Hydrogeology Journal
https://doi.org/10.1007/s10040-026-03102-6

PAPER

Risk assessment of tunnelling-induced hydrogeological interference on springs using a machine learning approach

Ernesto Pugliese[1] · Dany Lauzon[2] · Antonio Dematteis[3] · Stefano Rodani[4] · Gianluca Benedetti[4] · Alessandro Gargini[1]

Received: 15 September 2025 / Accepted: 21 May 2026

Abstract

Tunnel excavation through mountain ranges can alter groundwater regimes, affecting spring discharge and dependent ecosystems. Anticipating these interferences is essential for environmental risk management, cost–benefit analysis, and effective mitigation planning. Conventional assessment methods are largely parametric, not physically based, and heavily dependent on expert judgment, introducing subjectivity and uncertainty. This study proposes a data-driven approach based on machine learning (ML), specifically a Random Forest (RF) model, to evaluate hydrogeological risk and spring vulnerability during the preliminary design phase of tunnels. The method exploits geological information typically available in early stages and is validated using two well-documented Italian case studies: The Bologna–Florence high-speed railway line, excavated in fractured sedimentary formations of the Northern Apennines, and the Gran Sasso highway tunnels, driven in karstified carbonate rocks of the Central Apennines. Both datasets are published, hydrogeologically validated, and supported by detailed pre- and post-excavation monitoring, ensuring robustness and consistency. The ML framework demonstrates strong predictive performance across standard classification metrics, including accuracy, precision, recall, F1-score, Matthews correlation coefficient (MCC), and area under the curve (AUC). While it employs input parameters comparable to those of parametric methods such as the drawdown hazard index (DHI), it differs by deriving parameter weights directly from data. Moreover, the use of Shapley additive explanations (SHAP) values enhances interpretability, mitigates black-box behaviour, and allows expert knowledge to be effectively integrated. The innovation, therefore, lies in offering a transferable and reproducible tool to support early-stage tunnelling decisions, representing a clear improvement over traditional qualitative or semi-quantitative approaches.

Keywords Machine learning · Groundwater protection · Hydrogeological interference risk · Tunnel · Springs

Introduction

Estimating the impact of underground constructions against groundwater flow (Celico et al. 2005; Vincenzi et al. 2014; Preisig et al. 2014) and dependent ecosystems in mountainous areas is one of the key analyses carried out during preliminary investigations, both for environmental and socio-economic purposes (Yang et al. 2009). The water drainage produced by tunnelling, modifying the hydrogeological budget of an aquifer, can adversely affect groundwater discharge to springs and surface-water bodies and also induce a drawdown of hydraulic head with dewatering and impairment of specific discharge of wells (Sjölander-Lindqvist 2005; Attanayake and Waterman 2006; Vincenzi et al. 2009; Lv et al. 2020; Xu et al. 2021). Additionally, the induced drawdown and modification of groundwater flow patterns

✉ Ernesto Pugliese
ernesto.pugliese2@unibo.it

1 Biological, Geological and Environmental Sciences Department, Alma Mater Studiorum University of Bologna, 40126 Bologna, Italy

2 Department of Civil, Geological and Mining Engineering, Polytechnique Montréal, Station Centre-Ville, P.O. Box 6079, Montréal H3C 3A7, Canada

3 Lombardi SA, Via del Tiglio 2, P.O. Box 934, 6512 Bellinzona-Giubiasco, Switzerland

4 Gruppo Ferrovie dello Stato Italiane – Architecture, Environment & Territory Department – Geology Division, ITALFERR S.P.A., Via Galati 87, 00155 Roma, Italy

Published online: 13 June 2026

could also affect water quality and geochemistry (Mossmark et al. 2015), with undesired effects on ecosystems (Kværner and Snilsberg 2011; Li et al. 2016). A preliminary assessment of the tunnel drainage effects is also related to the proper management of drilling operations: significant and not adequately foreseen inflows can produce severe consequences in underground works, forcing the interruption of the advancement and even the modification of the designed tunnel route, with evident issues arising in terms of workers′ safety conditions and costs increase (Schwarz et al. 2006; Chiocchini and Castaldi 2011).

In order to reliably assess the potential impacts of dewatering hazard associated with tunnel drainage, preliminary (*ante operam*) environmental monitoring of potential receptors is essential. This allows for meaningful comparison with *post operam* conditions while accounting for other effects such as climatic variability or groundwater withdrawals (Attanayake and Waterman 2006; Gargini et al. 2008). Indeed, there are numerous cases of flow reductions in springs due to climate change (Casati et al. 2024; Filippini et al. 2024; Barbieri et al. 2023; Petitta et al. 2022; Sappa et al. 2018; 2019, Allocca et al. 2014; Fiorillo et al. 2015; Fiorillo and Guadagno 2012; Leone et al. 2021; Polemio and Casarano 2008) and these effects must be properly taken into account; however, the data collected during preliminary monitoring will be useful not simply as a testimonial reference point for any eventual future impact assessment but as an input for predictive models anticipating the impacts associated with excavation.

In the scientific literature, several cases address environmental risk assessment of tunnelling-induced hydrogeological interference; however, when it comes to risk analyses related to the impacts on springs, case studies are rather limited. This issue is often approached primarily in relation to the prediction of water inflows into tunnels, usually through numerical modelling. On the basis of the inflow estimated by the model, a corresponding drawdown of the piezometric surface is also considered, which in turn may lead to a reduction in spring discharge, or in some cases, to their complete drying up. Existing predictive models for the assessment of hydrogeological interference risk from tunnelling could be ranked as parametric, physically based (analytical or numerical), and statistical data-driven methods.

Parametric tools (PTs) offer a relatively robust, easy-to-use, and cost-effective (though subjective and not physically based) means to gain a preliminary understanding of potential hydrogeological interferences on receptors. They achieve this by categorizing variables that link the rock mass system to potential tunnel inflows and by relating these inflows to their likely impacts on receptors. One of the most adopted parametric tools is the drawdown hazard index (DHI), originally defined in 2001 (Dematteis et al. 2001) and then improved and fully explained in its intrinsic workflow by Dematteis et al. (2025), which, according to the general formulation of any environmental risk (Slejko 1993), relates the hazard, expressed by parameters of the rock mass subjected to excavation, the vulnerability, expressed by the hydrogeological connection between tunnel and springs, and the value of the receptors, expressed by the magnitude of the discharge of springs and the typology of the associated groundwater flow system. A probability impact degree for each spring, expressed by the DHI index, is associated with each receptor. Applications of DHI are reported by Torri et al. (2007) and Vincenzi et al. (2010). Filipponi and Jeannin (2010); Filipponi (2014) developed a parametric method, *KARSTAlea*, for the risk analysis in a karst setting, a particularly challenging task due to the difficulty in predicting hydrogeologic connections between underground work and springs through conduits. *KARSTAlea* involves constructing a 3D model and identifying areas more subjected to risk through the study of geomorphological elements and field observations. Alternative methods have also been developed to predict groundwater inflows into crossing-rocks tunnels. Most of these approaches rely on the geological characterization of the rock mass, with the aim of identifying geological features that are likely to store groundwater and act as preferential flow paths. However, such methods require extensive datasets and detailed subsurface investigations to achieve a high level of accuracy. Further research based on these approaches is therefore essential for improving the reliability of groundwater inflow assessments in tunnel engineering (Heidari et al. 2016; Katibeh and Aalianvari 2009; Zarei et al. 2011).

Physically based models (PBMs) bring objectivity to the analysis also if they could be too simplistic, like analytical ones, or too complex and deserving high resources or computational efforts, like numerical. Analytical and semianalytical methods are widely employed to predict groundwater inflows into rock tunnels and other subsurface openings. The development of these methods typically requires the adoption of simplifying assumptions and boundary conditions in order to formulate the governing equations (Frenelus et al. 2021). Several of them have been derived from Goodman's original equation (Goodman et al. 1965), incorporating different assumptions and simplifications to account for a wider range of hydrogeological conditions (El Tani 1999, 2003; Farhadian et al. 2016; Perrochet and Dematteis 2007; Karlsrud 2001; Gattinoni and Scesi 2010). Numerical methods have increasingly become powerful tools employed across diverse fields of engineering and scientific research. They are particularly valuable in the prediction and quantification of groundwater inflows into tunnels constructed in various rock masses. As noted, these methods rely on mathematical models that describe the hydrogeological and geomechanical characteristics of the medium under consideration (Molinero et al. 2002; Vincenzi et al. 2010; Amiri and Asgari-Nejad 2022; Golian et al. 2018, Surinaidu et al. 2014). A major drawback of numerical models is that they demand adequate data in both quantity and quality, which are often lacking during the preliminary design phase of a tunnel. According to Chiu and Chia (2012), numerical methods are frequently applied when hydrological conditions are complex, as they allow for a more reliable prediction of groundwater inflow. Indeed, in cases where geological and hydrogeological conditions are highly intricate, numerical approaches are often regarded as indispensable for approximating groundwater inflows into tunnels (Frenelus et al. 2021).

Data-driven models (DDMs) represent a third possible approach to the preliminary evaluation of impacts, where the relationships between hazard and impact indicators, even if not physically resolved, are empirically derived from site-specific or hydrogeological setting-specific real data such as the statistical study of Cesano et al. (2000). Achieving high precision with these methods requires vast amounts of relevant data. Consequently, as with numerical approaches, factors such as cost and time consumption may represent significant obstacles to obtaining accurate predictions (Li et al. 2017; Mahmoodzadeh et al. 2021).

Comparisons between different approaches are few: Vincenzi et al. (2010) compared PTs (original version of DHI, Dematteis et al. 2001), DDMs (Cesano's approach), and PBMs (by the use of MODFLOW code) for a preliminary design of a hydroelectric water derivation tunnel through a crystalline (metamorphic) setting in the Western Alps, but without the chance to retroactively validate the assessment. On the other hand, several comparisons between methods for calculating tunnel inflows are available, based on physically based analytical and numerical approaches (Butscher 2012; Liu et al. 2021; Mahmoodzadeh et al. 2021).

In order to overcome the intrinsic subjectivity of more popular methods, like in other applications of environmental risk assessment where the complexity of the relationships between input (hazard and vulnerability parameters) and output (impacts and associated probability) is not manageable by process-based models, artificial intelligence or machine learning (ML) could represent a viable alternative to DDMs, only in cases where robust validation and training datasets are available (Dal Seno et al. 2024; Liu et al. 2021; Mondini et al. 2023; Orland et al. 2020).

The application of ML techniques in hydrological modelling has grown significantly over the past decade, driven by advances in data availability and algorithmic development (Rajaee et al. 2019; Refsgaard et al. 2022; LaBianca et al. 2024). ML methods have also undergone development, moving from classical methods to more complex and powerful ones. The contribution of ML is paramount in addressing data with complex, nonlinear relationships not easily habdled by PBM tools owing to factors such as heterogeneity, overall simplicity of conceptual models, and paucity of available data (Gonzalez and Arsanjani 2021). Moreover, ML methods typically require less computational effort compared to PBMs.

Machine language methods can address classification and regression issues, the first with the aim to assign a label or category to an observation, the second with predicting a continuous value based on input data, and are categorized as unsupervised or supervised (Ayodele 2010). Unsupervised learning works without external guidance, using a data-driven algorithm to identify patterns and relationships between variables, thereby grouping the dataset into clusters. In supervised learning, the algorithm correlates independent variables (parameters) with a target variable, such as the presence or absence of any effect originating from any hazard against receptors, e.g. contamination or damage. Existing hydrogeological applications of ML methods include: identification of groundwater potential zones in mountainous regions (Dahal et al. 2023), groundwater level predictions (LaBianca et al. 2024; Aderemi et al. 2023, Barzegar et al. 2017; Chang et al. 2016; Sahoo and Jha 2013; Tapoglou et al. 2014), groundwater vulnerability assessment (Afshar et al. 2007; Sajedi-Hosseini et al. 2018), groundwater budgeting (Gorgij et al. 2017; Pradhan et al. 2019), and groundwater quality and quantity (Nourani et al. 2024). For a review of supervised and unsupervised ML methods in hydrogeological applications, refer to Dhapre et al. (2025).

Among the most advanced ML tools are methods based on deep neural networks such as "deep learning" (Goodfellow et al. 2016). These models are composed of many layers that allow the network to learn more complex and abstract representations of the data. For example, among applications relevant in relation to the presented research, could be considered the contributions of Dal Seno et al. (2024), who compared statistical methods, classical and advanced ML methods, including artificial neural networks, for evaluating rainfall thresholds for the activation of landslides in Northern Apennines (Italy).

The paper aims to develop a data-driven ML–based method, applicable to various hydrogeological settings of fractured/karst aquifers, through the training and validation compared to published and verified data on springs impact, and contrasting it with an existing widely adopted risk assessment tool (DHI). In order to address the raised issue of binary classification—impact versus no impact on receptors based on various indicators of the hazard and vulnerability conditions—a supervised Random Forest (RF) classification model was adopted. The data from environmental monitoring of drilled tunnels, both located in Italy, were used as training and validation sets for the model: the Florence-Bologna high-speed railway connection (Northern Apennines) and the Gran Sasso highway tunnel (Central Apennines). Figure 1 illustrates the workflow adopted in

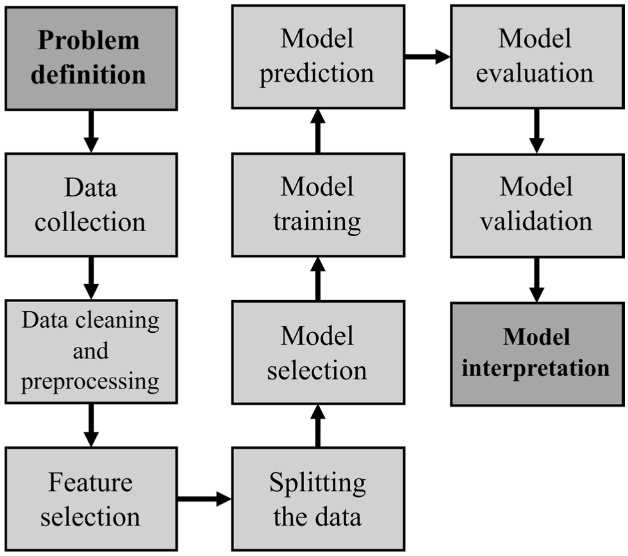


Fig. 1 Workflow for the development of a machine learning algorithm

this study, describing the steps followed for data processing and the training of the ML algorithm.

Materials and methods

In keeping with the main goal of the study, the availability of comprehensive pre- and post-construction monitoring datasets is essential for evaluating the real effects of tunnelling on groundwater flow and associated dependent ecosystems. Without such datasets, any meaningful or quantitative analysis would be severely limited or even impossible. Also, it must be emphasized that, for reasons of privacy and industrial confidentiality, it is not an easy task to identify free access datasets where spring impacts are examined according to objective scientific output.

Two main transportation tunnel projects, both located in Italy, have been chosen by the authors: the Bologna-Florence high-speed railway connection and Gran Sasso twin tunnels highway, both through the Apennines chain. Throughout the paper these two projects will be defined as BF (Bologna-Florence tunnel) or GS (Gran Sasso tunnels). The choice of BF and GS comes from the fact that they are among the few existing major tunnelling projects associated with detailed and published spring monitoring data both before and after construction. Both the general technical description of these two infrastructures and the results of spring monitoring are well documented by various papers and reports (Celico et al. 2005; Gargini et al. 2006, 2008; Petitta and Tallini 2002; Vincenzi et al. 2014), which provided an extensive database. The uniqueness and completeness of these data made them an ideal foundation for the authors' study, enabling a robust and scientifically sound evaluation of tunnelling impacts in complex geological settings.

The geological complexity and variability of the Apenninic chain have permitted the authors to apply the method in two very different hydrogeological settings. The Apennines represent a classic thrust-fold belt of sedimentary rocks, structured by a series of nappes that have been stacked upon each other as a result of the compressive forces generated by the collision between the African and Eurasian tectonic plates during the Alpine orogenesis (Boccaletti et al. 2011; Carminati and Doglioni 2012; Castellarin 2001; Vai and Martini 2013). Whereas the northern branch of the chain is lithologically dominated by sedimentary silico-clastic units (turbidites), along with clayey complexes (Ricci Lucchi 1978, 1981, 1986; Tinterri and Tagliaferri 2015), the central and southern branches denote a high proportion of karst carbonate aquifers (Celico et al. 2005; Boni et al. 1986). Both BF and GS were excavated according to a conventional blast-and-drill approach, thus they basically act as draining tunnels: BF affected groundwater flow systems of fractured sandstones of turbiditic units, whereas GS affected the main karst massif of the whole Apennine chain.

Bologna-Florence high-speed railway tunnel (BF)

The BF was drilled between 1996 and 2005. A total of seven main tunnels were excavated through and below the main divide of the northern Apennines, between Emilia-Romagna and Tuscany region for a total tunnel length of 73 km (Fig. 1). On the Emilia-Romagna side, the tunnels were built by direct excavation, owing to the more clayey and soft nature of the rocks, whereas on the Tuscany side they were built using the traditional blast-and-drill approach.

The dominant lithology along the line is represented by turbidites, alternance of arenitic and pelitic layers, both silicoclastic and marly-calcareous (Zuffa 1980; Ricci Lucchi 1986; Bruni et al. 1994; Zattin et al. 2000; Argnani and Ricci Lucchi 2001). These units were stacked, from Upper Cretaceous to Upper Pliocene, by the compression between Eurasian and African plates; subsequently, during the Late Pliocene to Quaternary, the chain was dissected by extensional faults, playing a significant role in the formation of intramontane basins, like the Mugello and Florence plains (Fig. 1). The conceptual model of groundwater flow systems in turbidites was proposed by Gargini et al. (2008) with the identification, through the aid of springs monitoring, of local (slope) and trans-watershed circulation patterns; with the latter often being controlled by post-orogenic extensional faults, which are the main drivers of tunnel drainage and spring impact (Vincenzi et al. 2009, 2014). The distribution of hydraulic conductivity in these units is highly heterogenous and anisotropic, being related to the pattern of fractures and to the ratio between arenitic (more fragile and prone to fracturing) and pelitic layers (A/P ratio; Antolini and Cremonini 1991; Benini et al. 1991; Gargini et al. 2006).

