Politecnico di Torino

Master's degree in environmental and Land Engineering

"Climate Change"

A.Y. 2025/2026

Integrated groundwater vulnerability assessment in fractured granitic media: coupling GIS indices with LiDAR-derived vegetation metrics

A case study of the Caldas da Cavaca system, central Portugal

Advisors:
Prof. Ganora Daniele
Co-advisors:
Prof. Meixedo João Paulo
Prof. Teixeira José Augusto Alves

Candidate:
Salaorni Tommaso

Index

1 Introduction

Groundwater is a critical freshwater resource, supplying drinking water, agricultural irrigation, and industrial activities globally. However, the sustainability of this resource is increasingly threatened by a combination of intensive anthropogenic activities and climate variability. In Southern Europe, and specifically in the Mediterranean region, aquifers face a dual challenge: they are subject to high extraction rates during seasonal drought periods while simultaneously being exposed to diffuse pollution sources, primarily from intensive agriculture.

The protection of these groundwater systems requires vulnerability assessment tools that can identify areas most susceptible to contamination. While traditional hydrogeological mapping provides a fundamental understanding of aquifer geometry, it often lacks the resolution required to characterize the complex spatial heterogeneity of hard-rock terrains. In granitic environments, groundwater circulation is strictly controlled by secondary porosity features, such as fractures and weathering profiles (Scesi and Gattinoni 2007), that are difficult to map using conventional cartography alone. Consequently, there is a need to integrate established vulnerability indices with high-resolution remote sensing technologies to refine the assessment of groundwater reservoirs.

The research focuses on the region surrounding Caldas da Cavaca in Central Portugal. This area represents a complex hydrogeological setting characterized by a fractured granitic basement and a shallow phreatic aquifer system developed within the weathered mantle (saprolite) and alluvial valleys (Meerkhan, et al. 2016).

The primary objective of this thesis is to assess the groundwater vulnerability of this region by conducting a comparative analysis of intrinsic and specific vulnerability models. Furthermore, a novel component of this research is the exploratory application of LiDAR-derived vegetation metrics. In hard-rock terrains, traditional vulnerability mapping (e.g., DRASTIC) relies heavily on interpolated geological data which may miss local structural nuances.

This study hypothesizes that in a water-limited Mediterranean environment, the spatial distribution of arboreal biomass—specifically canopy height and density—serves as a surface bio-indicator for subsurface conditions. We posit that anomalous tree growth correlates with

areas of deeper weathering (thicker regolith) and higher fracture density, which are critical factors for groundwater recharge. Therefore, LiDAR analysis is presented here not as a parameter within the indices, but as an independent, non-invasive line of evidence to validate the hydrogeological conceptual model.

1.1 Study area: Caldas da Cavaca

Caldas da Cavaca is a hard-rock hydrogeological system located in the municipality of Aguiar da Beira (Guarda district), Central Portugal, within the Beiras Variscan granitic belt and the Dão granitic complex, close to the Bragança-Vilaricá-Manteigas fault zone (≈40°46' N, 7°34' W). The climate is temperate-humid, corresponding to a Köppen-Geiger Cfb climate (Teixeira, Chaminé, et al., A comprehensive analysis of groundwater resources using GIS and multicriteria tools (Caldas da Cavaca, Central Portugal): environmental issues 2014), with a mean annual temperature of 13 °C and average precipitation of ≈ 1 250 mm yr^{-1}, the wettest month being January and a marked summer drought. (Teixeira, Chaminé, et al. 2014) (Meerkhan, et al. 2016).

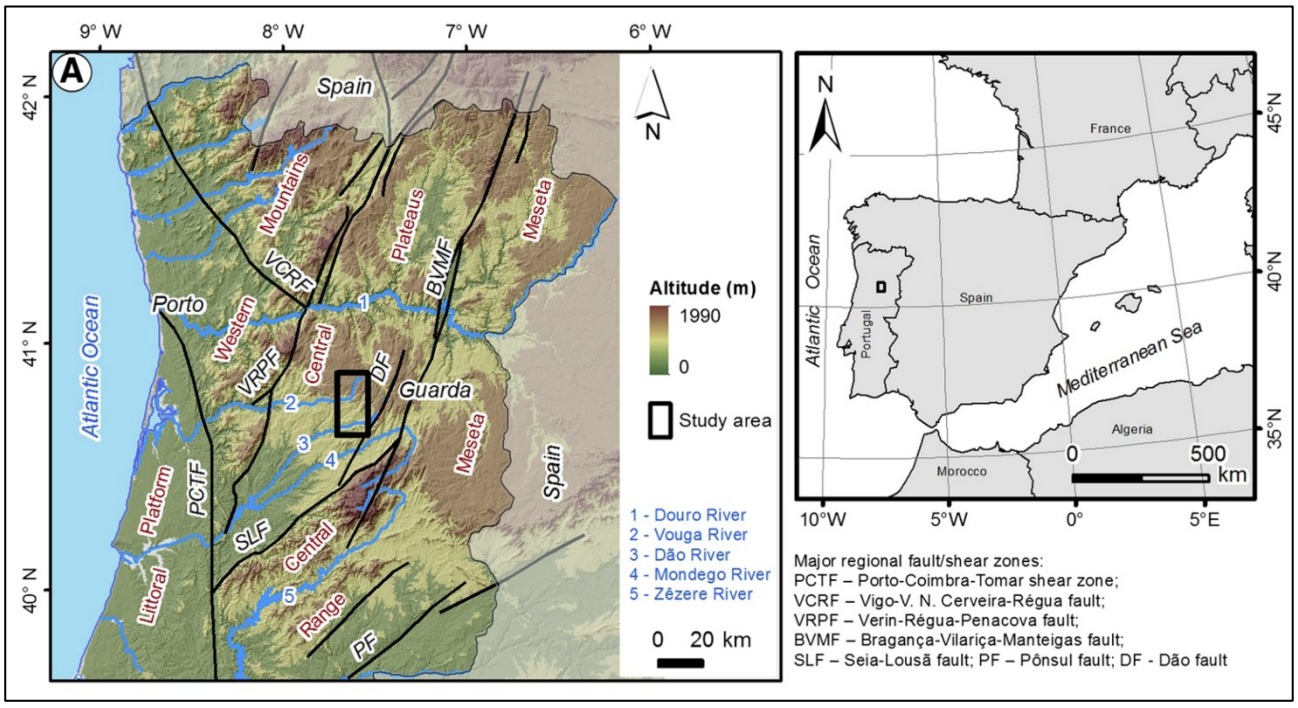

Figure 1: Regional framework of the Caldas da Cavaca hydromineral system, Aguiar da Beira municipality (Teixeira, Chaminé, et al. 2014)

