/or waste rockfill/earthfill (Blight, 2010; Kossoff et al., 2014). Tailings are typically hydraulically discharged as a low-density slurry into the impoundment, with the finer materials (slimes) settling further into the central core while the coarser materials make up the beach abutting the dam. To prevent the generation of surface dusts and acid mine drainage through oxidation processes, the tailings slimes may form the base of a supernatant pond consisting of process water used during mineral extraction (Figure 3-2; Blight, 2010; Kossoff et al., 2014). Tailings storage facilities may be constructed as partially enclosed impoundments on a slope, across a valley or as fully enclosed impoundments on relatively flat terrain. In keeping with the rising volume of the impounded material, tailings dams may be raised in the upstream direction (i.e. over the impounded tailings), in the downstream direction or vertically upwards using the centerline method (Figure 3-2; Vick, 1983; ICOLD, 2001; Blight, 2010).

It is worth noting that the rate of tailings deposition, consolidation and dam construction varies considerably depending on the site, hence impoundment age does not necessarily correlate with the total impounded volume (and by extension total outflow volume and flow mobility), as reflected in our dataset of tailings flows. Nevertheless, over time, tailings dams may impound tens of M m^3 of tailings, and in many cases supernatant ponds, over heights of tens of m, thus creating a long-term anthropogenic footprint on a previously natural landscape. Large impoundments thereby represent sources of significant potential energy capable of extreme downstream geomorphological, geochemical and socio-economic consequences in the event of catastrophic failure (Dobry and Alvarez, 1967; Takahashi, 1991; Kossoff et al., 2014; Macias et al., 2015; Cuervo et al., 2017; de Lima et al., 2020).

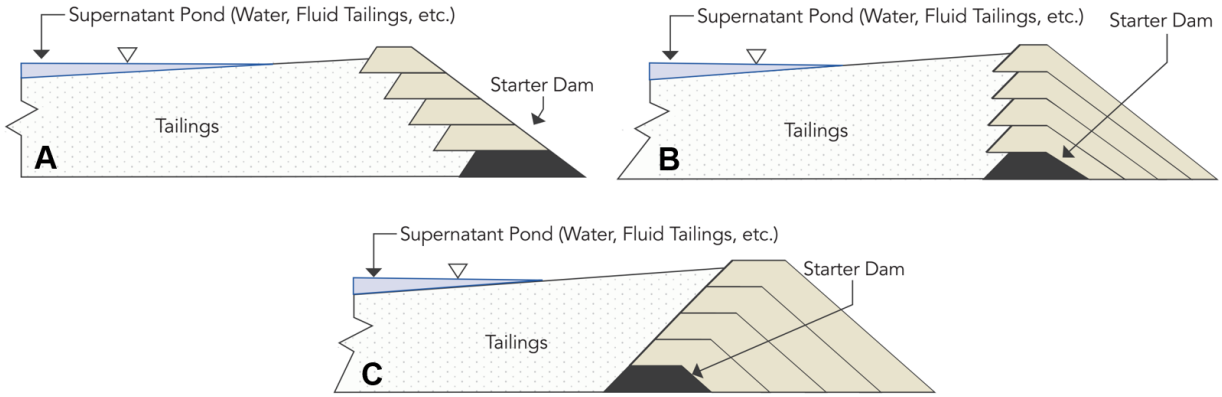

Figure 3-2: Schematic of the conventional construction (raising) methods of tailings dams (reproduced with permission from CDA, 2020). [A] is upstream, [B] is centerline and [C] is downstream-raised tailings dam.

3.3 Susceptibility to mass flow behavior

The practice of hydraulically disposing tailings into a large wet impoundment, without allowing for the necessary time for settlement and consolidation, may promote a loose, saturated tailings structure that is susceptible to brittle stress-strain behavior and liquefaction following a triggering mechanism (Martin and McRoberts, 1999; Blight, 2010; Jefferies and Been, 2016). Numerous case studies (e.g. Dobry and Alvarez, 1967; Fourie et al., 2001; Morgenstern et al., 2016; Robertson et al., 2019) indicate that a high propensity for tailings to liquefy *en masse* – whether induced statically or seismically – is typically congruent with high moisture contents approaching or exceeding the liquid limit (i.e. liquidity index > 1.0) and positive state parameters indicating contractive properties. The state parameter (ψ) is

the difference between the existing void ratio and the critical void ratio (i.e. the void ratio at which no volume change occurs during shear as determined from reconstituted samples) of the tailings at the same mean effective stress (Jefferies and Been, 2016). In saturated conditions, loose, contractive materials ($+\psi$) are potentially liquefiable upon undrained shearing whereas dilative materials ($-\psi$) are more compact and demonstrate greater resistance to liquefaction (Vick, 1983; Martin and McRoberts, 1999; Fourie et al., 2001; Blight, 2010; Morgenstern et al., 2016). During flow liquefaction, impounded tailings undergo rapid strain-softening (i.e. progressively lower shear stress is needed to induce the same level of strain) as the interstitial water begins to carry part of the load during material contraction until residual strength is attained. This initiates rapid flow-like movement of the liquefied mass as it escapes through the breach into the downstream environment (Morgenstern et al., 2016).

Several tailings flows have also initiated due to the uncontrolled or accidental release of a supernatant pond, irrespective of the liquefaction susceptibility of the impounded tailings. Cases such as 1937 Los Cedros and 2014 Mt. Polley have involved the prolonged, rapid discharge of free water with eroded non-liquefied tailings, displaying field behavior akin to a sediment-rich outburst flood (Macias et al., 2015; Morgenstern et al., 2015; Cuervo et al., 2017). When combining the effects of flow liquefaction with the availability of a voluminous pond, the resulting outflow may be typified by additional fluidization, catastrophic discharge rates, multiple flow surges, larger outflow volumes and widespread downstream inundation (CDA, 2020). Still, the 2015 Fundão, 2016 Luoyang and 2019 Feijão runouts have demonstrated that the occurrence of liquefaction is a sufficient condition to initiate catastrophic mass flows, even without the influence of free water (Reid and Fourie, 2017; Palu and Julien, 2019; de Lima et al., 2020).

Given this knowledge, the Canadian Dam Association (CDA) introduced a classification matrix that attempts to categorize breach events into one of four distinct classes based on the involvement of free water within the outflow and the occurrence of tailings liquefaction (Figure 3-3; Small et al., 2017; Martin et al., 2019; CDA, 2020; Gildeh et al., 2020). This sets a useful conceptual framework that may also be adopted as an indicator variable for tailings dam breach-runout prediction studies. The present work applies the CDA framework purely as a qualitative indicator variable to observe the influence of tailings liquefaction and supernatant pond on tailings flow behavior. It should be noted, however, that the task of assigning CDA classes is prone to subjectivity particularly for complex cases where, for

instance, (i) tailings liquefaction only occurred locally, (ii) the volume of free water was small enough to question the corresponding influence on runout behavior, (iii) the failure occurred in multiple stages of different CDA classes or (iv) rather than a standard dam breach, the failure manifested due to a deficient drainage system that produced a prolonged discharge into the downstream terrain.

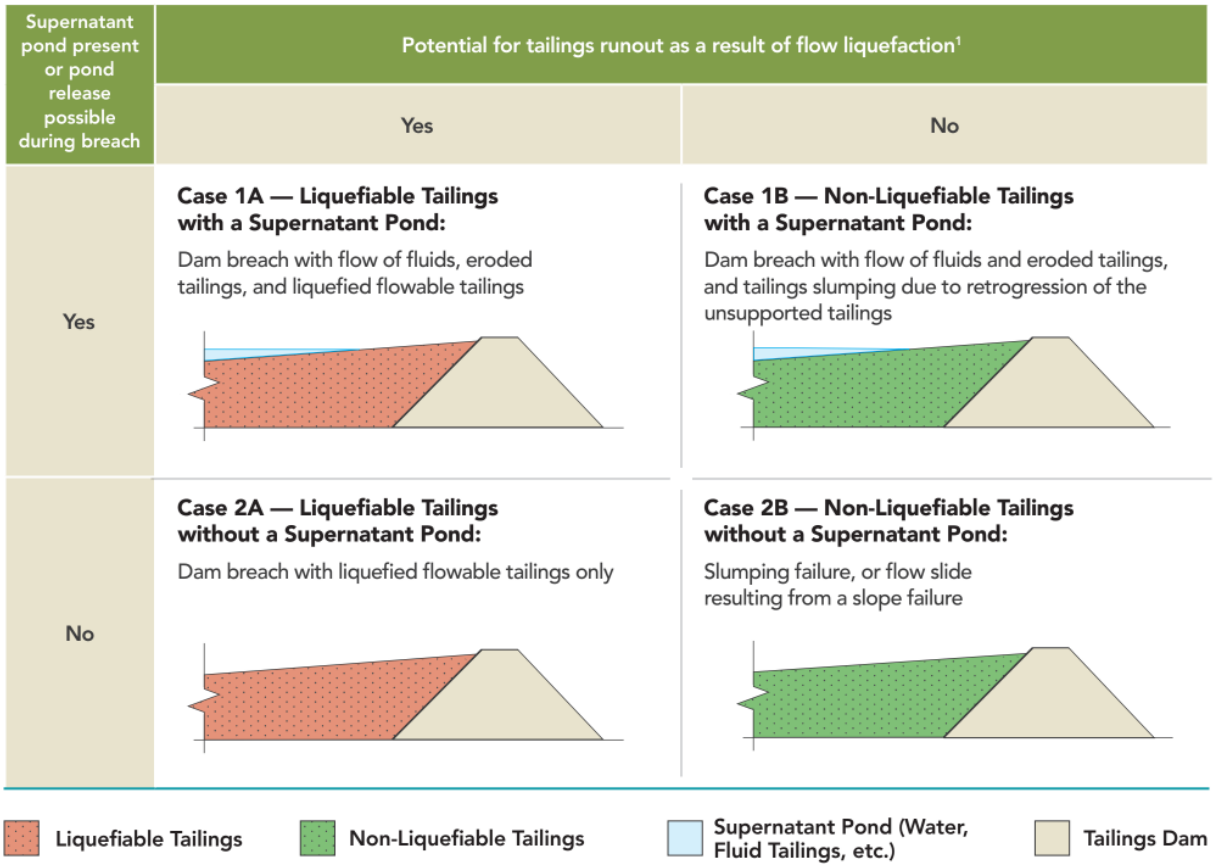

Figure 3-3: Schematic of the Canadian Dam Association (CDA) Tailings Dam Breach Analysis (TDBA) classification system for tailings dam failures (reproduced with permission from CDA, 2020). Note: (1) regardless of failure mode, flow liquefaction indicates the potential for tailings flow following a breach. Free water outflow induces erosive transport of tailings (whether liquefied or not) through the breach whereas liquefaction triggers a sudden loss of tailings strength through the mobilization of pore waters. As a result, both variables contribute to extremely rapid, highly mobile post-breach behavior.

3.4 Preconditioning and trigger variables

When tailings impoundments are in a precarious state, the failure process may be accelerated by one or multiple predisposing factors, resulting in a marginally stable dam that fails catastrophically upon the onset of even minor triggering mechanisms (Morgenstern et al., 2015, 2016; Robertson et al., 2019). Table 3-3 lists variables that may compromise tailings dam integrity with reference to corresponding adverse effects and failure modes. Whereas heavy rainfall, seismic liquefaction and foundation instability are common "external" trigger variables (Rico et al., 2008a; Azam and Li, 2010; Lyu et al., 2019), a metastable state may be preconditioned by "internal" anthropogenic root causes related to the engineering design, operations or regulations (Morgenstern et al., 2018). In particular, adequate drainage of interstitial and free water and sufficient compaction of the dam core are considered vital internal variables that may help mitigate the adverse effects of rain-induced overtopping, seismic/static liquefaction, internal erosion of the dam core and excessive groundwater seepage (Azam and Li, 2010; Lyu et al., 2019; Clarkson and Williams, 2021).

A significant percentage of failed tailings dams have been those that were raised in the upstream direction (Lyu et al., 2019; Franks et al., 2021), including 37 of the 46 tailings flow cases with reported dam raising methods in our database. The upstream design has thus received scrutiny, including a prohibition in Chile since 1970 after a series of earthquake-triggered failures (Dobry and Alvarez, 1967; Villavicencio et al., 2014; Valenzuela, 2016; Troncoso et al., 2017). Failures of some upstream-raised dams have been attributed to partial founding on soft, compressible, poorly consolidated slimes that consequently underwent lateral extrusion, undrained creep and development of excess pore water pressures to the point of failure (Dobry and Alvarez, 1967; Troncoso et al., 1993; AECOM, 2009; Morgenstern et al., 2016). However, as counter-argued by Vick (1992), Martin and McRoberts (1999) and Morgenstern (2018), the prominent causes of these instabilities may lie in improper design, construction and operation of the facility – including, but not limited to, the overabundance of contractive materials, insufficient drainage and seepage control and a loose embankment core – rather than the general upstream method itself.

Figure 3-4 shows the temporal bar chart distribution of the five major causative variables for 57 tailings dam failures that produced mass flows, as listed in our database. The variables are grouped as follows:

- “Weather hazards” include heavy rainfall, snowfall, storms and freeze-thaw cycles. These factors may lead to a number of failure mechanisms such as overtopping, internal erosion (piping), seepage, excess pore pressures and structural drainage issues.
- “Embankment deficiency” highlights structural issues of the tailings dam core (i.e. lack of compaction, use of improper construction materials and/or oversteepened slopes) that may promote internal erosion, piping, excess seepage and/or slope instability.
- “Seismic activity” refers to earthquake-induced liquefaction failures or seismic tremors that accelerate the impoundment towards failure.
- “Unstable foundation” refers to impoundment foundations that were hydrogeologically, geotechnically and/or topographically unsuitable to sustain high, rapid vertical loading.
- “Structural drainage deficiency” refers to deficiencies within the interstitial and supernatant pond drainage system, resulting in static liquefaction, overtopping or structural failure.

It is worth noting that some failures were triggered by a complex interplay of multiple variables, and thus assigning a singular trigger to a complex breach case may be a matter of subjectivity. In addition, Figure 3-4 excludes deep-rooted, mainly anthropogenic predisposal variables related to engineering, operational or regulatory issues (Morgenstern, 2018). Weather hazards have contributed to almost a third (32%) of the cases, followed by seismic activity (21%), structural drainage deficiency (19%), embankment deficiency (16%) and unstable foundation (12%). Tailings flows triggered by weather hazards have steadily increased in frequency over time, which is consistent with the findings of Rico et al. (2008a), Azam and Li (2010) and Lyu et al. (2019) in relation to tailings dam failures in general. The meteorological effect on tailings dam stability may be manifested in several ways. For instance, heavy rainfall events may trigger erosional failures of embankments through pond overtopping (Midgley, 1979; Wagener, 1997). Long-term increases in regional rainfall may trigger loss of suction in unsaturated zones and lead to undrained conditions (Robertson et al., 2019). Freeze-thaw cycling may cause structural drainage deficiency and loosening of the embankment core, while spring-time melting of snow and ice covers on the impoundment may reduce the freeboard between the dam crest and impounded materials and exacerbate undrained conditions (Chandler and Tosatti, 1995; UNEP and

OCHA, 2000). Extensive ice formation contributed to the recent catastrophic failure of the decant drainage system at the Luming tailings facility in China (Figure 3-1A; Petley, 2020a).

Table 3-3: Reported variables that may precondition or trigger tailings dam instability.

Variables	Adverse Effects	Case Examples	References
Seismic activity	- Tailings liquefaction - Acceleration of creep deformation	1928 Barahona; 1965 El Cobre; 2015 Fundão	Dobry and Alvarez (1967); Troncoso et al. (1993); Morgenstern et al. (2016)
Intense precipitation	- Higher liquefaction susceptibility (loss of suction) - Higher potential for pond overtopping - Structural failure of drainage system - Excessive seepage	1985 Stava; 1974 Bafokeng; 1994 Merriespruit	Chandler and Tosatti (1995); Midgley (1979); Wagener (1997)
Embankment deficiency	- Loose dam core may cause internal erosion (piping) and seepage - Oversteepened dam slope may trigger slope instability	1974 Bafokeng; 1995 Omai	Midgley (1979); Vick (1996)
Groundwater seepage	Excess pore pressure development and higher liquefaction susceptibility along base of impoundment	1966 Sgorigrad; 1985 Stava; 2019 Feijão	Lucchi (2009); Chandler and Tosatti (1995); Robertson et al. (2019)
Static liquefaction	- May increase shear stresses against dam to the point of catastrophic collapse - May occur in response to sudden loss of confinement (static unloading)	2015 Fundão; 2019 Feijão	Morgenstern et al. (2016); Robertson et al. (2019)
Foundation instability	Sudden undrained brittle failure of dam foundation triggering dam breach	1998 Aznalcóllar; 2014 Mt. Polley; 2018 Cadia	Alonso and Gens (2006a); Morgenstern et al. (2015); Jefferies et al. (2019)
Internal creep deformation	Undrained failure when peak strength of tailings is exceeded	2008 Kingston; 2019 Feijão	AECOM (2009); Robertson et al. (2019)
Tailings properties	- Strong bonding of stiff tailings may lead to higher peak strength and higher susceptibility to sudden brittle failure - Abundance of clay minerals may induce geochemical reactions and higher thixotropy - High-pH pore and free water may cause caustic corrosion of dam core	2009 Karamken; 2010 Ajka; 2015 Fundão; 2019 Feijão	Mecsi (2013); Morgenstern et al. (2016); Bánvölgyi (2018); Glotov et al. (2018); Robertson et al. (2019)
Anthropogenic factors	- Rapid tailings deposition may inhibit tailings consolidation - Mine blasting vibrations may accelerate creep deformation - Vandalism may cause structural issues - Engineering, operations and regulator mismanagement (e.g. inappropriate dam construction, inadequate drainage system and sub-standard monitoring)	1965 El Cobre New Dam; 2009 Karamken; 2015 Fundão	Dobry and Alvarez (1967); Glotov et al. (2018); Morgenstern et al. (2016); Morgenstern (2018)

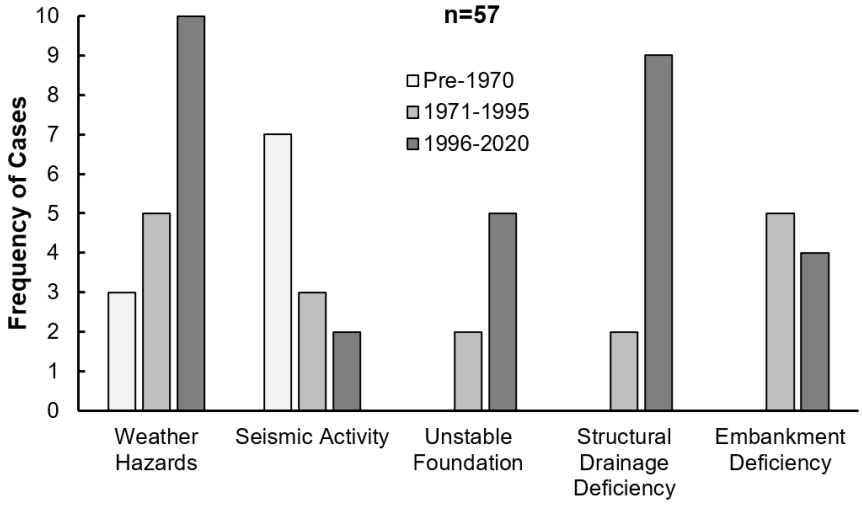

Figure 3-4: Temporal bar chart of major causative variables for 57 tailings dam failures that produced mass flows, as listed in our database. Refer to text for descriptions of each variable. It is worth noting that some failures were triggered by a complex interplay of multiple variables, and thus assigning a singular trigger to a complex breach case may be prone to subjectivity.

The stability of geological foundations underneath large impoundments has become of concern following the 1998 Aznalcóllar, 2014 Mt. Polley and 2018 Cadia incidents – all of which occurred with minimal advance warning (Gens and Alonso, 2006; Morgenstern et al., 2015; Jefferies et al., 2019). Such experiences have underscored the importance of identifying the weakest units in the foundation for consideration in tailings facility design. Moreover, the setting of large impoundments on mountain or valley slopes may yield significant changes to the local hydrogeologic regime, while inducing a perpetual state of high gravitational driving stresses against the embankment. The disasters of 1966 Sgorigrad, 1985 Stava and 2019 Feijão were all preconditioned in part by consistent groundwater seepage into the sloping foundations (Chandler and Tosatti, 1995; Lucchi, 2009; Robertson et al., 2019). Undetected groundwater upwelling was also a key contributory variable in the 1994 Tapo Canyon and 2010 Las Palmas dam breaches (Harder and Stewart, 1996; Villavicencio et al., 2014).

Tailings flows triggered by seismic activity have steadily reduced in frequency over time. A majority of early seismic liquefaction failure cases were in Chile and Japan (Torres and Brito, 1966; Dobry and Alvarez, 1967; Okusa et al., 1980; Ishihara, 1984); since then, both nations have advanced the state of operational practice such that tailings facilities have withstood powerful earthquakes (Yasuda et al.,

2013; Villavicencio et al., 2014; Ishihara et al., 2015; Valenzuela, 2016; Troncoso et al., 2017). A novel concern, however, is the potential role of sequential seismic tremors in accelerating the failure of marginally stable facilities (Agurto-Detzel et al., 2016; Morgenstern et al., 2016).

3.5 Magnitude–frequency

The documented total outflow volumes (V_F) in our database range from 20,000 m^3 to 32 M m^3, summing to a cumulative total of 136 M m^3. Based on our review of pre-existing datasets of tailings dam breaches and tailings flows (e.g. ICOLD, 2001; Rico et al., 2008b; Bowker and Chambers, 2015; Small et al., 2017; Larrauri and Lall, 2018; Ghahramani et al., 2020, WISE, 2020), we conclude that all tailings flows with $V_F \geq 200{,}000$ m^3 over the period 1965–2020 (totaling to 34 cases) have been publicly documented and are thus included in our database. Figure 3-5A illustrates the temporal cumulative V_F distribution of the 34 tailings flows that fall within this "complete interval". An estimated 40% of the cumulative V_F (~50 M m^3) occurred in the 49-year period between 1965 and 2013, whereas the remainder (60% or ~75 M m^3) occurred since 2014. This recent spike is dominated by the combined V_F of ~67 M m^3 from the 2014 Mt. Polley, 2015 Fundão and 2019 Feijão incidents.

A cumulative magnitude-frequency distribution of the complete interval of 34 cases is shown in Figure 3-5B. A rollover effect is observed whereby the slope of the power-law relation [Log F = -0.88 (log V_F) + 4.93] flattens below a V_F threshold of ~1 M m^3. This statistical rollover phenomenon has also characterized natural debris flows and is attributed to either data bias (undersampling) at smaller volumes or the manifestation of limiting physical conditions under which mass flows occur (Guthrie and Evans, 2004; Hungr et al., 2008). Using the power-law formula, we calculate that tailings dam breaches that have produced catastrophic mass flows with $V_F \geq 1$ M m^3 have occurred at a mean recurrence interval of 2–3 years over the period 1965–2020. For context and comparison, we also draw in Figure 3-5B a minimum-estimate magnitude-frequency curve of the 5-year planned volumes of 1,260 tailings facilities constructed globally over the period 1965–2019. This data is derived from a detailed but incomplete record in Franks et al. (2021) that was collated through voluntary survey responses from extractive companies. By extrapolation of the curve, it is estimated that a tailings facility with a planned 5-year storage volume of ≥ 200 M m^3 was constructed at least once every year over the period 1965–2019. The magnitude-frequency curves of tailings flows and constructed tailings facilities are noted to be similar in shape and form.

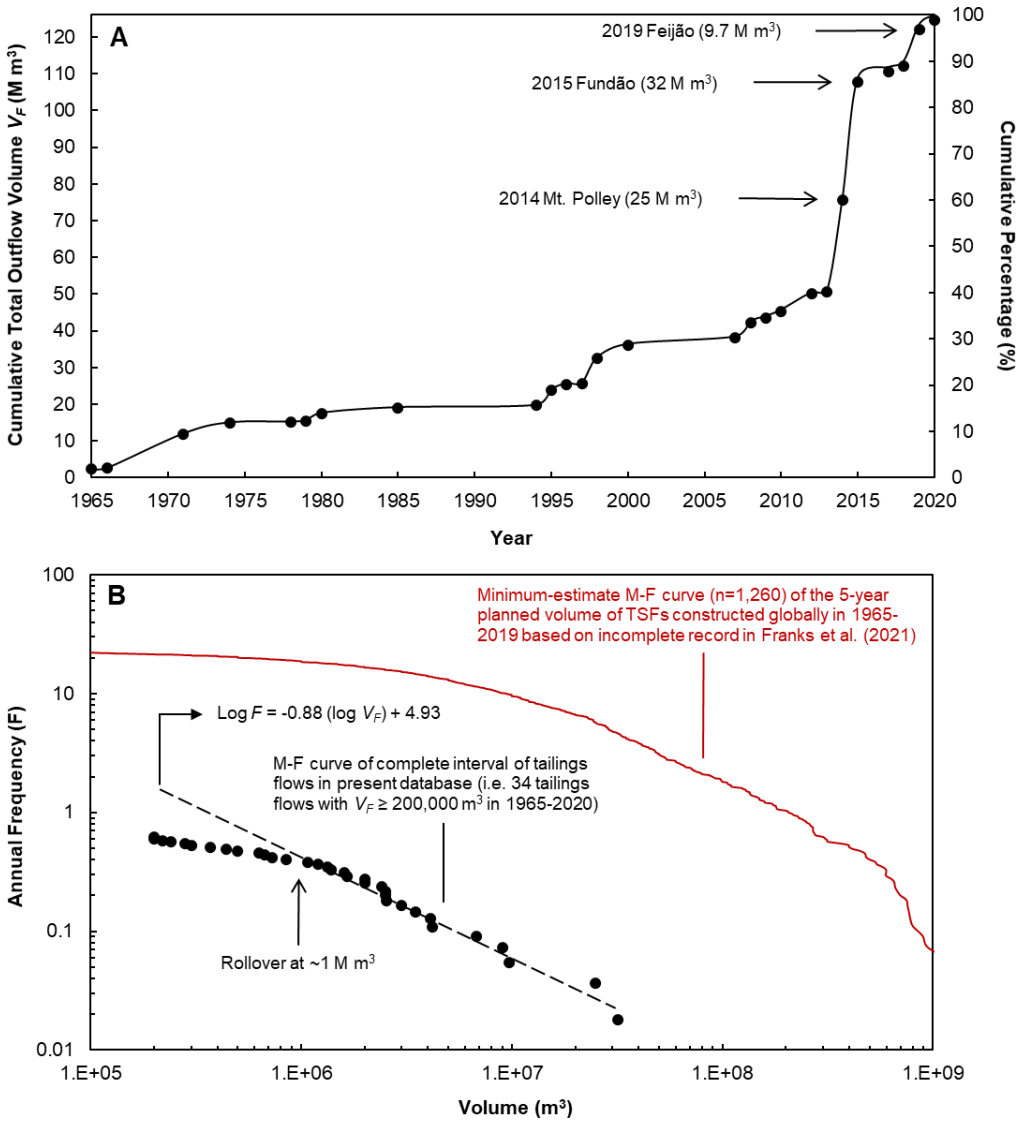

Figure 3-5: [A] Cumulative total outflow volume (V_F) distribution of the "complete interval" of our database (i.e. thirty-four tailings flows with $V_F \geq 200{,}000$ m³ in the period 1965–2020). The data points correspond to years in which one or more of the 34 events occurred, thus each data point represents the total sum of V_F of one or more of the 34 cases during that calendar year. [B] A comparison of the magnitude-frequency (M-F) curves of tailings flows and constructed tailings storage facilities (TSFs) since 1965. The 34 tailings flows comprising the complete interval of our database are represented by black solid circles. A rollover effect is observed at V_F ~1 M m³ below which the slope of the power-law relation flattens. The power-law formula for tailings flows with $V_F \geq 1$ M m³ is log F = -0.88 (log V_F) + 4.93 with an R^2 of 0.98. This implies that tailings dam breaches that have produced mass movements with $V_F \geq 1$ M m³ have occurred at a mean recurrence interval of 2–3 years over the period 1965–2020. The red solid line represents the minimum-estimate M-F curve of the 5-year planned volumes of 1,260 TSFs constructed globally over the period 1965–2019, with the data being derived from a detailed but incomplete record published by Franks et al. (2021). The M-F curves of tailings flows and constructed TSFs are noted to be similar in shape and form.

3.6 Initiation and field behavior

3.6.1 Breach development

An impending breach of a marginally stable tailings dam may be preceded by a prolonged period of small-scale deformations, cracking and/or seepage (Shakesby and Whitlow, 1991; Mecsi, 2013; Morgenstern et al., 2016; Jefferies et al., 2019; Rotta et al., 2020; Grebby et al., 2021). A catastrophic breach may be accompanied by a series of explosion-like noises and shockwaves signifying the detachment and collapse of large portions of the dam core (Torres and Brito, 1966; Blight and Fourie, 2005; Macias et al., 2015; de Lima et al., 2020). Global instability (i.e. near-complete failure of a dam) may occur rapidly in 10–20 s when triggered by liquefaction in upstream-raised, partially enclosed impoundments sitting on sloping terrain (typically ≥ 4º) (Dobry and Alvarez, 1967; Okusa et al., 1980; Colombo and Colleselli, 2003; Robertson et al., 2019; de Lima et al., 2020). In some cases, a vulnerable section of the dam may detach rapidly, slide some distance and progressively disintegrate in response to increased driving stresses from liquefied tailings (Torres and Brito, 1966; Alonso and Gens, 2006b; Jefferies et al., 2019). When the failure mode is overtopping and/or internal erosion of a ring-dyke facility, the breach may progressively develop over several hours prior to the catastrophic surge (Blight et al., 1981; Shakesby and Whitlow, 1991; Wagener, 1997). In such events, a fully formed breach channel typically attains the shape of an inverted trapezoid, achieving maximum width at the crest. In our database, we report the mean breach widths (defined in Table 3-2) as provided in case literature when applicable (e.g. Blight et al., 1981) or from our independent GIS measurements on available satellite imagery.

3.6.2 Outflow process

The post-breach outflow behavior of impounded materials is highly complex and depends on the failure mechanism and site conditions (Martin et al., 2015; CDA, 2020). Some insight into outflow processes can be gained from performing a case review that is categorized into CDA classes (Figure 3-3). As remarked earlier in Section 3.3, the effort of assigning CDA classes to breach events is prone to subjectivity depending on the application objective and case complexity. In this work, we apply the CDA framework purely as a qualitative indicator variable to observe the influence of tailings liquefaction and supernatant pond on tailings flow behavior.

3.6.2.1 1A Cases (Flow of Liquefied Tailings with a Supernatant Pond)

During 1A erosional breach cases, the mass flow may be preceded by a steady outflow of excess free water which may serve as a premonitory sign of impending failure (Midgley, 1979; Blight et al., 1981; Wagener, 1997; Blight and Fourie, 2005). In the 1994 Merriespruit case, where the total outflow volume was ~630,000 m^3, residents reported observing a strong flow of water passing through the village for ~2 hours prior to the catastrophic surge (Wagener, 1997). A significant proportion of the 70,000–100,000 m^3 of free water from the facility (30,000–50,000 m^3 of rainwater and 40,000–50,000 m^3 of supernatant pond) was incorporated within this initial steady flow (Wagener, 1997). The subsequent rapid surge consisted of statically liquefied tailings and residual free water. In the case of 1974 Bafokeng, following ~1.5 hrs of water seeping along the dam face, ~3 M m^3 of statically liquefied tailings (18% of Dam 1) were mobilized over 40–60 mins with a mean velocity and discharge of ~10 m/s and ~1,000 m^3/s, respectively, through the ~110 m wide breach (Blight et al., 1981).

During rapid 1A breach events, the total outflow ratios and peak discharges may vary widely depending on site conditions and failure modes. As a consequence of the 1965 Chilean earthquake, 45% of the El Cobre Old Dam (1.9 M m^3) escaped the facility whereas almost 100% of the freshly deposited, poorly consolidated tailings in the adjacent New Dam were mobilized (0.5 M m^3), resulting in a combined outflow of 2.4 M m^3. Eyewitnesses reported “waves of liquid columns” (interpreted as sand blows) on the New Dam tailings surface as a consequence of rapid pore pressure relief (Torres and Brito, 1966; Dobry and Alvarez, 1967). The 1985 Stava outflow, predisposed by inadequate drainage and the sloping foundation (12º–16º), achieved a catastrophic peak discharge of 28,160 m^3/s, indicating that it took just 13 s for 63% of the Upper and Lower Basins (~180,000 m^3) to be released (Takahashi, 1991, 2014). During the outflow sequence in such catastrophic 1A breaches, liquefaction flowsliding may continue for hours as a series of slope failures that regress backwards until stronger scarp materials are encountered, typically rendering steep, semi-circular scarps surrounding sub-horizontal (1º to 5º) wet terraces (Dobry and Alvarez, 1967; Shakesby and Whitlow, 1991; Blight and Fourie, 2005; Blight, 2010; Robertson et al., 2019; CDA, 2020).

A contrasting example was the 2009 Karamken incident (Russia) that was caused by heavy rainfall and lack of maintenance of the decommissioned facility (Glotov et al., 2018). Approximately 1.1 Mm^3 of floodwaters and 0.3 Mm^3 of liquefied tailings and dam materials outflowed through the ~50 m wide

breach at a relatively low peak discharge of ~10 m^3/s for ~30 minutes before reducing to ~1.4 m^3/s over the subsequent three days (Glotov et al., 2018).

3.6.2.2 1B Cases (Flow of Non-Liquefied Tailings with a Supernatant Pond)

In catastrophic 1B cases, the mass flow may occur in one or multiple surges lasting for several hours to a few days depending on free water availability (i.e. volume and distribution of the pond and, if applicable, the intensity and duration of rainfall). For example, the 1937 Los Cedros flow involved at least three major flow surges over a span of 48 hours with a total outflow volume of ~1.5 M m^3 (Macias et al., 2015). The flow pulses were triggered and sustained by a torrential rainstorm that lasted several days and provided a sufficient supply of free water to induce progressive softening and erosion of the impounded tailings. The peak discharge through the breach was retrospectively modelled to be ~8,000 m^3/s with a solids concentration of ~50% (Macias et al., 2015).

The Mt. Polley event also involved at least three flow surges (Morgenstern et al., 2015; Cuervo et al., 2017). The peak discharge through the ~160 m wide breach was calculated to be 5,130 m^3/s – about a thousand times higher than the probable maximum 200-year flood for Hazeltine Creek (Cuervo et al., 2017). Of the total outflow volume of 24.4 M m^3 (a third of the impoundment), 30% (7.3 M m^3) consisted of tailings solids, 27% (6.5 M m^3) was interstitial water and 43% (10.6 M m^3) comprised the supernatant pond (Morgenstern et al., 2015; Cuervo et al., 2017). Helicopter videos of the impoundment captured 8 hours after failure showed active sheet erosion of tailings, dendritic drainage channels and cirque-like arcuate headscarps (Figure 3-6; Morgenstern et al., 2015). As the steep scarps kept migrating headward due to fluvial erosion, the height drop in the headscarp enhanced pond drainage from farther ends of the facility, which in turn further promoted headscarp retreat akin to a geomorphic cascade sequence. In a natural landscape, this process may take hundreds to thousands of years whereas the Mt. Polley pond supply was exhausted in just a few hours post-breach (Morgenstern et al., 2015).

Occasionally in 1B cases, dam failure may occur not due to a breach but instead as a result of a deficient drainage system, triggering a prolonged single-stage discharge into the downstream environment. An example is the 1995 Omai case that was caused by blockage and damage of drainpipes that in turn induced internal erosion of the improperly constructed dam core (Vick, 1996). The resulting cyanide-laden leakage lasted for almost five days, with the initial peak discharge to the adjacent Omai River estimated to be ~50 m^3/s (Vick, 1996).

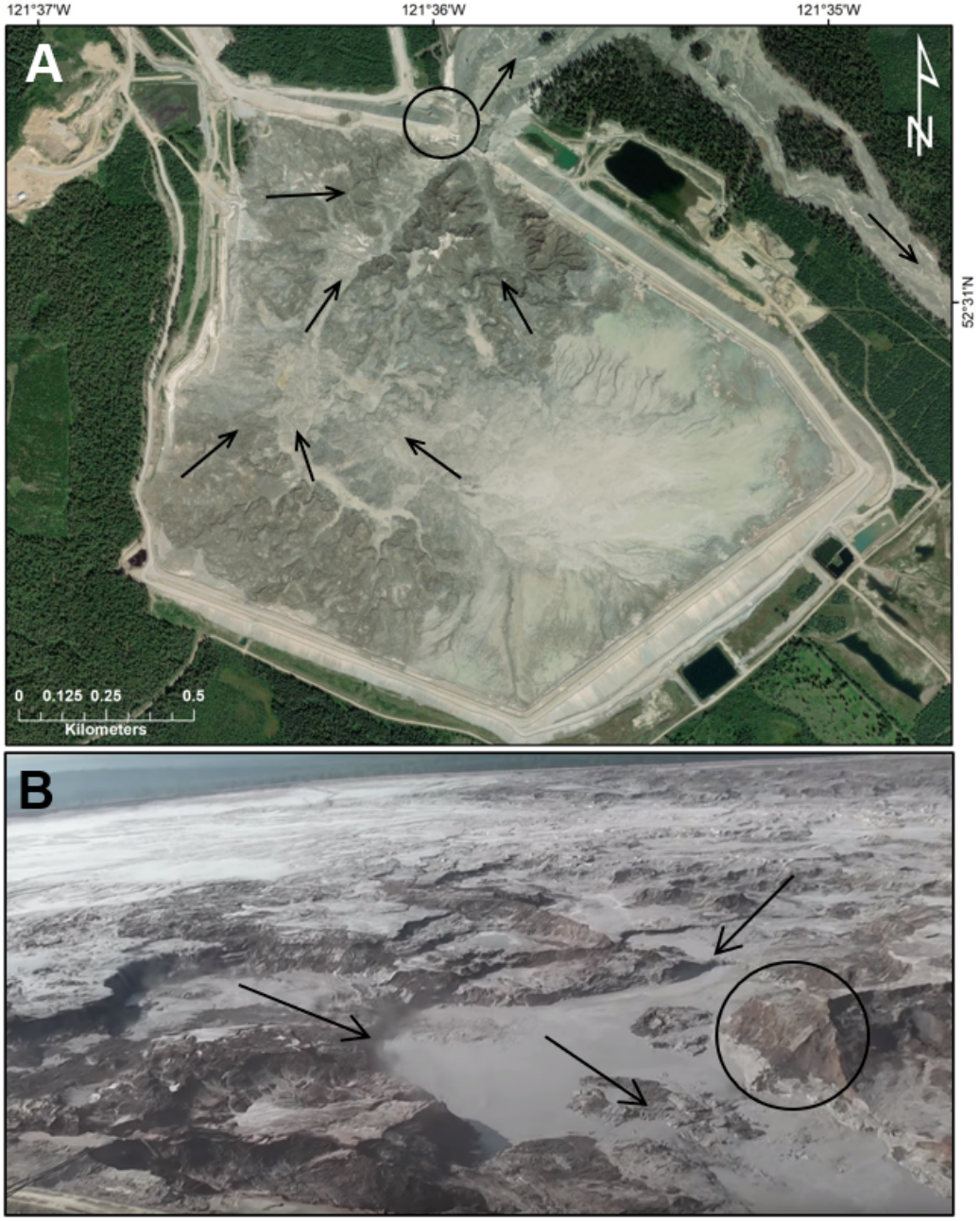

Figure 3-6: Post-erosional profile of the breached Mt. Polley facility. Image [A] is a satellite view of the failed impoundment and image [B] is an oblique aerial photograph viewing west towards the back of the facility. The left abutment of the failed dam is circled and the flow direction is indicated by black arrows. World imagery basemap [A] reprinted from ArcGIS Online maps under a CC BY license, with permission from Esri, original Copyright © 2020 Esri (Basemaps supported by Esri, DigitalGlobe, GeoEye, i-cubed, USDA FSA, USGS, AEX, Getmapping, Aerogrid, IGN, IGP, swisstopo, and the GIS User Community). Helicopter image [B] courtesy of the Cariboo Regional District of British Columbia, Canada.

3.6.2.3 2A Cases (Flow of Liquefied Tailings without a Supernatant Pond)

The 2015 Fundão and 2019 Feijão events (Morgenstern et al., 2016; Palu and Julien, 2019; Robertson et al., 2020; de Lima et al., 2020) are archetypes of catastrophic 2A cases. The outflow behaviors of 2A cases depend on the proportional volume of liquefiable tailings, which was atypically high in both the Fundão and Feijão cases despite the absence (or negligible volume) of impounded free water. Both sets of tailings contained high iron contents, corresponding to strong inter-particle bonding between oxidized grains which in turn elevated stiffness and specific gravity. Global instability of both facilities subsequently occurred as a sudden brittle response to the high peak strength being exceeded (Morgenstern et al., 2016; Robertson et al., 2019). The Fundão Dam collapse released 32 Mm^3 of liquefied tailings (61% of the impoundment), whereas 75% (9.7 M m^3) of the Feijão tailings underwent flow failure. Palu and Julien (2019) report an initial volumetric solids concentration of 57.5% in the Fundão outflow, while the peak discharge remains unknown. The Feijão outflow occurred as a singular surge with a peak height of 30 m that was completed in less than 5 mins (Robertson et al., 2019). Despite the decommissioned state of the Feijão facility, the contractive tailings were very wet owing to excess groundwater seepage and intense antecedent rainfall, leading to an interstitial water concentration of ~50% (Robertson et al., 2019). Figure 3-7 shows a post-failure view of the Feijão impoundment.

Figure 3-7: Oblique aerial photograph showing the failed Feijao impoundment on December 13, 2019 (almost 11 months after the disaster). Image purchased from Getty Images. The source area is characterized by a horseshoe-shaped steep scarp surrounding sub-horizontal terraces with drainage channels – a typical post-failure surface for catastrophic liquefaction (1A and 2A) cases.

