

# UC Berkeley

## UC Berkeley Previously Published Works

### Title

Dynamics of severe accidents in the oil & gas energy sector derived from the authoritative ENergy-related severe accident database

### Permalink

<https://escholarship.org/uc/item/1qz0f6br>

### Journal

PLOS ONE, 17(2)

### ISSN

1932-6203

### Authors

Mignan, Arnaud  
Spada, Matteo  
Burgherr, Peter  
[et al.](#)

### Publication Date

2022

### DOI

10.1371/journal.pone.0263962

### Copyright Information

This work is made available under the terms of a Creative Commons Attribution License, available at <https://creativecommons.org/licenses/by/4.0/>

Peer reviewed

## RESEARCH ARTICLE

# Dynamics of severe accidents in the oil & gas energy sector derived from the authoritative ENergy-related severe accident database

Arnaud Mignan<sup>1,2,3\*</sup>, Matteo Spada<sup>3,4</sup>, Peter Burgherr<sup>3</sup>, Ziqi Wang<sup>5</sup>, Didier Sornette<sup>1,6</sup>

**1** Institute of Risk Analysis, Prediction and Management (Risks-X), Academy for Advanced Interdisciplinary Studies, Southern University of Science and Technology (SUSTech), Shenzhen, China, **2** Department of Earth and Space Sciences, Southern University of Science and Technology (SUSTech), Shenzhen, China, **3** Laboratory for Energy Systems Analysis, Paul Scherrer Institute (PSI), Villigen, Switzerland, **4** Institute of Sustainable Development, Zurich University of Applied Sciences (ZHAW), Winterthur, Switzerland, **5** Department of Civil and Environmental Engineering, University of California, Berkeley, California, United States of America, **6** Department of Management, Technology and Economics, Swiss Federal Institute of Technology (ETH) Zurich, Zurich, Switzerland

\* [mignana@sustech.edu.cn](mailto:mignana@sustech.edu.cn)



## OPEN ACCESS

**Citation:** Mignan A, Spada M, Burgherr P, Wang Z, Sornette D (2022) Dynamics of severe accidents in the oil & gas energy sector derived from the authoritative ENergy-related severe accident database. PLoS ONE 17(2): e0263962. <https://doi.org/10.1371/journal.pone.0263962>

**Editor:** Giuseppe Vitiello, University of Naples Federico II, ITALY

**Received:** March 4, 2021

**Accepted:** January 31, 2022

**Published:** February 17, 2022

**Copyright:** © 2022 Mignan et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

**Data Availability Statement:** An archived folder of accident chains-of-events is available at the following public repository: <http://dx.doi.org/10.13140/RG.2.2.23051.41767>.

**Funding:** A.M. received funding from the National Natural Science Foundation of China (grant no. 42074054) as well as from the Swiss Competence Center for Energy Research - Supply of Electricity (SCCER SoE, <http://www.sccer-soe.ch>). P.B. and M.S. acknowledge additional support from the Future Resilient Systems (FRS) I program at the

## Abstract

Organized into a global network of critical infrastructures, the oil & gas industry remains to this day the main energy contributor to the world's economy. Severe accidents occasionally occur resulting in fatalities and disruption. We build an oil & gas accident graph based on more than a thousand severe accidents for the period 1970–2016 recorded for refineries, tankers, and gas networks in the authoritative ENergy-related Severe Accident Database (ENSAD). We explore the distribution of potential chains-of-events leading to severe accidents by combining graph theory, Markov analysis and catastrophe dynamics. Using centrality measures, we first verify that human error is consistently the main source of accidents and that explosion, fire, toxic release, and element rupture are the principal sinks, but also the main catalysts for accident amplification. Second, we quantify the space of possible chains-of-events using the concept of fundamental matrix and rank them by defining a likelihood-based importance measure  $\gamma$ . We find that chains of up to five events can play a significant role in severe accidents, consisting of feedback loops of the aforementioned events but also of secondary events not directly identifiable from graph topology and yet participating in the most likely chains-of-events.

## Introduction

Global affairs have become increasingly complex, with intertwined networks known to be vulnerable to cascading failures [1, 2]. Triggers can be natural perils, such as earthquakes, cyclones, or floods [3, 4], or anthropogenic in nature, such as events related to global warming [5]. Of all existing lifeline networks, energy flows (especially of the oil & gas sector) are the most critical for the good functioning of our modern society. Major energy accidents often have large financial consequences, whose losses are combinations of lost product, repair costs,

Singapore-ETH Centre (SEC), which was established collaboratively between ETH Zurich and Singapore's National Research Foundation (FI 370074011) under its Campus for Research Excellence and Technological Enterprise programme.

**Competing interests:** The authors have declared that no competing interests exist.

but also legal fees and fines by government agencies. For example, the 2005 BP Texas City refinery explosion killed 15 workers, injured 170, and led to financial losses of about \$1.5 billion [6, 7]. Some of those events are natural-technological (or Natech) accidents: The 2011 Tohoku earthquake and associated great tsunami, mainly remembered for causing the failure of the Fukushima nuclear plant, also shut almost a third of the Japanese refining capacity. Accidents at critical infrastructures can lead to toxic material release, business interruption, and some economic slowdown [8]. The extensive flooding due to record rainfall during the 2017 Hurricane Harvey damaged more than 40 industrial plants and triggered over 100 toxic spillage events [9].

The study of energy accidents has a long history [10–15] and reflects a high energy security risk [16]. Of existing accident data inventories, the Major Energy Accidents (MEA) database contains 279 major energy accidents spanning over the course of the 20<sup>th</sup> and early 21<sup>st</sup> centuries [17]. As for the ENergy-related Severe Accident Database (ENSAD), it lists ~23,500 unique accident records between 1970 and 2016, of which 4,067 are considered severe ( $\geq 5$  fatalities) [18, 19]. The two databases have been compared by Felder [12]. However, the dynamics of recorded accidents has yet to be better understood [20]. Energy models still generally do not allow for a deep exploration of extreme events and emergent endogenous phenomena [21], which consist of risk-amplifying cascading effects observed and already modelled in other complex systems [22–24].

Here, we explore the chains-of-events leading to severe accidents in the oil & gas industry using a data-driven approach based on ENSAD, which has been described as the most authoritative resource for comparative risk analysis of accidents in the energy sector [11, 13, 19]. This is the first study to apply graph theory and catastrophe dynamics modelling [25–28] to ENSAD, allowing us to describe the general topological properties of severe accidents at refineries [29], oil tankers [30] and gas networks [31], based on a rich database of more than a thousand events, spanning from 1970 to 2016, that includes information on the chains-of-events that led to these accidents. By describing severe accidents at critical infrastructures (CIs) with dependency risk graphs, we can first identify their main sources, catalysts, and sinks. Thanks to ENSAD's volume of data, we can then go beyond existing studies [32, 33] by exploring interactions between 27 different events and offer a novel approach for the systematic, data-driven, exploration of events that amplify severe accidents. We use the concept of fundamental matrix from Markov chain theory recently introduced to catastrophe dynamics [34] to explore the full space of potential chains-of-events. We also define a likelihood-based importance measure to identify the most critical combinations of transitional events. Although Markov models are commonly used in process accident and cascading effect modelling, they are generally concerned with critical events within the chains [35] or with the final state [25], not with the dynamics of the chains themselves that might provide richer information for energy risk mitigation.

