CHEMICAL ENGINEERING TRANSACTIONS

VOL. 99, 2023

A publication of AIDIC
The Italian Association of Chemical Engineering
Online at www.cetjournal.it

Guest Editors: Sauro Pierucci, Flavio Manenti

ISBN 978-88-95608-98-3; **ISSN** 2283-9216

DOI: 10.3303/CET2399049

A Dynamic Approach to Natech Risk Assessment Applied to an LPG Storage Facility in a Landslides Sensitive Italian Area

Tomaso Vairo[a, b]*, Stefania Magrì[a, c], Gabriele Mercurio[b], Margherita Pettinato[a], Maria Bolla[a], Bruno Fabiano[a]

[a] Department of Civil Chemical and Environmental Engineering, via Opera Pia 15, 16145, University of Genoa, Italy.
[b] ASL3, Workplace Safety and Prevention Unit (S.C. PSAL), via Bertani 4, 16125, Genoa, Italy.
[c] ARPAL (Ligurian Regional Environmental Protection Agency), Via Bombrini 8, 16149 Genova, Italy.
tomvairo@gmail.com

Due to the climate change, extreme weather phenomena are becoming increasingly intense and occur with higher frequencies, even in unusual areas. Nevertheless, historical data showed as Natech accidents can be triggered not only by natural disasters, like earthquakes or tornadoes, but even by natural phenomena that are considered of minor importance, such as rain and lightning. Only recently, the Natech issue has gained a great deal of attention, but there is still a lack of consolidated Natech risk-analysis methodologies and tools. The focus of this work is to include natural hazards into a dynamic risk assessment system beside the typical parameters of process risks. In Italy, rainfall represents the most common triggering factor for landslides. Generally, the determination of trigger and propagation can rely on physically-based approaches, which require the calibration of many parameters and are often difficult to apply, or on empirical correlations between rainfall and landslide built from historical data. On the other hand, by using a data driven approach, available data can be exploited to define the system state over time, anticipate the systems outcome, support decision-making, and adopt the most appropriate adjustments, allowing to enhance system resilience and knowledge. The actual capability of the proposed approach was evaluated on a simple case-study represented by an LPG storage facility located in landslides sensitive zone of Liguria Region.

1. Introduction

Rainfall-induced shallow landslides, also known as soil slips or shallow landslides, occur when the soil becomes unstable after intense or prolonged rainfall due to its saturation with water. These phenomena are generally dependent on a variety of factors, including the intensity and duration of the rainfall, the type of soil, the slope of the land, and the presence of vegetation (Park et al., 2013). Shallow landslides can significantly impact the communities and the environment, potentially leading to erosion, damages to infrastructures, and loss of life. They are particularly dangerous due to their sudden occurrence, often with little warning, and rapid spread (Petley, 2012). Therefore, it is important to understand how the main factors, such as geological conditions, weather patterns, and human activities, contribute to their occurrence and to develop strategies for mitigating the related risks and reducing the impacts. Several methods predicting landslides are currently discussed in literature. They usually rely on assessing geological data about the characteristics of the land or they triconcern the analysis of data about weather patterns and other triggering events, such as the intensity and duration of rainfall (Baum et al., 2010). Physically-based models are one of the most effective ways to predict landslides and their evolution over time under different conditions. However, these models are difficult to calibrate and validate, require a large amount of data and computational resources to run accurately, and are often used in conjunction with statistical, or empirical models to improve the accuracy of the predictions. Berti et al. (2012) proposed a method for evaluating rainfall thresholds based on Bayesian probability, returning a value of landslide probability for each combination of the selected rainfall variables. This work focuses on Natech risk assessment to evaluate the likelihood, the potential impacts, and the actions for risk reduction associated to natural hazards triggering technological scenarios involving hazardous materials,