The considered sector of the project is related to three of the main tunnels on the Tuscan side (Vaglia, Firenzuola, Raticosa), the longest ones of the route are associated with huge groundwater inflows and environmental impacts linked to turbiditic aquifer units (Vincenzi et al. 2009, 2014). During the excavation of the tunnelsm relevant water inflows occurred (Table 1). The tunnels were designed to abate the hydraulic head in order to guarantee the stability and durability of the concrete lining. During the excavation, the recorded specific linear drainage rate of the tunnel depends on the progressive intersection of the main tectonic lineaments with aquifer role, represented by post-orogenic, enlarged extensional faults and fractures (Fig. 2).

From Table 1, it is apparent how Firenzuola and Vaglia tunnels resulted in being more productive in hydrogeological terms: the first one is linked to the main extensional tectonic lineaments and to the occurrence of higher A/P value turbiditic units, while the second one, is assoicated with the dissolution effect and consequent enlargement of fractures in the local calcareous turbidites. The groundwater drainage produced scattered and locally relevant impacts, like spring and stream seasonal or perennial drying-up and induced drawdown on public and domestic wells (Vincenzi et al. 2014; Canuti et al. 2009). Based on detailed monitoring of 160 springs, both in *ante operam* and *post operam* conditions, it was ascertained that a total of 68 springs were impacted by the tunnel drainage with various intensity, producing either a decrease in flow rate, low flow season (summer for the local climatic setting) or complete irreversible desiccation.

Gran Sasso highway tunnel (GS)

The GS is located in central Italy (Abruzzo region) and consists of two parallel one-way tunnels (twin tunnels), excavated using the traditional blast-and-drill approach between 1969 and 1982 (Fig. 3), and having approximately the same length in the range 10,175–10,121 m. GS connects the Tyrrhenian and Adriatic sides of the Central Apennines, along the highway from Rome to Teramo, through the upper part of the Gran Sasso carbonate massif, at an elevation range of 931–894 m asl.

The Gran Sasso Massif has the highest elevation in the whole Apennines (maximum elevation at Corno Grande, 2914 m asl) with a tectonic framework characterized by an imbricated structural system, formed during the Apennine orogenesis in the Late Miocene to Early Pliocene (Vezzani and Ghisetti 1998), with the massif overthrusted eastward above less permeable turbiditic and silty units. Subsequently, during the Late Pliocene to Quaternary, this sector of the Apennines was also affected by extensional faults, associated with the opening of the Tyrrhenian Sea, causing a differential tectonic subsidence in the southwestern sector of the massif (Cavinato et al. 1994; D'Agostino and Tozzi 1995). The extensional faults played a significant role in the formation of intramontane basins (Cavinato and Celles 1999), which developed from the Late Pliocene onward and were filled with unconsolidated fluvio-lacustrine deposits. From a hydrogeological standpoint, the Gran Sasso Massif is identified as a calcareous-karst hydrogeological massif covering an area of approximately 700 km^2 (Petitta and Tallini 2002). The karst dissolution, associated with geomorphological factors (intramontane basins, development of epikarst), is a relevant factor regarding the direct and focused recharge and the very high bulk permeability (Adinolfi Falcone et al. 2008, 2012; Amoruso et al. 2011, 2013, 2014; Barbieri et al. 2005; Boni et al. 1986; Ford and Williams 1989; Lorè et al. 2002). The massif is hydraulically partitioned by faults acting as hydraulic barriers, with strong internal variation of the hydraulic head (Barbieri et al. 2005; Celico et al. 2005; Amoruso et al. 2012). The total spring discharge of the whole karst massif has been estimated as being up to 20–25 m^3/s (Tallini et al. 2014) and is distributed throughout many springs surrounding the block along the permeability thresholds and having contact with lower permeability units; the main base springs of the system (San Calisto, Capo Pescara) are located at the southern edge of the massif.

The drilling of the tunnels, with a relevant hydrostatic pressure of around 600 m of groundwater, involved the northern and upper sector of the massif. The active recharge, associated with the high permeability and hydraulic loading, induced several strong inrushes, with flow rates exceeding 20 m^3/s (Celico et al. 2005). After the completion of the drilling, the twin tunnels continue to drain about 1.5 m^3/s, with two thirds on the Adriatic side and one third on the Thyrrenian side of the massif. Several main springs were impacted throughout the area, even at distances of about 35 km from the tunnels. The global impairment of total spring discharge was estimated in 7% of the pristine conditions (Petitta and Tallini 2002; Petitta et al. 2010, 2015, 2022); the high permeability and size of the aquifer, associated with the location of the tunnels at a relevant elevation (actually GS is not a base tunnel like BF), induced a spring

Table 1 Data on tunnel lengths and water inrush on the Tuscan side. The mean monthly discharge is calculated from the completion of the tunnel up to the last recorded data point

Tunnel	Tunnel length (km)	Peak inrush (m^3/s)	Mean monthly discharge (m^3/s)	Specific discharge m^3/(s × km)
Raticosa	10.4	0.018	0.029	0.0028
Firenzuola	15.0	0.5	0.375	0.025
Vaglia	18.5	0.018	0.125	0.0068

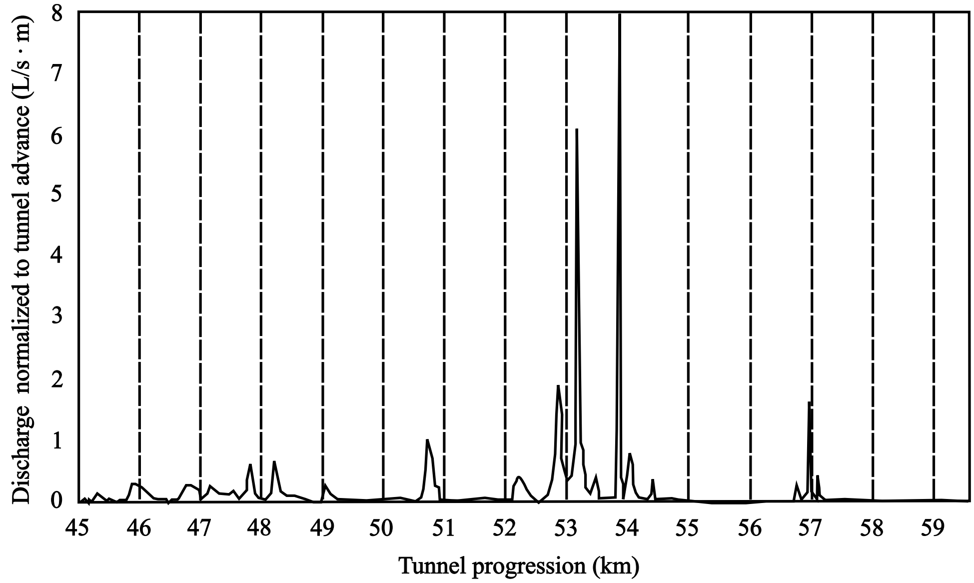


Fig. 2 Groundwater inflows during the excavation of the Firenzuola tunnel were observed at the intersections with tectonic structures. Modified from Canuti et al. (2009)

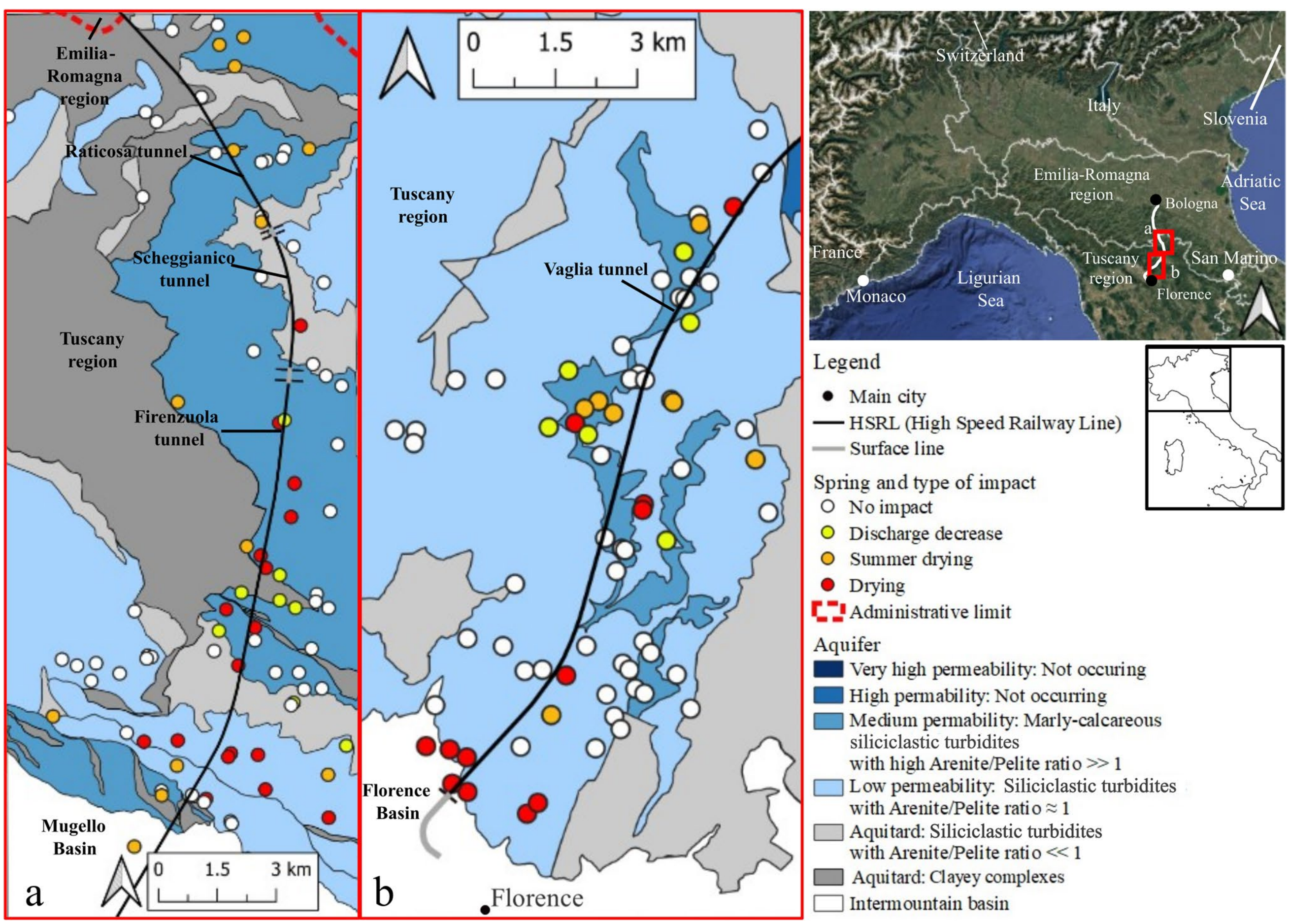


Fig. 3 Location of Bologna-Florence high-speed railway line, springs and hydrogeological map, hydrogeological interpretation based on the hydrogeological map 1:500,000 from WMS by Servizio Geologico d'Italia – ISPRA, with the degree of hydrogeological impact on springs as derived from environmental monitoring (geology redrawn, simplified, and modified from Vincenzi et al. 2014). **a** Map of the tunnels north of the Mugello basin. **b** Map of the tunnels south of the Mugello basin

discharge decrease but not a complete desiccation. In contrast to the BF project, environmental monitoring was not performed (the design is around 30 years older than BF), but it was possible to estimate the impacts through successive detailed hydrogeological studies comparing the springs' historical monitoring with modern collected data and groundwater budget analysis (Tallini et al. 2001). A total number of 213 springs were censored from historical data throughout the whole massif; however, only 50 of these were available for the impact evaluation by Tallini et al. (2001); of these, 28 springs were classified as affected by an impact (Fig. 3).

Choice of risk parameters

The hydrogeological interference between tunnelling (hazard associated with tunnel inflow) and springs (receptors at risk), acting in a complex and site-specific interplay between each other, is related to many different factors; these factors express the vulnerability of the receptors, e.g. the degree of hydrogeological connection between the hazard (tunnel) and the receptors. Lithological, sedimentological, tectonic, and hydrogeological elements concur, in a complex and setting-specific way, the determination of the associated hazard and vulnerability in relation to the intrinsic heterogeneity and anisotropy of the rock mass. On the other hand, in the preliminary stage of a tunnel design, the details of geological information and data are scarcel however, at the same time, it is urgent to appreciate in a robust way the possible associated risk.

The choice of parameters, or indicators, considered more relevant in affecting the risk associated with the impact of springs from tunnelling, was based on existing methods and the availability in the preliminary phase of tunnel designing. At this stage, only geologic and hydrogeological maps, along with geological sections, borehole logs and spring surveys, are available, whereas specific geotechnical and hydrologic observations during the excavation are completely missing. A total of 10 parameters were considered (Table 2), suggested mainly by the most adopted available parametric method—the DHI, as originally formulated by Dematteis et al. (2001) and successively reviewed or revisited by Dematteis et al. (2025), Attanayake and Waterman (2006), and Liu et al. (2015). Of the 10 parameters, four are related to hazard (tunnel inflow potential), and six are related to spring vulnerability (degree of hydrogeological connection between tunnel and spring). DHI, a parametric tool, was considered a starting point, owing to its completeness in the choice of parameters and the relative range; also, the aim of this study is to procure an assessment according to a fully data-driven process. A rating was assigned to each parameter according to a dimensionless scale ranging from 0.1 to 1 for hazard-related parameters and from 1 to 2 for vulnerability-related ones, as originally proposed by the DHI method. Two of the vulnerability parameters are dimensional and expressed in terms of their physical magnitude, while the last one (geology) is a categorical variable ranging from 1 to 5.

Hazard parameters

The hazard parameters drive the probability of a groundwater inrush in the tunnel. They refer to the event representing the hazard posing a risk to the receptor (spring). The trace of the tunnel is divided into 100-m-long sectors (as originally proposed by the DHI method), and the ratings are assigned to each parameter, considering the associated prevalence along the sector, as can be seen in the example in Fig. 4.

Fracture frequency (FF)

Fracture frequency (FF) influences both the bulk hydraulic conductivity of fractured aquifers (Snow 1969, 1970) and local increases in groundwater discharge associated with fault zones or intensely fractured rock. Major tunnel inflows often originate from enlarged fractures, fracture zones, or faults, which may act either as conduits or hydraulic barriers separating groundwater bodies with significant head differences. For example, during the GS tunnel construction, a sudden inrush encompassing tens of m^3/s occurred

Table 2 Parameters with the associated rating chosen as input for the ML method. All parameters, except EL and G, were initially considered by the DHI method, either in the original or revisited version (Dematteis et al. 2001,2025); some were modified or introduced by other authors (as indicated in the references)

Risk component	Parameter	Rating	Reference
Hazard	Fracture frequency (FF)	0.1–1	Liu et al. (2015)
	Rock mass permeability (MK)	0.1–1	-
	Overburden (OV)	0.1–1	Liu et al. (2015)
	Inflow potential (IP)	0.1–1	-
Vulnerability	Distance from the tunnel (DT)	Dimensional (m)	Attanayake and Waterman (2006)
	Differential elevation (EL)	Dimensional (m)	-
	Groundwater flow system (GFS)	1–2	Attanayake and Waterman (2006)
	Topographic effect (TE)	1–2	-
	Permeable channels (PC)	1–2	Attanayake and Waterman (2006)
	Geology (G)	1–5	Liu et al. (2015)

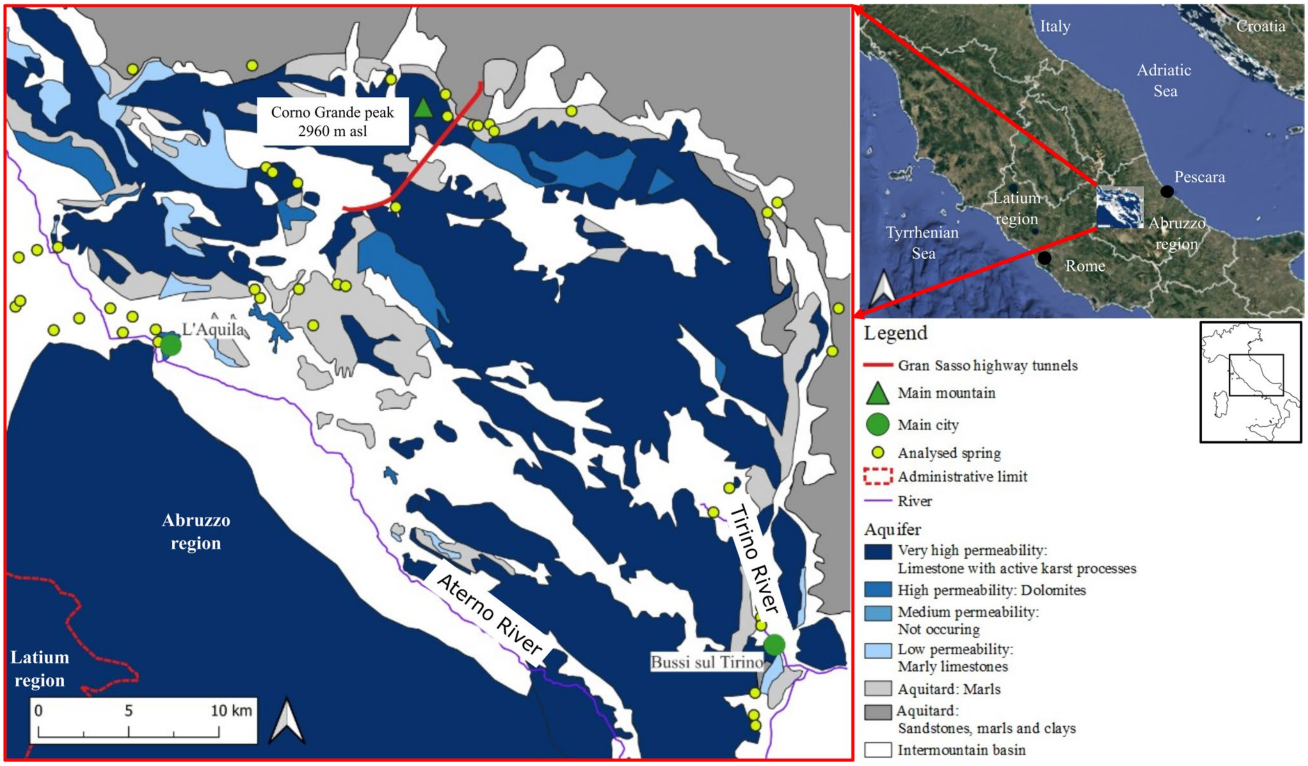


Fig. 4 Location of Gran Sasso highway tunnel, springs and hydrogeological map. The hydrogeological interpretation is based on the hydrogeological map 1:500,000 from WMS by Servizio Geologico d'Italia – ISPRA and Petitta and Tallini (2002)

at the Valle Fredda fault owing to a 200-m hydraulic head differential, causing casualties and complete tunnel flooding (Lunardi and Catalano 2006).