3.6.2.4 2B Cases (Flow of Non-Liquefied Tailings without a Supernatant Pond)

2B events are relatively under-reported in available literature. In 2B cases, the tailings are typically characterized by higher solids concentrations with a lower susceptibility to flow liquefaction. Failures of these tailings may be mobilized as slope failures or slumps with relatively limited downstream impact (Small et al., 2017; Martin et al., 2019; CDA, 2020).

3.6.3 Runout behavior

Large volumes of failed tailings are capable of exhibiting highly mobile, destructive behavior in the downstream environment. In addition to the properties of the flowing mass (i.e. contractive tendency and volumetric proportions of free water, pore water and solids), the downstream topography also exerts a strong control over the runout behavior. A case review is presented below that provides insight into the observed effects of topographic confinement, slope and substrate on the velocity, geomorphic impact and mobility of tailings flows. We note that our assessment focuses on the primary impact zone ("Zone 1" extent) of tailings flows (Figure 3-8) in order to evaluate the runout behavior over the downstream terrain, and above mean bankfull elevations if the flow extends into downstream watercourses. For further details, refer to Ghahramani et al. (2020) and the supplementary materials to the present manuscript.

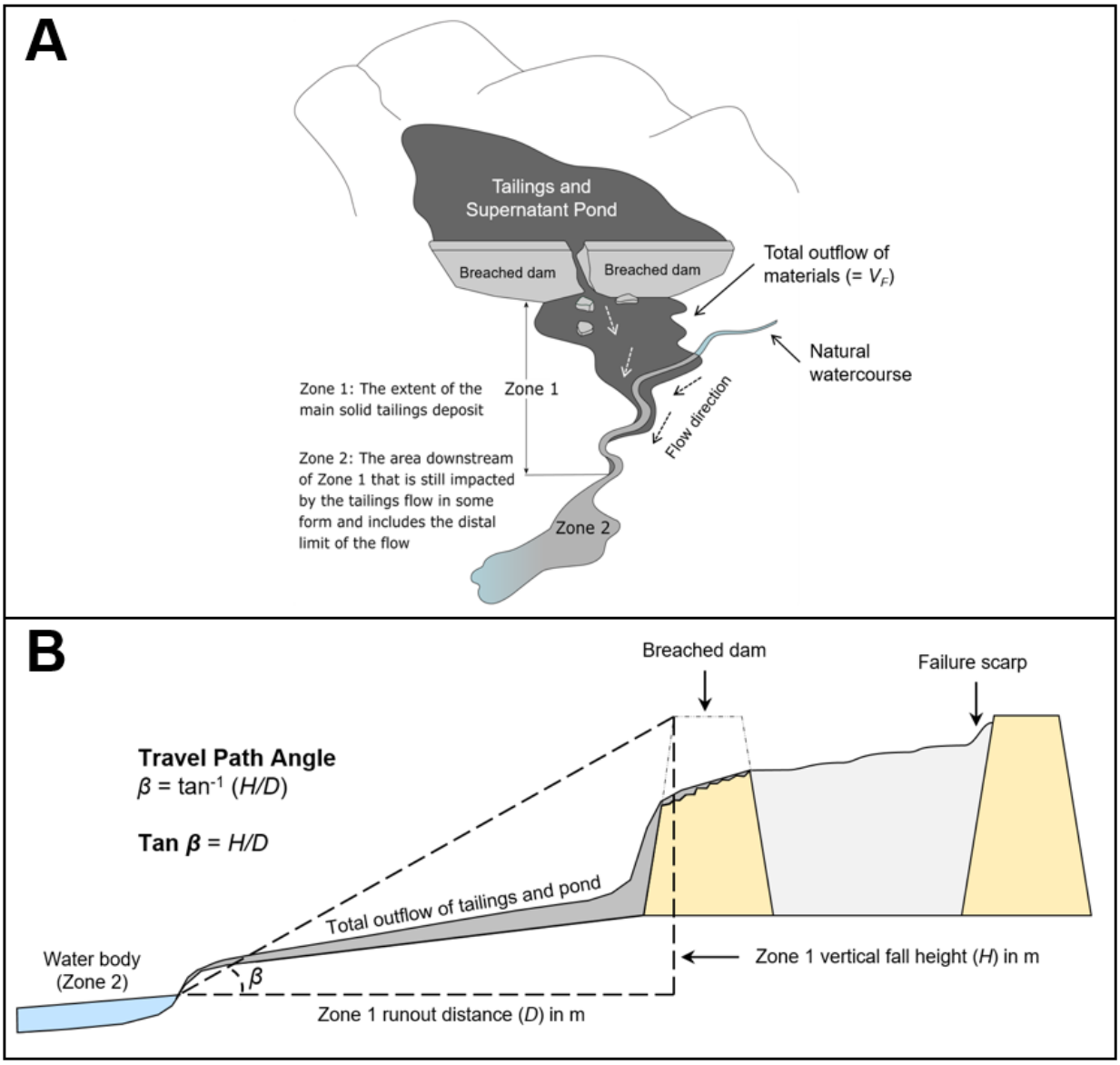

Figure 3-8: Idealized [A] oblique-view and [B] cross-section schematics of the runout zonation method adopted to independently measure the primary impact zone (Zone 1) runout distances, inundation areas and travel path angles of tailings flows with the support of remote sensing-GIS techniques (refer to Ghahramani et al., 2020 and the supplementary materials to the present manuscript for further details). In the given illustration, a supernatant pond is assumed to be present in the impoundment and incorporated within the outflow, although this is not always the case.

3.6.3.1 Channelized Runouts

Extremely rapid flow velocities (> 5 m/s; Cruden and Varnes, 1996; Hungr et al., 2001, 2014) and high kinetic energy have typically been exhibited by voluminous tailings flows meandering along relatively narrow creek or stream channels and/or relatively steep bed slopes. The seismically triggered Barahona flow in 1928 is among the first documented cases. The flow achieved a maximum superelevation of almost 60 m along laterally incised bends (indicating peak velocity of ~30 m/s) when flowing down the narrow Barahona creek channel with a bed slope of 4º–7º (Troncoso et al., 1993, 2017). The surge resulted in the deaths of 54 people and the destruction of a bridge, railway station and campsite (Troncoso et al., 1993, 2017). Nine years later, at least 300 people were killed during the Los Cedros disaster (Macias et al., 2015). The flow sustained velocities of over 20 m/s (peak of ~25 m/s) and depths of over 5 m (peak of ~20 m) for ~2 km (Macias et al., 2015). The extremely rapid field behavior was attributed to the location of the tailings facility that rested on a natural incline of ~22º in mountainous terrain, resulting in high gravitational driving stresses. Moreover, the mass flow travelled along the Tlalpujahua stream channel that had been further wettened by the triggering torrential rainstorm. Evidence of high mechanical energy included significant volumes of entrained debris including trees, roof tiles and detached building blocks (Macias et al., 2015).

The catastrophic field behavior of the Stava flow (268 deaths) is well-documented (Takahashi, 1991, 2014; Chandler and Tosatti, 1995; Colombo and Colleselli, 2003; Luino and De Graff, 2012). The initial flowslide travelled 600 m and crashed into the left bank of the Stava valley at a velocity of 8.3 m/s, destroying 31 buildings in the process (Chandler and Tosatti, 1995). The flowslide then accelerated along Stava Creek (bed slope of 2º–12º), sustaining a mean velocity of 25 m/s (peak of 31 m/s) until Tesero village 3.3 km downstream of the breach (Chandler and Tosatti, 1995; Takahashi, 1991, 2014). The flow depth progressively decreased from an initial high of 20 m to 5 m near Tesero (Takahashi, 1991, 2014). Almost 50,000 m^3 of debris – including soils, trees, and buildings – were entrained from both sides and the bottom of the Stava Creek valley which indicates an entrainment ratio of 0.28 (Hungr and Evans, 2004; Luino and De Graff, 2012). The flow entered the Rio Avisio 4.2 km downstream and subsequently dammed the river and formed a 500 m long lake upstream (Luino and De Graff, 2012).

A high erosional capacity was also demonstrated by the Mt. Polley flow along Hazeltine Creek valley (Cuervo et al., 2017). Although the tailings were not liquefied, a high volumetric proportion of free and

interstitial water (70% or 17 M m^3) enabled a 9 km long runout over a travel path angle of 1.3º. The mass would have travelled much further had it not been terminally confined by Quesnel Lake. Multiple knickpoint features up to 10 m in height were observed along the valley (Cuervo et al., 2017). Up to 0.6 M m^3 of creek sediments and vegetation were eroded and transported during the flow – almost equal to the entire outflow volume following the Merriespruit Dam breach. Steeper channel gradients along the creek corresponded to higher velocities and kinetic energy, which in turn exacerbated erosion (scouring depths of up to 7 m) while limiting deposition (less than 0.1 Mm^3) (Cuervo et al., 2017).

Catastrophic tailings flows may be capable of overwhelming creek channel banks to render widespread floodplain damage. An example is the seismically triggered El Cobre flow that killed over 200 people in 1965 (Figure 3-9; Torres and Brito 1966; Dobry and Alvarez, 1967). A peak velocity of ~13 m/s was estimated from arrival times and run-up heights (Torres and Brito, 1966). The tailings travelled along the El Cobre creek channel but carried sufficient energy and discharge to overwhelm both banks along the 11 km long runout, achieving a maximum width of 1 km. Similarly, the failure of the Plakalnitsa Dam (caused by intense groundwater seepage and rainfall) just a year later triggered a 6 km long runout along a creek channel through the village of Sgorigrad (Bulgaria). The tailings overwhelmed both banks and caused the destruction of 156 homes and the deaths of 107–488 people (Lucchi, 2009; Mossa and James, 2013).

The recent Feijão flow resulted in a death toll of 272 people (de Lima et al., 2020). de Lima et al. (2020) sectioned the field deposit into five energy zones to characterize the runout behavior (Figure 3-10). The flow initially attained a peak velocity of 28 m/s and height of 30 m along a relatively steep path (~10º), where the movement behaved similarly to a debris avalanche with extensive basal and lateral erosion and extreme destructive potential. Publicly available videos showed extremely rapid movement of the liquefied tailings as a coherent, homogenous mass (de Lima et al., 2020). The transition to a high-energy debris flow occurred ~700 m downstream where the ground slope angle declined to ~2º. In this zone, the flow continued to travel extremely rapidly (mean velocity of ~18 m/s) while entraining vast swaths of natural soils and vegetation (de Lima et al., 2020). Ground photographs highlight the very wet and mobile nature of the movement with a peak width of ~900 m at the ~2 km mark (Figure 3-10 inset photograph). The flow energy steadily declined as a consequence of the high volume of entrained material, transitioning to a high-energy mudflow about 4.8 km downstream where

velocities were calculated to be ≤ 7 m/s (de Lima et al., 2020). The mass entered the Paraopeba River at the 9 km mark as a low-energy mudflow (de Lima et al., 2020).

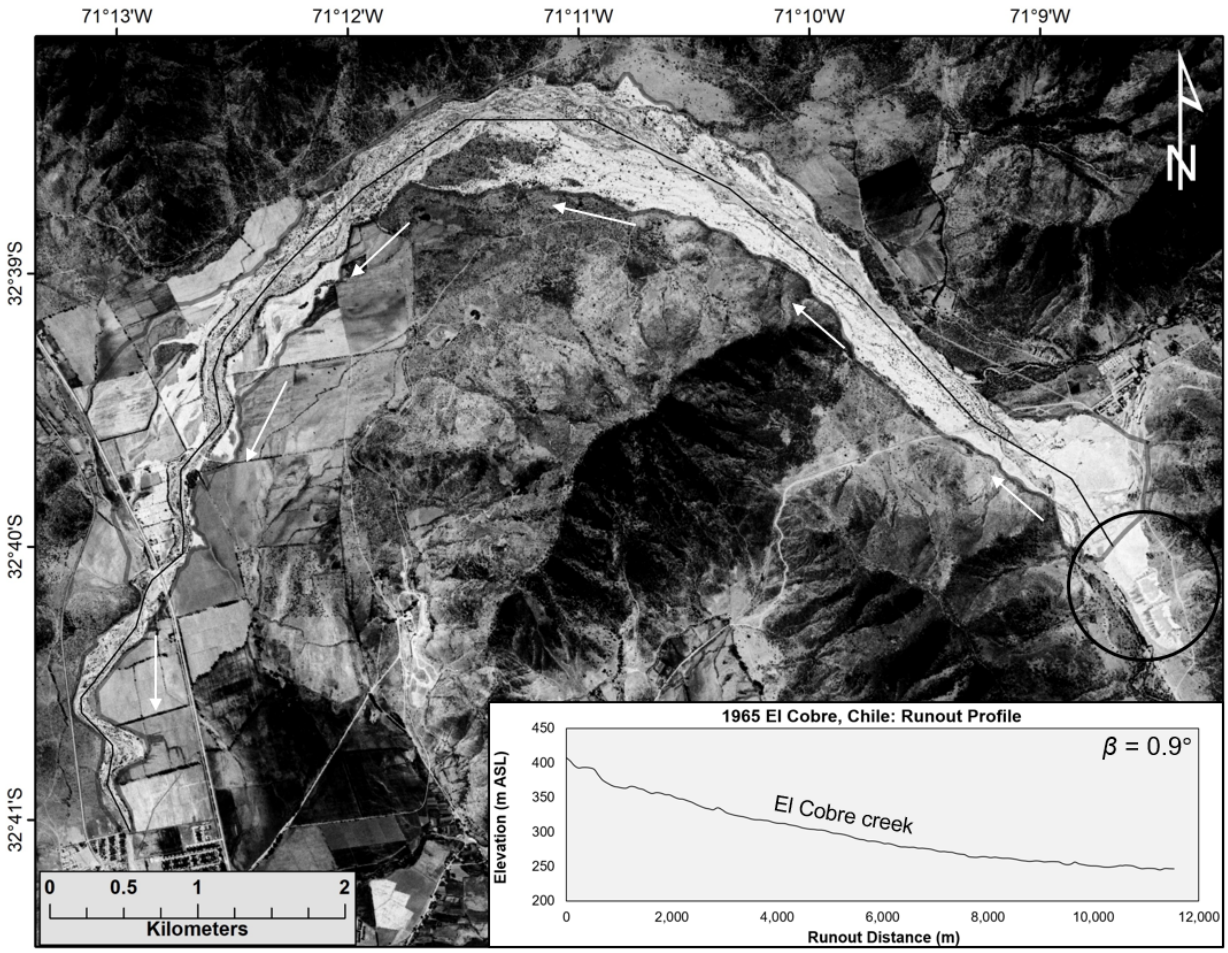

Figure 3-9: Corona (1.8 m) satellite image (captured November 1, 1968) of the March 28, 1965 seismically triggered El Cobre tailings flow (Zone 1 outlined in grey) in Chile. β represents the travel path angle of the runout path. The El Cobre Dams are circled and the flow direction is indicated by white arrows. Inset shows the SRTM runout profile with a vertical exaggeration of 13x. Raw image courtesy of USGS.

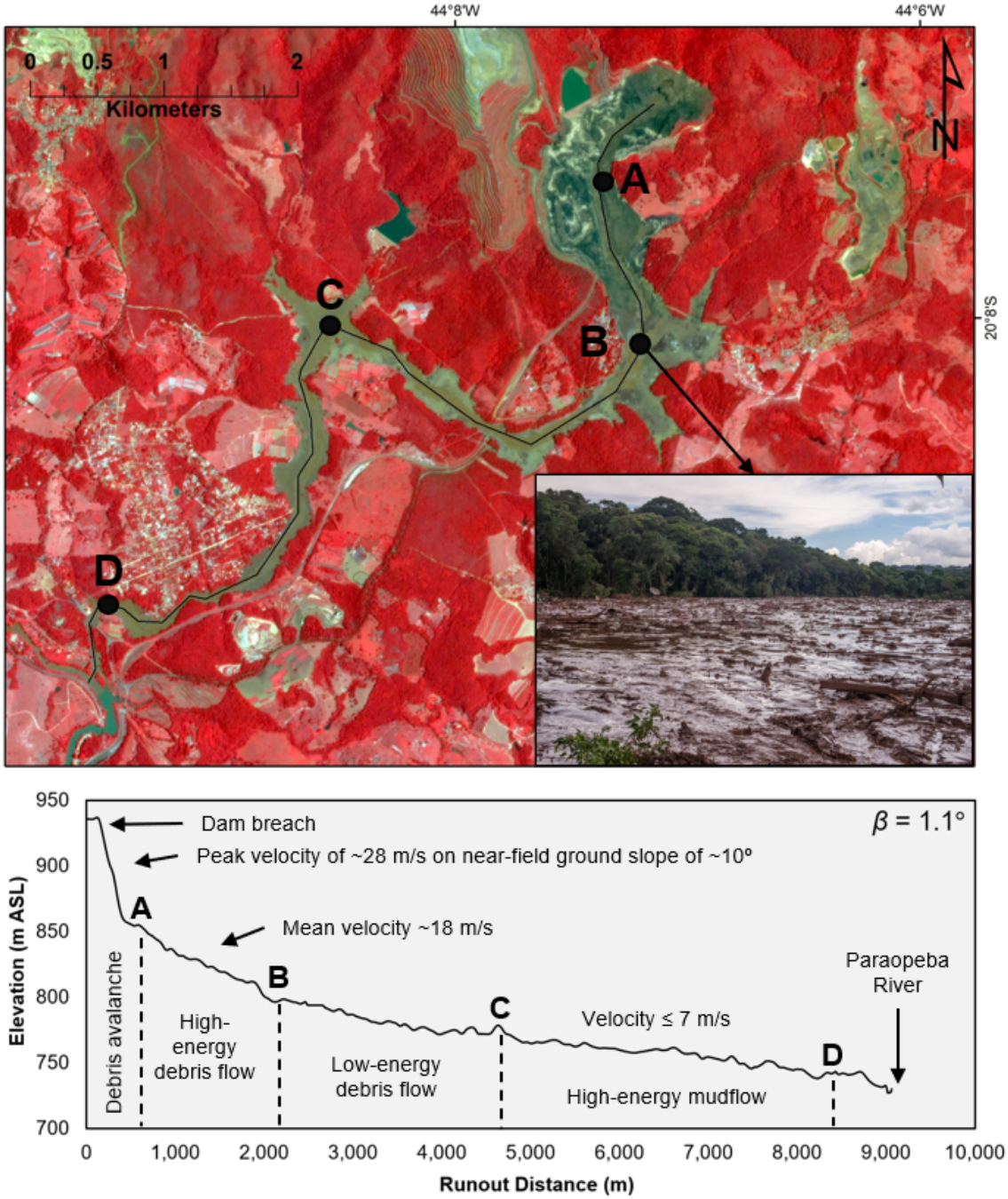

Figure 3-10: Infrared PlanetScope (3 m) satellite image (captured on January 29, 2019) of the January 25, 2019 Feijão flow in Brumadinho, Brazil. A 12m DEM runout profile is shown with a vertical exaggeration of 17x. The energy zones are based on field investigations performed by de Lima et al. (2020). Inset photograph highlights the very wet and mobile nature of the flow. Raw images and 10-m DEM courtesy of Planet Labs, Getty Images and Airbus.

Some of the lowest travel path angles (i.e. highest flow mobility), with values of < 1º, have corresponded to tailings flows along major rivers, with examples including the 1998 Aznalcóllar and 2015 Fundão runouts. The Aznalcóllar outflow initially developed into a 170 m long and 350 m wide alluvial fan that partially blocked the Agrio River (Gallart et al., 1999). The subsequent channelized flow elevated the peak river discharge to 800 m^3/s, equivalent to a 100-year local flood. The tailings travelled at a relatively low mean velocity of 1.9 m/s along steeper reaches and 0.7 m/s along gentler slopes within the active channel, reaching peak heights of 2 m above the banks (Gallart et al., 1999; McDermott and Sibley, 2000). Deposition was the primary mechanism as evidenced by the accumulation of crevasse splays and lack of erosional features (Gallart et al., 1999). Floodplain inundation was observed as far as 29 km downstream, indicating a low travel path angle of 0.1º along Zone 1 (Figure 3-11).

Figure 3-12 shows the near-field runout profile of the Fundão flow. In the early phase, the tailings flowed over and around the Santarem Dam. The 2-min travel time to the damaged conveyor 1.3 km downstream indicates a mean velocity of 11 m/s along this path (Morgenstern et al., 2016). The tailings swept through the town of Bento Rodrigues 6 km downstream before entering a major tributary to the Rio Doce at the 8 km mark. The volumetric solids concentration steadily reduced from an initial peak of 57.5% to 48.5% near Bento Rodrigues to 33% near the Candonga hydropower reservoir 118 km downstream (Palu and Julien, 2019). The travel time to the reservoir was 17 hours, indicating that the mean velocity had reduced to ~2 m/s during its runout along the major tributary. Almost 90% of the Fundão tailings were deposited along the 118 km long stretch of the floodplain (Palu and Julien, 2019). The Zone 1 runout distance is calculated to be 99 km (the longest Zone 1 runout in our database) with a low travel path angle of 0.35º.

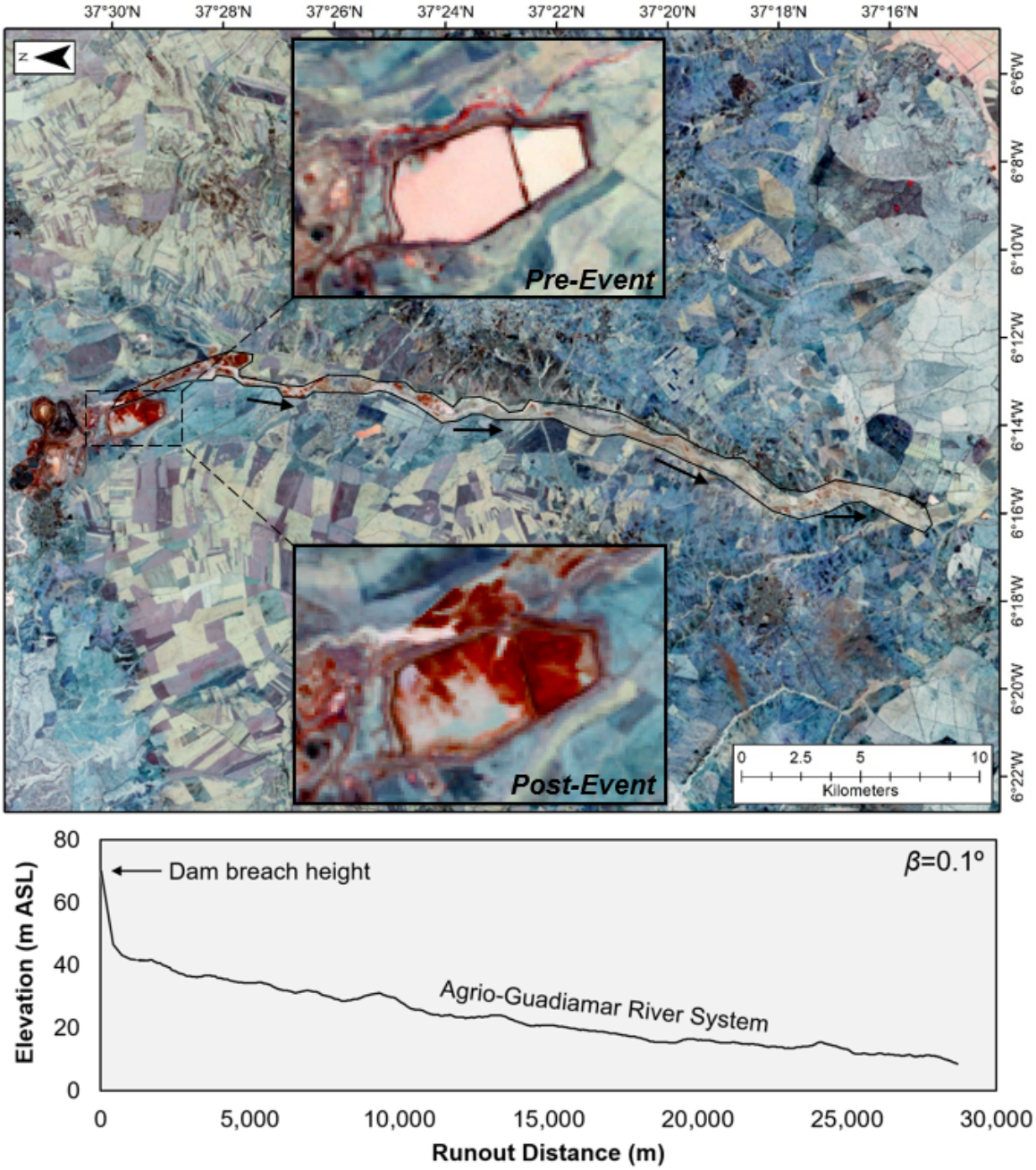

Figure 3-11: False-color Landsat 5 TM (30 m) satellite image (captured on June 13, 1998) of the April 25, 1998 Aznalcóllar tailings flow in Spain. Inset figures zoom into the pre- and post-failure dam. SRTM runout profile is shown with a vertical exaggeration of 100x. The red color indicates high wetness, which dominates the post-liquefaction, eroded surface of the impoundment. Raw image courtesy of USGS.

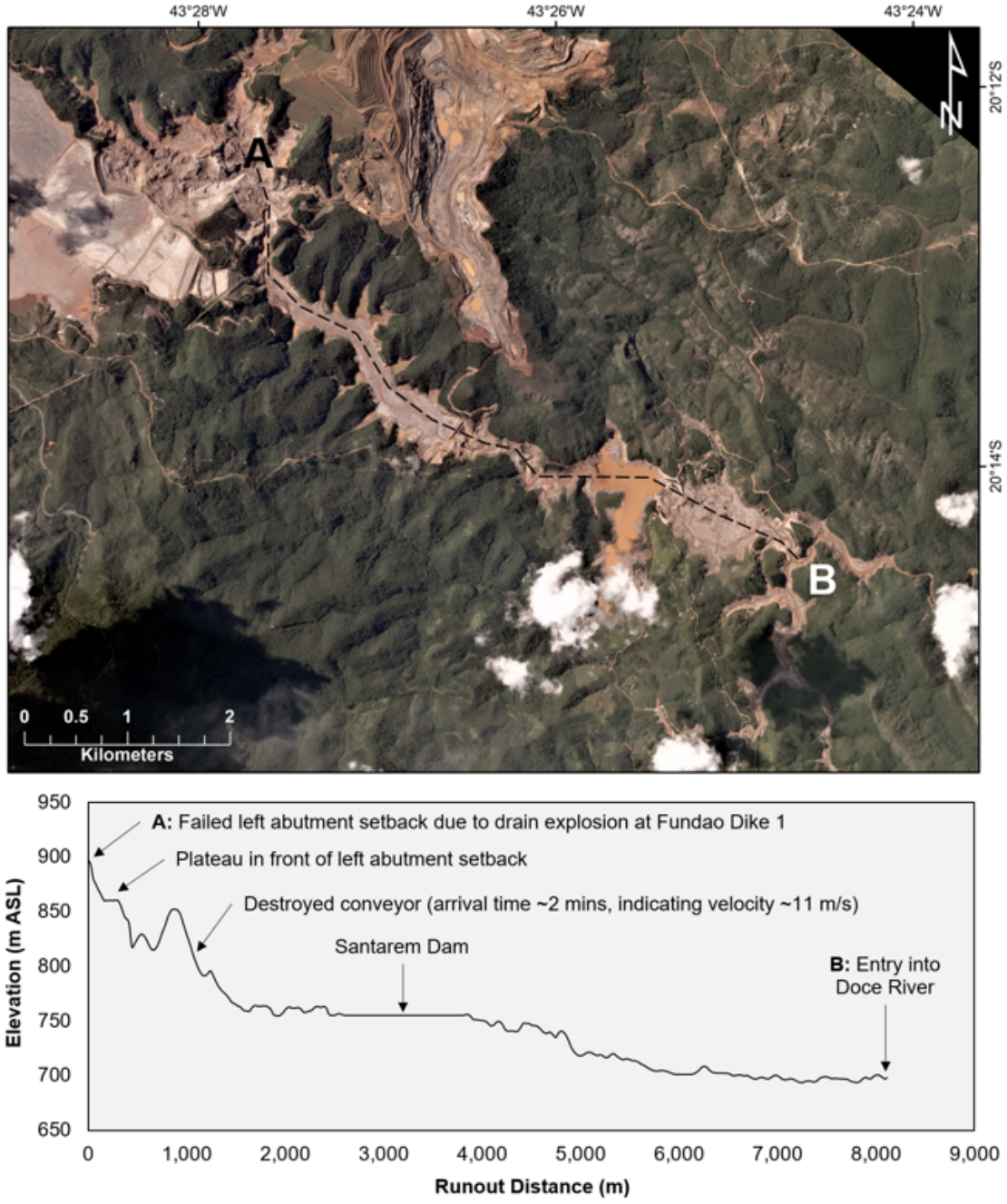

Figure 3-12: RapidEye (5 m) satellite image (courtesy of Planet Labs), captured on March 30, 2016, of the Fundão Dam failure in Brazil on November 5, 2015. Transect A-B delineates an SRTM runout profile of the flow (vertical exaggeration 10x) until its entry into a major tributary to the Rio Doce ~8 km downstream. The tailings continued to flow turbulently down-river, achieving a Zone 1 runout distance and travel path angle of 99 km and 0.35º respectively.

3.6.3.2 Unconfined Runouts

The inundation extents of tailings flows on unconfined terrain are mainly influenced by the rheological properties of the released materials (e.g. solids concentration) and travel path characteristics (e.g. substrate type and physical obstructions). The 2018 Cadia flowslide presents an unconfined 2A case where the inundation extent was rather limited due to its containment within the lower impoundment (Jefferies et al., 2019). Figure 3-13 illustrates two examples of slumping behavior following the Tonglvshan and Jharsuguda incidents in 2017. Although the trigger variables, geotechnical properties and CDA classes of these two cases are unknown, the relatively restricted movements may in part be ascribed to the absence of free water.

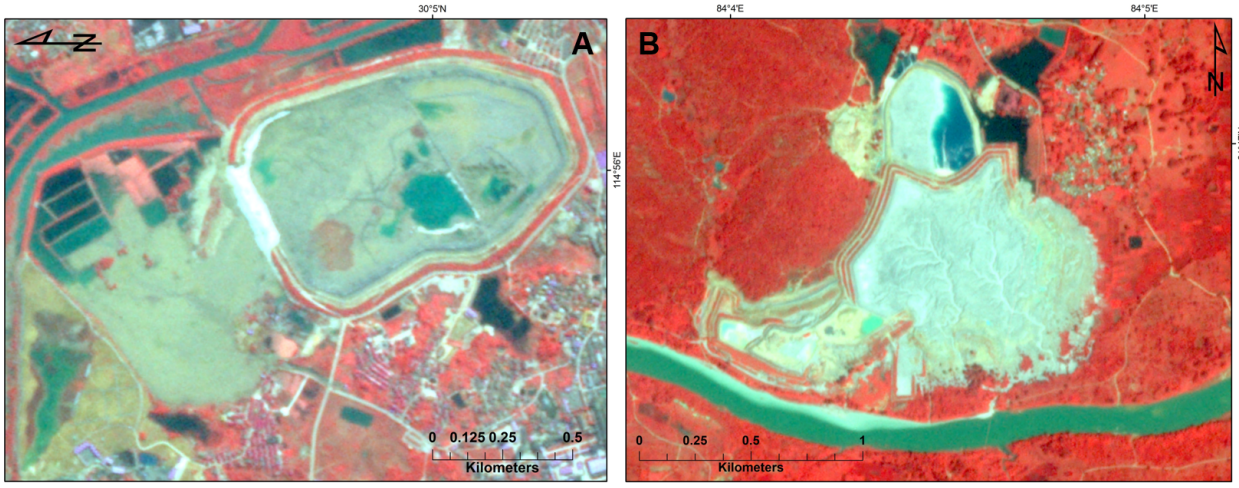

Figure 3-13: PlanetScope (3 m) satellite images (infrared band) of unconfined runouts following the [A] 12 March 2017 Tonglvshan (China) and [B] 28 August 2017 Jharsuguda (India) tailings dam breaches. Images were captured on 13 April 2017 and 25 September 2017 for Tonglvshan and Jharsuguda respectively. Free water was not incorporated in either event, which may have contributed to the observed slumping behavior. Note the enhanced width of the Tonglvshan slump due to terminal confinement by a raised agricultural field. Raw images courtesy of Planet Labs.

The Merriespruit and Ajka cases – both of which involved substantial volumes of free water – demonstrated highly mobile, expansive runout behavior on near-flat, unconfined, semi-urban terrain. The Merriespruit materials travelled through the eastern sector of the village where 17 people were killed (Blight, 1997; Wagener, 1997; Fourie and Papageorgiou, 2001; Fourie et al., 2001; Van Niekerk and Viljoen, 2005). The thickness of the tailings flow was ~2.5 m at 300 m downstream (Wagener,

1997). The surge achieved a maximum and mean flow width of 800 m and 400 m respectively on the relatively flat terrain (~1.4º). The inundation extent may have been enhanced by the rain-induced wet ground surface (Blight, 1997). It took ~5 mins for the tailings to reach a bird sanctuary pond 1.7 km away (Van Niekerk and Viljoen, 2005). The Zone 1 runout distance is estimated to be 2.2 km along a travel path angle of 1.3° (consistent with Blight, 1997).

The 2010 Ajka Dam breach released 1.64 M m^3 of materials consisting of roughly equal proportions of free water and statically liquefied bauxite tailings (Figure 3-14; Mecsi, 2013; Bánvölgyi, 2018). Mecsi (2013) reported very high moisture contents (> 150%) and void ratios (up to 6.0) in the freshly deposited tailings, while Bánvölgyi (2018) noted that the supernatant pond depth exceeded design criteria near the marginally stable embankment corner (Ghahramani et al., 2020). The Ajka flow achieved a maximum width of 1.5 km in the immediate vicinity of the breach before reducing to a mean width of 420 m. The mass flow swept through the towns of Kolontár and Devecser where 10 people were killed (Mecsi, 2013; Bánvölgyi, 2018). The liquefied mass proceeded to enter and travel along the Marcal River channel, achieving a Zone 1 runout distance of 17.3 km along a travel path angle of 0.3°.

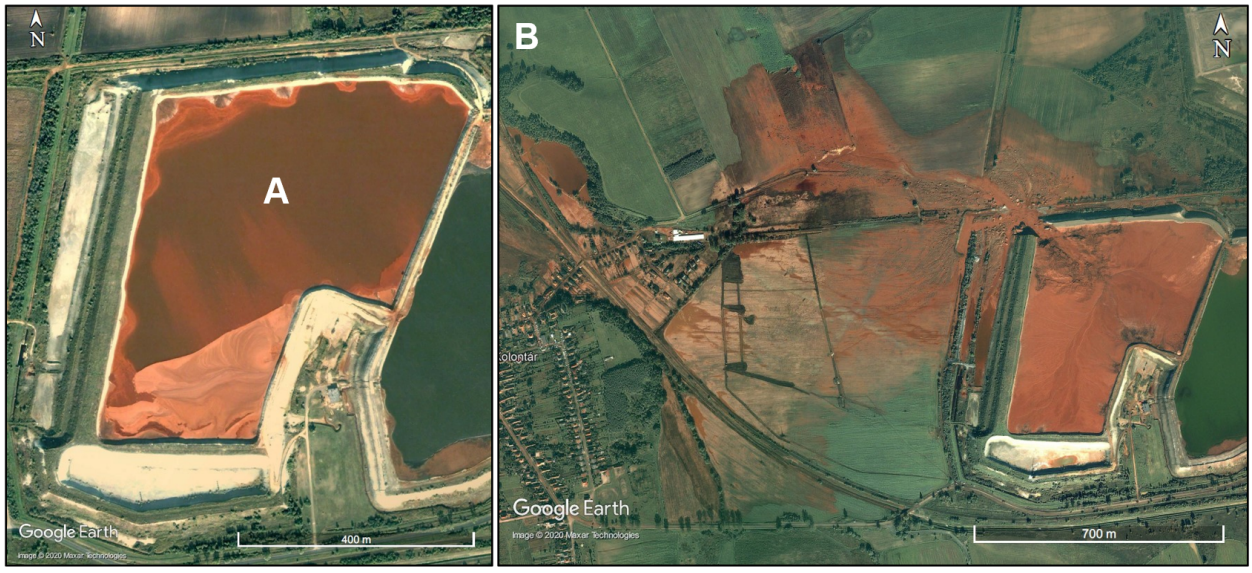

Figure 3-14: Worldview-2 satellite images (map data courtesy of Google and Maxar) showing the [A] pre-failure and [B] post-failure state of the Ajka bauxite tailings impoundment in Hungary that failed on October 4, 2010. Image was captured on October 10, 2010. Note the high volume of supernatant pond that promoted highly mobile and expansive runout behavior.

3.7 Implications for practice: empirical observations and limitations

Empirical assessment of tailings dam failures is a relatively recent and currently developing endeavour (Ghahramani et al., 2020). Previously published relationships are listed in Table 3-4. The primary focus has been on predicting the total outflow volume and subsequent inundation extent. While such efforts are appropriate for academic research, screening-level assessments and identifying case analogs, there is persisting concern regarding the perceived over-reliance on broad deterministic relationships to assist with practitioner-driven site assessments (Martin et al., 2015; Small et al., 2017; CDA, 2020; Gildeh et al., 2020). Breach-runout processes are complex and highly variable depending on local conditions that need to be accounted for when undertaking targeted prediction studies. As part of a critical review, we highlight below important site-specific complexities that require consideration yet remain unaddressed (or cannot be addressed) in broad empirical approaches, and thus, emphasize the importance of targeted investigations when undertaking site-specific breach-inundation studies.

Table 3-4: List of empirical relationships developed in preceding studies for tailings flows. A force-fit regression with a specified power-law slope of 2/3 for the A-V_F relationship developed by Ghahramani et al. (2020) is also shown. Notations are as follows: n is number of observations; R^2 is coefficient of determination of the best-fit line; V_F* denotes the previous definition of the total outflow volume (in M m^3) by Rico et al. (2008b), Small et al. (2017) and Larrauri and Lall (2018), whereas V_F is the revised definition of total outflow volume (in m^3) that Ghahramani et al. (2020) and the present work adopts (refer to text in Section 3.7.1 for the difference in definitions); D_{max} is the maximum runout distance in km including Zone 2 impact, which contrasts with the runout zonation method adopted by Ghahramani et al. (2020) and the present paper where only Zone 1 measurements of runout distance (D) in m are reported (see Figure 3-8 and supplementary material); h is dam height at breach location in m (not to be confused with upper-case H that denotes Zone 1 vertical fall height in m in this paper, as illustrated in Figure 3-8 and supplementary material); h·V_F* is the Dam Factor with units of M m^4; h_f is the modified Dam Factor (also in M m^4) that incorporates the total outflow ratio [V_F/V_T] as an additional multiplier; and A is Zone 1 inundation area in m^2.

Empirical Relationships	***n***	***R²***	**References**
$V_F^* = 0.354\ V_T^{\ 1.01}$	22	0.86	Rico et al. (2008b)
$D_{max} = 14.45\ (V_F^*)^{\ 0.76}$	23	0.56	Rico et al. (2008b)
$D_{max} = 0.05\ h^{\ 1.41}$	25	0.16	Rico et al. (2008b)
$D_{max} = 1.61\ (h \cdot V_F^*)^{\ 0.66}$	24	0.57	Rico et al. (2008b)
$V_F^* = 0.332\ V_T^{\ 0.95}$	29	0.89	Larrauri and Lall (2018)
$D_{max} = 3.04\ h_f^{\ 0.545}$	29	0.53	Larrauri and Lall (2018)
$A = 33\ V_F^{\ 0.73}$	33	0.58	Ghahramani et al. (2020)
$A = 80\ V_F^{\ 2/3}$	33	0.57	Ghahramani et al. (2020)

3.7.1 Total outflow volume

Rico et al. (2008b) performed the first set of linear regressions on a dataset of 29 cases from 1965 to 2000, which was later updated by Larrauri and Lall (2018) on an expanded dataset of 35 cases. To estimate the proportion of impounded materials released following a breach, both studies attempted to predict the total outflow volume of impounded materials (V_F) from the total pre-failure impounded volume of materials (V_T). It is worth noting that Rico et al. (2008b), Small et al. (2017) and Larrauri and Lall (2018) defined V_F as the "volume of tailings released", or in other words, the outflow volume of tailings only. In actuality, all three studies reported the "total outflow volume of impounded materials including tailings solids, interstitial water and free water". This revised definition of V_F is adopted by Ghahramani et al. (2020) and the present manuscript (see Table 3-4 caption).

The regressions of V_F from V_T produced well-constrained relationships with coefficients of determination (R^2) of 0.86 (Rico et al., 2008b) and 0.89 (Larrauri and Lall, 2018). However, this regression excludes the consideration of complex site-specific conditions. Given an impounded volume, the total outflow ratio (V_F / V_T) may vary widely from below 10% to over 90% – as indicated by the 95% prediction bounds in Figure 3-15 – depending on the interplay of several controlling variables, including (but not limited to) the following: (1) the volumetric proportion of impounded tailings solids, interstitial water and available free water; (2) the type of impoundment design and operational characteristics; (3) the geotechnical and rheological properties of impounded materials including the susceptibility to liquefaction (accounting for spatiotemporal variability); (4) the site-vicinity topography; (5) the foundation response to loading; (6) preconditioning and trigger variables; and (7) breach geometry and mechanisms. Incorporating all of these factors into simple regression models is impracticable due to limitations on the number of variables that may be considered, as well as the poor data quality and/or availability for many of the reported cases. Without the required contextual background, misconceptions about field processes may arise. Figure 3-15 attempts to address some of the uncertainty by classifying the data points according to their corresponding CDA classes, yet there are no distinct trends. It is also worth remarking that the task of assigning CDA classes to breach cases is prone to subjectivity (see Section 3.3). A more sophisticated empirical approach could be to compute (i) a dimensional analysis with the consideration of relevant attributes such as tailings properties and site setting, or (ii) a multiple linear regression of total outflow volume as a

function of impounded tailings (including pore water) volume and impounded free water volume. Both approaches, however, are hindered by data sparsity and/or poorly conditioned available data.