Paper Received: 9 December 2022; Revised: 3 March 2023; Accepted: 30 April 2023
Please cite this article as: Vairo T., Magri S., Mercurio G., Pettinato M., Bolla M., Fabiano B., 2023, A Dynamic Approach to Natech Risk Assessment Applied to an LPG Storage Facility in a Landslides Sensitive Italian Area, Chemical Engineering Transactions, 99, 289-294 DOI:10.3303/CET2399049

which is explicitly included in the last update of the so-called Seveso Directive on the control of major-accident hazards (Laurent et al., 2021). The growing severity of extreme natural phenomena, under the driving action of climate change, it is expected to increase the vulnerability to Natech risk. This multi-hazard risk involves different domains, including risk governance challenges, socio-economic context (Krausman et al., 2019), and implications from political/war instabilities. Recently, novel approaches have been proposed for multi-hazard natural disasters, potentially causing destructive damage to the process equipment (Huang et al., 2022). In this regard, machine learning techniques can be used to analyse data about past landslides and other accidents and to identify patterns or trends that can inform about disaster response and recovery efforts. The remainder of this paper outlines a novel framework, based on a Gradient Boosting algorithm to predict shallow landslides, integrated into a dynamic risk assessment system for early warning in major accident hazard facilities located in areas susceptible to landslides. A dedicated bow-tie approach is then adopted and the methodology is applied to a test case demonstrating current capabilities and future challenges.

2. Theoretical

2.1 Physically-based models

Physically-based models for predicting landslides are mainly based on the Richards equation, which derives from the principles of mass conservation and Darcy's law and describes the movement of water through the soil and other porous media. The behavior of landslides is simulated under different conditions and it's possible to forecast how landslides are likely to evolve over time (Mercurio, 2008). The use of physically-based approaches requires knowledge about saturation conditions and pore water pressure, as well as the calibration of many parameters, such as hydraulic and mechanical properties of soil, the effect of vegetation, and local rainfall variation in space and time, which makes their use over large territories difficult to apply.

2.2 Empirical models

Empirical models, based on the observed data, are used when mathematical models are not able to accurately represent complex systems, or phenomena. As stated by Berti et al. (2012), an empirical model based on a probabilistic analysis can associate a reliability to a given landslide threshold, giving a probability distribution of the forecast. In particular, probabilistic Bayesian approaches can be useful tools for providing effective weather forecasting (Vairo et al., 2019) and develop a natural risk index (Ancione and Milazzo, 2021). However, if the data used to build the model are biased, incomplete, or contain uncertainties, the results may be unreliable. These potential drawbacks can be addressed by using regression models, such as gradient boosted decision trees (GBDTs), to build up statistical models of the relationships between different variables. GBDTs are a type of machine learning algorithm that can be used to construct models of complex relationships between variables by combining the predictions of many simpler models. They are effective in handling large and complex datasets, even in presence of missing or incomplete data, without requiring preventive data cleaning or imputation. GBDTs can also be trained to handle noisy or unbalanced datasets, commonly encountered dealing with landslides. In this work, the open library *LightGBM* (Ke et al., 2017) was preliminarily tested and then applied for the developed approach.

3. Assessing Natech risk by a dynamic perspective

Climate change (with expected frequency increase) and energy transition (with possible novel hazards) evidence the need of advanced approaches to contain the Natech risk and correct land-use planning and management. As widely reported, Natech risk assessment (RA) is the process of evaluating the likelihood and potential impacts of natural hazards and technological disasters, such as earthquakes, landslides, and industrial accidents. The focus of risk assessment is to identify areas exposed to high risk for these types of events and to develop mitigation strategies This study aims to contribute to this challenge in relation with dynamic Natech RA of an LPG storage facility located in landslide sensitive area of located in Liguria (Italy).