At the preliminary design stage, FF can be estimated from geological logs, outcrop surveys, or proxies such as the intersection with major tectonic lineaments. The FF rating ranges from 0.1 to 1 (Table 3), with intermediate values allowed. A minimum value of 0.1 is adopted, following the DHI method, assuming that a completely unfractured rock mass is unrealistic. An example of FF rating assignment is shown in Fig. 4a.

Rock mass permeability (MK)

The MK parameter represents the relative bulk permeability (K) of the rock mass within each tunnel sector. In preliminary studies, direct permeability measurements are usually unavailable; therefore, MK is assigned based on expert judgment using qualitative geological indicators. For example, in turbiditic formations, the arenite/pelite ratio may be used as a proxy, whereas in carbonate formations, the degree of karstification is relevant. These interpretations may be supported by available borehole test data (e.g., Canuti et al. 2009).

The MK rating defines five discrete hydrogeological units (Table 4): aquitard, fractured aquifer with low to high

Table 3 Rating values for FF. The indicated values, along with the criterion, are an example derived from DHI

Criterion	Rating
Lack of faulting and absence of karst corrosion	0.1
Low intensity of fracture and/or hydrothermal weathering	0.4
High intensity of fracture and/or hydrothermal weathering	0.8
Main fault and/or damage zone with fractures	1

Table 4 Rating values for MK (Freeze and Cherry 1979; Domenico and Schwartz 1997; Fetter and Kreamer 2021)

Degree of permeability	K (m/s)	Rating
Very low (aquitard units)	$<1\times10^{-8}$	0.1
Low (aquitard units)	$1\times10^{-8}<\mathrm{MK}<1\times10^{-6}$	0.25
Medium (poorly permeable aquifer)	$1\times10^{-6}<\mathrm{MK}<1\times10^{-5}$	0.5
High (permeable fractured aquifer – Low karstification degree (diffuse stage)	$1\times10^{-5}<\mathrm{MK}<1\times10^{-3}$	0.75
Very high (highly permeable fractured aquifer: Active karst (mature stage)	$>1\times10^{-3}$	1

permeability, and karst aquifer. No intermediate values are allowed. Reference hydraulic conductivity values are reported in Table 4. Despite its simplification, this approach is consistent with common hydrogeological mapping practices. An example is shown in Fig. 4c.

Overburden (OV)

The rock overburden above the tunnel (OV) indirectly affects inflow hazard, being inversely correlated to the aperture of fractures, and thus affecting the aptitude to produce groundwater inflows (Bandis et al. 1983; Barton et al. 1985). The criterion of rating assignment to different ranges of overburden (Table 5) is related to the decreased rate of fracture aperture versus depth (Bandis et al. 1983). The assignment of the rating is based on the preliminary altimetric profile of the tunnel in relation to the morphological section along the trace (Fig. 4a).

Inflow potential (IP)

The IP-derived parameter was introduced by Dematteis et al. (2001) and quantifies the likelihood of the groundwater inflow occurrence in a certain sector of a tunnel. IP is calculated following Eq. (1):

$$\mathrm{IP} = \frac{41 \times \mathrm{FF} + 22 \times \mathrm{MK} + 17 \times \mathrm{OV} + 20 \times \mathrm{PZ}}{100} \quad (1)$$

where FF, MK and OV are the previously described parameters, and the plastic zone (PZ) is defined as the thickness of the enhanced permeability annular halo induced by the drilling around the tunnel. PZ is affected by the diameter and type of excavation and the quality of the rock mass; for example, the tunnel boring machine (TBM) drilling approach produces a thinner plastic zone compared to those that are drilled and blasted. In the original formulation of the DHI method, the assigned value for PZ was equal to 0 or 1 if, respectively, the diameter of the expected plastic zone was considered lower or higher than twice the diameter of the tunnel.

In the original DHI method, IP is a matrix-based parameter derived from the combination of the rock engineering system evaluation (RES, Hudson 1992) and the fully coupled model approach (Jiao 1995). Inside the rock mass, a set of interactions between various parameters potentially affecting IP occurs that is based on single binary relationships, and concurrence via the iteration of mutual relationships. IP, also if derived and not independent, was considered as the input parameter, along with FF, MK and OV, because ML allows for this adding process, referred to as "feature engineering" (Zheng and Casari 2018) and potentially capable of enhancing the predictive model's performance. However, please note that the presented study does not use PZ as an independent parameter, like FF, MK and OV. This is because of the fact that the analysed tunnels were both excavated with a drill and blast approach and because the parameter was considered not to be so physically based and relevant as the other three. For the derivation of the IP parameter, a value of PZ = 1 was used for all tunnel sectors. An example of the IP rating assignment can be seen in Fig. 3d.

Table 5 Rating values for OV

Overburden thickness (m)	Rating
≥ 200	0.1
$200 > \mathrm{OV} \geq 100$	0.25
$100 > \mathrm{OV} \geq 50$	0.5
< 50	1

Vulnerability parameters

Vulnerability parameters define and quantify the hydrogeological connection between a receptor (spring) and a sector of the tunnel according to topographic and hydrogeological factors encompassed by the parameters listed in the following.

Tunnel–spring spatial relationship (distance, DT; elevation, EL)

The spatial relationship between the tunnel and a given spring is provided by two parameters: the geometric projected horizontal distance between the tunnel and the spring (DT) and the differential elevation of both (EL). "Tunnel" is considered to be the centre of the sector specifically involved in the evaluation. DT (m) refers to the minimum linear distance between the spring (location of the orifice) and the centre of the tunnel sector; EL (m) refers to the vertical component of the minimum linear distance between the spring (location of the orifice) and the center of the sector. Different from the original DHI approach, where the geometrical distance was expressed by a discrete rating based on the range of distances, here a continuous numerical count is introduced to improve the accuracy (Table 6 shows the rating of DHI approach).

Table 6 Rating values for DT

Distance (m)	Rating
> 800	0
$600 < \mathrm{DT} \leq 800$	0.1
$400 < \mathrm{DT} \leq 600$	0.2
$200 < \mathrm{DT} \leq 400$	0.4
$100 < \mathrm{DT} \leq 200$	0.8
≤ 100	1

Typology of groundwater flow system (GFS)

The GFS characterizes the extent, depth, and flow pattern of the aquifer feeding the spring. Shallow and local systems are more sensitive to nearby tunnels, whereas deep and regional systems may be affected even at greater distances. In the DHI method, GFS is rated in a simplified manner (1 or 2), depending on the aquifer volume involved.

For this study, GFS classification was derived from hydrological analyses of the BF case (Gargini et al. 2008; Segadelli et al. 2021) and from hydrochemical and isotopic surveys of the GS case (Tallini et al. 2000; Petaccia and Rusi 2008). Five original GFS classes were mapped to ratings 1 (local), 2 (regional), or 1.5 for intermediate cases (Table 7).

Topographic effect (TE)

The TE parameter considers the geomorphological location of the spring in relation to the tunnel. In Fig. 5, the rationale for the rating assignment is explained, according to the relative position of the tunnel and the spring in regards to the main axis of the valley and slopes. The two cases presented in Fig. 5 differ in terms of the location of the tunnel in respect to the valley bottom: shallow tunnel route through the slope (case 1) or deep tunnel route below the valley (case 2). The rating attribution is based on the modification proposed by Dematteis et al. (2025) regarding the original formulation of Dematteis et al. (2001).

Permeability channels (PC)

This parameter takes into account the occurrence of main conductive (enlarged) faults, fracture zones or karst conduits connecting the tunnel's sector and the receptor. The connection is considered active if PC is intercepted by the sector and contemporarily is located inside a 200-m buffer zone in respect to the spring. The choice of distance is derived from the size of spring-head protection zones defined with geometric criterion in order to propose a buffer along which it is more likely to be an hydrogeological interference between the spring and the permeability channel (Table 8).

The PC vulnerability parameter does not duplicate the FF hazard parameter because the first one affects the connection between the tunnel and the receptor, whereas the second one is in relation to the inflow potential of the sector.

Table 7 Rating values for GFS (right-hand column) with a comparison with the hydrological (BF site) or geochemical (GS site) classification (see main text for references)

Site classification	Group	Groundwater flow system	Rating
BF springs	S_H	Local high-elevation system	1
	S_L	Local low-elevation system	1.5
	S_S	Local slope system	1
	T_H	Regional high-elevation system	1.5
	T_L	Regional low-elevation system	2
GS springs	2	Local system	1
	1–4	Mixed contributions between shallow deposits and deep circuits	1.5
	3–5	Regional system	2

Geology (G)

The parameter is categorical and indicates the lithology associated with the aquifer discharging to the given spring. The rating is related to the five main lithologies associated with any aquifer unit (Table 9). Springs could be located inside outcropping areas of aquitards or aquitard units (like shale or mudstone), which are likely associated with shallow debris or regolith; in such a case, the minimum rating will be assigned. In other instances, springs could be discharging from talus or debris covering a bedrock and the assignment to the type of aquifer (unconsolidated or hard-rock below) could be uncertain; in that case a rating of 2 will be considered, with the GFS independent rating contributing to a better definition.

The values indicated here are indicative of the most common and general lithologies encountered in the training datasets. Any other possible lithology, not listed in the table, could be accommodated, always taking into account the scale between 1 and 5.

Collinearity and parameter selection

Multicollinearity occurs when two or more predictor variables in a model are strongly correlated. In this study, multicollinearity was assessed through the analysis of the correlation matrix, which measures the strength of linear relationships among variables. The results are shown in Fig. 6.

Overall, no strong correlations are observed among most variables, with the notable exception of IP and FF, which exhibit a very high correlation coefficient (0.96), which is expected given the formulation of IP, in which FF is the dominant contributing factor (Eq. 1). Despite this, all parameters identified in the previous steps were retained as features in the model. Random Forest is a tree-based algorithm, and decision trees are robust to variable scaling because splits are performed using threshold values rather than absolute magnitudes (Hastie et al. 2009). As a result, data normalization

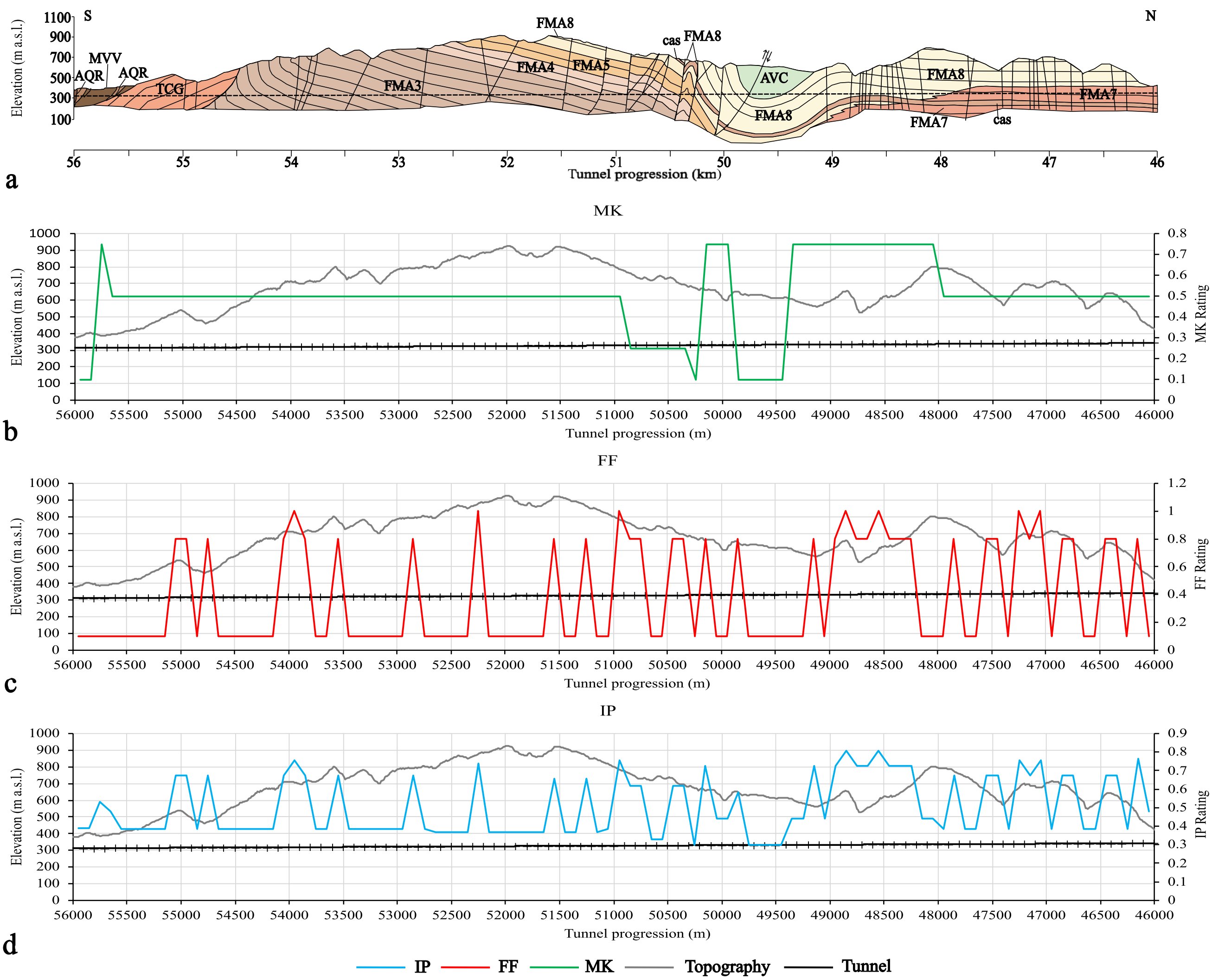


Fig. 5 Three hazard parameters from the DHI method applied to the Firenzuola tunnel (Canuti et al. 2009). The rock mass permeability (MK, green), fracture frequency (FF, red), and inflow potential (IP, blue) were calculated on 100-m sectors. Geology: FMA: Marnoso-Arenacea Formation; FMA3: Premilcuore member; FMA4: Galeata member; FMA5: Collina member; FMA7: Bassana member; FMA8: Nespoli member; MVV: Villore Varicoloured Marl Formation; AQR: Acquerino Formation; TCG: Carigiola Stream Formation; AVC: Varicoloured Clay-Limestone Formationcas: Casaglia chaotic Unit

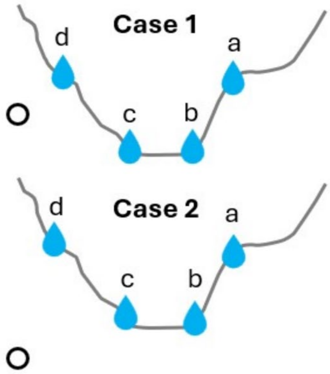
is generally unnecessary, and the algorithm is also relatively robust to multicollinearity. This robustness, however, may lead to some degree of informational redundancy. Nevertheless, all parameters were included to allow a direct comparison with the DHI methodology. The exclusion of additional geological and hydrogeological parameters (e.g., fracture orientation, recharge rate, aquifer thickness) that could play an important role is due both to their absence in the available datasets and to the intention of testing an approach closely aligned with that of the DHI.

Table 8 Rating values for PC rating

Permeability channels	Rating
No intersection	1
Intersection	2

Table 9 Rating values for G rating

Lithology	Rating
Not aquifer lithology (e.g. mudstone, shale, marls)	1
Unconsolidated deposits covering bedrock (debris, alluvium, colluvium)	2
Sedimentary or crystalline	3
Volcanic	4
Karst	5



Cases	Value
1a	1
1b	1
1c	1
1d	2
2a	1.5
2b	1.5
2c	2
2d	2

Fig. 6 Correlation matrix. IP: inflow potential; OV: overburden; MK: rock mass permeability; FF: fracture frequency; PC: permeability channel; GPS: groundwater flow system; G: geology; TE: topographic effect; EL: elevation; DT: distance

The Random Forest (RF) approach

To apply an ML-based data-driven tool for the risk assessment of a spring to impact induced from tunnelling, an RF algorithm was adopted. The process was applied in a Python environment. The RF algorithm is a ML-supervised learning method finalized for both regression and classification approaches. In this case, the classification algorithm was used.

Starting from an original dataset formed by specimens (elements) and the definition of possible target categories for each of them (for example, presence or absence of a certain consequence against a specimen), the RF algorithm will learn from a subset of the original dataset of elements, defined as the training set. It will be tested on the remaining part of the dataset, defined as the test set. In the specific application, the specimens or elements are represented by springs, and the binary target category is the presence or absence of an impact due to tunnelling.

Random Forest uses multiple decision trees (Fig. 7) and assigns the majority of the predictions to the predicted (more likely) category (Breiman 2001). The hyperparameters of the RF model are listed in Table 10. These parameters control the structure, randomness, and performance trade-offs of the forest, and are often tuned during model optimization. Their meaning and general ranges are presented, followed later by the main hyperparameters relevant to this study and the details of their fine-tuning process.

A certain ratio for splitting between the training and testing sets is chosen. The splitting proportion depends on the amount of available data and may range from 50/50 to 90/10% (Kuhn and Johnson 2013). In the presented case study, 70–30% is adopted (the dataset is of medium to small size).

