



ELSEVIER

Contents lists available at ScienceDirect

MethodsX

journal homepage: [www.elsevier.com/locate/mex](http://www.elsevier.com/locate/mex)

## Method Article

# A framework model to determine groundwater contamination risk based on a L-Matrix of aquifer vulnerability and hazardous activity indices



Fernando António Leal Pacheco<sup>a,\*</sup>, Victor Hugo Sarrazin Lima<sup>b</sup>,  
Teresa Cristina Tarlé Pissarra<sup>c</sup>, Renato Farias do Valle Junior<sup>d</sup>,  
Maytê Maria Abreu Pires de Melo Silva<sup>d</sup>, Marília Carvalho de Melo<sup>e</sup>,  
Carlos Alberto Valera<sup>f</sup>, João Paulo Moura<sup>g</sup>,  
Luís Filipe Sanches Fernandes<sup>g</sup>

<sup>a</sup> CQVR - Chemistry Centre of Vila Real, University of Trás-os-Montes and Alto Douro (UTAD), Ap. 1013,5001–801 Vila Real, Portugal

<sup>b</sup> University of Trás-os-Montes and Alto Douro (UTAD), Ap. 1013,5001–801 Vila Real, Portugal

<sup>c</sup> Faculty of Agrarian and Veterinary Sciences (FCAV), State University of São Paulo (UNESP), Via Prof. Paulo Donato Castellane, s/n, Jaboticabal SP 14884-900, Brazil

<sup>d</sup> Federal Institute of Triângulo Mineiro (IFTM), Uberaba Campus, Geoprocessing Laboratory, Uberaba, MG 38064-790, Brazil

<sup>e</sup> Secretaria de Estado de Meio Ambiente e Desenvolvimento Sustentável, Cidade Administrativa do Estado de Minas Gerais, Rodovia João Paulo II, 4143, Bairro Serra Verde, Belo Horizonte, Minas Gerais, Brazil

<sup>f</sup> Coordenadoria Regional das Promotorias de Justiça do Meio Ambiente das Bacias dos Rios Paranaíba e Baixo Rio Grande, Rua Coronel Antônio Rios, 951, Uberaba, MG 38061-150, Brazil

<sup>g</sup> CITAB - Centre for the Research and Technology of Agro-Environmental and Biological Sciences, University of Trás-os-Montes and Alto Douro (UTAD), Ap. 1013,5001–801 Vila Real, Portugal

## A B S T R A C T

The method presented in this study assesses groundwater contamination risk using a L-Matrix system approach. The L-Matrix in this case is a cartesian diagram where the XX-axis represents aquifer vulnerability ( $0 \leq V \leq 1$ ) determined by the well-known DRASTIC model, and the YY-axis represents the potential hazardousness ( $0 \leq H \leq 1$ ) of an activity (infrastructural development, industrial activities, livestock and agriculture) measured by a European Commission approach. The diagram is divided into four regions, the boundaries of which are set to  $V = 0.5$  and  $H = 0.5$ . Watersheds are represented in this diagram considering their V and H indices, and assigned a potential contamination risk if groundwater sites located within their limits show contaminant concentrations above legal limits for a given use. Depending on the region the watershed falls in the L-Matrix diagram, different management or contamination prevention actions are highlighted: activity development, activity monitoring, activity planning or activity inspecting. Watersheds located in the inspecting region and simultaneously evidencing contamination risk require immediate action, namely conditioning or even suspension

\* Corresponding author.

E-mail address: [fpacheco@utad.pt](mailto:fpacheco@utad.pt) (F.A.L. Pacheco).

of use. The method is tested in the Paraopeba River basin (Minas Gerais, Brazil), a densely industrialized basin that was recently affected by an iron-ore mine tailings dam break.

- The L-Matrix diagram highlights different groundwater susceptibility realities experienced by watersheds with different combinations of aquifer vulnerability and activity hazardousness, namely possibility for potential expansion of new hazardous activities but also the necessity to periodically inspect and eventually condition or suspend others.
- The L-Matrix diagram is likely a better approach to implement contamination prevention measures in watersheds, than the integrated contamination risk index used by most methods.

© 2022 The Author(s). Published by Elsevier B.V.

This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>)

---

## ARTICLE INFO

*Method name:* A framework model to determine groundwater contamination risk based on a L-Matrix of aquifer vulnerability and hazardous activity indices

*Keywords:* Aquifer, Potentially harmful activity, Contamination likelihood

*Article history:* Available online 16 September 2022

---

### Specifications table

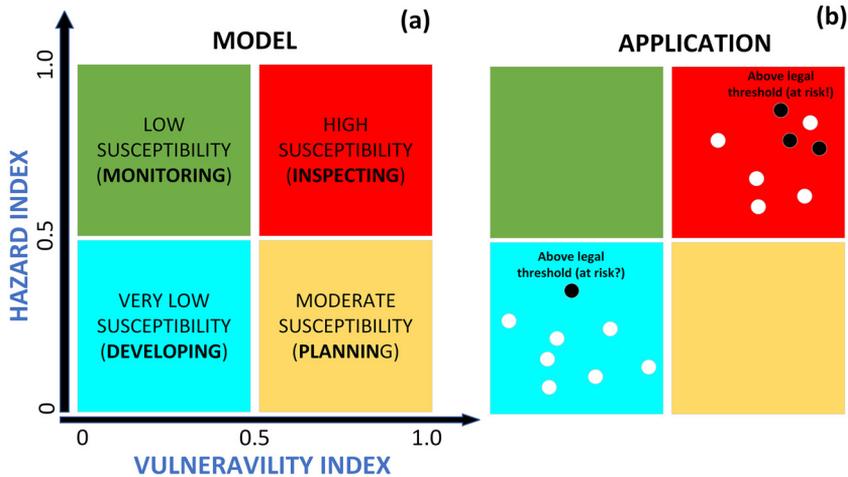
Subject area:	Environmental Science
More specific subject area:	Groundwater contamination risk
Name of your method:	A framework model to determine groundwater contamination risk based on a L-Matrix of aquifer vulnerability and hazardous activity indices
Name and reference of original method:	Not applicable
Resource availability:	Besides the Supplementary Materials provided with this submission, the authors are committed to provide additional data at the request of an interested scientist
Direct Submission or Co-Submission:	Direct Submission
	Co-submissions are papers that have been submitted alongside an original research paper accepted for publication by another Elsevier journal

---

## Method details

### Background

The assessment of groundwater contamination risk is the basis for groundwater management and contamination prevention, and there are methods capable to quantify and map the risk at the watershed scale. The general approach of most methods that were published in the last decade [1–7] is to estimate the groundwater contamination risk as product or sum of an intrinsic vulnerability index (V) and a hazard index (H) or contamination load potential. The V index is usually linked to media properties (e.g., aquifer hydraulic conductivity) and environmental settings (e.g., topography), and the H index to the distribution of likely contamination sources within the catchment as well as to the dangerousness of the source (e.g., toxicity, contaminant release likelihood and expected quantity, etc.). These methods are efficient while representing groundwater contamination risk on a single map, but may fail while effective groundwater protection tools. This is because the same risk can represent very different realities from the prevention of contamination standpoint. For example, the reality behind the product of a low vulnerability and a very high contamination potential is different from the product of a very high vulnerability and a low contamination potential, but the calculated risk is the same. In order to help improving groundwater contamination risk models as tools for contamination prevention, we developed a method that also uses V and H indices to assess groundwater contamination risk in a watershed, but where the risk is assigned by a L-Matrix system



**Fig. 1.** (a) conceptual groundwater contamination risk model. (b) the circles represent watersheds. The black color indicates contaminant concentration above a legal threshold imposed to a specific use, while the white color indicates concentration below that limit. The back circles represent risk of using the groundwater.

considering the region where the watershed falls in a *V versus H* plot. In addition, we complemented the assessment of risk with the potential consequences of a contamination event, linked to the compliance of contamination indicators with legal thresholds that determine the possibility to use the groundwater as resource. In general, the method proposed in this study follows the “pressures – pathway – receptor” framework proposed by the European Union [8]. Here, the pressures are the potential hazardous events caused by contaminant spills, the pathway is the intrinsic vulnerability of groundwater to contamination, and the receptor is groundwater as resource.