## The ENSAD database

We use the ENergy-related Severe Accident Database as sole input to our study [11, 13]. Severe accidents are defined in ENSAD as events leading to five or more fatalities, ten or more injuries, at least 200 evacuees, release of hydrocarbons exceeding 10,000 tons, enforced clean-up of land and water over areas of 25 km<sup>2</sup> or greater, or an economic loss equal or greater to \$5 million USD (inflation adjusted) [36]. Chains-of-events are extracted for the CIs part of the oil & gas sector. We considered 1,187 severe accidents in refineries (674), tankers (263), and natural gas networks (250). Offshore platforms for oil were not included for lack of readily available information on chains-of-accidents. However, the natural gas network includes the entire

energy chain, from exploration and production, storage and distribution, and power plants. While the distribution of accidents is worldwide for refineries and gas networks, for tanker accidents, only those in the Mediterranean Sea were considered (one of the most active and accident-prone area in the world) [30].

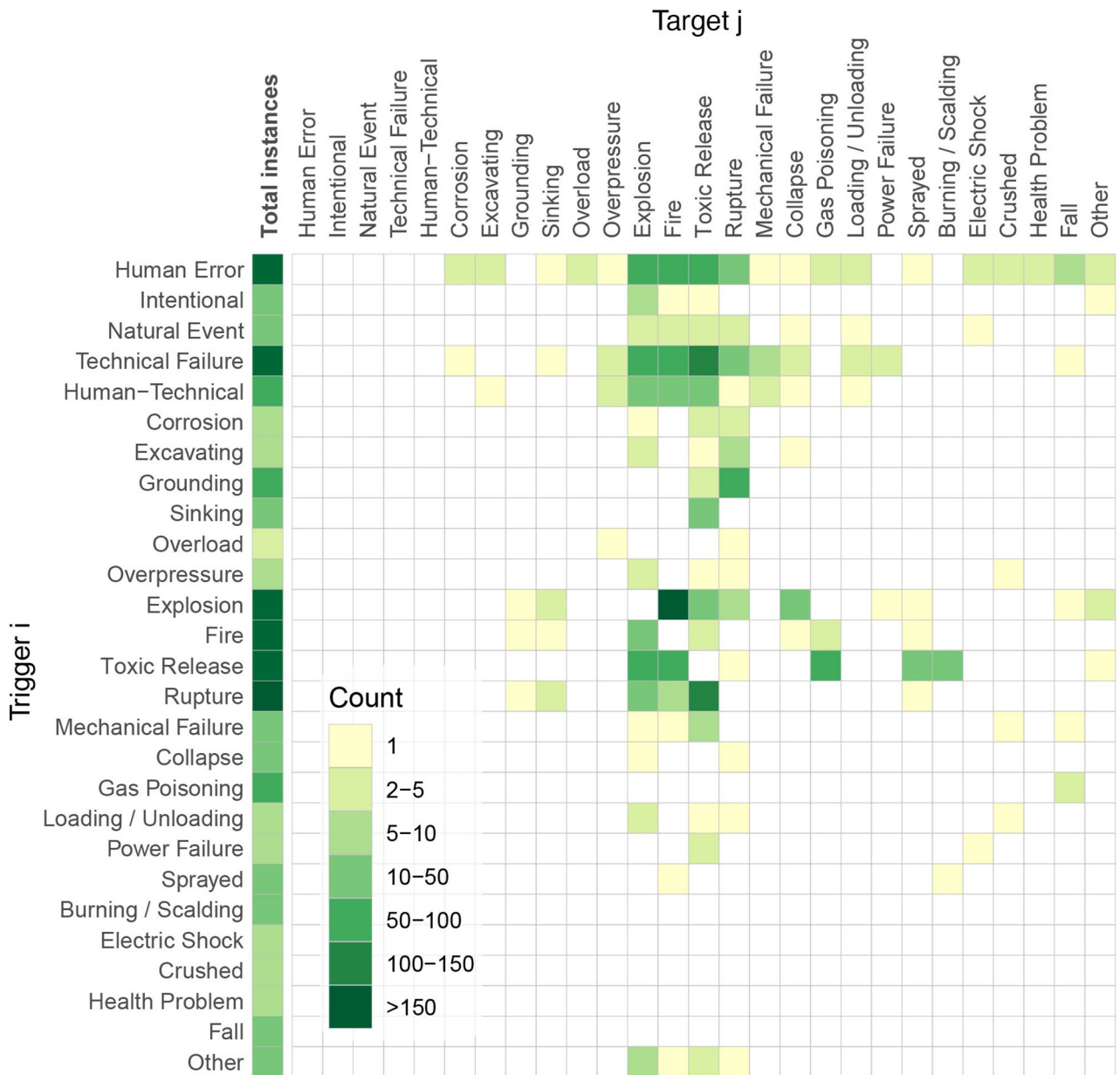
One important attribute of ENSAD, which has so far not been analysed, is the description of the chain-of-events leading to a severe accident. The goal of this study is to remediate to this omission. Each chain-of-events is composed of 1 to 5 events, up to 6 for refineries in which case the initial trigger attribute is separate from the rest of the chain-of-events by construction. The number of possible events was originally 30 for refineries, 12 for tankers and 12 for gas networks, merged into 27 events for the 3 CI types combined. Chains-of-events here represent causality chains between events participating to a severe accident, not limited to loss-generating events. Successive (loss-generating) accidents are often referred to as cascading effects or domino effects [32]. Our study is more closely related to fault tree analysis [37] although we can distinguish between both types of events. Indeed, the sub-accidents that generate losses are all the events that can act as a sink in ENSAD (see below).

In a preliminary step to chain-of-event encoding, we combined the events which were seldom observed in ENSAD into more representative categories for frequentist analysis. All natural-hazards (heavy rain, flood, landslide, earthquake, tsunami, lightning, (wind)storm, heatwave, and unknown natural triggers) were combined under the term Natural Event; machinery failure (1) and loss of power (1) were included in Technical Failure; personal health problem (3), radiation (1) and food poisoning (1) were combined in Health Problem; theft/scooping (1) (but not intentionally leading to an accident) and collision (37) were included in Human Error (if not already declared as Natural Event, e.g. due to fog). Several terms were also shortened to improve graph readability: 'Spill / Release / Leakage' to Release (which can be toxic and/or inflammable), 'Rupture / Crack / Break Up' to Rupture, 'Overflow / Overfitting / Overload' to Overload, 'Asphyxiation / Gas Poisoning / Suffocating' to Gas Poisoning, 'Sprayed / Sprinkled / Wetted' to Sprayed, and 'Crushed / Pushed / Pinned' to Crushed. It should be noted that ENSAD does not provide detailed causal factors for human errors, in contrast to the ones provided in the HFACS-OGI framework [38]. This is due to the lack of information retrievable for some accidents. Human error here implicitly includes violation of safety rules, violation of guidelines, management failure, *etc.* Fig 1 shows the total count  $n_i$  of cases per event type  $i$  as well as the count  $n_{ij}$  of one-to-one interactions, based on which the conditional probabilities to be used as input in section 3 are estimated with  $p_{ij} = n_{ij}/n_i$ . This frequentist approach provides a straightforward and informative, data-driven, inference.

It must be noted that our approach is subjected to the completeness, homogeneity, and uncertainties of ENSAD, as in any data-driven analysis. However, ENSAD has been shown to be more careful in its methodology compared to other available databases [12]. To avoid under-sampling, rare accident associations were preliminary merged into more representative categories, as described in the previous paragraph. The problem of heterogeneity between different CIs and the potential lack of comparability mentioned in the energy accident literature [12] is shown to be unfounded when considering chains-of-events attributes, as different oil & gas CIs share a number of hazardous elements. We will demonstrate below that refineries, tankers, and gas networks—despite their specificities—display similar accident network properties as when they are combined in one unique graph.