A high number of simulation models is run (100 in the presented case study), each trained with a different association of specimens (fixed as 70% of total population of specimens) and validated against the complementary test set, changing the random seed each time for the choice of the specimens (Fig. 8). The random seed process splits the specimens randomly between the training and the test sets. In this way, the potential information provided by each specimen is randomly used in order to avoid losing any information about situations that cause impacts or nonimpacts to the receptors. The choice of the limited number of simulation models (100) is due to the relatively low number of specimens. The final output is represented by 100 different probability distributions linked to each simulation. Thus, each specimen (spring) will have an average and a standard deviation of the RF probability related to the 100 simulations. Each simulation corresponds to a different ML model. The data splitting was performed in Python using the train_test_split function, and the proportion between class 1 (impact) and class 0 (no impact) was preserved between the training and test sets by using the stratify argument. This approach also helps to reduce prediction bias. In the dataset, impacted sources represent approximately 30% of the total, indicating a moderate class imbalance.

The most essential hyperparameters were selected, along with their corresponding ranges for this model, as shown in Table 11. Tuning was performed using the GridSearchCV function from the Python package scikit-learn (Table 12). The ranges represent all possible combinations that were analyzed, while the final value corresponds to the one that maximizes the scoring metrics. Through trial and error, the tested hyperparameters were found to be the most sensitive for classification. Other hyperparameters, which were not fine-tuned, were kept at their default values. Although these could be included in the GridSearchCV function, their potential gains are marginal.

A split is performed at each node of the decision tree (Fig. 9); each node corresponds to a parameter taken into consideration for the evaluation and potentially leading to the final target category assignment. Each slice tries to unify the dataset towards one of the target categories, creating homogeneous subsets (Breiman et al. 1984). The target categories indicate the result, which is the answer that was sought (impact – no impact). There are different slicing criteria that measure the quality of the split according to the RF classification: the Gini impurity (GI), the entropy (H), and the log-loss (Table 10). Gini impurity (GI) is the default criterion and was considered sufficient as the splitting criterion for the measurement of the quality of the split; it represents the likelihood of misclassifying a randomly selected element from the dataset, according to the distribution of target categories within the dataset.

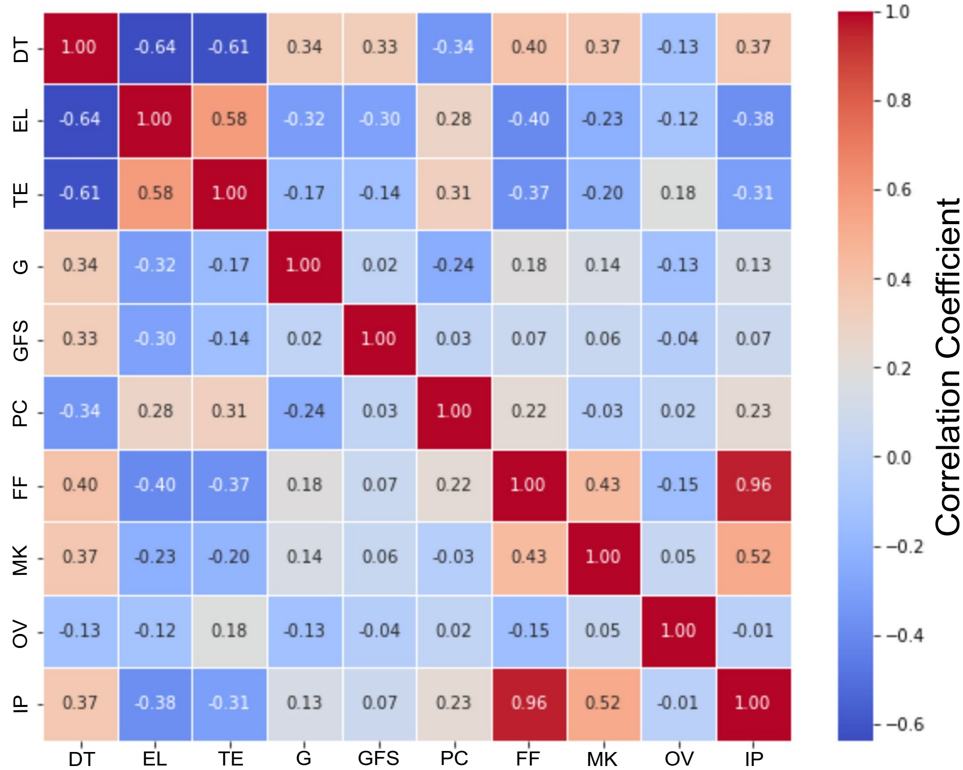


Fig. 7 A diagram summarizing the different steps for the creation and evaluation of a Random Forest model

The splitting in the decision tree aims to minimize GI, given by (2):

$$\mathrm{GI} = 1 - \sum_{i=1}^{n} p_i^2 \qquad (2)$$

where p_i is the probability of choosing the element i to be selected for splitting (Breiman 2001; Rodriguez-Galiano et al. 2014; Lauzon and Gloaguen 2024). At the end of the path, nodes are reached that indicate the outcome of one category over the other. The results and the classification of the predetermined number of trees were averaged.

A threshold must be chosen to indicate whether an element will be assigned the occurrence or the nonoccurrence of a certain consequence (impact or no impact). In the presented case study, a threshold of 0.5 has been adopted: this means that if > 50% of the trees leading to an adverse consequence (impact) were assigned to an element (spring), the given element was considered to be adversely impacted. With 100 models, the average of the RF probability values was computed. Some models indicated impact, while others indicated nonimpact regarding the same spring. Both pieces of information were considered when assessing the risk for the individual spring. For model evaluation, the standard scoring metrics listed in Table 12 were used. Given the presence of 100 different models, cross-validation was performed, and the mean and standard deviation of each scoring

Table 10 Overview of the main hyperparameters of the Random Forest Classifier (derived from Python package scikit-learn). Each hyperparameter is briefly described along with its recommended range of values

Hyperparameter	Description	Typical/default values
n_estimators	Number of trees in the forest	100 (default), common range: 100–1000
Slicing criterion	Function to measure split quality	‘gini’ (default), ‘entropy’, ‘log_loss’
max_depth	Maximum depth of a tree. Prevents overfitting	None, 10–100
min_sample_split	Minimum samples required to split a node	2 (default), 2–10
min_sample_leaf	Minimum samples required at a leaf node	1 (default), 1–5
min_weight_fraction_leaf	Minimum weighted fraction of sum of weights at a leaf node	0.0 (default), typically near 0.0
max_features	Max number of features to consider when splitting	‘sqrt’ (default), ‘log2’, float (0–1)
max_leaf_node	Grow trees with max number of leaf nodes	None (default)

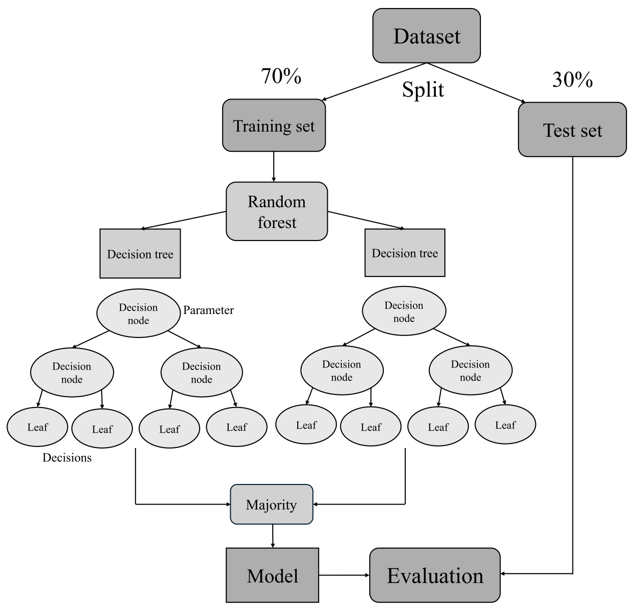


Fig. 8 Scheme of operation of the 100 simulations. The dataset is divided 100 times into training and test sets. Metrics are calculated for each simulation, and the mean and standard deviation are calculated. In addition, there will be 100 different probabilities of the RF, from which the mean and standard deviation will be calculated again. The final boxplot summarises the results of a single spring

metric were computed. The schematic representation is presented in Fig. 8.

Following the run of the model, the most relevant parameter to predict the final outcome category was chosen. There are several ways to calculate the feature importance and parameter effect of a model, both at a global level, with the aim of explaining the behaviour of the entire model across the whole data distribution, or at a local level, with the aim to explain individual predictions or classifications (Molnar et al. 2020).

At the global level, the permutation feature importance (PFI) was calculated. PFI is based on the idea of permuting the values of a parameter (referred to as "feature" in the definition) and observing how much the model's performance deteriorates (Breiman 2001). PFI is a method used

Table 11 Selected RF hyperparameters of RF with corresponding adopted values

Hyperparameter	Range (list)	Final value
n_estimators	440–500	490
max_depth	4–10	8
min_samples_split	2; 5; 10	2
min_samples_leaf	1; 2; 4	1
Max_features	"auto"; "sqrt"; "log2"	"sqrt"

Table 12 Main scoring criteria used for RF model tuning with GridSearchCV and for model robustness and evaluation

Scoring	Formula
Accuracy (*A*)	$\frac{TP+TN}{TP+FN+TN+FP}$
Precision (*P*)	$\frac{TP}{TP+FP}$
Recall (*R*)	$\frac{TP}{TP+FN}$
F1	$\frac{2\times P\times R}{P+R}$
Matthew correlation coefficient (MCC)	$\frac{TP\times TN-FP\times FN}{\sqrt{(TP+FP)(TP+FN)(TN+FP)(TN+FN)}}$

TP = true positive; TN = true negative; FP = false positive; FN = false negative

to assess the influence of each variable on a model's predictions. The idea is straightforward: each feature is taken in turn and its values in the test dataset are randomly shuffled, thereby breaking any relationship between that variable and the target. The model's performance is then recalculated; if the metric decreases significantly, it indicates that the model relies heavily on that feature, whereas a small change suggests a negligible impact. This approach provides an intuitive measure directly linked to model effectiveness and is applicable to any type of algorithm. In this case, accuracy was used as the performance metric. As a local method for evaluating feature effects, a Shapley additive explanations (SHAP) value was used. The SHAP value, derived from cooperative game theory, is a method for fairly distributing the contribution of each parameter in a predictive model. It quantifies the parameter's relevance by calculating its average marginal contribution across all possible combinations of parameters (Lundberg and Lee 2017). The dataset analysis followed a two-step process: first, the BF and GS monitoring datasets were analyzed separately; then, they were merged into a single dataset comprising 210 spring elements.

Comparison with drawdown hazard index (DHI)

As a comparison with the RF method, the DHI parametric tool was applied to the same datasets. The final output of the method is the DHI, derived as follows:

$$\mathrm{DHI} = \mathrm{IP} \times \frac{(\mathrm{GFS} \times \mathrm{DT} \times \mathrm{TE} \times \mathrm{PC})}{16} \quad (3)$$

where all the parameters have been previously described. DHI was evaluated for each spring according to the score indicated in Table 13. The thresholds were identified in Dematteis et al. (2001) by analyzing observed impact cases in the study areas. Both the RF and the DHI approaches rely on similar input data and adopt the same scoring framework.

For the comparison, the bootstrap confidence interval (CI) was adopted. The bootstrap CI is a statistical technique used to assess the variability of a metric without

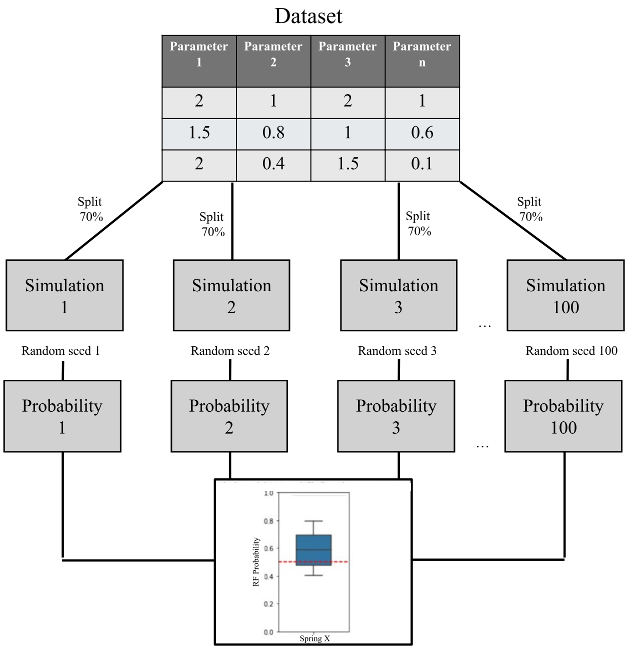


Fig. 9 Rating assignment for the GE parameter. Modified from Dematteis et al. (2025)

assuming a specific data distribution. It generates many resampled datasets with replacement, computes the metric on each, and derives an empirical distribution from which confidence intervals (typically 95%) are obtained. When comparing two methods, the bootstrap can be applied to the difference in metrics, providing a robust and intuitive way to determine whether one model significantly outperforms the other. Model performance was evaluated using F1, which balances precision and recall, and MCC, a robust metric for overall performance that handles class imbalance effectively.

Results

The dataset analysis followed a two-step process: a first application to the original datasets was treated separately, followed by a unified application to both datasets considered as a whole.

The probability distribution of impact associated with each application (probability mean and standard deviation for each spring), derived from the decision trees, was used to create the boxplots presented in Fig. 10. Three types of outcomes result in comparison with a 0.5 threshold (50% of decision trees addressing an impact):

- Group 1: springs with a distribution entirely above the 0.5 threshold (see Spring ID 2 in Fig. 10);
- Group 2: springs with a distribution entirely below the 0.5 threshold (see Spring ID 14 in Fig. 10);
- Group 3: springs with a mixed distribution (see Spring ID 17 in Fig. 10).

Figure 10 shows 30 springs, out of the total datasets; this is to show just one example of the results. The remainder of the boxplots are shown in Figs. S4–S11 of the electronic supplementary materials (ESM). The evaluation metrics of the application of the RF model to the BF and GS original datasets (first step) are presented in Table 14, along with the results of the merged dataset and the results of the DHI applied to the two case studies. The validation results for the two individual models are presented in Table 15. The first model (BF) was validated with data from the GS and vice versa.

The receiver operating characteristic (ROC) curve is a plot that illustrates the diagnostic ability of a binary classifier by showing the trade-off between the true positive rate (TPR) and the false positive rate (FPR) at different threshold settings. It helps compare models and evaluate the balance between sensitivity and specificity. The area under the curve (AUC) measures performance: the closer to 1, the better the model (Zweig and Campbell 1993).

The ROC curves related to the 100 models of the RF application, according to the two steps, are shown in Fig. 11, along with the quantification of the AUC of the 100 models' mean and standard deviation. In Fig. 12, the parameter importance and parameter effect analysis via the SHAP values and PFI values are shown. The graph associates high SHAP values with the impact category, while negative values are related to nonimpacts (the graph can also be set in reverse).

For the BF dataset, there are 48 springs from group 1, 81 springs from group 2, and 31 springs from group 3 (Fig. 10a). The AUC value is 0.90 ± 0.03 (Fig. 11a). The SHAP value (Fig. 12a) reveals the relevance of inflow potential (IP), followed by fracture frequency (FF), and, in third place, by the differential elevation (EL) parameter. High values of IP and FF and low values of EL are associated with positive SHAP values, corresponding to an impact outcome (Impact). Concerning the other parameters, a positive SHAP value is directly correlated with GFS, OV, and MK, whereas an inverse correlation is observed with DT values. The PFI method shows that IP, FF, and EL are the features that most affect accuracy (Fig. 12b).

Table 13 DHI classes and meaning (Dematteis et al. 2001)

DHI class	Index score	Probability of impact
1	≤ 0.16	Negligible
2	$0.16 < DHI \leq 0.24$	Low
3	> 0.24	High

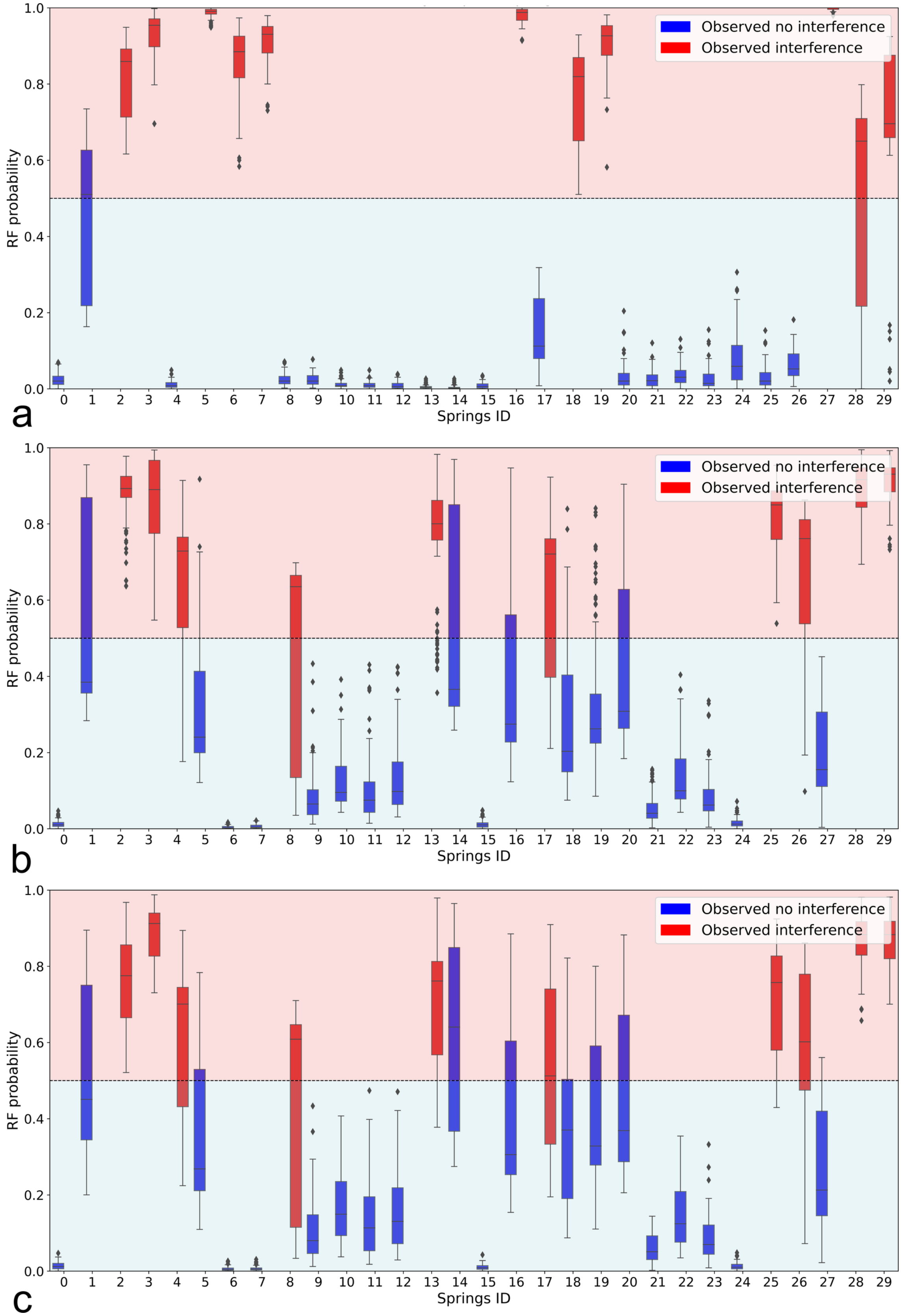


Fig. 10 Boxplot of the probability distribution of impact for each spring in the original dataset, based on 100 models. **a** First 30 springs of the BF dataset; **b** First 25 springs of the GS dataset; **c** First 30 springs of the whole (BF + GS) dataset. The actual observed impacts are indicated

