



Large scale risk assessment methodology and application to selected areas

Deliverable 2.2

AM

Document information

Grant Agreement Number	n° 101112889
Project Title	Information-based Strategies for LAND Remediation
Project Acronym	ISLANDR
Project Coordinator	Nike Luodes
Project Duration	1 May 2023 – 30 April 2026 (36 months)
Related Work Package	Work Package 2
Lead Organisation	BRGM
Contributing Partner(s)	GreenDecision, BRGM: Giulia Minolfi (GD) Fabio Rosada (GD) Elisa Giubilato (GD) Lisa Pizzol (GD) Alex Zabeo (GD) Stephane Belbeze (BRGM) Virginie Derycke (BRGM) Timo Tarvainen (GTK)
Submission Date	14.11.2025
Dissemination Level	Public

The content of this deliverable reflects only the author's view. The European Commission is not responsible for any use that may be made of the information it contains.

History

Date	Submitted by	Reviewed by	Version (Notes)
22/09/2025	Giulia Minolfi	GreenDecision staff, WP2 partners	1.1 – First Draft – Comments received from MC
14/11/2025	Giulia Minolfi	GreenDecision staff, WP2 partners	1.2 – Final Draft – modified according to MC comments

AWAITING APPROVAL BY THE EUROPEAN COMMISSION

Summary

Soils provide vital ecosystem services, such as food production, water regulation and biodiversity support, yet their protection has long been underestimated. In Europe, the lack of harmonized standards and monitoring approaches has led to fragmented policies, leaving soil vulnerable to degradation and pollution. Contamination arises from a wide variety of sources, including industrial activities, uncontrolled waste disposal, agriculture, mining, and emerging pollutants such as those identified in WP2.2. Contaminants can spread through air, water and food chains, posing risks to human health, ecosystems, and the economy.

To address these challenges, the European Union has placed soils at the core of its Green Deal and Zero Pollution Action Plan, with specific strategies such as the EU Soil Strategy for 2030 and the Soil Monitoring and Resilience Law (SML). These initiatives promote a risk-based approach to contaminated land management, requiring Member States to identify, assess and manage both point source and diffuse contamination, while ensuring sustainable land use.

The ISLANDR project supports these objectives by developing innovative, information-based strategies for soil remediation and land management. It aims to provide harmonized data, advanced risk-based methods, and spatial tools to improve the prioritization of contaminated and potentially contaminated sites across Europe. A metadata catalogue of soil contamination will be compiled and combined with novel interpolation techniques, forming the basis for regional-scale risk assessment methodologies.

Two complementary approaches are being developed in this Task:

- Prioritization of potentially contaminated sites, based on historical and current data on anthropogenic activities, to identify hotspots requiring further site-specific investigation or remediation;
- Assessment of diffuse contamination, using measured concentration data to identify areas where widespread pollution may pose risks to health and the environment.

Both approaches are implemented within a GIS-based framework, integrating hazards, pathways, and receptors to produce regional-scale risk maps. These outputs will provide decision-makers with scientifically robust tools to guide land-use planning, investment, and remediation strategies, ensuring that resources are directed towards the areas of highest risk.

This Deliverable presents a GIS-based framework for Regional Risk Assessment (RRA) that requires two key types of input data: an inventory of Potentially Contaminated Sites (PCSs) and geochemical background data for diffuse contamination. Both data categories are expected to be progressively available under the SML. The methodology follows a stepwise, risk-based logic consistent with Article 43b of the SML, enabling prioritisation before detailed investigations. The framework was successfully tested in the ISLANDR Toulouse Test Area (ITA), demonstrating its capacity to rank PCSs and identify diffuse contamination hotspots, supporting early-stage decision-making at regional scale.

Keywords

Risk assessment, regional, contaminated soil, point source, contaminated site, prioritization, GIS, spatial analysis, diffuse contamination, decision making.

Abbreviations and acronyms

Acronym	Description
WP	Work Package
C&D	Communication & Dissemination
ITA	ISLANDR Test Area
RRA	Regional Risk Assessment
PCSSs	Potentially contaminated sites

AWAITING APPROVAL BY THE EUROPEAN COMMISSION



Table of Contents

Document information	1
History	2
Keywords	3
Abbreviations and acronyms	4
List of Figures	6
List of Tables	6
Introduction.....	8
1. Regional Risk Assessment of Potentially Contaminated Sites	11
1.1. Methodology for RRA of PCSs	13
1.1.1. Definition of the conceptual model.....	13
1.1.2. GIS-based reclassification of parameters and attributes	15
1.1.3. Hazard analysis	18
1.1.4. Pathway analysis	20
1.1.5. Vulnerability analysis	23
1.1.6. Relative Regional Risk estimation.....	25
1.1.7. Model Builder in GIS	26
2. Regional Risk Assessment of diffuse contamination.....	26
2.1. Methodology for Regional Risk Assessment of diffuse contamination	27
2.1.1. Hazard definition.....	28
2.1.2. Pathway analysis for diffuse contamination.....	28
2.1.3. Vulnerability analysis for diffuse contamination	30
2.1.4. Total risk per contaminant.....	31
2.1.5. Cumulative risk for diffuse contamination.....	32
3. Regional Risk Assessment for Potential Contaminated Sites - application to ITA Toulouse Métropole	32
3.1. Study area.....	32
3.2 Hazard Analysis	33
3.3 Pathway Analysis	34
3.3.1. Groundwater and Surface water pathways	34
3.3.2. Air pathway.....	35
3.4 Vulnerability Analysis	36
3.4.1. Groundwater Vulnerability.....	36
3.4.2. Surface water Vulnerability	37
3.4.3. Human health Vulnerability	38
3.4.4. Protected Areas Vulnerability.....	39
3.5 Total Risk map	40
4. Regional Risk Assessment for diffuse contamination - application to ITA Toulouse Métropole	42

4.1. Hazard Analysis	42
4.2. Pathway Analysis	44
4.3. Vulnerability Analysis	45
4.4. Contaminants and total risk maps	45
5. Advantages and limits of the methodology	52
6. Conclusions	53
Bibliography	57

List of Figures

Figure 1. Structure of the Regional Risk assessment methodology for the prioritisation of contaminated sites and related assessment steps	13
Figure 2. Exposure diagram for regional risk assessment, with the selected pathways linking the contamination sources to the relevant receptors	14
Figure 3. Schematic representation of transport processes from a potentially contaminated site	21
Figure 4. Pathway-related (a) Groundwater, b) Surface water, c) Air, d) Direct contact) hazard scores for 40 PCSs in Toulouse Métropole	34
Figure 5. Examples of a) Groundwater, b) Surface water pathways for one PCS in Toulouse Métropole	35
Figure 6. Aridity Index for Toulouse Métropole. All the area has a score of 1 in the reclassified map (<25)	36
Figure 7. Reclassified Groundwater vulnerability for Toulouse Métropole. A score of 0.3 Is for industrial use of GW, 0.85 is for agricultural use, 0.1 for no use	37
Figure 8. Reclassified map for Surface Water vulnerability in the Toulouse Métropole	38
Figure 9. Reclassified Human health vulnerability map for Toulouse Métropole	39
Figure 10. Reclassified map for Protected areas vulnerability in Toulouse Métropole	40
Figure 11. Aggregated overall risk of the selected PCS in Toulouse Métropole	41
Figure 12. Pb hazard in Toulouse Métropole: a) absolute values; b) Pb hazard normalized with CSC thresholds; c) Pb normalized hazard in agricultural areas; d) Pb normalized hazard in residential areas; e) Pb normalized hazard in industrial/commercial areas	44
Figure 13. Reclassified Groundwater pathway map for diffuse contamination in Toulouse Métropole	45
Figure 14. Groundwater risk map from Pb diffuse contamination in Toulouse Métropole soils. Maximum risk is 9.6	46
Figure 15. Surface water risk map from Pb diffuse contamination in Toulouse Métropole soils. Maximum risk is 0.68. Pixels representing the intersection between the Surface water pathway and surface water bodies are not visible at this scale	47
Figure 16. Air risk map for human health due to Pb diffuse contamination in Toulouse Métropole soils. Maximum risk is 8.5	48
Figure 17. Air exposure risk map for protected areas due to Pb diffuse contamination in Toulouse Métropole soils. Maximum risk is 0.57	49
Figure 18. Direct contact risk map for human health due to Pb diffuse contamination in Toulouse Métropole soils. Maximum risk is 8.5	50
Figure 19. Total risk map due to Pb diffuse contamination in Toulouse Métropole soils	51

List of Tables

Table 1. Pathway parameters and scores	16
Table 2. Vulnerability attributes and scores for Groundwater receptor	17
Table 3. Vulnerability attributes and scores for Surface Water receptor	18
Table 4. Vulnerability attributes and scores for Human Health receptor	18



Table 5. Vulnerability attributes and scores for Protected areas receptor..... 18

Table 6. Fallback mapping for risk phrases. 19

Table 7. Selected characteristics for receptor’s vulnerability evaluation. 24

Table 8. Results for groundwater (GW), surface water (SW), direct contact (DC) and air risk (max and mean values) and total risk for 40 selected PCSs in Toulouse Métropole. HH: human health, PA: protected areas. 42

Table 9. CSC (Contamination Screening Concentrations, corresponding to the Italian Concentrazioni Soglia di Contaminazione as defined in D.Lgs. 152/2006) for considered contaminants in residential/green, agricultural and industrial/commercial areas..... 43

AWAITING APPROVAL BY THE EUROPEAN COMMISSION

Introduction

Four years after the Soil Monitoring and Resilience Law draft (SML, 2023) and after two months that the Law officially enters into force, Member States are expected to ensure that risk for human health and the environment of potentially contaminated and contaminated sites are identified, managed and kept to acceptable levels (article 12).

The effort required to identify and prioritize contaminated sites in Europe is considerable (Pérez et al., 2018). Determining whether a site is contaminated requires soil and groundwater investigations, but it is neither feasible nor cost-effective to assess all land across the EU within a short timeframe. In this perspective, a definition of “potentially contaminated site” is desirable: a “potentially contaminated site” is a “site where an activity is or has been operated that may have caused soil contamination” (Van-Camp et al., 2004). Maps, historical risks activities, historical archives, local knowledge, industrial permits and license records, administrative information, surveys of surface and groundwater quality and site inspections may provide indications that a site may be potentially contaminated (Pizzol et al., 2011). Once the list of potentially contaminated sites is available, a prioritization according to relative risk assessment procedures can be performed to select sites where a preliminary soil investigation is required first (Van-Camp et al., 2004). These risk assessment procedures are performed at larger scales (i.e. at regional scale) and should take into account that over a regional area, impacts caused by multiple sources are distributed and often multiplied on diverse receptors through different pathways. Therefore, the risk assessment process can be implemented following a tiered approach which distinguishes between an initial screening phase (Tier 1) for the identification of potentially *serious* contamination which triggers site specific investigation and the determination of *urgent* sites (Tier 2), if there is unacceptable risk for human health and environment which triggers remediation actions. Tier 1 is performed by using regional relative risk assessment, while Tier 2 is performed using site-specific risk assessment.

Regional risk assessment (RRA) can be defined as a quantitative and systematic procedure to estimate and compare the impacts of environmental problems that affect large geographic areas (Hunsaker et al., 1990); it deals with multiple habitats and multiple sources releasing multiple stressors which affect multiple endpoints (Landis, 2005). The result of a RRA is the ranking of risks in terms of their magnitude, rather than to produce absolute estimates of health/ecological impacts, in order to select those to be investigated more thoroughly. This risk approach is called relative risk assessment (Pizzol et al., 2011, Pizzol et al., 2016, Limasset et al., 2018).

To support prioritisation when dealing with a very large number of sites, the SML suggests following a risk-based and stepwise approach to identify and investigate potentially contaminated sites (43b). It also outlines that the evaluation of potential risk used in such prioritisation is much more generic than the site-specific risk assessment. RRA supports such a stepwise approach supporting Member States (MS) in securing prioritization of risk investigation measures.

Within this conceptual framework, the ISLANDR project contributes within this Deliverable, with two complementary large-scale methodologies that support the operationalisation of Tier 1 decision-making process:

1. A RRA for Potentially Contaminated Sites (PCSs), to identify and rank active or inactive potentially contaminating activities, thus guiding site specific investigation and further site-specific risk assessment on the highest priority sites.

2. A RRA for Diffuse Contamination, which requires contamination data to address widespread, low-level contamination and verify if, in some areas, it can lead to serious risk for human and ecosystems. In this case, the output supports land management and spatial planning decisions aimed at limiting potential long-term impacts rather than triggering intrusive investigations.

This dual methodological structure ensures that the ISLANDR framework remains consistent with the European Union's risk-based and proportionate approach, targeting resources and detailed characterisation only where evidence suggests a higher likelihood of significant impact.

To address these challenges, RRA requires a structured conceptual framework capable of linking contamination sources, environmental transfer mechanisms, and potential receptors. The methodological approach adopted in this work is grounded in the consolidated framework of RRA, originally proposed by Landis and colleagues (Landis, 2005; Landis & Wieggers, 1997) as a structured method to evaluate and compare multiple stressors acting across large geographic areas. The RRA framework emphasizes the spatial integration of sources, pathways, and receptors (S-P-R), allowing for a relative ranking of risks rather than absolute quantification, and has been widely applied in ecological and human health risk studies at regional scales.

Originally proposed by Holdgate (1979–80) to describe the transfer of pollutants from a source, through environmental pathways, to potential receptors, the S-P-R model has since become a widely adopted framework for environmental risk assessment. At the European level, a significant effort has been made to harmonize and operationalize such approaches through initiatives led by the European Commission and the European Environment Agency, culminating in the development of the PRA.MS methodology (EEA, 2005a, EEA, 2005b). This procedure provided a transparent and reproducible scoring system for hazard, pathway and receptor characterization, representing one of the earliest EU-level attempts to implement a risk-based approach for contaminated site management. This model has been further developed and adapted for application at regional scale (Pizzol et al., 2011, Pizzol et al., 2016, Limasset et al., 2018, Minolfi et al., 2018).

The methodology presented in this Deliverable, builds upon these consolidated foundations, updating and adapting them to current European regulatory and scientific contexts, particularly to the principles introduced by the SML. Rather than proposing a completely new framework, our objective was to refine and operationalize existing, well-established methods through their large-scale implementation in a GIS environment. This ensures both scientific robustness and policy alignment, while maintaining continuity with the EU's long-standing efforts toward risk-based and data-driven soil management.

The two complementary methodologies have been designed and applied to a case study in the ISLANDR Test Area (ITA), namely the metropolitan area of Toulouse (37 municipalities in southern France). This region presents diverse land uses and levels of urbanization - ranging from dense urban cores to dispersed rural landscapes - which makes it particularly suitable to demonstrate the flexibility of the approach.

The first methodology addresses PCSs, defined as locations where past or ongoing anthropogenic activities may have caused soil contamination. Where dedicated datasets are not available, PCSs can be identified through a combination of historical archives, industrial permits and licenses, administrative information, and local knowledge. The outcome of this methodology is a relative ranking of PCSs at regional scale, providing decision-makers with a prioritization tool to guide further investigations and, where appropriate, site-specific risk assessments or remediation actions.



The second methodology focuses on diffuse soil contamination, which arises from widespread anthropogenic emissions rather than from discrete point sources. This assessment requires quantitative input data, and in ISLANDR it is supported by regional datasets and probabilistic interpolation models developed in WP1 to handle sparse, imprecise, and clustered data. By integrating exposure pathways with receptor sensitivity, the methodology produces a relative ranking of areas affected by diffuse contamination, highlighting those that may pose risks to human health or the environment. These results can inform land-use planning, support regulatory development, and indicate priority areas for monitoring or, if needed, site-specific follow-up.

In both approaches, GIS-based spatial analysis plays a central role. Geospatial tools are used to represent the spatial distribution of hazards, pathways, and receptors, allowing the identification of high-priority sites or areas and supporting a robust and scalable regional risk assessment framework.

AWAITING APPROVAL BY THE EUROPEAN COMMISSION

1. Regional Risk Assessment of Potentially Contaminated Sites

Soil plays a crucial role in supporting both human life and ecosystems. A high-quality soil contributes to food security, clean water provision, sustainable agriculture, biodiversity conservation and climate change mitigation (EEA, 2023). However, a wide range of human activities - including industries, waste disposal in uncontrolled landfills, mining, the use of agricultural chemicals, sewage sludge, and livestock waste, environmental accidents - has left a legacy of contaminated sites all around the world (Ramón et al., 2019). Key contaminants include polycyclic aromatic hydrocarbons (PAHs), heavy metals, persistent organic pollutants (e.g. PCBs, dioxins), radionuclides, and pathogenic microorganisms are the main soil contaminants causing concerns. More recently, new emerging soil contaminants, such as nanomaterials, veterinary medicines, microplastics and surfactants, have raised further concerns. All these substances, both the well-known and the new emerging contaminants, may persist in soils, leading to long-term ecological degradation and spread through air, surface water and groundwater, leading to cross-borders impacts on ecosystems and human health. In some cases, contaminants are taken up by plants and may bioaccumulate in the food chain, compromising the safety of the food consumed by both humans and animals (FAO, 2018).

The systematic identification of contaminated sites, combined with risk-based remediation decision-making, is widely recognized as a necessary step to public health and guide sustainable land management (Bardos et al., 2018). Rather than assuming that all sites require immediate remediation, such approaches enable proportional and targeted actions where risks are likely to be unacceptable. European policy has increasingly emphasized soil protection, as evidenced by various soil-focused initiatives and strategies, including the European Green Deal, the Biodiversity Strategy, and the new EU Soil Strategy. In line with these frameworks, risk-based decision-making is central to achieve the objectives of the Zero Pollution Action Plan (European Commission, 2021). In this context, the proposed Soil Monitoring and Resilience Directive highlights the urgency to managing both point-source and diffuse contamination, with the goal of reducing soil pollution to non-harmful levels by 2050. The directive was approved by the European Parliament in April 2024, and as of September 2025, is awaiting the final adoption from the Council of the European. Article 12 of the directive calls on Member States to ensure that risks of potentially contaminated and contaminated sites are identified, managed and kept to acceptable levels. Within four years of the directive's entry into force, Member States must also implement a step wise risk-based framework for identifying, investigating, and managing potentially contaminated sites, including the definition of thresholds for unacceptable risks to human health and the environment. This process primarily concerns generic or screening-level risk assessment, which enables consistent prioritization and proportional management actions across the EU.

The scale and complexity of this task are considerable (Pérez et al., 2018). Since comprehensive soil and groundwater investigations cannot feasibly cover all EU land in a short timeframe, especially with limited resources, priority should be given to areas where the likelihood of contamination - and associated risk - is higher. In response, a new methodology for RRA designed to identify and rank potentially contaminated sites at regional scale, has been developed, especially for areas where detailed contamination data are lacking. The aim is not to replace site-specific assessments, but rather to support them by targeting locations that warrant further investigation.

Risk assessment methods for contaminated sites have been developed by several national environmental agencies (e.g., APAT, 2008; DEFRA, 2011; USEPA, 1991, 2005; Health Canada, 2010 a,b). At

the site-specific level, risk is often expressed as the product of a hazard (the intrinsic toxicity of contaminants and their environmental concentrations) and exposure (the probability that receptors - human or ecological - come into contact with contaminants through specific pathways). These assessments may involve sampling and lab analyses, often resulting in quantitative indices for both carcinogenic and non-carcinogenic effects (Chen et al., 2022). They provide absolute evaluations of risk and are typically tailored to individual receptors and land uses.

At a regional scale, however, applying these detailed methods is neither feasible nor appropriate, given the broader scope and different data availability. In regional applications, risk assessment often relies on simplified or generalized assumptions about exposure scenarios and land use, reflecting the limited availability of detailed site data at large scale. Consequently, such assessments typically aim to compare relative risk levels across the study area, rather than to quantify absolute risk levels. This distinction is crucial: while site-specific assessments aim to determine whether a specific location poses an unacceptable risk, RRA methods are designed to rank a large number of sites to support prioritization decisions in a stepwise approach as suggested in the SML (43b). This comparative approach has been widely adopted in frameworks such as the Relative Risk Model (Landis & Wieggers, 1997; Landis, 2005) and the PRA.MS methodology (EEA, 2005b). Spatial heterogeneity in contamination patterns, land use, and receptor distributions further complicates the picture (Albanese et al., 2014; Pizzol et al., 2011). Accordingly, regional assessments rely on simplified but robust frameworks that can guide further investigations or policy responses (Hunsaker et al., 1990; Landis, 2005).

Our methodology builds on well-established methods in this field (Pizzol et al., 2011, EEA 2005a, EEA 2005b), that has been extensively applied in recent years (Su et al., 2025, Jiang et al., 2024, Pizzol et al., 2016, Li et al., 2014) and has been applied to a complex case study: the metropolitan area of Toulouse, France, which includes 37 municipalities with diverse land uses - agricultural, forested, industrial, commercial, and transportation corridors. This area presents a range of environmental pressures and contamination legacies, making it a suitable testbed for evaluating the applicability and scalability of the methodology. The objective is to prioritize potentially contaminated sites, defined as those with historical or ongoing human activities that may have caused pollution. Where official datasets are incomplete, auxiliary sources such as local knowledge, historical archives, land use records, and site inspections provide supporting evidence.

The outcome of the methodology is a relative risk-based ranking of sites to inform which locations may merit further site-specific investigation. This supports more efficient allocation of efforts and resources, focusing attention where the potential risk is highest. The method includes a refinement of the PRA.MS scoring system (EEA, 2005b), updated to align with the current European CLP Regulation (EC No 1272/2008), thereby ensuring consistency with hazard classification based on the Globally Harmonised System (GHS). This updated hazard evaluation captures substance-specific behaviour across environmental compartments (soil, water, air), enhancing the precision of risk estimations.