Geologically the area consists of coarse-grained porphyritic granite, alluvial deposits and doleritic dykes; the granite is heavily weathered (up to 50 m depth) and extensively fractured,

especially along the Ribeira de Coja fault zone (Teixeira, Chaminé, et al., A comprehensive analysis of groundwater resources using GIS and multicriteria tools (Caldas da Cavaca, Central Portugal): environmental issues 2014) (Meerkhan, et al. 2016).

Hydrogeologically, three aquifer types are recognised:

- a shallow unconfined aquifer in the alluvia cover (pH 5–6.5, EC < 20 µS cm^{-1}) (Meerkhan, et al. 2016);
- a weathered-to-fractured granite aquifer (pH 5–6.5, EC 20–50 µS cm^{-1}) (Teixeira, Chaminé, et al., A comprehensive analysis of groundwater resources using GIS and multicriteria tools (Caldas da Cavaca, Central Portugal): environmental issues 2014) (Meerkhan, et al. 2016);
- a deep confined hydromineral aquifer hosted in the fault-controlled granite (temperature ≈ 29.8 °C, pH 8.3-8.6, silica ≈ 55 mg L^{-1}, fluoride up to 14 mg L^{-1}, EC 353-427 µS cm^{-1}, TDS 262-272 mg L^{-1}) (Meerkhan, et al. 2016).

Ground-water resources support thermal spa activities, mineral-water bottling and local irrigation. High infiltration potential occurs in the NW and SE strips where weathered granite, low slopes and dense fracture networks coincide. Intrinsic vulnerability mapping (GOD-S, DRASTIC-Fm, SINTACS, SI, and the DISCO index) shows the shallow aquifer to be high-to-very-high vulnerable, while the deep hydromineral reservoir is low-vulnerable; DISCO highlights the primary role of first-order lineaments in controlling contamination pathways (Meerkhan, et al. 2016) (Teixeira, Chaminé, et al. 2014).

Based on these results, Portuguese law-mandated wellhead protection zones have been revised, expanding intermediate zones to encompass the most vulnerable fault-controlled areas and restricting immediate zones around the new mineral-water wells (Meerkhan, et al. 2016) (Teixeira, Chaminé, et al. 2014).

1.2 State of the art

1.2.1 The Foundation of Vulnerability Mapping: The DRASTIC Model

The concept of groundwater vulnerability mapping was formalized to provide a standardized method for evaluating the potential of aquifers to transmit contaminants from the surface. The

foundational framework for this field is the DRASTIC system, developed for the U.S. Environmental Protection Agency (Aller, Lehr and Bennet 1987).

Aller et al. defined vulnerability as an intrinsic property of the hydrogeological system, independent of the specific nature of the pollutant. Their model established the standard seven hydrogeological parameters (namely Depth to Water, Net Recharge, Aquifer Media, Soil Media, Topography, Impact of the Vadose Zone, and Hydraulic Conductivity) that act as the primary filters controlling contaminant transport. This standardized weighting system remains the global benchmark for intrinsic vulnerability assessment (Aller, Lehr and Bennet 1987).

1.2.2 The Shift to Specific Vulnerability

While DRASTIC effectively characterizes natural defence mechanisms, subsequent research highlighted its limitations in agricultural regions where land management plays a dominant role in contamination. To address this, Ribeiro (2000) proposed the Susceptibility Index (SI), a modification of the DRASTIC method designed specifically for assessing vulnerability to agricultural pollution, particularly nitrates.

The distinction between these approaches was rigorously tested by Stigter, Ribeiro, and Carvalho Dill (2006) in the south of Portugal. Their comparative study demonstrated that while DRASTIC successfully identifies hydro geologically sensitive zones, it often underestimates the risk in areas with intensive land use. By introducing a "Land Use" parameter and removing purely hydraulic factors (like Hydraulic Conductivity), the SI provided a higher correlation with observed nitrate concentrations in groundwater. Stigter et al. concluded that for diffuse agricultural pollution, specific vulnerability indices (SI) provide a more realistic risk assessment than intrinsic models (DRASTIC).

1.2.3 Hydrogeology of Fractured Granitic Media

The application of these indices in the Caldas da Cavaca region is complicated by the nature of the geological substrate. Unlike sedimentary basins, groundwater circulation in crystalline rock masses is governed by secondary porosity. Scesi and Gattinoni (2007) emphasize that in

these "hard-rock" environments, flow is strictly controlled by the discontinuity network (fractures) and the degree of weathering.

The characterization of these rock masses requires specific geotechnical evaluation, as outlined by the ISRM suggested methods (Ulusay 2015).

1.2.4 Previous Research in Caldas da Cavaca

The specific hydrogeomorphology of the Caldas da Cavaca study area has been extensively characterized by Teixeira et al. in multiple studies.

Hydrogeomorphology and GIS: integrated hydrogeomorphological mapping with GIS tools to evaluate the region's water resources. They established the conceptual model of the area, distinguishing between the deep hydromineral circulation and the shallow phreatic systems (Teixeira, Chaminé, et al. 2008).

Environmental Issues: In a subsequent comprehensive analysis (Teixeira, Chaminé, et al. 2014), focused on the environmental pressures facing these resources. Their work highlighted the vulnerability of the shallow aquifers developed within the weathered granitic mantle and alluvial deposits. They documented the conflict between the region's natural hydrogeological fragility and the anthropogenic pressures of agriculture, validating the need for the dual (intrinsic vs. specific) assessment approach adopted in this thesis.