As such, performing simple statistical exercises on case histories in an attempt to provide a global physical explanation for complex post-breach behavior may produce inconsistent findings, potentially leading to flawed judgment in site assessments. Site-specific investigations through numerical models, laboratory tests and field observations are thereby needed to appropriately address uncertainties, particularly given that the predicted runout behavior is strongly influenced by the predicted total outflow volume.

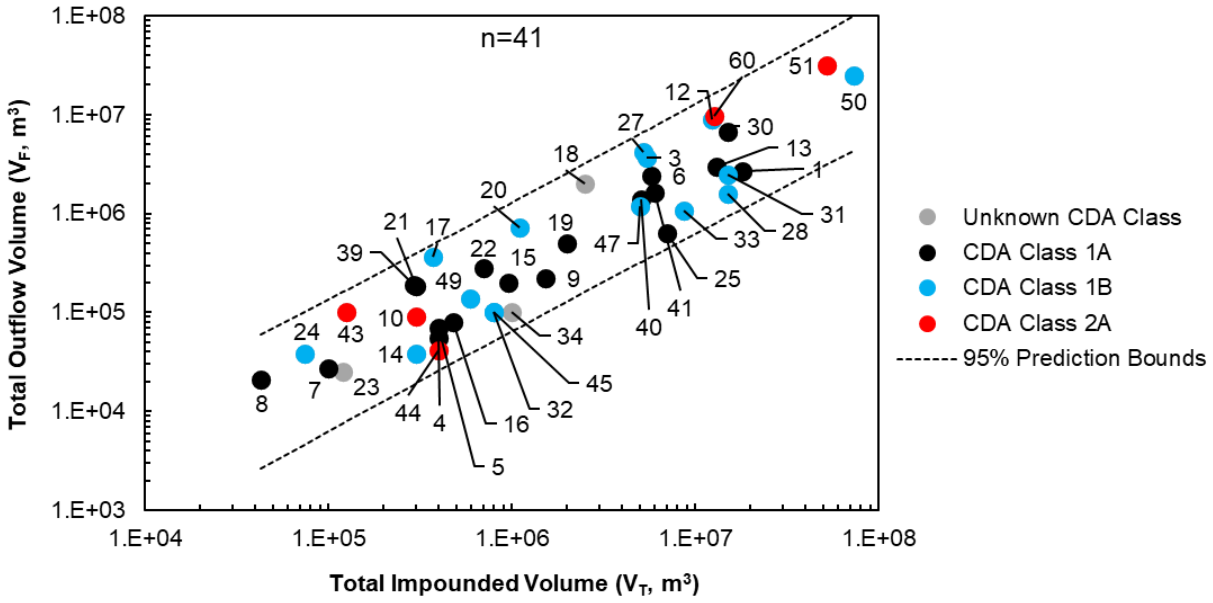

Figure 3-15: Observations of total outflow volume (V_F in m³) versus total impounded volume (V_T in m³) for 41 tailings dam breaches in log-log space. Cases are labelled according to corresponding IDs in Table 3-1. CDA class is incorporated as a qualitative indicator variable, but no distinct trends are observed. Twenty-four of the forty-one cases are upstream-raised dams whereas the rest are downstream-raised (3), centerline-raised (1), earth dam (1), open-pit storage (1) and unknown (11). Thirty-three of the forty-one cases consist of hard rock tailings whereas the rest are soft rock tailings (6) or unknown (2). The 95% prediction interval captures the observed data scatter, with the assumption that errors are normally distributed with zero mean. As established by the prediction bounds, the total outflow ratio may range from below 10% to over 90% depending on the interplay of several conditioning variables that are outlined in the text.

3.7.2 Downstream runout geometry

As shown in Table 3-4, Rico et al. (2008b) and Larrauri and Lall (2018) produced relationships of the reported maximum runout distance D_{max} (that includes Zone 2 impact) from V_F, dam height (H), a proxy energy expenditure index termed the Dam Factor $H \cdot V_F$ (adopted from Costa, 1985) and its modification H_f [$= H \cdot V_F \cdot (V_F/V_T)$]. Ghahramani et al. (2020) formulated a relationship between Zone 1 inundation area (A) and V_F with the support of a novel database of 33 tailings flows and a consistent runout zonation method (that the present paper adopts), where a force-fit regression with a slope of *2/3* was observed to be valid for tailings flows with a coefficient of 80.

However, the correlations have been poor to moderate (R^2 of 0.1 to 0.6) relative to analogous correlations developed for natural mass movements (R^2 of 0.7 to 0.9; Corominas, 1996; Legros, 2002; Crosta et al., 2003; Ghahramani et al., 2020; Petley, 2020c). The scatter is attributed to the site-specific variability of the following conditions: (1) the design of the tailings facility; (2) the hydro-geotechnical and rheological properties of the released materials, including the volumetric contributions of liquefied and/or non-liquefied tailings, interstitial water and free water; (3) the topography of the impoundment foundation (e.g. steeper foundations generate higher driving stresses); and (4) the nature of the travel path, including the slope angle, degree of confinement, substrate type (wet or dry, urban or vegetated) and presence of physical obstacles (lakes, trees and infrastructure).

As recommended by Small et al. (2017) and Ghahramani et al. (2020), we incorporate some of the qualitative uncertainties – namely CDA class, confinement, hard/soft rock tailings and travel path "wet" substrate – in Figure 3-16 and Figure 3-17. It is observed that tailings flows display, on average, higher mobility when travelling along major rivers and that channelization induces stronger correlations (compared to unconfined flows) particularly for inundation area (Figure 3-16A). Nonetheless, no distinct trends are visible for hard and soft rock tailings and CDA classes. The scarcity of quantitative data on the solids concentrations and entrainment capacities (with very few exceptions – e.g. 1985 Stava; Takahashi, 2014) further limits our ability to evaluate the mechanical energy and mobility of tailings flows. Numerous authors have thus advised prudent, professional judgment when interpreting the statistical scatter with the consideration of site conditions (Martin et al., 2015; 2019, Small et al., 2017; CDA, 2020; Ghahramani et al., 2020). This highlights the importance of site-specific

investigations (led by calibrated numerical models) in addressing uncertainties and enhancing breach-runout prediction efforts, ideally within a probabilistic framework.

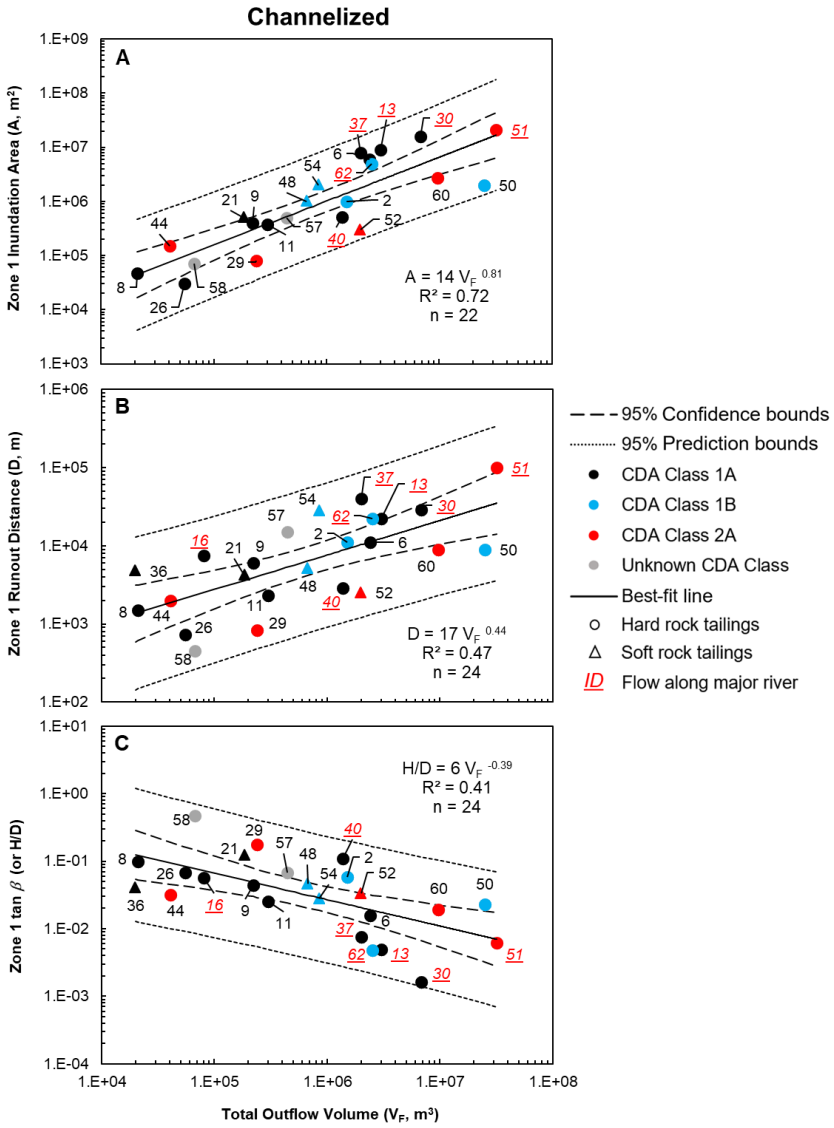

Figure 3-16: Observations of Zone 1 [A] inundation area (A in m^2), [B] runout distance (D in m) and [C] tan β (or H/D, dimensionless) versus total outflow volume (V_F in m^3) for channelized tailings flows in log-log space. Cases are labelled according to IDs in Table 3-1. The confidence interval captures the uncertainty about the best-fit line whereas the prediction interval captures the observation data scatter, with the assumption that errors are normally distributed with zero mean. Qualitative parameters such as CDA class, hard/soft rock tailings and travel path "wet" substrate (i.e. flow along major river channels) are incorporated as indicator variables. On average, higher mobility is observed for tailings flows along major rivers. A stronger correlation is obtained for inundation area. No distinct trends are observed for CDA classes and hard/soft rock tailings.

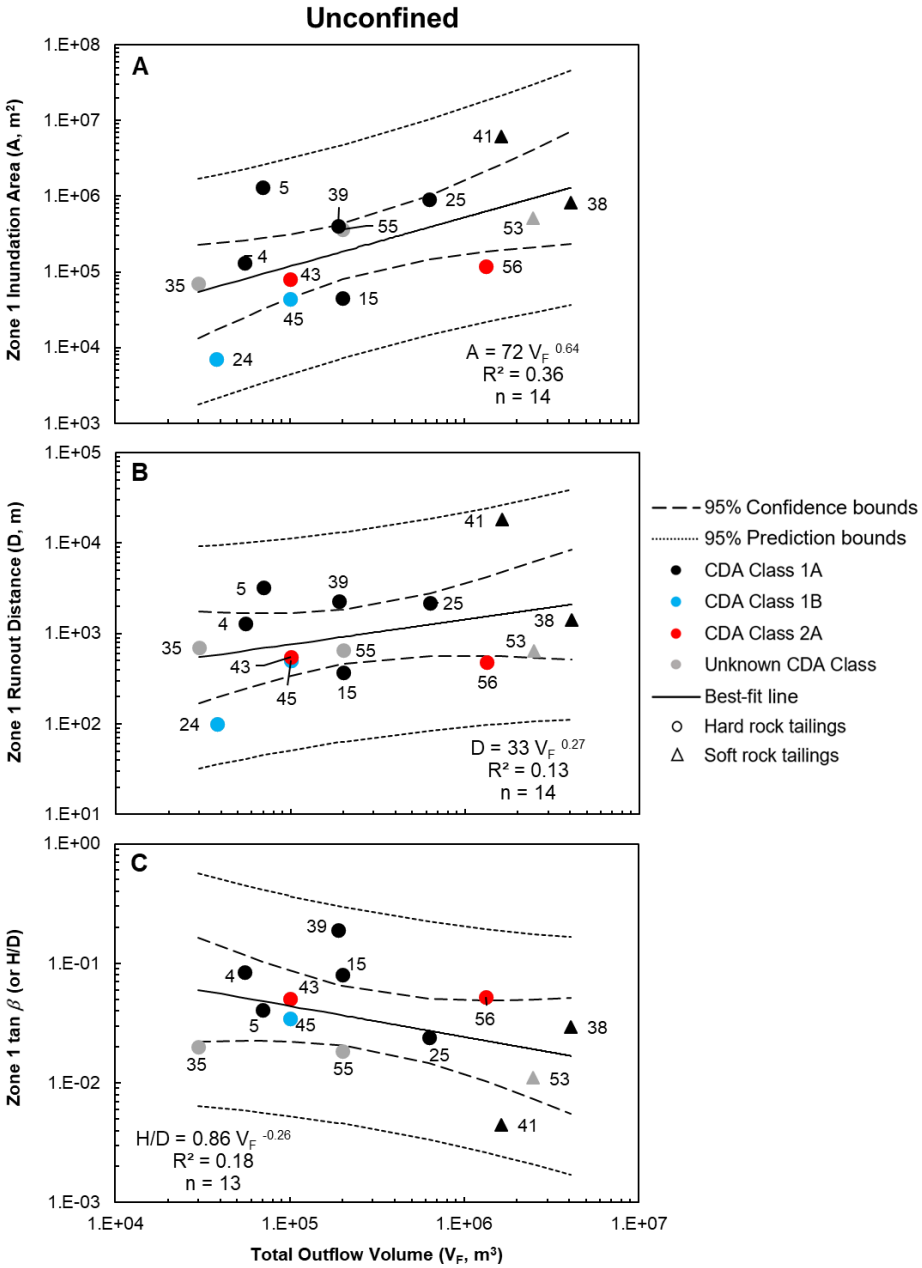

Figure 3-17: Observations of Zone 1 [A] inundation area (A in m^2), [B] runout distance (D in m) and [C] tan β (or H/D, dimensionless) versus total outflow volume (V_F in m^3) for unconfined tailings flows in log-log space. Cases are labelled according to IDs in Table 3-1. The confidence interval captures the uncertainty about the best-fit line whereas the prediction interval captures the observation data scatter, with the assumption that errors are normally distributed with zero mean. Other qualitative parameters such as the CDA class and hard/soft rock tailings are incorporated as indicator variables, but no distinct trends are observed and the correlations are relatively poor compared to channelized counterparts.

3.8 Summary and concluding remarks

This paper has provided a global review of the complex conditions and processes associated with the onset and behavior of catastrophic mass flows resulting from tailings dam failures. This work is timely considering the number of cases in recent decades that have resulted in loss of life, environmental and infrastructural damage and economic impact. With the support of a novel comprehensive database of tailings flows, this synthesis provides some insight into the influence of impoundment conditions, trigger variables, failure mechanisms and the downstream topography on tailings flow behavior. The dataset also sheds light on the limitations of data quality and availability in the public domain.

Tailings facilities are anthropogenic sources of energy capable of producing highly mobile mass movements when breaching initiates amid unfavorable conditions (i.e. when large volumes of tailings are contractive, saturated and potentially overlain by free water). Magnitude-frequency statistics indicate that tailings dam breaches that have produced mass flows with volumes of $\geq$ 1 M m^3 have occurred at a mean rate of 2–3 years since 1965. Meteorological hazards and impoundment drainage issues are identified as major causative variables. Tailings facilities are characteristically diverse depending on locality, and thus, the behavior of a tailings flow strongly depends on site-specific factors such as the type of failed facility, rheology of the flowing mass, nature of the runout path, triggering mechanism and failure mode. Nevertheless, the occurrence of liquefaction and/or the incorporation of free water are globally observed to be sufficient conditions to trigger extremely rapid, highly mobile runout behavior. Relative to unconfined terrains, channelized paths enhance the runout distance and inundation area – an effect that appears to be more pronounced along major rivers. Channelization and steeper bed slopes enhance flow velocities (peak of 25–30 m/s) which in turn exacerbate the downstream impact. Erosion is the dominant geomorphic mechanism during peak flow stages which indicates high mechanical energy. However, the entrainment capacities and solids concentrations (and conversely the interstitial and free water contents) of tailings flows remains quantitatively under-studied, suggesting a knowledge gap regarding the implications for tailings flow mobility and energy. Reductions in flow velocities are induced by gentler slopes and unconfined topography (peak of 6–16 m/s). Tailings flow velocities along major rivers steadily reduce to $\leq$ 5 m/s where the primary geomorphic mechanism transitions from erosion to deposition of coarser particles.

It is worth noting that the compiled database is not aimed to provide revised empirical correlations to be unconditionally applied in breach-runout analyses, but rather to provide assistance to practitioners performing site assessments identify case histories with characteristics and pre-failure conditions similar to the dam under consideration. Although general trends may be observed, empirical efforts to predict tailings outflow-runout are prone to considerable uncertainty due to (1) data scarcity or poorly conditioned available data for many of the selected cases and (2) the complexity and variability of the site and tailings conditions that are inherently challenging to incorporate into simple regression models. Broad statistical approaches are thus appropriate for generating scholarly insights, identifying case analogs and performing screening-level assessments, while site-specific studies should rely on targeted investigations through numerical models, laboratory tests and field observations to ascertain potential breach-inundation behavior, ideally within a probabilistic framework.

3.9 Data availability

The supplementary data for this study is stored here: https://doi.org/10.5683/SP2/NXMXTI, an open-source online data repository hosted at *Borealis* (Rana et al., 2021b).

3.10 Postscript

The purpose of this addendum is to acknowledge and briefly describe a relevant study by Piciullo et al. (2022) that was published after Rana et al. (2021a). Piciullo et al. (2022) built on the work of Rico et al. (2008b), Rourke and Luppnow (2015), and Larrauri and Lall (2018) by conducting linear regression exercises to predict released volume from stored volume, dam height, dam factor, dam raise method, failure mode, and "pond ratio" (i.e. ratio between the water pond area and the TSF area; Rourke and Luppnow, 2015) using a larger sample of 70 cases. However, the correlation between released and stored volume indicated a much lower R^2 coefficient (0.59) compared to previous studies (>0.80). This coefficient slightly increased to 0.64 when correlating released volume to dam factor, and to 0.75 when excluding 3 outlier cases. There was no correlation between released volume and pond ratio. Piciullo et al. (2022) concluded that there were evident issues in the quality of the reported data on stored volumes. These observations support the conclusion of Rana et al. (2021a) that statistical methods are not appropriate tools for predicting a complex parameter such as released volume due to its dependency on site-specific conditions and the poor accuracy or availability of data.

3.11 References

Given that this thesis is in part article-based, the content of Chapter 3 represents material published in the following citations:

Rana, N.M., Ghahramani, N., Evans, S.G., McDougall, S., Small, A., Take, W.A., 2021a. Catastrophic mass flows resulting from tailings impoundment failures. Eng. Geol. 292, 106262. https://doi.org/10.1016/j.enggeo.2021.106262.

Rana, N.M., Ghahramani, N., Evans, S.G., McDougall, S., Small, A., Take, W.A., 2021b. A comprehensive global database of tailings flows. Borealis v2. https://doi.org/10.5683/SP2/NXMXTI.

Rana, N.M. Ghahramani, N., Evans, S.G., McDougall, S., Take, W.A., 2021c. Analysis of downstream field behavior of long-runout tailings flows using remote sensing. In: Proceedings of the Mine Waste and Tailings Conference, Brisbane, Australia.

Chapter 4: Using Sentinel-1 InSAR to Monitor Tailings Dams and Predict Instability

Overview:

Tailings dams impound waste by-products from mining operations to ensure public safety, but recent breach incidents have prompted public calls for more robust monitoring programs. Satellite InSAR technology has grown in popularity due to its ability to detect millimeter-scale displacements in infrastructure. However, the literature on InSAR applications on tailings dams is relatively limited compared to other engineering sectors. This gap has led to limited understanding of (i) whether InSAR can be as accurate or representative as on-the-ground instrumentation, (ii) whether failures of tailings dams can be predicted in advance, and (iii) what conditions and variables govern the quality of InSAR results and monitoring/prediction efforts.

This study explores the capabilities and limitations of satellite InSAR for monitoring the geotechnical stability of tailings dams. This is achieved by the analysis of open-source Sentinel-1 data to undertake a ground-truth assessment at a test site and a forensic analysis of 7 failure cases. The methodology involves the use of a commercial software with an automated Persistent Scatterer (PS) workflow (SARScape Analytics) for all cases, with comparison to an advanced proprietary software (SqueeSAR) with a dual PS and Distributed Scatterer (DS) algorithm for the ground-truth test site and one forensic case study. For all cases, it was only possible to retrieve one-dimensional LOS displacement data. Key conclusions of practical value are as follows:

1. Commercial InSAR software with Sentinel-1 data provides reasonable accuracy when monitoring consolidation and lower-scale deformation (< 50 mm/yr) but is ineffective in larger or accelerating deformation regimes (> 50 mm/yr), where advanced proprietary algorithms are more appropriate.

2. Environmental conditions strongly influence the quality of Sentinel-1 PS-InSAR. The best results are in dry, bare-earth or urban terrains, whereas vegetation, water, and snow/ice cover worsen PS-InSAR data quality.

3. Sentinel-1 InSAR is a useful hazard-screening tool in active mine sites with large TSFs or multiple TSFs, or in legacy mines with abandoned TSFs, as it may help guide where to undertake targeted investigations.

4. Due to the polar orbit of SAR satellites, sub-horizontal movements in the north-south direction are generally poorly captured.

5. Not all TSF failures show clear warning signs for weeks in advance, and the revisit interval of SAR satellites prevents detection of instantaneous failure modes. Therefore, long-term

monitoring programs should ideally be integrated with a combination of remote sensing methods and field instrumentation.

Future avenues of research could involve case-study comparisons between medium-resolution Sentinel-1 and high-resolution TerraSAR-X data and ground-truth comparisons between different proprietary algorithms.

4.1 Introduction

4.1.1 Purpose Statement

This article explores the capabilities and limitations of satellite-based interferometric synthetic aperture radar (InSAR) technology for monitoring the geotechnical stability of mine tailings dams. This is achieved by undertaking a ground-truth assessment and several forensic analyses of failure cases. By this, practitioners will gain a better understanding of how environmental conditions, failure modes, and software/satellite selection can influence the quality and/or value of InSAR results.

4.1.2 Background

Tailings storage facilities (TSFs) impound fine-grained, wet, geochemically hazardous mine waste behind constructed dams in perpetuity for societal and environmental protection (Vick, 1983; Blight, 2010). TSFs can store considerable volumes of flowable material that, if released accidentally, could produce far-reaching and long-lasting consequences, as evidenced by a number of tailings dam failure incidents in recent years (e.g. Morgenstern et al., 2015, 2016; Robertson et al., 2019; Rana et al., 2021). Such events highlight the importance of implementing proactive monitoring systems at TSF sites to ensure safe performance.

A key objective in tailings dam monitoring is to observe for potential signs of instability by analyzing spatiotemporal rates of displacement – a variable that is of special concern in scenarios involving creep deformation, static liquefaction, or foundation deformations. In industry practice, the displacement rate has conventionally been monitored by field observations, in-situ instrumentation (e.g. monitoring prisms, inclinometers and extensometers), ground-based InSAR, ground-based or airborne Light Detection and Ranging (LiDAR), and/or photogrammetry. Another available technique is satellite InSAR, which has been utilized as a complementary, and potentially cost-effective, monitoring tool in

mining practice (Hu et al., 2017; Raspini et al., 2022). The central focus of this article is on the utility of satellite InSAR for monitoring tailings dams and predicting instability.

Satellite InSAR is a remote sensing technology for the geospatial detection and monitoring of features or phenomena on the Earth's surface. SAR satellites are polar orbits in either the ascending (south-to-north) or descending (north-to-south) direction. The satellites emit microwave signals that can penetrate cloud cover in day or night, thus overcoming the limitations of optical satellites. By conducting "interferometric" analysis of SAR images (i.e. InSAR), one can measure millimeter-scale displacements in the line-of-sight (LOS) direction of the SAR satellite, or in two dimensions (east-west horizontal and up-down vertical) if two satellite orbit tracks in opposite directions are overlapping spatially and temporally (e.g. Hu et al., 2017). Figure 4-1 shows a conceptual schematic of how satellite InSAR is used to detect displacements in dams.

The C-band Sentinel-1 satellite, commenced in mid-2015 by the European Space Agency (ESA), has become a popular SAR sensor in displacement monitoring studies due to the open-source data release and the revisit times of 6 or 12 days in most of the world. The X-band TerraSAR-X (2008-present) satellite is an alternative, at-cost, high-resolution option with near-global coverage and a revisit time of ~11 days. Comparisons between the two satellites have shown that TerraSAR-X can offer a higher point-density of displacement measurements and a lower standard deviation in displacement results (Bischoff et al., 2017; Holden et al., 2020; Colombo, 2021; Gama et al., 2022).

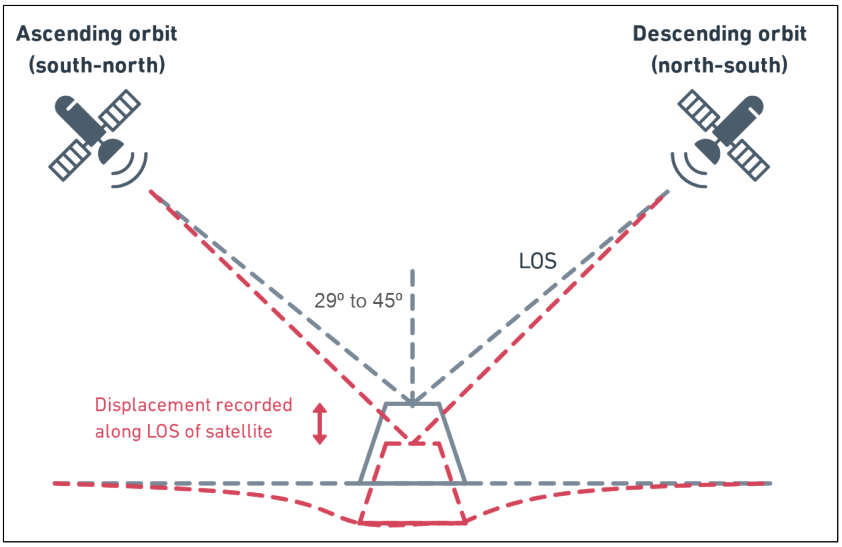

Figure 4-1: Conceptual schematic showing the application of satellite InSAR for monitoring the 1-D displacement of a dam, in this hypothetical case via foundation settlement, along the line-of-sight (LOS) of the satellite. The range of LOS incidence angles shown here (29º to 45º) is typical of the Sentinel-1 satellite. In this illustration, north is into the page. Note that SAR satellites are conventionally right-looking.

Among the numerous InSAR deformation analysis techniques (Aswathi et al., 2022), Persistent Scatterer (PS) has gained attention due to its powerful ability to produce highly precise, long-term displacement time-series mainly for human-built structures such as bridges, roads, buildings, and dams (Ferretti et al., 2001; Crosetto et al., 2016). By analyzing a stack of SAR images with reasonable temporal and geometric baseline consistency, one can evaluate the LOS displacement time-series across a PS point cloud, thus providing critical insights into deforming infrastructure (Crosetto et al., 2016). Small Baseline Subset (SBAS) InSAR is an alternative time-series approach that was designed to improve the spatial distribution and density of "Distributed Scatterer" (DS) observation points in vegetated study areas, albeit at reduced spatial resolution (Berardino et al., 2002; Casu et al., 2006; Lanari et al., 2007; Yunjun et al., 2019; Grebby et al., 2021; Aswathi et al., 2022). Comparisons between the PS and SBAS InSAR techniques have been presented in numerous application studies (Shanker et al., 2011; Yan et al., 2012; Grzovic and Ghulam, 2015; Gama et al., 2020; Cigna et al., 2021). Examples of software used to process PS or SBAS data for displacement monitoring include SqueeSAR (Ferretti et al., 2011; Carla et al., 2019a; Bischoff et al., 2020), GAMMA (Werner et al., 2000; Wang et al., 2020), SARPROZ (Perissin et al., 2011; Bakon et al., 2014), SARScape (Gama et

al., 2020), APSIS (Sowter et al., 2016; Grebby et al., 2021), the open-source HyP3/MintPy (Yunjun et al., 2019), and the open-source SNAP/SNAPHU (Chen and Zebker, 2002).

A key technical drawback of InSAR is the complicated and lengthy workflow that necessitates the use of computers and software with high data processing capacity, which can often render such studies unfeasible for users with limited resources. This obstacle, along with the conceptual complexity of advanced InSAR, has contributed to a limited archive of case studies on tailings dams. Hu et al. (2017) monitored displacements at the Kennecott TSF in the USA by integrating ENVISAT, ALOS Palsar-1, and Sentinel-1A data. Mazzanti et al. (2021) used over 400 Sentinel-1 images in the ascending and descending orbit direction to study displacements at the Zelazny Most TSF in Poland.

Three recent TSF breach cases have also been forensically analyzed using InSAR: 2018 Cadia, Australia (Carla et al., 2019a; Jefferies et al., 2019; Thomas et al., 2019; Hudson et al., 2021; Bayaraa et al., 2022), 2019 Feijao, Brazil (Gama et al., 2020; Holden et al., 2020; Rotta et al., 2020; Grebby et al., 2021), and 2022 Jiaokou, China (Duan et al., 2023). The common conclusion in these case studies was that satellite InSAR can be an effective monitoring tool for TSFs exhibiting slow, long-term deformations, but awareness of limitations is needed as it relates to the oblique geometry of 1-D LOS displacement measurements, the difficulty in predicting instantaneous failure mechanisms (as satellite InSAR is not a real-time technique), phase unwrapping errors, and loss of coherence.

4.1.3 Goal and Scope

To build on previous advancements, it is necessary for the case study inventory to be expanded in order to explore the capabilities and limitations of InSAR when monitoring tailings dams in diverse site conditions. With the support of Sentinel-1 data, this study helps fill this research gap in two ways. First, we undertake a ground-truth assessment at a test site where InSAR results are compared to in-situ monitoring prism data. Second, we examine the precursor displacements in 7 TSF failure cases (2017-2019) selected from published databases (Islam and Murakami, 2021; Rana et al., 2021, 2022) based on the following criteria: (i) their variability in reported failure mechanisms and site characteristics, and (ii) the spatial-temporal coverage of Sentinel-1 data over the sites.

For all cases, we use a commercial software that offers an automated workflow for PS analysis (SARScape Analytics). In addition, for the ground-truth test site and one forensic case study, the

proprietary algorithm SqueeSAR – integrated with an advanced PS-DS technique – is used for comparison. We note that, in all of the cases, it was possible to retrieve only the 1-D LOS displacements – i.e. spatially and temporally overlapping ascending and descending orbit tracks, which allow insights into east-west and vertical displacements, were not available. Furthermore, TerraSAR-X could not be incorporated into this study due to poor spatial-temporal coverage of the archived data over our case study sites as well as the lack of cost-effective accessibility to the data products.

The main goal of this article is to provide practical insights and considerations for geotechnical practitioners who may be considering InSAR as a long-term monitoring tool at their TSF sites. This goal is achieved by the following approach:

1. We analyze the accuracy of Sentinel-1 InSAR on a site-scale using multiple software.
2. We assess if unstable locations and accelerations in precursor displacements can be detected by the present approach. This allows us to:
    a. Identify high-deformation hotspots to match with the observed breach location and reported breach mechanism, which is important for hazard assessment;
    b. Check if the failure was preceded by accelerating movements, and if so, how many weeks in advance this trend was observed, which is important for risk management; or
    c. Identify errors in the results due to the inherent limitations of the selected software, the limitations of C-band Sentinel-1 data, or the LOS velocity threshold being exceeded.
3. We observe the influence of dam-satellite geometry, environmental conditions, and failure modes on the quality and interpretation of LOS displacement results.

This study is timely considering the public scrutiny on the mining industry in light of high-profile TSF failures and the ensuing calls by investors for accessible monitoring solutions for poorly instrumented or closed/abandoned TSFs in high-risk areas (Church of England, 2022). The significance of this work could also extend into other engineering sectors that would benefit from monitoring of critical infrastructure by remote sensing (e.g. reservoir dams, deforming roads or bridges, or oil and gas extraction fields). As such, the paper is written with the aim of being accessible to a broad target audience. **Error! Reference source not found.** lists the terminologies of key InSAR variables that are relevant to this study.

Table 4-1: Terminologies of key InSAR variables relevant to this study.

Term	Definition and description
Interferometric synthetic aperture radar (InSAR)	Interferometric analysis of SAR images (InSAR) is a remote sensing method used to map terrain elevations, flood/landslide inundation extents (change detection), and displacements. Radar sensors from spaceborne satellites record the amplitude and phase of the backscattered signals. The change in phase across multiple SAR images is measured in order to derive displacement measurements.
Persistent Scatterer (PS)	PS-InSAR uses dozens of SAR images over the same study area to generate a point-cloud of displacement measurements along the satellite's line-of-sight. PS-InSAR is mainly applicable when monitoring areas with slow (mm-scale), long-term deformation activity (e.g. subsidence, uplift, creep deformation, dam settlement). This level of precision is achievable because the PS algorithm estimates and removes certain phase components such as atmospheric contributions and topographic errors, whereby the error of the phase is in the order of the wavelength of the emitted signal (i.e. millimeters). PS-InSAR is mainly effective when monitoring urban infrastructure (dams, buildings, bridges, roads), and generally not natural terrains except bedrock outcrops due to the hard, reflective surfaces that allow for the preservation of pixels exhibiting a strong, constant reflection over a long period.
Distributed Scatterer (DS) and Small Baseline Subset (SBAS)	Unlike PS, DS points represent statistically homogenous groups of pixels corresponding to areas with relatively moderate coherence where phase values can be averaged, enhancing the signal-to-noise ratio (SNR). The SBAS-InSAR technique leverages DS points to show better spatial coverage of movements in natural terrain such as bare soil or short vegetation, where PS techniques can falter.
Coherence	A measure of the noise in the phase of the interferograms involved in the InSAR analysis. Coherence ranges from 0 to 1, whereby higher values indicate greater similarity in corresponding pixels across the SAR time-series imagery (i.e. a metric of higher quality results). The coherence value mainly depends on the ground characteristics (e.g. free water and vegetation reduce coherence, whereas hard, stable surfaces increase coherence), perpendicular and/or temporal baseline decorrelation, and atmospheric conditions. A minimum coherence value of 0.45 is selected in this study.
Baseline	There are two types of baseline criteria: (i) perpendicular (geometric) baseline, which is the distance between two acquisition locations perpendicular to the satellite's look direction; and (ii) temporal baseline, which is the difference in time between two acquisitions. In general, consistently short perpendicular and temporal baselines across a time-series enhance the coherence of the displacement results.
Band	A variable that represents the range of wavelengths captured by the sensor of a SAR satellite. In engineering-geoscience applications, there are 3 main bands: X-band (~3 cm), C-band (~6 cm), and L-band (~24 cm). A higher wavelength implies deeper penetration by the signal into the target surface. Only C-band Sentinel-1 data is processed in this study.

Continued from Table 4-1	
Orbit and look direction	The orbit direction of a SAR satellite is either in the ascending (roughly northward) or descending (roughly southward) direction. A SAR satellite is conventionally right-looking (i.e. looking eastward if orbiting northward).
Revisit interval	The frequency with which the SAR satellite revisits the same geographic location with the same acquisition geometry. This is typically 12 days, and sometimes 6 days, for Sentinel-1.
Line-of-sight (LOS) incidence angle	The oblique view angle between the satellite and the target object, measured from the vertical. In this article, cross-sectional diagrams are drawn for each case study to schematically illustrate the LOS.
LOS displacement	The 1-D InSAR "change-detected" results represent movements along the LOS of the satellite. These measurements often do not represent the maximum rate of movement experienced by the target object, nor the direction towards which the maximum rate of movement is occurring. The LOS displacement implies the "detected change in position away (expressed by negative values) or towards (positive values) the satellite".
LOS velocity	The calculated average rate of detected movements along the LOS of the satellite, typically expressed in terms of mm per year.
Phase unwrapping	One of the most important steps in SAR data processing that aims to reconstruct the signal's original phase values, thus allowing the retrieval of the final LOS displacement measurements. Phase unwrapping errors can occur when, for example, the threshold of maximum detectable displacement is exceeded (see below).
LOS velocity threshold	The maximum displacement rate per annum that can be reliably detected in an InSAR processing software. This threshold can depend on the software being used, the satellite, and the environmental conditions of the study area. Bayaraa et al. (2022) reported a threshold of 426 mm/yr when processing Sentinel-1 data using TerraMotion's ISBAS algorithm. TRE Altamira reports that their SqueeSAR (PS-DS) algorithm is best-suited for displacement rates of < 1,000 mm/yr.
Reference point/location	The location that serves as the reference site for phase unwrapping, assigned either manually by the user or, in this case, by the software by default. The reference point is typically selected in a stable area with high coherence.

4.2 SAR data processing software

We used ENVI SARScape Analytics (v. 5.6) to process Sentinel-1 data for the ground-truth test site and all the forensic case studies. It is a cost-saving commercial software that offers a user-friendly, streamlined workflow for PS-InSAR analysis, whereby each processing step is automated and the

analysis time is reduced substantially, which was convenient in this multi-case investigation. The Analytics package is a condensed, limited version of the entire SARScape software suite, which has been previously used to analyze the 2019 Feijao (Brazil) TSF failure using both the PS and SBAS techniques (Gama et al., 2020). The steps that are automated in the PS processing chain include co-registration, interferogram creation, coherence generation, height estimation, baseline refinement, noise filtering, and phase unwrapping. There are inherent drawbacks in this automated approach; for instance, it is not possible to produce or extract individual interferograms, to modify any parameters or steps in the workflow, to view or modify the reference location, or to manually filter out noise. For a technical background on the PS algorithm, we refer to Ferretti et al. (2001), Crosetto et al. (2016), and references therein.

SARScape Analytics requires a minimum of 20 scenes with consistent temporal and geometric baseline and satellite data characteristics in order to run a successful PS analysis. We downloaded open-source Sentinel-1 SLC IW scenes from the Alaska Satellite Facility (ASF) web platform. For all cases, we assigned a minimum coherence of 0.45 in the PS processing workflow, which is the same value adopted by Grebby et al. (2021) in their ISBAS analysis of the Feijao failure. Other input variables into the software were the geoid type (EGM2008), the base global digital elevation model (DEM) which we selected to be 30-m resolution SRTM-3 v4, and the area of interest in KML format (must be between 4 and 25 km^2).

For the ground-truth test site and the Cieneguita (2018 Mexico) case study, the SAR images were processed using the proprietary SqueeSAR algorithm, in order to facilitate a comparison of results for these two cases. The processing was done by TRE Altamira based on instructions on the study area and time-series duration provided by the lead author (N. Rana). SqueeSAR overcomes the limitations of alternative software packages by integrating both PS and DS points during analysis, thus enhancing the spatial density of point-cloud displacement data in most terrains. SqueeSAR has been used to study tailings dam failures (2018 Cadia, Australia), open-pit slope instabilities (e.g. Carla et al., 2019a), urban deformation (e.g. Bischoff et al., 2017, 2020), and natural landslides (e.g. Carla et al., 2019a,b). The technical framework of SqueeSAR is described in Ferretti et al. (2011). SqueeSAR is best-suited to monitor displacement rates of < 1,000 mm/yr.

4.3 Ground-truth assessment

4.3.1 Background and approach

As a complementary lead-up to the forensic case studies, we undertook a ground-truth (accuracy) assessment at a tailings dam situated in a cold-climate setting. Key identifier details of the mine and TSF were kept confidential to ensure site anonymity. We compared Sentinel-1B PS-InSAR LOS cumulative displacement results to in-situ data captured via 4 monitoring prisms (MP2, MP3, MP8, and MP9) over the same study period. The MPs were installed in late 2018 along the tailings dam. MP2 and MP3 lie at the crest and along the downstream slope, respectively, of the SE corner of the dam, whereas MP8 and MP9 are installed at the crest and along the downstream slope, respectively, at the approximate center of the embankment. The accuracy for these MPs is approximately ±10 mm. The errors are mainly due to setup issues of the total station, and the magnitude of the errors vary depending on how far the prism is located from the setup location.

The Sentinel-1B track used for analysis had an ascending (i.e. roughly south-to-north) orbit geometry and a LOS incidence angle of 31º. The ~1 km-long tailings dam trends north-south, whereby the downstream slope of the dam faces eastward away from the satellite but is still exposed at an acute angle to the satellite's LOS. The MP data indicates that certain sections of the dam crest exhibited some settlement deformation mainly towards the upstream (western) direction, whereas the downstream (east-facing) slope of the dam has shown relatively stable behaviour with minimal cumulative movement. The main objective here was to check if Sentinel-1 InSAR, as processed via both SARScape Analytics and SqueeSAR, shows reasonable consistency with the MP data and can provide adequate representations of low- and moderate-deformation areas (i.e. with displacement rates between 0 and 50 mm/yr).

On SARScape Analytics, we processed 36 Sentinel-1B images from May 2019 to October 2021, excluding the winter months from the processing stack because snow/ice is known to reduce the coherence of InSAR observations (Carla et al., 2019b; Kim et al., 2022). The SqueeSAR processing involved a total of 42 Sentinel-1B images over roughly the same study duration, also excluding the winter images. We confirmed the start- and end-dates of the winter months by checking for snow/ice cover on the TSF on high-resolution, high-frequency RapidEye and PlanetScope satellite imagery.

To facilitate a fair time-series comparison, we followed these steps:

1. The original MP datasets included results collected during the winter months, whereas Sentinel-1 images corresponding to winter months were excluded from the processing stack. To avoid temporal inconsistency, we baselined the cumulative displacement results from the Sentinel-1 and MP datasets to the start of the Spring season in each study year – i.e., the graphical comparisons encompass Spring-Fall of 2019, Spring-Fall of 2020, and Spring-Fall of 2021.