While many existing approaches are either descriptive or inventory-based, our framework is explicitly designed as a decision-support tool. It offers a pragmatic solution to the challenge of prioritizing contaminated land at scale and aligns with EU goals for environmental restoration and pollution prevention.

1.1. Methodology for RRA of PCSs

The proposed methodology for RRA of PCSs is structured in five sequential steps (Figure 1) to support prioritisation of risk identification, evaluation and management in a stepwise approach. The following paragraphs describe each step in detail.

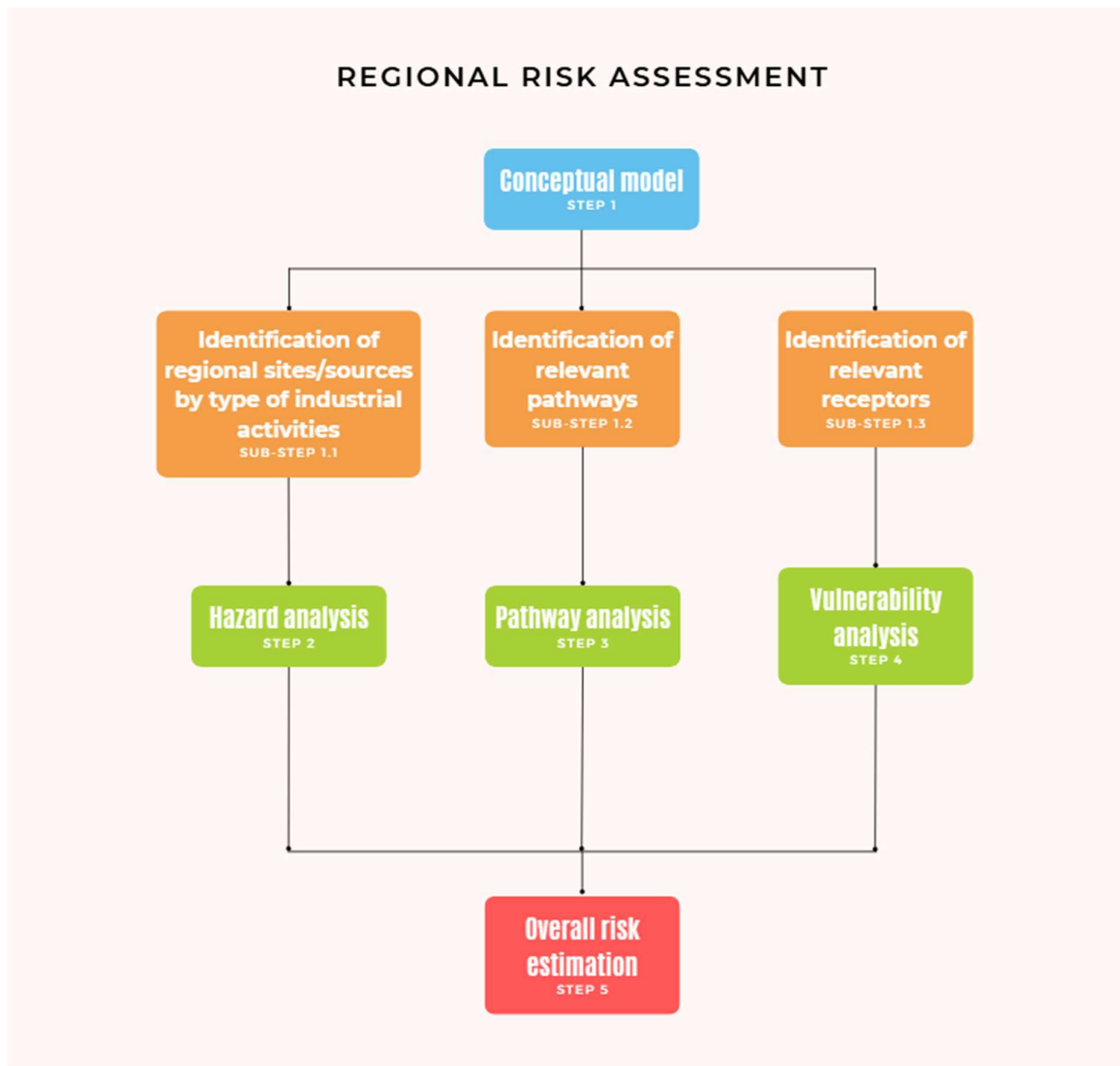


Figure 1. Structure of the Regional Risk assessment methodology for the prioritisation of contaminated sites and related assessment steps.

1.1.1. Definition of the conceptual model

The first step involves defining the conceptual model, by identifying:

1. the contamination sources (sub-step 1.1),
2. the main release mechanisms and transport pathways (sub-step 1.2), and
3. the potential receptors which may be in contact with the released contaminants (sub-step 1.3).

1.1.1.1. Identification of contamination sources

Contamination sources are identified based on the type and location of anthropogenic activities with known or suspected polluting potential. These include active and historical industrial sites, mining areas, waste disposal facilities, and sites affected by environmental incidents. The identification process draws on multiple data sources, including land use records, industrial permits, accident reports, and regulatory inventories such as activities listed in Annex I of Directive 2010/75/EU, Directive 2012/18/EU, and Annex III of Directive 2004/35/CE.

The methodology considers both ongoing and past activities, recognising that legacy pollution can result in current contamination. Consistent with international guidance (ASTM, 2021; USEPA, 2020; UNEP, 2019; ISO, 2015), all sites with potential contamination risk are included, even if direct measurements are not yet available.

1.1.1.2. Identification of Transport Pathways

Once the sources of contamination have been identified, the next step involves the environmental pathways through which contaminants may travel from the source to human or ecological receptors. These include leaching into groundwater, runoff into surface water, volatilization into the air, and direct contact with contaminated soil. Contaminant mobility is influenced by soil characteristics such as porosity and texture, which determine whether contaminants are retained, degraded, or mobilized. Figure 2 illustrates the main transport mechanisms considered in the methodology.

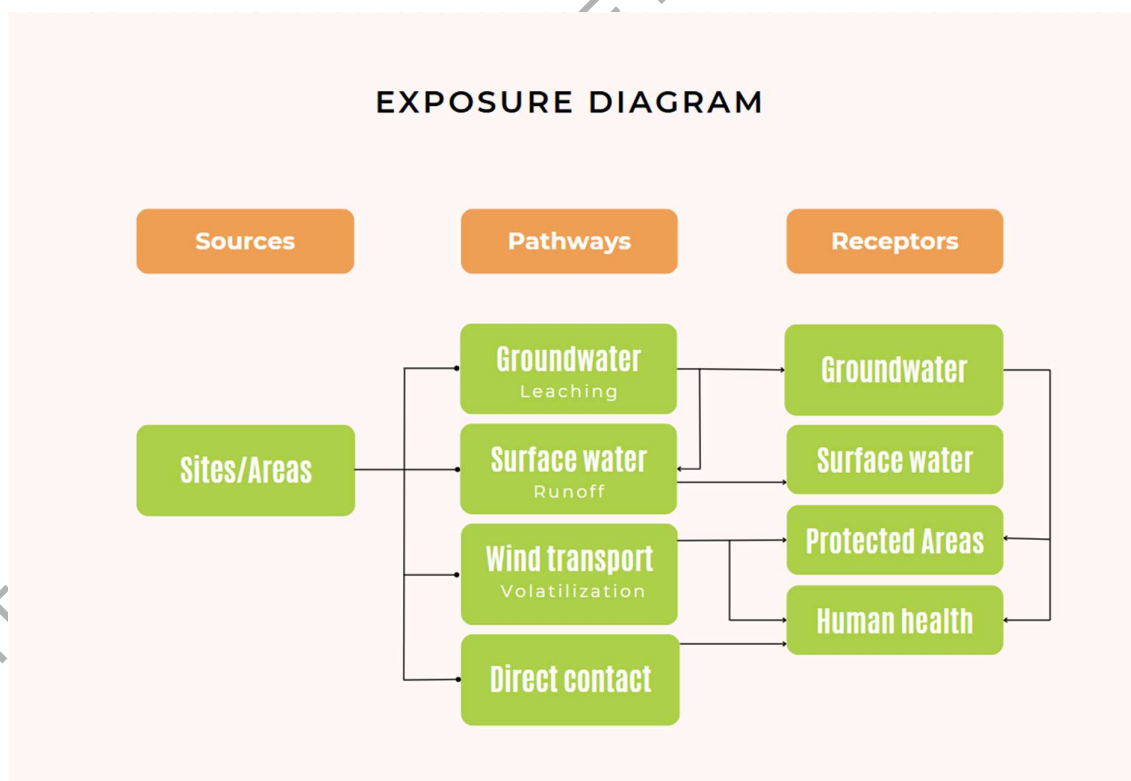


Figure 2. Exposure diagram for regional risk assessment, with the selected pathways linking the contamination sources to the relevant receptors.

1.1.1.3. Identification of Receptors

Potential receptors include:

- humans, that can be exposed to contaminants in soil through “direct contact”, i.e., ingestion (inadvertent eating), dermal exposure (skin contact) or inhalation (breathing);
- groundwater bodies;
- surface waters;
- protected natural areas, where sensitive ecological receptors are expected to be present.

Groundwater and surface water are treated both as receptors and as transport media. Ecological receptors are integrated during the vulnerability analysis phase but are not tracked through complex contaminant transport chains (e.g., groundwater to lake systems), in order to maintain manageability at a large scale.

1.1.2. GIS-based reclassification of parameters and attributes

For both pathway parameters and receptor attributes, spatial datasets need to be converted into raster format and reclassified to reflect their contribution to hazard propagation or vulnerability assessment. The reclassification assigns each class or value range a score representing its relative influence. This process ensures consistency across heterogeneous datasets by normalizing them to a common scoring framework.

Scores attributed to each parameter range, listed in Tables 1-5, are taken from Zabeo et al. (2011) and Minolfi et al. (2018), except for the flow accumulation parameter. In the GIS environment, the reclassification is implemented using attribute-based lookup tables or conditional mapping functions. Categorical datasets are assigned scores directly based on class codes. The resulting scored rasters serve as standardized inputs for subsequent spatial analyses, including pathway modelling (paragraph 1.1.4.) and receptor vulnerability estimation (paragraph 1.1.5.). By harmonizing all spatial inputs in this manner, the methodology enables the integration of environmental and socio-economic factors into a unified GIS-based risk assessment framework.

Flow accumulation represents the number of upstream cells that contribute runoff to a given cell, and it is commonly used in hydrological modeling to identify drainage networks, delineate watersheds, and estimate the intensity of surface runoff. In typical applications, higher flow accumulation values indicate areas where water from larger upstream catchments converge, thus highlighting preferential flow paths. However, in the case of this methodology, since the risk must be calculated for individual PCS, the outcome does not produce a cumulative pattern, but rather a single downstream path starting from each source point, without actual “accumulation,” and therefore characterized by constant values along the path. To account for this limitation while still capturing the hydrological contribution to risk, each PCS is first assigned the mean annual precipitation value at its location, which is then converted into a raster. Normalization is directly based on precipitation values: the maximum precipitation within the study area is assigned a score of 1, and the minimum a score of 0,2.

PATHWAY		Parameter	Value-Range	Score	
	Groundwater	Land Use	Agricultural areas/livestock		1
			Forest		0,8
			Park/school, Green urban areas and sport and leisure facilities		0,6
			Residential Urban fabric		0,4
		Industrial/commercial		0,2	
		Slope	<=2%		1
	2% - 5%			0,6	
	5% - 8%			0,4	
	>8%			0,2	
Infiltration Capacity	Very favorable		1		
	Favorable		0,5		
	Not very favorable		0,2		
	Not favorable		0		
Surface Water	Flow Accumulation	Max precipitation (mm/y)		1	
		Min precipitation (mm/y)		0,2	
Air	Aridity Index (Particulate mobility)	< 25		1	
		25-42		0,7	
		42-74		0,3	
		>74		0	

Table 1. Pathway parameters and scores.

Attribute	Value-Range	Score
GW Use	Potable - Municipal use	1
	Private well (use not specified)	0,9
	Irrigation (food vegetables)	0,85
	Other use (no food vegetables, parks...)	0,55
	Industrial	0,3
	Non Used	0,1
	No GW body	0
GW Quality	The main GW horizon does not exist/lack of information	1
	Water quality good and stable, does not require conditioning	1
	Water quality good, but can be unstable because of insufficient isolation layer, water does not require conditioning	0,75

Water quality moderate, water requires simple sanitation, (according to procedures used in standard water conditioning station in water supply systems: removal of iron, manganese, aeration), the quality is stable	0,5
Water quality moderate, water requires simple sanitation (according to procedures used in standard water conditioning stations in water supply systems: removal of iron, manganese, aeration), the quality is potentially unstable because of insufficient isolation layer	0,25
Water quality bad, water requires complicated sanitation (according to procedures beyond those used in the standard station of water conditioning in the water supply system)	0

Table 2. Vulnerability attributes and scores for Groundwater receptor.

Attribute	Value-Range	Score
SW Use	Potable - Municipal use	1
	Private well (use not specified)	0,9
	Irrigation (food vegetables)	0,85
	Other use (no food vegetables, parks...)	0,55
	Industrial	0,3
	Non Used	0,1
SW Quality	Point on the stream, water body or its part, where it was determined that the water meets drinking water standards, food industry needs or other industries requiring drinking water quality and the quality for living conditions of Salmonidae species	1
	Point on the stream, water body or its part where it was determined that the water is usable for living conditions for farming fish other than Salmonidae species, husbandry use, bathing activities and recreational activities.	0,6
	Point on the stream, water body or its part, where it was determined, that the water is usable for industry purposes except for the industries requiring drinking water quality, watering purposes in agriculture and horticulture purposes.	0,3
	Point on the stream, water body or its part, where it was determined that the water is not usable in any of the use categories because of its above limits contamination.	0
SW Order (Strahler classification)	1	1
	2	0,8
	3	0,6
	4	0,5
	5	0,3
	>=6	0,1

Table 3. Vulnerability attributes and scores for Surface Water receptor.

Attribute	Value-Range	Score
Land Use	Agricultural areas/livestock	1
	Forest	0,8
	Park/school, Green urban areas and sport and leisure facilities	0,8
	Residential Urban fabric	0,4
	Industrial/commercial	0,2
Population Density	High: ≥ 2000 people/km ²	1
	Medium: 1000 people/km ²	0,6
	Low: 250 people/km ²	0,2
Percentage of vulnerable groups	High: >40	1
	Medium-High: 32-40	0,7
	Medium: 22-31	0,5
	Low: <22	0,3

Table 4. Vulnerability attributes and scores for Human Health receptor.

Attribute	Value-Range	Score
Type of protection	Protected areas at state level	1
	Protected areas at regional level	0,8
	Sites identified by Natura 2000, database not included in the first two classes	0,6
	Ecological area protected at local level	0,5
	Biogenic natural monuments protected at local level	0,4
	Abiogenic natural monuments protected at local level	0,3
	Cultural heritage monuments protected at local level	0,2

Table 5. Vulnerability attributes and scores for Protected areas receptor.

1.1.3. Hazard analysis

The hazard analysis quantifies the potential hazard posed by each identified site, based on the characteristics of anthropogenic activities historically or currently performed at the location. Where multiple activities are present on a single site, their hazards are estimated individually and then aggregated. The ISLANDR approach aims to bridge the gap between desk-based inventories and site-specific investigations, thus it provides a consistent, information-based and spatially explicit tool for large-scale screening of potentially contaminating activities, supporting consistent prioritization across regions. In this sense, the methodology complements existing national frameworks by offering a harmonized approach for large-scale comparison and prioritization across regions and Member States.

The hazard score for each potential point source S is composed of the following aspects:

- the intrinsic hazard of the activity H_n , based on the substances potentially produced and released into the environment;
- a dimension A_s , expressed as the area of the site or estimated waste volumes;
- a duration in time T_a , defined as the number of years the activity was operational; we assumed a maximum time window of 100 years (1924 - 2024) following the reference methodologies.

A key innovation of the methodology lies in the estimation of intrinsic hazard for each activity. We used a database developed by BRGM (2018) that correlates 345 NAF/NACE-coded activities (Nomenclature of Economic Activities, the European statistical classification of economic activities) with 3501 hazardous substances, exploiting for each activity and substances an emission probability. The PRA.MS procedure (EEA, 2005) provides hazard levels to sources in terms of toxicity of contaminants (risk phrases); each contaminant can have one or more hazards according to the pathway/route of exposure typical of each chemical. NAF codes from the French database were matched to the corresponding NACE codes, while the CLP risk phrases were associated to the new Globally Harmonized System (GHS) hazard classes, ensuring compliance with current EU hazard classification. The intrinsic hazard of each activity was then calculated as the sum of the products of emission probability p_u and PRA.MS hazard score h_u across all substances linked to that activity. This procedure yields four pathway-specific hazard values for each activity.

However, the PRA.MS methodology does not provide hazard scores for several risk phrases, mainly those related to physical hazards (e.g., flammability or explosivity) and some specific ecotoxic effects (e.g., toxicity to soil organisms, flora, fauna, or bees). This exclusion reflects the original scope of PRA.MS, which focused on human health and general ecological endpoints, but it may result in activities being assigned zero hazard despite their potential environmental relevance.

To avoid this underestimation, a conservative fallback mapping was defined (Table 6). Physical hazards (e.g., R1–R19, R30, R44, R47) were assigned the minimum non-zero score (10), acknowledging their indirect contribution to contamination through accidental release. Aquatic toxicity classes (R50–R53) were aligned with CLP/GHS categories (EC No 1272/2008; ECHA, 2017), following a progressive scale (10–30). Specific ecotoxic phrases (R54–R56, R57) and ozone-depleting effects (R58–R59) were assigned intermediate scores (20), reflecting their ecological significance in line with ISO (2018) and recent guidance on soil and ecosystem protection (EFSA, 2013).

This precautionary approach ensures that activities linked exclusively to such risk phrases are not incorrectly classified as having zero intrinsic hazard, while remaining consistent with the PRA.MS scoring framework.

Risk phrase	CLP/H equivalent (indicative)	Main pathway(s)	Proposed fallback score	Rationale
R1–R6 (Explosive risk)	H200–H205	Direct contact / Air	10	Physical hazard (explosivity), may cause accidental releases.
R7–R9 (Fire/explosion risk)	H260–H261	Direct contact / Air	10	Reactive with water/air; precautionary score.
R10 (Flammable)	H226	Air / Direct contact	10	Fire hazard; possible release to environment.
R11 (Highly flammable)	H225	Air / Direct contact	10	Same as above, higher flammability.
R12 (Extremely flammable)	H220/H222	Air	10	Strong fire hazard, indirect environmental risk.
R13–R19 (Gas under pressure, reacts with water, forms explosive mixtures)	H280, H290, H271	Air / Direct contact	10	Physical-chemical hazards; low but non-zero environmental score.
R30, R44, R47 (May become highly flammable in use, risk of explosion if heated)	H280–H290	Air / Direct contact	10	Conditional hazards; low precautionary score.
R50–R52 (Very toxic / toxic / harmful to aquatic organisms)	H400, H401, H402	Surface water / Groundwater	30 (R50), 20 (R51), 10 (R52)	Classical ecotoxicity scale (acute aquatic toxicity).
R53 (May cause long-term adverse effects in aquatic environment)	H412/H413	Surface water / Groundwater	20	Chronic aquatic toxicity, precautionary intermediate score.
R54–R56 (Toxic to flora / fauna / soil organisms)	H410/H420	Surface water	20	Ecotoxic endpoints not covered; precautionary assignment.
R57–R59 (Toxic to bees / may cause ozone depletion)	H420, H420/H420A	Air	20	Specific ecotoxic effects; precautionary mid score.

Table 6. Fallback mapping for risk phrases.

The activity-specific intrinsic hazard H_n is calculated by aggregating the probability and hazard of all the substances associated with it (Equation 1):

$$H_n = \sum_{u \in n} p_u \cdot h_u \quad \text{Equation 1}$$

Where:

- H_n is the intrinsic hazard of a specific activity n among the 345 identified in the BRGM matrix;
- p_u is the emission probability of substance u belonging to the set of substances potentially released by the anthropogenic activity n ;
- h_u is the hazard score of substance u potentially released by the anthropogenic activity n which is estimated from the PRA.MS procedure (EEA, 2005b).

If in a potentially contaminated site, more than a single activity n has occurred over time, only the highest H_n is considered for the further calculations.

The overall hazard score H_s for a potentially contaminated site is then computed by weighing each activity's intrinsic hazard by its duration and area:

$$H_s = \sum_{a \in s} H_a \cdot T_a \cdot A_s \quad \text{Equation 2}$$

Where:

- H_s is the source hazard score;
- H_a is the max H_n ;
- s is the Source;
- T_a is the time duration of activity a in source s ;
- A_s is the area of the activity a in source s .

To preserve the site-specific information of each PCS while obtaining a continuous spatial representation in GIS, the Euclidean distance function was applied. This tool generates a raster in which each cell represents the distance from the PCS point. The resulting raster was then multiplied by the hazard score (H_s) of the site as a constant. The calculation was restricted to a maximum radius of 5 km, beyond which the influence of the contamination is assumed to be negligible. This threshold aligns with empirical findings showing that exposure estimates best correlate with pollutant levels when considering sources within approximately 5 km (Tenailleau et al., 2024). The subsequent steps of the RRA process were then executed in GIS through an automated loop for each PCS.

1.1.4. Pathway analysis

The movement of contaminants from point sources to receptors involves complex environmental processes, including leaching, runoff, volatilization, and atmospheric dispersion (Figure 3, Davis et al., 1993). In this methodology, we adopt a simplified, yet operational modelling approach tailored to large-scale applications, incorporating the key environmental pathways relevant across diverse land uses, including industrial, agricultural, and urban areas.