2 Materials and methods

2.1 LiDAR

LiDAR (Light Detection and Ranging) is an active remote sensing method that utilizes a laser to measure distances and determine the precise 3D location of targets (Carter, et al. 2012). The technology operates by emitting pulses of light from a laser scanner and measuring the time it takes for the reflection (echo) to return to the sensor; because the speed of light and the location of the emission are known, the XYZ coordinates of the reflection surface can be calculated (Melin, Shapiro e Glover-Kapfer 2017). Airborne LiDAR systems generally consist of the laser scanner integrated with a Global Positioning System (GPS) or Global Navigation Satellite System (GNSS) to track the aircraft's position, and an Inertial Measurement Unit (IMU) or Inertial Navigation System (INS) to measure the aircraft's attitude (roll, pitch, and yaw) (Melin, Shapiro e Glover-Kapfer 2017) (Wang, et al. 2024).

A key characteristic of modern LiDAR is the ability to record multiple returns from a single laser pulse (Carter, et al. 2012) (Lohani s.d.). As the pulse travels, it can penetrate semi-transparent surfaces like vegetation; the first return typically represents the top of the canopy, while the last return often represents the ground, allowing for the derivation of both forest structure and bare-earth topography (Melin, Shapiro e Glover-Kapfer 2017). In addition to range measurements, LiDAR sensors record the intensity of the returned signal, which indicates the reflectance strength of the target material. The resulting dataset is a high-density, georeferenced point cloud that provides a detailed 3D representation of the scanned area (Carter, et al. 2012).

Most publicly available datasets are discrete-return point clouds, although some systems record full-waveform data that capture the entire back-scattered signal (used for finer vegetation analysis) (Melin, Shapiro e Glover-Kapfer 2017)

After acquisition, the raw points are processed to produce derived raster products:

- Digital Terrain Model (DTM) – interpolated from ground-classified points
- Digital Surface Model (DSM) – generated from the highest returns
- Canopy-Height Model (CHM) – DSM minus DTM
- Intensity rasters, slope/aspect maps, and contour lines

These derivatives are commonly stored as GeoTIFF or ESRI GRID rasters, while vector outputs (e.g., contour lines, classified polygons) use standard shapefile or GeoPackage formats (Carter, et al. 2012).

2.1.1 Source of LiDAR data

The primary dataset utilized in this study consists of high-definition airborne LiDAR data covering mainland Portugal, acquired by the Portuguese government via the Direção-Geral do Território in 2024. The data is georeferenced to the EPSG:3763 coordinate system and was collected across four adjacent North-South plots (Lots 1 through 4), segmented by specific M-coordinate ranges from the south to the north of the territory. The acquisition campaigns employed different sensor configurations depending on the location: Lot 1 was captured using Teledyne Optech Galaxy T2000 and PhaseOne iXU-RS1000 sensors, while Lots 2, 3, and 4 utilized Riegl VQ-780II-S LiDAR sensors paired with PhaseOne iXM-RS 150F and iXM-RS 100 aerial photography units. (Direção-Geral do Território s.d.)

The fundamental component of this dataset is a LAS point cloud exhibiting a high point density of 10 points/m^2. These points are geometrically adjusted with a reported positional accuracy of 30 cm in the planimetric component and 10 cm in the altimetric component. Each point contains rich attribute data, including intensity values normalized to 16 bits and colour attributes spanning the Red, Green, Blue, and Near-Infrared (RGBNir) channels. The cloud is further categorized into nine specific classes: Other, Terrain, Low Vegetation (0–0.5 m), Medium Vegetation (0.5–2 m), Tall Vegetation (>2 m), Buildings, Noise, Water, and Bridges (Direção-Geral do Território s.d.).