## Chain-of-event encoding & modelling

**Graph definition.** As a first step in our analysis, we encode the chains-of-events given in ENSAD into a  $27 \times 27$  adjacency matrix  $A$  where each element  $a_{ij}$  is the conditional frequentist

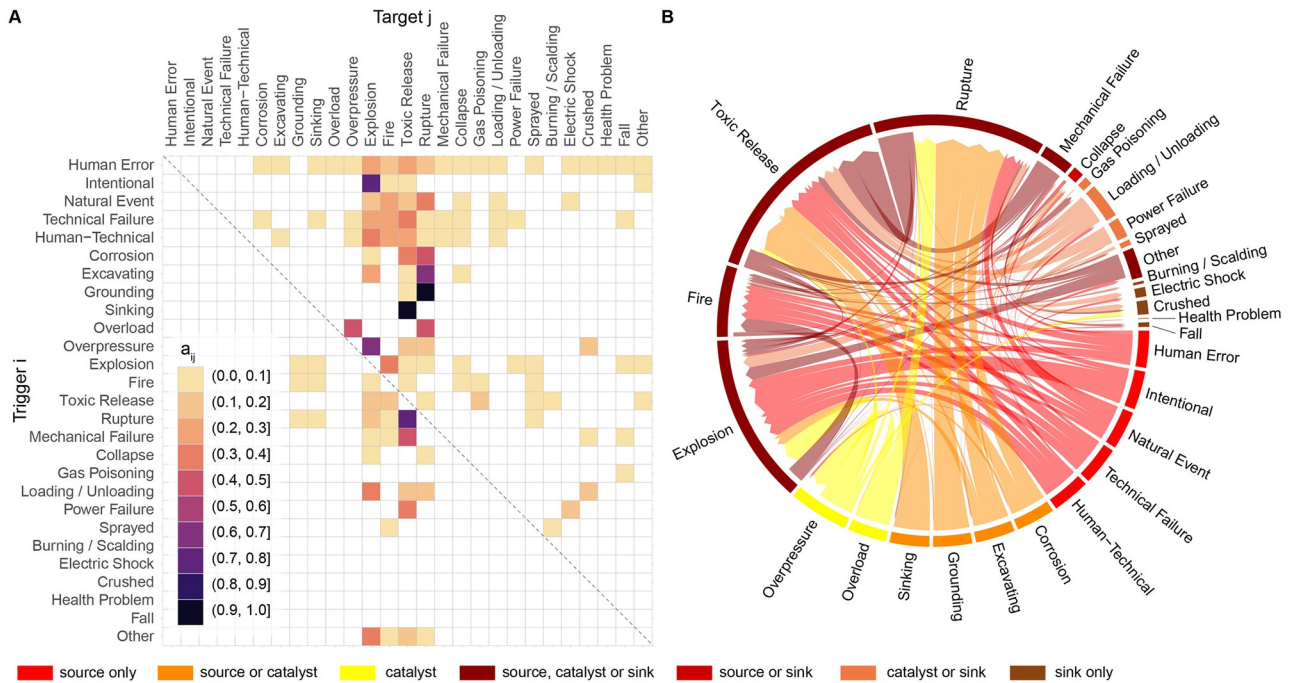


**Fig 1. Event counts, total versus conditional in ENSAD for the oil & gas energy sector, for a total of 1,187 severe accidents.** Considered critical infrastructures (CIs) are refineries, tankers, and gas networks.

<https://doi.org/10.1371/journal.pone.0263962.g001>

probability of event  $j$  given the occurrence of event  $i$ . The result—derived from the data of Fig 1—is shown here in Fig 2A. This accident interaction system is further represented by a chord diagram in Fig 2B to make the distinction between source events, catalysts and sink events, and by a directed weighted graph in Fig 3 for centrality measure analysis. We merged the causal networks of the different CI systems by assuming that similar CI elements play similar engineering roles at a generic level. This hypothesis will be shown to be reasonable in section 4 (with further plots given as S1 File).

We consider the three principal measures of centrality as explanatory metrics of the chains-of-events in severe accidents: degree (in- and out-degree), betweenness, and closeness [39].



**Fig 2. One-to-one interaction system leading to severe accidents in the energy sector.** (A) Adjacency matrix  $A$  with one-to-one conditional frequentist probabilities derived from ENSAD (refineries, tankers, and gas networks combined); (B) Chord diagram representing the complexity of the directed chains-of-events, as encoded in the adjacency matrix. Note that sub-accidents (i.e., loss generating events) are all the events that can act as a sink in ENSAD.

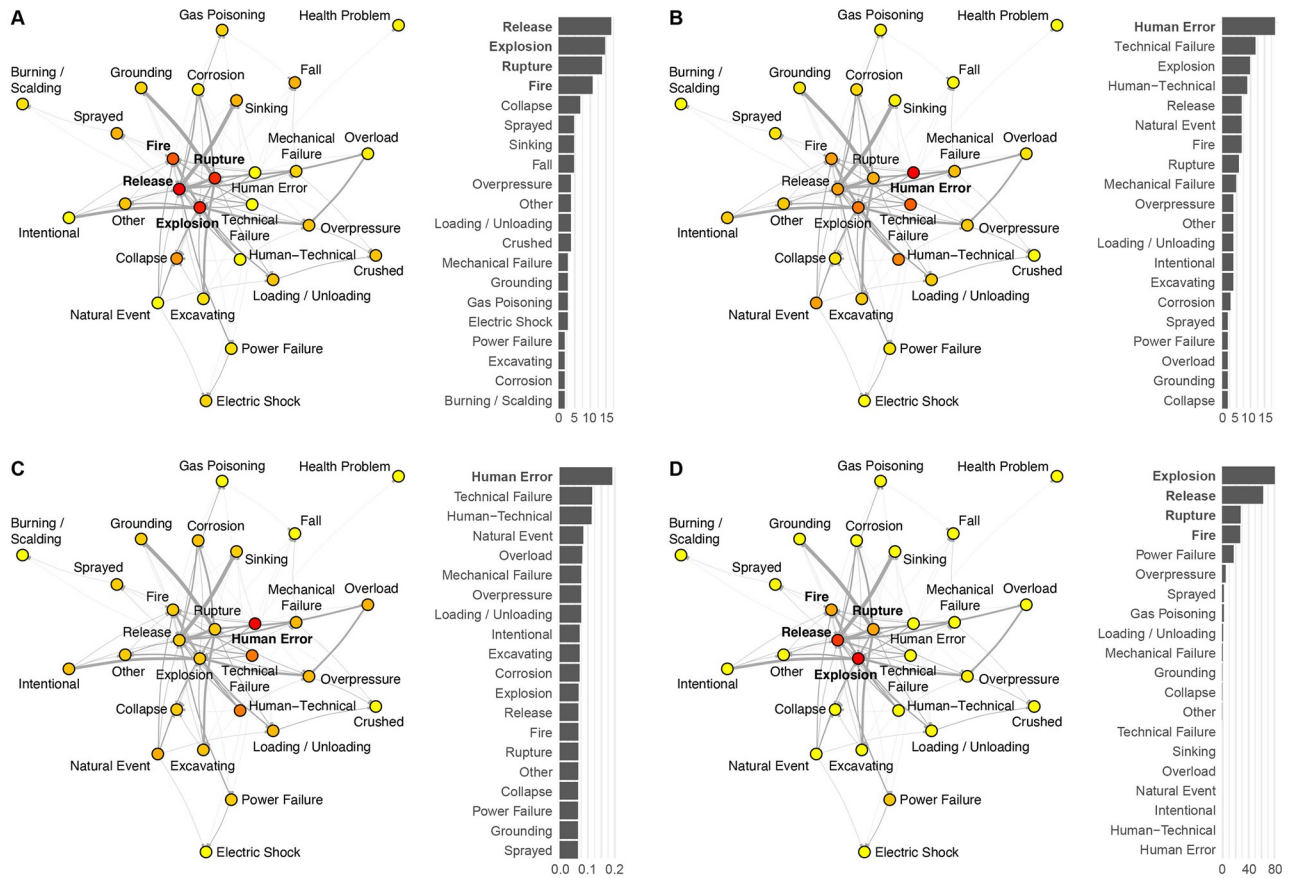
<https://doi.org/10.1371/journal.pone.0263962.g002>

Degree centrality of a vertex is the number of edges incident on the node; closeness centrality of a vertex is the average of the shortest path lengths from the vertex to all other vertices in the graph; betweenness centrality of a vertex is the number of the shortest paths that pass through that vertex. In catastrophe dynamics modelling [26–28], the main triggering events (or sources) are represented by high out-degree and closeness centralities. The main triggered events (or sinks) are represented by a high in-degree centrality. Catalysts of cascading effects are represented by a high betweenness centrality.