Table 14 Evaluation metrics for Random Forest (RF) and drawdown hazard index (DHI) application to original datasets. The metrics for the DHI are based on different thresholds

Method	Case study	Accuracy (%)	Precision (%)	Recall (%)	F1 (%)	MCC (%)
RF	BF	83 ± 4	77 ± 7	80 ± 9	78 ± 6	64 ± 9
	GS	92 ± 6	92 ± 7	94 ± 9	93 ± 6	85 ± 11
	BF + GS	85 ± 3	80 ± 5	86 ± 6	83 ± 4	70 ± 7
DHI	BF 0.24	70	70	48	57	37
	BF 0.16	70	61	74	67	41
	GS 0.24	62	100	32	48	41
	GS 0.16	68	100	32	60	50

For the GS application, there is an improvement in the evaluation metrics results regarding the BF case (Table 14). There are 25 springs from group 1, 20 springs from group 2, and 5 springs from group 3 (Fig. 10b), and the AUC value is 0.99 ± 0.02 (Fig. 11b). The SHAP analysis (Fig. 12c) results differ from the BF case dataset: the most relevant parameter is GFS, followed by DT and MK, whereas the least important ones are OV, FF, and G. A high GFS value is positively correlated with SHAP values, as is the case for MK and DT. The PFI method shows that accuracy changes the most with GFS, EL, and DT (Fig. 12d).

In the second step, the same procedure was carried out by combining the two datasets (Fig. 10c; Fig. 11c). The AUC value is 0.93 ± 0.03. The SHAP analysis (Fig. 12e) indicates that IP has major relative weight, followed by FF and then GFS. The least indicative value for prediction is OV. A direct correlation with the SHAP values is derived from IP, FF, GFS, CP, MK, TE, and OV. Regarding EL and DT, the values are more dispersed compared to the analysis of the first step on separate datasets. Concerning the sensitivity analysis (PFI), the parameters that most influence the A value are again IP, followed by GFS and FF (Fig. 12f).

For each spring, the associated impact risk was assessed according to two approaches: the first, considers the "impact" values exceeding the "high" probability threshold of DHI (0.24, Torri et al. 2007); the second, considers the "impact" values exceeding the "Low" probability threshold of DHI (0.16, Torri et al. 2007). In this way, on the basis of accuracy, precision, recall, F1, and MCC, it was possible to adopt the same evaluation metrics for both approaches.

The results of the comparison are shown in the lower half of Table 14. The evaluation metric values are higher when the 0.16 threshold of low probability is used. Figure 13b shows, using a Radar Chart, the graphical comparison between the ML model and the DHI approach. Bootstrap confidence intervals (CIs) were computed to assess the performance of the RF model against the DHI-based methods using thresholds of 0.16 and 0.24, with both the MCC and F1 score. The results indicate that RF consistently outperforms both DHI thresholds across all metrics. Specifically, for MCC, the mean differences were 0.286 (95% CI [0.119, 0.451]) for RF vs DHI, 0.16 and 0.276 (95% CI [0.111, 0.438]) for RF vs DHI 0.24, indicating statistically significant improvements. Similarly, for the F1-score, RF achieved a mean gain of 0.131 (95% CI [0.035, 0.224]) against DHI 0.16 and 0.119 (95% CI [0.028, 0.209]) against DHI 0.24. In all cases, the confidence intervals did not include the value "zero", confirming the superiority of RF. MCC was prioritized as a robust global measure, particularly suitable for unbalanced data, whereas F1 provided a balanced assessment of precision and recall. The results are summarized in Table 16.

The average ROC curve of the combined approach (the mean value of the 100 models) was subsequently used to select three relevant thresholds. Several threshold triplets were randomly generated, along with the associated partial ROC curve. The optimal set of thresholds was identified by maximizing the AUC value of the best-performing partial ROC curve segments (Titti et al. 2022). The values of the three selected thresholds are reported in Table 17, while the segment fractionation of the ROC curve is illustrated in Fig. 14. An average probability value (APV), with a corresponding risk class, is associated with each analysed element (spring) in relation to the established thresholds. The results of the classification for all springs can be found in Table S1 and Figs. S5–S11 of the ESM.

Table 15 Validation values of the two individual models against the other dataset

Case study	Accuracy (%)	Precision (%)	Recall (%)	F1 (%)	MCC (%)
BF validation	56 ± 5	56 ± 3	95 ± 9	70 ± 4	3 ± 15
GS validation	61 ± 2	52 ± 6	25 ± 8	33 ± 6	12 ± 4

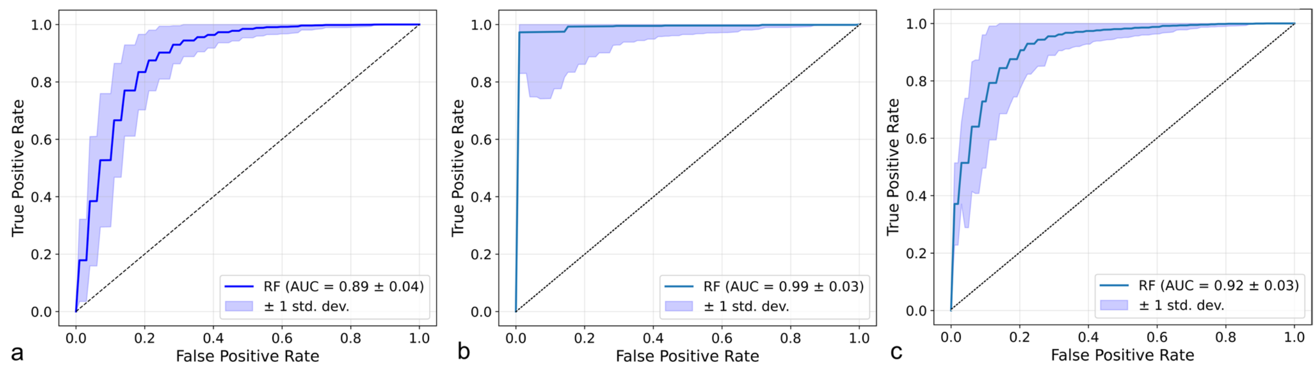


Fig. 11 ROC curve for the three RF model applications: **a** BF; **b** GS; **c** BF+GS. The blue line represents the mean of the 100 models, and the light blue fringe denotes the standard deviation envelope around the mean

Sensitivity analysis and discussion

The comparative analysis of the results of different RF models and of the relative relevance of parameters in affecting the impact probability sheds a data-driven light on the relationship between hazard and receptors in the risk assessment of hydrogeological interference induced from tunnelling.

The sensitivity (*S*) of each parameter in affecting the output of the three RF models and the DHI method is shown in Fig. 15. The sensitivity, or relative weight of the parameter, was calculated as follows: for the RF model, SHAP values were employed. The SHAP values were summed, and each parameter was subsequently normalized. For the DHI, the variation between the minimum and maximum value of each parameter was calculated and normalized with respect to the maximum possible DHI value.

Sensitivity analysis for BF model

The relative contribution to *S* from the hazard and vulnerability parameters is, respectively, 67.5 and 32.5%. In terms of hazard, the main role, and relatively major *S* value, derives from parameter IP and FF, with values respectively of 38 and 23%. This is hydrogeologically reasonable because, in the framework of the risk paradigm, they are connected to the magnitude of the hazard (e.g. higher likelihood of water influx in the tunnel strongly related to the degree of fracturing). The dominance of FF in affecting the hazard, in respect to MK (much less relevant with an *S* value of 2%), is likely due to the geological setting, dominated by turbiditic sedimentary aquifers. These aquifers consist of the alternance of arenitic and pelitic layers with a strong contrast in geomechanical properties where the first potential aquifer is more prone to brittle behaviour and the second is more ductile. Moreover, it was found that faults, with associated higher fracture frequency, are the elements that guide the more active groundwater inrushes in the tunnel crosscutting the effects of the permeability of single rock mass units. Vulnerability parameters appear to play a more marginal role, with relatively greater importance assigned to EL (9%) and DT (6%); by the way they are now differentiated from each other, in terms of relative weight, respect to the DHI approach. TE and MK emerge as the least influential parameters, with an S value lower than 3%. The relationship between each parameter and the impact probability is always directly proportional.

Sensitivity analysis for GS model

The relative contribution to *S* from the hazard and vulnerability parameters is, respectively, 28.5 and 71.5%; thus, owing to the karst nature of the setting, the hydrogeological connection plays the dominant role. This is important because it demonstrates that two very different geological settings have been analysed; thus, the final global approach could be considered more robust. The general relevance of hazard parameters decreases in respect to the BF model. The major relative weight is now provided by MK along with IP (both with an *S* value of 14%), with an associated decreased sensitivity for FF. The paramount role of karstification in the GS setting pushes MK to the upper rim of the specific *S* range for hazard with the main role in affecting the IP (a single maximum value of MK was assigned to all sectors of the tunnel owing to the relevance of dissolution in the carbonate units). Tunnel sectors are so far from many impacted springs that the relative weights of sector-specific hazard parameters, like FF and OV, is poorly relevant. A different combination of the relative weight of parameters is also observed for vulnerability. The most relevant parameters of all, including also the hazard, are GFS with an *S* value of 27% (in respect to 5% for the BF site); the typology of groundwater

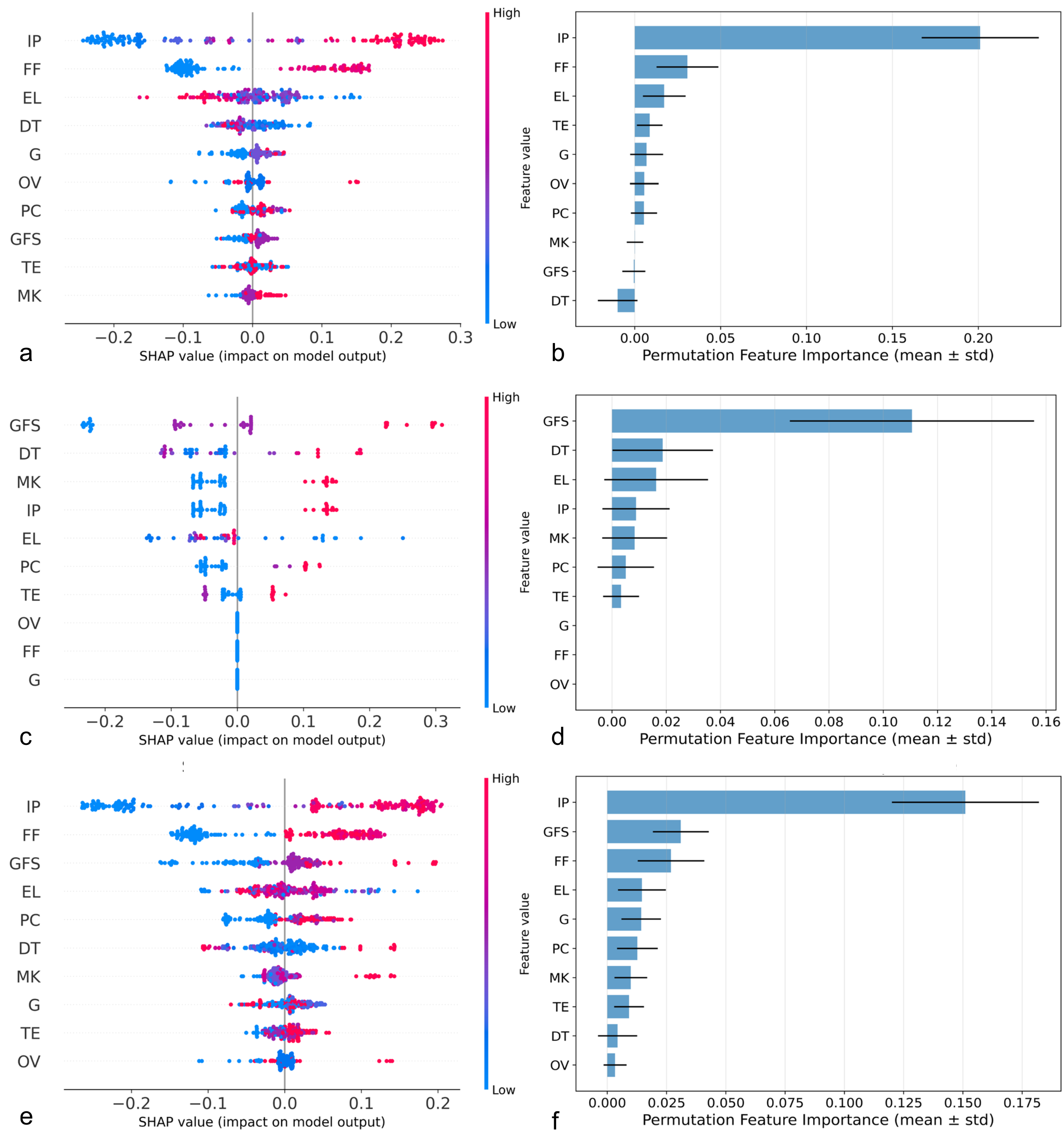


Fig. 12 SHAP values for the BF model (**a**), GS model (**c**), and the combined BF+GS model (**e**); and PFI values for the BF model (**b**), GS model (**d**), and the BF+GS model (**f**). IP: inflow potential; OV: overburden; MK: rock mass permeability; FF: fracture frequency; PC: permeability channel; GPS: groundwater flow system; G: geology; TE: topographic effect; EL: elevation; DT: distance

flow systems becomes the key vulnerability factor affecting impact to springs, regarding the simple geometrical distance between tunnel and receptors (still however relevant). Most of the more productive impacted springs are connected to the regional GFS and located far away from the tunnel at a distance in the range of 15–35 km. The relative weight for PC is also higher in respect to the BF setting (*S* value respectively 11 and 4%), which is the consequence of the unfeasibility of an univoque assignment of the spatial relationship between a sector and a spring (for

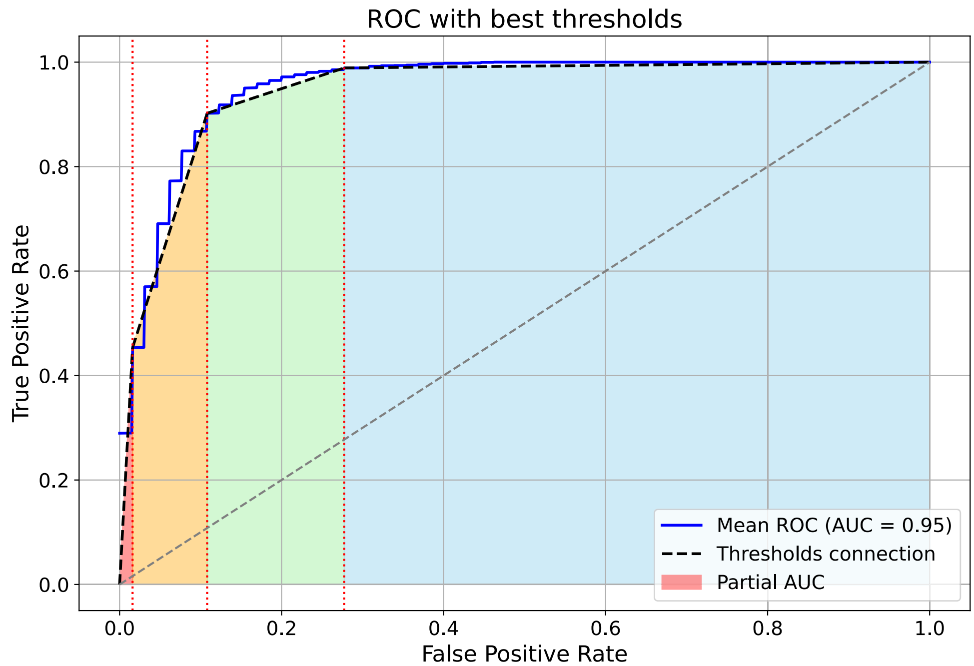


Fig. 13 Identification of the thresholds (red dotted lines) that maximise the sum of partial AUC values in relation to the average ROC curve

the high degree of hierarchization of the karst groundwater flow system). Thus, the most conservative approach was chosen by assigning the maximum rating to all the springs. The relationship between each parameter and the impact probability is always directly proportional with the exception of vulnerability parameters EL and DT (Fig. 12); this is due to the occurrence of negative EL values—i.e., when the tunnel is located at a higher elevation than the spring, with a consequent reduction of the impact probability, and the presence of impacts in the distal springs, at the termination of the flow system, is located far apart.

Sensitivity analysis for the "combined" model

The setting-specific role of each parameter is obviously smoothed down. The relative contribution to the *S* from hazard and vulnerability parameters is, respectively, 50.5 and 49.5%. In terms of hazard, IP and FF parameters are the most relevant, whereas OV is the least relevant. In terms of vulnerability, GFS is the most relevant parameter, followed by the geometrical ones (EL and DT). The relationship between each parameter and the impact probability is always directly proportional, with the exception of EL and DT, as derived from the GS application as previously explained.