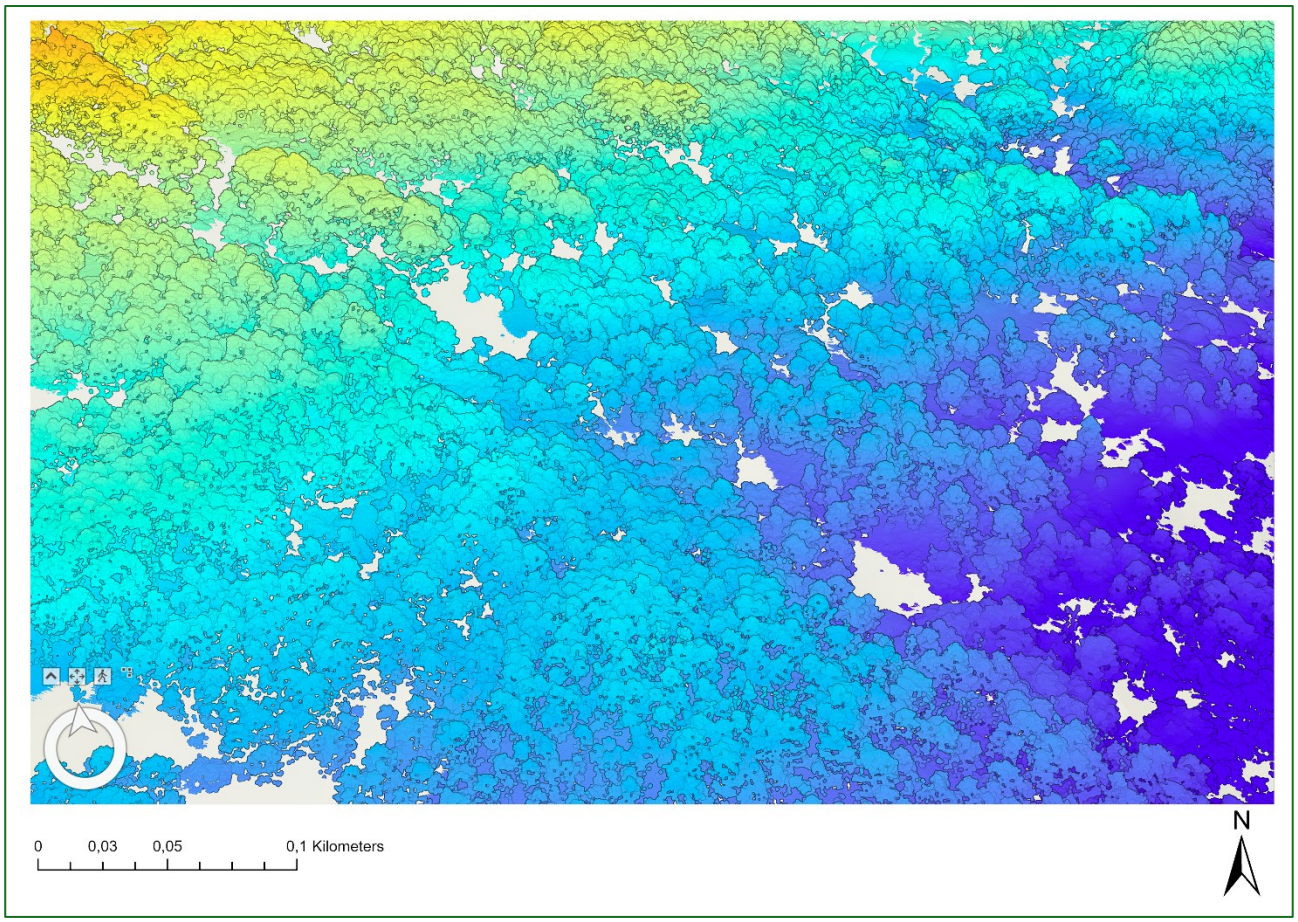

Figure 2: close up of the study area's point cloud

2.2 Geological data

The geological data was mainly attained from previous studies and surveys of the area.

2.2.1.1

2.2.2 Overview

The assessment of soil and groundwater vulnerability to contaminants requires a robust hydrogeological and geotechnical foundation. This chapter outlines the geological data used to characterize the study area, specifically the Caldas da Cavaca hydromineral system. The primary geological and hydrogeological datasets utilized in this research were not gathered during the current scope of work. Instead, this thesis leverages the extensive high-resolution datasets generated by prior investigations into the region's fractured rock masses and hydromineral resources.

The data used here were originally compiled through a multi-technical approach developed to support hydrogeological conceptualization and groundwater protection mapping. These studies provided a rigorously validated geodatabase that integrates cartographic information, remote sensing analysis, and site-specific field data (Ulusay 2015) (Teixeira, Chaminé, et al. 2008) (Teixeira, Chaminé, et al. 2014).

2.2.3 Hydrogeological field inventory and characterization

The inventory process involved a survey of all accessible water points within the study area, encompassing natural springs, dug wells, drilled boreholes, and horizontal water galleries. Each feature was georeferenced using GPS equipment (specifically a Trimble GeoExplorer) (Teixeira, Chaminé, et al. 2014).

At each identified water point, in-situ physicochemical characterization was conducted, using portable multiparametric equipment. Key parameters measured included water temperature, electrical conductivity (EC), and pH. These specific metrics were chosen to differentiate between shallow, unconfined circulation and deeper hydromineral pathways. For instance, low electrical conductivity (typically below 50 µS/cm) and temperatures correlated with ambient air conditions were used to identify shallow circulation in the weathered granite, whereas higher temperatures (near 30°C) and mineralization were indicators of deep-seated hydromineral fluids rising through major fault zones.

Furthermore, the inventory included the measurement of yield and piezometric levels where possible, providing data on the hydraulic behaviour of the system (Teixeira, Chaminé, et al. 2008) (Teixeira, Chaminé, et al. 2014) (Meerkhan, et al. 2016).

2.2.4 Geological and geotechnical characterization of rock masses

Because the movement of contaminants is governed almost entirely by the physical state of the rock, specifically its weathering grade and the density of its fractures, the geotechnical study of the region is of great importance. The type of lithology, structure and weathering grade implies different hydraulic conductivity, transmissivity and storage coefficient for diverse geological formations and affects the groundwater infiltration. In fact, the influence of fractures

in the propagation of fluids and contaminants is twofold: they may act as hydraulic systems by providing preferential flow paths, or as barriers that prevent flow across them. This, in turn, contributes to the transportation of the contaminants across and to the subsurface (Meerkhan, et al. 2016) (Scesi and Gattinoni 2007).

In this case study, the dominant granitic rocks and mafic dykes show varied weathering grades, ranging from fresh to highly weathered according to the ISRM classification, and the fracturing intercept degree is dominantly moderate to close. Furthermore, there is a small area of very low thickness made up of alluvium deposits (Teixeira, Chaminé, et al. 2008).

The characterization was primarily based on the standards set by the International Society for Rock Mechanics (ISRM), which allows for a consistent classification of the granitic bedrock into distinct units based on their hydraulic properties. The data gathering process segmented the rock mass into three main categories of weathering, which directly influence infiltration and vulnerability:

- Fresh to Slightly Weathered Granite (W_{1-2}): found usually at higher altitudes (600–700 m), these rocks are characterized by moderate (F3) to close (F4-5) fracturing degrees. They often form morphological highs and outcrops, acting as areas of lower permeability compared to the weathered valleys;
- Moderately Weathered Granite (W_3): located typically at intermediate altitudes (500–650 m), these zones appear as wide corridors aligned with major NE-SW tectonic trends;
- Highly Weathered Granite (W_{4-5}): typical of plateau areas and specific corridors, these rocks exhibit intense arenization (decomposition into sandy material), which significantly increases their storage capacity and infiltration potential. This weathering can reach depths of up to 50 meters in fault zones (Ulusay 2015) (Teixeira, Chaminé, et al. 2008) (Teixeira, Chaminé, et al. 2014) (Stigter, Ribeiro and Carvalho Dill 2005).