3.1 The predictive model

The predictive model is built adopting as predictors five features, namely: the soil moisture of the area; accumulated rain over 3, 6,12, 24 hours in the area; peak rain over 3 and 24 hours in the area; day of the year. The soil moisture is identified as the relevant predisposing factor, while the accumulated and peak rainfalls represent the triggering factor. Since a precise knowledge of the exact time of the landslide occurrence is not available, the accumulated and peak rain data are calculated from 00 UTC and over the whole day. The dataset used for the model training and validation, was provided by the Ligurian Regional Environmental Protection Agency (ARPAL). An example of a soil moisture map obtained at 00:00 UTC together with the location of the rain gauges in the explored pilot area is shown in Figure 1.

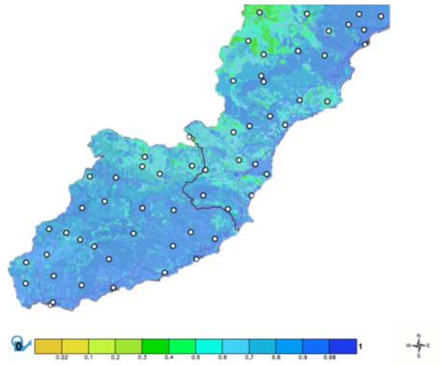

Figure 1: Example of a soil moisture map and location of rain gauges in the investigated West area of Liguria region, Italy. (https://omirl.regione.liguria.it/).

The whole province of Imperia was considered for the training and test of the model, to increase the number of observations. In fact, albeit the used dataset covered 6 years, from 2014 to 2019, occurrence of landslides was observed in 365 days only. Preliminary Data Analysis (PDA) shows that the landslides dataset is heavily unbalanced, since on most of the days no landslide events occurred, therefore, the number of events was classified for obtaining homogeneous classes, as shown in Table 1. The classified data are shown in Figure 2.

Table 1: classification of landslide events

Class (Label)	*Description*	*Hazard level*	*Nr events*
1	Few shallow landslides	Moderate hazard	0-1
2	Shallow landslides and rapid mud flows of limited size	Medium hazard	2-3
3	Slope instability	High hazard	3+

During training and validation, the performance of the model on different datasets is measured, exhibiting a good accuracy, as shown in Figure 2 left-hand side. The metric is calculated based on the predictions made by the model and the true values of the target variable. It is used to measure the accuracy of the model as it is being trained and to guide the training process. The performance of the model on the validation set can be used to tune the model's hyperparameters and to assess the model's generalization ability, or its ability to make accurate predictions on unseen data.

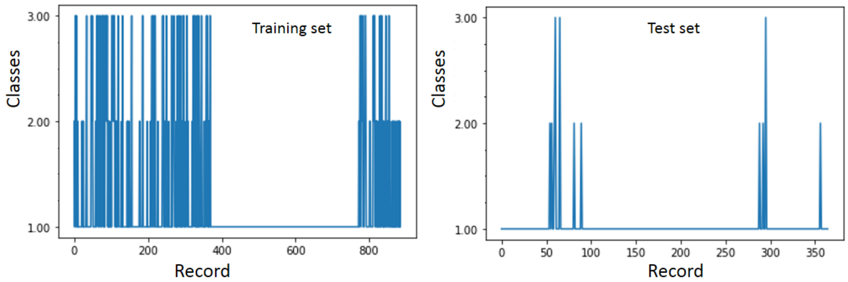

Figure 2: Classification of landslides events in the training and test set.

The most appropriate metric for the purpose of the implemented model is *multi log-loss*. Multi log-loss is used to evaluate the performance of a classification model. It is based on the concept of *cross-entropy loss*, which is a measure of the difference between the predicted probability distribution and the true probability distribution for a given set of data. Multi log-loss is used to compare the predicted probability distribution for each class with the true probability distribution, which is represented by a binary indicator variable that takes on a value of 1 for the true class and 0 for all other classes. As presented in result section, the prediction accuracy is represented by a confusion matrix presenting information about the true positive, true negative, false positive, and false negative predictions obtained by the model. The true positive rate (TPR) and false positive rate (FPR) can be calculated from the confusion matrix and are used to evaluate the model's performance. The consistency of the predictive model is evident by the calculation of the feature importance, which refers to the relative contribution of each feature (also known as a predictor, or input variable) to the model predictions (Vairo et al., 2023a). Features that are more important exert a larger impact on the model predictions and are more relevant to the problem-solving item. As reported in Figure 3, the feature importance is measured by the coverage, i.e., the number of times that the feature is used to split the data into the model's decision trees, providing in this way a clear assessment of the model explainability.