2. The MP data was originally reported as horizontal-easting and vertical. To avoid geometric inconsistency, we converted/projected the MP data to the Sentinel-1 LOS (31º eastward) using the following formula: *LOS-projected displacement = vertical displacement * cos(31º) – horizontal-easting displacement * sin(31º).*

3. For the comparisons, we selected the InSAR data point closest to the corresponding MP.

4. We inserted error bars for each data point, which are represented by the estimated average standard deviation of incremental LOS displacements in the InSAR data (±3 mm for SqueeSAR and ±6 mm for SARScape Analytics) and the approximate accuracy for the MP data as reported by the mine engineers (±10 mm).

4.3.2 Results

Figure 4-2 shows the LOS velocity maps (i.e. calculated mean displacement rate per annum) from SARScape Analytics and SqueeSAR, enabling a side-by-side visual comparison. Figure 4-2 also includes a cross-section diagram to elucidate the geometric relationship between the satellite, the dam, and the MPs. Areas where InSAR pixels are sparse due to low coherence are attributed to surface wetness, indicating the presence of tailings or water.

An important distinction is apparent between the two software: the level of noise in the results. This is confirmed in Figure 4-3 which shows histogram distributions of the LOS velocity and coherence values. The LOS velocities produced in SARScape Analytics are normally distributed with a mean of 3.0 mm/yr and a standard deviation of 10.9 mm/yr, but with two extreme tails that are indicative of noise. In comparison, the SqueeSAR results indicate a mean LOS velocity of -3.3 mm/yr with a lower standard deviation of 6.7 mm/yr without the extreme tails. Furthermore, the total count of data points in SARScape Analytics is much higher than SqueeSAR (5,144 vs. 3,859), whereas the mean coherence

is 0.64 for SARScape Analytics and 0.89 for SqueeSAR. This indicates a stronger noise-filtering proficiency, and consequently higher reliability, in the results produced in SqueeSAR.

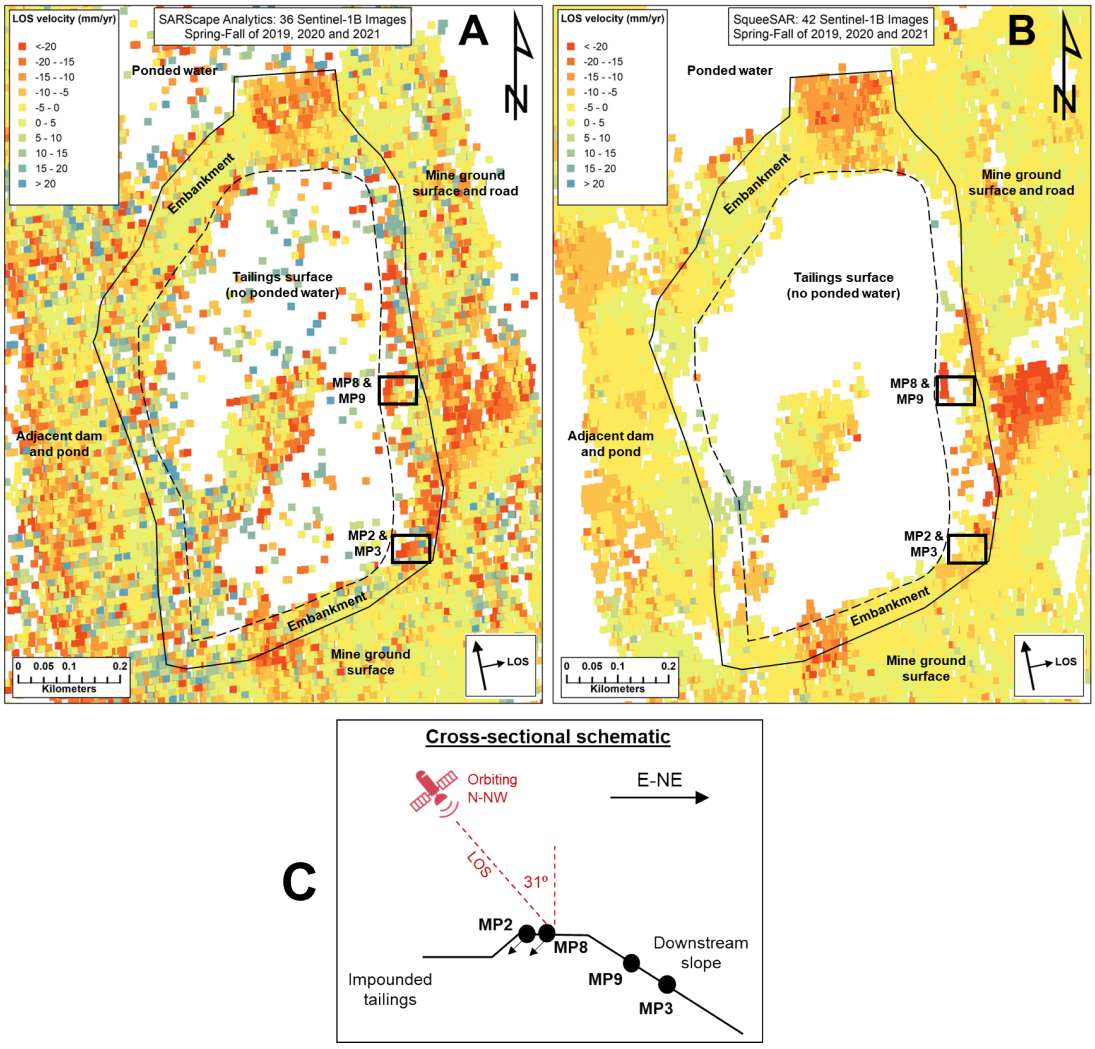

Figure 4-2: Sentinel-1 InSAR line-of-sight (LOS) velocity maps of the anonymized ground-truth TSF test site based on [A] SARScape Analytics and [B] SqueeSAR. The maps are annotated with the locations of the selected monitoring prisms (MPs), the surface of the exposed tailings (shaded outline), the embankment (solid outline), the mine ground surface, ponded water sites, and the adjacent dam. Negative (red) values indicate detected movements away from the satellite, positive (blue) values indicate detected movements towards the satellite, and green-yellow values indicate detected stable areas. [C] Cross-sectional schematic illustrating the geometric relationship between the satellite, the tailings dam, and the MPs. The small black arrows from MP points indicate the direction of bulk movement in the dam according to MP data (MP2 and MP8 at the crest), whereas the absence of small black arrows at MP points signify low rates of deformation according to MP data (MP9 and MP3 on the downstream slope).

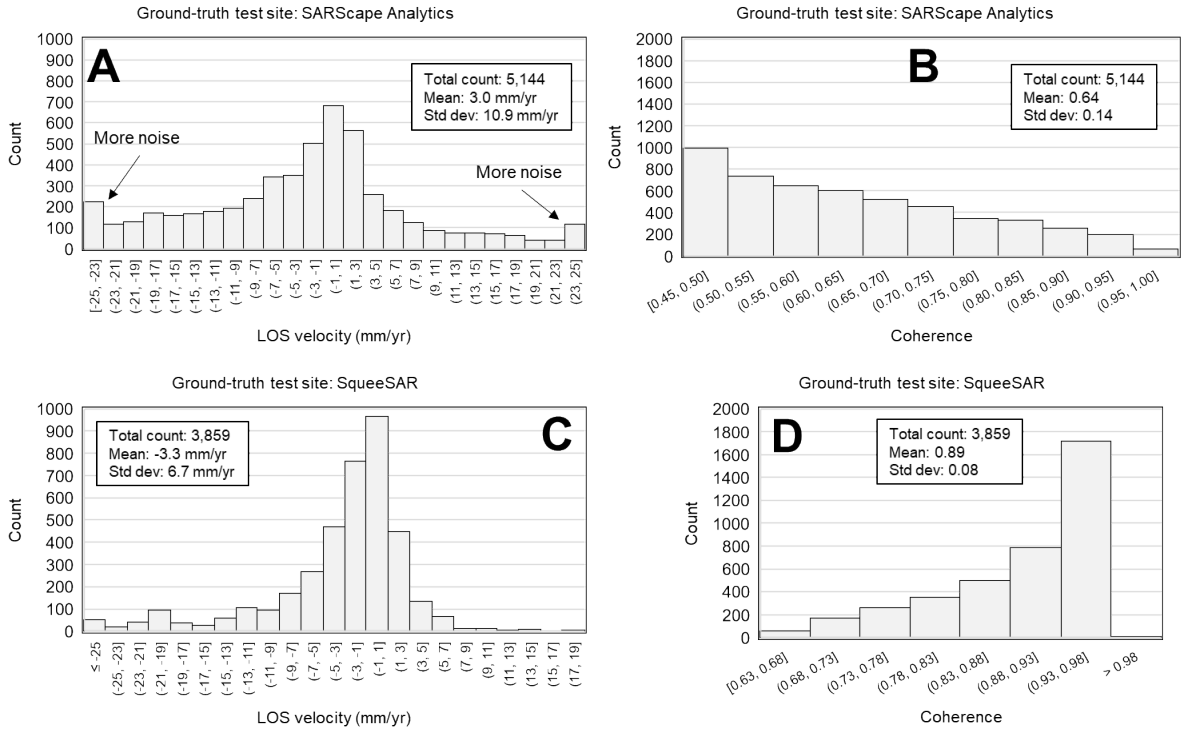

Figure 4-3: Histogram distributions of line-of-sight (LOS) velocities (mm/yr) and coherence values of the Sentinel-1 InSAR data for the ground-truth TSF test site. Results processed via SARScape Analytics are shown in [A] and [B] whereas SqueeSAR results are shown in [C] and [D].

Figure 4-4 shows the graphical comparisons for the dam crest (MP2 and MP8) over the Spring-Fall seasons of 2019-2021. The InSAR results generally show reasonable consistency with the MP data. There are, however, exceptions when the SqueeSAR data deviates towards positive displacements (2020 MP2 plot) and when there are potential errors in the MP8 data in 2019 and 2020. SqueeSAR exhibited a lower density of pixels along the dam crest compared to SARScape Analytics, which may be due to different thresholds or criteria of filtering out data points in the two software.

Figure 4-5 shows the graphical comparisons for the downstream slope of the dam (MP3 and MP9). The InSAR results generally lie close to the MP data, indicating local-scale representativeness. One notable exception is MP9 in 2020, where the SqueeSAR results are higher than those of SARScape Analytics, but there was no MP data collected during this period that would help with verification.

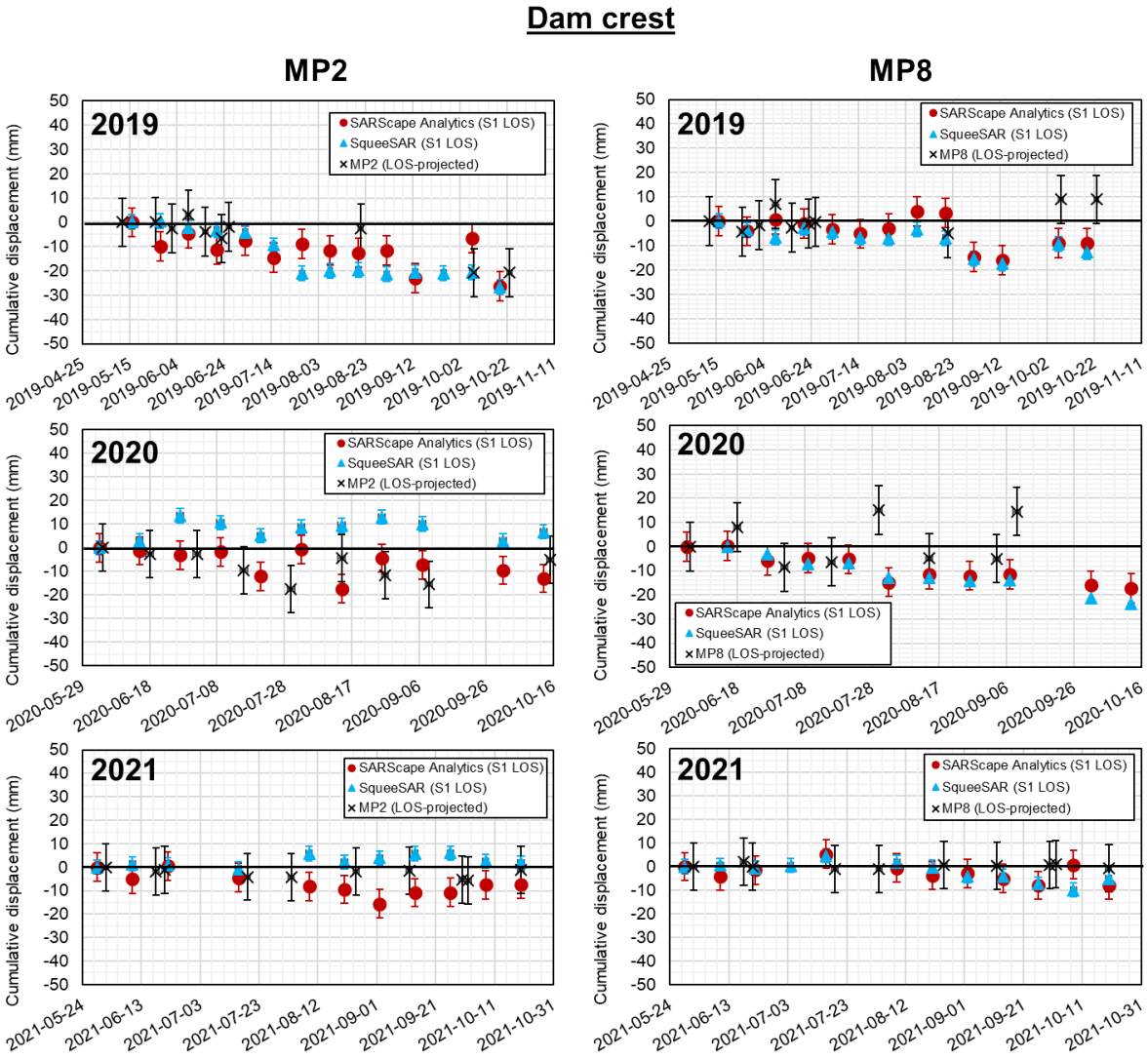

Figure 4-4: Ground-truth results comparing the Sentinel-1B line-of-sight (LOS) InSAR cumulative displacement data (processed on SARScape Analytics and SqueeSAR) to *in-situ* cumulative displacement data from monitoring prisms MP2 and MP8 situated on the crest of the tailings dam. The study period encompasses the Spring-Fall seasons of 2019-2021 (i.e. winter images were excluded from the processing stack). Error bars represent standard deviations of ±3 mm for SqueeSAR and ±6 mm for SARScape Analytics and an approximate accuracy of ±10 mm for the MPs.

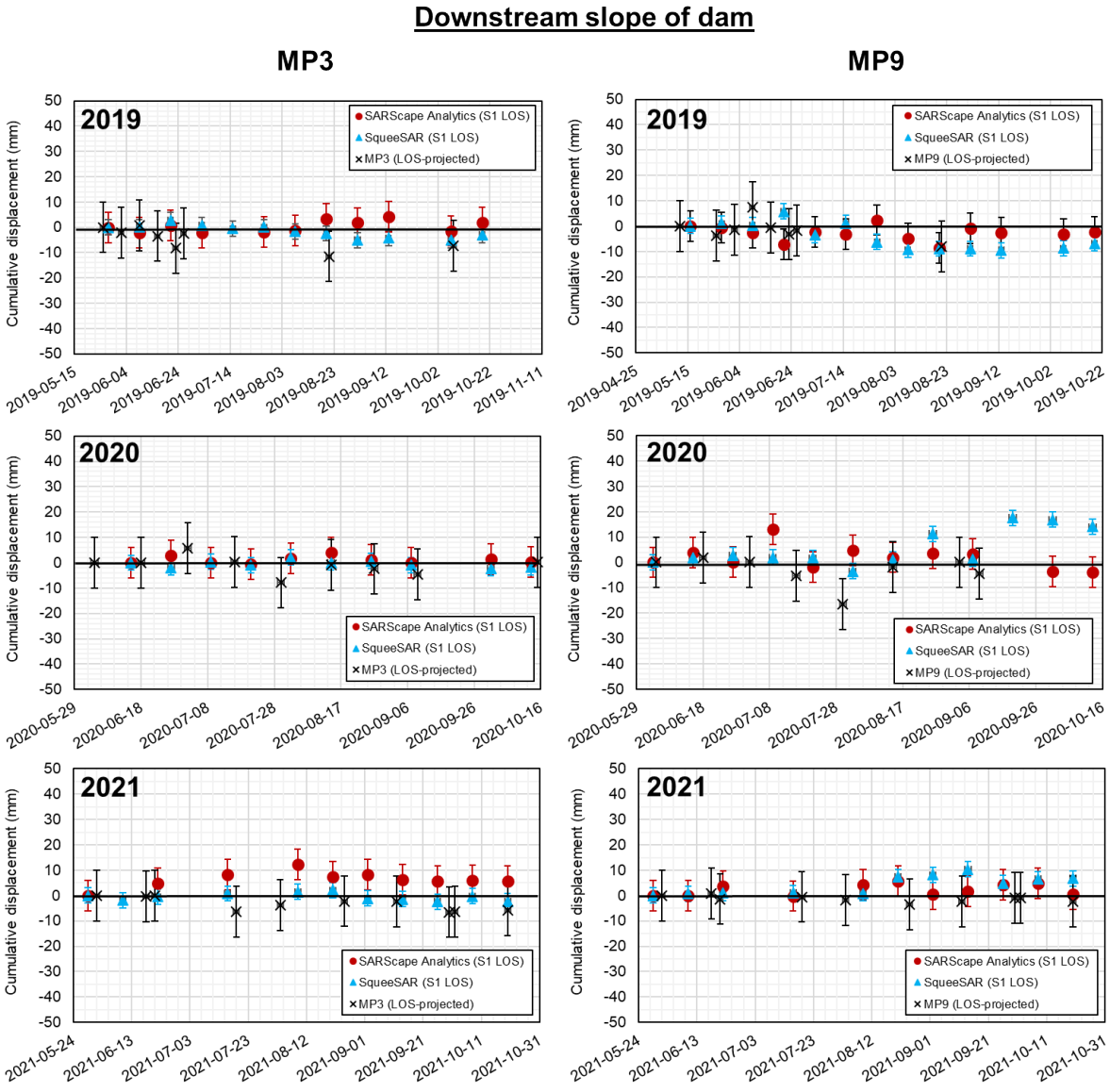

Figure 4-5: Ground-truth results comparing the Sentinel-1B line-of-sight (LOS) InSAR cumulative displacement data (processed on SARScape Analytics and SqueeSAR) to *in-situ* cumulative displacement data from monitoring prisms MP3 and MP9 situated on the downstream slope of the tailings dam. The study period encompasses the Spring-Fall seasons of 2019-2021 (i.e. winter images were excluded from the processing stack). Error bars represent standard deviations of ±3 mm for SqueeSAR and ±6 mm for SARScape Analytics and an approximate accuracy of ±10 mm for the MPs.

4.4 Forensic case studies

4.4.1 Background and approach

In this section, we investigate the precursor LOS displacements in 7 tailings dams that experienced a breach in the period 2017-2019. Background information on the case studies is listed in Table 4-2. Details on the Sentinel-1 image processing stacks for each case are provided in Table 4-3, with the exception of the ground-truth test site to maintain its site anonymity. All of the cases were processed using SARScape Analytics, while only the Cieneguita case was also processed using SqueeSAR to show a comparison. We selected these cases after screening databases of TSF failures (e.g. Islam and Murakami, 2021; Rana et al., 2021, 2022) based on two criteria. First, the cases encompassed a variety of breach mechanisms, allowing us to check for distinct displacement patterns depending on the failure mechanism (e.g. erosion vs. liquefaction vs. subsidence). Second, the 7 cases occurred in 7 countries with diverse climatic, topographic, and land-use regimes. This allowed us to check for environmental factors (e.g. wet or vegetated areas vs. dry desert or hard-surface sites) that influence the quality of the InSAR results.

For some of the cases, the predisposal variables (i.e. the underlying causal variables that preconditioned the precarious stability of the TSF), trigger mechanisms (i.e. a natural or anthropogenic activity that caused the breach to occur at a particular location at a given time), and failure modes (i.e. the mechanism by which the breach and outflow occurred) are poorly described in existing literature. These underlying factors form the background story of each case, therefore our knowledge of these factors has an influence on our judgment of the InSAR results. In Table 4-4, we assign qualitative levels of knowledge uncertainty for each case:

- “High uncertainty” means that the factor is virtually unknown due to lack of research.
- “Medium uncertainty” means that there are news articles that have provided basic information on the factor, or that there are satellite images that provide some basic insights into the factor.
- “Low uncertainty” means that the factor has been well-documented in scientific material.

Our main objective was to check whether the location and timing of the breach could have been predicted in advance using InSAR. For each case study, we adopted the following consistent approach:

1. We presented the LOS velocity map illustrating the calculated mean displacement rates per annum across the site.
2. We identified localized hotspots of detected high-velocity movements in the unstable section and, where applicable, in other sections of the dam exhibiting similar detected behavior.
3. We plotted the cumulative LOS displacement time-series for these hotspot locations. The plotted cumulative displacement results are bounded by shaded error bar regions that represent the estimated average standard deviation of incremental LOS displacement measurements for each case study site.
4. We checked if accelerations in the precursor displacement patterns could be observed.

Where precursor identifiers on the breach location and/or timing were not found, we provided explanations on the basis of the failure mechanism (if this information was available), the limitations of the software, and/or the limitations of medium-resolution Sentinel-1 data. By this, we demonstrated the capabilities and limitations of our approach to monitoring metastable tailings dams. Out of all the 7 case studies, only Cadia and Feijao have been forensically studied using InSAR (Carla et al., 2019a; Thomas et al., 2019; Gama et al., 2020; Rotta et al., 2020; Holden et al., 2020; Grebby et al., 2021; Hudson et al., 2021; Bayaraa et al., 2022), which allows us to compare our results versus previously published results for these two events.

Table 4-2: Basic descriptions of TSF failure cases analyzed in this study by Sentinel-1 InSAR.

TSF failure	Country	Site location (Lat, Long)	Failure date (Y-M-D)	Failure description	Case references
Tonglvshan	China	30.08, 114.95	2017-3-12	Subsidence in underground mine goaf triggered detachment of dam section.	Wang et al. (2018); Ghahramani et al. (2020); Rana et al. (2021a,b); Zhuang et al. (2021)
Mishor Rotem	Israel	31.06, 35.21	2017-6-30	Cause and trigger unknown; failure mode involved erosional breach and outflow of ponded water based on interpretation of satellite imagery and drone video.	Petley (2017); Ghahramani et al. (2020); Rana et al. (2021a,b, 2022a,b)
Cadia	Australia	-33.5, 148.99	2018-3-9	Failure along weak, compressible volcanic bedrock foundation triggered static liquefaction of tailings and a breach along a section of the dam.	Jefferies et al. (2019); Carla et al. (2019a); Thomas et al. (2019); Bayaraa et al. (2022)
Cieneguita	Mexico	27.12, -108.03	2018-6-4	Internal erosion of poorly compacted embankment. TSF was highly active according to time-series imagery.	Ghahramani et al. (2020); Rana et al. (2021a,b, 2022a,b)
Feijao	Brazil	-20.12, -44.12	2019-1-25	Static liquefaction of tailings and complete collapse of the dam, which may have been triggered by drilling activity on a metastable section of the dam.	Robertson et al. (2019); Gama et al. (2020); Rotta et al. (2020); Grebby et al. (2021); Arroyo and Gens (2021)
Hindalco	India	23.36, 85.86	2019-4-9	Breach of bauxite TSF that was dammed by poorly constructed, steep, gabion retaining walls. Trigger is unknown, and an undrained failure mechanism is suspected based on satellite images and local reports.	Ghahramani et al. (2020); Rana et al. (2021a,b, 2022a,b)
Cobriza	Peru	-12.58, -74.37	2019-7-10	Cause and trigger unknown; failure mode involved erosional breach in upper section of dam according to interpretation of satellite imagery.	Petley (2019); Ghahramani et al. (2020); Rana et al. (2021a,b, 2022a,b)

Table 4-3: Characteristics of the Sentinel-1 data acquired for each case study. SARScape Analytics was used as the processing software for all of these cases, while Cieneguita also involved an additional analysis using SqueeSAR.

TSF failure	SAR mission (band)	Product type	Beam mode	Orbit and look direction	Orbit track ID	Revisit time	Incidence line-of-sight (LOS) angle	Polarization	Time-series duration	No. of SAR images
Tonglvshan	Sentinel-1A/B (C-band)	Single Look Complex (SLC)	Interferometric Wide (IW)	Ascending (looking E-NE)	40	12 days	33.8º	VV+VH	2015-06-05 to 2017-03-02 (21 months)	49
Mishor Rotem					87	6 and 12 days	38.0º		2016-01-10 to 2017-06-27 (19 months)	64
Cadia				Descending (looking W-NW)	45	12 days	35.3º		2016-09-27 to 2018-02-25 (17 months)	45
Cieneguita					56		33.5º		2017-02-19 to 2018-06-02 (16 months)	39 and 40
Feijao					53 and 155		32.5º and 45º		2017-06-13 to 2019-01-22 and 2017-06-08 to 2019-01-17 (19 months)	50 and 50
Hindalco					121		38.2º		2017-09-04 to 2019-03-22 (20 months)	52
Cobriza					98		40.9º		2017-09-08 to 2019-06-30 (22 months)	54

Table 4-4: Assigned qualitative levels of knowledge uncertainty for each case study based on our literature review of the predisposal variables, trigger mechanisms, and failure modes (see definitions in text).

TSF failure case	Level of knowledge uncertainty		
	Predisposal variables	**Trigger mechanism**	**Failure mode**
2017 Tonglvshan, China	Medium	High	Low
2017 Mishor Rotem, Israel	Medium	High	High
2018 Cadia, Australia	Low	Low	Low
2018 Cieneguita, Mexico	Medium	High	Medium
2019 Feijao, Brazil	Low	Low	Low
2019 Hindalco, India	Medium	High	High
2019 Cobriza, Peru	Medium	High	Medium

4.4.2 Results

The following sub-sections present the SARScape Analytics results for each forensic case study. In addition, Section 4.2.4 also illustrates the SqueeSAR results for Cieneguita. The presentation of the InSAR results is preceded by a background review of the site conditions and reported failure description. For all of the cases except Tonglvshan, we constructed time-lapse videos using between 60 and 120 PlanetScope (3 m resolution) optical satellite images to show the evolution of the TSFs until their breach. The open-access URL links to these time-lapse videos are listed in Table A.1 in the Appendix.

4.4.2.1 2017 Tonglvshan, China

The Tonglvshan copper-iron TSF is located near Daye City in Hubei Province, China. Historical satellite imagery on Google Earth shows that the TSF was operational dating back to the early 2000s. The TSF is situated on part-urban, part-agricultural terrain that gently dips towards the north at ~0.3º. The broader mine site is interspersed with densely packed villages, including one along the western boundary of the TSF. The mine is bordered by a river to the north and to the east. The pre-failure storage volume was 15.8 M m3 of mostly tailings and some surface water (Zhuang et al., 2022). The surface area of the impoundment is ~500,000 m2, enclosed by a 3 km-long perimeter embankment.

On 12 March 2017, the northern section of the dam experienced a breach with an approximate geometry of 50 m width x 16 m height (Figure 4-6). The released volume was reportedly ~500,000 m^3

(Zhuang et al., 2022) that inundated an area of 300,000 m^2 (Ghahramani et al., 2020). The flow was terminally confined by a raised dike at the opposite end of the impacted field (Rana et al., 2021; Zhuang et al., 2022). The failure reportedly occurred due to the long-term effects of weathering and gravity in the roof granite in the mine goaf underneath the TSF. This caused the upper strata to fail, which led to the subsidence and brittle instability in the dam foundation (Zhuang et al., 2022). The tailings subsequently pooled against the northern section of the TSF (the lowest elevation in the TSF area) and caused a breach by the impact force, after which the dam rotationally slid ~45º clockwise without disintegrating (i.e. remained largely intact) (Zhuang et al., 2022). The incident led to 2 deaths and 6 injuries.

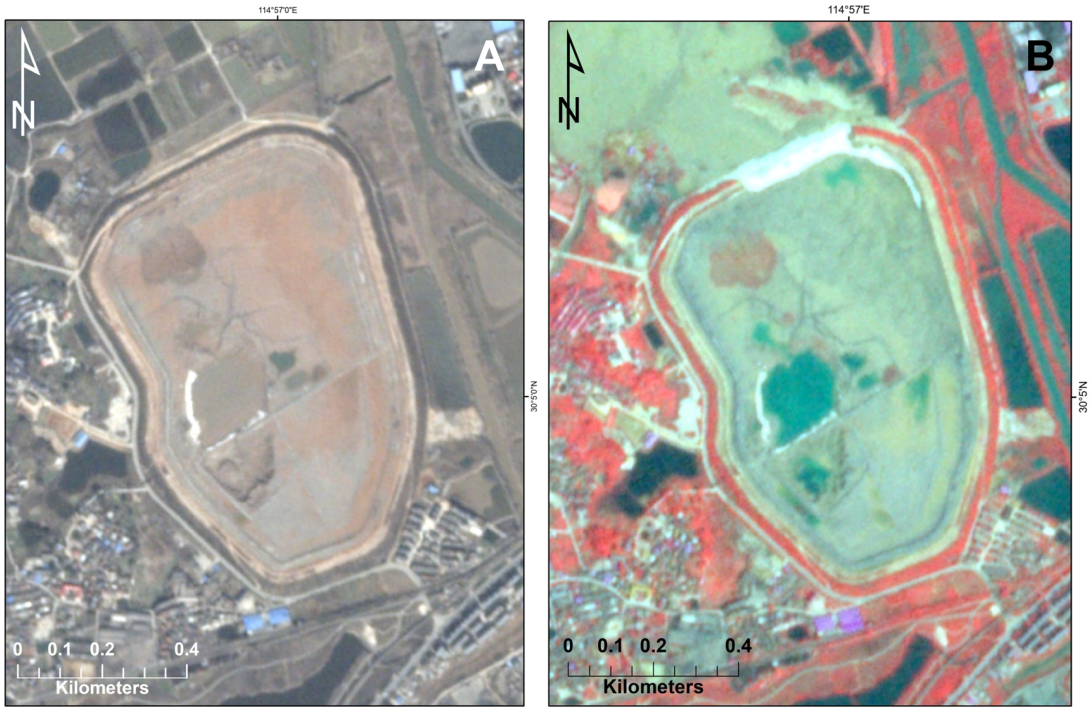

Figure 4-6: [A] Pre-failure (13 February 2017) and [B] post-failure (13 April 2017) PlanetScope (3 m resolution) images of the 12 March 2017 Tonglvshan tailings dam breach in China.

Figure 4-7 shows the results from our analysis of 49 Sentinel-1 images over a duration of 21 months. The satellite had an ascending orbit parallel to the length of the TSF with a LOS incidence angle of 33.8º towards E-NE. This implies that, at the northern section that experienced the breach, displacements in the downstream (N-NW) direction were along-track in relation to the satellite orbit

(Figure 4-7B) and thus are poorly captured. However, LOS components of vertical, settlement- and subsidence-related deformation can be measured. Figure 4-8 presents histograms of LOS velocities and coherence values. The LOS velocities are normally distributed around ~0 mm/yr with a standard deviation of 8.4 mm/yr, while the mean coherence is relatively low (0.55), which reflects the site heterogeneity, including the presence of vegetation and water ponds.

We observe an abundance of PS points along the perimeter embankment and on the adjacent infrastructure, but a sparsity in the impoundment due to high surface wetness, which reduces coherence (Figure 4-7A). The LOS velocities in the infrastructure surrounding the TSF were generally close to 0 mm/yr (i.e. stable). This contrasts with the tailings dam, where movements along the western half of the tailings dam were much more pronounced than those along the eastern half, where PS points were either sparse (along eastern wall) or indicated minimal movement (SE wall). This observation may be a reflection of the location and extent of the underground mine goaf that reportedly was a major cause of the breach.

We identify two concentrated hotspots of anomalously high LOS velocities: one along the northern section that eventually breached (Plot 1 area; LOS cumulative displacements plotted in Figure 4-7C) and another at the middle of the western wall that remained stable (Plot 2 area; Figure 4-7D). The behaviour of LOS deformation was similar in both locations, suggesting a consistent rate movement away from the satellite (i.e. settlement and potential subsidence activity), but no precursor accelerations are detected. In fact, the results in Figure 4-7C suggest that the unstable northern section experienced minimal cumulative movement in the 2 months preceding the collapse. However, the unstable northern section was geometrically unfavorable in relation to the satellite's LOS compared to the western wall that was directly facing the LOS, which influences our interpretation of these results. Furthermore, phase unwrapping errors may have been manifested at the unstable northern section due to high deformations approaching the threshold, but this cannot be confirmed via the present automated software.

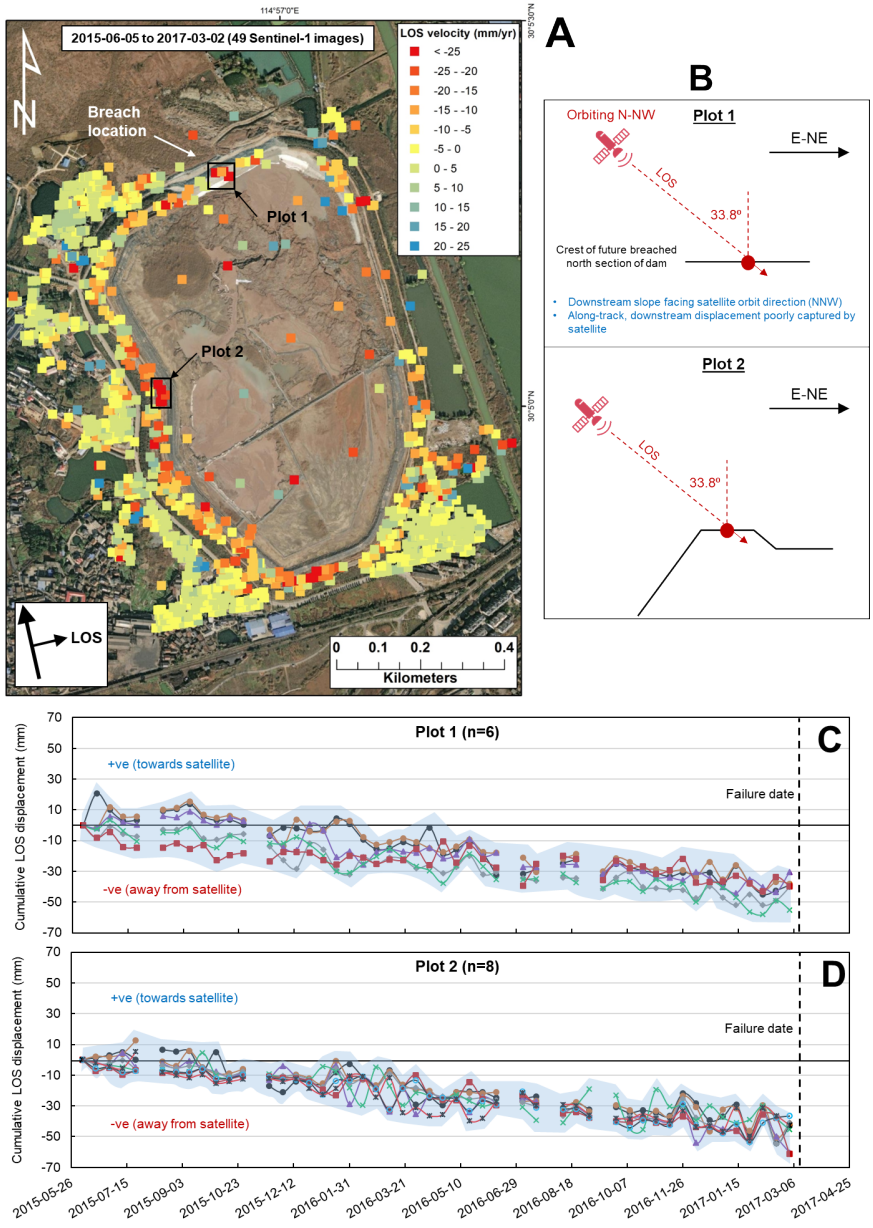

Figure 4-7: Sentinel-1 PS-InSAR results, processed on SARScape Analytics, for the 12 March 2017 Tonglvshan tailings dam breach in China. [A] Line-of-sight (LOS) velocity map, annotated with the breach location and the Plot 1 and Plot 2 areas that we selected for graphical analysis. Negative (red) values indicate detected movements away from the satellite, positive (blue) values indicate detected movements towards the satellite, and green-yellow values indicate detected stable areas. [B] Cross-sectional schematics for the Plot 1 and Plot 2 areas illustrating the geometric relationship between the satellite, the tailings dam, and the PS points selected for graphical analysis. [C-D] Cumulative LOS displacements over the time-series duration for the Plot 1 and Plot 2 areas. Error bar shading represents an estimated average standard deviation of ±7.0 mm for incremental LOS displacement measurements.

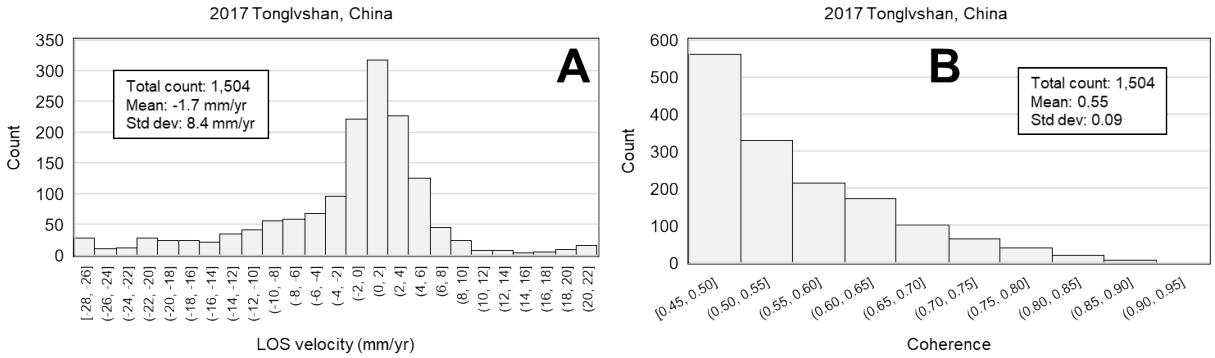

Figure 4-8: Histograms of [A] line-of-sight (LOS) velocities (mm/yr) and [B] coherence of the Sentinel-1 PS-InSAR data, processed on SARScape Analytics, for the 2017 Tonglvshan TSF breach in China.

4.4.2.2 2017 Mishor Rotem, Israel

The Mishor Rotem TSF is located about 30 km SW of the Dead Sea in the Negev Desert of Israel. The TSF consists of three compartments with a mean dam height of 60 m (Rana et al., 2021) and covers a surface area of ~500,000 m^2 with a perimeter dike length of 3 km. The TSF is one of several impoundments at the site storing highly acidic phosphogypsum tailings and supernatant wastewater sourced from the local open-pit mine, quarry, and power plant operations. The NE compartment of the TSF experienced a breach on 30 June 2017 (Figure 4-9), releasing at least 100,000 m^3 of tailings and wastewater that flowed for 28 km along a dry creek channel. We developed a time-lapse video of the Mishor Rotem TSF using 80 PlanetScope images between January 2016 and July 2017 (Table A.1). The time-lapse shows that there was a substantial volume of ponded wastewater abutting the breached section of the dam for a number of months preceding the breach (Rana et al., 2021). However, the predisposal and trigger variables and the failure mechanism remain unclear. A drone video of the breached dam is publicly available via YouTube (The Portal for Natural Agriculture and the Environment, 2017).

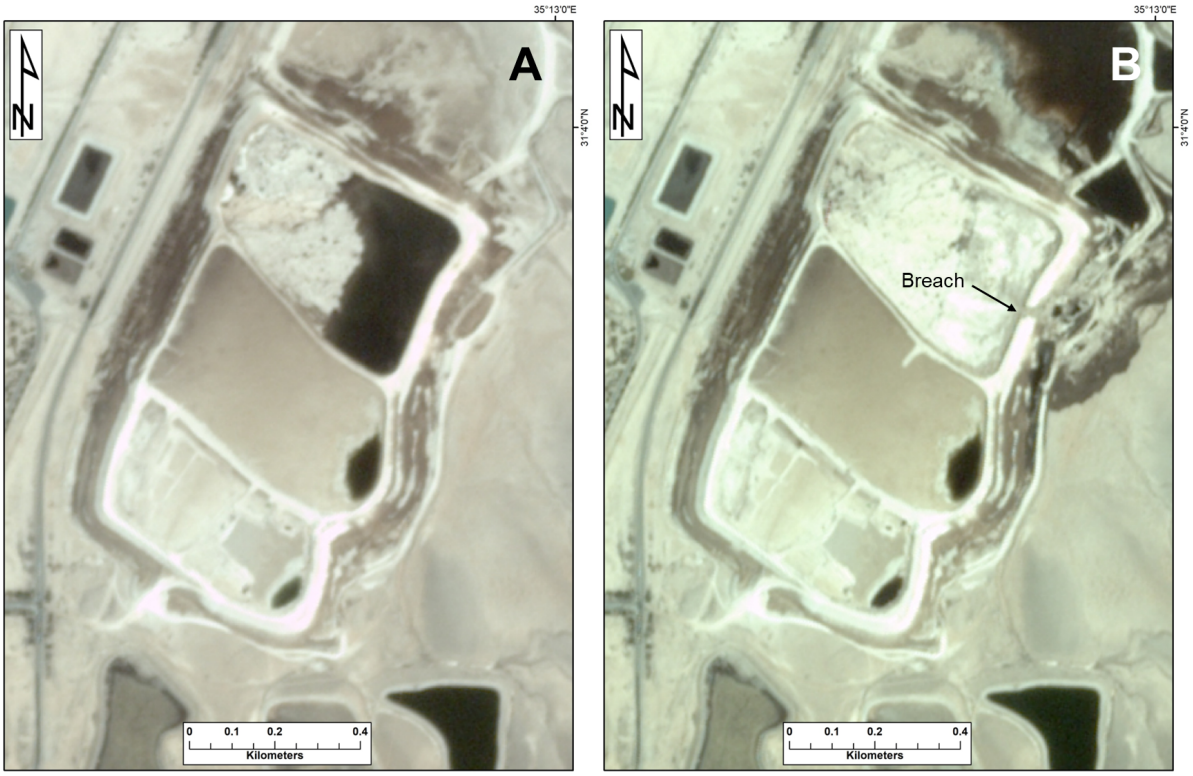

Figure 4-9: [A] Pre-failure (27 June 2017) and [B] post-failure (3 July 2017) PlanetScope (3 m resolution) images of the 30 June 2017 Mishor Rotem tailings dam breach in Israel.