#### *Model rationale and outline*

Let's *V* be an index that ranges from 0 to 1, with the middle point ( $V = 0.5$ ) defined as the threshold that separates aquifer vulnerability from non-vulnerability. And let's apply the same reasoning to the hazard index termed *H*. The plot of *V* and *H* on a L-Matrix diagram would resemble Fig. 1a. The red square represents the “high-high” region where both vulnerability and hazards are high, the reason why this is a high-susceptibility region concerning groundwater contamination. The light blue square represents the region where *V* and *H* are both low and hence is a very low-susceptibility region. Finally, the green and orange squares represent “low-high” and “high low” regions, respectively, meaning that they are hybrid areas. The green square accommodates the watersheds protected by low vulnerability but threatened by a high density of potentially harmful anthropic activities, whereas the orange square has a low occupation by these activities but is vulnerable to any accident or other related circumstance that exposes groundwater to contamination.

From a contamination prevention perspective, the four squares represent different realities on the field, which require distinct management approaches. Considering the condition of simultaneous vulnerability and hazard potential, the activities developed in the watersheds with *V* and *H* values falling in the red square are sought to be periodically and rigorously inspected for verification of their compliance with best practices, under the penalty of highly probable groundwater exposure to contamination in the case of a spill. These watersheds must also be monitored for specific indicators of contamination (e.g., nitrate for agriculture) in springs, dug wells or drilled wells with short periodicity. The watersheds falling in the green square are characterized by a dense occupation with potentially harmful activities (high *H*), but their location in the non-vulnerability region (low *V*)

anticipates low exposure of groundwater to contamination. In this case, periodical inspection of each activity may not be necessary. However, for precaution, springs, dug wells and drilled wells located within the watershed should be monitored for contamination, eventually with longer periodicity. The light blue and orange squares accommodate watersheds with low hazard potential, but they are vulnerable if plotted in the orange square and not vulnerable otherwise. In both cases, the low density of potentially harmful activities (low H) means that there is space for installing new ones. In the light blue case, this can be encouraged because the watersheds are non-vulnerable. If that occurs, these watersheds will move towards and eventually fall the green square after some time. In the orange case, it is necessary to verify if there are zones within the watersheds where  $V < 0.5$ . If that is the case, a watershed management plan could be implemented to allow increasing the value of H in those zones. Thus, in general, the management of susceptibility in the red, orange, green and light blue squares must be approached through inspection, planning, monitoring and developing strategies, respectively, to be implemented in the corresponding watersheds by the competent authorities. These management approaches are captured by the present method but not by most contamination risk methods.

The colored squares describe levels of groundwater susceptibility to contamination. When real watersheds are plotted in the L-Matrix diagram as function of their V and H indices, and represented with black or white color depending on whether a contamination indicator is above or below the legal threshold for a certain use, then the groundwater contamination risk can be assessed (Fig. 1b). The risk exists if the concentration is above the legal threshold (black circles), because the exposure to the contaminant will likely condition or even suspend the use of groundwater as resource, before mitigation or corrective measures are implemented to reduce the risk. Otherwise, irrespective of susceptibility, there is no risk because concentrations are conforming the legislation. Fig. 1b represents a number of hypothetical watersheds, most of them with white color but two of them with black color. One black circle plots in the red and the other in the light blue square. Technically, groundwater in both watersheds is at risk, but the interpretation of each case from a management or contamination prevention standpoint is different. The black circle plotted in the red square is expected because the susceptibility is high in this region. Thus, a contaminant spill will likely put groundwater at risk. In this case, the short-term action must be to identify the hazardous events responsible for the spill and remove the cause immediately. The light blue case is striking. Considering the protection provided by the low vulnerability, coupled with the low hazard potential, black circles are unexpected in this region of the L-Matrix diagram. When they occur, they probably mean circumstantial or localized conditions not detected in the appraisal of vulnerability and hence the short-term action should be of reassessment, namely: (a) verify if time dependent factors (e.g., depth of water table) varied substantially within the evaluation timeframe; (b) verify if local geological settings (e.g., preferential flow paths such as wide-open fractures) are capable to justify the contamination event; among others. In general, Fig. 1b allows getting a panoramic view over the susceptibility of watersheds from a studied region and if groundwater flowing across them have been exposed to contaminants endangering its use as resource. The method is applicable to any watershed even if groundwater is not being used as resource, but is more directed to those watersheds where groundwater uses are typified. Only in the later cases the method is fully useful as management and groundwater protection tool.

The method proposed in this study assesses vulnerability using the well-known DRASTIC model [9] and the hazard potential using the method proposed by the European Commission (EC) and described in its COST Action 620 report [10]. The description of both methods will be densified in the next section. In the application of Fig. 1 to the visualization of susceptibility, contrarily to the EC approach, the present study discriminates among the general type of hazards, namely those related with infrastructural development, industrial activities, and livestock and agriculture, meaning that in the end the method produces three L-Matrix diagrams, one for each hazard type, and not just one diagram representing the influence of all hazard types at once. The reason for this option is that management or contamination prevention measures are likely very different if groundwater contamination risk is related with agriculture hazards, mostly caused by diffuse propagation of nutrients, or if relates with infrastructure or industrial hazards mostly caused by direct discharges of domestic or industrial effluents into the streams, rivers and reservoirs.

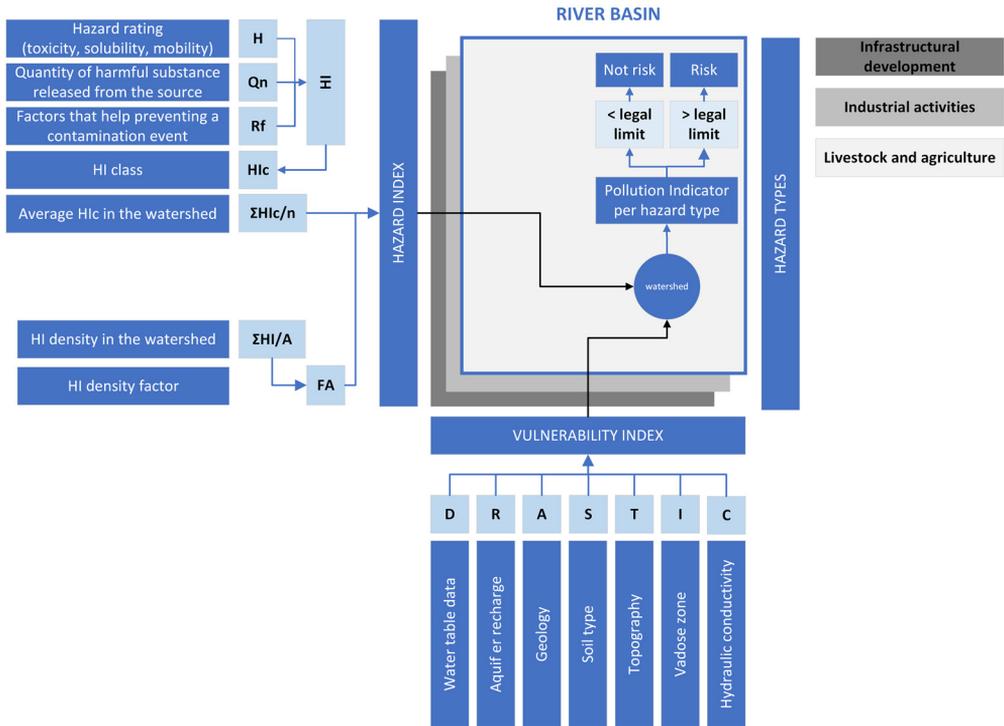


Fig. 2. Method workflow.