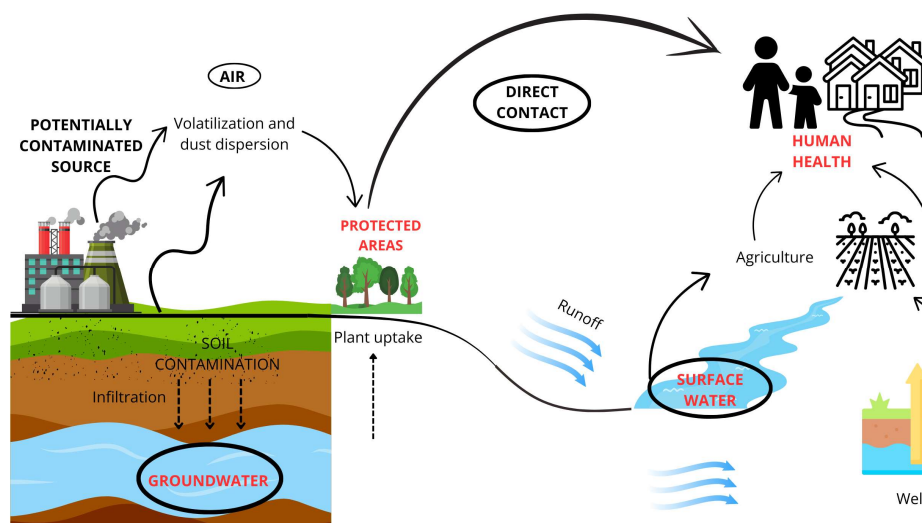


Figure 3. Schematic representation of transport processes from a potentially contaminated site

The behavior of contaminants through these pathways is partially captured in the intrinsic hazard scores assigned to each source (paragraph 1.1.3.), which differentiate potential impacts across four transport media: groundwater, surface water, air, and direct soil contact.

1.1.4.1. Groundwater Pathway

This pathway models contaminant transport via vertical leaching into the saturated zone. Contaminants may percolate through the unsaturated soil layer and reach the aquifer, which can further act as a secondary transport medium. Water and solutes could also flow through the saturated zone until eventually reaching their discharge point, such as a spring (Lerner et al., 2009).

The infiltration process is influenced by soil properties, land cover, terrain slope, and precipitation patterns (USEPA, 2013, Beck et al., 2017, Sun et al., 2018). Areas with high permeability, low slope, and permeable land uses (e.g. vegetated surfaces) are more prone to infiltration than urban or compacted soils. Conversely, steep slopes and low-permeability soils favor surface runoff.

The groundwater pathway score for each raster cell is calculated as a function of three key spatial parameters:

- Slope (derived from a Digital Elevation Model – DEM)
- Infiltration capacity Index
- Land use

These parameters are combined according to Equation 3:

$$Pg = S \cdot IC \cdot L \quad \text{Equation 3}$$

Where

- Pg is the groundwater pathway score
- S is the slope map calculated from the DEM
- IC is the infiltration capacity index map

- L is the land use score map.

1.1.4.2. Surface Water Pathway

Surface water contamination occurs through runoff and mobilization of contaminants during rainfall events, especially from impermeable surfaces or areas with low infiltration capacity (Braud et al., 2017). A flow accumulation map is derived from the terrain model (DEM) to simulate overland flow routes and identify zones where water (and associated contaminants) are likely to concentrate.

GIS tools are used to:

1. Compute flow direction and flow accumulation from the DEM.
2. Weight the flow accumulation with mean annual precipitation to account for hydrological intensity.
3. Initiate flow accumulation from PCS locations, to simulate contaminant transport from each source towards water bodies.

Equation 6 defines the surface water pathway score:

$$Ps_s = Fa_s$$

Equation 4

Where:

- Ps_s is the map of surface water pathway for source s
- Fa_s is the flow accumulation map for source s .

1.1.4.3. Air Pathway

Contaminants may be transported through the atmosphere via two mechanisms:

- Volatilisation of gaseous or semi-volatile compounds
- Particle-bound transport, where contaminants adhere to airborne dust or aerosols

This pathway is particularly relevant for substances with low solubility and high environmental persistence (Csavina et al., 2011; Prabhakar et al., 2014). Atmospheric transport is often long-range and influenced by climatic and meteorological conditions (Dragović et al., 2014; Chen et al., 2022).

We model this pathway using:

- Air hazard classes, derived from intrinsic hazard scores
- The De Martonne Aridity Index, which integrates annual temperature and precipitation, to refine dispersion potential (Minolfi et al., 2018; Punia, 2021)

The De Martonne Aridity Index (AI) is evaluated in each point using meteorological data (Minolfi et al., 2018). Other climatic factors such as temperature and precipitation, in fact, significantly influence the paths of contaminants' dispersal (Punia, 2021). The AI is estimated by the following formula:

$$AI = \frac{P}{T + 10}$$

Equation 5

Where:

- AI is the De Martonne Aridity Index;

- P is the mean annual precipitation map (mm/yr);
- T is the mean annual temperature map (°C).

1.1.4.4. Direct contact

This pathway represents the potential for human exposure through direct interaction with contaminated soil, particularly relevant in urban or agricultural areas where soil access is common. No intermediate transport process is considered; risk is determined solely by the spatial overlap between potential sources and human receptors.

Hazard propagation is modelled using Inverse Distance Weighting (IDW) from the contamination source, with no additional environmental modifiers (e.g., slope, land use).

1.1.5. Vulnerability analysis

Receptor vulnerability to contamination from potential point sources is defined as the susceptibility of a receptor to an imposed contaminant load, determined by its intrinsic characteristics (Pizzol et al., 2011). The methodology used to estimate the *receptor vulnerability score* follows the multi-step procedure described in Zabeo et al. (2011) and consists of five steps:

1. Identification of regional receptors;
2. Selection of attributes relevant for assessing receptor sensitivity to contamination;
3. Spatial assignment of attribute values to receptors using GIS analysis;
4. Normalization of attribute values to a common scale;

Steps 3 and 4 are well described in paragraph 1.1.2.

The vulnerability score V_r for a receptor type r is calculated by integrating its attribute through a receptor-specific vulnerability function F_v applied to each raster cell representing the receptor (Equation 6):

$$V_r = F_v(C_r)$$

Equation 6

Where:

- V_r is the receptor's vulnerability score;
- F_v is the vulnerability function;
- C_r are the selected receptors' attributes

The analysis considers four receptor categories:

- Groundwater
- Surface water
- Human health
- Protected areas

The selected attributes for each receptor are listed in Table 7.

Receptor	Attributes
Groundwater	Groundwater use; groundwater quality
Surface water	Surface water order; surface water quality; surface water use
Human health	Land use; population density; percentage of vulnerable population groups
Protected areas	Type of protection

Table 7. Selected characteristics for receptor's vulnerability evaluation.

1.1.5.1. Groundwater Vulnerability

Groundwater vulnerability is influenced by proximity to point sources and its role in potable water supply. The score V_g is calculated from groundwater quality and use parameters (Equation 7).

$$V_g = Q \cdot U \quad \text{Equation 7}$$

Where:

- V_g is the groundwater vulnerability score;
- Q is the groundwater quality parameter;
- U is the groundwater use parameter.

Areas where groundwater is used for drinking water supply or irrigation receive higher scores due to direct human and agricultural exposure risks.

1.1.5.2. Surface Water Vulnerability

Surface water vulnerability reflects its ecological and human use value and its hydrological characteristics. The score V_s is calculated as:

$$V_s = O \cdot Q \cdot U \quad \text{Equation 8}$$

Where:

- V_s is the surface water vulnerability score;
- O is the surface water order parameter;
- Q is the surface water quality parameter;
- U is the surface water use parameter.

Watercourses used for drinking water supply, irrigation, aquaculture, or recreational purposes receive higher scores due to the increased potential for human and ecological exposure. The surface water order according the Strahler classification indicates the hierarchical position of a stream within the drainage network. Lower-order streams receive higher vulnerability scores, as contaminant concentrations are generally higher and less diluted compared to those in higher-order streams.

1.1.5.3. Human Health Vulnerability

Human health vulnerability is determined by the land use, the population density and the percentage of vulnerable groups (e.g. children, elderly groups) according the following Equation 9:

$$Vh = L \cdot PD \cdot VG$$

Equation 9

Where

- Vh is human health vulnerability;
- L is the land use parameter
- PD is the population density parameter map;
- VG is the percentage of vulnerable group parameter map.

1.1.5.4. Protected Areas Vulnerability

Protected areas are scored according to their legal protection status, reflecting ecological sensitivity and conservation value. The score Vp is calculated as:

$$Vp = PAT$$

Equation 10

Where:

- Vp is the protected areas vulnerability;
- PAT is the protected area type parameter.

1.1.6. Relative Regional Risk estimation

The spatial source risk represents the combined effect of the hazard associated with each potential point source, the transport through environmental pathways, and the vulnerability of receptors. This step integrates the results from paragraph 1.1.3 – 1.1.5. into a single spatially explicit risk score.

Each potentially contaminated site (PCS) is first evaluated independently. For each source, the spatial source risk is computed by combining:

- The intrinsic hazard map (from paragraph 1.1.3.),
- The relevant environmental pathway score maps (from paragraph 1.1.4.), and
- The receptor vulnerability score maps (from paragraph 1.1.5.).

The integration follows Equation 11:

$$SRs = H \cdot P \cdot V$$

Equation 11

Where:

- SRs is the total risk for each source.
- H is the hazard score spatially propagated through all pathways
- P is the pathway score except for direct contact (which uses the point-source hazard map directly);
- V is the vulnerability score for the receptor under consideration.

For classification and prioritization purposes, these individual spatially distributed values need to be aggregated to a single representative value for each source. At first, risk values are obtained for each pathway (groundwater, surface water, direct contact, and air), calculating both maximum and mean risk values. The final total risk score for each PCS is calculated as the sum of SRs mean risk values, as they provide a more balanced representation of the overall exposure potential across the affected area, thus reporting the max risk values to capture localized hotspots.

The regional comparison of point sources does not aim to merge or replace existing site-specific conceptual models, but rather to provide a spatially consistent overview of potential risks across the study area. This screening-level approach enables regional prioritization, identifying sites where site

specific investigations would be most beneficial, without requiring data sharing from individual site or access to confidential investigation reports.

As such, it complements existing national contaminated site management frameworks by offering a harmonized, regional-scale layer of analysis that can support planning, resource allocation, and policy development.

1.1.7. Model Builder in GIS

The entire methodology was developed and implemented using the Model Builder tool in ArcGIS. Model Builder is a visual programming environment that allows the creation of workflows by combining geoprocessing tools, datasets, and logical operators in a graphical interface. This approach ensures transparency and reproducibility of the analysis, as each processing step is explicitly represented and can be inspected.

One of the main advantages of using Model Builder is its flexibility. The workflow designed for this study can be easily replicated in other geographical contexts: by simply replacing the input datasets (e.g., land cover, meteorological data, population density), the model automatically updates the subsequent analyses without requiring substantial modifications to the workflow.

This makes the methodology adaptable, transferable, and accessible also to users who may not have advanced programming skills, but who are familiar with the principles of GIS-based geoprocessing.

2. Regional Risk Assessment of diffuse contamination

Diffuse contamination can be considered as a blanket contamination resulting from large-scale atmospheric transport and mixing processes that contribute uniformly across extensive areas (Flem et al., 2022; Reimann et al., 2019). Unlike point source, diffuse contamination originates from multiple and heterogeneous anthropogenic activities rather than a single, discrete emission source. These activities release significant quantities of organic and inorganic compounds that can persist in soils for long periods before being transported and degraded in surface and groundwater systems. In addition to atmospheric deposition or application of agrochemicals, diffuse contamination can be linked to urban soil management.

The main causes of diffuse contamination are:

- Atmospheric deposition, covering wide areas which and difficult to delimit, often linked to vehicular traffic, highly urbanized areas, waste treatment plants, and energy production;
- Dispersed sources, such as fertilizer and pesticide application in agriculture; livestock farming and slurry management; sewage sludge disposal; shipping and railway activities; exceptional events like fires and flooding;
- Small-scale or artisanal activities, often historical, for which both source and cause data are usually unavailable (Van Gestel et al., 2022);
- Historical land use (e.g. reuse of excavated soil): elevated concentrations of metals in fillings must be taken into account in the re-use of soil and in land use planning (Tarvainen et al., 2023, Tarvainen et al., 2024).

In most countries, diffuse contamination of agricultural and urban soil is managed as a separate domain from site-specific contamination (Bardos et al., 2019). Urban soils retain a “long memory”: while local

geology exerts the primary control on chemical composition, this baseline is strongly by a long history of industrial activities, atmospheric deposition, reconstruction and backfilling with heterogenous materials of uncertain origin and composition (Belbèze et al., 2023).

Heavy metals illustrate the complexity of this issue, as they can derive from both geogenic and anthropogenic sources and may naturally occur at elevated levels in soils and groundwaters. The main difficulty in assessing diffuse contamination lies in the distinguishing the natural background variability from the anthropogenic inputs, which can extend over distances from a few meters (e.g. roadside pollution) to more than 100-200 km (e.g. emission from major industrial emitters). According to ISO (2018), background concentration in a soil can be defined as the natural geogenic variation without anthropogenic influences, or alternatively as the concentration of a substance in soils characteristic of an area, reflecting both natural sources and diffuse anthropogenic inputs such as atmospheric deposition. In some approaches, the second definition is known as geochemical baseline concentration. In the latter case, this is also referred to as the geochemical baseline concentration (Belbèze et al., 2023). Several methodologies have been developed to separate geogenic from anthropogenic contributions, including Compositional Data Analysis (Petrik et al., 2018, Sahoo et al., 2020), indices such as the Pollution Index or Enrichment Factor (Konadu et al., 2023; Gan et al., 2023), or long-term monitoring of atmospheric deposition through terrestrial mosses (Flem et al., 2022). Building on the work of Fabian et al. (2017), cumulative distribution function of contaminants in top and sub-soils have also been used at regional to continental scale to characterize diffuse contamination patterns.

Given the wide range of sources and substances, a detailed qualitative assessment of each potential source of diffuse contamination is not feasible at regional scale. Instead, a risk assessment of diffuse contamination begins with available soil contamination data, which are treated as hazard indicators. What we can consider is that, once contaminants are released into the environment and soil, they can migrate further. The assessment typically focuses on a limited number of key contaminants relevant for ecosystems and human health, such as heavy metals, polyaromatic hydrocarbons (PAHs), polychlorinated biphenyls (PCBs), mineral oils, and selected pesticides (e.g. hexachlorobenzene, HCB).

Risk areas could be identified based on soil quality studies and threshold levels set for metals that pose risks to the environment and human health (Solismaa et al., 2025). Given the complexity of distinguishing geogenic from anthropogenic contributions and the lack of harmonized European thresholds for soil quality, diffuse contamination requires an adapted risk assessment approach. Based on this, land use can be planned in a safe and controlled manner.

Building on the regional framework originally developed for point sources, the methodology can be extended to observed soil concentration data, using national threshold values (e.g., Italian regulatory limits) as reference proxies. This ensures comparability at larger scales while allowing the prioritization of areas most affected by diffuse contamination, which will be further detailed in the following methodological section.

2.1. Methodology for Regional Risk Assessment of diffuse contamination

The RRA framework applied to point source contamination was adapted to address diffuse contamination, maintaining the same overall structure based on hazard, pathways, and receptor analysis in a GIS environment. The main modification concerns the definition of the hazard component, while the pathways, receptors, and ranking procedure remain consistent with the previous approach.

2.1.1. Hazard definition

In the original framework developed for point source contamination, the hazard component was derived from land-use records and historical archives to identifying potentially contaminated sites and assign them a relative hazard score. In the case of diffuse contamination, the hazard is instead defined using measured concentrations of contaminants in soils. In this case, hazard evaluation does not rely on historical land-use or industrial records, but rather on observed concentration data interpolated into continuous surfaces. Interpolation data were obtained in Deliverable 1.4 of ISLANDR Project (BRGM).

For each contaminant u , hazard evaluation is performed by comparing its spatial concentration (C_u) against concentration thresholds (T_u) defined for different use categories. Three threshold levels were considered, representing the maximum acceptable concentration in residential, agricultural and industrial land uses. Thresholds can be derived from national legislation (e.g., Italian soil quality standards), and in the absence of harmonized European values, these may serve as suitable proxies for a broader application. The spatial assignment of thresholds was based on the Corine Land Cover (CLC) classification, allowing each raster cell to be evaluated against the threshold appropriate for its land-use category.

For each contaminant, a threshold function $Ft(C_u, T_u)$ is then applied according Equation 12 to reclassify the interpolated concentration maps (D 1.4). Cells with concentrations below the threshold are assigned a hazard value of zero, while concentrations above the threshold are normalized by dividing the concentration values by the corresponding threshold concentration (CSC), in order to obtain a dimensionless hazard index to enable comparison among different contaminants in subsequent summation steps. This approach ensures that only areas with concentrations exceeding acceptable values contribute to the risk calculation, while also making the outputs of different substances directly comparable in multi-contaminant analysis.

$$\widehat{C}_u = \left\{ \widehat{x}_i = \begin{cases} 0, & \text{if } x_i < T_u \\ \frac{x_i}{T_u}, & \text{if } x_i \geq T_u \end{cases}, \forall x_i \in C_u \right\} \quad \text{Equation 12}$$

Where:

– \widehat{C}_u is the normalized raster C_u given the threshold T_u , – x_i are the raster cells of C_u

2.1.2. Pathway analysis for diffuse contamination

In the case of diffuse contamination, the pathway analysis follows the same conceptual model adopted for PCSs. However, key adaptations are required due to the nature of the source: instead of discrete point sources, each contaminant's spatial concentration map described in the previous section, represents a spatially continuous hazard layer used as the starting point for all the pathway calculations.

The same environmental pathways considered for PCSs are applied:

- Leaching to groundwater,
- Runoff into surface waters,
- Air dispersion of particulates (excluding volatilization),
- Direct contact with contaminated soils.

As explained in paragraph 1.1.2., all pathway parameters and receptor attributes need to be converted into raster format and reclassified to reflect their contribution to hazard propagation or vulnerability assessment. The reclassification assigns each class or value range a score representing its relative influence.

2.1.2.1. Groundwater Pathway

As in the PCS-based methodology, the groundwater pathway simulates the potential downward migration of contaminants from the unsaturated zone to groundwater (Lerner et al., 2009). The driving factors remain slope, land use and infiltration capacity. The final groundwater pathway score is thus expressed as:

$$Pg = S \cdot IC \cdot L \quad \text{Equation 13}$$

Where

- Pg is the groundwater pathway score
- S is the slope map calculated from the *DEM*
- IC is the infiltration capacity index map
- L is the land use map

For further details see paragraph 1.1.4.1.

2.1.2.2. Surface Water Pathway

For diffuse contamination, overland flow processes are represented through flow accumulation starting from every raster cell. Since the hazard is already spatially distributed, and precipitation directly influences contaminant mobilization and transport (Braud et al., 2017), the pathway is simulated by computing flow accumulation using the product of the hazard raster and mean annual precipitation as a weight raster.

Considering that the cumulative nature of flow accumulation would pose an artefactual dominance of the SW pathway risk in the final total risk estimation, a normalization must be applied by dividing all cell values by the maximum raster value, reducing the effect of extreme peaks. This adjustment preserves the relative spatial pattern of SW risk along the drainage network, while aligning its overall magnitude with the other pathways for balanced aggregation.

$$PS_s = \frac{Fa_s(H \times P, Fd)}{\max Fa_s(H \times P, Fd)} \quad \text{Equation 14}$$

Where:

- PS_s is the map of surface water pathway for contaminant s
- Fa_s is the flow accumulation map weighted with contaminant hazard and precipitation maps.
- P is precipitation
- Fd is Flow direction (*DEM*)

2.1.2.3. Air Pathway

In contrast to PCS-based modelling, where atmospheric dispersion was limited in 5km buffer from individual sources, in diffuse contamination the continuous hazard layer across the study area is

maintained and its dispersion potential is modulated by the De Martonne Aridity Index, which expresses susceptibility to wind-driven particle transport under dry climatic conditions (Minolfi et al., 2018; Punia, 2021).

$$P_{A_s} = AI \quad \text{Equation 15}$$

Where:

- P_{A_s} is the map of Air pathway
- AI is Aridity index

For further details see paragraph 1.1.4.3.

2.1.2.4. Direct Contact

As for PCS-based assessments, this pathway accounts for exposure through direct soil contact. Since hazard is already spatially explicit, no decay or dispersion is modelled. The direct contact risk is obtained by overlaying the hazard raster with receptor layers (e.g., residential or agricultural areas).

2.1.3. Vulnerability analysis for diffuse contamination

Receptor vulnerability in the context of diffuse contamination represents the intrinsic susceptibility of a receptor to being adversely affected by contaminant exposure, depending on its environmental, ecological, or socio-economic characteristics (Pizzol et al., 2011).

The vulnerability score is again estimated following the multi-step procedure described by Zabeo et al. (2011), consisting of the following phases:

1. Identification of regional receptors;
2. Selection of attributes relevant for receptor sensitivity;
3. Spatial assignment of attribute values through GIS analysis;
4. Normalization of attribute values on a common scale;
5. Aggregation of attributes per receptor using receptor-specific vulnerability functions (Fv).

Each steps follow the same procedure described in paragraph 1.1.5. for the PCS-based assessment.

As in the previous case, four receptor categories are considered:

- Groundwater bodies,
- Surface waters,
- Human health,
- Protected natural areas.

Once vulnerability rasters are computed for each receptor category, they are subsequently combined with pathway-adjusted hazard rasters to produce the final risk maps (paragraph 2.1.4.).

2.1.3.1. Groundwater Vulnerability

Groundwater vulnerability is evaluated based on groundwater use (e.g. drinking water abstraction, irrigation) and aquifer typology or depth, reflecting the likelihood of exposure to diffuse contaminants.

Areas supplying drinking water are assigned higher vulnerability scores due to direct implications for human health.