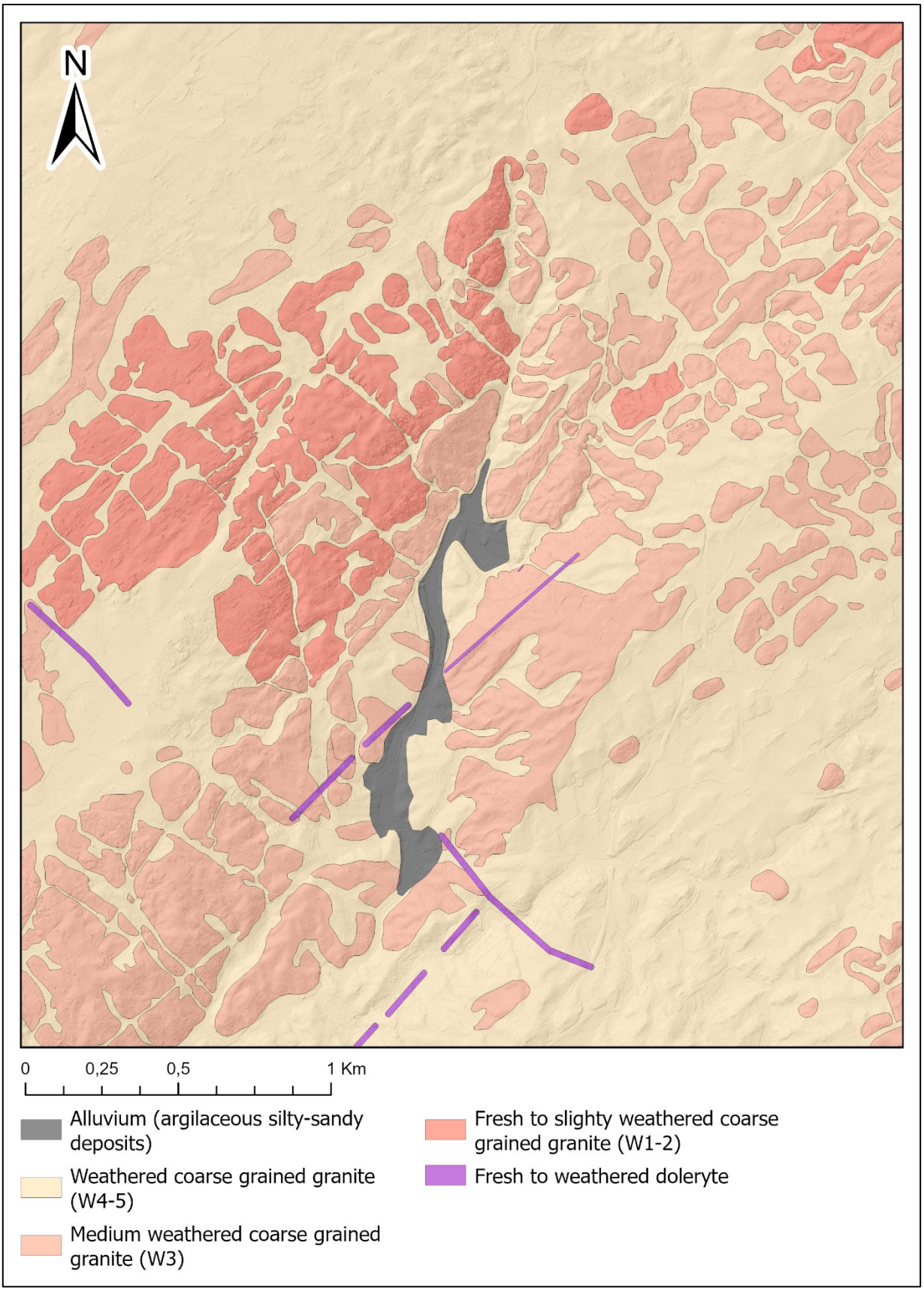

Figure 3: geological composition of the study area, categorised through the ISRM weathering scale

Tectonic lineaments were defined in the source studies as the simple and complex linear properties of geological structures, such as faults, fractures, joints, and discontinuity surfaces, arranged in a straight or slightly curved line pattern (O'Leary, Friedman and Pohn 1976). They were mapped using a combination of remote sensing followed by verification through geological maps and field work to eliminate interpretation errors from satellite and aerial pictures. They were then classified by order of importance for groundwater storage and infiltration potential (Teixeira, Chaminé, et al. 2008).

To synthesize this data for vulnerability modelling, the source studies translated these field observations into specific parameters. Table 1 summarizes how the geological observations were categorized for use in the vulnerability indices (e.g., DRASTIC-Fm and DISCO):

Table 1: summary of geological characterization according to ISRM

Geological feature	Classification (ISRM)	Vulnerability implication
Fracturing degree	Wide spacing (F1-2)	Lower permeability; lower vulnerability potential.
	Moderate (F3) to close (F4-5)	Higher secondary porosity; increased infiltration and vulnerability.
Weathering grade	Fresh to slightly weathered (W1-2)	Reduced storage; flow restricted to fractures.
	Highly weathered (W4-5)	High storage capacity; behaves similarly to porous media (arenization).
Tectonic lineaments	1st order lineaments	Major fault zones acting as primary flow conduits (High Vulnerability).
	2nd order lineaments	Secondary fractures with moderate influence on flow.

2.3 Quantification of Arboreal Vegetation

The first step in the analysis of the study area was the enumeration and the localization of arboreal vegetation.

2.3.1 Generating DTM and DSM

Following the compilation of raw files into a single LAS Dataset (.LASD), the point cloud was imported into ArcGIS Pro. The data adhered to the ASPRS standard classification scheme (Carter, et al. 2012), containing the following categories: Unassigned (1), Ground (2), Low Vegetation (3), Medium Vegetation (4), High Vegetation (5), Building (6), Noise (7), Water (9), and Reserved.

The Digital Terrain Model (DTM) was derived by filtering the point cloud for the "Ground" class (Class 2) and converting the data into a raster format using the Point to Raster tool. To determine the elevation value of each cell, the "Most Frequent" (ESRI s.d.) cell assignment type was selected. This methodology enforces a specific protocol for cells containing multiple LiDAR points: priority is given to the Z-value (elevation) possessing the most common attribute. In instances where point attributes differ but share the same frequency (a tie), or if no single attribute is dominant, the algorithm defaults to the point with the lowest Feature ID (FID) to ensure a deterministic value assignment.

To generate the Digital Surface Model (DSM), a specific subset of classes was utilized: Ground, Low Vegetation, Medium Vegetation, and High Vegetation. Notably, the "Building" class was intentionally excluded from the DSM generation process. As the ultimate objective was the creation of a Normalized Digital Surface Model (nDSM) for individual tree segmentation, excluding anthropogenic structures at this stage prevented the introduction of non-vegetative outliers into the canopy height analysis.