### Model development

The method's workflow is illustrated in Fig. 2. It comprises the Vulnerability, Hazard and Risk modules, as described below. To get full acquaintance of all potential risks to groundwater in a watershed, the hazard module must be replicated for infrastructural development, industrial activities and livestock and agriculture types.

### The vulnerability module

The DRASTIC model is well-known and widely used in environmental studies for decades. It is a multi-criteria analysis that set a linear relationship between aquifer vulnerability and seven environmental factors, namely: *D* – Depth to the water table, *R* – Annual recharge, *A* – Aquifer material, *S* – Soil type, *T* – Topographic slope, *I* – Impact of the vadose zone, and *C* – Hydraulic conductivity. Taken altogether, this forms the DRASTIC acronym. Some of these factors are expressed on a numerical scale (e.g., *D*, *R*) while the others are expressed as categories (e.g., *A*, *S*). Irrespective of being numerical or categorical, the environmental factors are evaluated within the target area and subsequently recast into a shared dimensionless scale (ratings) as to allow consistent aggregation (i.e., estimating the DRASTIC index from the parent factors) and mutual comparison. The ratings vary from 1 (lowest vulnerability) to 10 (highest vulnerability). The aggregation is a weighted sum of these ratings that equates the DRASTIC index (*V*):

$$V = (D_w \times D_r) + (R_w \times R_r) + (A_w \times A_r) + (S_w \times S_r) + (T_w \times T_r) + (I_w \times I_r) + (C_w \times C_r) \quad (1)$$

where subscript  $r$  represents the factor rating and subscript  $w$  a factor weight. The original weights were set up by an EPA – Environmental Protection Agency's committee using a consensus (Delphi) approach. The weights were assumed constant, with values  $D_w=5$ ,  $R_w=4$ ,  $A_w=3$ ,  $S_w=2$ ,  $T_w=1$ ,  $I_w=5$ , and  $C_w=3$ . Notwithstanding the criticism received thereafter [11,12], because of being constant and hence not adjustable to local settings, these weights were adopted in the present method article.

The tables used by Aller et al. [9] to convert the environmental factors into ratings are reproduced in Table A1 (Appendix). The conversion of numerical factors (i.e., D, R, T and C) is straightforward because classes in this case are bounded by specific values (e.g., all water table depths below 1.5 m are allocated a vulnerability rating D = 10). This objectivity is lacking when the categorical factors are appraised, because the classes in this case fail to accommodate local peculiarities about geology, soil cover or vadose zone properties. For these factors, adjustment to local settings is frequent and hence the classes and corresponding ratings presented in Table A1 are indicative.

The original DRASTIC V (Eq. (1)) vary from 23 and 226. Before being used in the L-Matrix diagram (Fig. 1), the calculated values must be recast in a 0 – 1 scale.

### The hazards module

The hazard index (HI) describes the harmfulness degree of a hazardous activity and has been equated to [10]:

$$HI = H \times Q_n \times R_f \quad (2)$$

where  $H$  is the hazard's rating related with the toxicity, solubility and mobility of associated harmful substances;  $Q_n$  is a ranking factor (0.8 – 1.2) related with the quantity of those substances, which can be released and further seep into the underground; and  $R_f$  is a reduction factor related with issues that can hinder the occurrence of a contamination event, such as proper maintenance of equipment or security measures implemented along with the hazardous activity. If no information exists on these issues,  $R_f$  should be set to 1.

The ratings assigned to infrastructural development, industrial activities and livestock and agriculture hazards by the European Commission [10] are depicted in Table A2 (Appendix) and were adopted for the present method. They vary from a minimum of 10 and a maximum of 100. Considering the applicable ranking factors between 0.8 and 1.2 and assuming  $R_f = 1$  (Eq. (2)), the HI values will vary between 8 and 120. While drawing hazard maps, this range is simplified through the recast of HI into five qualitative classes ( $HI_c$ ), as follows:  $HI_c = 1$  – very low ( $HI \leq 24$ ),  $HI_c = 2$  – low ( $24 < HI \leq 48$ ),  $HI_c = 3$  – moderate ( $48 < HI \leq 72$ ),  $HI_c = 4$  – high ( $72 < HI \leq 96$ ),  $HI_c = 5$  – very high ( $96 < HI \leq 120$ ).

In the present method, the application of ranking factors ( $Q_n$ ) to the hazard ratings resorted to an environmental classification of activities that was regulated in the state of Minas Gerais, Brazil, through publication of COPAM Normative Resolution no. 217, of December 6, 2017. This legislation attributed a class between 1 and 6 to all sorts of activities, as function of increasing dimension and potential threat for air, soil or water. Thus, the ranking factors in the present method were set to: 0.8 (for Class 1 activities), 1 (Class 2), 1.05 (Class 3), 1.1 (Class 4), 1.15 (Class 5) and 1.2 (Class 6). A very complete inventory of activities with this environmental classification incorporated is public and can be downloaded in GIS format from the IDE-SISEMA dataset (<https://idesisema.meioambiente.mg.gov.br/webgis>). To be used in this method, the IDE-SISEMA database was updated with the hazard ratings of Table A2, after an exercise of correspondence between the class names in both sources.

The target area for application of our method is a river basin and their sub-basins termed watersheds. The  $HI_c$  value of a watershed considers the number of hazardous activities of a certain type (e.g., industrial) developed within its boundaries and takes the average of all  $HI_c$ :

$$HI_c(\text{watershed}) = \frac{\sum HI_c}{n} \quad (3)$$

Then, in order to differentiate watersheds with the same average  $HI_c$  but different spatial hazard densities, the result of Eq. (3) is multiplied by a density factor (FA) arbitrarily set between  $FA_{\min} = 0.8$

and  $FA_{max} = 1.2$  as the ranking factor. First, the hazard index density is calculated by:

$$D_{HI} = \frac{\sum HI}{A} \quad (4)$$

Where  $A$  is the watershed area. Then, the  $D_{HI}$  values representing the various watersheds that compose the river basin are analyzed to verify if they fit a normal, log-normal or other distribution. In heterogeneous basins, where hazardous activities are frequently concentrated in some regions, the distribution of  $D_{HI}$  within the river basin is likely log-normal. In case the distribution is normal or not log-normal, the FA factor is estimated as follows:

$$FA = FA_{max} - (FA_{max} - FA_{min}) \times \frac{D_{HI_{max}} - D_H}{D_{HI_{max}} - D_{Damin}} \quad (5)$$

Where  $D_{HI_{max}}$  and  $D_{HI_{min}}$  represent the maximum and minimum hazard densities in the river basin, respectively. If the distribution is log-normal, then the densities in Eq. (5) are replaced by their natural logarithm counterparts. Finally, the average hazard index of a watershed, considering the hazard density, is obtained as follows:

$$HI_c(\text{watershed, hazard density}) = HI_c(\text{watershed}) \times FA \quad (6)$$

And recast to a 0 – 1 scale before being used in the L-Matrix diagram (Fig. 1).