2.1.3.2. Surface Water Vulnerability

Surface water vulnerability depends on ecological relevance, anthropic use, and hydrological characteristics. Water bodies used for drinking water supply, irrigation, aquaculture, or recreational purposes are assigned higher vulnerability scores. Furthermore, consistent with the Horton-Strahler stream order concept, lower-order streams are considered more vulnerable, as they are generally more sensitive to pollutant loads and subject to lower dilution capacity compared to higher-order rivers.

To evaluate the Surface water vulnerability, only the intersections between the Surface water pathway and the water bodies are considered.

2.1.3.3. Human Health Vulnerability

Human health vulnerability is derived from a combination of land use, population density, and the proportion of vulnerable groups, such as children (0–14 years) and elderly populations (>75 years). Areas characterized by residential or agricultural use and high population density are assigned higher vulnerability.

2.1.3.4. Protected Areas Vulnerability

Protected areas are assigned vulnerability scores based on their legal protection status and conservation value. National and regional nature reserves, ecological areas, and other environmentally sensitive zones receive higher scores.

2.1.4. Total risk per contaminant

The risk estimation for diffuse contamination follows the same conceptual structure as for PCS-based risk assessment, integrating hazard, pathway behaviour, and receptor vulnerability within a spatially explicit Source–Pathway–Receptor (S–P–R) framework. However, in contrast to PCS analysis, where risk is produced at the source level and later aggregated to a representative score per site, diffuse contamination generates fully spatial risk outcomes that must be preserved across the entire study area. In this case, the risk is computed directly at the raster cell level, without any need for spatial aggregation to a single value. Instead of assigning intrinsic hazard values to point sources, the interpolated concentration-based hazard raster H_u is multiplied by the normalized pathway modifier (derived from slope, land cover and infiltration index). For each contaminant u , a pathway- and receptor-specific risk raster is computed as:

$$R_{u,x,y} = H_u \times P_x \times V_y \quad \text{Equation 15}$$

Where:

- H_u is the hazard map of the contaminant u
- P_x is the map of the path x and $x \in \{g, s, a, dc\}$ (g = groundwater, s = surface water, a = air, dc = direct contact)
- V_y is the map of the vulnerability of the receptor y and $y \in \{g, s, pa, hh\}$ (g = groundwater, s = surface water, pa = protected areas, hh = human health)

This produces a set of risk maps for each contaminant and for each pathway/receptor combination. To obtain a total risk raster per contaminant, risk contributions across the four pathways are summed:

$$R_u = R_{u,g,g} + R_{u,s,s} + R_{u,a,pa} + R_{u,a,hh} + R_{u,dc,hh} \quad \text{Equation 16}$$

2.1.5. Cumulative risk for diffuse contamination

Finally, assuming additive effects as a conservative screening assumption, a multi-contaminant cumulative risk map is generated by summing all contaminant-specific risk rasters:

$$R = \sum_{u \in U} R_u \quad \text{Equation 17}$$

Where:

- U is the set of all contaminants

The cumulative risk map is obtained by summing the individual risk rasters for each contaminant. This additive approach is adopted as a conservative screening assumption, consistent with international practices for multi-contaminant risk assessment (e.g., WHO, 2001; ECHA, 2016). In screening-level analyses, risks are typically assumed to be additive unless specific toxicological interaction data are available. Therefore, summation provides a precautionary estimate of overall risk potential, ensuring that areas influenced by multiple contaminants are appropriately highlighted for further investigation or land-use management considerations.

This approach enables the production of a harmonised, multi-contaminant risk map that supports spatial prioritisation of areas potentially affected by diffuse contamination. While it does not allow back-tracking contamination to individual responsible activities or sites, it provides a strategic overview useful for land-use planning, defining monitoring needs, applying preventive measures (e.g., agricultural management practices), and identifying areas where more detailed site-specific investigations may be required.

3. Regional Risk Assessment for Potential Contaminated Sites - application to ITA Toulouse Métropole

3.1. Study area

The selected study area is the Toulouse Métropole ITA, situated in the Haute-Garonne department, in the Occitanie region of southwestern France. Covering an area of approximately 465 square kilometers, the metropolitan region comprises 37 municipalities, including the city of Toulouse, which serves as its administrative and economic center. With a population exceeding 800,000, Toulouse Métropole is the fourth largest region and one of the fastest-growing urban areas in France, driven by its position as a leading hub for the aerospace industry, technology sectors, and higher education institutions, such as the University of Toulouse (INSEE, 2023).

The land use within the metropolitan area is highly diverse, combining dense urban zones, industrial areas, and peri-urban agricultural lands. Since the late 1970s, in fact, Toulouse's urban fabric expanded rapidly into rural areas, creating a vast metropolitan region (Duvernoy, 2018). This variety of land use

contributes to a complex contamination profile and introduces different contamination risks, from industrial emissions and waste disposal to agricultural runoff and urban sprawl. The presence of pollutants is often exacerbated by legacy contamination from past industrial practices, which can persist in urban soils long after the cessation of harmful activities (Pouyat et al., 2015).

The region's geological composition includes alluvial plains along the Garonne River, which pose specific risks for contaminant transport due to their high permeability. Toulouse experiences a temperate oceanic climate with Mediterranean influences, characterized by mild winters and warm summers. The region receives an average annual rainfall of approximately 600-700 mm, with precipitation patterns that could influence the leaching and migration of soil contaminants (Météo-France, 2023).

Given its industrial history, rapid urbanization, and agricultural practices, Toulouse Métropole presents a complex and dynamic environment for soil risk assessment. Studies have shown that urban agriculture, while beneficial for food security, can pose risks if soil contamination is not adequately managed (Moskal & Berthrong, 2018). The need for a multi-faceted approach to soil contamination risk assessment in Toulouse is underscored by the complex interplay of urban development, environmental health, and community well-being (Cheng et al., 2021). This makes it an ideal case study for evaluating contamination pathways, identifying vulnerable receptors, and developing targeted risk management strategies to support sustainable urban planning and public health protection.

A key consideration in interpreting the results is how representative the Toulouse case is of other European urban areas. The selected test area encompasses a heterogeneous landscape, including a major urban center, peri-urban and industrial zones, and surrounding agricultural and natural environments. This diversity mirrors the spatial and functional complexity found in many European metropolitan regions, where legacy industrial activities coexist with expanding residential and agricultural land uses. Therefore, while local characteristics may influence specific results, the overall configuration of Toulouse provides a relevant and transferable setting for testing the large-scale applicability of the proposed risk assessment framework.

3.2 Hazard Analysis

In this study, hazard scores were calculated exclusively on the basis of the intrinsic hazard of activities (H_n), which reflects the substances potentially produced and released into the environment according to their NACE classification. Information on site dimension or operational duration was not available and therefore not included in the scoring.

For this study, we used the French BASIAS database, which is the national repository for regional historical inventories of industrial sites and activities, managed by BRGM. Other databases, such as the IHU (Urban Historical Inventory), could also provide relevant information; however, we were not able to secure an agreement in time, as BRGM is still negotiating with the metropolitan council for approval.

The original BASIAS dataset contained 3,943 PCSs in the Toulouse metropolitan area, both active and inactive, based on historical inventories. Since some PCSs lacked geolocation data, the final working database included 2,195 sites. To evaluate the methodology, the GIS-based application was applied to a sample of 40 PCSs, selected to represent a range of hazard values, activity types, and geographic settings.

Each PCS is classified according to the French NAF code. These were converted into the corresponding European NACE codes, which were then matched with the activities-pollutants matrix. This procedure

provided four hazard scores for each PCS, basing on the PRA.MS scores: each activity is associated with four pathway-specific hazard values, since the related contaminants may present one or more hazards depending on the exposure pathway characteristic of each substance. For sites where more than one activity was recorded, we retained only the activity associated with the highest hazard score.

Results are shown in Figure 4.

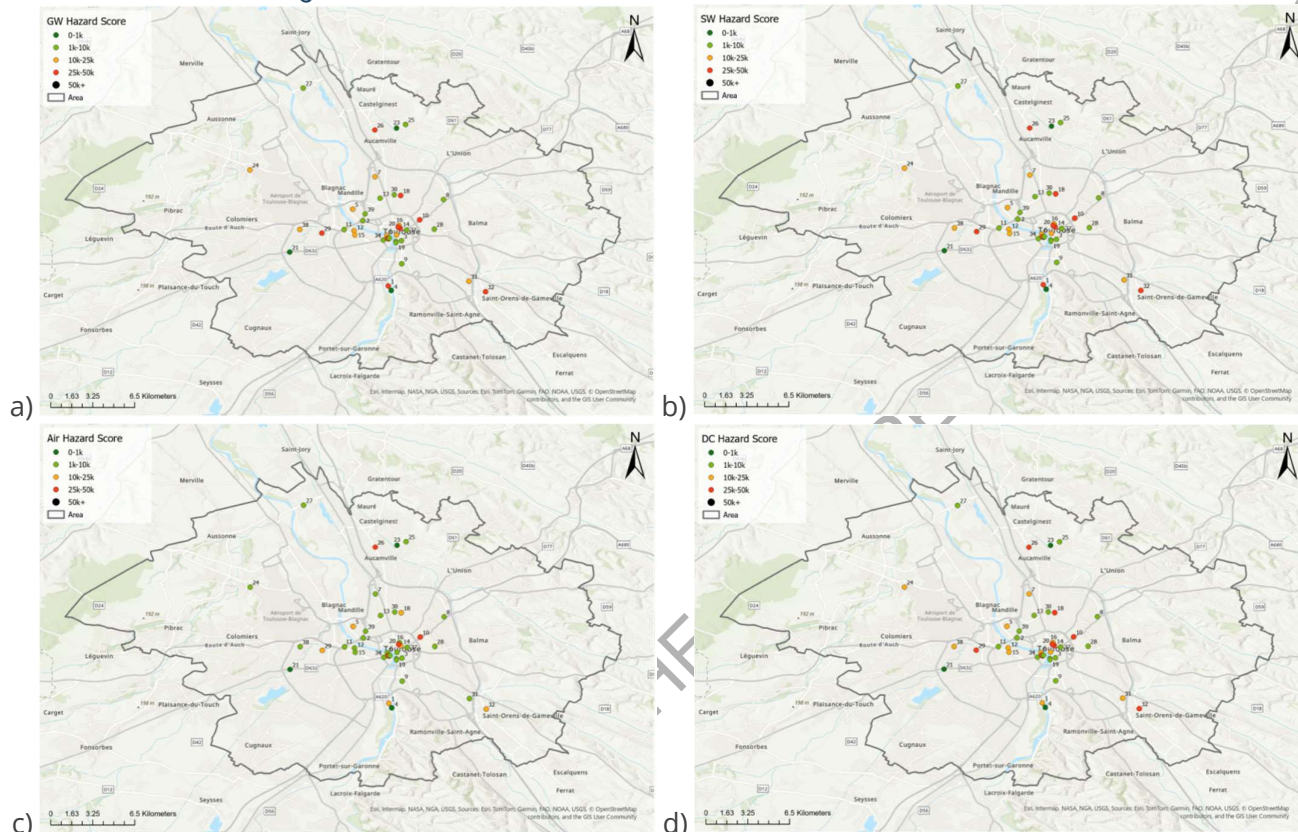


Figure 4. Pathway-related (a) Groundwater, (b) Surface water, (c) Air, (d) Direct contact) hazard scores for 40 PCSs in Toulouse Métropole.

3.3 Pathway Analysis

3.3.1. Groundwater and Surface water pathways

BRGM provided the following datasets for the GW and SW pathway analysis:

- Corine Land Cover
- Digital Elevation Model (DEM)
- Infiltration capacity Index, based on 277 geotechnical surveys conducted in the Toulouse area (BRGM, 2022).

The Corine Land Cover map included 20 classes that were aggregated to fit in the land use parameter scores.

The slope raster derived from the DEM was generated using the *Slope* tool with the PERCENT_RISE option, in order to express the elevation change as a percentage and to assign the corresponding scores.

The DEM was also used to generate the Flow Accumulation map, required for the computation of the SW pathway. The model accounts for the actual amount of water accumulated during precipitation events. Moreover, by integrating the geolocation of PCSs into the Flow Direction analysis, it was possible to identify the potential pathways and locations of contaminants transported by surface water runoff from these sites.

Regarding the Infiltration Capacity Index (ICI), three were the values found in the studied area corresponding to very favorable, not very favorable and not favorable areas to infiltration. Those areas were then reclassified according to the ICI parameter scores (1, 0.2 and 0 respectively).

Examples of Groundwater and Surface Water pathways for one PCS in Toulouse Métropole is given in Figure 5.

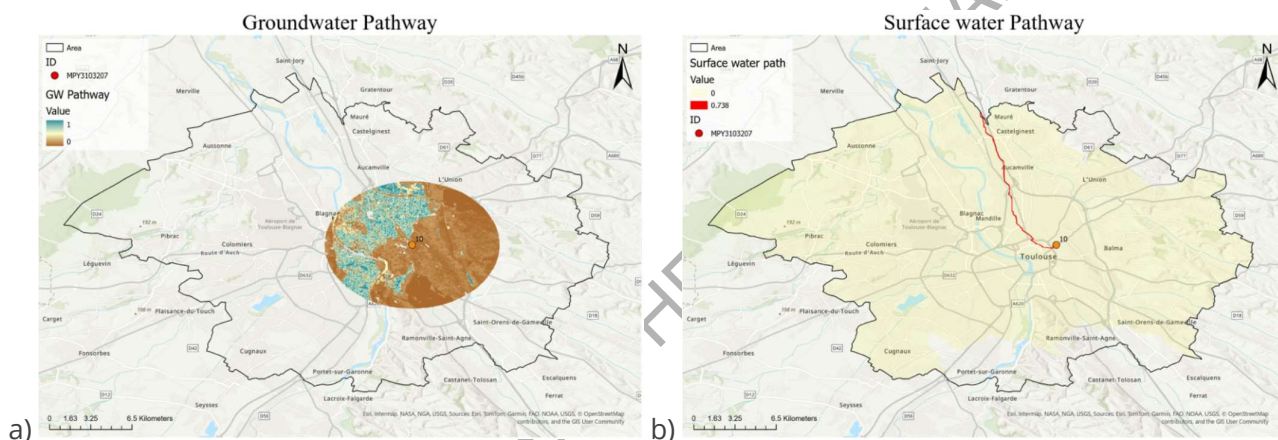


Figure 5. Examples of a) Groundwater, b) Surface water pathways for one PCS in Toulouse Métropole.

3.3.2. Air pathway

Meteorological data (temperature and precipitation) were obtained from the open dataset of Toulouse Métropole (Toulouse Métropole, 2025). The following weather stations were used:

- 01 – Toulouse Metropole
- 03 – Toulouse Busca
- 05 – Toulouse Nakache
- 38 – Toulouse Parc Jardin des Plantes
- 45 – Toulouse St. Exupéry
- 58 – Toulouse Fondeyre
- 62 – Toulouse Parc Maourine

These data were used both for the calculation of the Aridity Index (AI) and as input weights in the Flow Accumulation analysis. The AI was found to be less than 25 in all the area, giving the highest score (Figure 6).

Aridity Index

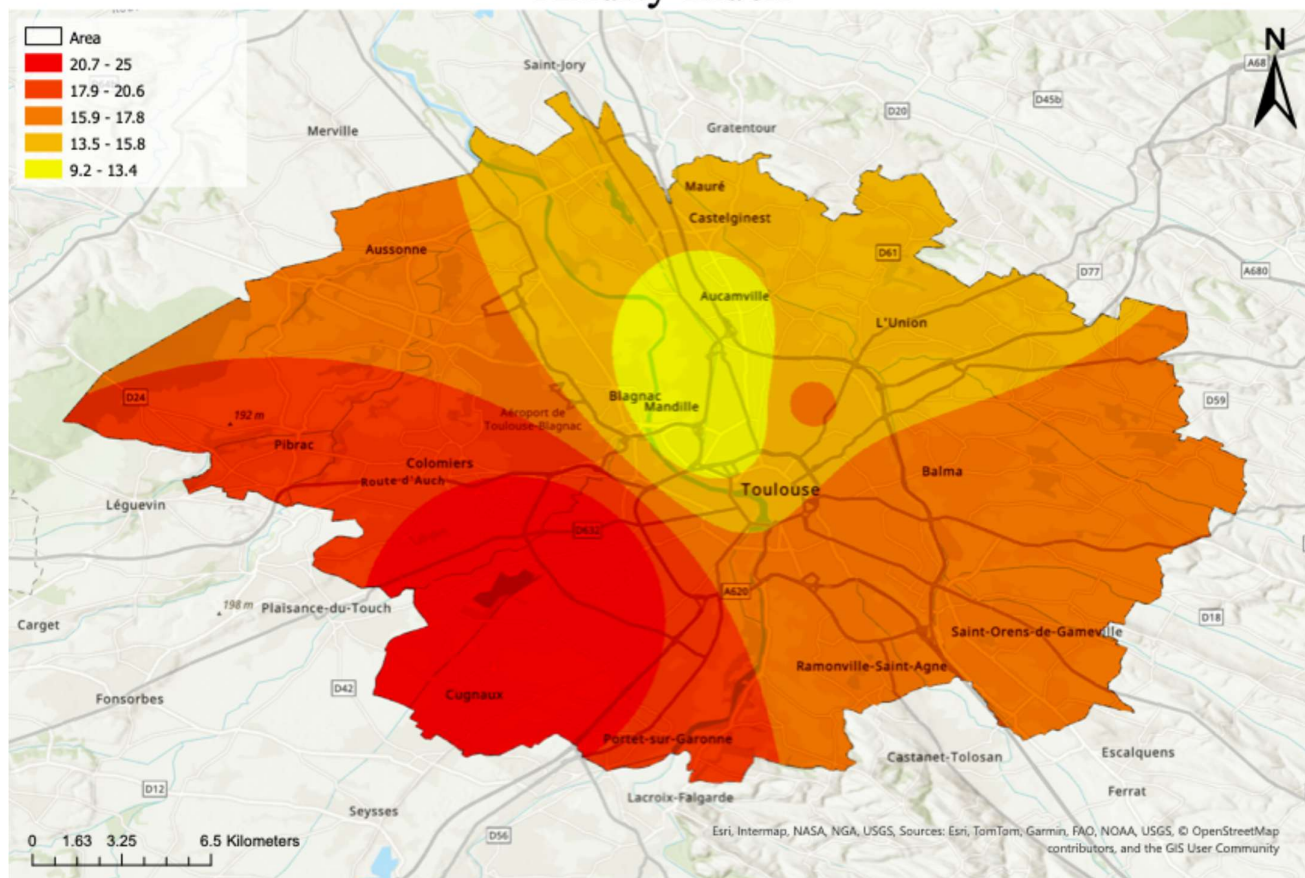


Figure 6. Aridity Index for Toulouse Métropole. All the area has a score of 1 in the reclassified map (<25).

3.4 Vulnerability Analysis

3.4.1. Groundwater Vulnerability

BRGM provided the dataset of groundwater abstraction points for drinking, agricultural, and industrial uses. We then used the shapefile downloaded from the SANDRE Atlas Catalogue, which contains information on groundwater bodies. We focused on the shallowest aquifers, reclassifying the different areas according to the abstraction points located within them. Where multiple points were present, the highest value was retained. Areas lacking information on aquifers as well as those where no abstraction points were identified within an aquifer, were conservatively assigned the score corresponding to potable water use. Results are shown in Figure 7

Groundwater Vulnerability

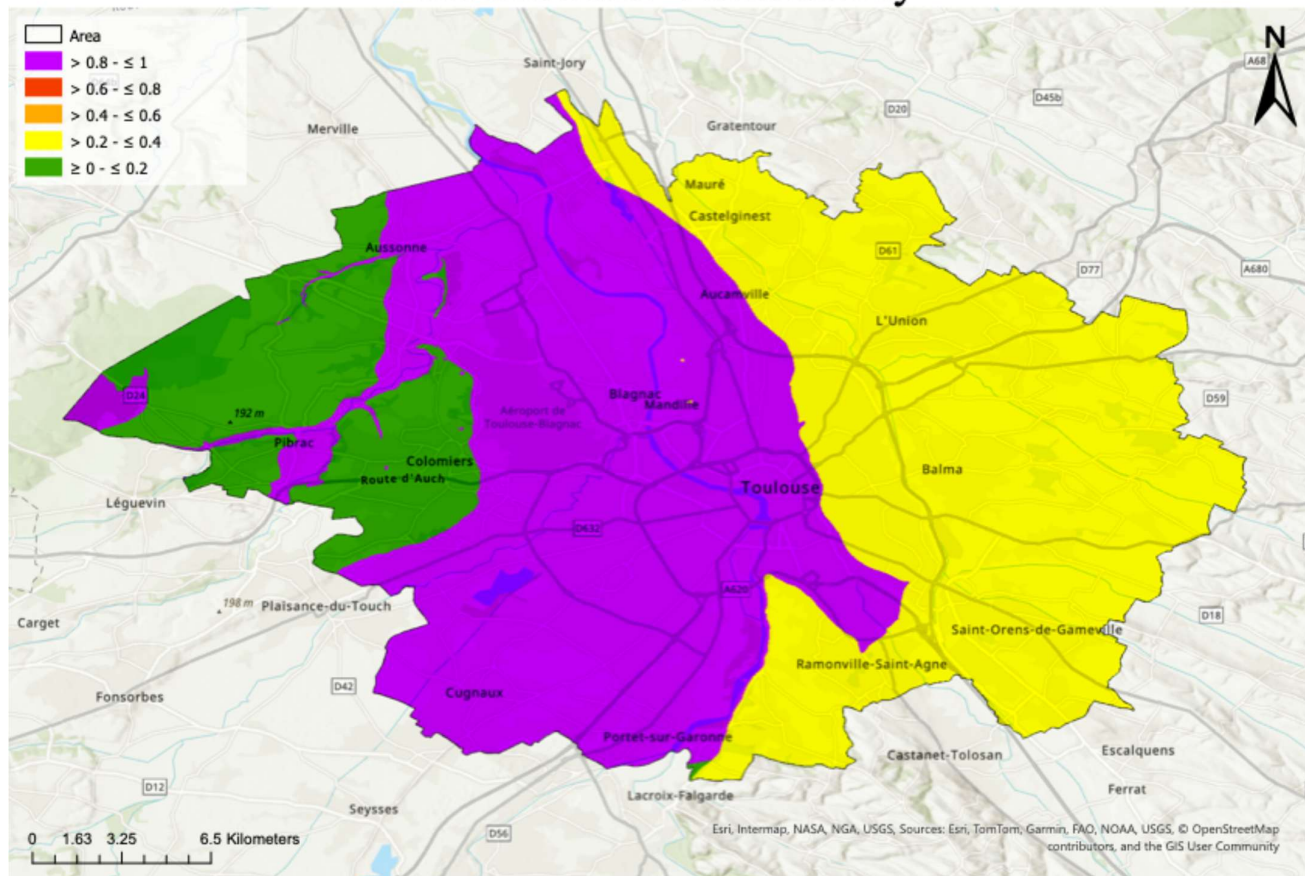


Figure 7. Reclassified Groundwater vulnerability for Toulouse Métropole. A score of 0.3 is for industrial use of GW, 0.85 is for agricultural use, 0.1 for no use.