**Markov chain theory.** With severe accidents defined as the succession of individual events that interact with each other, we may assume that the dynamical process is described by the following linear differential equation

$$\frac{d\vec{X}(t)}{dt} = \mathbf{Q}\vec{X}(t) \tag{1}$$

in which the dynamics of the vector  $\vec{X}$  of events is described by the transition rate matrix  $\mathbf{Q}$  that encodes all possible one-to-one direct interactions. The propagator for  $\vec{X}(t) \rightarrow \vec{X}(t + \tau)$  can be written as  $\mathbf{P}(\tau) = \exp(\mathbf{Q}\tau)$  with the entries  $p_{ij}(\tau)$  interpreted as the conditional probability of event  $j$  being triggered by event  $i$  for a period of  $\tau$  and with  $\mathbf{A} := \mathbf{P}(\tau = 1)$ . We define an outflow event as a last (implicit) matrix entry for conservation of probabilities with conditional probability  $1 - \sum_{j=1}^n p_{ij}$ . We also fix the last row of  $\mathbf{A}$  as  $p_{n+1} = (0, \dots, 0, 1)$ , i.e. the outflow event does not trigger any real event but itself, which is an absorbing state in Markov chain parlance. These 28<sup>th</sup> column and row are not displayed on the adjacency matrix plots for



**Fig 3. Centrality measures of the severe accident chains-of-events.** (A) In-degree centrality; (B) out-degree centrality; (C) closeness centrality; (D) betweenness centrality. Influential nodes (or hazardous events) are highlighted in bold per centrality measure. Yellow-to-red colours of the vertices represent increasing centrality estimates, following the scales and ranking shown in the histograms.

<https://doi.org/10.1371/journal.pone.0263962.g003>

convenience. As so, we do not directly apply Eq (1) but operate with the adjacency matrix, as shown in the next subsection.

Note that the proposed approach has two main limitations: first, the system dynamics considers one-to-one interactions (reflected in Eq 1) and not triplets or higher-order interactions although triggering may be caused or enhanced by the combination of multiple variables; second, the true temporal evolution of a chain-of-events is not modelled, only the ranks or succession of events with their cumulative count acting as an effective time. Eq 1 provides a first-order description of the process and the rate matrix, defined from the adjacency matrix, is directly encodable from ENSAD data. Including *n*-to-one interactions would require physical modelling which is out of the scope of the present empirical study. Moreover, the memorylessness property of Markov theory is a very common and well accepted assumption. Finally, the chains-of-events generated in the present study are implicitly assumed to occur in a natural physical time scale that is simplified into the consecutive trials in a Markov chain. Inclusion of common time intervals between different types of events would require details that are not available in the ENSAD.

**Fundamental matrix definition.** While the analysis of the matrix power  $A^T$  would suffice to explore the full space of possible interactions in a topological sense, we use the concept of fundamental matrix *N* to capture the frequency, or intensity, of each transition event before

resting at the absorbing state:

$$N := (I - A)^{-1} = I + \sum_{\tau=1}^{\infty} A^{\tau} = I + A + A^2 + \dots = I + M \tag{2}$$

where the entry  $n_{ij}$  is the expected number of times the chain-of-events reaches event  $j$  given that the accident starts with event  $i$ ;  $I$  is the identity matrix,  $\tau$  the number of pseudo-time steps and  $M$  the interaction matrix [34], which excludes the step  $\tau = 0$ . Eq 2 describes all possible combinations of chains-of-events via the matrix power  $A^{\tau}$  including amplification via the sum operator. Elements  $m_{ij}(\tau)$  are interpretable as the total set of scenarios leading to event  $j$  for a chain-of-events starting from event  $i$  and of length 1 to  $\tau$ .

This study is the second application of the fundamental matrix to catastrophe dynamics. The first one, by Mignan and Wang [34], analysed the role of adjacency matrix topology on the matrix  $M$ . It described the non-trivial behaviour of cascading patterns in terms of the space of possibilities covered, and of interaction amplification by feedback loops. That study further tested the approach on a small database of 29 historical catastrophes known for their domino effects. The present work is the first application of the fundamental matrix to the energy sector and to such a large database (more than one thousand accidents from ENSAD). Notice that the approach proposed earlier by Helbing and Kühnert [25] models the final state that is equivalent to  $A^{\tau}$ , but based on the rate matrix  $Q$ . In contrast, the interaction  $M$  provides some information on all the transition states of the chain.

**Importance measure definition.** For the purpose of risk mitigation, an importance measure of a chain of  $m$  events  $Z(m) = (z_1, z_2, \dots, z_m)$  is defined using the Markov property. It is derived from the likelihood function

$$\mathcal{L}(z_1, z_2, \dots, z_m) = \prod_{i=1}^m p(z_i|z_{i-1}) \tag{3}$$

where we set  $p(z_1|z_0) \equiv p(z_1)$ . The conditional probability can be directly read from  $A$ , i.e.  $p(z_i|z_{i-1}) = a(z_{i-1}, z_i)$ . Since any chain that has not yet reached the absorbing state has the potential to further grow, it is useful to define the concept of a *complete chain* as any chain that rests at the absorbing state. Formally, let the absorbing state be denoted by  $z_a$ , a chain  $(z_1, z_2, \dots, z_m)$  is said to be complete iff  $z_m = z_a$ , and  $z_i \neq z_a$  for  $i \in \{1, 2, \dots, m-1\}$ . An importance measure for a complete chain with  $m$  events can then be defined as:

$$\gamma(z_1, \dots, z_{m-1}, z_a) = \frac{\mathcal{L}(z_1, \dots, z_{m-1}, z_a)}{\sum_{Z'(m) \in \sigma_z(m)} \mathcal{L}(Z'(m))} = \frac{\prod_{i=1}^a a\{z_{i-1}, z_i\}}{\sum_{Z'(m) \in \sigma_z(m)} \mathcal{L}(Z'(m))} \tag{4}$$

where  $\mathcal{L}$  is the likelihood function,  $z_m = z_a$  the absorbing (final) state, and  $\sigma_z(m)$  the set of all possible complete chains with  $m$  events. Note that  $\gamma$  can be interpreted as the probability of a specific chain among chains-of- $m$ -events. It can thus be used for severe accident ranking.

### ENSAD application

The ENSAD accident interaction encoding is first described in Figs 1 and 2. Counts of one-to-one interactions go from one instance for the rarest interactions in ENSAD to more than 150 in the case of the most common triggering process, which is an explosion triggering a fire (Fig 1). We observe a rich variety of 117 potential one-to-one interactions. Since most accidents considered in ENSAD are described as chains-of-events, the conditional frequentist probabilities  $a_{ij}$  can be relatively high, with 36% of cases above 0.1 (Fig 2). Two interactions have a probability greater than 0.9 of occurring, which are grounding (i.e., impact to seabed or waterway side, also known as stranding) to rupture, and tanker sinking to toxic material release. All interactions are constrained by the underlying assumption that an accident is severe in



ENSAD. Hence, a tanker sinking implicitly means that a toxic spill follows for the accident to be severe. The same rule applies for all other interactions.