Figure 4-10 illustrates the PS results from the processing of 64 Sentinel-1 images over a 19-month duration, with a revisit time of mostly 6 days and sometimes 12 days. The satellite's orbit track was ascending with a LOS incidence angle of 38º. The unstable section of the dam was exposed to the satellite's LOS at an oblique angle (Figure 4-10B), which makes the geometry complex and unfavorable when interpreting the LOS precursor displacement results at the breach location. Figure 4-11 shows that the LOS velocity results are normally distributed around ~0 mm/yr with a standard deviation of 10 mm/yr, and that the mean coherence is relatively high at 0.68. The high quality of the PS results is attributed to the ideal environmental conditions (dry desert) and the large number of processed Sentinel-1 images with a shorter mean revisit time.

There are several key observations to be noted here:

1. Compared to the surrounding power plant infrastructure, water storage dams, and the natural desert, the perimeter embankment of the Mishor Rotem TSF exhibited anomalously high LOS velocities over the pre-failure study period.

2. The western corner of the perimeter dam, which was directly facing the satellite (thus making it geometrically ideal), was moving outward (i.e. westward towards the satellite) at a rate of >25 mm/yr. This peak LOS velocity was also observed along the NW slopes of the TSF as well as the unstable eastern section that eventually breached.

3. There are signs of apparent rotational deformation along the NW slopes and the southern dam face, where negative LOS velocities along and near the crest indicated movement away from the satellite and positive LOS velocities in the central and near-toe portions of the embankment indicated movement towards the satellite.

4. PS points were generally missing along the crest of the eastern half of the TSF. In contrast, the western half of the TSF contained a sufficient density of PS points along the crest. This could be due to the dam crest along the eastern half undergoing intense construction/raise activity, but the poor frequency of high-resolution, pre-failure satellite images covering the Mishor Rotem mine site prevented confirmation of this hypothesis.

5. The PS points closest to the breach location were in the bottom-half of the dam face (which, to reiterate, was obliquely exposed to the satellite's LOS). These points indicated movement towards the satellite, which appears counterintuitive. It is difficult to explain this deformation behaviour without possessing sufficient knowledge of the background story of this event. However, we infer that the likelihood of failure was exacerbated by the large supernatant pond that was abutting this unstable section of the dam (the base of the now-released pond appears black-colored in Figure 4-10A).

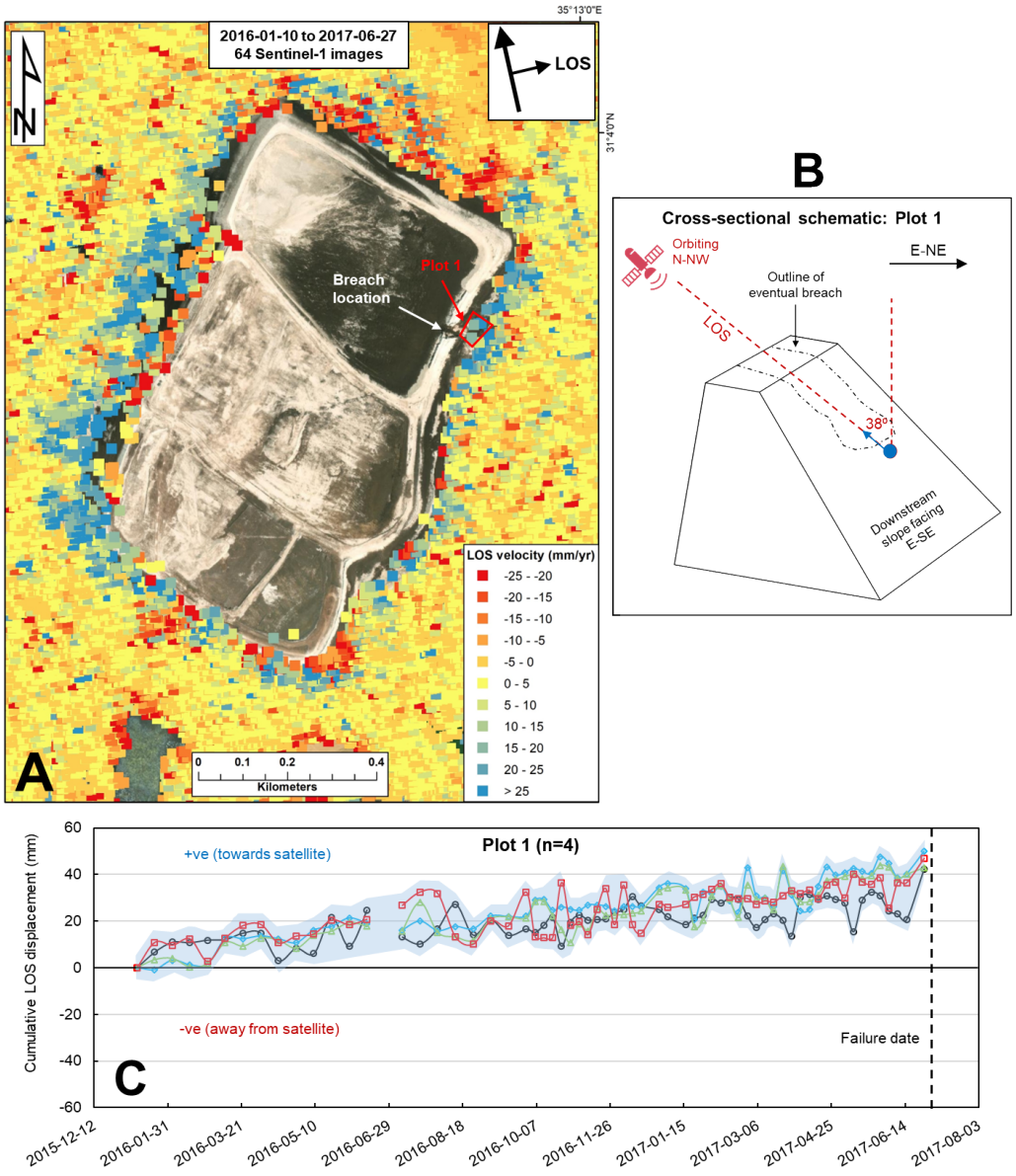

Figure 4-10: Sentinel-1 PS-InSAR results, processed on SARScape Analytics, for the 30 June 2017 Mishor Rotem tailings dam failure in Israel. [A] Line-of-sight (LOS) velocity map, annotated with the breach location and the Plot 1 area selected for graphical analysis. The basemap represents a post-failure satellite image. Negative (red) values indicate detected movements away from the satellite, positive (blue) values indicate detected movements towards the satellite, and green-yellow values indicate detected stable areas. [B] Cross-sectional schematics for the Plot 1 area illustrating the geometric relationship between the satellite, the tailings dam, and the PS points selected for graphical analysis. The small blue arrow indicates the direction of LOS movement (in this case, towards the satellite). [C] Cumulative LOS displacements over the time-series duration for the Plot 1 area. Error bar shading represents an estimated average standard deviation of ±4.9 mm for incremental LOS displacement measurements.

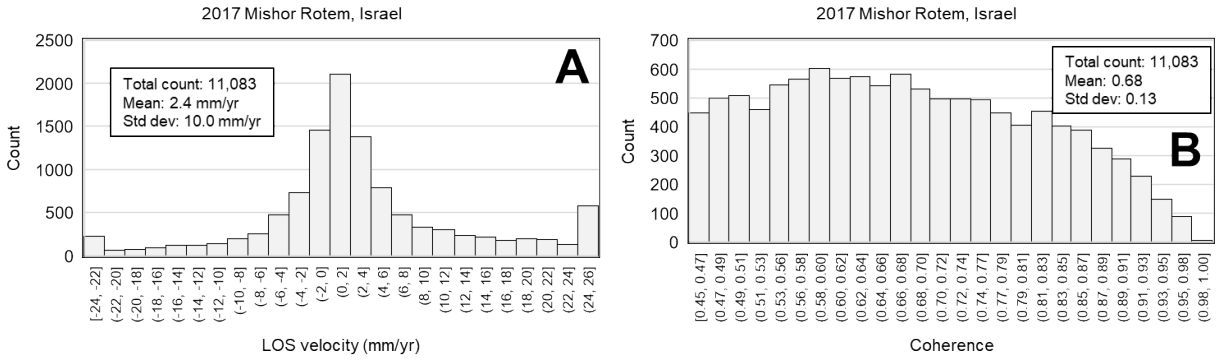

Figure 4-11: Histograms of [A] LOS velocities (mm/yr) and [B] coherence of the Sentinel-1 PS-InSAR data, processed on SARScape Analytics, for the 2017 Mishor Rotem TSF failure in Israel.

4.4.2.3 2018 Cadia, Australia

The Cadia gold mine in New South Wales, Australia operates two large TSFs: the Northern TSF (NTSF) and the Southern TSF (STSF), both of which are upstream-constructed, stepped side-hill impoundments (Jefferies et al., 2019). The NTSF was commenced in 1997 and covers an area of ~5 km^2 with a perimeter of ~10 km and a dam height of 94 m. The STSF covers ~4.5 km^2 with a perimeter of ~9 km. The storage volumes in the two TSFs are unknown. The TSFs are situated on rural land with minimal residential populations in the vicinity. According to SRTM data, the ground surface underneath the TSFs gently dips in the S-SW direction.

In March 2018, a ~300 m wide breach formed in the SW section of the NTSF (Figure 4-12). We created a time-lapse video of the Cadia dam using 120 PlanetScope images between September 2016 and September 2018 (Table A.1). The video shows that the failure process was two-staged: the initial breach occurred on 9 March 2018, and another secondary event occurred at the same location on 11 March 2018.

A volume of 1.3 M m^3 of statically liquefied tailings flowed into the lower STSF, which remained stable, thus preventing environmental contamination outside the mine perimeter. The Independent Expert Panel (Jefferies et al., 2019) identified the main failure cause to be a low-density, highly compressible Forest Reef Volcanics (FRV) layer in the foundation underneath the SW wall of the NTSF. This previously unidentified material was strain-weakening – that is, the unit became brittle

when subjected to high loads. Accelerated displacements over this foundation unit triggered static liquefaction in the loose, saturated tailings in the NTSF. The Cadia failure has been forensically examined by InSAR using proprietary processing software in several publications (Carla et al., 2019a; Jefferies et al., 2019; Thomas et al., 2019; Hudson et al., 2021; Bayaraa et al., 2022). These studies identified accelerating deformation in the 3 months preceding the dam collapse.

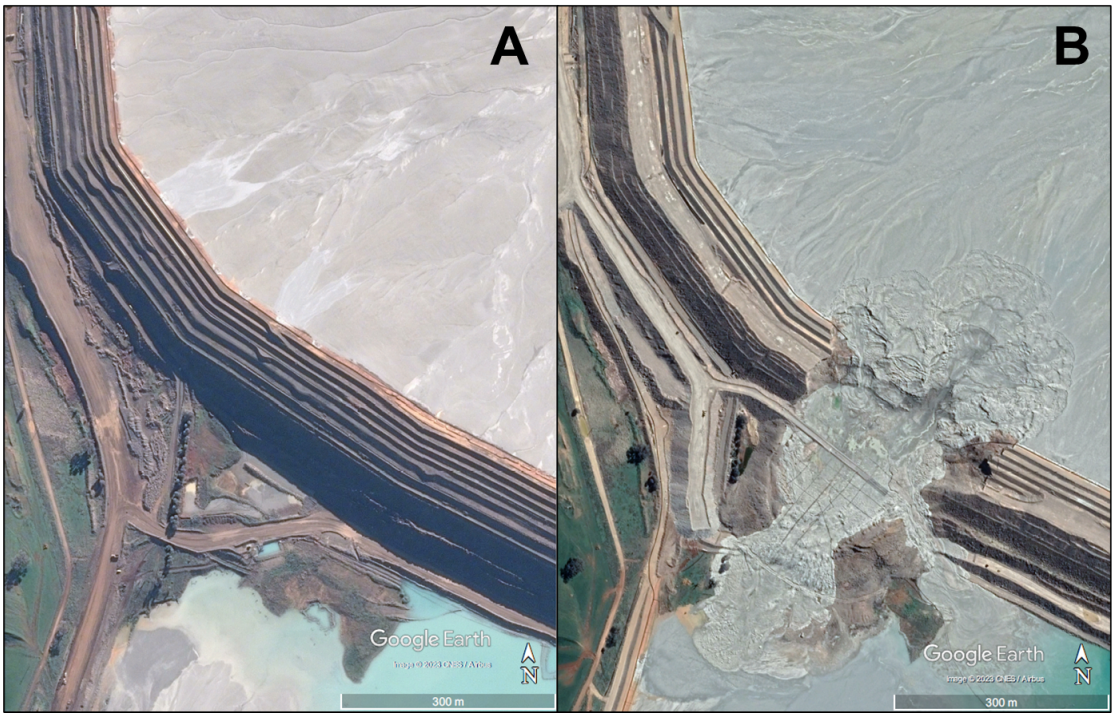

Figure 4-12: [A] Pre-failure (11 June 2016) and [B] post-failure (12 September 2018) Google Earth Worldview-2 (0.5 m resolution) images of the 9 March 2018 Cadia tailings dam breach in Australia.

Figure 4-13 shows our PS-InSAR results after processing 45 Sentinel-1 images over a 17-month duration. The satellite had a descending orbit track with a LOS incidence angle of 35º. The dam-satellite geometric relationship was not an influential factor here, because although the unstable dam face was obliquely exposed to the satellite's LOS, the vertical settlements along the dam were well-captured. Figure 4-14 shows that the mean coherence was relatively high at 0.65, reflecting the ideal environmental conditions, and that the LOS velocities are right-skewed due to the high density of settlement areas (with negative LOS velocity values) along the dam crest.

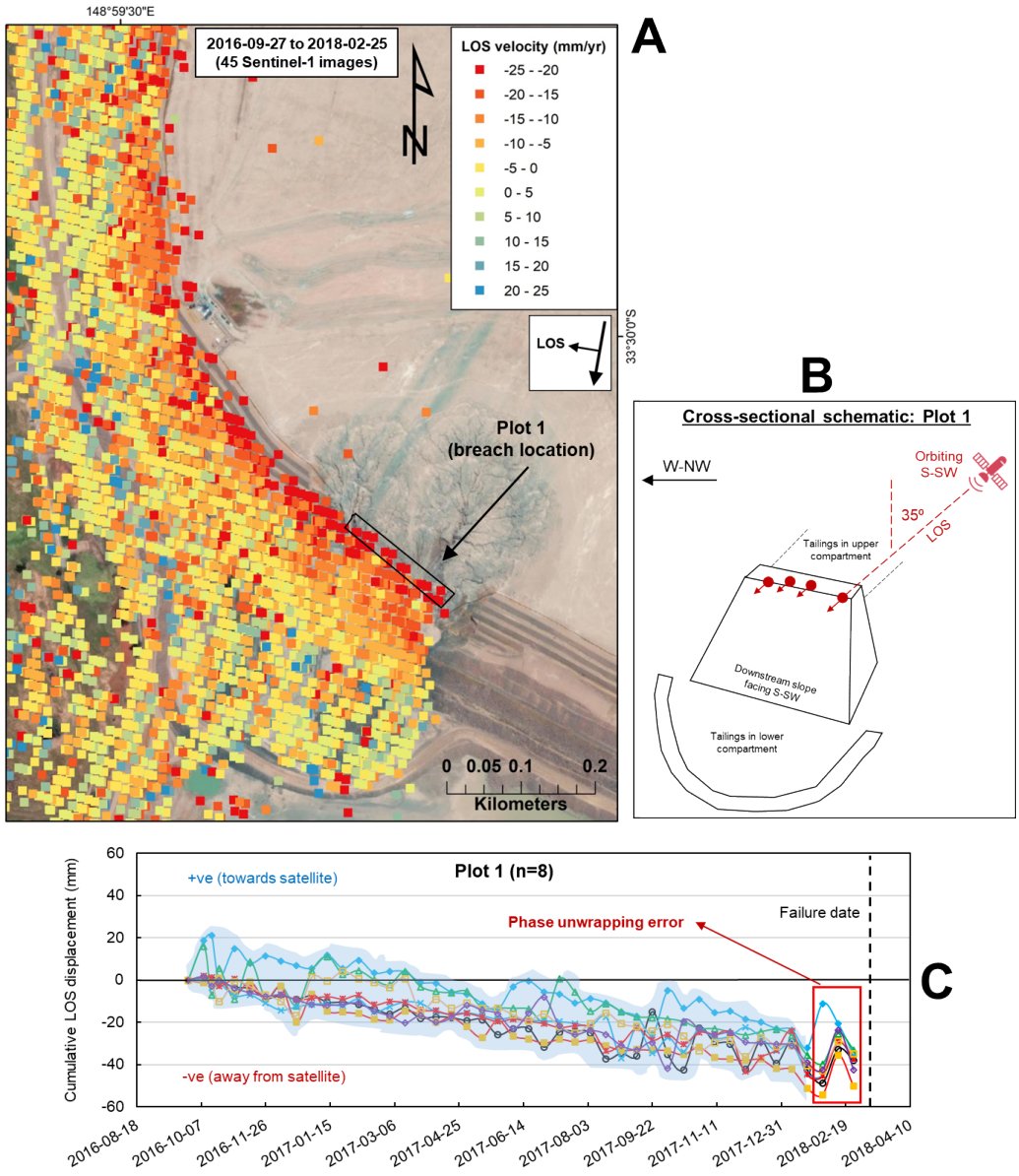

Figure 4-13: Sentinel-1 PS-InSAR results, processed on SARScape Analytics, for the 9 March 2018 Cadia tailings dam failure in Australia. [A] Line-of-sight (LOS) velocity map, annotated with the breach location and the Plot 1 area selected for graphical analysis. Negative (red) values indicate detected movements away from the satellite, positive (blue) values indicate detected movements towards the satellite, and green-yellow values indicate detected stable areas. [B] Cross-sectional schematic for the Plot 1 area illustrating the geometric relationship between the satellite, the tailings dam, and the PS points selected for graphical analysis. The small red arrows indicate the direction of LOS movement, in this case away from the satellite. [C] Cumulative LOS displacements over the time-series duration for the Plot 1 area. Error bar shading represents an estimated

average standard deviation of ±5.8 mm for incremental LOS displacement measurements until the phase unwrapping errors detected towards the end.

The LOS velocity patterns show that the vicinity of the unstable section, especially along the dam crest, represented a hotspot of anomalously high movements (Figure 4-13A). These observations resemble those seen in previous publications on the Cadia event. However, Figure 4-13C shows the manifestation of phase unwrapping errors. While the commencement of the acceleration phase is clearly observed in January 2018, the errors are manifested in the results from February 2018 up to the final acquisition date before 9 March 2018. This is a key technical limitation associated with commercial SAR processing software when monitoring relatively large, accelerating deformations. It has major implications for the ability to make time-to-failure predictions using cost-effective commercial software, as these errors were not observed when using proprietary software (Carla et al., 2019a; Bayaraa et al., 2022).

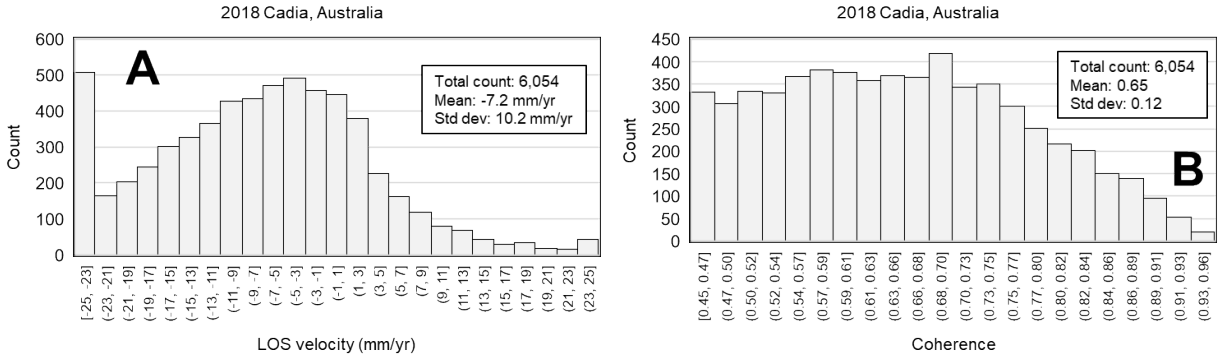

Figure 4-14: Histograms of [A] line-of-sight (LOS) velocities (mm/yr) and [B] coherence values of the Sentinel-1 PS-InSAR data, processed on SARScape Analytics, for the 2018 Cadia TSF breach in Australia.

4.4.2.4 2018 Cieneguita, Mexico

The Cieneguita gold-silver tailings dam in Chihuahua, Mexico was breached catastrophically on 4 June 2018 (Figure 4-15). The total released volume was ~440,000 m^3, including ~250,000 m^3 of tailings and ~190,000 m^3 of embankment and construction materials (Rana et al., 2021, 2022). According to historical satellite imagery on Google Earth, the tailings deposition commenced at this site in mid-2013. We constructed a time-lapse video of the Cieneguita TSF using 60 PlanetScope images spanning the

period January 2017 to July 2018 (Table A.1). The video shows that the TSF was undergoing high rates of depositional and construction activity.

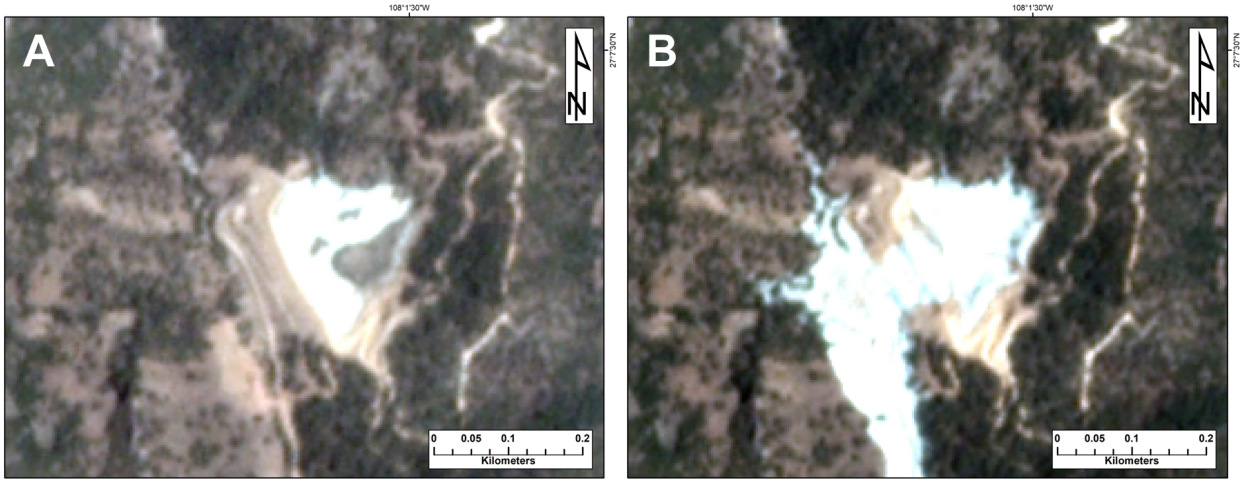

Figure 4-15: [A] Pre-failure (3 June 2018) and [B] post-failure (5 June 2018) PlanetScope (3 m resolution) images of the 4 June 2018 Cieneguita tailings dam breach in Mexico.

The size of the TSF was very small, covering an area of ~35,000 m^2 with a perimeter of ~750 m and a dam crest length of about 150 m around the date of failure. The natural ground slope at the base of the TSF dips ~7° towards Canitas Creek. Post-failure ground photographs show that the gold-silver tailings were impounded in a bowl-like depression behind a dike made of compacted alluvial sands. According to local reports, premonitory signs included extensive cracking along the downstream face of the dam about 4 months prior to failure, likely indicative of ongoing internal erosion and a weakened state of the embankment in response to rapid loading on the sloping impoundment (Rana et al., 2021, 2022). The breached portion of the dam extended over half of the total crest length of ~170 m. Site photographs showed that the breached portion of the dam disintegrated after falling into the creek bed. At least three mine workers were killed by the collapse. The tailings flow achieved a runout distance of 15 km along Canitas Creek while inundating an area of 500,000 m^2 (Ghahramani et al., 2020).

Figure 4-16 shows the PS results from the processing of 40 Sentinel-1 images over a 16-month duration. The satellite's orbit track was descending with a LOS incidence angle of 33.5º. The geometric relationship between the satellite and the dam was not ideal as the downstream slope of the dam was

facing away from the satellite's LOS and was marginally exposed to the satellite's view. The displacements in the dam towards the downstream direction were along-track of the satellite and thus were poorly captured in the results.

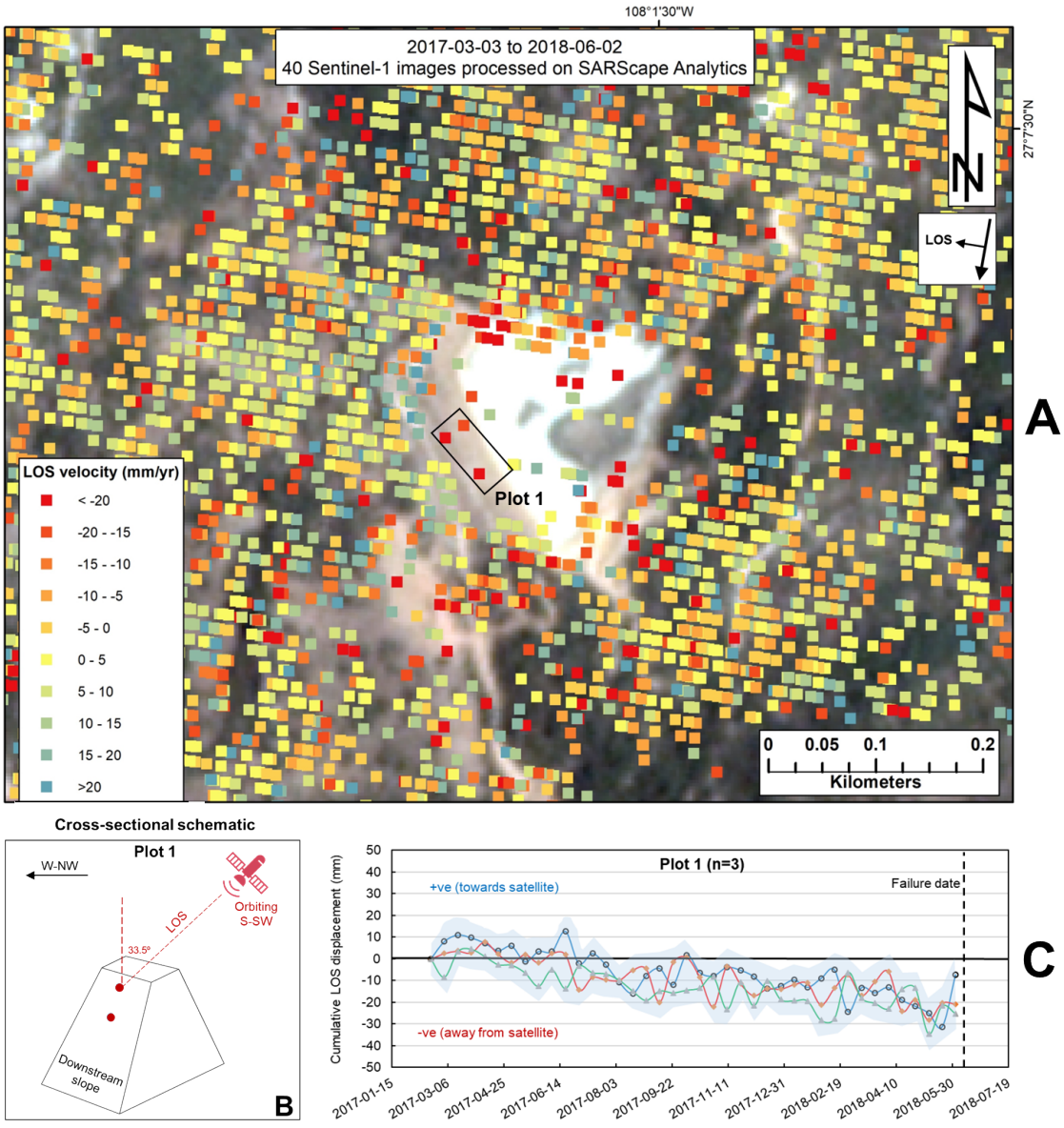

Figure 4-16: Sentinel-1 PS-InSAR results, processed on SARScape Analytics, for the 4 June 2018 Cieneguita tailings dam failure in Mexico. [A] Line-of-sight (LOS) velocity map, annotated with the Plot 1 area selected for graphical analysis. Negative (red) values indicate detected movements away from the satellite, positive (blue) values indicate detected movements towards the satellite, and green-yellow values indicate detected

stable areas. [B-C] Inset images of the active TSF, showing the state of the TSF at the beginning of the time-series (February 2017) and the post-failure view of the TSF (June 2018). [D] Cross-sectional schematic for the Plot 1 area illustrating the geometric relationship between the satellite, the tailings dam, and the PS points selected for graphical analysis. [F-G] Cumulative time-series LOS displacements for the Plot 1 area. Error bar shading represents an estimated average standard deviation of ±7.2 mm for incremental LOS displacement measurements.

We observe a low density of PS points in the TSF, which is due to the high rate of construction activity over the study period. The PS results in the surrounding forested vicinity also exhibit strong scatter. For graphical analysis, we selected 3 PS points on the unstable section of the dam (Plot 1 area). We observe that the cumulative LOS displacements reached ~30 mm in the Plot 1 area, but the results exhibit considerable variability and noise.

To address this issue, the proprietary SqueeSAR PS-DS software, known to produce robust, high-coherence results even in vegetated terrains (Ferretti et al., 2011), was used to process 39 Sentinel-1 images encompassing the same study period (Figure 4-17). The SqueeSAR results still show a low density of points in the TSF, but with significantly reduced scatter and noise in the results, including in the surrounding forest. For analysis, we selected 4 points along the freeboard between the dam crest and the tailings (Plot 1 area). Here, the cumulative LOS displacements ranged between 50 and 90 mm – much higher than the values produced using SARScape Analytics. However, the SqueeSAR results do not detect any signs of accelerating precursor LOS movements. In fact, in the 10 months preceding failure, the dam appeared relatively stable compared to the first 6 months of the study period.

The histograms in Figure 4-18 show that the mean coherence was 0.72 in the SqueeSAR results compared to 0.50 for SARScape Analytics. The standard deviation in the LOS velocity results was high for both SqueeSAR (14.6 mm/yr) and SARScape Analytics (15.2 mm/yr), but the former appeared closer to a normal distribution. The extreme tail towards the negative in both datasets results from the concentration of away-from-sensor movements detected in the TSF.

It is clear that the environmental conditions and the intense construction activity affected the quality of the InSAR results for this case study. In addition, we infer that the size of the TSF and the medium resolution of the C-band Sentinel-1 data (5 x 20 m) were also influential factors, as each pixel covered a significant portion of the small dam. The use of high-resolution TerraSAR-X (1-3 m) data would have likely resulted in a higher density of points in the TSF with a lower standard deviation (e.g. Gama et al., 2022).

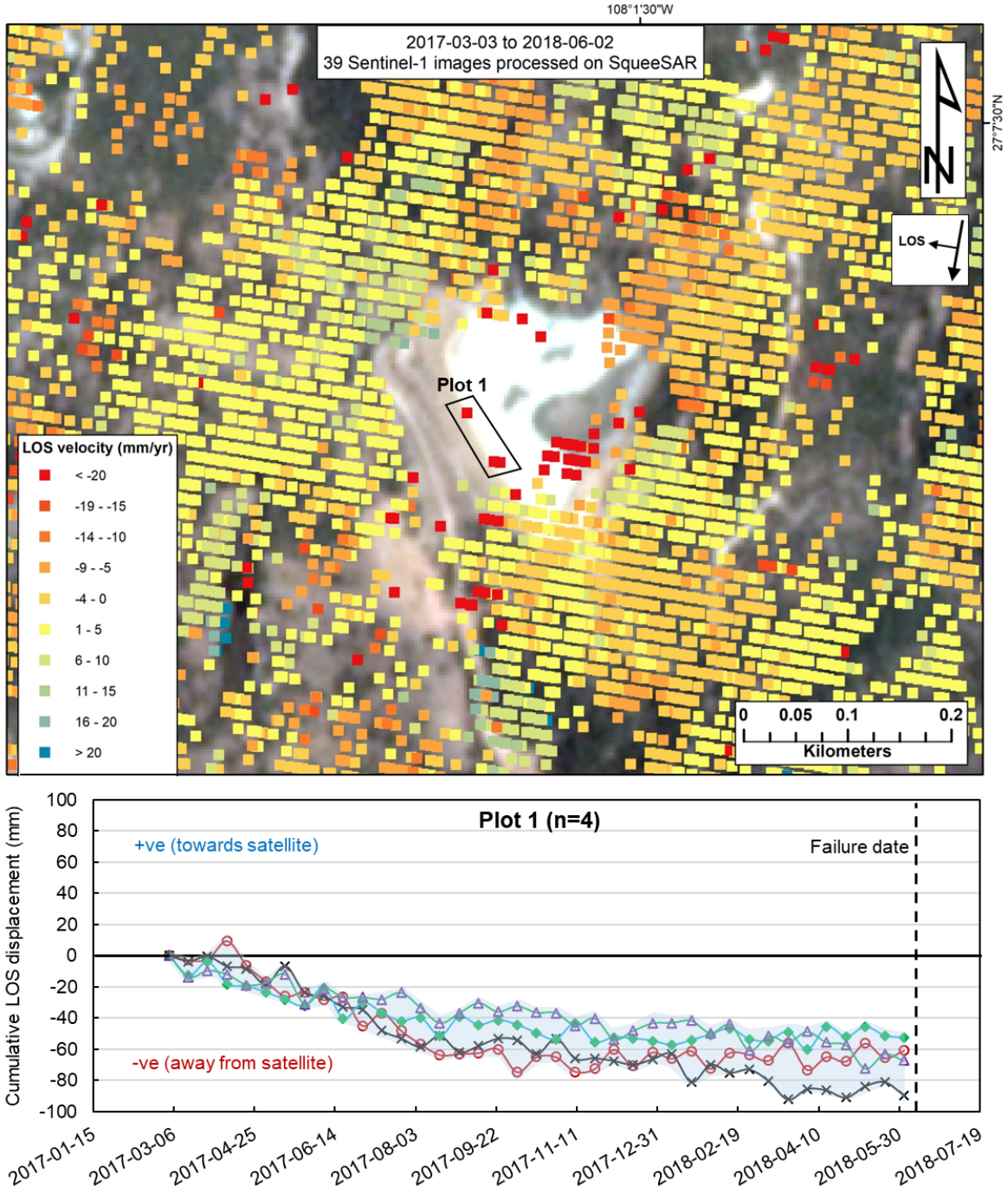

Figure 4-17: Sentinel-1 PS+DS InSAR results, processed on SqueeSAR, for the 4 June 2018 Cieneguita TSF failure in Mexico. Error bar shading represents an estimated standard deviation of ±2.7 mm for each measurement. Negative (red) values indicate detected movements away from the satellite, positive (blue) values indicate detected movements towards the satellite, and green-yellow values indicate detected stable areas.

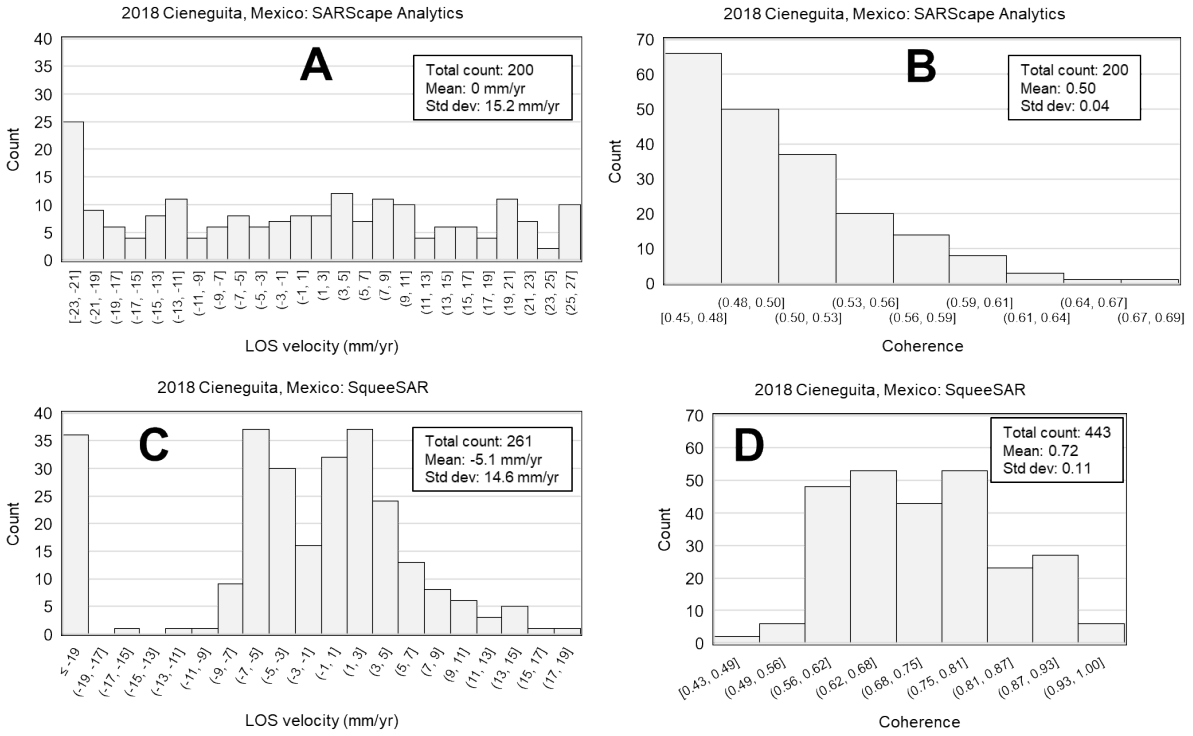

Figure 4-18: Histogram distributions of LOS velocities (mm/yr) and coherence values to show the quality of the Sentinel-1 InSAR data, processed on SARScape Analytics [A-B] and SqueeSAR [C-D], for the 2018 Cieneguita TSF failure in Mexico.

4.4.2.5 2019 Feijao, Brazil

The Feijao TSF is located near Brumadinho, Brazil. The dam was raised upstream over the course of 37 years, ultimately reaching a height of 87 m (Robertson et al., 2019). The dam collapsed on 25 January 2019, releasing 9.7 M m3 of tailings, equivalent to 75% of the pre-failure impounded volume (Figure 4-19). The failure was predisposed by several factors, as reported by the Independent Expert Panel (Robertson et al., 2019): i) the application of the upstream raise method with a steep slope; ii) the deposition of fine, weak tailings near the crest of the dam due to a lack of water management; iii) a setback in construction, which caused the upper portions of the dam to overlie weaker, finer-grained tailings; iv) the lack of effective horizontal drainage, groundwater seepage, and high rainfall that led to high internal water levels; v) a loss of suction in the unsaturated portion of the tailings, leading to a sudden loss of strength; and vi) high iron content in the tailings resulting in particle bonding via iron

oxidation, causing brittle behaviour in the tailings. These issues preconditioned the occurrence of static liquefaction on the date of failure, likely triggered by drilling activity on a metastable section of the dam (Arroyo and Gens, 2021). The collapse produced a flow that travelled 9 km downstream before entering and polluting the Paraopeba River (de Lima et al., 2020; Rana et al., 2021). The disaster resulted in 272 deaths and rendered long-lasting environmental and socio-economic effects in the region, prompting major updates to global industry standards in tailings management (ICMM et al., 2020). We constructed a time-lapse video of the Feijao TSF using 90 PlanetScope satellite images between June 2017 and February 2019 (Table A.1), which confirm that the TSF was in an inactive state during this period.

A number of publications have conducted InSAR investigations of the Feijao event (Gama et al., 2020; Holden et al., 2020; Grebby et al., 2021). The differences in data interpretation and conclusions between these studies are summarized as follows:

- Gama et al. (2020) used the SBAS and PS algorithms in SARScape (the complete software, not the automated and limited Analytics package that is used in this study) to process 26 Sentinel-1 images and detected a mild acceleration phase in the weeks preceding the collapse. The authors concluded that their confidence in their inverse-velocity prediction results was low due to the wide error distributions that were not centered around the actual failure date.
- Holden et al. (2020) from 3vGeomatics Inc. used proprietary software to analyze Sentinel-1 (>3 years of images over two orbit tracks), TerraSAR-X (~2 years of images), and COSMO-SkyMed data (30 images) and concluded that the precursor acceleration was not statistically significant or anomalous enough to have been a reliable warning sign. They used the findings of Robertson et al. (2019), who noted “no apparent signs of distress prior to failure”, as a geotechnical justification for this conclusion.
- Grebby et al. (2021) processed 45 Sentinel-1 images from two satellite orbit tracks (both descending) using the ISBAS algorithm in the Punnet (now APSIS) software (Terra Motion Limited). Based on their inverse-velocity analysis of 4-5 data points that showed a prediction interval of ~40 days around the failure date, the authors concluded that the collapse was foreseeable.