3.4.2. Surface water vulnerability

For the SW order parameter, we used a shapefile from the SANDRE Atlas Catalogue containing information on the width of different sections of the river network. According to Horton–Strahler laws, the stream order is directly related to discharge, depth, and width. We therefore assigned the following scores based on river width:

- 0–15 m: score 1 (as a precautionary measure)
- 15–50 m: score 0.5
- >50m: score 0.3
- NC: score 0.8

For the SW use parameter, shapefiles containing abstraction points for drinking water and agricultural use were downloaded from the *Système d'Information sur l'Eau du bassin Adour-Garonne* of the Agence de l'Eau Adour-Garonne. Since the precise location of the abstraction points is not provided (in accordance with Vigipirate security regulations, which place them at the center of the municipality), a 100 m buffer was created around the river network to include most of the points. Those located further away, and therefore too imprecise, were not considered in subsequent analyses. We used Thiessen polygons to assign vulnerability scores to river segments, based on the use of the abstraction points

falling within each segment. Where two points with different uses falls within the same polygon, the most conservative score was retained. Results are shown in Figure 8.

Surface Water Vulnerability

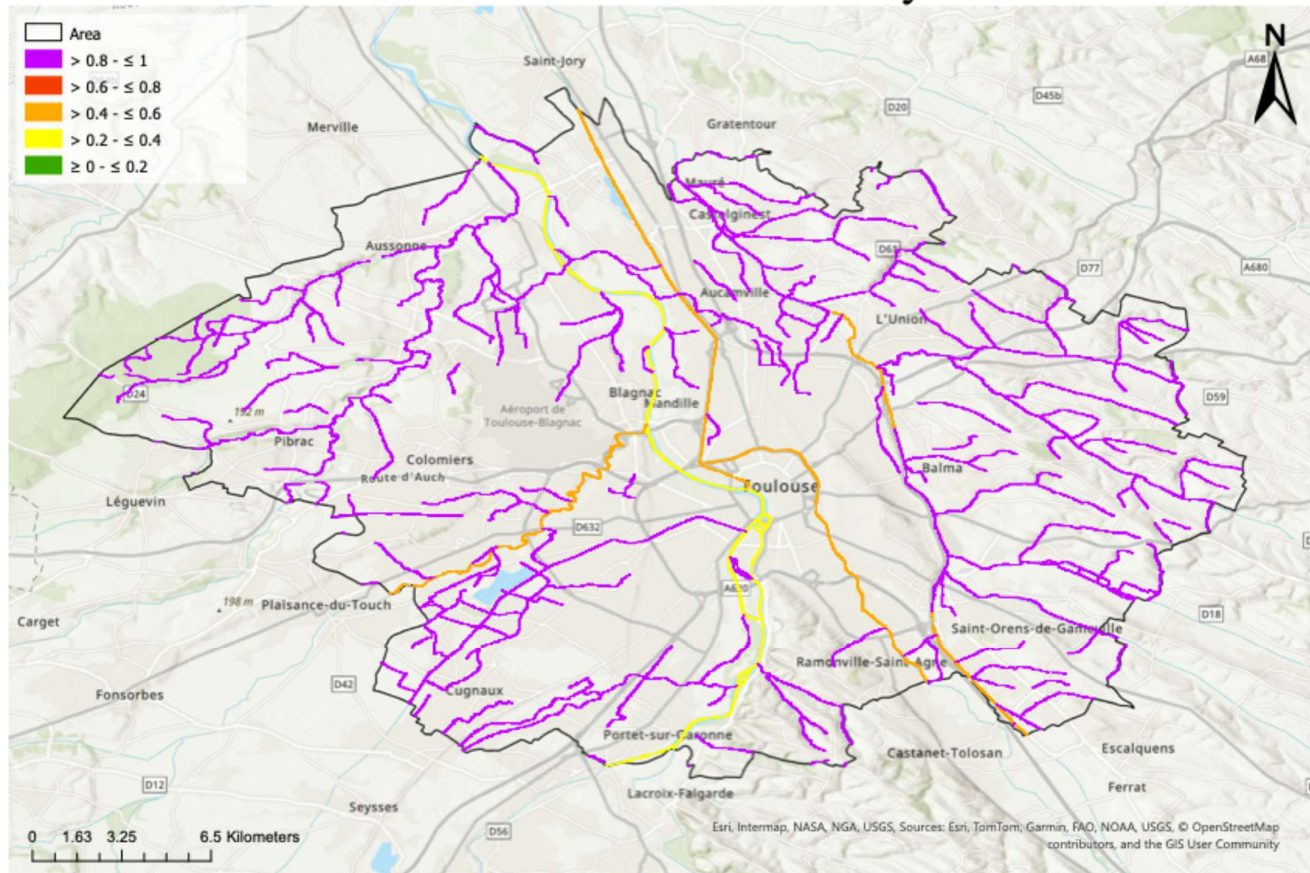


Figure 8. Reclassified map for Surface Water vulnerability in the Toulouse Métropole.

3.4.3. Human health Vulnerability

For the human health vulnerability analysis, we again used the Corine Land Cover map, together with the IRIS database provided by BRGM. The IRIS database (“Ilots Regroupés pour l’Information Statistique”) is the smallest census unit defined by INSEE (2023) and contains detailed demographic information at the sub-municipal level. It provides data on population density for each area as well as information on vulnerable groups, which in our analysis included children aged 0–15 and elderly people over 75 years.

The total human health vulnerability reclassified map is shown in Figure 9.

Human Health Vulnerability

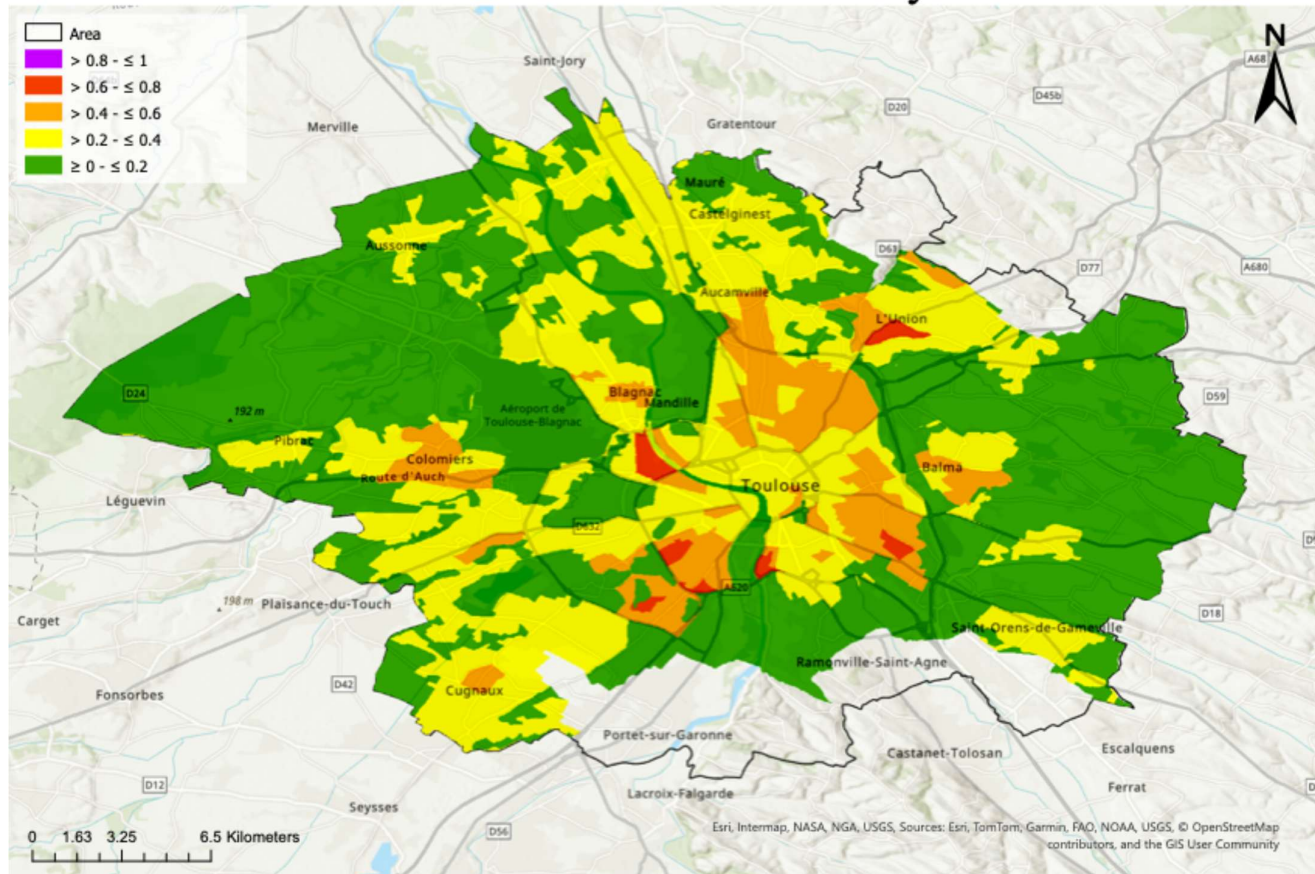


Figure 9. Reclassified Human health vulnerability map for Toulouse Métropole.

3.4.4. Protected Areas Vulnerability

For the protected areas receptor, we used three different datasets: National Nature Reserves (score 1; Muséum national d'Histoire naturelle [MNHN], 2024), Regional Nature Reserves (score 0.8; Région Occitanie, 2024), and Sensitive Natural Areas (ENS, local ecological sensitive areas) (score 0.5; Observatoire Régional de l'Environnement Occitanie [OREO] / OpenIG, 2022). These datasets were merged into a single raster, and each category was assigned the corresponding score in order to generate the final protected areas receptor map (Figure 10).

Protected Areas Vulnerability

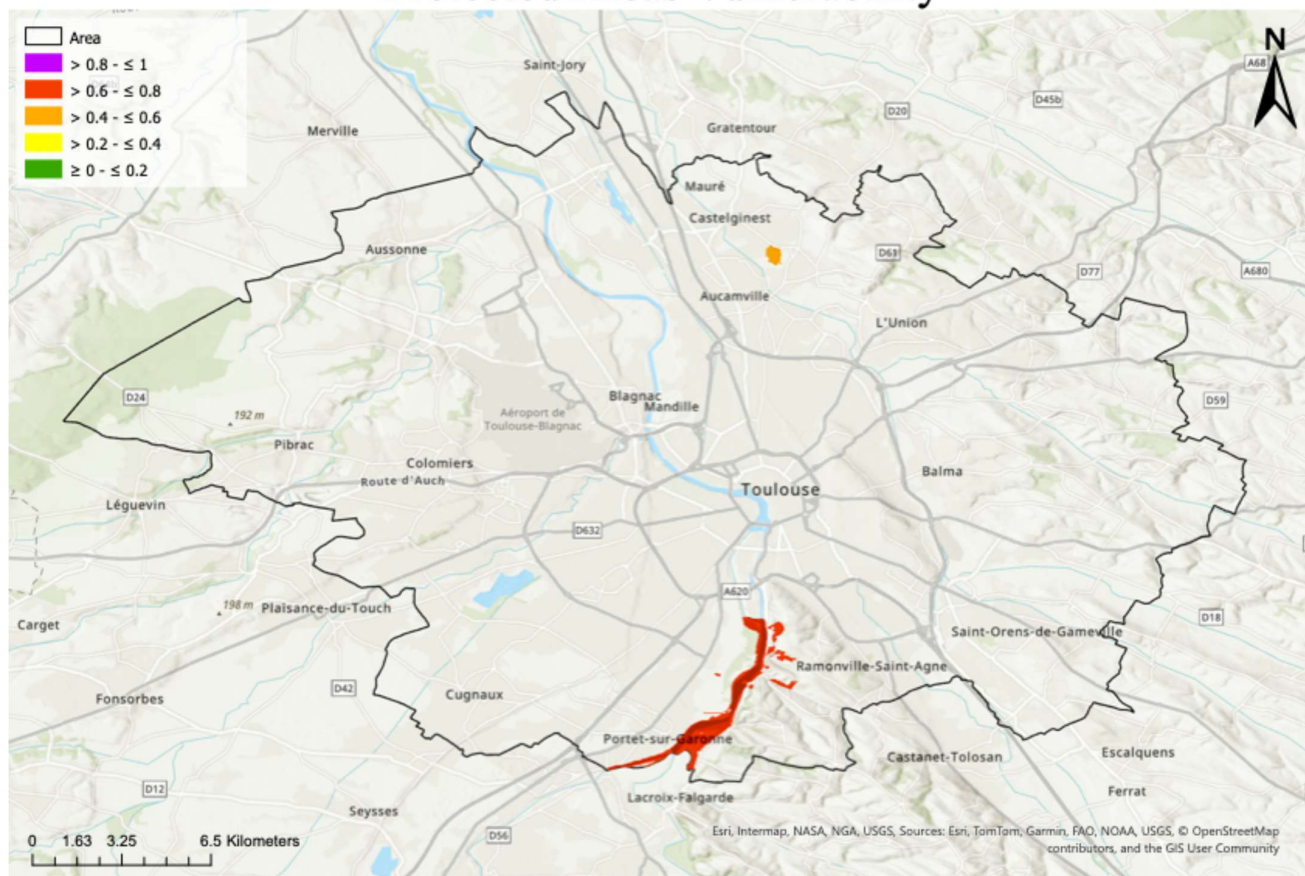


Figure 10. Reclassified map for Protected areas vulnerability in Toulouse Métropole.

3.5 Total Risk map

The resulting risk maps extend over a defined area, depending on the pathways and receptors considered. For classification and prioritization purposes, these spatially distributed values need to be aggregated to a single representative value for each source. For each pathway (groundwater, surface water, direct contact, and air), both maximum and mean risk values were calculated. When aggregating the pathways to obtain the total risk, we used the mean values, as they provide a more balanced representation of the overall exposure potential across the affected area. Nevertheless, maximum values were also reported in Table 8, since they reflect a precautionary perspective by highlighting the highest possible risk associated with each source. This dual reporting allows the methodology to capture both the average conditions relevant for regional-scale prioritization and the extreme cases that may require further site-specific investigation. Figure 11 presents the spatial results of aggregated overall risk for the selected PCSs.

Potentially Contaminated Sites Risk

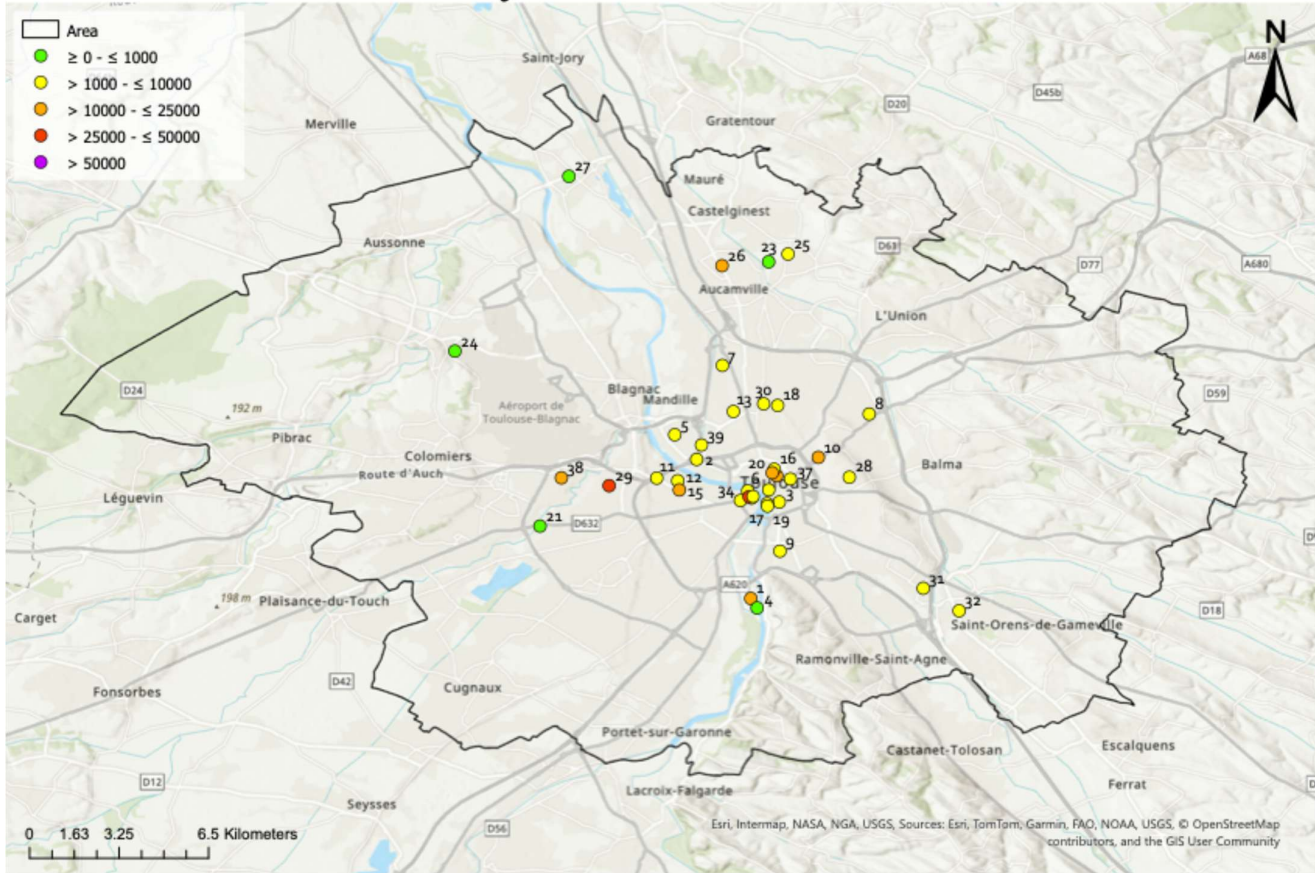


Figure 11. Aggregated overall risk of the selected PCS in Toulouse Métropole.

AWAITING APPROVAL BY THE COMMISSION



OBJECTID	ID	NACE	GW Hazard Score	SW Hazard Score	Air Hazard Score	DC Hazard Score	GW mean risk	GW max risk	SW risk	Air hh mean risk	Air hh max risk	Air pa mean risk	Air pa max risk	DC hh mean risk	DC hh max risk	Risk TOT
1	MPY3100173	C21.10Z	27650	27650	14510	23500	757	21586	9940	1297	8505	4940	10618	1010	6860	17944
2	MPY3100197	C20.18Z	3600	3600	2040	3180	132	2631	1531	206	1254			898	6860	2766
3	MPY3100241	D35.2	9100	9100	4060	7280	265	6923	1667	443	1937	380	642	873	6860	3628
4	MPY3100376	E37.00Z	0	0	0	0								997	6860	997
5	MPY3101818	C25.62B	21390	21390	11210	17580	737	14879	6854	1028	6746			742	6860	9362
6	MPY3101841	G47.30Z	13090	13090	5880	10430	451	11095	6793	642	2668	299	627	931	6860	9117
7	MPY3101989	V89.01Z	14400	14400	7740	11460	546	7355	4390	732	3549	406	657	283	4461	6357
8	MPY3102103	S96.01	9020	9020	4880	7380	140	2016	3164	467	2440			214	3786	3985
9	MPY3102283	G45.21B	5530	5530	2380	4340	148	4072	420	240	1600	541	1037	942	6860	2291
10	MPY3103207	C20.30Z	48870	48870	27070	38050	1179	24487	14199	2977	12092			608	6860	18963
11	MPY3103636	C18.1	6650	6650	3010	5180	210	4851	3065	279	2094			948	6860	4503
12	MPY3103902	V89.01Z	14400	14400	7740	11460	500	10400	6773	756	4990			970	6860	8998
13	MPY3104117	V89.03Z	5670	5670	3780	4900	220	2478	705	394	1890			542	6860	1861
14	MPY3104273	C20.20Z	39130	39130	21170	31510	1213	27917	9174	2375	9703	62	207	814	6860	13638
15	MPY3104329	H49.39	14220	14220	8220	11520	498	9454	8035	809	4938	153	415	992	6860	10487
16	MPY3104433	B07.29Z	13040	13040	6560	10640	416	8997	3220	737	2863			801	6860	5174
17	MPY3104713	G45.21A	7920	7920	4380	6240	247	6224	1419	475	2077	445	773	913	6860	3500
18	MPY3104897	E38.31Z	36360	36360	20220	29220	1268	12362	4939	2187	10110			437	5110	8832
19	MPY3104943	S96.01	9020	9020	4880	7380	280	7215	1575	528	2372	529	924	916	6860	3828
20	MPY3104948	C20.30Z	48870	48870	27070	38050	1565	34848	11751	3036	11951			818	6860	17170
21	MPY3105147	E37.00Z	0	0	0	0								717	6860	717
22	MPY3105348	H49.39	14220	14220	8220	11520	449	11372	2945	910	3680	383	782	876	6860	5563
23	MPY3105471	E37.00Z	0	0	0	0								0	99	0
24	MPY3109992	V89.01Z	14400	14400	7740	11460	46	4841		288	1950			5	716	339
25	MPY3110151	V89.03Z	5670	5670	3780	4900	107	1307	2844	332	1427	1711	1812			4994
26	MPY3110152	C20.30Z	48870	48870	27070	38050	1411	18724		2235	10936	9479	10065	1	394	13126
27	MPY3110235	G45.21A	7920	7920	4380	6240	253	4136		300	1314					553
28	MPY3111316	S96.01	9020	9020	4880	7380	145	3911		511	2440			520	6860	1176
29	MPY3111362	E38.31Z	36360	36360	20220	29220	808	19860	30010	1682	11668			885	6860	33386
30	MPY3111785	S96.01	9020	9020	4880	7380	332	3067	1351	524	2440			449	5215	2657
31	MPY3111886	V89.01Z	14400	14400	7740	11460	55	2517		565	4038	576	1401	185	4143	1381
32	MPY3112198	E38.31Z	36360	36360	20220	29220	98	4489		1222	7581	534	1412	84	3139	1938
33	MPY3112536	S96.01	9020	9020	4880	7380	304	7314	5036	530	2259	388	726	939	6860	7197
34	MPY3112598	S96.01	9020	9020	4880	7380	315	7042	5242	525	2403	421	794	965	6860	7468
35	MPY3112636	C20.30Z	48870	48870	27070	38050	1664	39643	26942	2942	12586	1981	3792	941	6860	34470
36	MPY3112638	V89.03Z	5670	5670	3780	4900	190	4665	3110	412	1726	279	527	933	6860	4925
37	MPY3112802	G45.21A	7920	7920	4380	6240	224	5407	1779	491	2147	31	78	785	6860	3311
38	MPY3114097	E38.11Z	14960	14960	7760	11840	189	5489	10035	581	3315			685	6860	11489
39	MPY3115079	A01.6	4130	4130	1680	3150	155	2813	1937	169	929			819	6860	3080

Table 8. Results for groundwater (GW), surface water (SW), direct contact (DC) and air risk (max and mean values) and total risk for 40 selected PCSs in Toulouse Métropole. HH: human health, PA: protected areas.