2.3.2 Generating the nDSM

Following the generation of the DTM and DSM, the next objective was to isolate the vertical structure of the arboreal vegetation by removing the influence of the underlying topography. To

achieve this I derived a Normalized Digital Surface Model (nDSM), often referred to as a Canopy Height Model (CHM) in forestry applications (Melin, Shapiro e Glover-Kapfer 2017).

The raw DSM represents absolute elevations above the vertical datum (e.g., mean sea level), but the inventory analysis required object heights scaled relative to the ground surface. Consequently, the nDSM was calculated by mathematically subtracting the DTM from the DSM on a pixel-by-pixel basis. This algebraic operation effectively normalizes the dataset, resulting in a raster where cell values correspond directly to the height of vegetation Above Ground Level (AGL) rather than absolute elevation (Melin, Shapiro e Glover-Kapfer 2017). This step was executed using the Raster Calculator tool in ArcGIS Pro.

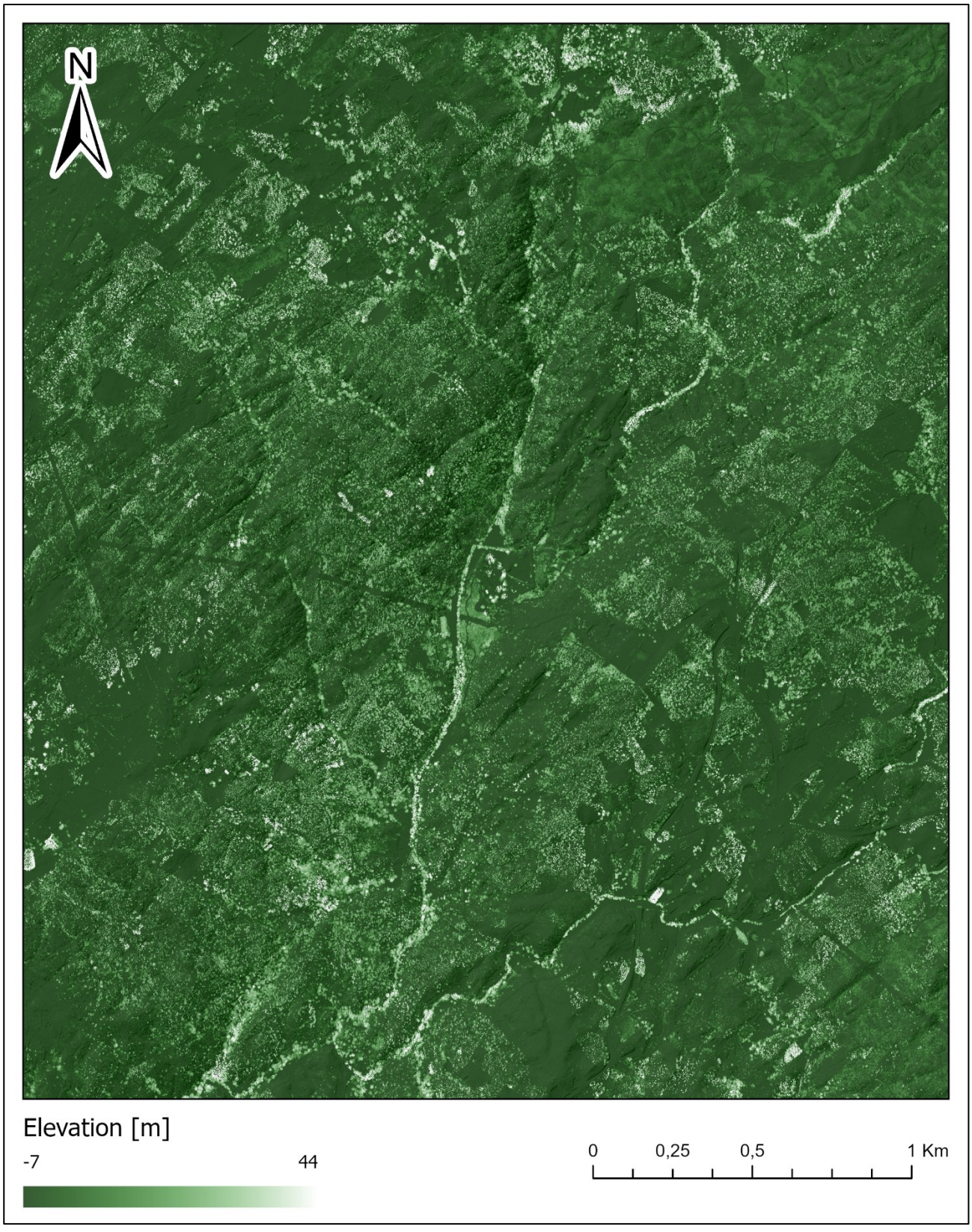

Figure 4: nDSM of the study area

2.3.3 Vegetation height thresholding

To restrict the inventory analysis to fully grown and mature arboreal vegetation, a vertical height threshold was applied to the Normalized Digital Surface Model (nDSM): all cell values below the 2m threshold were reclassified as *NoData*. This filtration step eliminated understory features, such as shrubs, bushes, and low-lying vegetation, ensuring that the subsequent segmentation algorithms operated exclusively on the canopy layer of mature trees.

2.3.4 Canopy Peak Detection via Focal Statistics

To identify individual treetops within the forest canopy, a local maximum algorithm was applied to the filtered nDSM using the Focal Statistics tool. Conceptually, this algorithm iterates through every cell in the input raster, referred to as the processing cell, and calculates a specific statistic based on the values of the surrounding cells that fall within a specified neighbourhood shape. For every processing cell, the calculated value is assigned to the corresponding location in the output raster (ESRI s.d.).

The *Maximum* statistic type was utilized within a square neighbourhood configuration. This assigns the highest elevation value found within the defined square window to the processing cell.

To detect the tree peaks, the output of the Focal Statistics function was compared against the original nDSM. This was executed using the Con (Conditional) function. The Con function sets the value of output pixels based on an if/else evaluation of each input pixel. It operates by returning values from a specified "True Raster" if a conditional evaluation is true, or values from a "False Raster" if the condition is false (ESRI s.d.).