Sensitivity analysis for DHI method

The comparison with the RF models is shown in Fig. 15. The relative contribution to the *S* from hazard and vulnerability parameters is, respectively, 32 and 68%. The inflow potential (IP) is the most relevant in terms of hazard, whereas, among the vulnerability parameters, all (TE, GFS, DT, PC) have the same *S* value (17%), but are much higher in respect to the RF approach (every model). The DHI method dampens out all vulnerability parameters at the same weight, with all the four parameters having values ranging from 1 to 2, which affects the final DHI value in the same proportion.

Results of the SHAP and PFI plots

The two case studies provide unique insights. In the BF case, the SHAP plot and PFI indicate that the most informative

Table 16 Results of the bootstrap confidence intervals between the RF and the two DHI thresholds

Metric	DHI threshold	ΔRF-DHI (mean)	95% CI
MCC	0.24	0.286	(0.119, 0.451)
MCC	0.16	0.276	(0.111, 0.438)
F1-score	0.24	0.131	(0.035, 0.224)
F1-score	0.16	0.119	(0.028, 0.209)

Table 17 Threshold values and corresponding categorization into risk classes

Class	APV threshold values range
A	0–0.16
B	0.16–0.45
C	0.45–0.78
D	0.78–1

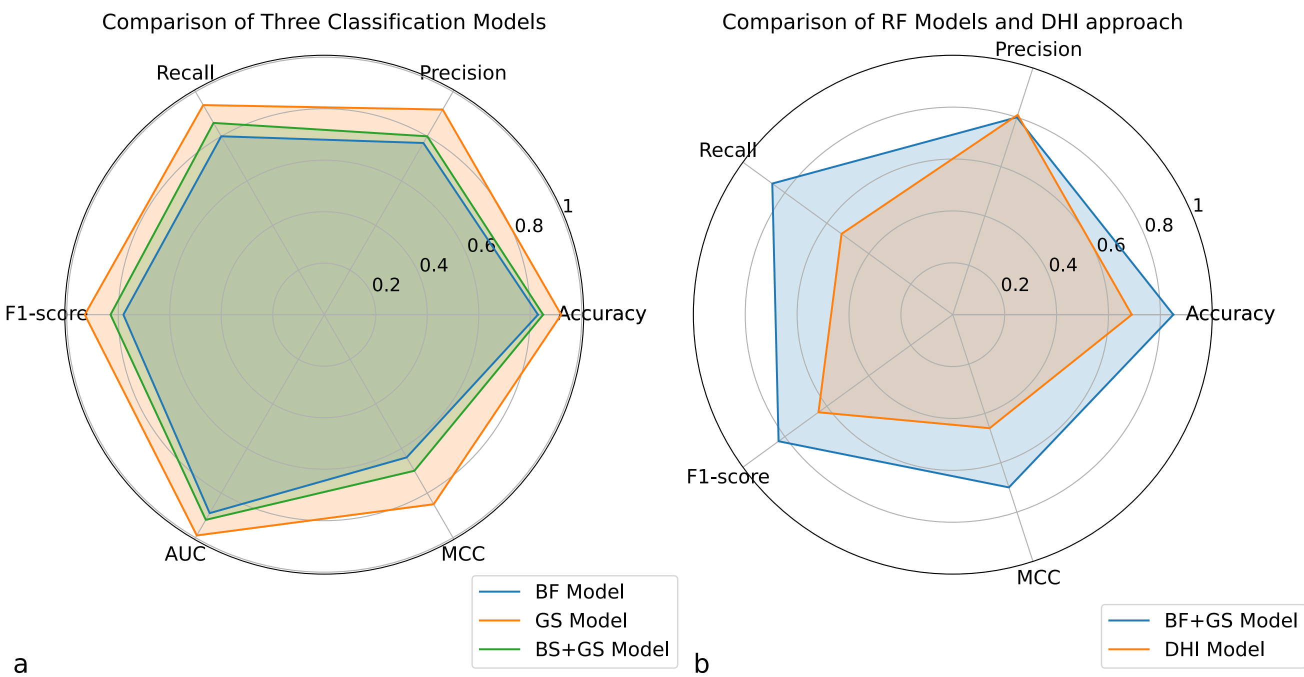


Fig. 14 Radar chart relative to: **a** Comparison of three RF classification models; **b** Comparison between the global (BF + GS) RF classification model and DHI parametric method

parameters are IP and FF. These are associated with relatively low-permeability units (alternation of arenitic and silty-clayey layers), but with a relevant occurrence of extensional fault zones, which increase the FF locally and frequently, and, consequently, the IP values. Most groundwater flow paths are therefore established within these fractures; when the tunnel intersects them, hydraulic connections are created, leading to impacts on the springs. In the

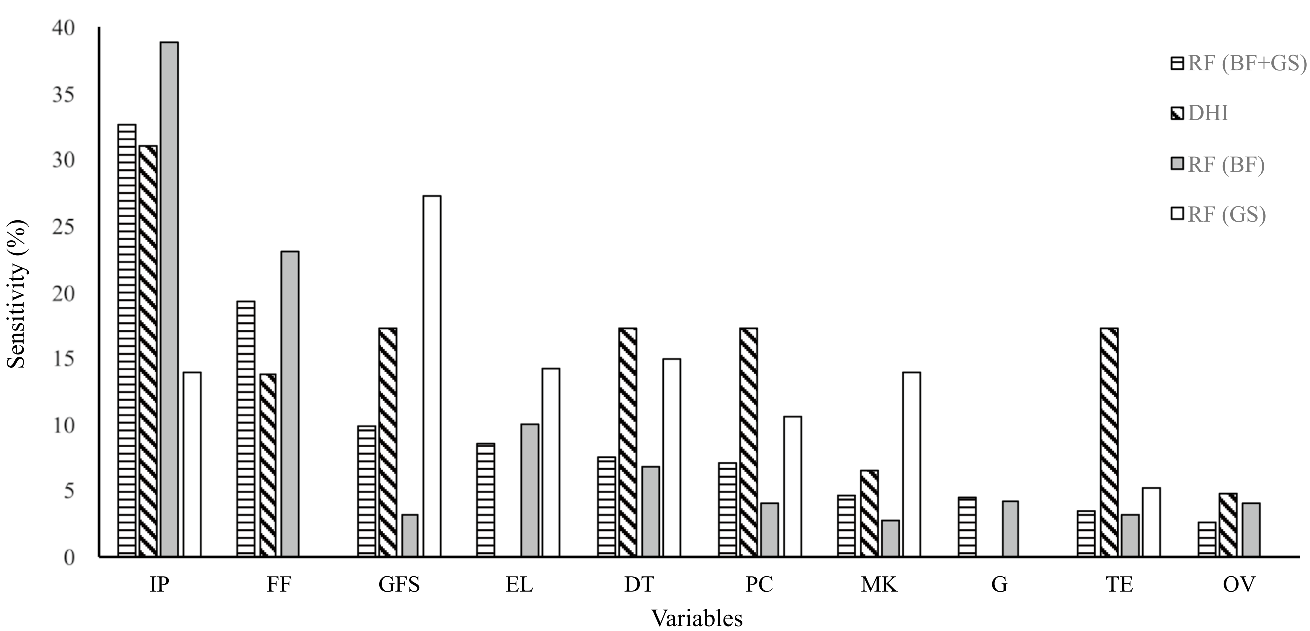


Fig. 15 Sensitivity levels of each parameter in the three RF models and the DHI. Some parameters are absent in the DHI, while others are absent in the GS because they have no decision-making weight in the models. IP: inflow potential; OV: overburden; MK: rock mass permeability; FF: fracture frequency; PC: permeability channel; GPS: groundwater flow system; G: geology; TE: topographic effect; EL: elevation; DT: distance

GS case, the most influential parameters are GFS and DT. This does not imply that IP and FF are unimportant; rather, in the GS case study, their values are consistently high, resulting in limited variability because they are categorical variables. Accordingly, when the model trained on the BF dataset is applied to the GS case, it predicts all springs as impacted, reflecting the uniformly high IP and FF values. Under conditions of similarly high IP and FF, the other parameters become discriminative between those springs impacted and those nonimpacted.

For almost all parameters, a single directional relationship is observed: in the SHAP plots, increasing parameter values correspond to higher impact probabilities. In contrast, DT and EL exhibit a bidirectional behaviour. For DT, this is related to specific cases, such as the Tirino springs, located approximately 35 km south of the tunnel, which nonetheless experienced an impact on discharge. Although impact probability generally decreases with increasing distance, the Gran Sasso tunnel affects a regional karst aquifer, allowing even distant springs to be affected. EL, by contrast, is inversely related to impact probability: Smaller vertical distances are associated with higher likelihood of impact. However, this relationship is also affected by the horizontal distance from the tunnel: proximity increases the plausibility of hydraulic connection, whereas greater distances reduce the likelihood that a spring belongs to the same flow system affected by the tunnel.

Model performance evaluation

The metrics value of the RF model results is higher with respect to the DHI application with both chosen thresholds (Table 14). Concerning DHI, if there is also an increase in recall as the impact threshold is lowered, a parallel decrease in accuracy, and thus in the occurrence of false positives, is observed. This indicates that an increase in the DHI value corresponds to a higher probability of impact; however, a high DHI value alone does not necessarily indicate a source that can be impacted. Failing to distinguish a high DHI value for one parameter from that of another may result in high predicted probabilities where no actual impact is observed. The analysis of the ROC curves (Fig. 11) indicates a high level of model performance across the three case studies, as expressed by the proximity to the unity value for AUC. The BF validation shows better metrics, with a high recall. For the GS model, the ROC curve shows exceptionally good performance with really high evaluation metrics, but the effect of model overfitting must be taken into account, employing caution in the analysis. Substantially, the model is able to detect almost all impacted springs (high recall), but with quite low precision in other situations (Table 15), and the validation yields extremely low values for various metrics (accuracy, precision, F1, MCC). In this situation, usually the model would be considered overfitted, because it fails to generalize its application to case studies other than its own.

Nevertheless, the reciprocal validation of the two separate models does not provide good metrics (Table 15), owing to the fact that the associated impacts are controlled by different processes. The need to unify the models arises from the necessity to merge together the contributions of parameters with different relative relevance in the different settings. The ROC curve of the unified model is slightly better than the single BF case and worse than the GS case. Unlike GS, however, the model does not appear to be overfitted, as inferred from the shape of the ROC curve.

Risk class

The classification of the different springs involved in the analysis into distinct impact-risk classes is necessary for the cost–benefit assessment of an underground project. The potential occurrence of an impact must be addressed either through the substitution of the resource (if exploited) or by introducing modifications to the project design. To this end, the ROC curve was used to identify the different thresholds. The division of the average ROC curve is based on changes in its slope, indicating points where the true positive rate slows down, while the false positive rate increases. The three thresholds and the corresponding three classes in Table 17 were used to define four distinct risk classes (Table 18). Furthermore, by using these thresholds, it is possible to assess the position of each spring in the probability boxplot and to quantify the associated uncertainty.

Conclusions

A data-driven machine learning Random Forest (RF) model for the risk assessment of hydrogeological interference (impact) against springs derived from tunnelling has been provided. The model is based on entry values of parameters related either to the hazard, e.g. the probability of having a relevant water inflow in the tunnel, or of the vulnerability of the receptors exposed to risk (springs), related to the degree of hydrogeological connection between the tunnel and the receptor. The parameters are derived from existing,

Table 18 Risk classes. Classes are taken from Table 17

Class	Risk
A	Very low
B	Low
C	Medium
D	High

not physically based parametric approaches commonly used in professional applications (DHI method), alongside the introduction of additional ones, not taken into account by DHI, but considered relevant regarding the possibility of affecting both hazard and vulnerability.

The model was trained and calibrated independently, versus two datasets originating from the environmental monitoring of spring discharge associated with the tunnel excavation in two different geological settings: a sedimentary fractured aquifer and a carbonatic karst aquifer. In both cases, actual hydrogeological interference was verified by different approaches through a comparison with *ante operam* (before excavation) observed conditions. Site-specifically trained models show good performance, but they appear to be site-dependent and overfitted. This is particularly evident in the karst case application, where the model, on its own, fails to have predictive utility.

The consequent merging of the two separate models into a unified one helps to balance out discrepancies, also with the inclusion of a new parameter that is related to the lithological composition of the aquifer and contributes to improving the final predictability score. The predictive power of the unified model is good, as evidenced by the metrics in Table 14; furthermore, the model demonstrates a strong performance as indicated by a high AUC. The most relevant parameters affecting the hazard are the IP, strongly controlled in a combined way by fracture frequency and rock mass permeability, whereas the most relevant parameters in affecting vulnerability are the typology of the groundwater flow system discharging to the spring and the geometrical relationships (distance, relative elevation) between the spring and the tunnel. The results of the ML model have been compared with the output of the parametric DHI, unravelling the original subjectivity of weights assigned to different parameters.

Machine learning generally demonstrates higher accuracy compared to the DHI, according to the key evaluation metrics. Both approaches (data-driven and parametric) exhibit strong and comparable sensitivity of parameters related to the hazard component of hydrogeological risk, particularly within the IP. Conversely, a marked difference arises in relation to the affecting vulnerability parameters: While the relative weight provided by the DHI method to the four parameters is equal, the ML approach emphasizes the role of groundwater flow systems associated with springs over other purely topographical or geometrical parameters, related to the distance or differential elevation between the spring and tunnel. These differences lead to divergent prediction outcomes between the two approaches.

To the best of current knowledge, this represents the first release of a data-driven general-purpose model to assess impacts from tunnelling. The presented approach, useful for a preliminary assessment of the associated risk of impact, deserves further validation with other settings but has the advantage of being data-driven; thus, in some way, it is physically based on the complex hydrogeological processes affecting the transfer of hydrogeologic impact from the tunnel to the receptors (springs). The adopted parameters are tailored to base tunnels that could interact with deep groundwater flow systems. For the method to be applied to different scenarios, the parameters should be reweighted accordingly—for example, a shallow tunnel may have a greater impact on shallower groundwater flow systems compared to deeper ones. Nevertheless, the method provides a reliable risk analysis that supports the designer in the early stages of the project, enabling the effective identification of high-risk areas where further investigations can be subsequently focused. The presented approach could be considered a starting point subjected in the future to continual improvement with the addition of new datasets, coming from different geological settings.

Although ML is often considered a black box, and so not able to identify and quantify the physical process explaining a given observation, in many applications involving the hydrology of groundwater flow, where complexity, anisotropy and heterogeneity are dominant, ML is surely a powerful tool to unravel the mystery associated with the assessment. For the case of hydrogeological interference by a tunnel, up to now, in the early stage of the design, approached only by subjective parametric methods, the objective quantification of the relative weights of parameters that affect the risk is relevant and contributes to identifying which are the key indicators deserving better consideration in the early investigations. In a complex scenario such as tunnel excavation, finding ways to simplify information and identify connections between different parameters is extremely important.

Limitations and future work

The analysed case studies are linked to the complex geology of the Italian territory and are constrained by the availability of well-documented and objective investigations on actual impacts and by the disclosure agreements of main projects. The Bologna-Florence (BF) case is associated with the hydrogeology of turbiditic complexes, while the Gran Sasso (GS) case involves a carbonate system. In studies set in metamorphic and/or volcanic environments, the methodology may produce potential false positives and false negatives. Nevertheless, it is believed that in similar geological settings, the approach could achieve reliable results. The presented research should be considered a starting point of an ongoing job; any further availability of well documented and verified investigations is welcome in order to validate the model, both in cases with comparable geological conditions and in settings with different geology. In the latter scenario, if performance is low, a detailed analysis could identify which parameters are most informative. The underlying idea is to demonstrate that, with appropriate data, the model can outperform one of the most commonly used methods. Improving pre- and post-operational monitoring would therefore contribute to the creation of a higher-quality dataset.

Another critical aspect is data uncertainty and accessibility. The two case studies benefited from comprehensive hydrogeological and geological data coverage. High-quality geological sections and environmental monitoring allowed for precise identification of the relevant parameters for each spring. In studies with more limited environmental data, there is a significant risk of error, particularly regarding the GFS parameter. This parameter uncertainty must be coupled with careful interpretation by the user. In ambiguous situations, adopting the most conservative parameter estimates to assess impact probability may be advisable, subsequently guiding further data collection in higher-risk areas.

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/s10040-026-03102-6.

Author contribution All authors contributed to the study conception and design. Material preparation, data collection and analysis were performed by Ernesto Pugliese, Alessandro Gargini and Dany Lauzon. The first draft of the manuscript was written by Ernesto Pugliese and all authors commented on previous versions of the manuscript. All authors read and approved the final manuscript. Alesandro Gargini was the general supervisror of the entire project.

Funding Open access funding provided by Alma Mater Studiorum - Università di Bologna within the CRUI-CARE Agreement. This work belongs to the PhD project of Ernesto Pugliese “Sviluppo e calibrazione di protocolli e metodi per la quantificazione del rischio di interferenza idrogeologica di grandi opere in sotterraneo in differenti contesti geologici”, and has been supported by the University of Bologna (PhD project 38th cycle). The project is funded by a research agreement between ITALFERR S.p.A. and BiGeA dept. (Contratto per il conferimento do borse di dottorato attivate ai sensi del DM 352 del 9 aprile 2022–38° ciclo – A.A. 2022/2023). Funding came from the European Union–NextGenerationEU, Mission 4, Component 2, Investment 3.3 (Ministerial Decree no. 117 of 2 March 2023) CUP J33C22001540009.

Data availability The data are available in the following repository: https://doi.org/10.5281/zenodo.15390291.

The code produced is publicly vailable on GitHub at the following URL: https://github.com/NestEr96/Random_Forest_hydrogeological_interfernce For the reproducibility of the method, the parameter selection methodology is described in the various subsections of section “Materials and Methods”, with examples provided in the ESM. Using the data available in the library, it was possible to train the Random Forest model using the hyperparameters indicated in Table 10.

Declarations

Ethics approval and consent to participate Not applicable.

Consent for publication Not applicable.

Competing interests The authors declares that they have no competing interests.