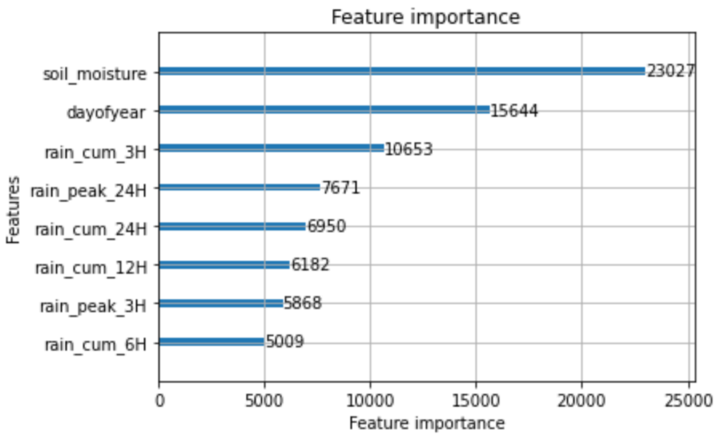

Figure 3: Model explainability assessment based on feature importance.

4. Results and discussion

Figure 4(a) presents the performance results of the training phase of the model. Figure 4(b), shows a graphical representation of the prediction accuracy by the confusion matrix. TPR, also known as sensitivity or recall, is the number of true positive predictions made by the model divided by the total number of positive cases in the data. FPR is the number of false positive predictions by the model divided by the total number of negative cases in the data. The landslide predictive model has a fairly good accuracy, even though a slight underestimation in identifying class 2 events is evident in the confusion matrix. As a refinement, an enlarged dataset is currently under testing, adopting different models and a wider and more detailed dataset, as long as new processed data are available.

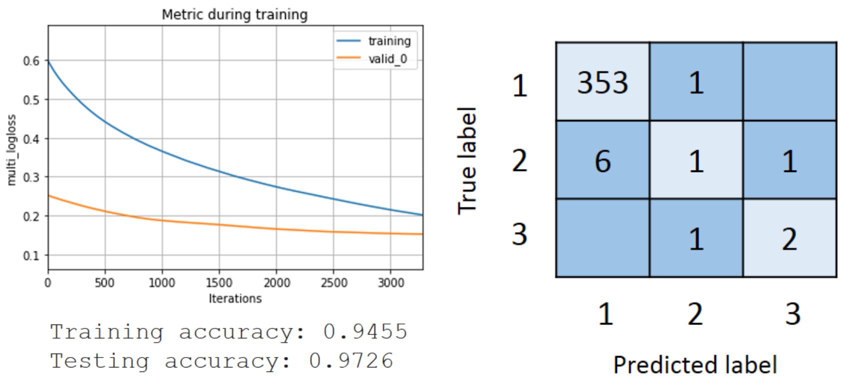

Figure 4: (a) left-hand side: metric during the model training. (b) right-hand side: confusion matrix.

For assessing the Natech risk for an LPG storage facility in a landslide sensitive zone, starting from the framework outlined in Vairo et al. (2023b), the following interdependencies need to be thoroughly evaluated.

1. How the *Adversity* affects the *System*.
2. How the *System*, when subjected to the Adversity, delivers the *Capability of interest*.
3. What is the quality of the *Delivered Capability* gaged against the *Required Capability*.