The ENSAD already contains some examples of long historical chains-of-events [19]. For example, in the case of refineries, one of the worst industrial disasters in recent U.S. history happened on the 23<sup>rd</sup> of March 2005. On that day, at the BP Texas City Refinery (USA), due to an error, a raffinate splitter tower has been overfilled. Once the isomerization (ISOM) unit was started, the overfill caused the opening of the pressure relief devices, which resulted in a flammable liquid geyser from a blowdown stack that was not equipped with a flare. The release of flammables led to an explosion followed by fires. The latter killed 15 workers located close to the source of the accidents and injured another 180, alarmed the community (with 43,000 people ordered to remain indoors), and resulted in financial losses exceeding \$1.5 billion, due to the damages to the refinery and the houses in a radius of 750 meters from it. Other examples include (i) a release followed by fire from a depropanizer, which caused the injury of a worker at the Alkylation unit of the Delaware City refinery (USA) in 2015, (ii) an explosion followed by fire, which caused the injury of 6 workers during maintenance of a tank at the Aspropyrgos refinery in Greece in 2015.

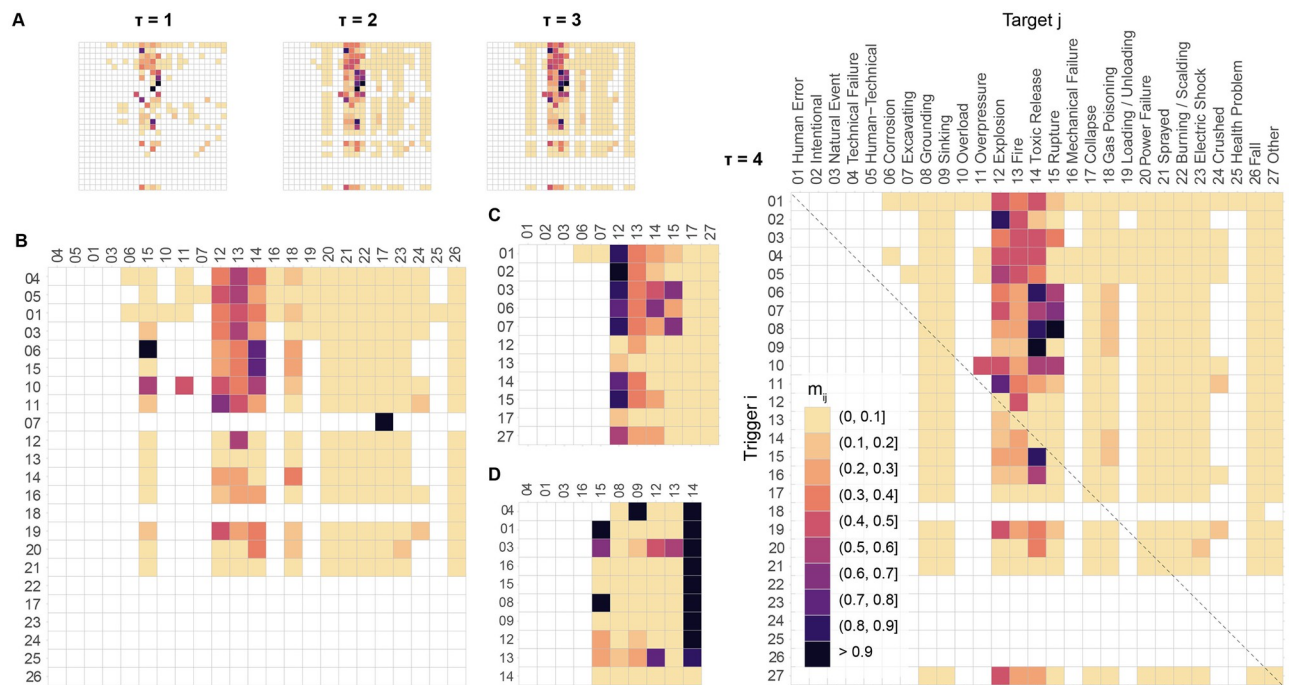
In the case of gas networks, an example taken from ENSAD is the accident that happened 50 km south-west from Brussels (Belgium) in 2004. In an excavation process during a construction work, a scraper ruptured a major underground high-pressure natural gas pipeline. Immediately, fire fighters and police rushed to secure the area, but an explosion occurred followed by fire, which killed 17 persons and injured 120. Other examples are the explosion of a pipeline caused by corrosion and the consequent release of natural gas, followed by fires that killed 22 people and injured 40 in India in 2014; the explosion of a CNG station caused by an intentional attack, which caused 25 fatalities and 100 injuries in Pakistan in 2011. Some of those chains-of-events are not limited to gas networks. Some events are however associated to CI-specific elements, such as excavation when pipelines are involved.

For tankers, one of the worst accidents in the Mediterranean Sea happened near the port of Genoa in 1991. The Cypriot oil tanker Haven, anchored off Genoa, loaded with 144,000 tons of crude oil, caught fire, exploded, and broke into three parts. One of them sank on the spot, the other parts were taken hold of and sank during their towing, releasing the entire loaded crude oil in the sea. In addition, the accident caused 6 fatalities and 30 injuries among the people on the tanker. Other examples are the Ottoman Integrity tanker, which released 3300 tons of crude oil during shipping in the Peloponnesus area in Greece in 2015; a tanker grounded at the port of Tarragona (Spain), which sank and released the remaining crude oil in the vessel (around 25 tons) in 2008. All those examples were encoded as one-to-one interactions in the previous section but can be retrieved via Markov theory.

Here, we generalise the process of interactions by applying the above theory to the ENSAD. Based on the adjacency matrix shown in Fig 2, centrality measures are computed and plotted on the matching graphs in Fig 3. They are then ranked in histograms to identify the main sources, sinks and catalysts of severe accidents. The main triggers (or sources) are found to be human error followed by technical failure as they have the highest out-degree and closeness centralities. We also observe that the main triggered events (or sinks), represented by a high in-degree centrality, are toxic (including inflammable) release, explosion, element rupture and fire. These events are also the four main catalysts for cascading effects, being also represented by the highest betweenness centralities. These results are consistent across CI types, supporting our approach to combine all CIs in one graph. Figures similar to Fig 3 but for the individuals CIs are given as S1 File. S1 Fig in S1 File shows the accident network and matching centrality measures for refineries. Results are very similar to the combined oil & gas CI network since this class is dominating the full dataset with 57% of cases. S2 Fig in S1 File shows the same

plots for tankers. For this CI, sinking replaces explosion in the top sinks (highest in-degree centrality), the other sinks remaining the same (release, rupture, fire). Human error remains the dominant initial trigger, the only event being at the highest rank for both out-degree and closeness centralities. Similar catalysts are found except again for sinking replacing explosion (highest betweenness centrality). S3 Fig in S1 File, finally, shows the results for gas networks. The accident network topology is similar to the one for refineries and all oil & gas CIs combined. Quantified by a globally comprehensive data source, our results confirm the existing literature that most accidents are due to human errors [38, 40–42] but also that fire, explosion, and toxic release are the major hazards in the industry [43, 44], as well as the main sources for further domino effects [32].