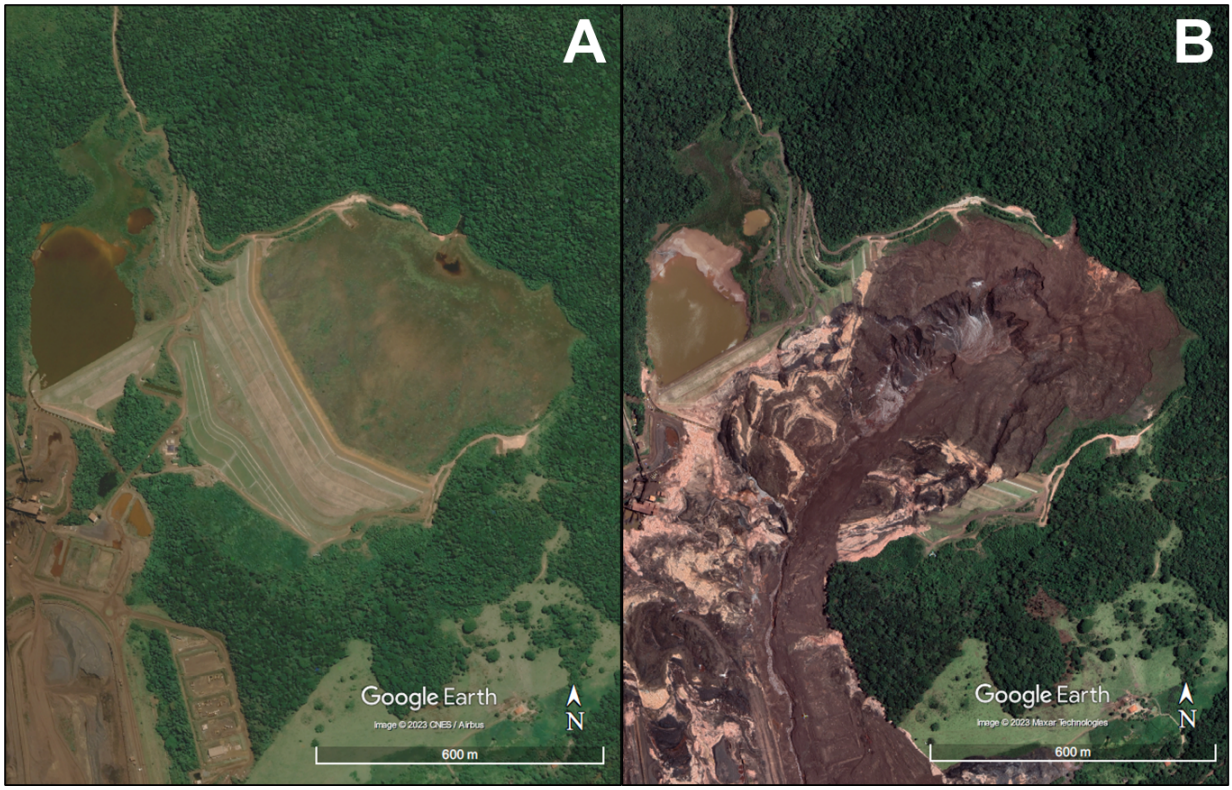

Figure 4-19: [A] Pre-failure (early January 2019) and [B] post-failure (early February 2019) Google Earth Worldview-2 (0.5 m resolution) images of the 25 January 2019 Feijao tailings dam breach in Brazil.

Figure 4-20 shows our PS results from the processing of 50 Sentinel-1 images over two orbit tracks (155 and 53, both descending) over a 19-month duration. Track 155 has a relatively high LOS incidence angle (45º) compared to Track 53 (32.5º) (Figure 4-20C). This implies that the PS results of Track 53 are more sensitive to vertical deformations compared to those of Track 155, which are more sensitive to sub-horizontal displacements. The dam face was exposed to the satellite's LOS at an oblique angle (Figure 4-20C); this makes it a non-ideal geometry to estimate the rate of movement in the downstream direction, which is almost parallel to the orbit track.

The results for both tracks clearly show that the impounded tailings were deforming away from the sensor, whereas LOS displacements recorded in the dam were not as pronounced and showed greater variability (Figure 4-20A/B). These observations are consistent with the PS results shown in Gama et al. (2020). Figure 4-20D/E shows that the long-term LOS deformation results at the selected PS points

in the tailings are variable and likely subject to noise, and the cumulative LOS displacements were between 40 and 60 mm over the 19-month study period. Figure 4-20F shows that the average of these cumulative LOS displacements in the tailings approached 50 mm for both tracks. This compares to averages of 27 mm (SBAS) and 35 mm (PS) over a 10-month duration reported by Gama et al. (2020), an average of 20 mm (ISBAS) over 17 months reported by Grebby et al. (2021), and averages of 40 mm (Track 53) and 60 mm (Track 155) over 27 months reported by Holden et al. (2020). In Figure 4-20F, we observe a mild acceleration phase, particularly in Track 155, in the ~45 days preceding the collapse, but this trend does not visually appear to be anomalous enough to serve as a reliable, proactive warning sign, as argued by Holden et al. (2020).

Figure 4-21 shows that the LOS velocity distributions for both tracks were strongly skewed towards the negative with an extreme tail, which represents the abundance of PS points in the tailings indicating movements away from the sensor. The standard deviation of the LOS velocity results was 11.9-12.2 mm/yr. The mean coherence was relatively low at 0.53-0.54, which reflects the effects of environmental factors, specifically the forested area surrounding the TSF, on the quality of the PS results.

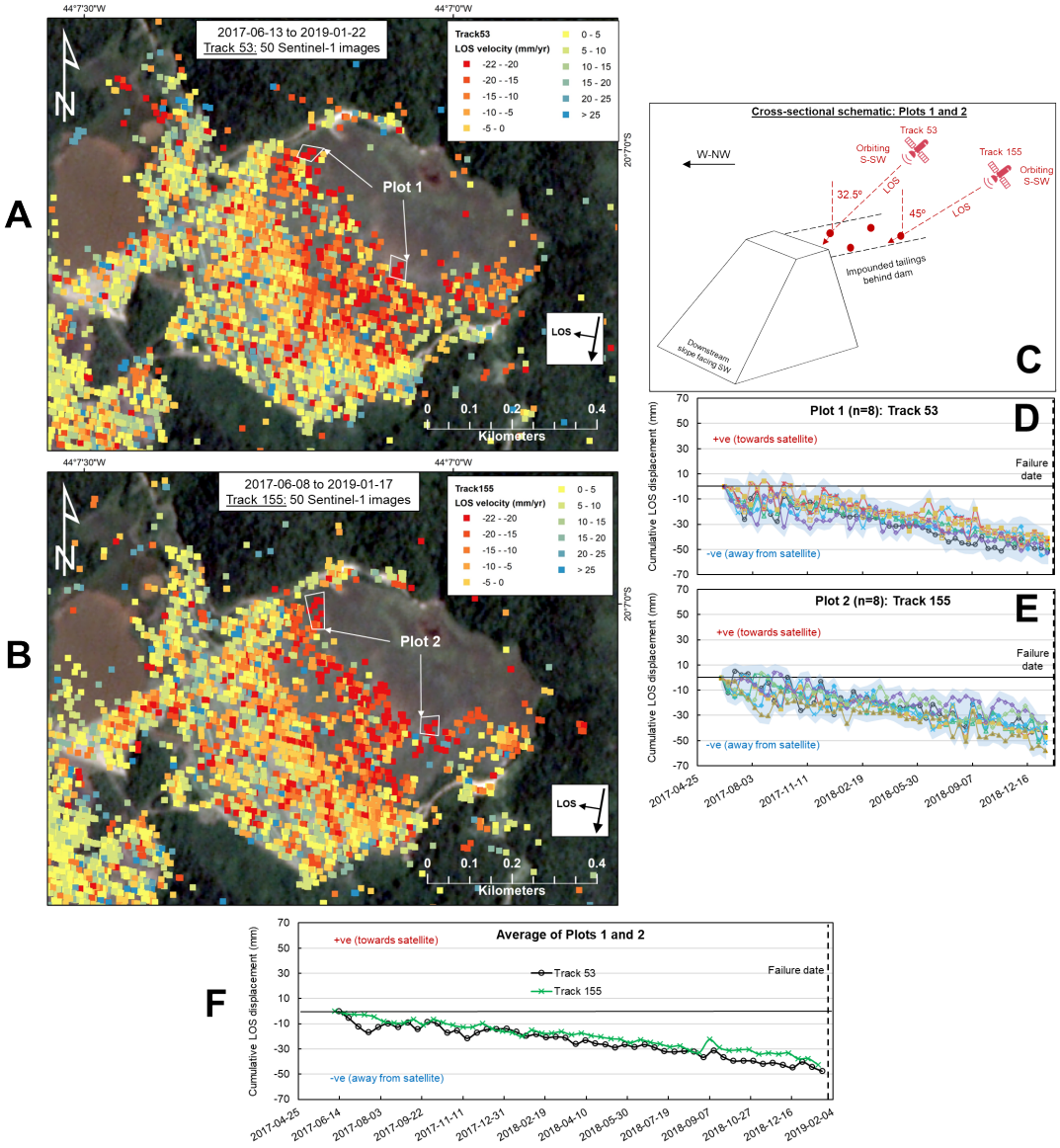

Figure 4-20: Sentinel-1 PS-InSAR results, processed on SARScape Analytics, for 2 satellite orbit tracks over the site of the 25 January 2019 Feijao tailings dam failure in Brazil. [A-B] Line-of-sight (LOS) velocity map for Track 53 and Track 155, annotated with the Plot 1 and Plot 2 areas selected for graphical analysis. Negative (red) values indicate detected movements away from the satellite, positive (blue) values indicate detected movements towards the satellite, and green-yellow values indicate detected stable areas. [C] Cross-sectional schematic for the Plot 1 and Plot 2 areas illustrating the geometric relationship between the satellite, the tailings dam, and the PS points selected for graphical analysis. The small red arrows indicate the direction of detected LOS displacement, in this case away from the satellite. [D-E] Graphical analysis of cumulative LOS displacements over the time-series duration for the Plot 1 (Track 53) and Plot 2 (Track 155) areas. Error bar shading represents an estimated average standard deviation of ±6.7 mm for incremental LOS displacement measurements. [F] Average cumulative LOS displacements from Plot 1 and Plot 2.

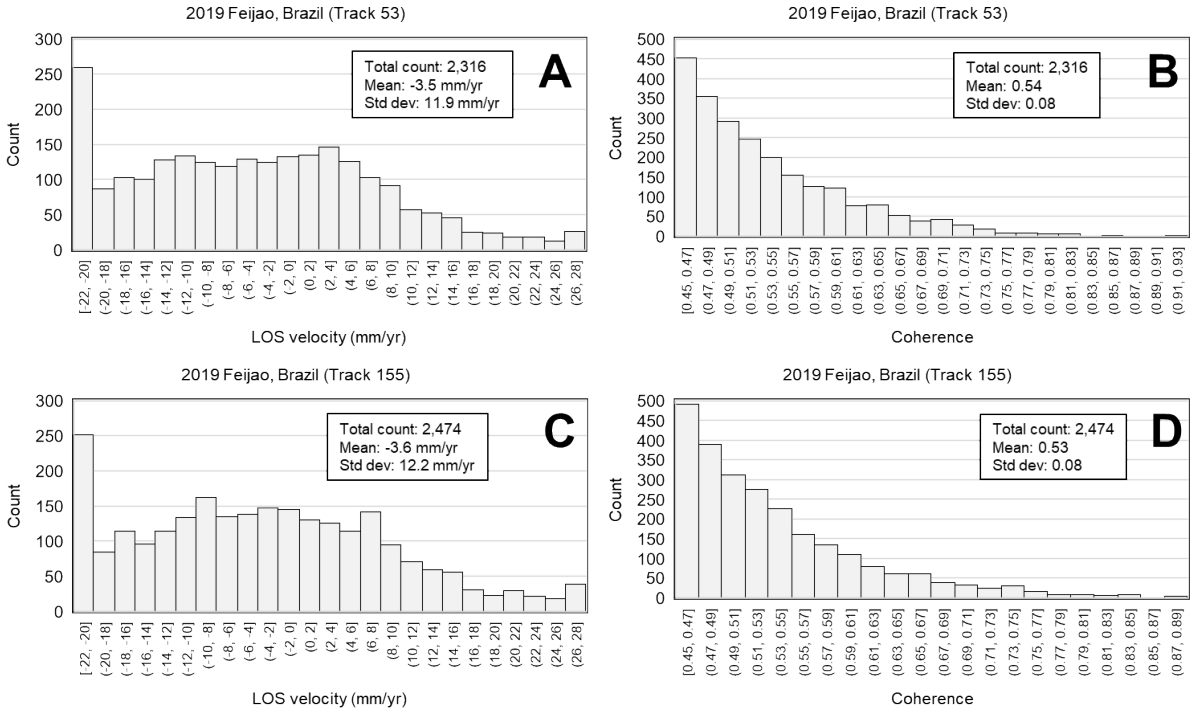

Figure 4-21: Histograms of LOS velocities and coherence of the Sentinel-1 PS-InSAR data, processed on SARScape Analytics, for the 2019 Feijao TSF failure in Brazil. [A-B] Track 53. [C-D] Track 155.

4.4.2.6 2019 Hindalco, India

Hindalco Industries operates a bauxite residue TSF near the village of Muri in Jharkhand, India. The TSF covers a surface area of ~300,000 m^2 and was dammed by ~5 m high gabion retaining walls with a perimeter of over 2 km. Historical satellite images on Google Earth show that a water pond covered the SW portion of the impoundment for several years until sometime in the period 2011-2014, when tailings were deposited into the pond. There were also two water storage ponds as distinct compartments in the TSF.

On 9 April 2019, about 600 m of the SW section of the gabion wall was breached (Figure 4-22). The failed materials, covering an area of ~100,000 m^2, comprised the entire extent of the former pond area, as well as one of the water storage ponds. The flowslide was bounded by railway tracks towards the west and travelled for a few hundred meters southward along the margin of the tracks towards the village. The flowslide headscarp was relatively steep and approximately 600 m long. The trigger

mechanism of the event is unclear, but local reports have pointed to the poorly constructed gabion wall as a failure cause, and to a potential undrained failure mechanism based on field observations, eyewitness accounts, and analysis of a publicly available video of the breach area (Kumar, 2019; Rana et al., 2021). We created and posted a time-lapse video of the Hindalco TSF using 120 PlanetScope images captured over the period September 2017 to September 2019 (Table A.1).

Figure 4-22: [A] Pre-failure (9 April 2019) and [B] post-failure (9 May 2019) satellite images of the 9 April 2019 Hindalco tailings dam breach in India. Image A is PlanetScope (3 m resolution) and image B is Google Earth Worldview-2 (0.5 m resolution).

Figure 4-23 shows the PS-InSAR results for the Hindalco case. We processed 52 Sentinel-1 images spanning 18.5 months. The satellite was on a descending orbit track with a LOS angle of 38.2º. Although the downstream face of the western gabion wall was facing away and partially hidden from the satellite's LOS, the crest was fully exposed (Figure 4-23C), which allowed the retrieval of LOS deformations in the downstream (westward) direction, parallel to the LOS direction.

We observe a concentration of higher-velocity movements (up to 24 mm/yr) away from the sensor (i.e. in the downstream direction) in the tailings abutting the dam at the northern edge of the breach section, which may represent the area of initiation (marked as "Plot 1"). The maximum LOS displacement recorded at this location was 50 mm over the study period (Figure 4-23C). However, we also observe a similar concentration towards the north in the tailings deposited close to the interior edge of the gabion wall (marked as "Plot 2"), where the peak LOS displacement approached 60 mm (Figure 4-23D). Figure 4-23E illustrates that the mean LOS displacement behaviour was near-identical at both locations with four notable stages of deformation: (i) stable behaviour between September 2017 and February 2018; (ii) increase in velocity between February and June 2018 (equivalent to ~50 mm/yr in this period); (iii) stable behaviour between June and December 2018; and (iv) increase in velocity between January and March 2019 (equivalent to ~50 mm/yr in this period). This raises the question of why the the breach occurred in the area of Plot 1, whereas the area of Plot 2 remained stable. Without greater knowledge of site-specific conditions, answering this question with validity remains challenging despite the availability of InSAR-derived observations.

Figure 4-24 shows histogram distributions of the LOS velocities and the coherence values for the PS points spanning the extent of the TSF. The LOS velocity results were quasi-normally distributed but with a skew to the negative and an extreme tail due to the higher frequency of negative-value (movement away from sensor) observations. The standard deviation of the LOS velocity results was 11.2 mm/yr. The mean coherence was 0.56, and high-coherence observations were generally scarce, which we primarily attribute to the abundance of vegetation in the vicinity of the TSF.

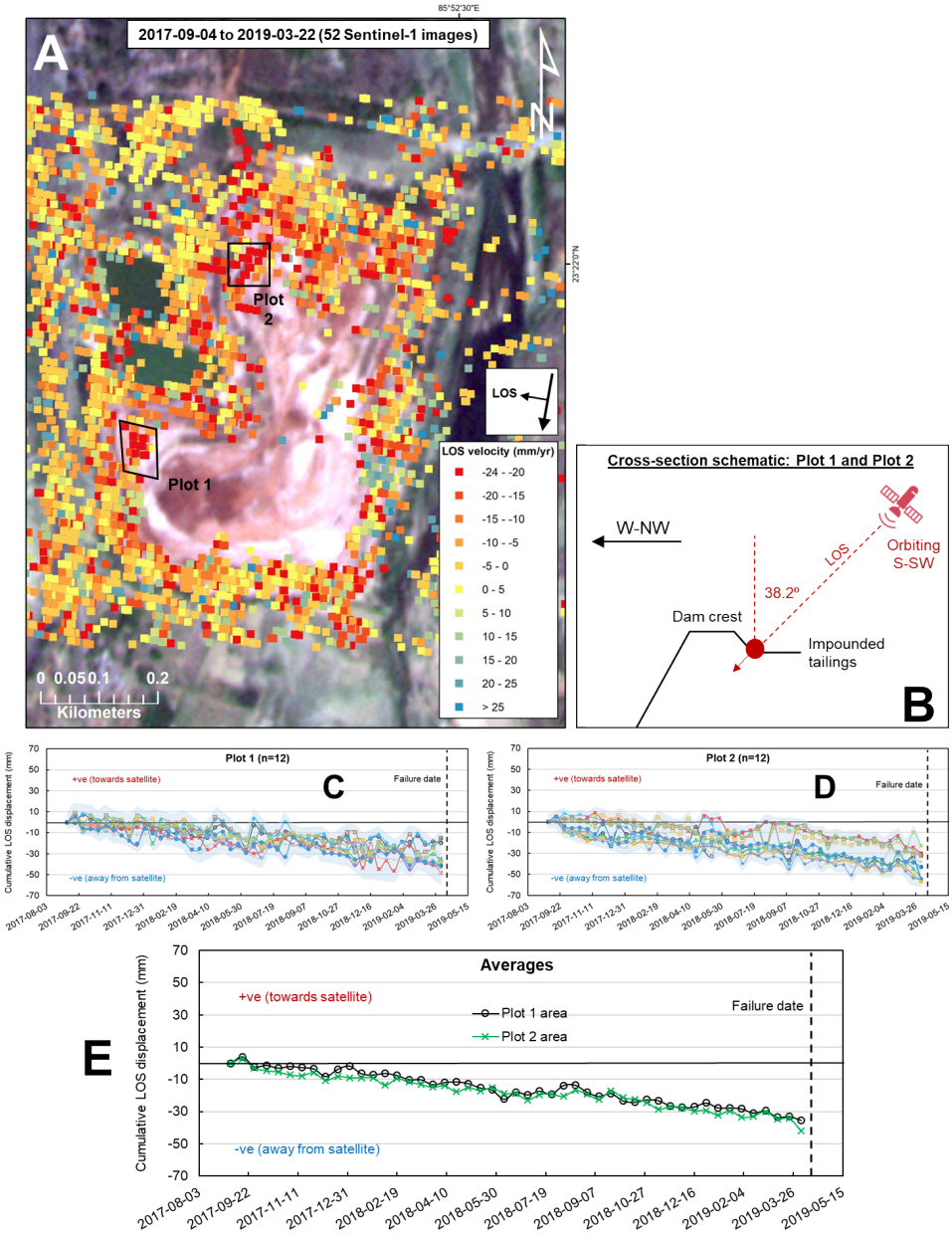

Figure 4-23: Sentinel-1 PS-InSAR results, processed on SARScape Analytics, for the 9 April 2019 Hindalco tailings dam failure in India. [A] Line-of-sight (LOS) velocity map, annotated with the Plots 1 and 2 areas selected for graphical analysis. [B] Cross-sectional schematics for the Plots 1 and 2 areas illustrating the geometric relationship between the satellite, the tailings dam, and the PS points selected for graphical analysis. The small red arrow indicate the direction of detected LOS movement, which in this case is away from the satellite. [C-D] Cumulative LOS displacements over the time-series duration for the Plots 1 and 2 areas. [E] Average cumulative LOS displacements for the Plots 1 and 2 areas. Error bar shading represents an estimated average standard deviation of ±6.5 mm for incremental LOS displacement measurements.

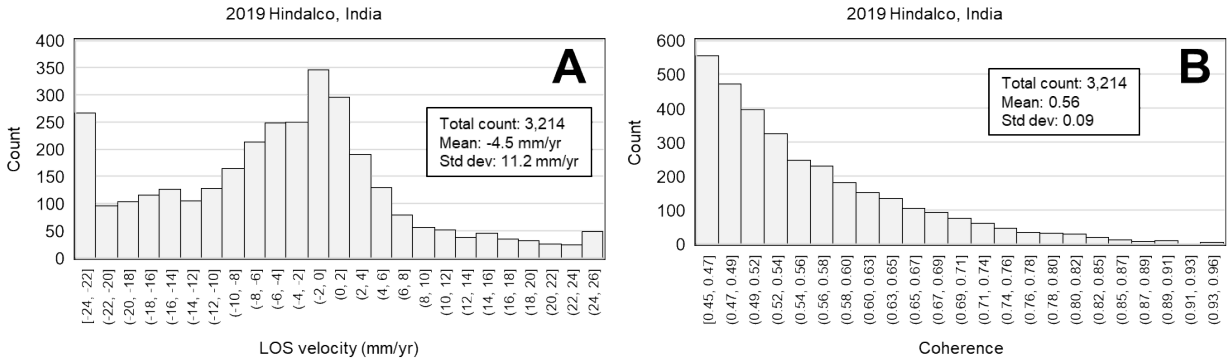

Figure 4-24: Histograms of [A] line-of-sight (LOS) velocities (mm/yr) and [B] coherence of the Sentinel-1 PS-InSAR data, processed on SARScape Analytics, for the 2019 Hindalco TSF failure in India.

4.4.2.7 2019 Cobriza, Peru

The Cobriza tailings dam in Huancavelica, Peru sits on a steeply dipping mountain valley in the high Andes (elevation 2,200 to 2,300 m above sea level). The mountain rock slope has a dip of almost 30° eastwards towards the Rio Mantaro. According to local reports, the foundation slope underneath the then-proposed dam was flattened to about 16° during construction of the mine infrastructure. Our review of historical satellite imagery on Google Earth suggests that the dam has been in place dating back to at least 2010. There is also another tailings dam at this mine that is located just 250 m south of the unstable dam.

The dam was breached on 10 July 2019, releasing toxic copper tailings into the worker camp and subsequently into the Rio Mantaro (Figure 4-25). The dam was breached at its southern corner at an elevation of ~2,240 m ASL, releasing about 67,500 m^3 of copper tailings and ponded water through a ~50 m wide breach channel (Rana et al., 2021). We created a time-lapse video of the Cobriza mine using 80 PlanetScope images over the period September 2017 to July 2019 (Table A.1). The video shows that, beginning early May 2019, there was discernible discoloration in the section of the dam crest that eventually breached a month later. This discoloration coincided with the rising pond water level that appeared to be close to the dam crest.

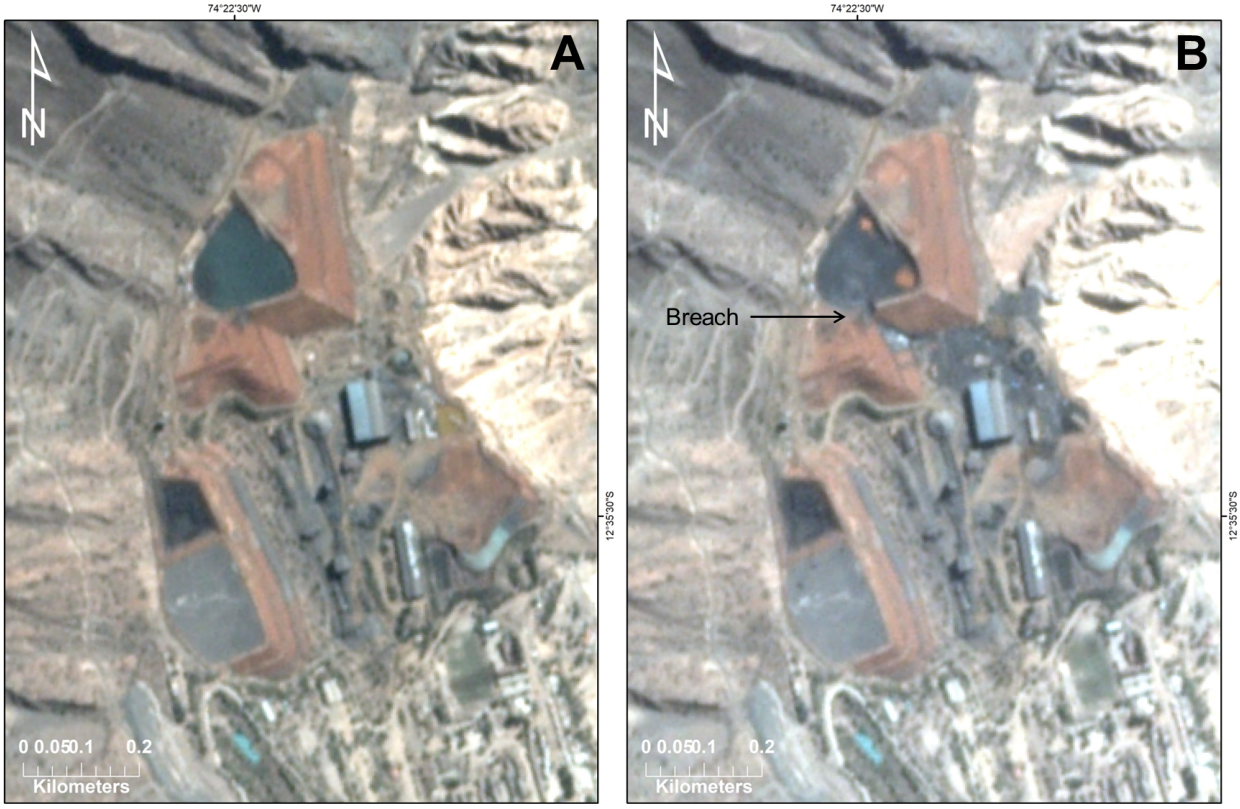

Figure 4-25: [A] Pre-failure (9 July 2019) and [B] post-failure (11 July 2019) PlanetScope (3 m resolution) images of the 10 July 2019 Cobriza tailings dam breach in Peru.

Figure 4-26 shows our PS results after processing 54 Sentinel-1 images over a 20-month duration. The PS results cover the Cobriza mine including the unstable tailings dam, the stable tailings dam towards the south, a waste rock pile towards the SE, and mine infrastructure. The satellite had a descending orbit with a relatively high LOS incidence angle of 40.9º. This implies that, compared to previous cases (except Feijao Track 155), the results derived from this satellite are more sensitive to sub-horizontal movements. The crest and southern and eastern slopes of the dam were exposed to the satellite's LOS (Figure 4-26D).

We observe concentrations of movements towards the satellite along the dam crest where the breach occurred (marked as Plot 1 area on Figure 4-26A). This section spatially overlaps the precursor discoloration observed on optical satellite imagery (Figure 4-26B/C). We note a similar hotspot in a

section along the eastern slope of the TSF. In comparison, the other tailings dam to the south showed stable behaviour with no anomalies in deformation patterns. The only other site in the mine that was experiencing high rates of movement was the waste rock pile just SE of the unstable Cobriza dam.

The cumulative LOS displacements recorded along the unstable section of the dam crest were ~50 mm over the study duration (Figure 4-26E). The general patterns indicate a near-constant rate of movements in the downstream direction. Beginning mid-May 2019, we observe a month-long period of accelerated movements; however, in the final 2 weeks preceding the breach, the apparent direction of movement changes, where we now observe away-from-satellite displacements in 4 out of 5 PS points that are plotted. This appears counterintuitive to our expectations, but without access to site-specific data or to proprietary-software InSAR results over the site (e.g. SqueeSAR), it remains challenging to confirm whether this is a true representation or is a symptom of phase unwrapping, similar to Cadia.

Figure 4-27 shows that the LOS velocity results were normally distributed around ~0 mm/yr with a standard deviation of 8.8 mm/yr. The coherence values were higher on average with a mean of 0.64, which we attribute to the ideal environmental conditions (dry, hard surfaces without vegetation) at the mine site.

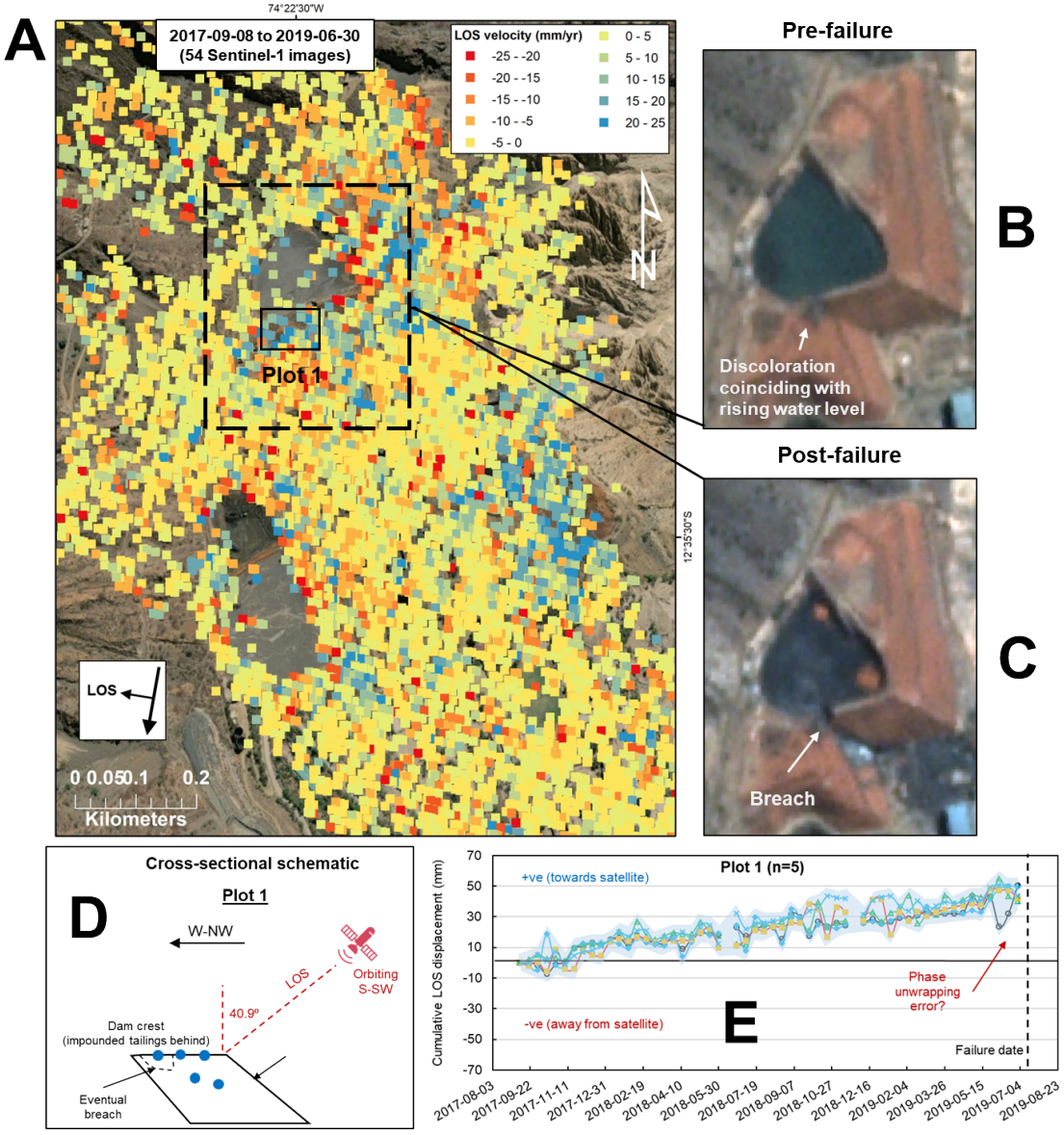

Figure 4-26: Sentinel-1 PS-InSAR results, processed on SARScape Analytics, for the 10 July 2019 Cobriza tailings dam failure in Peru. [A] Line-of-sight (LOS) velocity map, annotated with the Plot 1 area selected for graphical analysis. Negative (red) values indicate detected movements away from the satellite, positive (blue) values indicate detected movements towards the satellite, and green-yellow values indicate detected stable areas. [B] Inset pre-failure and [C] post-failure PlanetScope images of the TSF. [D] Cross-sectional schematics for the Plot 1 area illustrating the geometric relationship between the satellite, the tailings dam, and the PS points selected for graphical analysis. The small red arrows indicate the direction of detected LOS movement, which in this case is towards the satellite. [E] Cumulative LOS displacements over the time-series duration for the Plot 1 area. Error bar shading represents an estimated average standard deviation of ±5.4 mm for incremental LOS displacement measurements.

Figure 4-27 illustrates that the LOS velocity results were normally distributed around ~0 mm/yr with a standard deviation of 8.8 mm/yr. The coherence values were higher on average with a mean of 0.64, which we attribute to the ideal environmental conditions (dry, hard surfaces without vegetation) at the mine site.

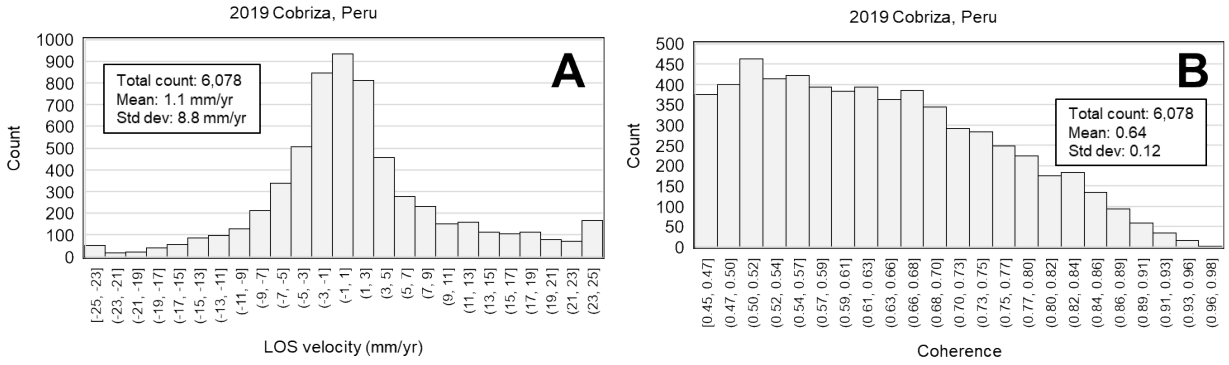

Figure 4-27: Histograms of [A] LOS velocities (mm/yr) and [B] coherence of the Sentinel-1 PS-InSAR data, processed on SARScape Analytics, for the 2019 Cobriza tailings dam failure in Peru.

4.5 Discussion

The main outcomes of each case study are summarized in **Error! Reference source not found.**. Our investigations produce key insights into how the reliability of Sentinel-1 InSAR results for a tailings dam may be influenced by several variables, listed in the following sub-sections, that have practical implications for data interpretation.

Table 4-5: Summary of main outcomes from the ground-truth assessment and each forensic case study.

Case study	**Summarized outcome**
Ground-truth test site	InSAR observations using SARScape Analytics and SqueeSAR were generally comparable to *in-situ* monitoring prism data, though SqueeSAR exhibited some deviations from the general trend. SqueeSAR's higher threshold for filtering out data points resulted in a lower density of points along the dam crest. However, SqueeSAR exhibits stronger proficiency in reducing noise and increasing coherence compared to SARScape Analytics, which implies higher reliability.
2017 Tonglvshan, China	Breach section showed high displacements, but so did another location on the dam. However, the unstable dam face was not exposed to the satellite's view, which resulted in poor capture of movements towards the failure direction.
2017 Mishor Rotem, Israel	Clear observations of high LOS displacements in the breach section, but the non-ideal oblique angle between the satellite's LOS and the dam face resulted in incomplete understanding of the pre-failure deformation behaviour.
2018 Cadia, Australia	In the 3 months preceding the breach, rapid acceleration was not detected, which is inconsistent with other studies of this case. Phase unwrapping errors were clearly detected due to exceedance of the LOS velocity threshold in SARScape Analytics, whereas this was not an issue in the proprietary software used in other studies.
2018 Cieneguita, Mexico	Significant noise in the SARScape Analytics results due to forested terrain. SqueeSAR results had significantly less noise and suggested higher LOS displacements, but provided no sign of precursor acceleration; in fact, the rate of movement was relatively low in the months preceding failure. The TSF size was small, thus the high-resolution alternative of TerraSAR-X data would be convenient here to offer a higher density of InSAR points.
2019 Feijao, Brazil	Results show reasonable consistency with previous studies. Mild acceleration detected in the ~45 days preceding failure, but does not appear to be anomalous enough to serve as a reliable warning sign in proactive scenarios.
2019 Hindalco, India	Breach section showed high displacements, but so did another location. Without background knowledge of the site conditions, it remains unclear why one section failed whereas the other did not.
2019 Cobriza, Peru	Breach section showed clear patterns of high LOS displacements, especially compared to the nearby stable tailings facility. Results in the 2 weeks preceding failure exhibited especially noisy behaviour (similar to Cadia) – perhaps an indication of phase unwrapping error.

4.5.1 Influential variables

4.5.1.1 Environmental conditions

Vegetated terrains reduce coherence and exacerbate noise in PS results. This is evident in the cases of Cieneguita and, to some extent, Feijao and Hindalco, despite the processing stack containing a sufficient number of images. This limitation is overcome when using the PS-DS technique and advanced noise-filtering algorithm implemented in SqueeSAR, as shown in our Cieneguita case study. Snow/ice cover also adversely affects the quality of InSAR results (Carla et al., 2019b) and therefore was a key factor in our ground-truth assessment, in which we removed winter images from the processing stack. This has implications for the continuous satellite-based monitoring of TSFs in cold-climate regimes. In contrast, the quality of PS data was observed to be consistently robust in dry, bare-surface environments such as deserts (e.g. Mishor Rotem) or high-mountain bedrock (e.g. Cobriza).

4.5.1.2 Orientation of satellite in relation to TSF geometry

The satellite orbit direction and the LOS angle have an important influence on the magnitude of InSAR-derived displacement results and the subsequent data interpretation. This effect is particularly applicable to TSFs with multiple dams of variable orientations. This effect is also generally universal irrespective of the selected processing software, as SAR satellites are on polar orbits with right-looking orientation. The Mishor Rotem case study is an example where, despite the higher-quality PS results, the obliqueness of the unstable dam face in relation to the satellite's view resulted in incomplete understanding of the pre-failure deformation behaviour. However, LOS components of vertical deformation on the dam crest can still be well-detected irrespective of the satellite orbit direction, as observed in the Cadia case. The tracking of vertical versus horizontal deformation (when using only a single satellite rather than multiple overlapping satellites) is sensitive to the satellite's LOS angle, whereby a smaller angle indicates stronger sensitivity to vertical movements.

4.5.1.3 Size of TSF versus selection of satellite

When monitoring a smaller-sized TSF (e.g. Cieneguita), the resolution of the open-access Sentinel-1 data (20 x 5 m) may be too coarse, as a single pixel may cover a significant part of the tailings dam being studied. A potential solution to this is the TerraSAR-X satellite for which the spatial resolution can be 1-3 m depending on the imaging mode. A limitation of our study is that a comparison between

Sentinel-1 and TerraSAR-X could not be presented because the TerraSAR-X data archive does not offer the spatial-temporal coverage as that of Sentinel-1, as well as due to the high cost per TerraSAR-X scene. To our knowledge, only Holden et al. (2020) have conducted a comparison between the two satellites for a TSF failure (Feijao), whereas Gama et al. (2022) and a few studies in other sectors (e.g. Bischoff et al., 2017; Colombo, 2021; Wang et al., 2021) have commented on the higher point density and lower standard deviation offered by TerraSAR-X.

The positional accuracy of observation points also varies depending on the satellite. According to general insights from SqueeSAR case studies, the approximate point-elevation accuracy is ±1.5 m for TerraSAR-X compared to ±8 m for Sentinel-1. Furthermore, the approximate north-south and east-west point-location accuracy is ±1 m and ±3 m, respectively, for TerraSAR-X, compared to ±8 m and ±12 m for Sentinel-1. These differences are particularly important considerations when monitoring smaller-sized tailings dams.