In the analyzed case study, the *Adversity* is a landslide, the *System* is an LPG storage facility, and the *Capability of interest* is the system safety. Natech risk reduction implies reducing both likelihood and severity, in particular by implementing the efficiency of safety barriers and emergency response measures. The Natech hazard identification, showing the interaction between the landslide and the given technological installation including final outcomes, is schematized in Table 2 according to the form of unconditional event tree. Here the role of personnel and management should be properly expanded to cover a number of relevant items e.g., procedures, education, accountability and motivation (Fabiano et al., 2022).

Table 2: Landslide triggered events on an LPG storage facility.

Off-site event	*Operator error*	*Abnormal load*	*Failures*	*Management*	*Loss of service*	*Triggered initiating event*	*Consequences*
Landslide	Containment degraded Failure to respond correctly to an alarm	Internal temperature or pressure outside design limit Pressurization / under pressure	Safety system degraded. Control system degraded. Containment system degraded	Inadequate materials or specification Hidden defect in containment system Failure to detect dangerous situation. Failure of process controls.	Loss of cooling water / nitrogen Loss of compressed air	Rupture of pipe on a pressurized storage system Sudden catastrophic failure of vessels Failure of an excess flow control valve on demand Failure of an automatic shutoff valve closure Failure of a level / flow sensor	Catastrophic failure - fireball and flash fire or pool fire Localised failure of a pressure vessel – jet flame and flash fire and possible explosion Pipe failures BLEVE of vessels Vaporiser leak jet fire, flash fire, and explosion. Leak inside cylinder filling plant - confined explosion

The landslide initiating events presented in Table 1 are used to determine the landslide triggered events, by the means of a Bow-Tie centered on the pivotal event as depicted in Figure 5. Based on the bow-tie approach, it is possible to determine the initiating causes of an accidental event and the failure or operational disruptions possibly leading final damage outcomes.

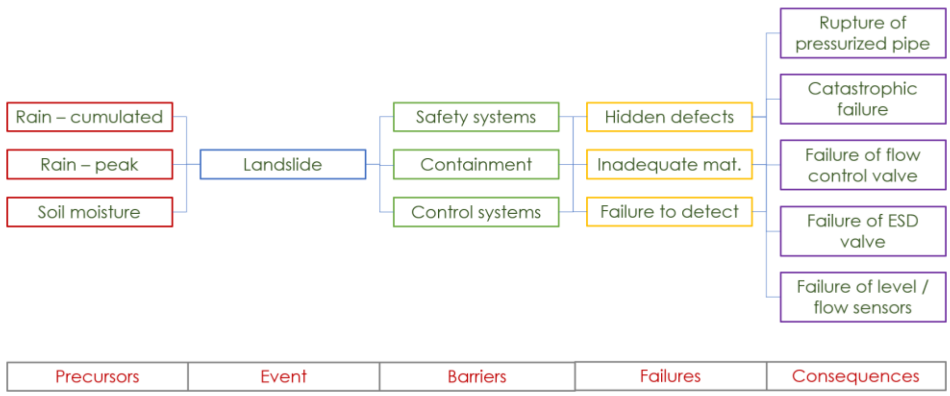

Figure 5: Simplified Bow-Tie centred on landslide event.

Table 3: Bow-Tie elements – landslide consequences / initiating events.

Cause	*Expected pdf value*
Rupture of pressurized pipe	1E-6
Catastrophic failure	1E-9
Failure of flow control valve	1E-5
Failure of ESD valve	1E-6
Failure of level / flow sensors	1E-3

As summarized in Table 3, the probabilities are evaluated by a MCMC sampling from the probability distribution determined from the predictive model. Each element is updated at new observation of the identified precursors i.e., Rain-cumulated, Soil moisture, and Rain-peak.

5. Conclusions

The landslide predictive model provided in general, a good accuracy, even if an underestimation is verified in identifying class 2 events, due to incomplete raw data. It was shown that of landslide event prediction has a considerable impact on the top events probabilities and the dynamically update of the risk precursors represents a relevant safety parameter allowing an early detection of a potential threat. As a refinement of the method, further investigation of the tuning parameters of the predictive model is underway, with a thorough sensitivity analysis covering a wider range of situations/configurations for improving the early warning ability.