### The risk module

Having the V and H indices been calculated for a group of watersheds comprising the studied river basin, and a particular hazard category (infrastructural development, industrial activities and livestock and agriculture), the risk module uses a contamination indicator to represent the category and plots the watersheds in the L-Matrix diagram using black or white color if the level is higher or lower than the limit for the water use, respectively. The contamination indicators used in the method refer to concentrations in groundwater, of chloride for infrastructural development, metals for industrial activities, and nitrate for livestock and agriculture. These indicators are not universal and can be customized if a specific study recognizes better indicators for its target river basin. The contamination thresholds depend on the use of groundwater, namely drinking water, industry or agriculture. The L-Matrix diagram populated with black and white circles representing the risky and non-risky watersheds, respectively, allows a comprehensive view over the susceptibility and groundwater contamination risk across the entire river basin, enabling the implementation of specific mitigation measures, if necessary, which will depend on the hazard category and on whether the watersheds fall in the red, orange, green or blue squares. As already emphasized, mitigation measures are likely very different for the livestock and agriculture, industry and infrastructural development categories. For example, inspection of sub-basins (those located in the red square and represented with a black circle) can signify the verification of excessive fertilizer use in agriculture fields, if the category is livestock and agriculture, or the verification of leaks in the sewage systems if the category is infrastructural development. The L-matrix diagram of a category is presumably a forthright management tool and hence must immediately highlight the potential need to implement adequate measures in the threatened sub-basins specifically related with that category. On the other hand, taken the three L-matrix diagrams altogether the global threats over the sub-basins are highlighted and the full list of potential mitigation measures anticipated. It could be possible to develop a L-Matrix diagram capable to represent all categories at once, but we found no obvious advantage to build it, as it would be more complicated to interpret and hence to use as management tool.

### Data requirements and preparation

The present groundwater contamination risk method is implemented in a GIS – Geographic Information System environment (e.g., ArcGIS / ArcMap or ArcGIS Pro), and uses a software with capability to draw graphs (e.g., Microsoft Excel) and generate the L-Matrix diagram (Fig. 1). The GIS software is not only used to prepare the base information and use it to estimate the V and H

indices based on the workflow (Fig. 2), but also to produce vulnerability and hazard/risk maps for the watersheds of a river basin. It uses georeferenced data such as shapefiles or geodatabases, and a number of tools to process them.

There is no general way to indicate the types and sources of data used to build the vulnerability and hazard georeferenced databases, because they will differ from place to place according to the available information. Besides, as the tool names, data requirements, and approaches to execute a specific function differ among GIS computer packages, it's not useful to explain data processing for vulnerability of hazard assessment based on the tools of a specific software. Thus, in the next paragraphs, only general guidance is presented about data types and procedures adopted to build the V and H databases and generate the corresponding outputs.

The preparation of D parameter (DRASTIC module) starts with the creation of a point shapefile that will gather the locations of groundwater sites (dug or drilled wells) and the depths to the water table in the corresponding attribute table. These depths are then interpolated over the river basin, reclassified as D factor (Table A1) and finally mapped in raster format. The preparation of an R database depends on the method selected to estimate the annual recharge [13], but the most appropriate is likely the Recession-Curve Displacement Method [14] because the output is an yearly recharge valid for the watershed scale. This method uses long-term stream flow discharge records to estimate aquifer recharge between each pair of consecutive hydrologic years and calculates an average value over the timeframe. The base information for the R database will be a point shapefile with the average recharges allocated to the watersheds' outlet points, and the path to build the R map will be the same used to generate the D map. A note is given here about the interpolation method we adopted to estimate D and R across the river basin. Both parameters relate with shallow groundwater flow, which is frequently topographically driven. For that reason, we adopted the "Topo to Raster" method of ArcGIS Pro, because that tool was specifically designed to create a surface that more closely represents a natural drainage surface and better preserves both ridgelines and stream networks from input contour data. The A and S factors are usually prepared from geological and soil maps, respectively. From the GIS standpoint, these maps provide the factor's geometrical component in the form of polygon shapefiles. The A and S ratings will be added to the corresponding attribute tables considering the geological and soil types (Table A1). Finally, using shapefile to raster conversion tools embedded in the GIS software, the factor maps are generated. This general path is often adjusted when more specific information is available about the aquifer materials or soil types. For example, in watersheds characterized solely by igneous and metamorphic rocks, the use of Table A1 ratings would result in a map where  $A = 3$  (the typical index) everywhere. In this case, fracture density ( $d$ ) within the watershed can be used to refine the map by setting the A ratings between 2 and 5 (the range in Table A1) as function of increasing  $d$  values [12]. The reason is because fracture density influences groundwater flow and hence aquifer vulnerability. The preparation of T factor starts with the processing of a digital elevation model to obtain terrain slopes across the watershed, using the appropriate GIS tools. Then, the slopes are reclassified to obtain the T classes (Table A1) and the result is used to generate the T factor map in raster format. Specific information on the vadose zone is difficult to obtain, because studies on the hydrologic and material characterization of this layer are much scarcer than, for example, on soil characterization. A close inspection of Table A1 exposes the resemblance between the conversion tables of factors I and A, probably indicating that the DRASTIC authors considered adequate to use geologic maps as base information for the I factor, as they did for the A factor. This is striking because the vadose zone in most landscapes is essentially the bottom of soil profiles, and hence the I factor of a watershed would be better represented by a soil map. In the present method, it is proposed to set the I ratings as function of material (texture) characterization of soil profiles, namely texture characterization of deeper horizons (right above the bedrock). Then, these ratings should be added to the attribute table of a polygon shapefile containing the soil type boundaries within the watershed, and the I factor map generated through shapefile to raster conversion. Information on the C factor is also essentially scarce. This is because representative values of aquifer hydraulic conductivity are mostly the output of pumping tests conducted on drilled wells, and these results are often inexistent or unavailable for a reasonable number of points within the watershed. Public institutions often get access to all existing data (<https://idesisema.meioambiente.mg.gov.br/webgis>) and in some cases