4.5.1.4 Selection of processing software

Each InSAR software is founded on algorithms that filter and convert raw radar satellite data into point-cloud displacement data to allow for interpretation by practitioners. The strengths and limitations of these algorithms differ according to each software. A demonstration of these differences was presented using the ground-truth test site and the Cieneguita case. The commercial SARScape Analytics software allows automated data processing, which enables faster run-times and streamlines multi-site or regional-scale assessments. Yet, there are conspicuous errors when encountering heavily vegetated terrains (e.g. Cieneguita) and when processing large, accelerating displacements (e.g. phase unwrapping error with Cadia). In fact, in all of the cases examined herein, the recorded LOS velocity was always below 50 mm/yr, which appears to represent the approximate detection threshold for Sentinel-1 processing on SARScape Analytics. In comparison, SqueeSAR can detect displacement rates of up to 1,000 mm/yr, whereas the maximum detectable threshold when processing Sentinel-1 data on TerraMotion's ISBAS algorithm is 426 mm/yr (Bayaraa et al., 2022).

The automated approach in SARScape Analytics also prevents the user from manually assigning or locating the reference point – an important parameter in a standard InSAR analysis. Based on our literature review of InSAR case studies, some of these limitations may be overcome in advanced commercial software (e.g. the complete SARScape package used in Gama et al., 2020), whereas all of

these limitations are overcome in proprietary software (e.g. ground-truth test site and Cieneguita in this study; Cadia in Carla et al., 2019a and Bayaraa et al., 2022). The capabilities and limitations vary even between the various company-specific proprietary software, but to our knowledge, a practical case-study comparison for TSFs is yet to be demonstrated in scientific literature.

The positional accuracy of observation points is another parameter that can depend on the processing software. In satellite InSAR, although the precision of displacement measurements is millimetric, the position of observation points is known with a meter-scale accuracy. This can have implications for interpreting whether the InSAR observations correspond to the crest, slope or toe of a dam, particularly for smaller dams. In the ground-truth case, the average standard deviation of height (elevation) measurements for points in the TSF was 3.2 m when using SqueeSAR. When using SARScape Analytics, this value increased to 4.8 m. In Cieneguita, the average standard deviation of height measurements for points in the TSF was 7.3 m when using SqueeSAR, compared to 7.9 m when using SARScape Analytics. For the other forensic case studies processed using SARScape Analytics, these values ranged from 5.7 m (Feijao Track 53) to 7.9 m (Hindalco), with the exception of Feijao Track 155 which had an anomalously large value of 12.0 m, likely due to the high LOS incidence angle of 45º.

4.5.2 Implications for monitoring

4.5.2.1 Accuracy

We conducted a ground-truth assessment at a TSF test site using SARScape Analytics and SqueeSAR to demonstrate the accuracy of Sentinel-1 as a monitoring tool. With the support of in-situ data from monitoring prisms as a comparison benchmark, we observed that both software are able to represent low- to moderate-deformation regimes (displacement rates of 0-50 mm/yr) on a local scale with reasonable accuracy, though the SARScape Analytics results had a higher standard deviation (±6 mm) compared to SqueeSAR (±3 mm). SqueeSAR's threshold criteria for filtering out data points resulted in a lower density of points along the dam crest, but SqueeSAR exhibits stronger proficiency in reducing noise and increasing coherence compared to SARScape Analytics, which implies higher reliability.

Another example of this outcome is illustrated in the Cieneguita case. The PS results exhibited strong scatter due to the vegetated terrain and recorded cumulative LOS displacements of 30-40 mm, whereas

the PS-DS results resolved this issue and suggested that the cumulative LOS displacements were much higher (50-90 mm).

This study highlights the complementary role that satellite InSAR can play in long-term monitoring programs at TSF sites. InSAR can be especially effective in active mines containing multiple TSFs or a large TSF, or in legacy mines with abandoned TSFs, where installing and maintaining monitoring instrumentation can pose practical challenges. However, some precautionary notes are as follows:

- 1-D InSAR results often do not represent the maximum rate of movement that the dam is actually experiencing, nor does the LOS represent the true 3-D direction towards which the maximum rate of movement is occurring. This is where the availability of satellite data of overlapping orbits can be important in retrieving 2-D displacements (vertical and east-west horizontal), which was not possible for all cases analyzed in this study.
- Nevertheless, sub-horizontal movements along the north-south direction are generally poorly captured due to the polar orbits of SAR satellites.
- Conducting routine ground-truth exercises with the support of *in-situ* data provides value by confirming that the InSAR results are representative.

4.5.2.2 Prediction of instability (location and timing)

When attempting to predict TSF instability using InSAR, there are two components that require attention: breach location and breach timing. Based on our selected commercial software, our forensic case studies generally suggest that the breach location is easier to predict than the failure date. However, in the cases of Tonglvshan and Hindalco, there was at least one another hotspot of deformation other than the breach location. Given that these case studies were founded on relatively poor background knowledge (Table 4), explaining why the breach occurred where it did was challenging.

With respect to breach timing, the Feijao case is an interesting example where several studies (including the present) have conducted forensic InSAR investigations using different software and techniques and have obtained reasonably similar rates of precursor deformation, yet have arrived at different conclusions on whether the failure timing was foreseeable. This is because the precursor deformation patterns indicated minor accelerations which were subject to user-specific judgment, unlike the Cadia case that exhibited anomalous, exponential accelerations in the 3 months preceding

the breach. Ultimately, the benefit of hindsight is an important factor in how the pre-failure InSAR data has been perceived in some forensic case studies, including the ones presented here.

When using proprietary software such as SqueeSAR, prediction capabilities for both breach location and timing are significantly enhanced due to improved noise-filtering proficiency in diverse terrains (e.g. Cieneguita) and higher detection thresholds for annual LOS velocity rates (e.g. Cadia; Carla et al., 2019a; Bayaraa et al., 2022).

Not all TSF failures show clear geotechnical signs of impending collapse. Some TSF failures have reportedly occurred without advance warning, either due to brittle collapse of the tailings structure (e.g. Feijao; Robertson et al., 2019) or sudden anthropogenic disturbances or localized triggers (Rana et al., 2021). The main issue is that the temporal frequency of SAR satellite data (i.e. satellite revisit interval) remains too low for InSAR to be able to identify the specific triggering mechanism or capture the deformation behaviour in the hours preceding a breach. This underscores the value of keeping continuous, accessible records of in-situ data, and highlights why InSAR is a useful hazard-screening technology that can complement, but not substitute, on-the-ground observations.

4.6 Lessons learned

This study explored the capabilities and limitations of satellite InSAR to monitor the geotechnical stability of tailings dams. This research is timely considering the increased reliance on satellite technology in the mining sector and the need for more case study applications to enhance technical knowledge of InSAR. The goal of this study was to generate practical insights and considerations chiefly from an engineer's perspective. This was achieved by undertaking a ground-truth assessment and by conducting forensic analyses of 7 failure cases. The methodology involved the use of a commercial software with an automated PS workflow (SARScape Analytics) for all cases, with comparison to an advanced proprietary software (SqueeSAR) implemented with a PS-DS algorithm for the ground-truth test site and one forensic case study.

We summarize our conclusions as follows:

1. Commercial InSAR software with Sentinel-1 data provides reasonable accuracy when monitoring consolidation processes and lower-scale deformation (< 50 mm/yr) but is

ineffective in larger or accelerating deformation regimes (> 50 mm/yr), where advanced proprietary algorithms are more appropriate.

2. Environmental conditions strongly influence the quality of PS-InSAR. The best results are in dry, bare-earth or urban terrains, whereas vegetation, snow/ice cover, and free water worsen PS-InSAR data quality.

3. Sentinel-1 InSAR is a useful hazard-screening tool in mine sites with large TSFs or multiple TSFs, as it may help guide where to undertake targeted investigations. Sentinel-1 InSAR may also provide value in monitoring legacy mines with abandoned TSFs where installing and maintaining instrumentation may pose practical challenges.

4. Due to the polar orbit of SAR satellites, sub-horizontal movements in the north-south direction are generally poorly captured.

5. Not all geotechnical failures of TSFs show clear warning signs for weeks in advance. The revisit interval of SAR satellites prevents the detection of instantaneous failure modes. Therefore, long-term monitoring programs should ideally be integrated with a combination of remote sensing methods and field instrumentation.

Future avenues of research could involve case-study comparisons between medium-resolution Sentinel-1 and high-resolution TerraSAR-X data and ground-truth comparisons between different proprietary algorithms.

4.7 References

Given that this thesis is in part article-based, the content of Chapter 4 represents material published in the following citations:

Rana, N.M., Delaney, K., Evans, S.G., Deane, E., Small, A., Adria, D., McDougall, S., Ghahramani, N., Take, W.A. Analysis of Sentinel-1 PS-InSAR data using an automated commercial software to monitor tailings dams and predict instability. To be submitted to *Remote Sensing of Environment* in May 2023.

Chapter 5: Synthesis and Outlook

5.1 Introduction

This thesis presented a comprehensive body of research on tailings dam failures, some of which have resulted in loss of life, environmental contamination, and socioeconomic challenges. These occurrences have led to intense public scrutiny on safety standards in mining practice. This thesis studied these phenomena using advanced statistical and remote sensing techniques. The specific research focus was on quantifying the magnitude-frequency and failure rate statistics, geospatially documenting and analyzing breach-runout behavior, and mapping pre-failure deformation patterns using satellite InSAR. By this, the thesis successfully advances the state of knowledge on the background hazard-risk of tailings dams, the common causal and failure mechanisms for their failures, the potential downstream consequences, and the capabilities of satellite technology in monitoring and instability-prediction applications.

The main conclusions from each chapter are summarized in Sections 5.2-5.4. Section 5.5 summarizes the novel methods and scientific contributions in this thesis. Section 5.6 concludes this thesis by providing a forward-looking commentary on the progress of the tailings industry in achieving its goal of zero-harm.

5.2 Background hazard-risk of tailings dams

Chapter 2 statistically analyzed new datasets and estimates of the construction and failures of tailings dams (TSFs) and large water-retention dams (LRFs) worldwide in the period 1965–2020. This comparison served to contextualize the background hazard-risk of TSFs and address pre-existing misconceptions about their failure rates. Long-standing data gaps on LRF failures in China, and subsequently worldwide, and on constructed TSFs worldwide are addressed by estimation/extrapolation approaches to illustrate the range of uncertainty in the results.

The total number of LRF failures is estimated to have been between 394 and 608. The annual numbers of newly constructed and failed LRFs declined near-proportionally, thus the cumulative failure rate of LRFs stayed fairly constant, falling in the range of 1.2% to 1.8% as of end-of-2020. The rate drops to at least 0.7% when excluding China. The cumulative fatality rate of LRFs reduced over time

to 1.2 deaths per constructed facility, and falls in the range of 64 to 98 deaths per failure, as of end-of-2020. Failures of LRFs with very high storage capacities (>200 million m^3) have continued to occur since 2016.

In comparison, the annual number of TSF failures stayed relatively constant, whereas the annual construction rate of TSFs is estimated to have increased by ~3x, thus the cumulative failure rate of TSFs declined over time. When assuming the lower-estimate of the number of constructed TSFs (6,810), the cumulative failure rate is ~4.4% as of end-of-2020. When adopting the upper-estimate (20,230 TSFs), a rate of ~1.5% is obtained, which falls in the same order as the corresponding rate of LRFs. A review of published estimates of existing TSFs worldwide indicates that the "true" rate is much lower than 4.4% and closer to 1.5%. The cumulative fatality rates of TSF failures reduced over time to 0.1–0.3 death per constructed facility and 6 deaths per failure as of end-of-2020, which are lower than those of LRFs. However, the size and the environmental impact of TSF failures have increased on average worldwide, especially since 2014. The rising global rate of failed tailings volumes has been approximately proportional to the rising global rate of tailings production since the 1990s.

Heavy rainfall events and intensifying precipitation patterns are statistically important causative variables for the failures of both LRFs and TSFs. This has implications for the design and management of storage capacity, freeboard, facility drainage and spillways under climate change conditions. Our results are applicable broadly on a global scale and are conditioned by uncertainties in the data and the methods used to address data gaps. A more comprehensive public disclosure effort is necessary to improve the robustness of future statistical assessments, particularly with respect to reservoir facility failures in China and constructed TSFs worldwide.

5.3 Tailings flow characteristics

Chapter 3 narrowed the research focus to downstream mass flows from failed tailings dams (i.e. tailings flows). In particular, the chapter explored the site-specific variables that precondition the generation and control the behavior of catastrophic tailings flows following a breach. This work was supported by a new global database of 63 tailings flow cases remotely analyzed through a compilation of satellite imagery, digital elevation models and literature.

The chapter concludes that TSFs are anthropogenic sources of energy capable of producing mass flows when breaching initiates in unfavorable conditions – i.e. the occurrence of liquefaction and/or the incorporation of free water are sufficient conditions to trigger extremely rapid, highly mobile runout behavior. Magnitude-frequency statistics indicate that tailings flows with volumes of ≥ 1 M m^3 have occurred at a mean rate of 2–3 years since 1965. Weather hazards and drainage issues are identified as major causal variables for TSF failures that produce tailings flows. The behavior of a tailings flow depends on site-specific factors such as the geometry of the TSF, rheology of the flowing mass, nature of the runout path (i.e. substrate wetness, channelization, and slope), triggering mechanism and failure mode. Nevertheless, relative to unconfined terrains, channelized paths enhance the runout distance and inundation area – an effect that appears to be more pronounced along major rivers. Channelization and steeper bed slopes enhance flow velocities (peak of 25–30 m/s) which in turn exacerbate the downstream impact. Erosion is the dominant geomorphic mechanism during peak flow stages which indicates high mechanical energy. However, the entrainment capacities and solids concentrations (and conversely the interstitial and free water contents) of tailings flows remains quantitatively under-studied, suggesting a knowledge gap regarding the implications for tailings flow mobility and energy. Reductions in flow velocities are induced by gentler slopes and unconfined topography (peak of 6–16 m/s). Tailings flow velocities along major rivers steadily reduce to ≤ 5 m/s where the primary geomorphic mechanism transitions from erosion to deposition of coarser particles.

Although general trends may be observed, empirical efforts to predict tailings outflow-runout are prone to considerable uncertainty due to data scarcity or poorly conditioned available data for many of the selected cases and the complexity and variability of the site and tailings conditions that are inherently challenging to incorporate into simple regression models. Broad statistical approaches are thus appropriate for generating scholarly insights, identifying case analogs and performing screening-level assessments, while site-specific studies should rely on targeted investigations through numerical models, laboratory tests and field observations to ascertain potential breach-inundation behavior, ideally within a probabilistic framework.

5.4 Satellite InSAR as a monitoring tool for tailings dams

Chapter 4 uses open-source Sentinel-1 InSAR data and a cost-saving, automated Persistent Scatterer (PS) software to undertake a ground-truth assessment at a test site and to conduct forensic investigations of 7 tailings dam breaches (2017-2019). For the ground-truth site and one forensic case study, the results are compared to those developed through the proprietary SqueeSAR algorithm, which is integrated with a dual PS and Distributed Scatterers (DS) algorithm with advanced noise-filtering proficiency. The analysis generated key lessons of value to practitioners:

1. Commercial InSAR software with Sentinel-1 data provides reasonable accuracy when monitoring consolidation processes and lower-scale deformation (< 50 mm/yr) but is ineffective in larger or accelerating deformation regimes (> 50 mm/yr), where advanced proprietary algorithms are more appropriate.

2. Environmental conditions strongly influence the quality of PS-InSAR. The best results are in dry, bare-earth or urban terrains, whereas vegetation, snow/ice cover, and free water worsen PS-InSAR data quality.

3. Sentinel-1 InSAR is a useful hazard-screening tool in mine sites with large TSFs or multiple TSFs, as it may help guide where to undertake targeted investigations. Sentinel-1 InSAR may also provide value in monitoring legacy mines with abandoned TSFs where installing and maintaining instrumentation may pose practical challenges.

4. Due to the polar orbit of SAR satellites, sub-horizontal movements in the north-south direction are generally poorly captured.

5. Not all geotechnical failures of TSFs show clear warning signs for weeks in advance. The revisit interval of SAR satellites prevents the detection of instantaneous failure modes. Therefore, long-term monitoring programs should ideally be integrated with a combination of remote sensing methods and field instrumentation.

Future avenues of research could involve case-study comparisons between medium-resolution Sentinel-1 and high-resolution TerraSAR-X data and ground-truth comparisons between different proprietary algorithms.

5.5 Novel methods and contributions

This thesis adopted a number of new approaches and techniques to produce impactful research contributions to this field.

In Chapter 2, to overcome limitations in data availability, a new statistical method was introduced to develop the first representative estimate of constructed TSFs worldwide (~20,000) based on the analysis of a limited data sample and mineral production data. This new estimate has now been verified by a compiled "geo-census" (confidential file) that documents the names, locations, and basic information where available, for ~11,000 TSFs (as of March 2023), despite the missing/incomplete information for mining-intensive regions such as India, eastern Europe (incl. Russia), central Asia (Kazakhstan, Kyrgyzstan, Tajikistan, Uzbekistan), SE Asia (Indonesia and Philippines), and the central African sub-continent. This data compilation has had major implications beyond the scope of this thesis: the ICOLD Global TSF Registry initiative has been restarted with the goal of reporting the basic information of TSFs worldwide, and to update this "live" registry periodically (as with the ICOLD Registry of Large Water-Retention Dams). The technical component of this initiative is being led by the present author (N. Rana) with support from the ICOLD Tailings Committee, Canadian Dam Association, and other interested groups.

In Chapters 2 and 3, post-event satellite images of 90+ historical TSF failures were acquired, post-processed by GIS, and published in the public domain. Data from the first high-resolution Earth observation instrument, namely the Corona (Keyhole) spy satellite, was accessed through declassified U.S. government archives in order to recover images of case histories dating back to 1965, including the El Cobre (1965 Chile), Sgorigrad (1966 Bulgaria), and Certej (1971 Romania) disasters that collectively killed almost 800 people. After the 1927 Barahona (Chile) and 1938 Tlalpujahua (Mexico) disasters, these were the next benchmark events that prompted engineers to further study the geotechnical properties of tailings dams and to adopt better breach-prevention methods. This imagery compilation effort has enabled the possibility of conducting detailed re-examination studies by generating high-quality insights into flow behavior and by supporting field site visits in cross-checking how the site conditions have changed since the flow event.

In Chapter 3, a new runout classification method was implemented to map the downstream runout geometry of tailings flows, and to subsequently develop new validated empirical relationships to predict

these parameters from the released volume. This runout zone classification was first introduced and described in Ghahramani et al. (2020) to map inundation areas, and in Chapter 3 (Rana et al., 2021a) the technique was extended to cover runout distance and travel path angle. This methodological framework is important in forensic studies, simulation models, and consequence classifications by distinguishing between zones where impacts outside of the channelized floodplain are observed and zones where impacts are confined to the floodplain only.

In Chapter 4, an automated PS-InSAR software was used for the first time to study tailings dams and their failures. The streamlined workflow enabled the most comprehensive forensic investigation of tailings dams to date in scientific literature, but is also underlain by key technical limitations. On that note, a comparison between the results of multiple software was also introduced in tailings dam research, showing that not all software produce the same quality of results. These insights delivered practical considerations to support decision-making on the selection of the satellite and the InSAR software in monitoring projects depending on the TSF characteristics and environmental conditions.

5.6 Concluding outlook

This thesis was written in the aftermath of the September 2022 Jagersfontein TSF disaster in South Africa. A satellite image of the incident is shown in Figure 5-1. The event led to 3 deaths, over 40 injuries, evacuation of multiple villages, and major environmental and socioeconomic losses. The runout distance and inundation area are measured to be 48.5 km and 14.6 km^2 in Zone 1 using the methods outlined in Chapter 3. The event offers a tragic, sobering reminder that much work remains to be done to reach the aspirational goal of "zero-harm" in the mining industry, set by the ICMM after the Feijao disaster.

Following Jagersfontein, mining investors from the Church of England (2022) issued an urgent press release calling for (i) a global tailings repository to include location and ownership details for all TSFs, whether in public or private hands, and (ii) a global tailings monitoring system that integrates satellites as well as ground sensors (where available) to independently monitor the "highest-risk" dams, including those not held by public companies. Special attention was also directed to orphaned, abandoned, and legacy tailings sites.

These goals are indeed an ideal next step for the tailings industry. In fact, the central objectives and outcomes in this thesis have naturally aligned with these goals. As alluded to earlier in Section 5.5, the ICOLD Global TSF Registry initiative is underway and currently accounts for ~11,000 TSFs in the form of a geo-census, with several geographic regions yet to be explored. As concluded in Chapter 3, TSFs are anthropogenic sources of energy (i.e. hazards), and with little knowledge of where these hazards are, it is not possible to begin implementing proactive, integrated monitoring programs, especially for abandoned TSFs, TSFs in regions with poor data reporting, or TSFs in private ownership. Successfully creating a global registry of TSFs will require a long-term diplomatic effort involving collaboration between companies, regulators, technical experts (including engineers and remote sensing scientists), and investors.

With regards to the second goal (global monitoring system), satellite InSAR has indeed become a powerful tool, as comprehensively documented in Chapter 4. Although there are key technical challenges associated with InSAR that require consideration, the InSAR industry is rapidly evolving with new companies, modified processing algorithms, enhanced satellite data, and improved accessibility to the public. To build on the comprehensive study in Chapter 4, it is important to reiterate here the concluding recommendations for future research: (i) comparisons between Sentinel-1 and TerraSAR-X data to study the effects of different spatial resolution, (ii) case-study comparisons between the monitoring capabilities of different algorithms/software (perhaps in the form of anonymized round-robin exercises to avoid potential issues of bias and commercialization), and (iii) additional ground-truth assessments to study the accuracy/precision of radar satellite data in comparison to *in-situ* observations. Beyond these potential research avenues, a continuing challenge will be to strategize the risk prioritization approach for satellite monitoring – i.e. to define what is meant by "highest-risk" and identify tailings dams that match those criteria.

While the Jagersfontein incident offers a sobering reminder that much work remains to be done, a point of optimism is that thesis is just one of many ongoing research projects on tailings management and safety (e.g. liquefaction susceptibility, dam breach-runout modelling, water management, design criteria) – all of which are necessary to facilitate a safer, more sustainable mining industry.

Figure 5-1: Worldview images on Google Earth, captured on 15 September 2022, of the Jagersfontein TSF disaster in South Africa on 11 September 2022. [A] The breached TSF. [B] The downstream runout from the failed TSF (indicated by rectangle).

References

Adiansyah, J.S., Rosano, M., Vink, S., Keir, G., 2015. A framework for a sustainable approach to mine tailings management: disposal strategies. J. Clean. Prod. 108, 1050–1062. https://doi.org/10.1016/j.jclepro.2015.07.139.

Adiansyah, J.S., Rosano, M., Vink, S., Keir, G., 2015. A framework for a sustainable approach to mine tailings management: disposal strategies. J. Clean. Prod. 108, 1050–1062. https://doi.org/10.1016/j.jclepro.2015.07.139.

Adria, D., Ghahramani, N.G., Rana, N.M., McDougall, S., Evans, S.G., Take, W.A., 2021. Advancements in the modelling of tailings dam breaches. In: Proceedings of the Tailings and Mine Waste 2021 Conference, Alberta, Canada.

AECOM, 2009. Root cause analysis of TVA Kingston Dredge Pond failure on December 22, 2008. Tennessee Valley Authority, USA.

Agurto-Detzel, H., Bianchi, M., Assumpção, M., Schimmel, M., Collaço, B., Ciardelli, C., Barbosa, J.R., Calhau, J., 2016. The tailings dam failure of 5 November 2015 in SE Brazil and its preceding seismic sequence. Geophys. Res. Lett. 43 (10), 4929–4936. https://doi.org/10.1002/2016GL069257.

Ahmaruzzaman, M., 2010. A review on the utilization of fly ash. Prog. Energy Combustion Sci. 36(3), 327-363. https://doi.org/10.1016/j.pecs.2009.11.003.

Alonso, E.E., Gens, A., 2006a. Aznalcóllar dam failure. Part 1: Field observations and material properties. Géotechnique 56 (3), 165–183. https://doi.org/10.1680/geot.2006.56.3.165.

Alonso, E.E., Gens, A., 2006b. Aznalcóllar dam failure. Part 3: Dynamics of the motion. Géotechnique 56 (3), 203–210. https://doi.org/10.1680/geot.2006.56.3.203.

Amnesty International, 2017. The human rights impacts of the Mount Polley Mine Disaster, British Columbia, Canada.

Arroyo, M., Gens, A., 2021. Computational analyses of Dam I failure at the Corrego de Feijao mine in Brumadinho (Investigation report commissioned by Vale S.A.).

Association of State Dam Safety Officials (ASDSO), 2020. Spencer Dam Failure Investigation Report.

Association of State Dam Safety Officials (ASDSO), 2021. Investigation of failures of Edenville and Sanford Dams (interim).

Aswathi, J., Binojkumar, R.B., Oommen, T., Bouali, E.H., Sajinkumar, K.S., 2022. InSAR as a tool for monitoring hydropower projects: A review. Energy Geosci. 3(2), 160-171. https://doi.org/10.1016/j.engeos.2021.12.007.

Australia-New Zealand Committee on Large Dams (ANCOLD, 1994). Guidelines on Risk Assessment 1994.

Azam, S., Li, Q., 2010. Tailings dam failures: a review of the last one hundred years. Geotech. News 28 (4), 50–54.

Baecher, G.B., Christian, J.T., 2003. Reliability and statistics in geotechnical engineering. John Wiley & Sons.

Bakon, M., Perissin, D., Lazecky, M., Papco, J, 2014. Infrastructure non-linear deformation monitoring via satellite radar interferometry. Procedia Tech. 16, 294-300. https://doi.org/10.1016/j.protcy.2014.10.095.

Bánvölgyi, G.G., 2018. The failure of the embankment of the red mud reservoir at Ajka (Hungary). In: Proceedings of the 36th International ICSOBA Conference, Belem, Brazil.

Baxter, R.M., 1977. Environmental effects of dams and impoundments. Annu. Rev. Ecol. System. 8 (1), 255-283. https://doi.org/10.1146/annurev.es.08.110177.001351.

Bayaraa, M., Sheil, B., Rossi, C., 2022. InSAR and numerical modelling for tailings dam monitoring – the Cadia failure case study. Géotechnique, 1-19. https://doi.org/10.1680/jgeot.21.00399.

Behera, S.K., Ghosh, C.N., Mishra, D.P., Singh, P., Mishra, K., Buragohain, J., Mandal, P.K., 2020. Strength development and microstructural investigation of lead-zinc mill tailings based paste backfill with fly ash as alternative binder. Cem. Concr. Composite 109, 103553. https://doi.org/10.1016/j.cemconcomp.2020.103553.

Berardino, P., Fornaro, G., Lanari, R., Sansosti, E., 2002. A new algorithm for surface deformation monitoring based on small baseline differential SAR interferograms. IEEE Trans. Geosci. Remote Sens. 40(11), 2375-2383. https://doi.org/10.1109/TGRS.2002.803792.

Best, J., 2019. Anthropogenic stresses on the world's big rivers. Nat. Geosci. 12(1), 7-21. https://doi.org/10.1038/s41561-018-0262-x.

Bischoff, C.A., Basilico, M., Ferretti, A., Molinaro, D., Giannico, C., Ghail, R.C., Mason, P.J. (2017). A comparison between TerraSAR-X and Sentinel-1 PSInSAR data for infrastructure monitoring in London, UK. GRSG 28th International Annual Conference "Applied Geological Remote Sensing.

Bischoff, C.A., Ferretti, A., Novali, F., Uttini, A., Giannico, C., Meloni, F., 2020. Nationwide deformation monitoring with SqueeSAR® using Sentinel-1 data. Proc. Int. Assoc. Hydrol. Sci. 382, 31-37. https://doi.org/10.5194/piahs-382-31-2020.

Bishop, A.W., 1973. The stability of tips and spoil heaps. Q. J. Eng. Geol. Hydrogeol. 6 (3-4), 335–376. https://doi.org/10.1144/GSL.QJEG.1973.006.03.15.

Blight, G.E., 1997. Destructive mudflows as a consequence of tailings dyke failures. Proc. Inst. Civil Eng. Geotech. Eng. 125 (1), 9–18. https://doi.org/10.1680/igeng.1997.28992.

Blight, G. E., 2010. Geotechnical Engineering for Mine Waste Storage Facilities. CRC Press, London.

Blight, G.E., Robinson, M.J., Diering, J.A.C., 1981. The flow of slurry from a breached tailings dam. J. South. Afr. Inst. Min. Metall. 81 (1), 1–8.

Blight, G.E., Fourie, A.B., 2005. Catastrophe revisited – disastrous flow failures of mine and municipal solid waste. Geotech. Geol. Eng. 23 (3), 219–248. https://doi.org/10.1007/s10706-004-7067-y.

Boshoff, E.T., Morkel, J., Naude, N., 2018. Identifying critical parameters in the settling of African kimberlites. Miner. Process. Extr. Metall. Rev. 39(2), 136-144. https://doi.org/10.1080/08827508.2017.1391248.

Boulange, J., Hanasaki, N., Yamazaki, D., Pokhrel, Y., 2021. Role of dams in reducing global flood exposure under climate change. Nat. Comm. 12(1), 1-7. https://doi.org/10.1038/s41467-020-20704-0.

Bowker, L.N., Chambers, D.M., 2015. The risks, public liability and economics of tailings storage facility failures. University of Maine Press, pp. 1–55.

Bowker, L.N., Chambers, D.M., 2017. In the dark shadow of the supercycle tailings failure risk & public liability reach all time highs. Environ. 4 (4), 75. https://doi.org/10.3390/environments4040075.

Briaud, J.L., 2020. Failure has consequences. GeoStrata Magazine Archive 24(1), 18-20.

British Geological Survey (BGS), 2021. World Mineral Production Data. https://www2.bgs.ac.uk/mineralsuk/statistics/home.html (accessed December 2021).

Bullock, L.A., James, R.H., Matter, J., Renforth, P., Teagle, D.A., 2021. Global carbon dioxide removal potential of waste materials from metal and diamond mining. Frontiers in Climate 77. https://doi.org/10.3389/fclim.2021.694175.

Canadian Dam Association (CDA), 2013. Dam Safety Guidelines 2007 (Revised 2013).

Canadian Dam Association (CDA), 2020. Technical bulletin: guidelines for tailings dam breach analyses (draft).

Carlà, T., Intrieri, E., Raspini, F., Bardi, F., Farina, P., Ferretti, A., Colombo, D., Novali, F., Casagli, N., 2019a. Perspectives on the prediction of catastrophic slope failures from satellite InSAR. Sci. Rep. 9(1), 1-9. https://doi.org/10.1038/s41598-019-50792-y.

Carlà, T., Tofani, V., Lombardi, L., Raspini, F., Bianchini, S., Bertolo, D., Thuegaz, P., Casagli, N., 2019b. Combination of GNSS, satellite InSAR, and GBInSAR remote sensing monitoring to improve the understanding of a large landslide in high alpine environment. Geomorph. 335, 62-75. https://doi.org/10.1016/j.geomorph.2019.03.014.

Carrivick, J.L., Tweed, F.S., 2016. A global assessment of the societal impacts of glacier outburst floods. Glob. Planet. Change 144, 1-16. https://doi.org/10.1016/j.gloplacha.2016.07.001.

Casu, F., Manzo, M., Lanari, R., 2006. A quantitative assessment of the SBAS algorithm performance for surface deformation retrieval from DInSAR data. Remote Sens. Environ. 102(3-4), 195-210. https://doi.org/10.1016/j.rse.2006.01.023.

Chandler, R.J., Tosatti, G., 1995. The Stava tailings dams failure, Italy, July 1985. Proc. Inst. Civil Eng. Geotech. Eng. 133, 67–79. https://doi.org/10.1680/igeng.1995.27586.

Chovan, K., Julien, M. R., Ingabire, E. P., James, M., Masengo, E., Lépine, T., Lavoie, P., 2021. A risk assessment tool for tailings storage facilities. Can. Geotech. J. https://doi.org/10.1139/cgj-2020-0329.

Church of England, 2022. Urgent Investor Call for Global Tailings Monitoring System & Tailings Legacy Fund. Retrieved from: https://www.churchofengland.org/media-and-news/press-releases/urgent-investor-call-global-tailings-monitoring-system.

Cigna, F., Esquivel Ramírez, R., Tapete, D., 2021. Accuracy of Sentinel-1 PSI and SBAS InSAR Displacement Velocities against GNSS and Geodetic Leveling Monitoring Data. Remote Sens. 13(23), 4800. https://doi.org/10.3390/rs13234800.

Clarkson, L., Williams, D., 2021. An Overview of Conventional Tailings Dam Geotechnical Failure Mechanisms. Min. Metall. Explor. 1–24. https://doi.org/10.1007/s42461-021-00381-3.

Colombo, D. Why InSAR monitoring in mining should be "high resolution". Retrieved from: https://www.linkedin.com/pulse/why-insar-monitoring-mining-should-high-resolution-davide-colombo/ (accessed 10 November 2022).

Colombo, P., Colleselli, F., 2003. The collapse of the Stava tailings basins – analysis of the event and observations. In: Tosatti, G. (Ed.), A Review of Scientific Contributions on the Stava Valley Disaster (Eastern Italian Alps), 19th July 1985. Pitagora Editrice, Bologna, pp. 335–358.

Corominas, J., 1996. The angle of reach as a mobility index for small and large landslides, Can. Geotech. J. 33 (2), 260–271. https://doi.org/10.1139/t96-005.

Costa, J.E., 1985. Floods from dam failures. U.S. Geol. Survey Open-File Report 85–560.

Courchesne, B., Schindler, M., Mykytczuk, N.C., 2021. Relationships between the microbial composition and the geochemistry and mineralogy of the cobalt-bearing legacy mine tailings in Northeastern Ontario. Front. Microbiol. 12. https://doi.org/10.3389/fmicb.2021.660190.

Crosetto, M., Monserrat, O., Cuevas-González, M., Devanthéry, N., Crippa, B., 2016. Persistent scatterer interferometry: A review. ISPRS J. Photogramm. Remote Sens. 115, 78-89. https://doi.org/10.1016/j.isprsjprs.2015.10.011.

Crosta, G.B., Cucchiaro, S., Frattini, P., 2003. Validation of semi-empirical relationships for the definition of debris-flow behavior in granular materials. In: Proceedings of the 3rd International Conference on Debris-Flow Hazards Mitigation: Mechanics, Prediction and Assessment, Davos, Switzerland, pp. 821–831.

Cruden, D.M., Varnes, D.J., 1996. Landslide types and processes. In: Turner, A.K., Schuster, R.L. (Eds.), Landslides investigation and mitigation. Transportation research board, US National Research Council. Special Report 247, Chapter 3, pp. 36–75.

Cuervo, V., Burge, L., Beaugrand, H., Hendershot, M., Evans, S.G., 2017. Downstream geomorphic response of the 2014 Mount Polley tailings dam failure, British Columbia. In: Workshop on World Landslide Forum, pp. 281–289.

Davies, M., Martin, T., Lighthall, P., 2000. Mine tailings dams: when things go wrong.

Davies, M.P., 2002. Tailings impoundment failures: Are geotechnical engineers listening? Waste Geotech. 20, 1–36.

Davies, W.E., Bailey, J.F., Kelly, D.B., 1972. West Virginia's Buffalo Creek Flood: A Study of the Hydrology and Engineering Geology. U.S. Geol. Survey Circular 667.

Dawson, R. F., Morgenstern, N. R., Stokes, A. W., 1998. Liquefaction flowslides in Rocky Mountain coal mine waste dumps. Can. Geotech. J. 35(2), 328-343. https://doi.org/10.1139/t98-009.

DeNeale, S.T., Baecher, G.B., Stewart, K.M., Smith, E.D., Watson, D.B., 2019. Current state-of-practice in dam safety risk assessment. Oak Ridge National Laboratory, USA.

Di Baldassarre, G., Wanders, N., AghaKouchak, A., Kuil, L., Rangecroft, S., Veldkamp, T.I., Garcia, M., van Oel, P.R., Breinl, K., Van Loon, A.F., 2018. Water shortages worsened by reservoir effects. Nat. Sustain. 1 (11), 617-622. https://doi.org/10.1038/s41893-018-0159-0.

Dobry, R., Alvarez, L., 1967. Seismic failures of Chilean tailings dams. J. Soil Mech. Found. Div. 93, 237–260.

Dong, L., Deng, S., Wang, F., 2020. Some developments and new insights for environmental sustainability and disaster control of tailings dam. J. Clean. Prod. 269, 122270. https://doi.org/10.1016/j.jclepro.2020.122270.

Duan, H., Li, Y., Jiang, H., Li, Q., Jiang, W., Tian, Y., Zhang, J., 2023. Retrospective monitoring of slope failure event of tailings dam using InSAR time-series observations. Nat. Haz. 1-17. https://doi.org/10.1007/s11069-023-05946-7.

Dutta, T., Kim, K.H., Uchimiya, M., Kwon, E.E., Jeon, B.H., Deep, A., Yun, S.T., 2016. Global demand for rare earth resources and strategies for green mining. Environ. Res. 150, 182-190. https://doi.org/10.1016/j.envres.2016.05.052.

Evans, S.G., 2006. The formation and failure of landslide dams: an approach to risk assessment. Ital. J. Eng. Geol. Environ. 1, 15-20.

Evans, S.G., Delaney, K.B., and Rana, N.M., 2021. The occurrence and mechanism of catastrophic mass flows in the mountain cryosphere. In Snow and Ice-Related Hazards, Risks, and Disasters (pp. 541-596). Elsevier. https://doi.org/10.1016/B978-0-12-817129-5.00004-4.

Evans, S.G., Hermanns, R.L., Strom, A., Scarascia-Mugnozza, G., 2011. Natural and artificial rockslide dams. Springer.

Fan, H., He, D., Wang, H., 2015. Environmental consequences of damming the mainstream Lancang-Mekong River: A review. Earth-Sci. Rev. 146, 77-91. https://doi.org/10.1016/j.earscirev.2015.03.007.

Feng, S., Li, Y., 2021. Study on coal fly ash classified by bulk density. In: Journal of Physics: Conference Series. IOP Publishing. Shanghai, China.

Ferretti, A., Prati, C., Rocca, F., 2001. Permanent scatterers in SAR interferometry. IEEE Trans. Geosci. Remote Sens. 39(1), 8-20. https://doi.org/10.1109/36.898661.

Ferretti, A., Fumagalli, A., Novali, F., Prati, C., Rocca, F., Rucci, A., 2011. A new algorithm for processing interferometric data-stacks: SqueeSAR. IEEE Trans. Geosci. Remote Sens. 49(9), 3460-3470. https://doi.org/10.1109/TGRS.2011.2124465.

Foster, M., Fell, R., Spannagle, M., 2000. The statistics of embankment dam failures and accidents. Can. Geotech. J. 37 (5), 1000-1024. https://doi.org/10.1139/t00-030.

Fourie, A.B., Blight, G.E., Papageorgiou, G., 2001. Static liquefaction as a possible explanation for the Merriespruit tailings dam failure. Can. Geotech. J. 38 (4), 707–719. https://doi.org/10.1139/t00-112.

Fourie, A.B., Papageorgiou, G., 2001. Defining an appropriate steady state line for Merriespruit gold tailings. Can. Geotech. J. 38 (4), 695–706. https://doi.org/10.1139/t00-111.

Franks, D.M., Stringer, M., Torres-Cruz, L.A., Baker, E., Valenta, R., Thygesen, K., Matthews, A., Howchin, J., Barrie, S., 2021. Tailings facility disclosures reveal stability risks. Sci. Rep. 11 (1), 1–7. https://doi.org/10.1038/s41598-021-84897-0.

Gabriel, F.Â., Ferreira, A.D., Queiroz, H.M., Vasconcelos, A.L.S., Ferreira, T.O., Bernardino, A.F., 2021. Long-term contamination of the Rio Doce estuary as a result of Brazil's largest environmental disaster. Perspect. Ecol. Conserv. 19(4), 417-428. https://doi.org/10.1016/j.pecon.2021.09.001.

Gallart, F., Benito, G., Martín-Vide, J.P., Benito, A., Prió, J.M., Regüés, D., 1999. Fluvial geomorphology and hydrology in the dispersal and fate of pyrite mud particles released by the Aznalcóllar mine tailings spill. Sci. Total Environ. 242 (1–3), 13–26. https://doi.org/10.1016/S0048-9697(99)00373-3.

Galvão, E.S., Feroni, R.D.C., Silveira, A.B., 2020. Trends in air quality and hospital admissions due to respiratory diseases since the standstill of an industrial plant in Brazil. Environ. Sci. Pollut. Res. 27(19), 24452-24465. https://doi.org/10.1007/s11356-020-08787-0.

Gama, F., Mura, J.C., R. Paradella, W., de Oliveira, C.G., 2020. Deformations prior to the Brumadinho dam collapse revealed by Sentinel-1 InSAR data using SBAS and PSI techniques. Remote Sens. 12(21), 3664. https://doi.org/10.3390/rs12213664.

Gauthier-Manuel, H., Radola, D., Choulet, F., Buatier, M., Vauthier, R., Morvan, T., Chavanne, W., Gimbert, F., 2021. A multidisciplinary approach for the assessment of origin, fate and ecotoxicity of metal (loid) s from legacy coal mine tailings. Toxics 9(7), 164. https://doi.org/10.3390/toxics9070164.

Ge, W., Qin, Y., Li, Z., Zhang, H., Gao, W., Guo, X., Song, Z., Li, W., van Gelder, P., 2020. An innovative methodology for establishing societal life risk criteria for dams: A case study to reservoir dam failure events in China. Int. J. Disaster Risk Reduct. 49, 101663. https://doi.org/10.1016/j.ijdrr.2020.101663.

Gens, A., Alonso, E.E., 2006. Aznalcóllar dam failure. Part 2: Stability conditions and failure mechanism. Géotechnique 56 (3), 185–201. https://doi.org/10.1680/geot.2006.56.3.185.