4. Regional Risk Assessment for diffuse contamination - application to ITA Toulouse Métropole

4.1. Hazard Analysis

The application of the RRA methodology for diffuse contamination was tested using the interpolated concentration maps developed in Task 1.3 and included in Deliverable D1.2, which served as the basis for defining the spatial hazard component. Several algorithms have been tested and tuned to propose an innovative anomaly detection algorithm especially adapted to cases where data is sparse (< 30), clustered and uncertain.

Toulouse Métropole has been chosen by the French Ministry of the Economy, Finance and Industry as a pilot agglomeration to demonstrate the operational feasibility of the recommendations made by the “Urban Geochemical Background Values” working group, and to provide support for the methodology for excavated soil reuse. Analyses of surface soil samples available for sensitive facilities (40 samples) were supplemented by those obtained during site diagnostics (1442 samples) commissioned by the Metropole, and during two sampling campaigns carried out by BRGM. These campaigns resulted in the collection of 140 high quality surface soil samples and 100 deep soil samples taken every meter in 20 boreholes 5 m deep. Analyses covered 24 parameters, while in this application we focused on: metallic

trace elements (As, Cd, Cr, Cu, Hg, Ni, Pb, Zn); PAHs; PCB; Sum of light hydrocarbons C5-C10 hydrocarbons; sum of C10-C40 hydrocarbons; and dioxins. For more information, see: Belbeze et al. (2019).

The Thresholds of contaminants in soil selected for maps reclassification were derived from the Italian national legislation (i.e., Italian soil quality standards, called CSC - Contamination Screening Concentrations, corresponding to the Italian *Concentrazioni Soglia di Contaminazione* as defined in D.Lgs. 152/2006) and reported in Table 9. To enable comparability with the Italian CSC, French hydrocarbon fractions were harmonised as follows: CNT (\approx C6–C10) was compared against the CSC for light hydrocarbons (C<12), whereas C10–C40 or HCT was compared against the CSC for heavy hydrocarbons (C>12). To address the C10–C12 overlap between definitions, the C10–C12 band was conservatively attributed to the heavy fraction.

Italian CSC are used strictly as non-binding benchmarks to support regional screening and prioritisation; exceedance of a CSC triggers further investigation in the Italian framework but in itself does not demonstrate an unacceptable risk. Results should therefore be interpreted as screening indicators rather than definitive site-specific conclusions.

For illustrative purposes, we present the results obtained for Pb, which showed the greatest exceedance of CSC among all analysed contaminants (Figure 12).

Parameter	Measure unit	Italy (CSC, D.Lgs. 152/06 – res./green) [mg/kg]	Italy (CSC, D.Lgs. 152/06 – agricol) [mg/kg]	Italy (CSC, D.Lgs. 152/06 – industr) [mg/kg]
As	mg/kg	20	30	50
Cd	mg/kg	2	5	15
Cr	mg/kg	150	150	800
Cu	mg/kg	120	200	600
Hg	mg/kg	1	1	5
Ni	mg/kg	120	120	500
Pb	mg/kg	100	100	1000
Zn	mg/kg	150	300	1500
Heavy hydrocarbons (C>12)	mg/kg	50	50	750
Somme 16HAP	mg/kg	10	10	100
Somme 7PCBi	mg/kg	0,06	0,02	5
Sum dioxines-furanes	TEQ	10	1	100
Light hydrocarbons (C<12)	mg/kg	10	10	250

Table 9. CSC (Contamination Screening Concentrations, corresponding to the Italian *Concentrazioni Soglia di Contaminazione* as defined in D.Lgs. 152/2006) for considered contaminants in residential/green, agricultural and industrial/commercial areas.

Hazard Pb

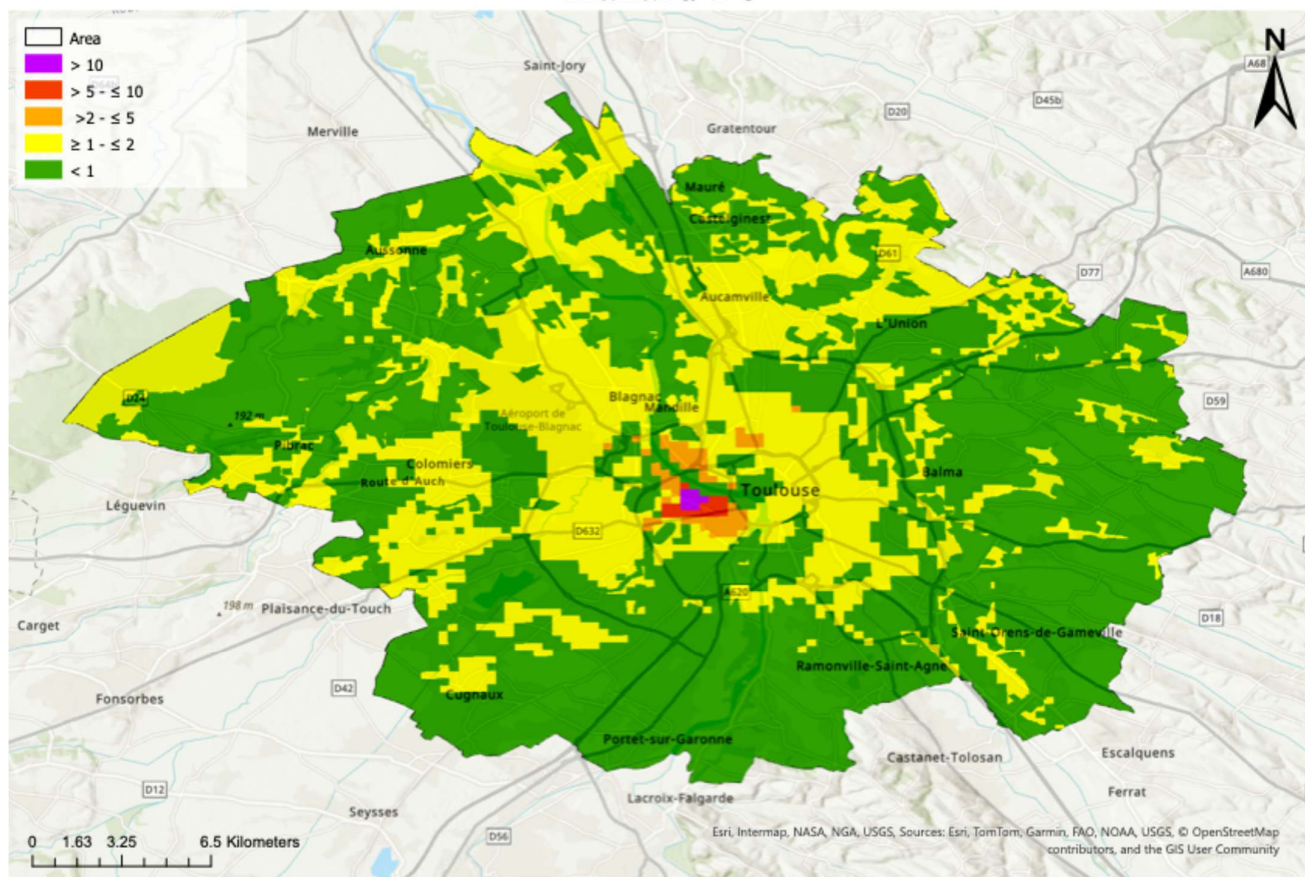


Figure 12. Pb hazard in Toulouse Métropole: a) absolute values; b) Pb hazard normalized with CSC thresholds; c) Pb normalized hazard in agricultural areas; d) Pb normalized hazard in residential areas; e) Pb normalized hazard in industrial/commercial areas.

The hazard map obtained for each contaminant represents only the exceedance of the regulatory threshold and not an absolute concentration. The raster was reclassified into classes according to the magnitude of exceedance, allowing the identification of areas where concentrations approach or moderately exceed the CSC, as well as clear hotspots. In particular, cells where concentrations exceeded the CSC by more than a factor of 10 were highlighted, as such exceedances are generally considered indicative of potentially unacceptable risk and may warrant further investigation.

4.2. Pathway Analysis

Results for the GW pathway map are shown in Figure 13. The same map is used for all contaminants: for details regarding the methodology to obtain the Groundwater pathway map see paragraph 3.3.1.

Groundwater pathway

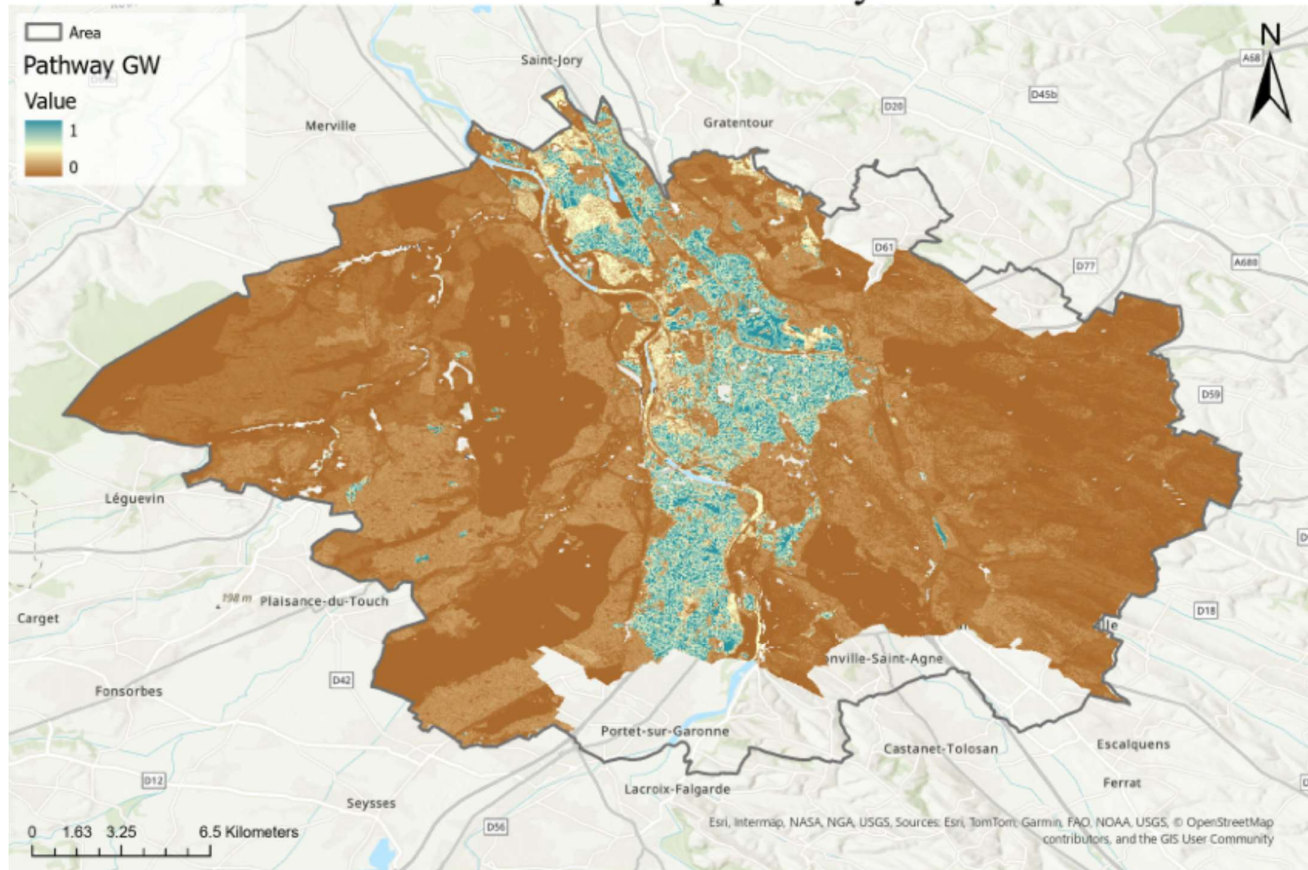


Figure 13. Reclassified Groundwater pathway map for diffuse contamination in Toulouse Métropole.

4.3. Vulnerability Analysis

For details on the vulnerability analysis methodology, see paragraph 3.4. The same receptor maps used for groundwater, surface water, human health, and protected areas (Figure 7-Figure 10) were applied.

4.4. Contaminants and total risk maps

Figure 14-Figure 18 illustrate the risk results for Pb contamination in soils, for the Groundwater, Surface Water, Air pathways, which includes both human health and protected areas as receptors, and for the Direct Contact pathway in relation to human health.

Figure 19 presents the total risk associated with Pb contamination, while Figure 20 shows the cumulative risk resulting from the summation of all contaminants analysed in the study area.

Groundwater Risk - Pb

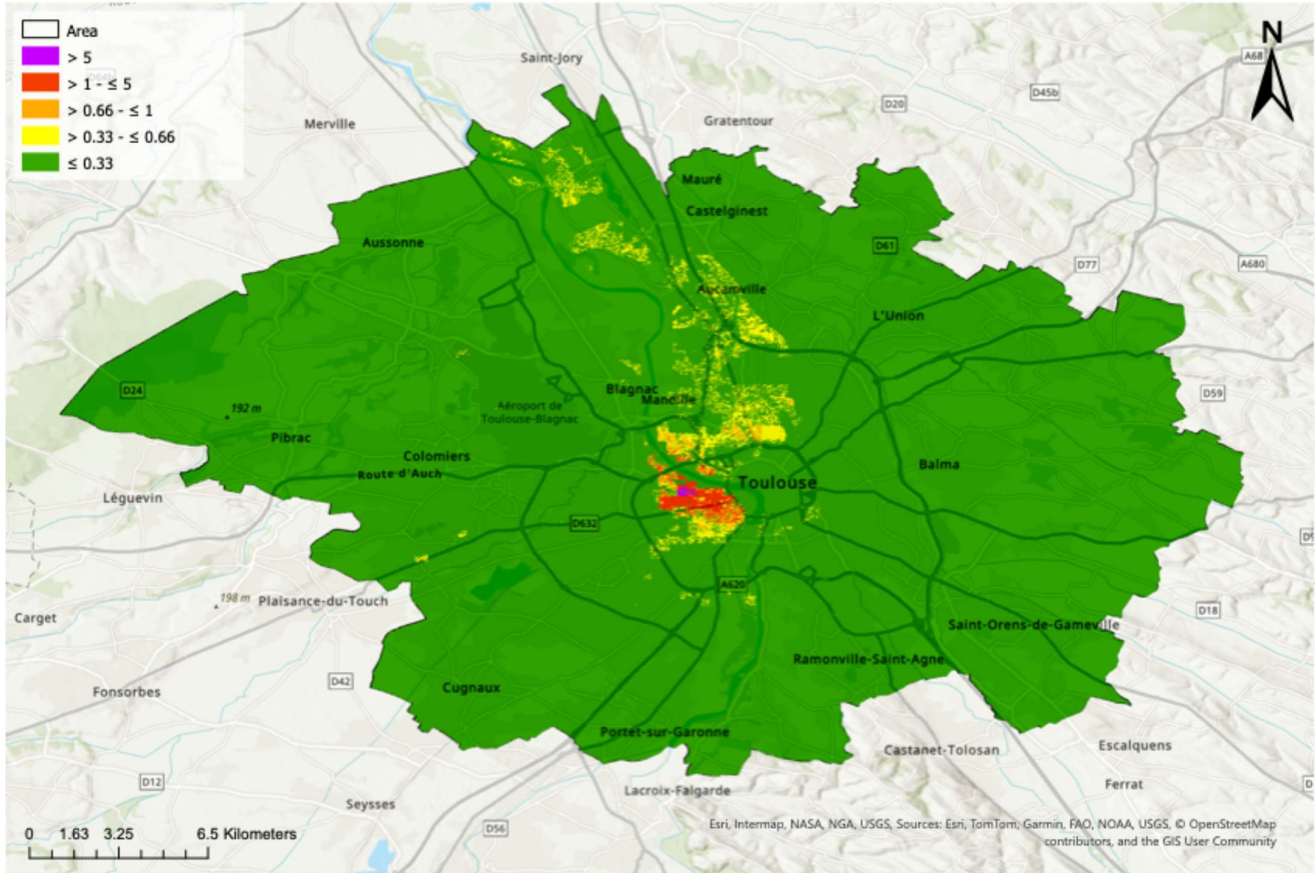


Figure 14. Groundwater risk map from Pb diffuse contamination in Toulouse Métropole soils. Maximum risk is 9.6.

AWAITING APPROVAL BY

Surface Water Risk - Pb

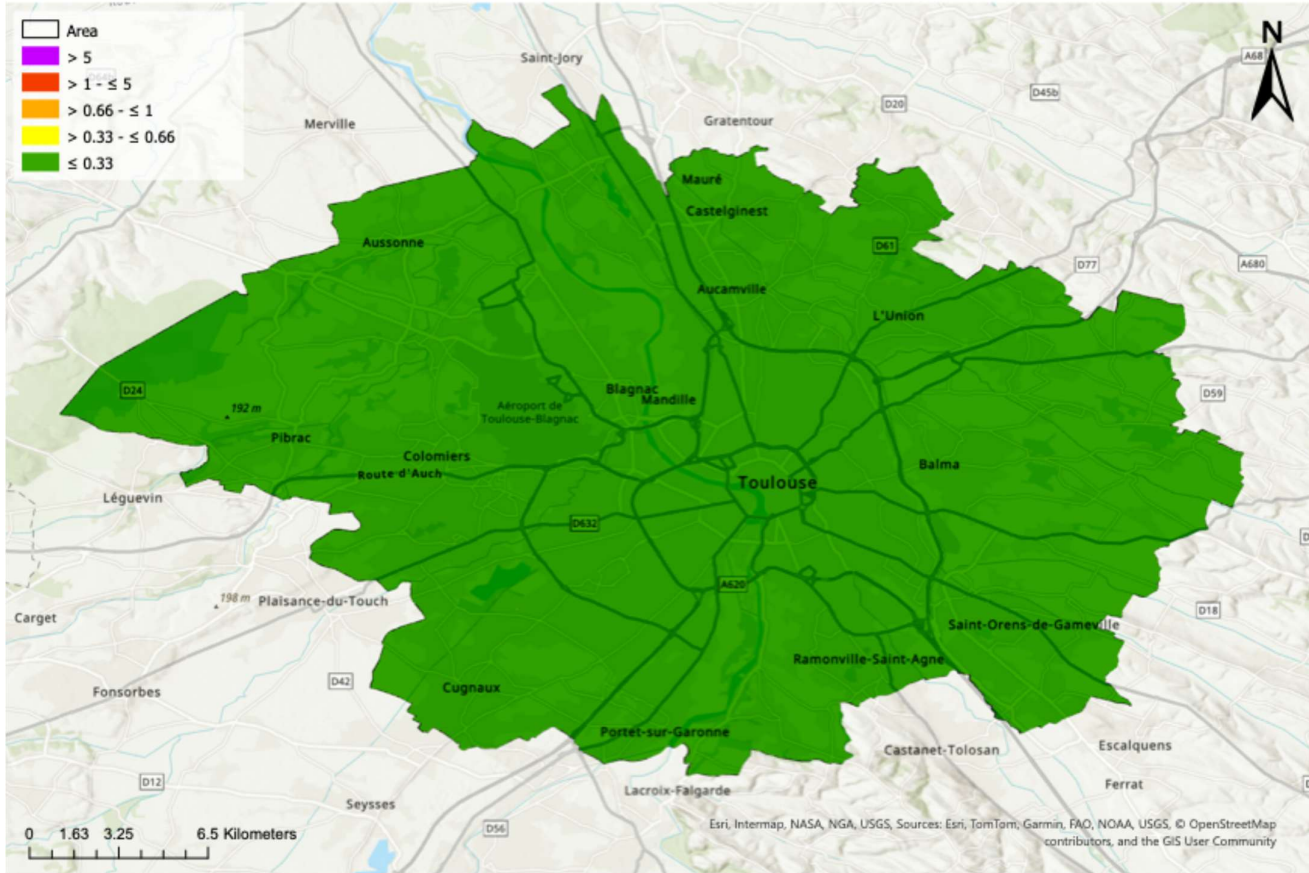


Figure 15. Surface water risk map from Pb diffuse contamination in Toulouse Métropole soils. Maximum risk is 0.68. Pixels representing the intersection between the Surface water pathway and surface water bodies are not visible at this scale.