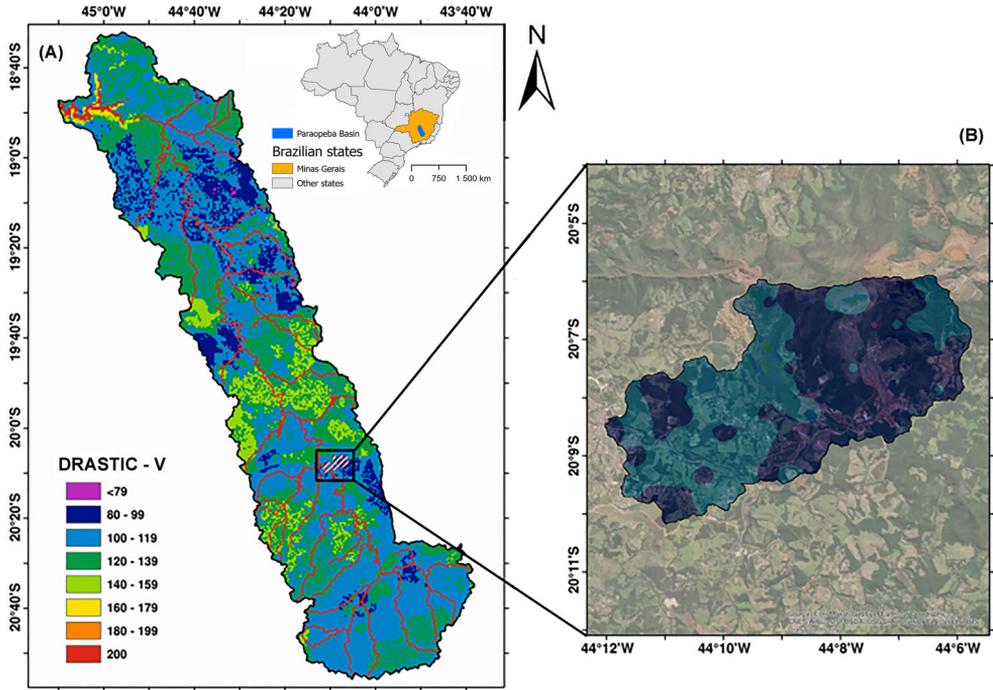
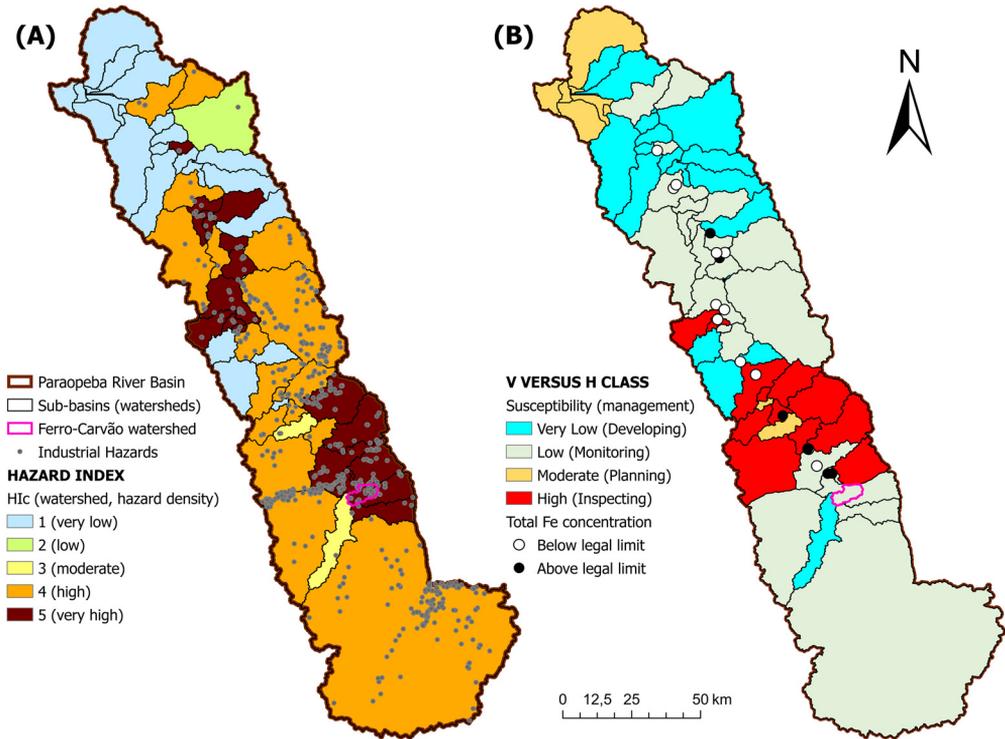


Fig. 3. (a) Vulnerability of Paraopeba River basin to groundwater contamination, determined by the DRASTIC model (b) Detail relative to the Ferro-Carvão watershed.

have used hydraulic conductivities and other relevant information to produce hydrogeologic domain maps. When available, these maps can be used as geometric component of C factor maps. In addition, material or hydraulic conductivity characterization contained in the corresponding attribute tables can be used to assign C ratings and generate the C maps through shapefile to raster conversion.

The assessment of hazard indices of a certain category (infrastructural development, industrial activities, and livestock and agriculture) and production of related maps, requires the preparation two shapefiles: one polygon shapefile with the delineation of a river basin and the corresponding sub-basins (watersheds); one point shapefile containing representative locations of hazardous activities (e.g., <https://idesisema.meioambiente.mg.gov.br/webgis>) and their characterization for hazard rating ( $H$ ; Table A2) and activity class (used to set up the  $Q_n$  factor). Having this data pre-processing done, the GIS software calculates the hazard index of each activity ( $HI$ ; Equation 2; assuming  $R_f = 1$ ) and recasts the result as hazard index class ( $HI_c$ ), using proper attribute calculation and reclassification tools, respectively. The  $HI$  and  $HI_c$  values are added to the point shapefile's attribute table and then the two shapefiles are compared through an attribute-table's intersection procedure, which will quantify the number of hazards per watershed ( $n$ ) and the corresponding  $\Sigma HI$  and  $\Sigma HI_c$  values. These results are added to the polygon shapefile through an attribute-table's join procedure, and then the  $\Sigma HI_c$  and  $n$  values are used to calculate the  $HI_c(\text{watershed})$  (Eq. (3)), while the  $\Sigma HI$  values and watershed areas are used to calculate  $D_{HI}$  (Eq. (4)),  $FA$  (Eq. (5)) and finally  $HI_c(\text{watershed, hazard density})$  (Eq. (6)). It is worth mentioning a note on the point shapefile representing the hazardous activities. In the industrial activity case, point shapefiles are often the available format, but infrastructures and livestock/agriculture hazardous activities are frequently represented by line and polygon shapefiles,

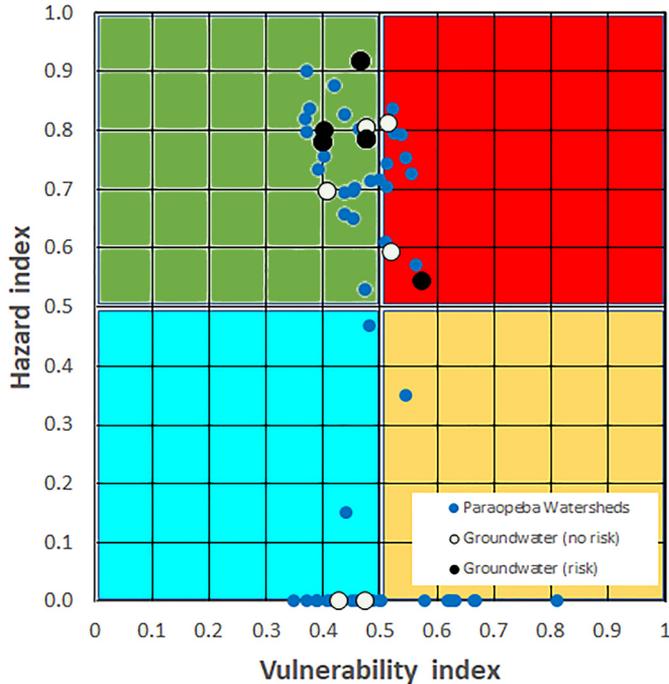


**Fig. 4.** (a) Spatial distribution of industrial hazards and the corresponding hazard index in the Paraopeba River basin; (b) Spatial distribution of susceptibility / management classes, overlapped with levels of total iron (Fe) concentrations in groundwater sites, above or below the Brazilian's legal threshold for drinking water (0.3 mg/L).

respectively. In these cases, the source format is first converted into a point shapefile whereby the lines or polygons are represented by their centroids. However, before conversion the line and polygon shapefiles representing the hazardous activities are clipped by the polygon shapefile containing the watersheds to ensure that activities occupying more than one watershed are represented in all of them.