The logic applied in this analysis was:

$$\text{Con}(\text{"nDSM"} == \text{"Focal\_Statistics"}, \text{"nDSM"}, \text{None})$$

In this expression, the tool evaluated whether the original pixel value in the nDSM was identical to the maximum value found in its neighbourhood. If the condition was true (indicating the pixel was a local peak), the tool assigned the original height value to that location (pixel) in the output raster; if false, that location was assigned *NoData*. This process stripped away all non-peak

pixels, leaving a raster consisting solely of local maxima representing the highest points of individual tree crowns (ESRI s.d.).

Because the size of the tree crowns varies, selecting an appropriate neighbourhood size is important in order to avoid commission errors (identifying multiple peaks on a single tree) or omission errors (skipping smaller trees). To determine the optimal parameter, a sensitivity analysis was conducted using four distinct neighbourhood sizes: 5, 10, 20, and 30 pixels. Given the spatial resolution of 0.25 m, these neighbourhoods corresponded to ground distances of 1.25 m, 2.5 m, 5.0 m, and 7.5 m, respectively.

2.4 Groundwater Vulnerability Assessment: DRASTIC and SI Evaluation

The next phase of this work focuses on the evaluation of groundwater vulnerability within the study area using the DRASTIC and SI (Susceptibility Index) frameworks.

Each of these factors is subdivided into ranges or significant media types that are rated between 1 and 10 according to their relative impact on the pollution potential, as shown in Table 2.

2.4.1 The DRASTIC index

The DRASTIC index, originally developed for the U.S. Environmental Protection Agency (Aller, Lehr and Bennet 1987), is a standardized framework for evaluating groundwater pollution potential. It is designed to assess intrinsic vulnerability, defined as the inherent geological and hydrogeological characteristics that determine the ease with which contaminants can migrate from the surface to the aquifer, independent of specific pollutant types or anthropogenic land use.

The method assumes that contaminants are introduced at the ground surface and flushed into the groundwater by precipitation (Aller, Lehr and Bennet 1987). To quantify vulnerability, the method uses a weighted sum:

$$DRASTIC = D \cdot w_D + R \cdot w_R + A \cdot w_A + S \cdot w_S + T \cdot w_T + I \cdot w_I + C \cdot w_C$$

Each parameter is assigned a fixed weight (w_i) ranging from 1 to 5, reflecting its relative importance in the contamination process. The parameters are instead assigned a value ranging

from 1 to 10, according to the physical value assumed by each of them. This reclassification process is explained in Table 2.

The DRASTIC index considers the following conditions as contributing to a high pollution potential:

- Shallow depth to groundwater (D), related to a short travel time of the contaminant in the unsaturated zone and thus only having a small chance for attenuation (e.g., oxidation or other interactions with the surrounding media);
- High net aquifer recharge (R), the main vehicle for leaching contaminants to the aquifer;
- Permeable aquifer media (A) showing no reactivity about the contaminant, hence allowing a quick spreading through the aquifer;
- Soil media (S) lacking (non-shrinking and non-aggregated) clay and organic material, conferring a low attenuation capacity and increasing the mobility of the contaminant;
- Flat topography (T), decreasing the likelihood of surface runoff and erosion and hence facilitating infiltration;
- Permeable vadose (unsaturated) zone media (I) showing no reactivity regarding the contaminant, thus creating conditions for leaching towards the aquifer;
- High hydraulic conductivity of the aquifer (C), allowing quick spreading throughout the aquifer (though this also depends on hydraulic gradient) (Stigter, Ribeiro and Carvalho Dill 2005);

Table 2: ranges (left column) and ratings (right column) for the DRASTIC and SI parameters

D[a] (m)		R[a] (mm)		T[a] (%)		S		C (m/day)	
<1.5	10	<51	1	<2	10	Thin or absent	10	<4.1	1
1.5-4.6	9	51-102	3	2-6	9	Gravel	10	4.1-12.2	2
4.6-9.1	7	102-178	6	6-12	5	Sand	9	12.2-28.5	4
9.1-15.2	5	187-254	8	12-18	3	Peat	8	28.5-40.7	6
15.2-22.9	3	>254	9	>18	1	Shrinking and/or aggregated clay	7	40.7-81.5	8
22.9-30.5	2					Sandy loam	6	>81.5	10
>30.5	1					Loam	5		
						Silty loam	4		
						Clay loam	3		
						Muck	2		
						Non-shrinking and non-aggregated clay	1		

A[ab]		I[b]	
Massive shale	1-3 (2)	Confining layer	1
Metamorphic/igneous	2-5 (3)	Silt/clay	2-6 (3)
Weathered metamorphic/igneous	3-5 (4)	Shale	2-6 (3)
Glacial till	4-6 (5)	Limestone	2-5 (3)
Bedded sandstone, limestone and shale sequences	5-9 (6)	Sandstone	2-7 (6)
Massive sandstone	4-9 (6)	Bedded limestone, sandstone, shale	4-8 (6)
Massive limestone	4-9 (8)	Sand and gravel with significant silt and clay	4-8 (6)
Sand and gravel	4-9 (8)	Sand and gravel	4-8 (8)
Basalt	2-10 (9)	Basalt	2-10 (9)
Karst limestone	9-10 (10)	Karst limestone	8-10 (10)

[a] For SI the ratings are multiplied by 10

[b] Typical ratings between brackets

Table 3: definition and weights of the DRASTIC and SI parameters

Letter	Meaning	Weight (w_i)	
		DRASTIC	SI
D	Depth to water	5	0.186
R	Net recharge	4	0.212
A	Aquifer media	3	0.259
S	Soil media	5	-
T	Topography	3	0.121
I	Impact of the vadose zone media	4	-
C	Hydraulic conductivity of the aquifer	2	-
LU	Land use	-	0.222

2.4.2 The Susceptibility Index (SI)

Adapted from the DRASTIC method, the Susceptibility Index (SI) evaluates vulnerability to diffuse agricultural pollution at scales ranging from 1:50,000 to 1:200,000 (Ribeiro 2000). The primary innovation of the SI is the integration of a land use (LU) parameter, which moves the assessment beyond purely intrinsic vulnerability, towards the effects of diffuse agricultural pollution (Stigter, Ribeiro and Carvalho Dill 2005). The term susceptibility is used here to denote the lack of capacity to resist groundwater contamination (Vrba e Zaporozec 1994).