To go beyond the identification and ranking of individual critical events by centrality measures, we investigate the potential amplification of chains-of-events via the interaction matrix  $M$  previously defined (Eq 2) and plotted in Fig 4. We are here concerned with event  $j$  being a sub-accident, meaning a loss-generating sink event, as defined by the colour scheme in Fig 2B. We find that the dynamical process stabilises rapidly, around  $\tau = 4$ , and expands the space of possibilities significantly compared to direct interactions  $a_{ij} = m_{ij} (\tau = 1)$ . This is captured in Fig 4A in which we see the number of possible interactions (i.e., non-zero cells) increase at each step to then remain constant at  $\tau = 4$ . The values  $m_{ij}$  can still evolve at this stage but changes are dwarfed relative to the summation of the previous steps, which is due to the fact that the matrix power  $A^\tau$  tends to zero for increasing  $\tau$ . Of all trigger events (27 minus 5 sink-only) and triggerable events (27 minus 5 source-only), 64% (14/22) of triggerable events can be triggered, directly or indirectly, by 95% (21/22) of trigger events. This suggests the potential for surprising accidents that were never experienced before. Longer chains-of-events ( $\tau > 4$ ) can be considered insignificant as their contribution, terms  $A^\tau$  for  $\tau > 4$  in Eq 2, becomes negligible. The interaction matrix  $M$  also shows which accidents  $j$  are most prone to occur due to

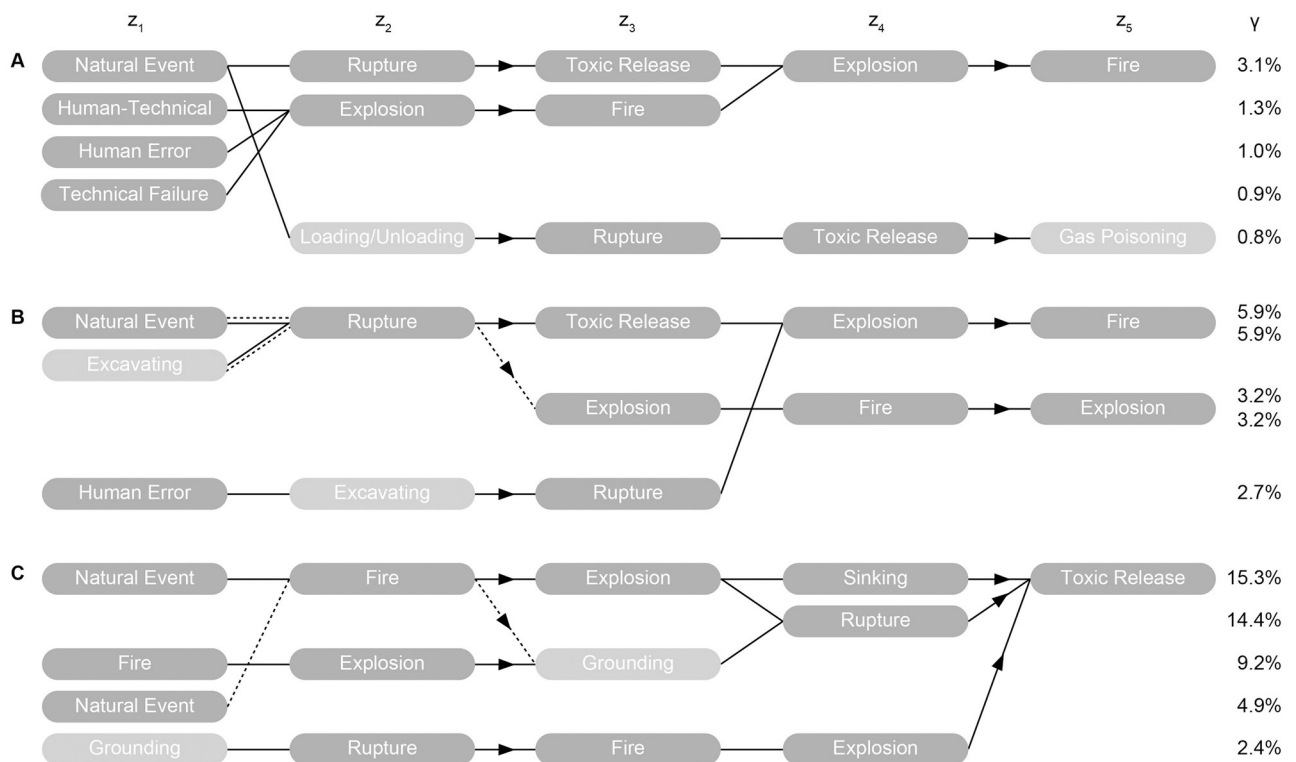


**Fig 4. Interaction matrix  $M$  encoding chains-of-events of length up to  $\tau$ .** (A) All CIs combined with  $M$  evolving from  $\tau = 1$  (i.e., adjacency matrix  $A$ ) to  $\tau = 4$ . (B)  $M(\tau = 4)$  for refineries. (C)  $M(\tau = 4)$  for gas networks. (D)  $M(\tau = 4)$  for tankers.

<https://doi.org/10.1371/journal.pone.0263962.g004>

a multiplicity of chains-of-events, turning towards purple-to-black as  $\tau$  increases. Those events are again fire, explosion, toxic release, and rupture, which is consistent with our previous measures of high betweenness and in-degree centralities. Those patterns are more pronounced when Eq 2 is applied to individual CIs as shown in Fig 4B–4D for refineries, gas networks and tankers, respectively. This is due to the smaller number of instances leading to some high conditional probabilities at  $\tau = 1$  in ENSAD. Those results remain consistent with the measures of high betweenness and in-degree centralities described in the S1 File. However, it shows that merging all CIs in one graph yields more stable patterns.

Chains-of- $m$ -events with  $m = 5$  (i.e.  $\tau = 4$ ) are finally ranked per CI by Eq 4, with the respective top-5 chains shown in Fig 5. We do not show smaller chains but recall that they are also included in  $M$ . We chose  $\tau = 4$  based on the results presented above. The approach provides a simple way to systematically explore emergent chains-of-events, including feedback loops. The best example observed here is the chain *explosion* → *fire* → *explosion* → *fire*. While the four critical events (explosion, fire, release, and rupture) drive the domino effects (shown in dark grey), we see other events, such as loading/unloading, gas poisoning, excavating or grounding, participate to the emergence of such chains (shown in light grey). Note that more would be found by analysing lower- $\gamma$  scenarios. Such events were only secondary in terms of centrality measures and their relative important may have been missed if solely using graph theory (see Fig 3). The metric  $\gamma$ , being interpretable as the probability of a specific chain among chains-of- $m$ -events, allows prioritizing mitigation measures not only based on critical events but on critical chains-of-events that contain more information. This is exemplified by the ENSAD example in Fig 5.



**Fig 5. Ranking of chains-of-5-events from the importance measure  $\gamma$  for different oil & gas CIs.** (A) Refinery case. (B) Gas network case. (C) Tanker case. Secondary events that were not highly ranked from centrality measures are represented in light grey.

<https://doi.org/10.1371/journal.pone.0263962.g005>

## Concluding remarks

Severe accidents in the oil & gas industry can have important repercussions on the global energy flow, and thus on the economy. Due to the complexity of the critical infrastructures involved, those accidents generally result from complex chains-of-events. The critical events that participate to the build-up to catastrophe (such as human error, explosion, fire) are well-known and identifiable from graph theory. However, the systematic exploration of plausible chains-of-events must provide richer information for risk mitigation. This was feasible in the present study via the analysis of the authoritative ENergy-related Severe Accident Database. We used the concept of fundamental matrix and defined an importance measure to identify the complexity of interactions and rank emergent behaviours, respectively.

The topological analysis of more than a thousand oil & gas accidents from the ENSAD confirmed the critical role of fire, explosion, and toxic release (but also element rupture) in chains-of-events leading to severe accidents, as well as human error as being the main source of such accidents, globally. Our dynamical analysis further demonstrated the role of long chains-of-events (up to 5 events), the significant extension of the space of possible indirect interactions, the feedback loops (*fire* ↔ *explosion*) and the role of secondary events on top of the four critical events (fire, explosion, release, rupture) in the most likely chains-of-events. This complexity was modelled via the interaction matrix  $M$  and ranked via the importance measure  $\gamma$ . Although the method is applicable to any adjacency matrix  $A$  of one-to-one interactions, only large databases such as ENSAD can lead to its full appreciation. Considering CIs individually, which reduced the number of instances, showed the important role of sample size. Interestingly, the full exploration of potential chains-of-events based on a narrower set of observed chains-of-events is one approach to model downward counterfactuals [45].