#### Model validation

The groundwater contamination risk method developed in this study is now tested in the Paraopeba River basin, located in the state of Minas Gerais, Brazil. This basin was selected because this study is part of a scientific project focused on the consequences for water quality, which resulted from the collapse of B1 tailings dam in the “Córrego do Feijão” iron-ore mine owned by Vale mining company and located in the Brumadinho municipality. The catastrophe occurred in 25 January 2019 and released 11.7 million tons of metal-rich tailings, which flowed along the Ferro-Carvão stream, a tributary of Paraopeba River. Part of this mud flow (nearly 2.8 million tons) even entered the Paraopeba's channel 10 km downstream of the dam, immediately raising the concentrations of various metals in the stream water. This alarming circumstance triggered the suspension of this river as source of drinking water supply to the Metropolitan Region of Belo Horizonte, the capital of Minas Gerais state.



**Fig. 5.** L-Matrix diagram of Paraopeba River sub-basins (67 watersheds) and groundwater sites (21 drilled wells; see locations in Fig. 4b).

Various studies integrated on this project have already reported water quality and other environmental impacts caused by the dam break event [15–18], but groundwater have not been investigated so far. In this method article, we present a brief description of some contamination risk results, namely those related with the industrial activities hazards. The full assessment of contamination risks, including the infrastructural development, industrial activities and livestock and agriculture, adequately framed in a literature review and properly discussed and compared with other similar cases, will be published elsewhere.

The dataset supporting the current presentation of results is provided in the Supplementary Materials. The window of time dependent data is 2019–2022. These materials comprise an ArcGIS / ArcMap geodatabase with four feature classes. The first is a polygon feature class composed of 67 Paraopeba River sub-basins (watersheds) characterized for area. The second is a point feature class that includes 747 industrial hazards retrieved from the IDE-SISEMA dataset (<https://idesisema.meioambiente.mg.gov.br/webgis>), characterized for hazard rating ( $H$ ; Table A2) and IDE-SISEMA's activity class. The attribute table also contains the calculated ranking factor ( $Q_n$ ), hazard index ( $HI$ ) and hazard index class ( $HI_c$ ), for reader's own verification of these results while replicating the method with the provided  $H$  and activity class data. The processing of feature classes 1 and 2 as indicated in the methodological sections allows generating the third polygon feature class, which contains the hazard index results assessed at the watershed scale, as well as the concomitant susceptibility / management classes. The attribute table of this feature class also contains the final  $V$  index determined by the DRASTIC model for the entire Paraopeba River basin, but no data is provided relative to the parent factors, because the DRASTIC method is understood and widely applied for a long time. Finally, the fourth is a point feature class containing the concentrations of various elements measured in 21 groundwater sites by the Vale company, used to evaluate the groundwater contamination risk.

The vulnerability of Paraopeba River basin is displayed in Fig. 3a. With exception of an area located near the outlet that is characterized by high DRASTIC vulnerabilities, the Paraopeba River basin is characterized by low to medium V scores. A detail about the DRASTIC vulnerability in the Ferro-Carvão stream is illustrated in Fig. 3b. In this case, the vulnerability is low.

The spatial distribution of industrial hazards within the Paraopeba River basin is shown in Fig. 4a, along with the hazard index class determined by Eq. (6) ( $HIC$  (watershed, hazard density)). A major portion of central and upper Paraopeba's watersheds, fall into the high or very high hazard index category, including the Ferro-Carvão stream watershed.

The combination of Figs. 3a and 4a via the L-Matrix diagram approach (Fig. 1a), allowed generating Fig. 4b where groundwater susceptibility to contamination and general watershed management approaches are indicated. A cluster of watersheds located in the central Paraopeba immediately downstream of Ferro-Carvão watershed, were assigned high susceptibility and included in the inspecting management category (red color sub-basins). The Ferro-Carvão watershed, irrespective of being classified with the very high hazard index label (Fig. 4a), fell in the low susceptibility and monitoring categories (Fig. 4b), because vulnerability is low (Fig. 3a). The red-colored watersheds in Fig. 4b are also the places where 6 among 8 groundwater sites showed total iron concentrations above the legal threshold set for drinking water (0.3 mg/L). The other two sites belong to watersheds labeled as low susceptibility / monitoring. These results are as expected (Fig. 1b) and the most likely interpretation has been detailed before in this document. The L-Matrix diagram used to generate Fig. 4b is displayed in Fig. 5.

### Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

### CRedit authorship contribution statement

**Fernando Antônio Leal Pacheco:** Conceptualization, Methodology, Writing – original draft, Writing – review & editing, Project administration. **Victor Hugo Sarrazin Lima:** Data curation, Investigation, Formal analysis, Visualization. **Teresa Cristina Tarlé Pissarra:** Validation, Resources, Project administration. **Renato Farias do Valle Junior:** Validation, Resources, Project administration. **Marília Carvalho de Melo:** Funding acquisition, Supervision. **Carlos Alberto Valera:** Funding acquisition, Supervision. **João Paulo Moura:** Software. **Luís Filipe Sanches Fernandes:** Validation, Resources, Project administration.

### Acknowledgments

This study was funded by the contract nº5500074952 / 5500074950 / 5500074953, signed between the Vale S.A. company and the following research institutions: Fundação de Apoio Universitário; Universidade de Trás-os-Montes e Alto Douro; and Fundação para o Desenvolvimento da Universidade Estadual Paulista Júlio de Mesquita Filho. The author Renato Farias do Valle Junior received a productivity grant from the CNPq – Conselho Nacional de Desenvolvimento Científico e Tecnológico. For the author integrated in the CITAB research centre, this work was further supported by National Funds of FCT – Fundação para a Ciência e Tecnologia, under the project UIDB/04033/2020. The author integrated in the CITAB research centre is also integrated in the Inov4Agro – Institute for Innovation, Capacity Building and Sustainability of Agri-food Production. The Inov4Agro is an Associate Laboratory composed of two R&D units (CITAB & GreenUPorto). For the author integrated in the CQVR, the research was additionally supported by National Funds of FCT – Fundação para a Ciência e Tecnologia, under the projects UIDB/00616/2020 and UIDP/00616/2020.

### Appendix

Table A1 and Table A2

**Table A1**

DRASTIC conversion tables. Adapted from [9]. Values placed inside brackets are termed typical ratings.