AWAITING APPROVAL BY

Air Exposure Risk - Human Health Receptor - Pb

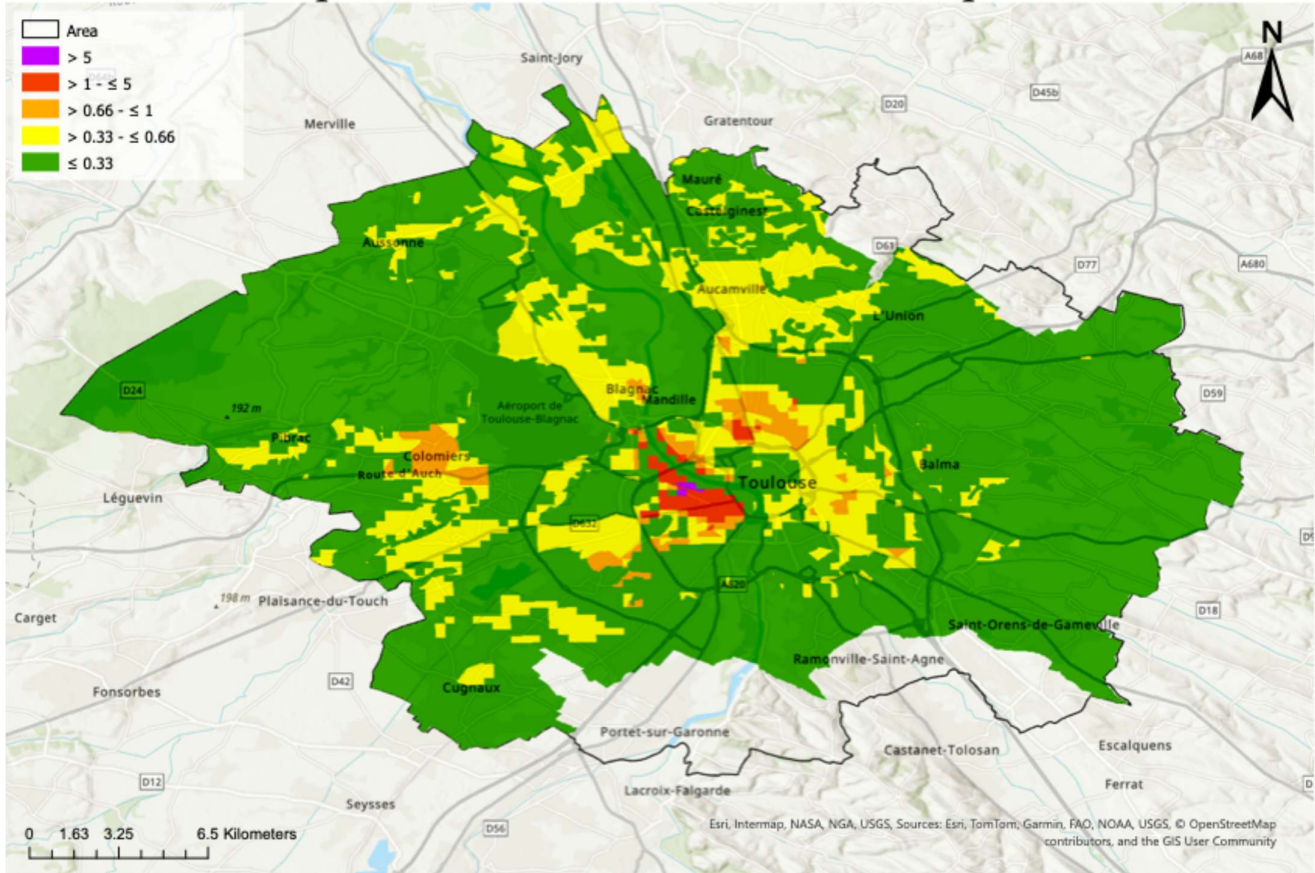


Figure 16. Air risk map for human health due to Pb diffuse contamination in Toulouse Métropole soils. Maximum risk is 8.5.

AWAITING APPROVAL BY

Air Exposure Risk - Protected Areas Receptor - Pb

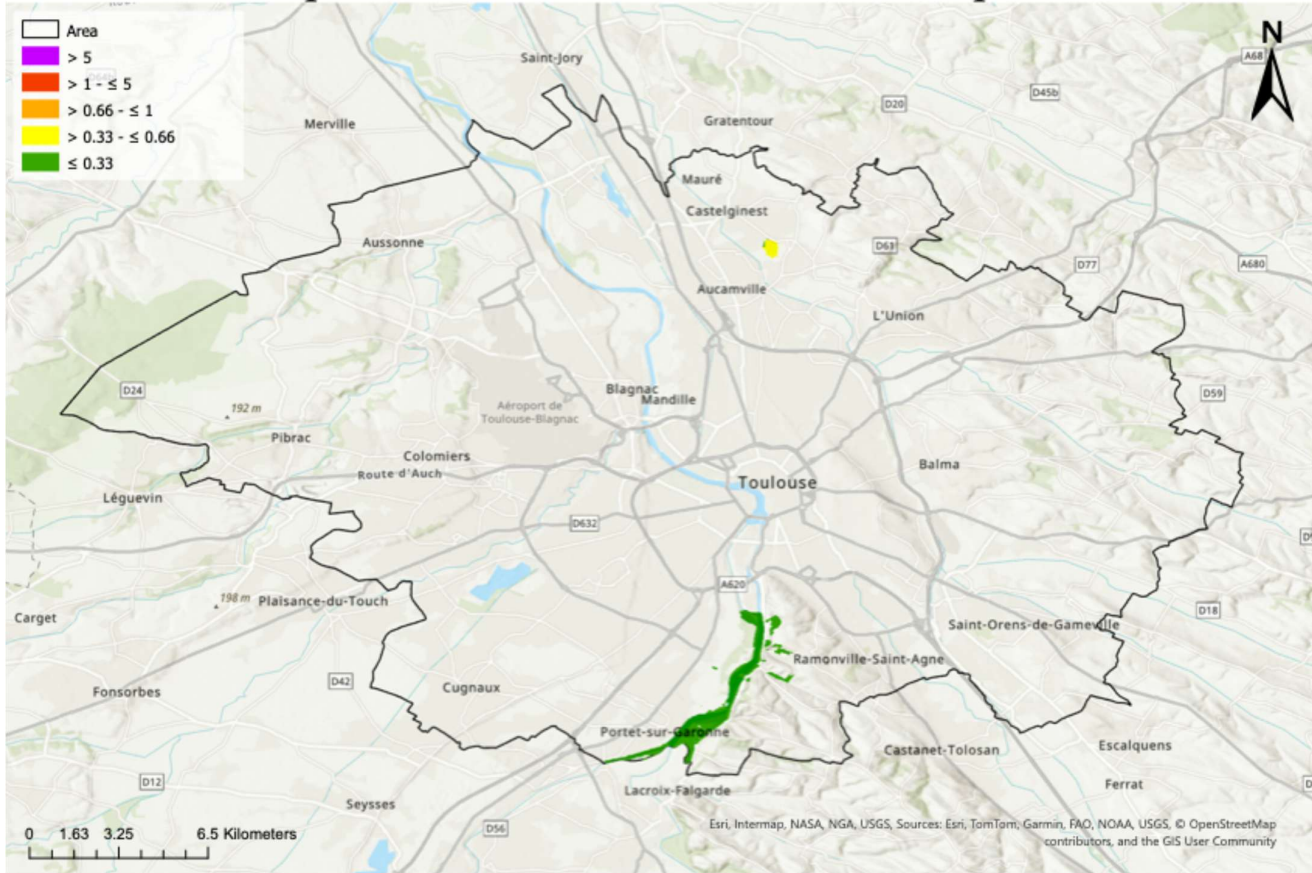


Figure 17. Air exposure risk map for protected areas due to Pb diffuse contamination in Toulouse Métropole soils. Maximum risk is 0.57.

AWAITING APPROVAL BY

Total Risk - Pb

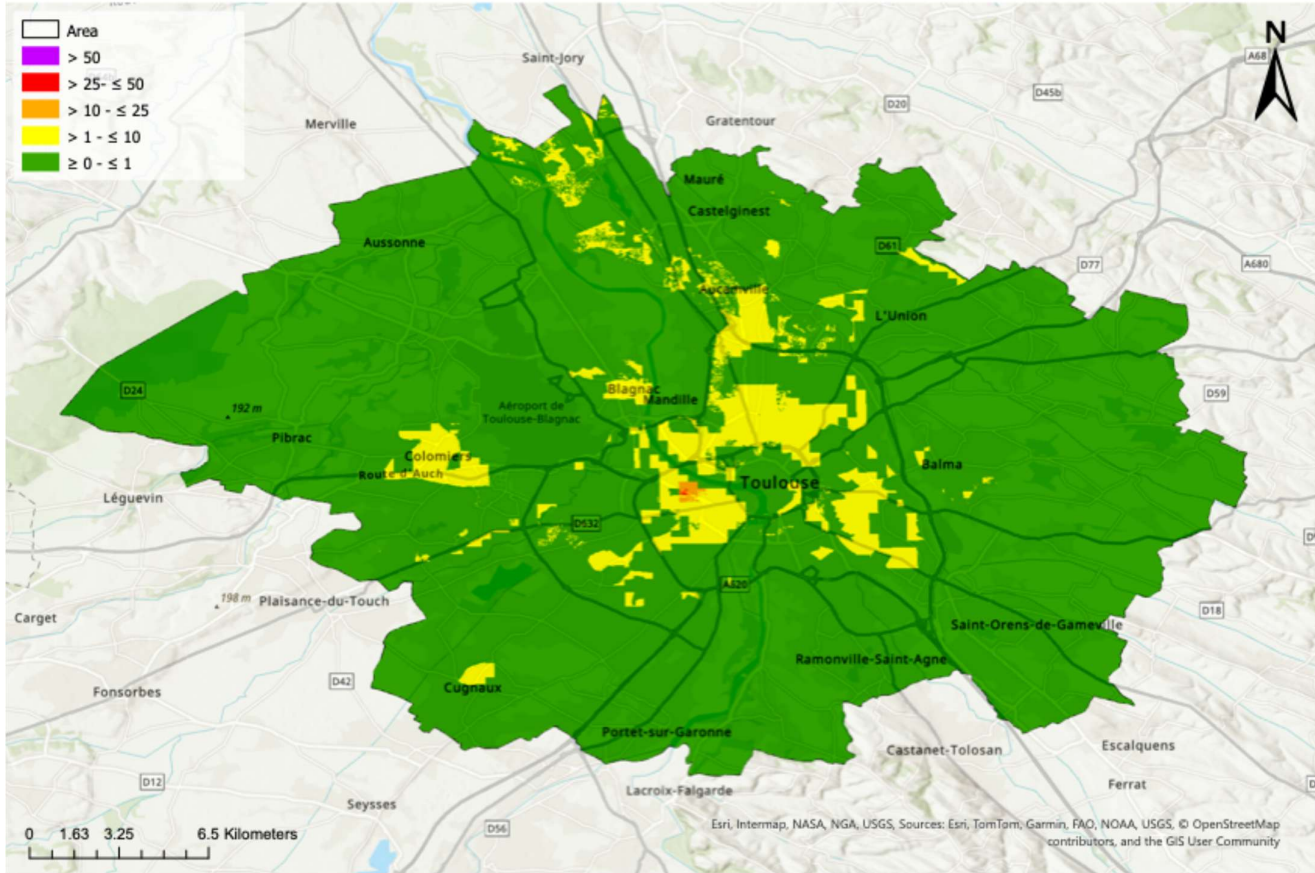


Figure 19. Total risk map due to Pb diffuse contamination in Toulouse Métropole soils.

AWAITING APPROVAL BY THE

Cumulative Risk

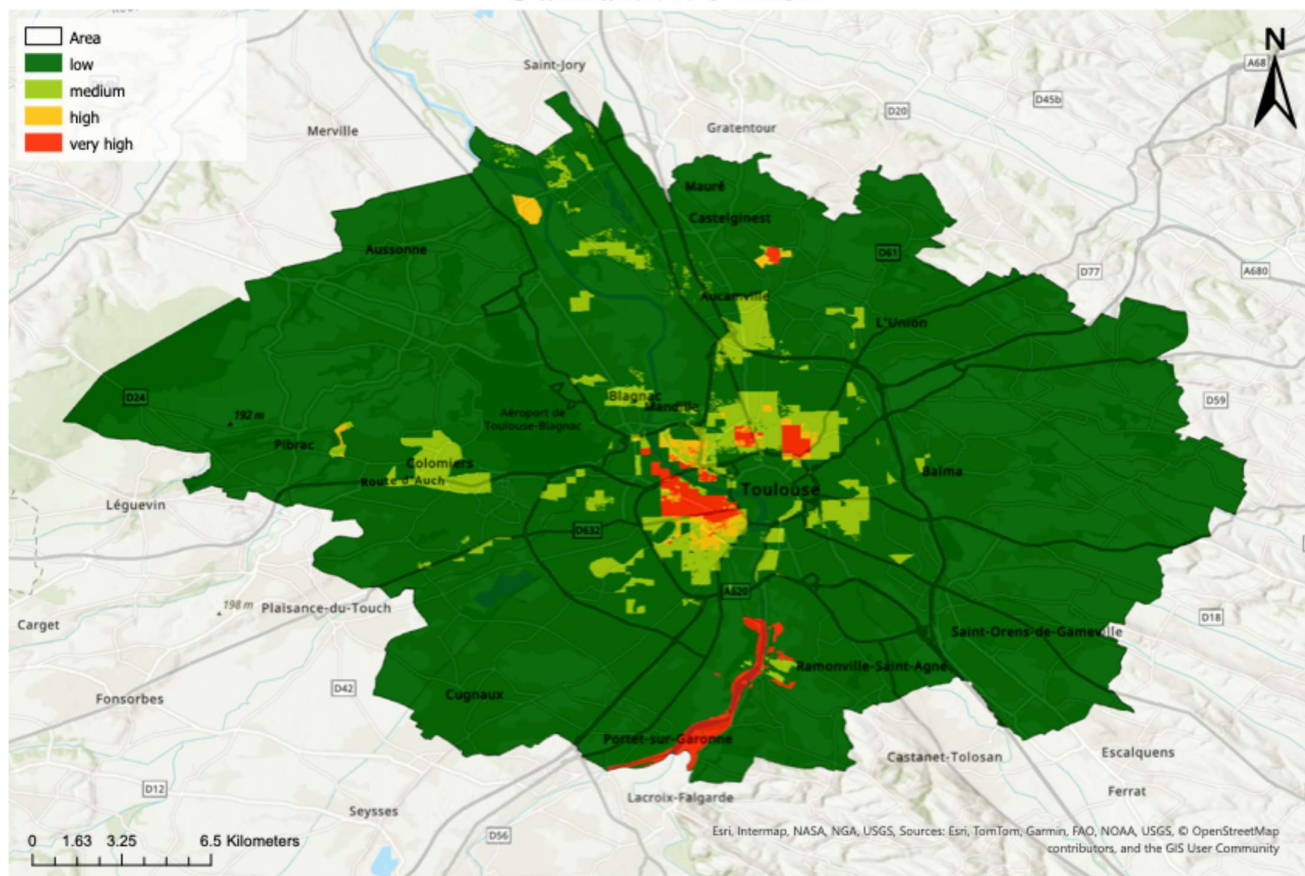


Figure 20. Cumulative risk map from diffuse contamination in Toulouse Métropole soils obtained by summing the risk contributions from all analysed contaminants. The resulting values were represented in 4 classes of intervention priority based on the Natural Breaks classification of the aggregated risk scores distribution, allowing differentiation between low priority areas and hotspot zones where further assessment or management measures may be required.

5. Advantages and limits of the methodology

The proposed methodologies support the implementation of the SML, which requires Member States to establish harmonised soil monitoring systems and apply a risk-based, stepwise approach (Articles 12 and 43b). While the Directive does not mandate the collection of all possible anthropogenic or historical data, it emphasises the need for harmonised soil indicators, contamination screening, and prioritisation tools to guide subsequent investigations. The proposed RRA frameworks directly contributes to these objectives by providing a consistent and scalable screening-level methodology for identifying areas at risk.

Specifically, the methodologies support the stepwise risk assessment approach promoted in Article 43b of the SML, which requires a structured transition from broad-scale preliminary screening to more detailed investigations only when necessary. In this context, the methodologies provide an effective basis for identifying areas or potential point sources that may warrant further investigation, thus

optimizing resource allocation and avoiding unnecessary site-specific assessments when risk levels are low.

The regional-scale application would allow Member States and competent authorities to:

- prioritise investigation steps within a legally compliant tiered framework;
- integrate both point-source and diffuse contamination within a unified conceptual model;
- anticipate remediation needs or preventive land management measures under Article 43b;
- facilitate strategic planning and policy decision-making at regional or national level.

The two methodological frameworks presented in this Deliverable are complementary. The PCS-based RRA allows the prioritisation of individual sites for further investigation, while the diffuse contamination approach provides an integrated view of widespread pollution patterns. Depending on the data available and the management objectives, the two approaches can be applied independently or combined to deliver a more complete regional risk picture and to support stepwise, proportionate decision-making.

The methodologies particularly the GIS application, could still need further refinement to ensure accuracy across different contexts. The main challenge in applying the methodologies lie in the availability and consistency of data, particularly when dealing with large-scale analyses. The approach may need to be adapted case by case depending on the type and quality of the data available. For this reason, even with a model builder available, it is essential that a GIS expert carry out the implementation. It is crucial to understand the nature of the input datasets and to verify the coherence of the results obtained.

In our case, one of the issues encountered was the delimitation of the study area, which followed a “political” boundary. As a result, in the PCSs RRA, we observed that, especially in the application of the pathway and receptor vulnerability analyses, the sources located in the central part of the study area produced more realistic outcomes compared to those located near the borders, where the flow paths were truncated. This limitation was particularly evident in the analysis of the surface water pathway.

An approximation introduced in the application concerns the parameters for the analysis of surface water pathway. Another limitation of this approach is that Thiessen polygons, used in the SW vulnerability evaluations, are based solely on Euclidean distance and do not account for hydrological boundaries or flow directions. Therefore, the assigned influence areas may not accurately reflect the actual catchment zones of the abstraction points. This limitation could be overcome by integrating hydrologically defined catchment areas, where available, which would provide a more realistic representation of the actual zones of influence of each abstraction point.

For all cases with imprecise information or missing data, a conservative approach was adopted by assigning the maximum risk scores.

6. Conclusions

The Deliverable developed and applied a comprehensive framework for RRA that integrates both point source and diffuse soil contamination within a scalable, GIS-based Source-Pathway-Receptor (S-P-R) structure. The first part of the methodology aims at ranking PCSs based on intrinsic hazard scoring,



pathway-specific transport mechanisms and receptor vulnerability, producing comparable risk indicators that support regional-level screening and prioritisation. The second part extends the same conceptual framework to diffuse contamination, adapting the hazard component to measured soil concentrations and enabling multi-contaminant accumulation through pathway–receptor interactions.

The case study application to the Toulouse metropolitan area demonstrated the operability of the approach using heterogeneous datasets, highlighting contamination hotspots and supporting decision-making under a stepwise, risk-based logic as recommended by the SML.

Overall, both methodologies provide an operational tool for early-stage prioritisation of contaminated areas, enabling authorities to identify PCSs or areas where further site-specific risk assessment, monitoring efforts or land management measures may be needed. The flexibility ensures that the methodology can be adapted to different regulatory contexts, datasets, and planning needs. Future improvements may refine pathway weighting for diffuse contamination, incorporate Contaminants of Emerging Concerns in the PCSs hazard evaluation and new soil health indicators, and strengthen alignment with expanding national soil monitoring frameworks.

AWAITING APPROVAL BY THE EUROPEAN COMMISSION

Annex 1. The activity–pollutants BRGM matrix – List of hazardous activities and the associated contaminating substances

As already introduced, one of the main challenges for the potentially contaminated site hazard estimation is to identify the substances that are potentially released by a specific industrial or economic activity. The identification and collection of contaminants which are potentially release by a specific economic activity has been already performed by the Bureau de Recherches Géologiques et Minières (BRGM, the French geological service, partner in ISLANDR project). BRGM developed an extensive matrix providing a list of hazardous activities and industries with the associated contaminating substances that can be emitted/released.

To develop this matrix, BRGM in 2014 carried out a huge work of revision of potential contaminated sites at national level in the context of a new soil pollution management policy guided by the Direction General de la Prevention des Risques du Ministère de l'Ecologie. They analysed the existent national databases, and, through statistic treatments, they identified the substances most widely used by industries and activities. In 2018 they published a second version containing a comparison with other information sources, obtaining a wider number of activity-substances couples. Finally, the last revision of the matrix has been released in 2022.

The main outcome of this work was a matrix containing the correspondence among activities and the chemical substances known as potentially dangerous for human health and the environment and that can be spread in the different media (soil, soil gases and water).