These insights should inform the optimal allocation of resources and needed level of maintenance in oil & gas CIs. With CI risk management crucially dependent on the analysis of causes and effects of failures [20, 23, 46, 47], especially of the evaluation of the entire accident sequence [33], decision-makers and first responders alike could use the ENSAD-based adjacency and interaction matrices to estimate the *a priori* cascading risk for the stress-testing of their CIs [48]. A Bayesian approach could be applied to update the ENSAD prior with CI-specific characteristics [4, 49]. Future improvements in accident databases shall further improve this approach. In general, we believe that data-driven dynamics will help minimize the impact of surprising accidents in the energy sector.

## Supporting information

**S1 File. Centrality measures of the severe accident chains-of-events at different types of critical infrastructures.** Refineries (S1 Fig), tankers (S2 Fig) and gas networks (S3 Fig). (PDF)

## Acknowledgments

We thank three anonymous reviewers for their comments, which helped improve this article.

## Author Contributions

**Conceptualization:** Arnaud Mignan, Matteo Spada, Peter Burgherr, Ziqi Wang.

**Data curation:** Matteo Spada, Peter Burgherr.

**Formal analysis:** Arnaud Mignan.

**Funding acquisition:** Arnaud Mignan.

**Investigation:** Arnaud Mignan, Matteo Spada, Peter Burgherr.

**Methodology:** Arnaud Mignan, Ziqi Wang, Didier Sornette.

**Validation:** Arnaud Mignan, Ziqi Wang.

**Visualization:** Arnaud Mignan.

**Writing – original draft:** Arnaud Mignan, Ziqi Wang.

**Writing – review & editing:** Arnaud Mignan, Matteo Spada, Peter Burgherr, Ziqi Wang, Didier Sornette.

## References

1. Buldyrev SV, Parshani R, Paul G, Stanley HE, Havlin S. Catastrophic cascade of failures in interdependent networks. *Nature*. 2010; 464: 1025–1028. <https://doi.org/10.1038/nature08932> PMID: 20393559
2. Sornette D, Maillart T, Kröger W. Exploring the limits of safety analysis in complex technological systems. *Int. J. Disaster Risk Reduction*. 2013; 6: 59–66. <https://doi.org/10.1016/j.ijdr.2013.04.002>
3. Mignan A, Scolobig A, Sauron A. Using reasoned imagination to learn about cascading hazards: a pilot study. *Disaster Prevention and Management*. 2016; 25 (3): 329–344. <https://doi.org/10.1108/DPM-06-2015-0137>
4. Spada M, Burgherr P. A Hierarchical Approximate Bayesian Computation (HABC) for accident risk in the energy sector triggered by natural events. *Proc. 29th European Safety and Reliability Conf., Singapore*. 2019; 1423–1430.
5. Otto C, Piontek F, Kalkuhl M, Frieler K. Event-based models to understand the scale of the impact of extremes. *Nature Energy*. 2020; 5: 111–114. <https://doi.org/10.1038/s41560-020-0562-4>
6. Khan FI, Amyotte PR. Modeling of BP Texas City refinery incident. *J. Loss Prevention in the Process Industries*. 2007; 20: 387–395. <https://doi.org/10.1016/j.jlp.2007.04.037>
7. Isimite J, Rubini P. A dynamic HAZOP case study using the Texas City refinery explosion. *J. Loss Prevention in the Process Industries*. 2016; 40: 496–501. <https://doi.org/10.1016/j.jlp.2016.01.025>
8. Krausmann E, Cruz AM. Impact of the 11 March 2011, Great Japan earthquake and tsunami on the chemical industry. *Nat. Hazards*. 2013; 67: 811–828. <https://doi.org/10.1007/s11069-013-0607-0>
9. Qin R, Khakzad N, Zhu J. An overview of the impact of Hurricane Harvey on chemical and process facilities in Texas. *Int. J. Disaster Risk Reduction*. 2020; 45: 101453. <https://doi.org/10.1016/j.ijdr.2019.101453>
10. Inhaber H. What is the Total Risk of Energy Systems? In: Berg GG, Maillie HD (eds). *Measurement of Risks*, Environmental Science Research. 1981; 21: 295–319.
11. Hirschberg S, Burgherr P, Spiekerman G, Dones R. Severe accidents in the energy sector: comparative perspective. *J. Haz. Mat.* 2004; 111: 57–65. <https://doi.org/10.1016/j.jhazmat.2004.02.009> PMID: 15231348
12. Felder FA. A critical assessment of energy accident studies. *Energy Policy*. 2009; 37: 5744–5751. <https://doi.org/10.1016/j.enpol.2009.08.059>
13. Burgherr P, Hirschberg S. Comparative risk assessment of severe accidents in the energy sector. *Energy Policy*. 2014; 74: 545–556. <https://doi.org/10.1016/j.enpol.2014.01.035>
14. Spada M, Burgherr P, Boutinard Rouelle P. Comparative risk assessment with focus on hydrogen and selected fuel cells: Application to Europe. *Int. J. Hydrogen Energy*. 2018; 43: 9470–9481. <https://doi.org/10.1016/j.ijhydene.2018.04.004>
15. Spada M, Sutra E, Burgherr P. Comparative accident risk assessment with focus on deep geothermal energy systems in the Organization for Economic Co-operation and Development (OECD) countries. *Geothermics*. 2021; 95: 102142. <https://doi.org/10.1016/j.geothermics.2021.102142>
16. Farrell AE, Zerriffi H, Dowlatabadi H. Energy Infrastructure and Security. *Annu. Rev. Environ. Resour.* 2004; 29: 421–469. <https://doi.org/10.1146/annurev.energy.29.062403.102238>
17. Sovacool BK. The cost of failure: a preliminary assessment of major energy accidents, 1907–2007. *Energy Policy*. 2008; 36: 1802–1820. <https://doi.org/10.1016/j.enpol.2008.01.040>
18. Burgherr P, Spada M, Kalinina A, Vandepaer L, Lustenberger P, Kim W. Comparative risk assessment of accidents in the energy sector within different long-term scenarios and marginal electricity supply mixes. *Proc. 29th European Safety and Reliability Conf., Singapore*. 2019; 1525–1532.