DRASTIC factor	Weight	Ranges and ratings in general	
		Range	Rating
D - Depth to the water table (m)	5	<1.5	10
		1.5-4.6	9
		4.6-9.1	7
		9.1-15.2	5
		15.2-22.9	3
		22.9-30.5	2
		>30.5	1
R - Recharge (mm/year)	4	<51	1
		51-102	3
		102-178	6
		178-254	8
		>254	9
A - Aquifer material	3	Clay shale	1-3(2)
		Metamorphic and igneous rocks	2-5(3)
		Altered metamorphic and igneous rocks	3-5(4)
		Glacial till	4-6(5)
		Sandstone/limestone/clay	5-9(6)
		Massive sandstone	4-9(6)
		Massive limestone	4-9(6)
		Sand/gravel	4-9(8)
		Basalt	2-10(9)
		Karst limestone	9-10(10)
S - Soil type	2	Thin or absent	10
		Ballast	10
		Sand	9
		Peat	8
		Aggregate and/or expandable clay	7
		Sandy loam	6
		Light	5
		Light silt	4
		Clayish Soil	3
		Silt	2
		Not aggregate and/or not expandable clay	1
T - Topographic slope (%)	1	<2	10
		2-6	9
		6-12	5
		12-18	3
		>18	1
I - Impact of the vadose zone	5	Confining layer	1
		Clay/Silt	2-6(3)
		Clayish Schist/mudstone	2-5(3)
		Metamorphic and igneous rock	2-7(6)
		Sandstone	4-8(6)
		Sandstone, limestone, mudstone	4-8(6)
		Sand/Ballast (with more % of silt and clay)	4-8(6)
		Limestone	2-8(4)
		Sand/Ballast	6-9(8)
		Basalt	2-10(9)
Karst limestone	8-10(10)		
C - Hydraulic conductivity (m/day)	3	<4.1	1
		4.1-12.2	2
		12.2-28.5	4
		28.5-40.7	6
		40.7-81.5	8
		>81.5	10

**Table A2**

Hazard rating tables per category. Adapted from [10].

Code	Category: Infrastructural development	Rating	Code	Category: Industrial activities	Rating	Code	Category: Livestock and Agriculture	Rating
1	Hazardous activity		2	Hazardous activity		3	Hazardous activity	
1.1	Waste water		2.1	Mining (in operation and abandoned)		3.1	Livestock	
1.1.1	Urbanization (leaking sewer pipes and sewer systems)	<b>35</b>	2.1.1	Mine, salt	<b>60</b>	3.1.1	Animal barn (shed, cote, sty)	<b>30</b>
1.1.2	Urbanization without sewer systems	<b>70</b>	2.1.2	Mine, other non-metallic	<b>70</b>	3.1.2	Feedlot	<b>30</b>
1.1.3	Detached houses without sewer systems	<b>45</b>	2.1.3	Mine, ore	<b>70</b>	3.1.3	Factory farm	<b>30</b>
1.1.4	Septic tank, cesspool, latrine	<b>45</b>	2.1.4	Mine, coal	<b>70</b>	3.1.4	Manure heap	<b>45</b>
1.1.5	Sewer farm and waste water irrigation system	<b>55</b>	2.1.5	Mine, uranium	<b>80</b>	3.1.5	Slurry storage tank or pool	<b>45</b>
1.1.6	Discharge from an inferior treatment plant	<b>35</b>	2.1.6	Outdoor stock piles of hazardous raw material	<b>85</b>	3.1.6	Area of intensive pasturing	<b>25</b>
1.1.7	Surface impoundment for urban waste water	<b>60</b>	2.1.7	Ore milling and enrichment facilities	<b>70</b>	3.2	Agriculture	
1.1.8	Runoff from paved surfaces	<b>25</b>	2.1.8	Mine waste heap and dirt refuse	<b>70</b>	3.2.1	Open silage (field)	<b>25</b>
1.1.9	Waste water discharge into surface water courses	<b>45</b>	2.1.9	Ore tailings	<b>70</b>	3.2.2	Closed silage	<b>20</b>
1.1.10	Waste water injection well	<b>85</b>	2.1.10	Mine drainage	<b>65</b>	3.2.3	Stockpiles of fertilizers and pesticides	<b>40</b>
1.2	Municipal Waste		2.1.11	Tailing pond	<b>65</b>	3.2.4	Intensive agriculture area (with high demand of fertilizers and pesticides)	<b>30</b>
1.2.1	Garbage dump, rubbish bin, litter bin	<b>40</b>	2.2	Excavation sites		3.2.5	Allotment garden	<b>15</b>
1.2.2	Waste loading station and scrap yard	<b>40</b>	2.2.1	Excavation and embankment for development	<b>10</b>	3.2.6	Greenhouse	<b>20</b>
1.2.3	Sanitary landfill	<b>50</b>	2.2.2	Gravel and sand pit	<b>30</b>	3.2.7	Waste water irrigation	<b>60</b>
1.2.4	Spoils and building rubble depository	<b>35</b>	2.2.3	Quarry	<b>25</b>			
1.2.5	Sludge from treatment plants	<b>35</b>	2.3	Oil and gas exploitation				
1.3	Fuels		2.3.1	Production wells	<b>40</b>			
1.3.1	Storage tank, above ground	<b>50</b>	2.3.2	Reinjection wells	<b>70</b>			
1.3.2	Storage tank, underground	<b>55</b>	2.3.3	Loading station	<b>55</b>			
1.3.3	Drum stock pile	<b>50</b>	2.3.4	Oil pipeline	<b>55</b>			
1.3.4	Tank yard	<b>50</b>	2.4	Industrial plants (none mining)				
1.3.5	Fuel loading station	<b>60</b>	2.4.1	Smelter	<b>40</b>			
1.3.6	Gasoline station	<b>60</b>	2.4.2	Iron and steel works	<b>40</b>			
1.3.7	Fuel storage cavern	<b>65</b>	2.4.3	Metal processing and finishing industry	<b>50</b>			
1.4	Transport and traffic		2.4.4	Electroplating works	<b>55</b>			

(continued on next page)

**Table A2** (continued)

Code	Category: Infrastructural development	Rating	Code	Category: Industrial activities	Rating	Code	Category: Livestock and Agriculture	Rating
1	Hazardous activity		2	Hazardous activity		3	Hazardous activity	
1.4.1	Road, unsecured	<b>40</b>	2.4.5	Oil refinery	<b>85</b>			
1.4.2	Road tunnel, unsecured	<b>40</b>	2.4.6	Chemical factory	<b>65</b>			
1.4.3	Road hauler depot	<b>35</b>	2.4.7	Rubber and tyre industry	<b>40</b>			
1.4.4	Car parking area	<b>35</b>	2.4.8	Paper and pulp manufacture	<b>40</b>			
1.4.5	Railway line	<b>30</b>	2.4.9	Leather tannery	<b>70</b>			
1.4.6	Railway tunnel, unsecured	<b>30</b>	2.4.10	Food industry	<b>45</b>			
1.4.7	Railway station	<b>35</b>	2.5	Power plants				
1.4.8	Marshalling yard	<b>40</b>	2.5.1	Gasworks	<b>60</b>			
1.4.9	Runway	<b>35</b>	2.5.2	Caloric power plants	<b>50</b>			
1.4.10	Pipeline of hazardous liquids	<b>60</b>	2.5.3	Nuclear power plant	<b>65</b>			
1.5	Recreational facilities		2.6	Industrial storage				
1.5.1	Tourist urbanization	<b>30</b>	2.6.1	Stock piles of raw materials and chemicals	<b>60</b>			
1.5.2	Camp ground	<b>30</b>	2.6.2	Containers for hazardous substances	<b>70</b>			
1.5.3	Open sport stadium	<b>25</b>	2.6.3	Cinder tip and slag heaps	<b>70</b>			
1.5.4	Golf course	<b>35</b>	2.6.4	Non-hazardous waste site	<b>45</b>			
1.5.5	Skiing course	<b>25</b>	2.6.5	Hazardous waste site	<b>90</b>			
1.6	Diverse hazards		2.6.6	Nuclear waste site	<b>100</b>			
1.6.1	Graveyard	<b>25</b>	2.7	Diverting and treatment of waste water				
1.6.2	Animal burial	<b>35</b>	2.7.1	Waste water pipelines	<b>65</b>			
1.6.3	Dry cleaning premises	<b>35</b>	2.7.2	Surface impoundment for industrial waste water	<b>65</b>			
1.6.4	Transformer station	<b>30</b>	2.7.3	Discharge of treatment plants	<b>40</b>			
1.6.5	Military installations and dereliction	<b>35</b>	2.7.4	Waste water injection well	<b>85</b>			

## Supplementary materials

Supplementary material associated with this article can be found, in the online version, at doi:[10.1016/j.mex.2022.101858](https://doi.org/10.1016/j.mex.2022.101858).