Declaration of generative AI and AI-assisted technologies in the writing process During the preparation of this work, the author(s) used ChatGPT in order to improve the readability of the paper and correct grammatical errors. After using this tool/service, the author(s) reviewed and edited the content as needed and take full responsibility for the content of the published article.

References

Adinolfi Falcone R, Falgiani A, Parisse B, Petitta M, Spizzico M, Tallini M (2008) Chemical and isotopic (δ18O‰, δ2H‰, δ13C‰, 222Rn) multi-tracing for groundwater conceptual model of carbonate aquifer (Gran Sasso INFN underground laboratory, central Italy). J Hydrol 357:368–388. https://doi.org/10.1016/j.jhydrol.2008.05.016

Adinolfi Falcone R, Carucci V, Falgiani A, Manetta M, Parisse B, Petitta M, Rusi S, Spizzico M, Tallini M (2012) Changes in groundwater flow and hydrochemistry of the Gran Sasso carbonate aquifer after 2009 L’ Aquila earthquake. Ital J Geosci 131:459–474. https://doi.org/10.3301/IJG.2011.34

Aderemi BA, Olwal TO, Ndambuki JM, Rwanga SS (2023) Groundwater levels forecasting using machine learning models: a case study of the groundwater region 10 at Karst Belt, South Africa. Syst Soft Comput 5:200049. https://doi.org/10.1016/j.sasc.2023.200049

Ackermann MR, Blömer J, Kuntze D, Sohler C (2014) Analysis of agglomerative clustering. Algorithmica 69:184–215. https://doi.org/10.1007/s00453-012-9717-4

Afshar A, Mariño MA, Ebtehaj M, Moosavi J (2007) Rule-based fuzzy system for assessing groundwater vulnerability. J Environ Eng 9372:532–540. https://doi.org/10.1061/(ASCE)0733-9372(2007)133:5(532)

Alfaro C, Wallace M (1994) Origin and classification of springs and historical review with current applications. Environ Geol 24:112–124. https://doi.org/10.1007/BF00767884

Allocca V, Manna F, De Vita P (2014) Estimating annual groundwater recharge coefficient for karst aquifers of the Southern Apennines (Italy). Hydrol Earth Syst Sci 18(2):803–817. https://doi.org/10.5194/hess-18-803-2014

Amiri V, Asgari-Nejad M (2022) Hydrogeological assessment and estimation of groundwater inflow into the water transmission tunnel to Urmia Lake, northwestern Iran. Bull Eng Geol Environ 81(3):111. https://doi.org/10.1007/s10064-022-02612-3

Amoruso A, Crescentini L, Petitta M, Rusi S, Tallini M (2011) Impact of the 6 April 2009 L’Aquila earthquake on groundwater flow in the Gran Sasso carbonate aquifer, central Italy. Hydrol Process 25:1754–1764. https://doi.org/10.1002/hyp.7933

Amoruso A, Crescentini L, Petitta M, Tallini M (2013) Parsimonious recharge/discharge modeling in carbonate fractured aquifers: the groundwater flow in the Gran Sasso aquifer (central Italy). J Hydrol 476:136–146. https://doi.org/10.1016/j.jhydrol.2012.10.026

Amoruso L, Crescentini S, Martino M, Petitta TM, Tallini M (2014) Correlation between groundwater flow and deformation in the fractured carbonate Gran Sasso aquifer (INFN underground laboratories, central Italy). Water Resour Res 50:4858–4876. https://doi.org/10.1002/2013WR014491

Antolini P, Cremonini G (1991) Litostratigrafia e tettonica della Formazione Marnoso-Arenacea romagnola fra le valli del Lamone e del Savio (Appennino Settentrionale) [Lithostratigraphy and Tectonics of the Romagna Marnoso-Arenacea Formation between the Lamone and Savio Valleys (Northern Apennines)]. Mem Descr Carta Geol D'Ital 46:447–458

Argnani A, Ricci Lucchi F (2001) Tertiary silicoclastic turbidite systems of the Northern Apennines. Anatomy of an orogen, the Apennines and adjacent Mediterranean basin. Kluwer, Dordrecht, The Netherlands, pp 327–350. https://doi.org/10.1007/978-94-015-9829-3_19

Attanayake PM, Waterman MK (2006) Identifying environmental impacts of underground construction. Hydrogeol J 14:1160–1170. https://doi.org/10.1007/s10040-006-0037-0

Ayodele TO (2010) Types of machine learning algorithms. New Adv Mach Learn 3:19–48. https://doi.org/10.5772/9385

Bandis SC, Lumsden AC, Barton NR (1983) Fundamentals of rock joint deformation. Int J Rock Mech Min Sci Geomech Abstr 20(6):249–268. https://doi.org/10.1016/0148-9062(83)90595-8

Barbieri M, Boschetti T, Petitta M, Tallini M (2005) Stable isotope (2H, 18O and 87Sr/86Sr) and hydrochemistry monitoring for groundwater hydrodynamics analysis in a karst aquifer (Gran Sasso, central Italy). Appl Geochem 20(11):2063–2081. https://doi.org/10.1016/j.apgeochem.2005.07.008

Barbieri M, Barberio MD, Banzato F, Billi A, Boschetti T, Franchini S, Gori F, Petitta M (2023) Climate change and its effect on groundwater quality. Environ Geochem Health 45(4):1133–1144. https://doi.org/10.1007/s10653-021-01140-5

Barton N, Bandis SC, Bakhtar K (1985) Strength, deformation and conductivity coupling of rock joints. Int J Rock Mech Min Sci Geomech Abstr 22(3):121–140. https://doi.org/10.1016/0148-9062(85)93227-9

Barzegar R, Fijani E, Asghari Moghaddam A, Tziritis E (2017) Forecasting of groundwater level fluctuations using ensemble hybrid multi-wavelet neural network-based models. Sci Total Environ. https://doi.org/10.1016/j.scitotenv.2017.04.189

Bendkik A, Boccaletti M, Bonini M, Poccianti C, Sani F (1994) Structural evolution of the outer Apennine chain (Firenzuola-Città di Castello sector and Montefeltro area, Tuscan-Romagnan and Umbro-Marchean Apennine). Mem Soc Geol Ital 48:515–522

Benini A, Farabegoli E, Martelli L, Severi P (1991) Stratigrafia e paleogeografia del Gruppo di S. Sofia. Mem Descr Carta Geol D'ital 46:231–244

Boccaletti M, Gianelli G, Sani F (1997) Tectonic regime, granite emplacement and crustal structure in the inner zone of the Northern Apennines (Tuscany, Italy): a new hypothesis. Tectonophysics 1951(96):127–143. https://doi.org/10.1016/S0040-1951(96)00177-1

Boccaletti M, Corti G, Martelli L (2011) Recent and active tectonics of the external zone of the Northern Apennines (Italy). Int J Earth Sci 100(6):1331–1348. https://doi.org/10.1007/s00531-010-0545-y

Boni C, Bono P, Capelli G (1986) Schema idrogeologico dell'Italia Centrale [Hydrogeological scheme of central Italy]. Mem Soc Geol It 35(2):991–1012

Breiman L (2001) Random forests. Mach Learn 45:5–32. https://doi.org/10.1023/A:1010933404324

Breiman L, Friedman J, Stone CJ, Olshen RA (1984) Classification and regression trees. Chapman and Hall/CRC. https://doi.org/10.1201/9781315139470

Bruni P, Cipriani N, Pandeli E (1994) New sedimentological and petrographical data on the Oligocene-Miocene turbiditic formations of the Tuscan Domain. Mem Soc Geol Ital 48:251–260

Butscher C (2012) Steady-state groundwater inflow into a circular tunnel. Tunn Undergr Space Technol 32:158–167. https://doi.org/10.1016/j.tust.2012.06.007

Canuti P, Ermini L, Gargini A, Martelli L, Piccinini L, Vincenzi V (2009) Le gallerie TAV attraverso l'Appennino toscano: impatto idrogeologico ed opere di mitigazione [High-speed rail tunnels across the tuscan apennines: hydrogeological impact and mitigation measures]. SFERA. https://sfera.unife.it/handle/11392/1382315

Carminati E, Doglioni C (2012) Alps vs. Apennines: the paradigm of a tectonically asymmetric Earth. Earth-Sci Rev 112(1–2):67–96. https://doi.org/10.1016/j.earscirev.2012.02.004

Casati T, Navarra A, Filippini M, Gargini A (2024) Assessing the long-term trend of spring discharge in a climate change hotspot area. Sci Total Environ 957:177498. https://doi.org/10.1016/j.scitotenv.2024.177498

Castellarin A (2001) Alps-Apennines and Po plain-frontal Apennines relations. Anatomy of an orogen: the Apennines and adjacent Mediterranean basins. Springer, Dordrecht, The Netherlands, pp 177–195. https://doi.org/10.1007/978-94-015-9829-3_13

Cavinato GP, Celles PD (1999) Extensional basins in the tectonically bimodal central Apennines fold-thrust belt, Italy: response to corner flow above a subducting slab in retrograde motion. Geology 27(10):955–958. https://doi.org/10.1130/0091-7613(1999)027%3c0955:EBITTB%3e2.3.CO;2

Cavinato GP, Cosentino D, De Rita D, Funiciello R, Parotto M (1994) Tectonic-sedimentary evolution of intrapenninic basins and correlation with the volcano-tectonic activity in central Italy. Mem Descrit Carta Geol D'Italia 49:63–76

Celico P, Fabbrocino S, Petitta M, Tallini M (2005) Hydrogeological impact of the Gran Sasso motor-way tunnels (central Italy). G Geol Appl 1:157–165. https://doi.org/10.1474/GGA.2005-01.0-16.0016

Cesano D, Olofsson B (2000) Parameters regulating groundwater inflows into hard rock tunnels-a statistical study of the Dolmen tunnel in southern Sweden. Tunn Undergr Space Technol 7798(00):153–165. https://doi.org/10.1016/S0886-7798(00)00043-2

Chang FJ, Huang CW, Chang LC, Kao IF (2016) Prediction of monthly regional groundwater levels through hybrid soft-computing techniques. J Hydrol 541:965–976. https://doi.org/10.1016/j.jhydrol.2016.08.006

Chiocchini U, Castaldi F (2011) The impact of groundwater on the excavation of tunnels in two different hydrogeological settings in central Italy. Hydrogeol J 19(3):651. https://doi.org/10.1007/s10040-010-0702-1

Chiu YC, Chia Y (2012) The impact of groundwater discharge to the Hsueh-Shan tunnel on the water resources in northern Taiwan. Hydrogeol J 20(8):1599–1611. https://doi.org/10.1007/s10040-012-0895-6

D'Agostino N, Tozzi M (1995) Rotazione di blocchi su assi orizzontali al margine meridionale del massiccio del Gran Sasso [Rotation of blocks on horizontal axes at the southern edge of the Gran Sasso massif]. Stud Geol Camerti 2:183–189

Dal Seno ND, Evangelista E, Piccolomini M, Berti M (2024) Comparative analysis of conventional and machine learning techniques for rainfall threshold evaluation under complex geological conditions. Landslides 21(12):2893–2911. https://doi.org/10.1007/s10346-024-02336-3

Dahal K, Sharma S, Shakya A, Talchabhadel R, Adhikari S, Pokharel A, Sheng Z, Pradhan AMS, Kumar S (2023) Identification of groundwater potential zones in data-scarce mountainous region using explainable machine learning. J Hydrol 627:130417. https://doi.org/10.1016/j.jhydrol.2023.130417

Dhapre M, Jadhav S, Das D, Khan J, Kim Y, Chiao S, Danielson T (2025) A systematic review of machine learning in groundwater monitoring. Environ Model Softw. https://doi.org/10.1016/j.envsoft.2025.106549

Dematteis A, Kalamaras G, Eusebio A (2001) A systems approach for evaluating springs drawdown due to tunneling. AITES/ITA World Tunnel Congress 2001, Milan, 9–12 June 2001

Dematteis A, Pugliese E, Gargini A (2025) Evaluating the impact of tunnelling on surface water resources: update of the Drawdown Hazard Index (DHI) method. Q J Eng Geol Hydrogeol. https://doi.org/10.1144/qjegh2025-041

Domenico PA, Schwartz FW (1997) Physical and chemical hydrogeology. Wiley, Chichester, England

El Tani M (1999) Water inflow into tunnels. In: Proceedings of the World Tunnel Congress ITA-AITES (pp. 61–70). International Tunneling and Underground Space Association, Châtelaine, Switzerland

El Tani M (2003) Circular tunnel in a semi-infinite aquifer. Tunn Undergr Space Technol 18(1):49–55. https://doi.org/10.1016/S0886-7798(02)00102-5

Falcone RA, Falgiani A, Parisse B, Petitta M, Spizzico M, Tallini M (2008) Chemical and isotopic (δ^{18}O‰, δ^{2}H‰, δ^{13}C‰, ^{222}Rn) multi-tracing for groundwater conceptual model of carbonate aquifer (Gran Sasso INFN underground laboratory, central Italy). J Hydrol 357(3):368–388. https://doi.org/10.1016/j.jhydrol.2008.05.016

Farhadian H, Katibeh H, Huggenberger P (2016) Empirical model for estimating groundwater flow into tunnels in discontinuous rock masses. Environ Earth Sci 75(6):471. https://doi.org/10.1007/s12665-016-5332-z

Fetter CW, Kreamer D (2021) Applied hydrogeology. Waveland, Long Grove,IL

Filipponi M (2014) Karst risk assessment for underground engineering: comparison of the KarstALEA method with a random karst distribution approach. Engineering geology for society and territory-volume 5: urban geology, sustainable planning and landscape exploitation. Springer, Cham, Switzerland, pp 603–607. https://doi.org/10.1007/978-3-319-09048-1_117

Filipponi M, Jeannin PY (2010) Karst-ALEA: a scientific based karst risk assessment for underground engineering. Advances in Research in Karst Media. Springer, Heidelberg, Germany, pp 435–440. https://doi.org/10.1007/978-3-642-12486-0_67

Filippini M, Segadelli S, Dinelli E, Failoni M, Stumpp C, Vignaroli G, Casati T, Tiboni B, Gargini A (2024) Hydrogeological assessment of a major spring discharging from a calcarenitic aquifer with implications on resilience to climate change. Sci Total Environ 913:169770, 169770. https://doi.org/10.1016/j.scitotenv.2023.169770

Fiorillo F, Guadagno FM (2012) Long karst spring discharge time series and drought occurrence in southern Italy. Environ Earth Sci 65(8):2273–2283. https://doi.org/10.1007/s12665-011-1495-9

Fiorillo F, Petitta M, Preziosi E, Rusi S, Esposito L, Tallini M (2015) Long-term trend and fluctuations of karst spring discharge in a Mediterranean area (central-southern Italy). Environ Earth Sci 74(1):153–172. https://doi.org/10.1007/s12665-014-3946-6

Ford DC, Williams PW (1989) Karst geomorphology and hydrology, vol 601. Hyman, London

Freeze RA, Cherry JA (1979) Groundwater. Prentice-Hall, Englewood Cliffs, NJ, pp 161–177

Frenelus W, Peng H, Zhang J (2021) Evaluation methods for groundwater inflows into rock tunnels: a state-of-the-art review. Int J Hydrol 5(4):152–168

Gargini A, Piccinini L, Martelli L, Rosselli S, Bencini A, Messina A, Canuti P (2006) Hydrogeology of turbidites: a conceptual model derived by the geological survey of Tuscan-Emilian Apennines and the environmental monitoring for the high speed railway tunnel connection between Florence and Bologna. Boll Soc Geol Ital 125(3):293–327

Gargini A, Vincenzi V, Piccinini L, Zuppi GM, Canuti P (2008) Groundwater flow systems in turbidites of the Northern Apennines (Italy): natural discharge and high speed railway tunnel drainage. Hydrogeol J 16(8):1577. https://doi.org/10.1007/s10040-008-0352-8

Gattinoni P, Scesi L (2010) An empirical equation for tunnel inflow assessment: application to sedimentary rock masses. Hydrogeol J 18(8):1797–1810. https://doi.org/10.1007/s10040-010-0674-1

Golian M, Teshnizi ES, Nakhaei M (2018) Prediction of water inflow to mechanized tunnels during tunnel-boring-machine advance using numerical simulation. Hydrogeol J 26(8):2827–2851. https://doi.org/10.1007/s10040-018-1835-x

Goodfellow I, Bengio Y, Courville A, Bengio Y (2016) Deep learning, vol 1, no. 2). MIT, Cambridge, MA

Goodman R, Moye D, Schalkwyk A, Javendel I (1965) Ground-water inflow during tunnel driving. Eng Geol Bull IAEG 2:39–56

Gonzalez RQ, Arsanjani JK (2021) Prediction of groundwater level variations in a changing climate: a Danish case study. ISPRS Int J Geo-Inf 10(11):792. https://doi.org/10.3390/ijgi10110792

Gorgij AD, Moghaddam AA, Kisi O (2017) Groundwater budget forecasting, using hybrid wavelet-ANN-GP modelling: a case study of Azarshahr plain, East Azerbaijan, Iran. Hydrol Res 48:455–467. https://doi.org/10.2166/nh.2016.202

Heidari M, Sharafi M, Khazaei S (2016) Study of morphology fractures in prediction of high local groundwater flow into tunnels using ASTER satellite images. J Indian Soc Remote Sens 44(2):253–268. https://doi.org/10.1007/s12524-015-0483-z

Hosna A, Merry E, Gyalmo J, Alom Z, Aung Z, Azim MA (2022) Transfer learning: a friendly introduction. J Big Data 9(1):102. https://doi.org/10.1186/s40537-022-00652-w

Hudson JA (1992) Rock engineering systems: theory and practice. Horwood, Chichester, England

Jiao Y (1995) Formalising the systems approach to rock engineering. PhD Thesis, Imperial College London, London

Karlsrud K (2001) Water control when tunneling under areas in the Oslo region. NFF Publ 12:4–27

Katibeh H, Aalianvari A (2009) Development of a new method for tunnel site rating from. J Appl Sci 9(8):1496–1502

Kværner J, Snilsberg P (2008) The Romeriksporten railway tunnel: drainage effects on peatlands in the lake Northern Puttjern area. Eng Geol 101(3–4):75–88. https://doi.org/10.1016/j.enggeo.2008.04.002