19. Kim W, Burgherr P, Spada M, Lustenberger P, Kalinina A, Hirschberg S. Energy-related Severe Accident Database (ENSAD): cloud-based geospatial platform. *Big Earth Data*. 2019; 2: 368–394. <https://doi.org/10.1080/20964471.2019.1586276>
20. Moura R, Beer M, Patelli E, Lewis J, Knoll F. Learning from accidents: Interactions between human factors, technology and organisations as a central element to validate risk studies. *Safety Science*. 2017; 99: 196–214. <https://doi.org/10.1016/j.ssci.2017.05.001>
21. McCollum DL, Gambhir A, Rogelj J, Wilson C. Energy modellers should explore extremes more systematically in scenarios. *Nature Energy*. 2020; 5: 104–107. <https://doi.org/10.1038/s41560-020-0555-3>
22. Mignan A, Wiemer S, Giardini D. The quantification of low-probability-high-consequences events: part I. A generic multi-risk approach. *Nat. Hazards*. 2014; 73: 1999–2022. <https://doi.org/10.1007/s11069-014-1178-4>
23. Zio E. Challenges in the vulnerability and risk analysis of critical infrastructures. *Reliability Engineering and System Safety*. 2016; 152: 137–150. <https://doi.org/10.1016/j.res.2016.02.009>
24. Rocha JC, Peterson G, Bodin Ö, Levin S. Cascading regime shifts within and across scales. *Science*. 2018; 362: 1379–1383. <https://doi.org/10.1126/science.aaf7850> PMID: 30573623
25. Helbing D, Kühnert C. Assessing interaction networks with applications to catastrophe dynamics and disaster management. *Physica A*. 2003; 328: 584–606. [https://doi.org/10.1016/S0378-4371\(03\)00519-3](https://doi.org/10.1016/S0378-4371(03)00519-3)
26. Khakzad N, Reniers G. Using graph theory to analyze the vulnerability of process plants in the context of cascading effects. *Reliab. Eng. Syst. Saf.* 2015; 143: 63–73. <https://doi.org/10.1016/j.res.2015.04.015>
27. Stergiopoulos G, Kotzanikolaou P, Theocharidou M, Gritzalis D. Risk mitigation strategies for critical infrastructures based on graph centrality analysis. *Int. J. Critical Infrastructure Protection*. 2015; 10: 34–44. <https://doi.org/10.1016/j.ijcip.2015.05.003>
28. Tang P, Xia Q, Wang Y. Addressing cascading effects of earthquakes in urban areas from network perspective to improve disaster mitigation. *Int. J. Disaster Risk Reduction*. 2019; 35: 10165. <https://doi.org/10.1016/j.ijdrr.2019.101065>
29. Burgherr P, Cinelli M, Spada M, Blaszczyński J, Stowinski R, Pannatier Y. Risk assessment of worldwide refinery accidents using advanced classification methods: Effects of refinery configuration and geographic location on outcome risk levels. In: Haugen S, Barros A, van Gulijk C, Kongsvik T, Vinnem JE (eds). *Safety and Reliability—Safe Societies in a Changing World*, London, UK. 2018; 1681–1688.
30. Spada M, Ferretti V. Toward the integration of uncertainty and probabilities in spatial multi-criteria risk analysis: An application to tanker oil spills. In: Haugen S, Barros A, van Gulijk C, Kongsvik T, Vinnem JE (eds). *Safety and Reliability—Safe Societies in a Changing World*, London, UK. 2018; 1673–1680.
31. Cinelli M, Spada M, Kadzinski M, Miebs G, Burgherr P. Advancing Hazard Assessment of Energy Accidents in the Natural Gas Sector with Rough Set Theory and Decision Rules. *Energies*. 2019; 12: 4178. <https://doi.org/10.3390/en12214178>
32. Khan FI, Abbasi SA. Models for Domino Effect Analysis in Chemical Process Industries. *Process Safety Progress*. 1998; 17 (2): 107–123. <https://doi.org/10.1002/prs.680170207>
33. Baalisampang T, Abbassi R, Garaniya V, Khan F, Dadashzadeh M. Modelling an integrated impact of fire, explosion and combustion products during transitional events caused by an accidental release of LNG. *Proc. Saf. Env. Protect*. 2019; 128: 259–272. <https://doi.org/10.1016/j.psep.2019.06.005>
34. Mignan A, Wang Z. Exploring the Space of Possibilities in Cascading Disasters with Catastrophe Dynamics. *Int. J. Environ. Res. Public Health*. 2020; 17: 7317. <https://doi.org/10.3390/ijerph17197317> PMID: 33036402
35. Cochran JK, Murugan A, Krishnamurthy V. Generic Markov models for availability estimation and failure characterization in petroleum refineries. *Computers & Operational Res*. 2001; 28: 1–12. [https://doi.org/10.1016/S0305-0548\(99\)00082-9](https://doi.org/10.1016/S0305-0548(99)00082-9)
36. Burgherr P, Hirschberg S. A Comparative Analysis of Accident Risks in Fossil, Hydro, and Nuclear Energy Chains. *Human and Ecological Risk Assessment*. 2008; 14: 947–973. <https://doi.org/10.1080/10807030802387556>
37. Khan FI, Abbasi SA. Techniques and methodologies for risk analysis in chemical process industries. *J. Loss Prev. Proc. Ind*. 1998; 11: 261–277. [https://doi.org/10.1016/S0950-4230\(97\)00051-X](https://doi.org/10.1016/S0950-4230(97)00051-X)
38. Nwankwo CD, Arewa AO, Theophilus SC, Esenowo VN. Analysis of accidents caused by human factors in the oil and gas industry using the HFACS-OGI framework. *Int. J. Occupational Safety and Ergonomics*. 2021; 1–13. <https://doi.org/10.1080/10803548.2021.1916238> PMID: 33849397
39. Freeman LC. Centrality in Social Networks Conceptual Clarification. *Social Networks*. 1978/79; 1: 215–239. [https://doi.org/10.1016/0378-8733\(78\)90021-7](https://doi.org/10.1016/0378-8733(78)90021-7)

40. Ugurlu Ö, Yildirim U, Basa E. Analysis of Grounding Accidents Caused by Human Error. *J. Marine Sci. Tech.* 2015; 23: 748–760. <https://doi.org/10.6119/JMST-015-0615-1>
41. Chernov D, Sornette D. *Man-made Catastrophes and Risk Information Concealment*. Springer International Publishing. 2016; 342 pp.
42. Zheng W, Chen J. Human errors are behind most oil-tanker spills. *Nature*. 2018; 560: 161–163. <https://doi.org/10.1038/d41586-018-05852-0> PMID: 30082696
43. Khan FI, Abbasi SA. Major accidents in process industries and an analysis of causes and consequences. *J. Loss Prev. Proc. Ind.* 1999; 12: 361–378. [https://doi.org/10.1016/S0950-4230\(98\)00062-X](https://doi.org/10.1016/S0950-4230(98)00062-X)
44. Yuan C, Cui H, Wang W, Ma S. Cause Factors in Emergency Process of Fire Accidents for Oil and Gas Storage and Transportation Systems Based on ISM and AHP. *J. Hazard. Toxic Radioact. Waste.* 2019; 23: 04018038. [https://doi.org/10.1061/\(ASCE\)HZ.2153-5515.0000432](https://doi.org/10.1061/(ASCE)HZ.2153-5515.0000432)
45. Woo G, Mignan A. Counterfactual Analysis of Runaway Earthquakes. *Seismol. Res. Lett.* 2018; 89 (6): 2266–2273. <https://doi.org/10.1785/0220180138>
46. Swanson L. Linking maintenance strategies to performance. *Int. J. Production Economics.* 2001; 70, 237–244. [https://doi.org/10.1016/S0925-5273\(00\)00067-0](https://doi.org/10.1016/S0925-5273(00)00067-0)
47. Cheng C-W, Yao H-Q, Wu T-C. Applying data mining techniques to analyze the causes of major occupational accidents in the petrochemical industry. *J. Loss. Prev. Proc. Ind.* 2013; 26: 1269–1278. <https://doi.org/10.1016/j.jlp.2013.07.002>
48. Esposito S, Stojadinovic B, Babic A, Dolsek M, Iqbal S, Selva J, et al. Risk-Based Multilevel Methodology to Stress Test Critical Infrastructure Systems. *J. Infrastruct. Syst.* 2020; 26: 04019035. [https://doi.org/10.1061/\(ASCE\)IS.1943-555X.0000520](https://doi.org/10.1061/(ASCE)IS.1943-555X.0000520)
49. Meel A, O'Neill LM, Levin JH, Seider WD, Oktem U, Keren N. Operational risk assessment of chemical industries by exploiting accident databases. *J. Loss Prev. Proc. Ind.* 2007; 20: 113–127. <https://doi.org/10.1016/j.jlp.2006.10.003>