## References

- [1] S. Saidi, S. Bouri, H. Ben Dhia, B. Anselme, Assessment of groundwater risk using intrinsic vulnerability and hazard mapping: application to Souassi aquifer, Tunisian Sahel, *Agric. Water Manag.* 98 (2011) 1671–1682, doi:[10.1016/j.agwat.2011.06.005](https://doi.org/10.1016/j.agwat.2011.06.005).
- [2] J. Wang, J. He, H. Chen, Assessment of groundwater contamination risk using hazard quantification, a modified DRASTIC model and groundwater value, Beijing Plain, China, *Sci. Total Environ.* 432 (2012) 216–226, doi:[10.1016/j.scitotenv.2012.06.005](https://doi.org/10.1016/j.scitotenv.2012.06.005).
- [3] E. Nixdorf, Y. Sun, M. Lin, O. Kolditz, Development and application of a novel method for regional assessment of groundwater contamination risk in the Songhua River Basin, *Sci. Total Environ.* 605–606 (2017) 598–609, doi:[10.1016/j.scitotenv.2017.06.126](https://doi.org/10.1016/j.scitotenv.2017.06.126).
- [4] A. Matzeu, R. Secci, G. Uras, Methodological approach to assessment of groundwater contamination risk in an agricultural area, *Agric. Water Manag.* 184 (2017) 46–58, doi:[10.1016/j.agwat.2017.01.003](https://doi.org/10.1016/j.agwat.2017.01.003).
- [5] M.A. Jenifer, M.K. Jha, Comprehensive risk assessment of groundwater contamination in a weathered hard-rock aquifer system of India, *J. Clean. Prod.* 201 (2018) 853–868, doi:[10.1016/j.jclepro.2018.08.005](https://doi.org/10.1016/j.jclepro.2018.08.005).
- [6] Y. Teng, R. Zuo, Y. Xiong, J. Wu, Y. Zhai, J. Su, Risk assessment framework for nitrate contamination in groundwater for regional management, *Sci. Total Environ.* 697 (2019) 134102, doi:[10.1016/j.scitotenv.2019.134102](https://doi.org/10.1016/j.scitotenv.2019.134102).
- [7] Q. Zhang, P. Li, Q. Lyu, X. Ren, S. He, Groundwater contamination risk assessment using a modified DRATICL model and pollution loading: a case study in the Guanzhong Basin of China, *Chemosphere* 291 (2022) 132695, doi:[10.1016/j.chemosphere.2021.132695](https://doi.org/10.1016/j.chemosphere.2021.132695).
- [8] European Commission, WFD Pressures and impacts assessment methodology: guidance on pressures and impacts methodology, (2003) 53.
- [9] L. Aller, T. Bennet, J.H. Lehr, R.J. Petty, G. Hackett, DRASTIC: a standardized system for evaluating groundwater pollution potential using hydro geologic settings., USEPA Doc. No. EPA/600/2-85-018. (1987).
- [10] F. Zwahlen, COST Action 620: Vulnerability and risk mapping for the protection of carbonate (karst) aquifers, European Commission Report EUR 20912, Official Publications of the European Communities, Luxembourg, 320 p, ISBN 92-894-6416-X, 2003.
- [11] F.A.L. Pacheco, L.M.G.R. Pires, R.M.B. Santos, L.F. Sanches Fernandes, Factor weighting in DRASTIC modeling, *Sci. Total Environ.* 505 (2015) 474–486, doi:[10.1016/j.scitotenv.2014.09.092](https://doi.org/10.1016/j.scitotenv.2014.09.092).
- [12] F.A.L. Pacheco, L.F. Sanches Fernandes, The multivariate statistical structure of DRASTIC model, *J. Hydrol.* 476 (2013) 442–459, doi:[10.1016/j.jhydrol.2012.11.020](https://doi.org/10.1016/j.jhydrol.2012.11.020).
- [13] A.L. Coes, T.B. Spruill, M.J. Thomasson, Multiple-method estimation of recharge rates at diverse locations in the North Carolina Coastal Plain, USA, *Hydrogeol. J.* 15 (2007) 773–788, doi:[10.1007/s10040-006-0123-3](https://doi.org/10.1007/s10040-006-0123-3).
- [14] W.-P. Chen, C.-H. Lee, Estimating ground-water recharge from streamflow records, *Environ. Geol.* 44 (2003) 257–265, doi:[10.1007/s00254-002-0753-2](https://doi.org/10.1007/s00254-002-0753-2).
- [15] F.A.L. Pacheco, R.F. do Valle Junior, M.M.A.P. de Melo Silva, T.C.T. Pissarra, M. Carvalho de Melo, C.A. Valera, L.F. Sanches Fernandes, Prognosis of metal concentrations in sediments and water of Paraopeba River following the collapse of B1 tailings dam in Brumadinho (Minas Gerais, Brazil), *Sci. Total Environ.* 809 (2022) 151157, doi:[10.1016/j.scitotenv.2021.151157](https://doi.org/10.1016/j.scitotenv.2021.151157).
- [16] F.A.L. Pacheco, M.D. de Oliveira, M.S. Oliveira, M. Libânio, R.F. do Valle Junior, M.M.A.P. de Melo Silva, T.C.T. Pissarra, M.C. de Melo, C.A. Valera, L.F.S. Fernandes, Water security threats and challenges following the rupture of large tailings dams, *Sci. Total Environ.* 834 (2022) 155285, doi:[10.1016/j.scitotenv.2022.155285](https://doi.org/10.1016/j.scitotenv.2022.155285).
- [17] R.G. Mendes, R.F. do Valle Junior, M.M.A.P. de Melo Silva, L.F. Sanches Fernandes, A.C. Pinheiro Fernandes, T.C.T. Pissarra, M.C. de Melo, C.A. Valera, F.A.L. Pacheco, A partial least squares-path model of causality among environmental deterioration indicators in the dry period of Paraopeba River after the rupture of B1 tailings dam in Brumadinho (Minas Gerais, Brazil), *Environ. Pollut.* 306 (2022) 119341, doi:[10.1016/j.envpol.2022.119341](https://doi.org/10.1016/j.envpol.2022.119341).
- [18] T.C.T. Pissarra, R.C.A. Costa, R.F. do Valle Junior, M.M.A.P. de M. Silva, A.M. da Costa, L.F. Sanches Fernandes, M. Carvalho de Melo, C.A. Valera, F.A.L. Pacheco, Role of mine tailings in the spatio-temporal distribution of phosphorus in river water: the case of B1 Dam break in Brumadinho, *Water* 14 (2022) 1572, doi:[10.3390/w14101572](https://doi.org/10.3390/w14101572).