The starting point of the BRGM work is a simplified activity-contaminant database elaborated in 1995 and published in 2007 in which 70 substances are correlated with 290 activities. Then, a new correlation matrix has been implemented in 2014 by integrating the main existing national datasets since 90's. The main used databases are here explained:

- BASIAS, the national database for the regional historical inventories of industrial sites and activities. managed by BRGM. It contains 400.000 activities in 262.000 sites and 580.000 references of associated products. Some sites include more than one activity, and all the relative substances are taken into account;
- BASOL is the French national database for contaminated sites and soils or that need monitoring or remediation actions. It includes 3700 sites with 84 activities and the identification in both soil and groundwater of 25 substances (metals and organic);
- ETS is the database regarding the sensitive sites, such as schools located on ancient industrial sites.
- ADES is the French national bank regarding the groundwater quality, including qualitative and quantitative analysis. 25000 sampling sites are scattered among 3872 sites reaching 4,7 million analysis. More than 300 substances produced by 320 different activities are detected;
- IREP, the French register for contaminants emissions in 3500 industrial sites. Data are detected by installations for 100 water contaminants, 50 air contaminants and 400 waste disposals. This is a useful instrument for the population as well as to monitoring the quality of their living environment;

- ARIA is a register for accidents from factories, deposits, construction sites or animal husbandry and so on in France and abroad.

In 2017 the national datasets were integrated by the inclusion of data from other sources, including European and international data:

- the RSDE (research and reduction of dangerous substances released in the water) is a national action aimed to a monitoring and quantification of the substances released in discharged waters by industries, based on the analysis of 106 substances measured in 3000 waste waters samples. The results were used to redact the annex 1 of 2009 legislative decree;
- a table by INERIS containing more than 300 polluting substances and describing the production and use practices in industries;
- the BREF (Best available techniques reference documents), the new European directive for the industrial emissions;
- Other international documents, such as the Guidance for the Safe Development of Housing on Land Affected by Contamination (UK Environment Agency, 2000), the list of potential contaminants by industrial and commercial sectors affecting soil and groundwater (Canada), the Hazardous Activities and Industries List (HAIL) contaminants (New-Zealand).

With all this information, they first had to harmonize the data coming from different sources. In particular, the 345 revised activities have been codified according to the national statistic nomenclature, the NAF (Nomenclature des Activités Français), which include 5 different levels of detail. The 3501 substances were identified with their formula, the SANDRE code and their CAS number. The SANDRE code is the French reference for the chemical substances' codification, while the CAS number is an international registration number. In the SANDRE code, the substances are then divided into groups and sub-groups according to their use and production into the industrial context and these are reported into the matrix. The final step consisted of the creation of a correlation matrix among activities and substances. Indicators of the possible presence/absence of a substance in relation to an activity is based on the frequency correlation in the reviewed database and range between 1 and 8.

For adapting the BRGM matrix to a practical use for the RRA methodology, we modified and adapted it through Microsoft Office Applications. In Virtual Basic we were able to filter the matrix and to maintain only the activity-substance couples with index higher or equal to 6, meaning that they have the highest probability to occur in association with certain activities. We further skimmed the matrix in the following steps, effectively maintaining only the correlation indexes of 7 and 8, meaning that a certain correlation activity-substance is found not only in French datasets but in international documents as well, and we kept the correlation of 6 only when 7 and 8 were not available. Then the NAF industry codes should be correlated to the NACE codes in order to have a common international reference for the activities' classification, and information about the hazard of each substance according to CLP regulation is added for each substance through the CAS number. The final matrix contains 477 substances and 260 activities.

Bibliography

Albanese, S., De Vivo, B., Lima, A., Frattasio, G., Křibek, B., Nyambe, I., & Majer, V. (2014). Prioritising environmental risk at the regional scale by a GIS aided technique: The Zambian Copperbelt Province case study. *Journal of Geochemical Exploration*, 144, 433–442. <https://doi.org/10.1016/j.gexplo.2014.02.004>

APAT – Agenzia per la Protezione dell'Ambiente e per i Servizi Tecnici. (2008). Criteri metodologici per l'applicazione dell'analisi assoluta di rischio ai siti contaminati (Revisione 2, 156 pp.). Roma: APAT. <http://www.isprambiente.gov.it/files/temi/siti-contaminati-02marzo08.pdf>

Bardos, R. P., Thomas, H. F., Smith, J. W. N., Harries, N. D., Evans, F., Boyle, R., Howard, T., Lewis, R., Thomas, A. O., & Haslam, A. (2018). The Development and Use of Sustainability Criteria in SuRF-UK's Sustainable Remediation Framework. *Sustainability*, 10(6), 1781. <https://doi.org/10.3390/su10061781>

Bardos, P., Kastrati, S., & Baraku, T. (2019). Kosovo: Healing land for the future.

Beck, N. G., Conley, G., Kanner, L., & Mathias, M. (2017). An urban runoff model designed to inform stormwater management decisions. *Journal of Environmental Management*, 193, 257–269. <https://doi.org/10.1016/j.jenvman.2017.02.007>

Belbèze, S., Rohmer, J., Négrel, P., & Guyonnet, D. (2023). Defining urban soil geochemical backgrounds: A review for application to the French context. *Journal of Geochemical Exploration*, 254, 107298. <https://doi.org/10.1016/j.gexplo.2023.107298>

Belbèze, S., Djemil, M., Béranger, S., & Stochetti, A. (2019). Détermination de FPGA - Fonds Pédogéochimiques Anthropisés urbains. Agglomération pilote: Toulouse Métropole (public No. RP-69502-FR; p. 347). BRGM.

Braud, I., Desprats, J.-F., Ayrat, P.-A., Bouvier, C., & Vandervaere, J.-P. (2017). Mapping topsoil field-saturated hydraulic conductivity from point measurements. *Journal of Hydrology and Hydromechanics*, 65(3), 264–275. <https://doi.org/10.1515/johh-2017-0017>

BRGM. (2022). CAPacité d'infiltration des eaux pluviales du territoire de Toulouse Métropole (CAPITOUL): Rapport final, Version 1, 24 mai 2022. BRGM/RP71904-FR.

Canadian Council of Ministers of the Environment (CCME). (1997, updated periodically). A protocol for the derivation of environmental and human health soil quality guidelines. CCME.

Cheng, Z., Hettiarachchi, G., & Kim, K. (2021). Urban soils research: Suitma 10. *Journal of Environmental Quality*, 50(1), 2–6. <https://doi.org/10.1002/jeq2.20191>

Chen, M., Cai, H., Wang, L., & Lei, M. (2022). Grid-scale regional risk assessment of potentially toxic metals using multi-source data. *ISPRS International Journal of Geo-Information*, 11(8), 427. <https://doi.org/10.3390/ijgi11080427>

Csavina, J., Landazuri, A., Wonaschutz, A., Rine, K., Rheinheimer, P., Barbaris, B., Conant, W., Saez, A. E., & Betterton, E. A. (2011). Metal and metalloid contaminants in atmospheric aerosols from mining

operations. *Water, Air, and Soil Pollution*, 221(1–4), 145–157. <https://doi.org/10.1007/s11270-011-0777-x>

Davis, P. A., Sheppard, M. I., & Andres, T. H. (1993). A simple model to predict the long-term fate of contaminants in unsaturated soils. *Waste Management*, 13, 25–40.

DEFRA – Department for Environment, Food and Rural Affairs. (2011). Guidelines for environmental risk assessment and management: Green Leaves III. Defra & Cranfield University, U.K.

DoE – Department of Environment. (1995). A guide to risk assessment and risk management for environmental protection. London: HSMO.

Dragović, R., Gajić, B., Dragović, S., Đorđević, M., Đorđević, M., Mihailović, N., & Onjia, A. (2014). Assessment of the impact of geographical factors on the spatial distribution of heavy metals in soils around the steel production facility in Smederevo (Serbia). *Journal of Cleaner Production*, 84, 550–562. <https://doi.org/10.1016/j.jclepro.2014.03.060>

Duvernoy, I., Zambon, I., Sateriano, A., & Salvati, L. (2018). Pictures from the other side of the fringe: Urban growth and peri-urban agriculture in a post-industrial city (Toulouse, France). *Journal of Rural Studies*, 57, 25–35. <https://doi.org/10.1016/j.jrurstud.2017.10.007>
Dijkstra, J. J., Meeussen, J. C. L., & Comans, R. N. J. (2004). Leaching of heavy metals from contaminated soils: An experimental and modeling study. *Environmental Science & Technology*, 38, 4390–4395.

ECHA (2017). Guidance on the Application of the CLP Criteria. European Chemicals Agency.

ECHA (2016). Guidance on Information Requirements and Chemical Safety Assessment. Chapter R.14: Occupational Exposure Estimation.

EEA. (2005). Provisional Risk Assessment Methodology for the Prioritisation of Contaminated Sites (PRA.MS). European Environment Agency.

EFSA (2013). Guidance on the risk assessment of plant protection products on bees. European Food Safety Authority.

European Commission. (2006). Thematic strategy for soil protection (COM(2006) 231). Brussels: European Commission.

European Commission (2008). Regulation (EC) No 1272/2008 on classification, labelling and packaging of substances and mixtures (CLP). Official Journal of the European Union.

European Commission. (2021). Communication ... EU soil strategy for 2030: Reaping the benefits of healthy soils for people, food, nature and climate (COM(2021) 699 final). Brussels: European Commission.

European Commission. (2023). Proposal for a directive ... Soil monitoring and resilience (Soil Monitoring Law) (COM(2023) 416 final). Brussels: European Commission.

Fabian, K., Reimann, C., & De Caritat, P. (2017). Quantifying diffuse contamination: method and application to Pb in soil. *Environmental science & technology*, 51(12), 6719-6726.

Food and Agriculture Organization of the United Nations (FAO). (2018). Global Symposium on Soil Pollution. FAO.

Flem, B., Reimann, C., & Fabian, K. (2022). Excess Cr and Ni in top soil: Comparing the effect of geology, diffuse contamination, and biogenic influence. *Science of the Total Environment*, 843, 157059. <https://doi.org/10.1016/j.scitotenv.2022.157059>

Gan, C., Tang, Y., Chen, L., & Yang, J. (2023). Assessment of metal pollution in agricultural soils and associated human health risk in Shifang City, China. *Clean: Soil, Air, Water*, 51(11). <https://doi.org/10.1002/clen.202300072>

Health Canada. (2010a). Federal contaminated site risk assessment in Canada. Part I: Guidance on human health preliminary quantitative risk assessment (PQRA) (Version 2.0). Ottawa, ON: Health Canada.

Health Canada. (2010b). Federal contaminated site risk assessment in Canada. Part II: Health Canada toxicological reference values (TRVs) and chemical-specific factors (Version 2.0). Ottawa, ON: Health Canada.

Holdgate, M. W. (1979–1980). A perspective of environmental pollution. Cambridge: Cambridge University Press. <https://doi.org/10.1177/004711788000600518>

Institut national de la statistique et des études économiques – INSEE. (2023). Annual report 2023. Paris: INSEE.

Institut national de la statistique et des études économiques - INSEE (2023). Population en 2020 – Recensement de la population – Base infracommunale (IRIS). Institut national de la statistique et des études économiques. Available at: <https://www.insee.fr/fr/statistiques/7704076>

International Organization for Standardization (ISO). (2015). ISO 14015: Environmental management — Environmental assessment of sites and organizations (EASO). Geneva: ISO.

International Organization for Standardization (ISO). (2018). Soil quality – Guidance on the determination of background values (NF EN ISO 19258:2018(E)). Geneva: ISO.

ISO (2018). Soil quality — Guidance on the determination of background values. International Organization for Standardization.

Jiang, Y., Guo, X., Ye, Y., Xu, Z., Zhou, Y., Xia, F., & Shi, Z. (2024). Spatiotemporal assessment and scenario simulation of the risk potential of industrial sites at the regional scale. *Science of The Total Environment*, 906, 167537. <https://doi.org/10.1016/j.scitotenv.2023.167537>

Pérez, A. P., & Rodriguez Eugenio, N. (2018). Status of local soil contamination in Europe (JRC Technical Report). Luxembourg: Publications Office of the European Union. <https://doi.org/10.2760/093804>

Konadu, F. N., Gyamfi, O., Ansah, E., Sheringham Borquaye, L., Agyei, V., Dartey, E., Dodd, M., Obiri-Yeboah, S., Darko, G. (2023). Human health risk assessment of potentially toxic elements in soil and air particulate matter of automobile hub environments in Kumasi, Ghana. *Toxicology Reports*, 11, 261–269. <https://doi.org/10.1016/j.toxrep.2023.09.010>

Landis, W. G. (2005). Regional scale ecological risk assessment: Using the relative risk model. CRC Press.

Landis, W. G., & Wieggers, J. K. (1997). Design considerations and a suggested approach for regional and comparative ecological risk assessment. *Human and Ecological Risk Assessment*, 3(3), 287–297.

Lerner, D. N., & Harris, B. (2009). The relationship between land use and groundwater resources and quality. *Land Use Policy*, 26S, S265–S273.

Li, D., Zhang, C., Pizzol, L., Critto, A., Zhang, H., Lv, S., & Marcomini, A. (2014). Regional risk assessment approaches to land planning for industrial polluted areas in China: The Hulunbeier region case study. *Environment International*, 65, 16–32.

Li, K., Wang, J., & Zhang, Y. (2022). Heavy metal pollution risk of cultivated land from industrial production in China: Spatial pattern and its enlightenment. *Science of the Total Environment*, 828, 154382. <https://doi.org/10.1016/j.scitotenv.2022.154382>

Limasset, E., Pizzol, L., Merly, C., Gatchett, A. M., Le Guern, C., Martinát, S., Klusáček, P. & Bartke, S. (2018). Points of attention in designing tools for regional brownfield prioritization. *Science of the Total Environment*, 622, 997–1008.

Minolfi, G., Albanese, S., Lima, A., Tarvainen, T., Fortelli, A., & De Vivo, B. (2018). A regional approach to the environmental risk assessment: Human health risk assessment case study in the Campania region. *Journal of Geochemical Exploration*, 184(B), 400–416. <https://doi.org/10.1016/j.gexplo.2016.12.010>

Moskal, B., & Berthrong, S. (2018). Novel soil barrier systems potentially protect urban growing beds from legacy soil contamination and improve soil health. *Urban Agriculture & Regional Food Systems*, 3(1), 1–9. <https://doi.org/10.2134/urbanag2018.06.0003>

Muséum national d'Histoire naturelle (MNHN). (2024). Réseau des Réserves Naturelles Nationales (RNN) [Shapefile].

National Research Council (NRC). (2009). Science and decisions: Advancing risk assessment. Washington, DC: National Academies Press. http://www.nap.edu/catalog.php?record_id=12209

Observatoire Régional de l'Environnement Occitanie (OREO) / OpenIG. (2022). Espaces Naturels Sensibles (ENS) Occitanie [Shapefile].

Petrik, A., Thiombane, M., Lima, A., Albanese, S., Buscher, J. T., & De Vivo, B. (2018). Soil contamination compositional index: A new approach to quantify contamination demonstrated by assessing compositional source patterns of potentially toxic elements in the Campania Region (Italy). *Applied Geochemistry*, 96, 264–276. <https://doi.org/10.1016/j.apgeochem.2018.07.014>

Pizzol, L., Zabeo, A., Klusáček, P., Giubilato, E., Critto, A., Frantál, B., Martinát, S., Kunc, J., Osman, R. & Bartke, S. (2016). Timbre Brownfield Prioritization Tool to support effective brownfield regeneration. *Journal of environmental management*, 166, 178-192.

Pizzol, L., Critto, A., Agostini, P., & Marcomini, A. (2011). Regional risk assessment for contaminated sites Part 2: Ranking of potentially contaminated sites. *Environment international*, 37(8), 1307-1320.

Pouyat, R., Szlávecz, K., Yesilonis, I., Wong, C., Murawski, L., Marra, P., Casey, R. E., & Lev, S. (2015). Multi-scale assessment of metal contamination in residential soil and soil fauna: A case study in the Baltimore–Washington metropolitan region, USA. *Landscape and Urban Planning*, 142, 7–17. <https://doi.org/10.1016/j.landurbplan.2015.05.001>

Prabhakar, G., Sorooshian, A., Toffol, E., Arellano, A. F., & Betterton, E. A. (2014). Spatiotemporal distribution of airborne particulate metals and metalloids in a populated arid region. *Atmospheric Environment*, 92, 339–347. <https://doi.org/10.1016/j.atmosenv.2014.04.044>

Punia, A. (2021). Role of temperature, wind, and precipitation in heavy metal contamination at copper mines: A review. *Environmental Science and Pollution Research*, 28, 4056–4072. <https://doi.org/10.1007/s11356-020-11580-8>

Ramón, F., & Lull, C. (2019). Legal measures to prevent and manage soil contamination and to increase food safety for consumer health: The case of Spain. *Environmental Pollution*, 250, 883–891. <https://doi.org/10.1016/j.envpol.2019.04.074>

Région Occitanie. (2024). Réserves Naturelles Régionales (RNR) [Dataset].

Reimann, C., Fabian, K., & Flem, B. (2019). Cadmium enrichment in topsoil: Separating diffuse contamination from biosphere-circulation signals. *Science of the Total Environment*, 651, 1344–1355. <https://doi.org/10.1016/j.scitotenv.2018.09.272>

Sahoo, P. K., Dall’Agnol, R., Salomão, G. N., Ferreira Junior, J. da S., Silva, M. S., Souza Filho, P. W. M., ... Siqueira, J. O. (2020). Regional-scale mapping for determining geochemical background values in soils of the Itacaiúnas River Basin, Brazil: The use of compositional data analysis (CoDA). *Geoderma*, 376, 114504. <https://doi.org/10.1016/j.geoderma.2020.114504>

Sandre. (2025). Atlas des masses d’eau souterraines [Data set]. Service d’Administration Nationale des Données et Référentiels sur l’Eau. Retrieved from <https://www.sandre.eaufrance.fr/atlas>

Solismaa, S., Loukola-Ruskeeniemi, K., Nuottimäki, K., Tolvanen, H., Järvinen, K. & Müller, I. (2025.) Historical mining towns: The establishment of ‘Soil Planning Areas’ for the risk management of contaminated soil. *Journal of Hazardous Materials* 486 (2025) 136962 <https://doi.org/10.1016/j.jhazmat.2024.136962>

Su, C., Yang, Y., Lei, M., Li, H. A., Dang, L., & Wu, P. (2025). Large-Scale Spatiotemporal Risk Assessment Method Based On Source–Pathway–Receptors: A 30-Year Case Study of Nonferrous Metal Industrial Sites. *Environmental Science & Technology*.

Sun, D., Yang, H., Guan, D., Yang, M., Wu, J., Yuan, F., ... Zhang, Y. (2018). The effects of land use change on soil infiltration capacity in China: A meta-analysis. *Science of the Total Environment*, 626, 1394–1401. <https://doi.org/10.1016/j.scitotenv.2018.01.104>

Tarvainen, T., Lehtonen, M., Lahaye, Y. & Jarva, J. (2023.) Analytical workflow to trace lead sources in fill-derived soils in Turku, Southwest Finland. *Applied Geochemistry* 156, 105735. <https://doi.org/10.1016/j.apgeochem.2023.105735>

Tarvainen, T., Lehtonen, M., Lahaye, Y., Jarva, J., Belbeze, S. & Nuottimäki, K. (2024). Urban anthropogenic soils - potential urban diffuse soil contamination. *Book of Abstracts. Joint Conference of ISEH ICEPH & ISEG on Environment and Health. University of Galway, Ireland, August 11 - 18, 2024.* <https://www.universityofgalway.ie/media/ineh/files/ISEH-ICEPH-ISEG-Book-of-Abstracts.pdf>

Tenailleau, Q., Lanier, C., Prud'homme, J., Cuny, D., Deram, A., & Ocelli, F. (2024). Distance-based indicators for evaluating environmental multi-contamination and related exposure: how far should we go?. *Environmental Science and Pollution Research*, 31(38), 50642-50653.

United Nations Environment Programme (UNEP). (2019). Guidelines for the identification and management of contaminated sites. UNEP.

U.S. Environmental Protection Agency (USEPA). (1991). Risk assessment guidance for Superfund: Volume I — Human health evaluation manual (Part B, Development of risk-based preliminary remediation goals). Washington, DC: USEPA.

U.S. Environmental Protection Agency (USEPA). (2005). Guidelines for carcinogen risk assessment. Washington, DC: USEPA.

U.S. Environmental Protection Agency (USEPA). (2013). Our built and natural environments: A technical review of the interactions between land use, transportation, and environmental quality (2nd ed.). EPA 231-K-13-001. Washington, DC: USEPA.

U.S. Environmental Protection Agency (USEPA). (2014). Framework for human health risk assessment to inform decision making. <https://www.epa.gov/risk/framework-human-health-risk-assessment-inform-decision-making>

U.S. Environmental Protection Agency (USEPA). (2020). Superfund: CERCLA overview. Washington, DC: USEPA.

U.S. Environmental Protection Agency (USEPA). (2021). Risk assessment guidance for Superfund (RAGS). Washington, DC: USEPA.

Van Gestel, G., Dedecker, D., & Ceenaeme, J. (2022). Diffuse soil contamination: Inventory of data sources and proposed approach.

Zabeo, A., Pizzol, L., Agostini, P., Critto, A., Giove, S., & Marcomini, A. (2011). Regional risk assessment for contaminated sites Part 1: Vulnerability assessment by multicriteria decision analysis. *Environment international*, 37(8), 1295-1306.



WHO (2001). Principles for the Assessment of Risks to Human Health from Exposure to Chemicals.

World Bank Group, & Technical University of Denmark (DTU). (2023). Global Wind Atlas 3.0 [Data set]. Retrieved from <https://globalwindatlas.info>

World Health Organization (WHO). (2008). Guidelines for assessing health risks from contaminated sites. Geneva: WHO.

AWAITING APPROVAL BY THE EUROPEAN COMMISSION