Geotechnical Engineering Office, Hong Kong Government (GEO), 1998. Landslides and Boulder Falls from Natural Terrain: Interim Risk Guidelines. Government of the Hong Kong Special Administrative Region, Geotechnical Engineering Office, GEO Report No.75.

Ghahramani, N., Mitchell, A., Rana, N.M., McDougall, S., Evans, S.G., Take, A., 2020. Tailings-flow runout analysis: Examining the applicability of a semi-physical area–volume relationship using a novel database. Nat. Hazards Earth Syst. Sci. https://doi.org/10.5194/nhess-2020-199.

Ghahramani N., Rana, N.M., Liu, S., Chen, H. J., Clohan, D., McDougall, S., Evans, S.G., Take, W.A. 2021. A preliminary benchmarking numerical modelling exercise for tailings flow runout analysis. In: Proceedings of Mine Waste and Tailings Conference 2021. 1-2 July, Brisbane, Australia, 433-443.

Ghahramani, N., Chen, H.J., Clohan, D., Liu, S., Llano-Serna, M., Rana, N.M., McDougall, S., Evans, S.G., Take, W.A., 2022. A benchmarking study of four numerical runout models for the simulation of tailings flows. Sci. Tot. Environ. 154245. https://doi.org/10.1016/j.scitotenv.2022.154245.

Gildeh, H.K., Halliday, A., Arenas, A., Zhang, H., 2020. Tailings Dam Breach Analysis: A Review of Methods, Practices, and Uncertainties. Mine Water Environ. 1-23. https://doi.org/10.1007/s10230-020-00718-2.

Glotov, V.E., Chlachula, J., Glotova, L.P., Little, E., 2018. Causes and environmental impact of the gold-tailings dam failure at Karamken, the Russian Far East. Eng. Geol. 245, 236–247. https://doi.org/10.1016/j.enggeo.2018.08.012.

Graham, W.J., 1999. A Procedure for Estimating Loss of Life Caused by Dam Failure. U.S. Department of Interior.

Grebby, S., Sowter, A., Gluyas, J., Toll, D., Gee, D., Athab, A., Girindran, R., 2021. Advanced analysis of satellite data reveals ground deformation precursors to the Brumadinho Tailings Dam collapse. Commun. Earth & Environ. 2 (2). https://doi.org/10.1038/s43247-020-00079-2.

Grzovic, M., Ghulam, A., 2015. Evaluation of land subsidence from underground coal mining using TimeSAR (SBAS and PSI) in Springfield, Illinois, USA. Nat. Hazards 79(3), 1739-1751. https://doi.org/10.1007/s11069-015-1927-z.

Gutenberg, B., Richter, C.F., 1954. Seismicity of the Earth and Associated Phenomena. Princeton University Press.

Guthrie, R.H., Evans, S.G., 2004. Magnitude and frequency of landslides triggered by a storm event, Loughborough Inlet. British Columbia. Nat. Hazards Earth Syst. Sci. 4, 475–483. https://doi.org/10.5194/nhess-4-475-2004.

Guthrie, R., 2015. The catastrophic nature of humans. Nat. Geosci. 8(6), 421-422. https://doi.org/10.1038/ngeo2455.

Halabi, A.L.M., Siacara, A.T., Sakano, V.K., Pileggi, R.G., Futai, M.M., 2022. Tailings Dam Failures: A Historical Analysis of the Risk. J. Fail. Anal. Prev. 1-14. https://doi.org/10.1007/s11668-022-01355-3.

Harder Jr., L.F., Stewart, J.P., 1996. Failure of Tapo Canyon tailings dam. J. Perform. Constr. Facil. 10 (3), 109–114. https://doi.org/10.1061/(ASCE)0887-3828(1996)10:3(109).

He, X.Y., Wang, Z.Y., Huang, J.C., 2008. Temporal and spatial distribution of dam failure events in China. Int. J. Sediment Res. 23(4), 398-405. https://doi.org/10.1016/S1001-6279(09)60010-X.

Herrington, R., 2021. Mining our green future. Nat. Rev. Mater. 6(6), 456-458. https://doi.org/10.1038/s41578-021-00325-9.

Holden, D., Donegan, S., Pon, A., 2020. Brumadinho Dam InSAR study: analysis of TerraSAR-X, COSMO-SkyMed and Sentinel-1 images preceding the collapse. In Slope Stability 2020: In: Proceedings of the 2020 International Symposium on Slope Stability in Open Pit Mining and Civil Engineering, Australia, pp. 293-306.

Hu, X., Oommen, T., Lu, Z., Wang, T., Kim, J.W., 2017. Consolidation settlement of Salt Lake County tailings impoundment revealed by time-series InSAR observations from multiple radar satellites. Remote Sens. Environ. 202, 199-209. https://doi.org/10.1016/j.rse.2017.05.023.

Hudson, R., Sato, S., Morin, R. McParland, M.A., 2021. Comparison of Sentinel-1 and Radarsat-2 data for monitoring of tailings storage facilities. EUSAR 2021: 13th European conference on synthetic aperture radar, Frankfurt, Germany.

Hungr, O., Evans, S.G., Bovis, M.J., Hutchinson, I.N., 2001. A review of the classification of landslides of the flow type. Environ. Eng. Geosci. 7 (3), 221-238.

Hungr, O., Evans, S.G., 2004. Entrainment of debris in rock avalanches; an analysis of a long run-out mechanism. Geo. Soc. Am. Bull. 116 (9–10), 1240–1252. https://doi.org/10.1130/B25362.1.

Hungr, O., McDougall, S., Wise, M., Cullen, M., 2008. Magnitude-frequency relationships of debris flows and debris avalanches in relation to slope relief. Geomorphol. 96 (3–4), 355–365. https://doi.org/10.1016/j.geomorph.2007.03.020.

Hungr, O., Leroueil, S., Picarelli, L., 2014. The Varnes classification of landslide types, an update. Landslides 11 (2), 167–194. https://doi.org/10.1007/s10346-013-0436-y.

Independent Forensic Team Report, 2018. Oroville Dam Spillway Incident.

Innis, S., Ghahramani, N., Rana, N.M., McDougall, S., Take, W.A., Kunz, N.C., 2022. The development and demonstration of a semi-automated regional scale hazard mapping tool for tailings storage facility failures. Resources 11(10), 82. https://doi.org/10.3390/resources11100082.

International Commission on Large Dams (ICOLD), 1996. A Guide to Tailings Dams and Impoundments. Design, Construction, Use and Rehabilitation. Bulletin 106, Paris. 242 p.

International Commission on Large Dams (ICOLD), 2001. Tailings dams – risk of dangerous occurrences, lessons learnt from practical experiences. Bulletin 121, Paris. 146 p.

International Commission on Large Dams (ICOLD), 2020. World Register of Dams. https://www.icold-cigb.org/GB/world_register/general_synthesis.asp (accessed October 2021).

International Commission on Large Dams (ICOLD), 2023. Tailings Dam Safety. Bulletin 194.

International Council on Mining and Metals (ICMM), United Nations Environment Programme (UNEP), Principles for Responsible Investment (PRI), 2020. Global Industry Standard on Tailings Management. https://globaltailingsreview.org/ (accessed 1 September 2020).

International Institute for Environment and Development (IIED), 2002. Breaking New Ground: Mining, Minerals and Sustainable Development. 450 p.

Ishihara, K., 1984. Post-earthquake failure of a tailings dam due to liquefaction of pond deposit. In: Proceedings of the 1st International Conference on Case Histories in Geotechnical Engineering, Missouri, USA.

Ishihara, K., Ueno, K., Yamada, S., Yasuda, S., Yoneoka, T., 2015. Breach of a tailings dam in the 2011 earthquake in Japan. Soil Dyn. Earthq. Eng. 68, 3–22. https://doi.org/10.1016/j.soildyn.2014.10.010.

Islam, K., Murakami, S., 2021. Global-scale impact analysis of mine tailings dam failures: 1915–2020. Glob. Environ. Change 70, 102361. https://doi.org/10.1016/j.gloenvcha.2021.102361.

Jefferies, M., Been, K., 2016. Soil liquefaction: a critical state approach, second ed. CRC Press, Florida.

Jefferies, M., Morgenstern, N.R., Van Zyl, D.V., Wates, J., 2019. Report on NTSF Embankment Failure. Cadia Valley Operations, for Ashurst Operations, South Orange, NSW, Australia.

Jones, H., Boger, D.V., 2012. Sustainability and waste management in the resource industries. Ind. Eng. Chem. Res. 51, 10057–10065. https://doi.org/10.1021/ie202963z.

Joshi, R.C., Lohita, R.P., 1997. Fly ash in concrete: production, properties and uses (Vol. 2). CRC Press.

Kim, J., Coe, J.A., Lu, Z., Avdievitch, N.N., Hults, C.P., 2022. Spaceborne InSAR mapping of landslides and subsidence in rapidly deglaciating terrain, Glacier Bay National Park and Preserve and vicinity, Alaska and British Columbia. Remote Sens. Environ. 281, 113231. https://doi.org/10.1016/j.rse.2022.113231.

Kossoff, D., Dubbin, W.E., Alfredsson, M., Edwards, S.J., Macklin, M.G., Hudson-Edwards, K.A., 2014. Mine tailings dams: Characteristics, failure, environmental impacts, and remediation. Appl. Geochem. 51, 229–245. https://doi.org/10.1016/j.apgeochem.2014.09.010.

Kraus, U., Wiegand, J., 2006. Long-term effects of the Aznalcóllar mine spill—heavy metal content and mobility in soils and sediments of the Guadiamar river valley (SW Spain). Sci. Total Environ. 367(2-3), 855-871. https://doi.org/10.1016/j.scitotenv.2005.12.027.

Kumar, R.M., 2019. Muri hindalco red mud blasting(2). https://www.youtube.com/watch?v=8K63D70b4CU&ab_channel=R.Mkumar (accessed 10 April 2019).

Lanari, R., Casu, F., Manzo, M., Zeni, G., Berardino, P., Manunta, M., Pepe, A., 2007. An overview of the small baseline subset algorithm: a DInSAR technique for surface deformation analysis. Pure Appl. Geophys. 164 (4), 637–661.

Larrauri, P.C., Lall, U., 2018. Tailings dams failures: updated statistical model for discharge volume and runout. Environ 5 (2), 28. https://doi.org/10.3390/environments5020028.

Latrubesse, E.M., Park, E., Sieh, K., Dang, T., Lin, Y.N., Yun, S.H., 2020. Dam failure and a catastrophic flood in the Mekong basin (Bolaven Plateau), southern Laos. Geomorphol. 362, 107221. https://doi.org/10.1016/j.geomorph.2020.107221.

Legros, F., 2002. The mobility of long-runout landslides. Eng. Geol. 63 (3–4), 301–331. https://doi.org/10.1016/S0013-7952(01)00090-4.

Liang, T., Werner, T.T., Heping, X., Jingsong, Y., Zeming, S., 2021. A global-scale spatial assessment and geodatabase of mine areas. Glob. Planet. Change 204, 103578. https://doi.org/10.1016/j.gloplacha.2021.103578.

de Lima, R.E., de Lima Picanço, J., da Silva, A.F., Acordes, F.A., 2020. An anthropogenic flow type gravitational mass movement: the Corrego do Feijao tailings dam disaster, Brumadinho, Brazil. Landslides 1–12. https://doi.org/10.1007/s10346-020-01450-2.

Liu, H., Probst, A., Liao, B., 2005. Metal contamination of soils and crops affected by the Chenzhou lead/zinc mine spill (Hunan, China). Sci. Total Environ. 339(1-3), 153-166. https://doi.org/10.1016/j.scitotenv.2004.07.030.

Liu, W., Carling, P.A., Hu, K., Wang, H., Zhou, Z., Zhou, L., Liu, D., Lai, Z., Zhang, X., 2019. Outburst floods in China: A review. Earth-Sci. Rev. 197, 102895. https://doi.org/10.1016/j.earscirev.2019.102895.

Lucchi, G., 2009. Sgorigrad–Stava identical disasters. Arca Edizioni, Lavis.

Luino, F., De Graff, J.V., 2012. The Stava mudflow of 19 July 1985 (Northern Italy): a disaster that effective regulation might have prevented. Nat. Hazards Earth Syst. Sci. 12 (4), 1029-1044 https://doi.org/10.5194/nhess-12-1029-2012.

Lyu, Z., Chai, J., Xu, Z., Qin, Y., Cao, J., 2019. A comprehensive review on reasons for tailings dam failures based on case history. Adv. Civ. Eng. https://doi.org/10.1155/2019/4159306.

MacDonald, T. C., Langridge-Monopolis, J., 1984. Breaching characteristics of dam failures. J. Hydraul. Eng. 110 (5), 567-586. https://doi.org/10.1061/(ASCE)0733-9429(1984)110:5(567).

Macías, J.L., Corona-Chávez, P., Sanchez-Nuñez, J.M., Martínez-Medina, M., Garduño-Monroy, V.H., Capra, L., Garcia-Tenorío, F., Cisneros-Máximo, G., 2015. The 27 May 1937 catastrophic flow failure of gold tailings at Tlalpujahua, Michoacán, Mexico. Nat. Hazard. Earth. Syst. Sci. 15(5), 1069–1085. https://doi.org/10.5194/nhess-15-1069-2015.

Macklin, M.G., Brewer, P.A., Balteanu, D., Coulthard, T.J., Driga, B., Howard, A.J., Zaharia, S., 2003. The long term fate and environmental significance of contaminant metals released by the January and March 2000 mining tailings dam failures in Maramureş County, upper Tisa Basin, Romania. Appl. Geochem. 18(2), 241-257. https://doi.org/10.1016/S0883-2927(02)00123-3.

Mahmoud, A.A., Wang, J.T., Jin, F., 2020. An improved method for estimating life losses from dam failure in China. Stochastic Environ. Res. Risk Assess. 34, 1263-1279. https://doi.org/10.1007/s00477-020-01820-1.

Martin, T.E., McRoberts, E.C., 1999. Some considerations in the stability analysis of upstream tailings dams. In: Proceedings of the 6th International Conference on Tailings and Mine Waste, Colorado, USA, pp. 287–302.

Martin, V., Al-Mamun, M., Small, A, 2019. CDA technical bulletin on tailings dam breach analyses. In: Proceedings of the 2019 International Commission on Large Dams Conference, Ottawa, Canada.

Martin, V., Fontaine, D., Cathcart, J., 2015. Challenges with conducting tailings dam breach assessments. In: Proceedings of the Tailings and Mine Waste, Vancouver, Canada, pp. 314–328.

Matyas, E.L., Welch, D.E., Reades, D.W., 1984. Geotechnical parameters and behavior of uranium tailings. Can. Geotech. J. 21(3), 489-504. https://doi.org/10.1139/t84-052.

Maus, V., Giljum, S., Gutschlhofer, J., da Silva, D.M., Probst, M., Gass, S.L., Luckeneder, S., Lieber, M., McCallum, I., 2020. A global-scale data set of mining areas. Scientific data 7(1), 1-13. https://doi.org/10.1038/s41597-020-00624-w.

McDermott, R.K., Sibley, J.M., 2000. The Aznalcóllar tailings dam accident – a case study. Min. Res. Eng. 9 (1), 101–118. https://doi.org/10.1142/S0950609800000111.

McGregor, R.G., Blowes, D.W., 2002. The physical, chemical and mineralogical properties of three cemented layers within sulfide-bearing mine tailings. J. Geochem. Explor. 76(3), 195-207. https://doi.org/10.1016/S0375-6742(02)00255-8.

McKenna, G., Van Zyl, D., 2020. Chapter VIII closure and reclamation. In: Oberle, B., Brereton, D., Mihaylova, A. (Eds.), Towards Zero Harm: A Compendium of Papers Prepared for the Global Tailings Review. https://globaltailingsreview.org/. (Accessed November 2020).

Mecsi, J., 2013. Technical analyses and lessons of the embankment failure at the Ajka red mud reservoir. In: Proceedings of the 7th International Conference on Soil Mechanics and Geotechnical Engineering 66, Paris.

Midgley, D.C., 1979. Hydrological aspects and a barrier to further escape of slimes: the failure of a slimes dam at Bafokeng. Civ. Eng. South Afr. 21 (6), 151–154.

Milanez, B., Ali, S.H., de Oliveira, J.A.P., 2021. Mapping industrial disaster recovery: lessons from mining dam failures in Brazil. Extr. Ind. Soc. 8(2), 100900. https://doi.org/10.1016/j.exis.2021.100900.

Morgenstern, N.R., 2012. Oil Sands Mine Closure – the End Game: An Update. In Proceedings of the 3rd International Oil Sands Tailings Conference, Edmonton, Alberta, 2-5 December 2012.

Morgenstern N.R., Vick, S.G., Viotti, C.B., Watts, B.D., 2016. Fundão tailings dam review panel: report on the immediate causes of the failure of the Fundão dam.

Morgenstern, N.R., 2018. Geotechnical risk, regulation, and public policy. Soils and Rocks 41 (2), 107–129.

Morgenstern, N.R., Vick, S.G., Van Zyl, D., 2015. Report on Mount Polley tailings storage facility breach. Report of independent expert engineering investigation and review panel for the Government of British Columbia and the Williams Lake and Soda Creek Indian Bands (Canada).

Morgenstern, N.R., Vick, S.G., Viotti, C.B., Watts, B.D., 2016. Fundao Tailings Dam Review Panel: Report on the Immediate Causes of the Failure of the Fundao Dam.

Mossa, J., James, L.A., 2013. Impacts of mining on geomorphic systems. In: Schroder, J.F. (Ed.), Treatise on Geomorphology. Academic Press, San Diego, pp. 74–95.

Mudd, G.M., 2007. Global trends in gold mining: towards quantifying environmental and resource sustainability. Resources Policy 32 (1–2), 42–56. https://doi.org/10.1016/j.resourpol.2007.05.002.

Mudd, G.M., Boger, D.V. 2013. The ever growing case for paste and thickened tailings – towards more sustainable mine waste management. J. Aust. Inst. Min. Metall. 2, 56–59.

Mulligan, M., van Soesbergen, A., Sáenz, L., 2020. GOODD, a global dataset of more than 38,000 georeferenced dams. Scientific Data 7(1), 1-8. https://doi.org/10.1038/s41597-020-0362-5.

Nilsson, C., 2009. Reservoirs. In: Likens, G.E (Ed.), Encyclopedia of inland waters. pp. 625-633. http://dx.doi.org/10.1016/B978-012370626-3.00039-9.

Ochieng, G.M., Seanego, E.S., Nkwonta, O.I, 2010. Impacts of mining on water resources in South Africa: A review. Sci. Res. Essays 5(22), 3351-3357.

Okereafor, U., Makhatha, M., Mekuto, L., Uche-Okereafor, N., Sebola, T., Mavumengwana, V., 2020. Toxic metal implications on agricultural soils, plants, animals, aquatic life and human health. Int. J. Environ. Res. Public Health 17(7), 2204. https://doi.org/10.3390/ijerph17072204.

Okusa, S., Anma, S., Maikuma, H., 1980. Liquefaction of mine tailings in the 1978 Izu-Ohshima-Kinkai earthquake, central Japan. In: Proceedings of the 7th World Conference on Earthquake Engineering 3, Istanbul, Turkey.

Olalla, C., Cuellar, V., 2001. Failure mechanism of the Aznalcóllar Dam, Seville, Spain. Géotechnique 51 (5), 399–406. https://doi.org/10.1680/geot.2001.51.5.399.

Oliver-Smith, A., 1996. Anthropological research on hazards and disasters. Annu. Rev. Anthr. 25 (1), 303-328. https://doi.org/10.1146/annurev.anthro.25.1.303.

Owen, J.R., Kemp, D., Lèbre, É., Svobodova, K., Murillo, G.P., 2020. Catastrophic tailings dam failures and disaster risk disclosure. Int. J. Disaster Risk Reduct. 42. https://doi.org/10.1016/j.ijdrr.2019.101361.

Palu, M.C., Julien, P.Y., 2019. Modeling the sediment load of the Doce River after the Fundão tailings dam collapse, Brazil. J. Hydraul. Eng. 145 (5). https://doi.org/10.1061/(ASCE)HY.1943-7900.0001582.

Pan, H., Zhou, G., Cheng, Z., Yang, R., He, L., Zeng, D., Sun, B., 2014. Advances in geochemical survey of mine tailings project in China. J. Geochem. Explor. 139, 193–200. http://dx.doi.org/10.1016/j.gexplo.2013.07.012.

Paronuzzi, P., Rigo, E., Bolla, A., 2013. Influence of filling–drawdown cycles of the Vajont reservoir on Mt. Toc slope stability. Geomorph. 191, 75-93. https://doi.org/10.1016/j.geomorph.2013.03.004.

Perissin, D., Wang, Z., Wang, T., 2011. The SARPROZ InSAR tool for urban subsidence/manmade structure stability monitoring in China. Proc. ISRSE, Sidney, Australia, 1015.

Petley, D., 2017. Mishor Rotem – another tailings dam failure, this time in Israel. https://blogs.agu.org/landslideblog/2017/07/07/mishor-rotem-1/ (accessed 1 September 2020).

Petley, D., 2019. Cobriza, Peru: another significant tailings dam failure. https://blogs.agu.org/landslideblog/2019/07/16/cobriza-mine-1/ (accessed 1 September 2020).

Petley, D., 2020a. The Luming Mine tailings accident in China. https://blogs.agu.org/landslideblog/2020/03/31/luming-mine-tailings-accident/ (accessed 1 April 2020).

Petley, D., 2020b. A major fly ash failure and release at Singrauli in India – Planet Labs images. https://blogs.agu.org/landslideblog/2020/04/23/a-major-fly-ash-failure-and-release-at-singrauli-in-india-planet-labs-images/ (accessed 24 April 2020).

Petley, D., 2020c. The exceptional mobility of tailings dam failures. https://blogs.agu.org/landslideblog/2020/02/10/the-exceptional-moblity-of-tailings-dam-failures/ (accessed 8 October 2020).

Petley, D., 2020d. Cobriza, Peru: another significant tailings dam failure. https://blogs.agu.org/landslideblog/2019/07/16/cobriza-mine-1/ (accessed 10 October 2020).

Piciullo, L., Storrøsten, E.B., Zhongqiang, L., Farrokh, N., Lacasse, S., 2022. A new look at the statistics of tailings dam failures. Eng. Geol. 106657. https://doi.org/10.1016/j.enggeo.2022.106657.

Rana, N.M., Ghahramani, N., Evans, S.G., McDougall, S., Small, A., Take, W.A., 2021a. Catastrophic mass flows resulting from tailings impoundment failures. Eng. Geol. 292, 106262. https://doi.org/10.1016/j.enggeo.2021.106262.

Rana, N.M., Ghahramani, N., Evans, S.G., McDougall, S., Small, A., Take, W.A., 2021b. A comprehensive global database of tailings flows. Borealis, v2. https://doi.org/10.5683/SP2/NXMXTI.

Rana, N.M. Ghahramani, N., Evans, S.G., McDougall, S., Take, W.A., 2021c. Analysis of downstream field behavior of long-runout tailings flows using remote sensing. In: Proceedings of the Mine Waste and Tailings Conference, Brisbane, Australia.

Rana, N.M., Ghahramani, N., Evans, S.G., McDougall, S., Take, W.A., 2021d. The magnitude-frequency of tailings flows resulting from failures of tailings impoundments. In: Proceedings of the 2021 International Tailings and Mine Waste Conference, Vancouver, Canada.

Rana, N.M., Ghahramani, N., Evans, S.G., Small, A., Skermer, N., McDougall, S., Take, W.A., 2022a. Global magnitude-frequency statistics of the failures and impacts of large water-retention dams and mine tailings impoundments. Earth-Sci. Rev. 104144. https://doi.org/10.1016/j.earscirev.2022.104144.

Rana, N.M., Ghahramani, N., Evans, S.G., Small, A., Skermer, N., McDougall, S., Take, W.A., 2022b. Supplementary materials to "Global magnitude-frequency statistics of the failures and impacts of large water-retention dams and mine tailings impoundments". Borealis, v3. https://doi.org/10.5683/SP3/K7SANT.

Raspini, F., Caleca, F., Del Soldato, M., Festa, D., Confuorto, P., Bianchini, S., 2022. Review of satellite radar interferometry for subsidence analysis. Earth-Sci. Rev. 104239. https://doi.org/10.1016/j.earscirev.2022.104239.

Reid, D., Fourie, A.B., 2017. Back analyses of the August 2016 Luoyang red mud tailings failure, in: Proceedings of the 2017 Tailings and Mine Waste Conference, Alberta, Canada.

Rico, M., Benito, G., Salgueiro, A.R., Díez-Herrero, A., Pereira, H.G., 2008a. Reported tailings dam failures: a review of the European incidents in the worldwide context. J. Hazard. Mater. 152 (2), 846–852. https://doi.org/10.1016/j.jhazmat.2007.07.050.

Rico, M., Benito, G., Diez-Herrero, A., 2008b. Floods from tailings dam failures. J. Hazard. Mater. 154 (1–3), 79–87. https://doi.org/10.1016/j.jhazmat.2007.09.110.

Robertson, P.K., de Melo, L., Williams, D.J., Wilson, G.W., 2019. Report of the expert panel on the technical causes of the failure of Feijão Dam I. Investigation report commissioned by Vale S.A.

Rose, N.D., Hungr, O., 2007. Forecasting potential rock slope failure in open pit mines using the inverse-velocity method. Int. J. Rock Mech. Min. Sci. 44 (2), 308–320. https://doi.org/10.1016/j.ijrmms.2006.07.014.

Rotta, L.H.S., Alcantara, E., Park, E., Negri, R.G., Lin, Y.N., Bernardo, N., Mendes, T.S.G., Filho, C.R.S., 2020. The 2019 Brumadinho tailings dam collapse: possible cause and impacts of the worst human and environmental disaster in Brazil. Int. J. Appl. Earth Obs. Geoinf. 90, 102119. https://doi.org/10.1016/j.jag.2020.102119.

Rourke, H., Luppnow, D., 2015. The risks of excess water on tailings facilities and its application to dam-break studies. In: Tailings and Mine Waste Management 2015 proceeding: https://www.srk.com/en/publications/the-risks-of-excess-water-on-tailings-facilities-and-its-application-to-dam-break-studies.

Saito, M., 1965. Forecasting the time of occurrence of slope failure. In: Proceeding of the 6th International Conference on Soil Mechanics and Foundation Engineering, Montreal. vol. 2. pp. 537–541.

Saito, M., 1969. Forecasting Time of Slope failure by Tertiary Creep. In: Proceedings of the 7th International Conference on Soil Mechanics and Foundation Engineering, Mexico City. vol. 2. pp. 677–683.

Santamarina, J.C., Torres-Cruz, L.A., Bachus, R.C., 2019. Why coal ash and tailings dam disasters occur. Sci. 364 (6440), 526–528.

Schafer, H.L., Slingerland, N., Macciotta, R., Beier, N.A., 2018. Overview of current state of practice for closure of tailings dams. In: Proceedings of the 6th International Oil Sands Tailings Conference, Edmonton, Alberta, Canada.

Seed, H.B., Idriss, I.M., Lee, K.L., Makdisi, F.I., 1975. Dynamic analysis of the slide in the Lower San Fernando Dam during the earthquake of February 9, 1971. J. Geotech. Eng. Div. 101(9), 889-911. https://doi.org/10.1061/AJGEB6.0000195.

Shakesby, R.A., Whitlow J.R., 1991. Failure of a mine waste dump in Zimbabwe: causes and consequences. Environ. Geol. Water Sci. 18, 143–153. https://doi.org/10.1007/BF01704668.

Shandro, J., Jokinen, L., Stockwell, A., Mazzei, F., Winkler, M.S., 2017. Risks and impacts to First Nation health and the Mount Polley mine tailings dam failure. Int. J. Indig. Health 12(2), 84-102. https://doi.org/10.18357/ijih122201717786.

Shanker, P., Casu, F., Zebker, H.A., Lanari, R., 2011. Comparison of persistent scatterers and small baseline time-series InSAR results: A case study of the San Francisco bay area. IEEE Geoscience and Remote Sensing Letters, 8(4), 592-596. https://doi.org/10.1109/LGRS.2010.2095829.

Small, A., James, M., Al-Mamun, M., 2017. Advancing the state of practice for tailings dam breach assessment using empirical correlations. In: Proceedings of the Canadian Dam Association Annual Conference, Kelowna, British Columbia, Canada.

Sowter, A., Amat, M.B.C., Cigna, F., Marsh, S., Athab, A., Alshammari, L., 2016. Mexico City land subsidence in 2014–2015 with Sentinel-1 IW TOPS: Results using the Intermittent SBAS (ISBAS) technique. Int. J. Appl. Earth Obs. Geoinf. 52, 230–242 (2016). https://doi.org/10.1016/j.jag.2016.06.015.

Spencer, D.L., Hatton, C.N., Bareither, C.A., Ward, K.J., Scalia IV, J., 2021. Deck hands needed! Experience necessary – addressing the impeding qualified tailings professional resource shortage. In: Proceedings of the International Mine Waste and Tailings Conference, Brisbane, Australia.

Strouth, A., McDougall, S. 2021. Societal risk evaluation for landslides: historical synthesis and proposed tools. Landslides 18, 1071–1085. https://doi.org/10.1007/s10346-020-01547-8

Takahashi, T., 1991. The big mud flow in the Stava River, Italy. In: Tosatti, G. (Ed.), A Review of Scientific Contributions on the Stava Valley Disaster (Eastern Italian Alps), 19th July 1985. Pitagora Editrice, Bologna, pp. 169–177.

Takahashi, T., 2014. Debris flow: mechanics, prediction and countermeasures. CRC Press, London.

Tang, L., Liu, X., Wang, X., Liu, S., Deng, H., 2020. Statistical analysis of tailings ponds in China. J. Geochem. Explor. 216, 106579. https://doi.org/10.1016/j.gexplo.2020.106579.

The Portal for Natural Agriculture and the Environment, 2017. The collapse of the pool of the Rotem plant that caused the pollution of Nahal Ashlim (translated from Hebrew). https://www.youtube.com/watch?v=B-Dkm5ZHg28&list=PLKe5zo1afI66aSpWagaU2BWDPRNXFfBTv&index=4&ab_channel=%D7%94%D7%A4%D7%95%D7%A8%D7%98%D7%9C%D7%9C%D7%97%D7%A7%D7%9C%D7%90%D7%95%D7%AA%D7%98%D7%91%D7%A2%D7%95%D7%A1%D7%91%D7%99%D7%91%D7%94 (accessed 30 December 2022).

Thomas, A., Edwards, S.J., Engels, J., McCormack, H., Hopkins, V., Holley, R., 2019. Earth observation data and satellite InSAR for the remote monitoring of tailings storage facilities: a case study of Cadia Mine, Australia. In: Proceedings of the 22nd International Conference on Paste, Thickened and Filtered Tailings, Australia, pp. 183-195.

Torres, N.A., Brito, A.M., 1966. Effects of the March 1965 earthquake on the tailings dams of El Soldado mine, Province of Valparaiso (in Spanish). Bull. Inst. Geol. Investig. Chile 20, 21–42.

Troncoso, J.H., Verdugo, R., Valenzuela, L., 2017. Seismic performance of tailings sand dams in Chile. In: Proceedings of the 16th World Conference in Earthquake Engineering, Santiago, Chile.

Troncoso, J.H., Vergara, A., Avendaño, A., 1993. Seismic failure of Barahona tailings dam. In: Proceedings of the 3rd International Conference on Case Histories in Geotechnical Engineering 34, St. Louis, USA.

U.S. Army Corps of Engineers (USACE), 2019. National inventory of dams. https://nid.sec.usace.army.mil/ (accessed August 2021).

U.S. Bureau of Reclamation (USBR), 2015. Reclamation consequence estimating methodology: dam failure and flood event case history compilation (interim). U.S. Department of the Interior, Denver, USA.

U.S. Geological Survey (USGS), 2021. Mineral commodity summaries. https://www.usgs.gov/centers/national-minerals-information-center/mineral-commodity-summaries (accessed December 2021).

United Nations Environmental Programme (UNEP) and Office for the Co-ordination of Humanitarian Affairs (OCHA), 2000. Spill of liquid and suspended waste at the Aurul S.A. Retreatment Plant in Baia Mare (Romania).

Vaidya, R.A., Molden, D.J., Shrestha, A.B., Wagle, N., Tortajada, C., 2021. The role of hydropower in South Asia's energy future. Int. J. Water Resour. Dev. 37 (3), 367-391. https://doi.org/10.1080/07900627.2021.1875809.

Valenzuela, L., 2016. Design, construction, operation and the effect of fines content and permeability on the seismic performance of tailings sand dams in Chile. Obras y proyectos, 19 (6–22). http://dx.doi.org/10.4067/S0718-28132016000100001.

Van Niekerk, H.J., Viljoen, M.J., 2005. Causes and consequences of the Merriespruit and other tailings-dam failures. Land Degrad. Dev. 16 (2), 201–212. https://doi.org/10.1002/ldr.681.

Versteeg, M.F., 1988. External safety policy in the Netherlands: an approach to risk management. J. Hazard. Mater. 17(2), 215-222. https://doi.org/10.1016/0304-3894(88)80006-2.

Vick, S.G., 1983. Planning, design, and analysis of tailings dams. John Wiley & Sons, New York.

Vick, S.G., 1992. Stability evaluation during staged construction – a discussion. J. Geotech. Eng. 118 (8), 1283–1288. https://doi.org/10.1061/(ASCE)0733-9410(1992)118:8(1283).

Vick, S.G., 1996. Failure of the Omai tailings dam. Geotech. News 14, 34–40.

Villavicencio, G., Espinace, R., Palma, J., Fourie, A., Valenzuela, P., 2014. Failures of sand tailings dams in a highly seismic country. Can. Geotech. J. 51 (4), 449–464. https://doi.org/10.1139/cgj-2013-0142.

Wagener, F., 1997. The Merriespruit slimes dam failure: overview and lessons learned. J. South Afr. Inst. Civ. Eng. 39 (3), 11–15.

Wahl, T.L., 1998. Prediction of embankment dam breach parameters: a literature review and needs assessment. Dam Safety Rep. No. DSO-98-004, Bureau of Reclamation. U.S. Department of the Interior, Denver, USA.

Wahler & Associates, 1973. Analysis of Coal Refuse Dam Failure: Middle Fork Buffalo Creek, Saunders, West Virginia. U.S. Bureau of Mines, Department of the Interior.

Wang, K., Yang, P., Hudson-Edwards, K.A., Lyu, W., Yang, C., Jing, X., 2018. Integration of DSM and SPH to Model Tailings Dam Failure Run-Out Slurry Routing Across 3D Real Terrain. Water 10 (8), 1087. https://www.doi:10.3390/w10081087.

Wang, L., Deng, K., Zheng, M., 2020. Research on ground deformation monitoring method in mining areas using the probability integral model fusion D-InSAR, sub-band InSAR and offset-tracking. Int. J. Appl. Earth Obs. Geoinf. 85, 101981. https://doi.org/10.1016/j.jag.2019.101981.

Wang, L., Ji, B., Hu, Y., Liu, R., Sun, W., 2017. A review on in situ phytoremediation of mine tailings. Chemosphere 184, 594–600. https://doi.org/10.1016/j.chemosphere.2017.06.025.

Wang, Y., Bai, Z., Zhang, Y., Qin, Y., Lin, Y., Li, Y., Shen, W., 2021. Using TerraSAR X-band and Sentinel-1 C-band SAR interferometry for deformation along Beijing-Tianjin intercity railway analysis. IEEE J. Select. Topics Appl. Earth Obs. Remote Sens. 14, 4832-4841. https://doi.org/10.1109/JSTARS.2021.3076244.

Werner, C., Wegmüller, U., Strozzi, T., Wiesmann, A., 2000. Gamma SAR and interferometric processing software. In: Paper Presented at Proceedings of the ERS-Envisat Symposium (Gothenburg, Sweden).

Werner, T.T., Mudd, G.M., Schipper, A.M., Huijbregts, M.A., Taneja, L., Northey, S.A., 2020. Global-scale remote sensing of mine areas and analysis of factors explaining their extent. Glob. Environ. Change 60, 102007. https://doi.org/10.1016/j.gloplacha.2021.103578.

Williams, D.J., 2021. Lessons from Tailings Dam Failures – Where to Go from Here? Minerals, 11(8), 853. https://doi.org/10.3390/min11080853.

World Information on Service Energy (WISE), 2020. Chronology of major tailings dam failures (online database). https://www.wise-uranium.org/mdaf.html (accessed 25 July 2020).

World Mine Tailings Failures (WMTF), 2020. Estimate of World Tailings Portfolio 2020. https://worldminetailingsfailures.org/estimate-of-world-tailings-portfolio-2020/ (accessed October 2021).

Xiao, R., Jiang, M., Li, Z., He, X., 2022. New insights into the 2020 Sardoba dam failure in Uzbekistan from Earth observation. Int. J. Appl. Earth Obs. Geoinform. 107, 102705. https://doi.org/10.1016/j.jag.2022.102705.

Xu, Y., Zhang, L., Jia, J., 2008. Lessons from catastrophic dam failures in August 1975 in Zhumadian, China. In: Proceedings of Geocongress 2008: Geosustainability and geohazard mitigation (pp. 162-169).

Yan, Y., Doin, M. P., Lopez-Quiroz, P., Tupin, F., Fruneau, B., Pinel, V., Trouvé, E., 2012. Mexico City subsidence measured by InSAR time series: Joint analysis using PS and SBAS approaches. IEEE J. Select. Topic. Appl. Earth Obs. Remote Sens. 5(4), 1312-1326. https://doi.org/10.1109/JSTARS.2012.2191146.

Yang, S.L., Milliman, J.D., Xu, K.H., Deng, B., Zhang, X.Y., Luo, X.X., 2014. Downstream sedimentary and geomorphic impacts of the Three Gorges Dam on the Yangtze River. Earth-Sci. Rev. 138, 469-486. https://doi.org/10.1016/j.earscirev.2014.07.006.

Yang, Y., Wei, Z., Chen, Y.L., Li, Y., Li, X., 2017. Utilizing phosphate mine tailings to produce ceramisite. Constr. Build. Mater. 155, 1081-1090. https://doi.org/10.1016/j.conbuildmat.2017.08.070.

Yasuda, S., Towhata, I., Ishii, I., Sato, S., Uchimura, T., 2013. Liquefaction-induced damage to structures during the 2011 great east Japan earthquake. J. Jpn. Soc. Civ. Eng. 1 (1), 181–193. https://doi.org/10.2208/journalofjsce.1.1_181.

Yu, S., Zhang, Q., Chen, Z., Hao, J., Wang, L., Li, P., Zhong, Q., 2021. Study of the Sheyuegou dam breach–Experience with the post-failure investigation and back analysis. Eng. Fail. Anal. 125, 105441. https://doi.org/10.1016/j.engfailanal.2021.105441.

Yunjun, Z., Fattahi, H., Amelung, F., 2019. Small baseline InSAR time series analysis: Unwrapping error correction and noise reduction. Comput. Geosci. 133, 104331. https://doi.org/10.1016/j.cageo.2019.104331.

Zhang, L., Peng, M., Chang, D., Xu, Y., 2016. Dam failure mechanisms and risk assessment. John Wiley & Sons, Singapore.

Zhong, Q., Wang, L., Chen, S., Chen, Z., Shan, Y., Zhang, Q., Ren, Q., Mei, S., Jiang, J., Hu, L., Liu, J., 2021. Breaches of embankment and landslide dams – State of the art review. Earth-Sci. Rev. 216, 103597. https://doi.org/10.1016/j.earscirev.2021.103597.

Zhong, Q.M., Chen, S.S., Fu, Z.Z., Shan, Y.B., 2020. New empirical model for breaching of earth-rock dams. Nat. Hazards Rev. 21 (2), 06020002. https://doi.org/10.1061/(ASCE)NH.1527-6996.0000374.

Zhuang, Y., Jin, K., Cheng, Q., Xing, A., Luo, H., 2021. Experimental and numerical investigations of a catastrophic tailings dam break in Daye, Hubei, China. Bull. Eng. Geol. Environ. 81(1), 1-16. https://doi.org/10.1007/s10064-021-02491-0.

Zmemla, R., Sdiri, A., Naifar, I., Benjdidia, M., Elleuch, B., 2020. Tunisian phosphogypsum tailings: Assessment of leaching behavior for an integrated management approach. Environ. Eng. Res. 25(3), 345-355. https://doi.org/10.4491/eer.2019.046.

Zubkova, O., Alexeev, A., Polyanskiy, A., Karapetyan, K., Kononchuk, O., Reinmöller, M., 2021. Complex Processing of Saponite Waste from a Diamond-Mining Enterprise. Appl. Sci. 11(14), 6615. https://doi.org/10.3390/app11146615.

Appendix

Table A.1: Time-lapse videos showing the evolution of the TSF failure sites until their breach events (referred to in Chapter 4.4).

TSF failure case	**Number of PlanetScope images**	**Timespan**	**Open-access URL link**
2017 Tonglvshan, China	N/A	N/A	N/A
2017 Mishor Rotem, Israel	80	Jan 2016 – July 2017	https://www.planet.com/stories/mishor-rotem-tsf-evolution-and-breach-WjEJhzoVR
2018 Cadia, Australia	120	Sept 2016 – Sept 2018	https://www.planet.com/stories/cadia-tsf-evolution-and-breach-AZZeTRo4g
2018 Cieneguita, Mexico	60	Jan 2017 – July 2018	https://www.planet.com/stories/cieneguita-mexico-tsf-evolution-and-collapse-snnoZ62Vg
2019 Feijao, Brazil	90	June 2017 – Feb 2019	https://www.planet.com/stories/feijao-tsf-failure-_57SR6h4g
2019 Hindalco, India	120	Sept 2017 – Sept 2019	https://www.planet.com/stories/hindalco-tsf-evolution-zM7Se32Vg
2019 Cobriza, Peru	80	Sept 2017 – July 2019	https://www.planet.com/stories/cobriza-peru-tsf-evolution-and-breach-4bv3Me2VR