Aknowledgements

This research was funded by INAIL, within the frameworks of the calls BRIC/2019/ID2 (Project DYN-RISK) and BRIC/2021/ID03 (project DRIVERS). Authors gratefully acknowledge contribution of the staff of ARPAL, SATURN Department for providing the training and testing data set.

References

Ancione, G., Milazzo, MF., 2021, The management of Na-tech risk using bayesian network, Water 13, 142, 1966.

Baum, R.L., Godt, J.W., 2013, Early warning of rainfall-induced shallow landslides and debris flows in the USA Landslides, 7, 259-272.

Berardi, R., Mercurio, G., Bartolini, P., Cordano, E., 2005, Dynamics of saturation phenomena and landslide triggering by rain infiltration in a slope, International Conference on Landslide Risk Management, 212-219, Taylor & Francis Ltd., Vancouver, Canada.

Berti, M., Martina, M. L. V., Franceschini, S., Pignone, S., Simoni, A., Pizziolo, M., 2012, Probabilistic rainfall thresholds for landslide occurrence using a Bayesian approach, Journal of Geophysical Research, 117, F04006.

Fabiano, B., Pettinato, M., Currò, F., Reverberi, A.P., 2022, A field study on human factor and safety performances in a downstream oil industry, Safety Science, 153, 105795.

Fredlund, D. G., Morgenstern, N. R., Widger, R. A., 1978, The shear strength of unsaturated soils, Canadian Geotechnical Journal, 15, 313-321.

Huang, M., Chen, G., Yang, P., Hua, K., Zhou, L., Men, J., Zhao, J. 2022, Multi-hazard coupling vulnerability analysis for buckling failure of vertical storage tank: Floods and hurricanes. Process Saf. Environ. Prot., 161, 528-541.

Ke, G., Meng Q., Finley, T., Wang, T., Chen, W., Ma, W., Ye, Q., Liu, T. Y., 2017, LightGBM: A highly efficient gradient boosting decision tree. Advances in neural information processing systems, 30, 3146-54.

Krausmann, E., Girgin, S., Necci, A., 2019, Natural hazard impacts on industry and critical infrastructure: Natech risk drivers and risk management performance indicators, Int. J. of Disaster Risk Reduction, 40, 101163.

Laurent, A., Pey, A., Gurtel, P., Fabiano, B., 2021, A critical perspective on the implementation of the EU Council Seveso Directives in France, Germany, Italy and Spain, Process Saf. Environ. Prot., 148, 47–74.

Mercurio, G., 2008, Infiltration processes and triggering mechanisms of shallow landslides in partially saturated soils (in Italian), PhD thesis, University of Genova, Genoa, Italy.

Park, H.J., Lee, J.H., Woo, I.K., 2013, Assessment of rainfall-induced shallow landslide susceptibility using a GIS-based probabilistic approach, Engineering Geology, 161, 1-15.

Petley, D., 2012, Global patterns of loss of life from landslides, Geology, 40, 927-930.

Vairo, T., Lecca, M., Trovatore, E., Reverberi, A.P., Fabiano B., 2019, A Bayesian belief network for local air quality forecasting, Chemical Engineering Transactions, 74, 271-276.

Vairo, T., Gualeni, P., Reverberi, A.P., Fabiano, B., 2021, Resilience dynamic assessment based on precursor events: Application to ship LNG bunkering operations, Sustainability, 13(2), 6836.

Vairo T., Cademartori, D., Clematis, D., Carpanese, M.P., Fabiano B., 2023a, Solid oxide fuel cells for shipping: A machine learning model for early detection of hazardous system deviations. Process Saf. Environ. Prot., 172, 184-194.

Vairo T., Pettinato, M., Reverberi, A.P., Milazzo, M.F., Fabiano B., 2023b, An approach towards the implementation of a reliable resilience model based on machine learning, Process Saf. Environ. Prot., 172, 632-641.