Kværner J, Snilsberg P (2011) Groundwater hydrology of boreal peatlands above a bedrock tunnel: drainage impacts and surface water groundwater interactions. J Hydrol 403(3–4):278–291. https://doi.org/10.1016/j.jhydrol.2011.04.006

LaBianca A, Koch J, Jensen KH, Sonnenborg TO, Kidmose J (2024) Machine learning for predicting shallow groundwater levels in urban areas. J Hydrol 632:130902. https://doi.org/10.1016/j.jhydrol.2024.130902

Lauzon D, Gloaguen E (2024) Quantifying uncertainty and improving prospectivity mapping in mineral belts using transfer learning and Random Forest: a case study of copper mineralization in the Superior Craton Province, Quebec, Canada. Ore Geol Rev. https://doi.org/10.1016/j.oregeorev.2024.105918

Leone G, Pagnozzi M, Catani V, Ventafridda G, Esposito L, Fiorillo F (2021) A hundred years of Sanit'a Spring discharge measurements: trends and statistics for understanding water resource availability under climate change. Stoch Env Res Risk A 35(2):345–370. https://doi.org/10.1007/s00477-020-01908-8

Li L, Tu W, Shi S, Chen J, Zhang Y (2016) Mechanism of water inrush in tunnel construction in karst areas. Geom Nat Hazards Risk 7:35–46. https://doi.org/10.1080/19475705.2016.1181342

Li SC, He P, Li LP, Shi SS, Zhang QQ, Zhang J, Hu J (2017) Gaussian process model of water inflow prediction in tunnel construction and its engineering applications. Tunn Undergr Space Technol 69:155–161. https://doi.org/10.1016/j.tust.2017.06.018

Liu J, Liu D, Song K (2015) Evaluation of the influence caused by tunnel construction on groundwater environment: a case study of Tongluoshan tunnel, China. Adv Mater Sci Eng 2015(1):149265. https://doi.org/10.1155/2015/149265

Liu JH, Zhang WB, Li XJ, Liu BL, Xu B, Xi Y (2021) A numerical method for predicting groundwater inflow into a mountain tunnel, taking into account rainfall recharge and lithological distribution over a large area. In: IOP Conference Series. Earth and Environmental Science. https://doi.org/10.1088/1755-1315/861/7/072040

Lorè A, Magaldi D, Tallini M (2002) Morphology and morphometry of the Gran Sasso (central Italy) surface karst. Geogr Fis Din Quat 25(2):123–134

Lunardi P, Catalano P (2006) Gran Sasso: il traforo autostradale [Gran Sasso: the motorway tunnel]. Gangeni, Rome

Lundberg SM, Lee SI (2017) A unified approach to interpreting model predictions. In: Advances in neural information processing systems, NIPS, Long Beach, CA, December 2017, 30 pp

Lv Y, Jiang Y, Hu W, Cao M, Mao Y (2020) A review of the effects of tunnel excavation on the hydrology, ecology, and environment in karst areas: current status, challenges, and perspectives. J Hydrol 586:124891. https://doi.org/10.1016/j.jhydrol.2020.124891

Mahmoodzadeh A, Mohammadi M, Noori KMG, Khishe M, Ibrahim HH, Ali HFH, Abdulhamid SN (2021) Presenting the best prediction model of water inflow into drill and blast tunnels among several machine learning techniques. Autom Constr 127:103719. https://doi.org/10.1016/j.autcon.2021.103719

Maillet E (1905) Essai d'hydraulique souterraine et fluviale [Essay of groundwater and river hydraulics]. Lib. Scient. Herman, New York City

Molnar C, König G, Herbinger J, Freiesleben T, Dandl S, Scholbeck CA, et al. (2020) General pitfalls of model-agnostic interpretation methods for machine learning models. In: International Workshop on Extending Explainable AI Beyond Deep Models and Classifiers. Springer, Cham, Switzerland, pp 39–68

Mondini AC, Guzzetti F, Melillo M (2023) Deep learning forecast of rainfall-induced shallow landslides. Nat Commun 14(1). https://doi.org/10.1038/s41467-023-38135-y

Monjoie A (1980) Prévision et contrôle des caractéristiques hydrogéologiques dans les tunnels du Gran Sasso (Apennin-Italie) [Forecasting and monitoring hydrogeological characteristics in the tunnels of Gran Sasso (Apennines, Italy)].

Molinero J, Samper J, Juanes R (2002) Numerical modeling of the transient hydrogeological response produced by tunnel construction in fractured bedrocks. Eng Geol 64(4):369–386. https://doi.org/10.1016/S0013-7952(01)00099-0

Mossmark F, Ericsson LO, Norin M, Dahlström LO (2015) Hydrochemical changes caused by underground constructions: a case study of the Kattleberg rail tunnel. Eng Geol 191:86–98. https://doi.org/10.1016/j.enggeo.2015.03.004

Nourani V, Ghaffari A, Behfar N, Foroumandi E, Zeinali A, Ke CQ, Sankaran A (2024) Spatiotemporal assessment of groundwater quality and quantity using geostatistical and ensemble artificial intelligence tools. J Environ Manage 355:120495. https://doi.org/10.1007/s10040-023-02686-7

Orland E, Roering JJ, Thomas MA, Mirus BB (2020) Deep learning as a tool to forecast hydrologic response for landslide-prone hillslopes. Geophys Res Lett 47(16). https://doi.org/10.1029/2020GL088731

Perrochet P, Dematteis A (2007) Modeling transient discharge into a tunnel drilled in a heterogeneous formation. Groundwater 45(6):786–790. https://doi.org/10.1111/j.1745-6584.2007.00355.x

Petaccia R, Rusi S (2008) Idrogeologia delle sorgenti del Ruzzo (Gran Sasso d'Italia) [Hydrogeology of the Ruzzo springs (Gran Sasso d'Italia)]. G Geol Appl 8(1):17–28

Petitta M, Tallini M (2002) Idrodinamica sotterranea del massiccio del Gran Sasso (Abruzzo): muove indagini idrologiche, idrogeologiche e idrochimiche (1994–2001) [Underground hydrodynamics of the Gran Sasso massif (Abruzzo): hydrological, hydrogeological and hydrochemical investigations (1994–2001)]. Boll Soc Geol Ital 121(3):343–363

Petitta M, Scarascia Mugnozza G, Barbieri M, Bianchi Fasani G, Esposito C (2010) Hydrodynamic and isotopic investigations for evaluating the mechanisms and amount of groundwater seepage through a rockslide dam. Hydrol Process 24(24):3510–3520. https://doi.org/10.1002/hyp.7773

Petitta M, Caschetto M, Galassi DM, Aravena R (2015) Dual-flow in karst aquifers toward a steady discharge spring (Presciano, Central Italy): influences on a subsurface groundwater dependent ecosystem and on changes related to post-earthquake hydrodynamics. Environ Earth Sci 73:2609–2625. https://doi.org/10.1007/s12665-014-3440-1

Petitta M, Banzato F, Lorenzi V, Matani E, Sbarbati C (2022) Determining recharge distribution in fractured carbonate aquifers in central Italy using environmental isotopes: snowpack cover as an indicator for future availability of groundwater resources. Hydrogeol J 30(5):1619–1636. https://doi.org/10.1007/s10040-022-02501-9

Piccinini L, Vincenzi V (2010) Impacts of a railway tunnel on the streams baseflow verified by means of numerical modelling. Acque Sotterranee - Ital J Groundwater 10(1). https://doi.org/10.7343/as-2021-503

Polemio M, Casarano D (2008) Climate change, drought and groundwater availability in southern Italy. Geol Soc Spec Publ 288:39–51. https://doi.org/10.1144/SP288.4

Pradhan S, Kumar S, Kumar Y, Sharma HC (2019) Assessment of groundwater utilization status and prediction of water table depth using different heuristic models in an Indian interbasin. Soft Comput 23:10261–10285. https://doi.org/10.1007/s00500-018-3580-4

Preisig G, Dematteis A, Torri R, Monin N, Milnes E, Perrochet P (2014) Modelling discharge rates and ground settlement induced by tunnel excavation. Rock Mech Rock Eng 47:869–884. https://doi.org/10.1007/s00603-012-0357-4

Rajaee T, Ebrahimi H, Nourani V (2019) A review of the artificial intelligence methods in groundwater level modeling. J Hydrol 572:336–351. https://doi.org/10.1016/j.jhydrol.2018.12.037

Refsgaard JC, Stisen S, Koch J (2022) Hydrological process knowledge in catchment modelling – lessons and perspectives from 60 years development. Hydrol Process 36(1). https://doi.org/10.1002/hyp.v36.110.1002/hyp.14463

Ricci Lucchi F (1978) Turbidite dispersal in a Miocene deep-sea plain. Geol Mijnb 57:559e–5576

Ricci Lucchi F (1981) The Marnoso arenacea turbidites, Romagna and Umbria Apennines. In: Ricci Lucchi F (ed), Excursion guidebook, with contribution on sedimentology of some Italian Basins, 2nd IAS Eur. Meeting, Bologna, Italy, pp 229–303

Ricci Lucchi F (1986) The Oligocene to Recent foreland basins of the northern Apennines. Foreland Basins, pp 103–139. https://doi.org/10.1002/9781444303810.ch6

Rodriguez-Galiano VF, Chica-Olmo M, Chica-Rivas M (2014) Predictive modelling of gold potential with the integration of multisource information based on random forest: a case study on the Rodalquilar area, southern Spain. Int J Geogr Inf Sci

28(7):1336–1354. https://doi.org/10.1080/13658816.2014.885527

Sahoo S, Russo T, Elliott J, Foster I (2017) Machine learning algorithms for modeling groundwater level changes in agricultural regions of the US. Water Resour Res 53:3878–3895. https://doi.org/10.1002/2016WR019933

Sahoo S, Jha MK (2013) Groundwater-level prediction using multiple linear regression and artificial neural network techniques: a comparative assessment Prédiction du niveau de la nappe par les techniques de régression linéaire multiple et de kréseau neuronal artificiel : évaluation comparative Predicción de niveles de agua subterránea usando regresión linear múltiple y técnicas de redes neuronales artificiales: una evaluación comparativa 利用多元线性回归和人工神经网络技术来预测地下水位 Predição dos níveis de água subterrânea usando técnicas de regressão linear múltipla e redes neuronais artificiais: uma avaliação comparativa. Hydrogeol J 21(8):1865–1887. https://doi.org/10.1007/s10040-013-1029-5

Sajedi-Hosseini F, Malekian A, Choubin B, Rahmati O, Cipullo S, Coulon F, Pradhan B (2018) A novel machine learning-based approach for the risk assessment of nitrate groundwater contamination. Sci Total Environ 644:954–962. https://doi.org/10.1016/j.scitotenv.2018.07.054

Sappa G, Ferranti F, Iacurto S, De Filippi FM (2018) Effects of climate change on groundwater feeding the Mazzoccolo and Capodacqua di Spigno Springs (central Italy): first quantitative assessments. International Multidisciplinary Scientific GeoConference: SGEM, 18(3.1), pp 219–226. https://doi.org/10.5593/sgem2018/3.1/S12.029

Sappa G, De Filippi FM, Ferranti F, Iacurto S (2019) Environmental issues and anthropic pressures in coastal aquifers: a case study in Southern Latium Region. Acque Sotterranee-Ital J Groundw 8(1). https://doi.org/10.7343/as-2019-373

Scesi L, Gattinoni P (2007) Roughness control on hydraulic conductivity in fractured rocks. Hydrogeol J 15:201–211. https://doi.org/10.1007/s10040-006-0076-6

Schmidhuber J (2014) Deep learning in neural networks: an overview. Neural Netw 61:85–117. https://doi.org/10.1016/j.neunet.2014.09.003

Schwarz L, Reichl I, Kirschner H, et al (2006) Risks and hazards caused by groundwater during tunnelling: geotechnical solutions used as demonstrated by recent examples from Tyrol, Austria. Environ Geol 49:858–864. https://doi.org/10.1007/s00254-006-0181-9

Segadelli S, Filippini M, Monti A, Celico F, Gargini A (2021) Estimation of recharge in mountain hard-rock aquifers based on discrete spring discharge monitoring during base-flow recession. Hydrogeol J 29(3):949–961. https://doi.org/10.1007/s10040-021-02317-z

Servizio Geologico d'Italia - ISPRA (2009) Carta Geologica d'Italia alla scala 1:500.000 (WMS) [Geological map of Italy at scale]. Geoportale Nazionale, MASE, Rome

Slejko D (1993) Considerations on the seismic risk. Ann Geophys. https://doi.org/10.4401/ag-4312

Sjölander-Lindqvist A (2005) Conflicting perspectives on water in a Swedish railway tunnel project. Environ Values 14(2):221–239. https://doi.org/10.3197/09632710540849

Snow DT (1969) Anisotropic hydraulic conductivity of fractured media. Water Resour Res 5(6):1273–1289

Snow DT (1970) The frequency and apertures of fractures in rock. Int J Rock Mech Min Sci Geomech Abstr 7(1):23–40. https://doi.org/10.1016/0148-9062(70)90025-2

Surinaidu L, Rao VG, Rao NS, Srinu S (2014) Hydrogeological and groundwater modeling studies to estimate the groundwater inflows into the coal mines at different mine development stages using MODFLOW, Andhra Pradesh, India. Water Resour Ind 7:49–65. https://doi.org/10.1016/j.wri.2014.10.002

Tallini M, Ranalli D, Petitta M, Tersigni S (2000) Karst aquifer characterization using physico-chemical spring data, Gran Sasso, central Italy. In: Groundwater: past achievements and future challenges. IAH Congress on Groundwater, Balkema, The Netherlands, pp 629–634

Tallini M, Petitta M, Ranalli D (2001) Caratterizzazione chimico-fisica e idrologica delle acque sotterranee del Gran Sasso: monitoraggio biologico del Gran Sasso [Chemical-physical and hydrological characterization of the Gran Sasso groundwater: biological monitoring of the Gran Sasso]. 2:225–276

Tallini M, Adinolfi Falcone R, Carucci V, Falgiani A, Parisse B, Petitta M (2014) Isotope hydrology and geochemical modeling: new insights into the recharge processes and water–rock interactions of a fissured carbonate aquifer (Gran Sasso, central Italy). Environ Earth Sci 72:4957–4971. https://doi.org/10.1007/s12665-014-3364-9

Tapoglou E, Trichakis IC, Dokou Z, Nikolos IK, Karatzas GP (2014) Groundwater-level forecasting under climate change scenarios using an artificial neural network trained with particle swarm optimization. Hydrol Sci J/J Sci Hydrol 59:1225–1239. https://doi.org/10.1080/02626667.2013.838005

Tinterri R, Tagliaferri A (2015) The syntectonic evolution of foredeep turbidites related to basin segmentation: facies response to the increase in tectonic confinement (Marnoso-Arenacea Formation, Miocene, Northern Apennines, Italy). Mar Pet Geol 67:81–110. https://doi.org/10.1016/j.marpetgeo.2015.04.006

Titti G, Sarretta A, Lombardo L, Crema S, Pasuto A, Borgatti L (2022) Mapping susceptibility with open-source tools: a new plugin for QGIS. Front Earth Sci 10:842425. https://doi.org/10.3389/feart.2022.842425

Torri R, De Matteis A, Delle Piane L (2007) Drawdown hazard of springs and wells in tunneling: predictive model and verification. In: Proceedings of XXXV IAH Congress, Groundwater and Ecosystems, vol 10, 17–21 September 2007, Lisbon

Toth J (1963) A theoretical analysis of groundwater flow in small drainage basins. J Geophys Res 68(16):4795–4812

Vai F, Martini IP (eds) (2013) Anatomy of an orogen: the Apennines and adjacent Mediterranean basins. Springer, Heidelberg, Germany

Vezzani L, Ghisetti F (1998) Carta geologica dell'Abruzzo, a scala 1:100.000. S. EL.CA.

Vincenzi V, Gargini A, Goldscheider N (2009) Using tracer tests and hydrological observations to evaluate effects of tunnel drainage on groundwater and surface waters in the Northern Apennines (Italy). Hydrogeol J 17(1):135–150. https://doi.org/10.1007/s10040-008-0371-5

Vincenzi V, Piccinini L, Gargini A, Sapigni M (2010) Parametric and numerical modeling tools to forecast hydrogeological impacts of a tunnel. Acque Sotter - Ital J Groundw 11(1):xx–xx. https://doi.org/10.7343/as-2022-558

Vincenzi V, Gargini A, Goldscheider N, Piccinini L (2014) Differential hydrogeological effects of draining tunnels through the Northern Apennines, Italy. Rock Mech Rock Eng 47:947–965. https://doi.org/10.1007/s00603-013-0378-7

Xu H, Li X, Gokdemir C (2021) Modeling and assessing the impact of tunnel drainage on terrestrial vegetation. Tunn Undergr Space Technol 116:104097. https://doi.org/10.1016/j.tust.2021.104097

Yang FR, Lee CH, Kung WJ, Yeh HF (2009) The impact of tunneling construction on the hydrogeological environment of "Tseng-Wen Reservoir Transbasin Diversion Project" in Taiwan. Eng Geol 103(1–2):39–58. https://doi.org/10.1016/j.enggeo.2008.07.012

Zarei HR, Uromeihy A, Sharifzadeh M (2011) Evaluation of high local groundwater inflow to a rock tunnel by characterization of geological features. Tunn Undergr Space Technol 26(2):364–373. https://doi.org/10.1016/j.tust.2010.11.007

Zattin M, Landuzzi A, Picotti V, Zuffa GG (2000) Discriminating between tectonic and sedimentary burial in a foredeep succession, Northern Apennines. J Geol Soc 157(3):629–633. https://doi.org/10.1144/jgs.157.3.629

Zheng A, Casari A (2018) Feature engineering for machine learning: principles and techniques for data scientists. O'Reilly Media, Sebastopol, CA

Zuffa GG (1980) Hybrid arenites; their composition and classification. J Sediment Res 50(1):21–29. https://doi.org/10.1306/212F7950-2B24-11D7-8648000102C1865D

Zweig MH, Campbell G (1993) Receiver-operating characteristic (ROC) plots: a fundamental evaluation tool in clinical medicine. Clin Chem 39(4):561–577. https://doi.org/10.1093/clinchem/39.4.561