To quantify vulnerability, the SI uses a weighted sum (Stigter, Ribeiro and Carvalho Dill 2005):

$$SI = D \cdot w_{D,SI} + R \cdot w_{R,SI} + A \cdot w_{A,SI} + T \cdot w_{T,SI} + LU \cdot w_{LU}$$

The weighting ratings assigned to each factor are different from the ones used in DRASTIC (Ribeiro 2000), as shown in

Table 3. The land use categorization and assigned values are presented in Table 4.

Three DRASTIC parameters were left out of the construction of the Susceptibility Index: soil media (S), Impact of the vadose zone media (I), and the conductivity of the aquifer (C).

Two of these include the soil (S) and unsaturated zones (I), thus suggesting that their direct influence on the contamination linked to agricultural practices is of little importance. The influence of soil type in particular is debated: some sources suggest that it is indirectly represented by land use (Francés, et al. 2002). However other sources recognize that the soil can have a certain amount of attenuation potential, especially when rich in clay minerals and organic matter, so it wouldn't necessarily be wise or obvious to leave these soil-contaminants interactions out of the assessment. On the other hand, an additional justification can be given by the fact that, due to ploughing, tillage and many other techniques applied to improve the soil structure and fertility, the natural soils are frequently disturbed during cultivation of land so that they lose much of their original characteristics (Stigter, Ribeiro and Carvalho Dill 2005). The unsaturated zone can also have a certain degree of attenuation capacity and prevent leaching of the contaminant to the groundwater but, in the case of persistent and mobile contaminants, its role is one of introducing a time-lag, given the low amount of attenuation that takes place (Stigter, Ribeiro and Carvalho Dill 2005). The last DRASTIC parameter not incorporated in the SI is the hydraulic conductivity of the aquifer (C), especially due to the difficulty of measurement: there rarely exist enough data to paint and accurate picture (Stigter, Ribeiro and Carvalho Dill

2005). Furthermore, the aquifer media parameter (A) already gives a qualitative measure of hydraulic conductivity.

The SI has limitations in cases where the aquifer displays a low residence time. A high recharge rate, combined with a thin aquifer, creates a dynamic where the groundwater contents are flushed out relatively quickly. In this scenario, the actual vulnerability is effectively lower, as the high turnover rate ensures that contamination is short-lived. The SI, however, operates on the opposite premise: it assumes that a high recharge rate increases the volume of contaminants leaching into the water table. By failing to account for the dilution and flushing effects of a low residence time, the SI erroneously predicts high vulnerability in these conditions. Therefore, the application of this index is limited to hydrogeological settings that do not exhibit exceedingly low residence times (Francés, et al. 2002).

The Land Use (LU) parameter serves as the primary determinant of specific vulnerability, acting as a proxy for the anthropogenic stresses placed on the hydrogeological system. Unlike the intrinsic parameters, the rating assignment for Land Use is governed by the potential for nutrient loading and the physical alteration of the soil structure. The highest vulnerability ratings are assigned to intensive agricultural classes, such as irrigated arable land and paddy fields. These categories typically receive near-maximum ratings due to the "double threat" they pose: the extensive application of nitrogen-based fertilizers combined with artificial irrigation, which accelerates the leaching of nitrates into the saturated zone.

Furthermore, the high ratings for these categories reflect the reduced attenuation capacity of the soil: agricultural practices such as ploughing and tillage frequently disturb the natural soil matrix, causing it to lose its original protective characteristics and facilitating rapid contaminant transport (Stigter, Ribeiro and Carvalho Dill 2005). In contrast, semi-natural land cover types, such as Forests and Scrubland, are assigned significantly lower ratings. In these zones, the absence of tillage preserves the soil's organic matter and clay structure, maintaining a high natural attenuation potential and preventing significant nitrate infiltration.

2.4.3 Index parameters

Most of the data needed to evaluate the DRASTIC and SI indexes were drawn from literature, with the exception of slope data and land use.

2.4.3.1 Depth of water (D)

The Caldas da Cavaca hydro-geological system comprises three aquifer units that differ in their hydraulic confinement and therefore in the way a water-table is expressed.

- Shallow unconfined (phreatic) aquifer: This unit occupies the valley bottom and the alluvial cover. Groundwater occurs at or just below the land surface, typically at depths of < 5 m, and its level follows the gentle topography (low-slope, weathered granite) and seasonal recharge from the 1 250 mm yr^{-1} precipitation. Because the aquifer is unconfined, a water-table is present; its position fluctuates seasonally, rising during the winter-spring wet period and falling during the summer (June-September). The areas of high infiltration potential promote rapid recharge, while the low-slope, highly fractured granite enhances hydraulic connectivity, making the water-table highly responsive to surface processes;
- Weathered-to-fractured granite aquifer: situated beneath the shallow cover, this layer is still unconfined to semi-confined. Its water-table lies a few metres below the surface but is less directly linked to topography because fracture density controls flow pathways. It shows transmissivity lower than 1 m2/day and very low water yields (less than 0.05 L·s−1);
- Deep confined hydromineral aquifer**:** hosted in the fresh granite of the Ribeira de Coja deep fault zone, with transmissivity ranging from 27 to 136 m2/day. this unit is sealed by overlying low-permeability layers, so the water is under pressure, and no surface water-table is observable; it is a confined system (Meerkhan, et al. 2016) (Teixeira, Chaminé, et al. 2014)

Thus, only the shallow alluvial unit is truly phreatic, with a classic water-table that mirrors recharge-discharge dynamics, while the deeper aquifers are either semi-confined or fully confined and lack a surface water-table.

Figure 5: depth of water (D) classification over the study area

2.4.3.2 Net recharge (R)

The net recharge (R) parameter quantifies the fraction of precipitation that actually reaches the aquifer after losses to evapotranspiration and surface runoff. For the Caldas da Cavaca catchment the source data adopted a uniform net-recharge value of 20 % of the annual rainfall for most of the study area, which translates to roughly 175 mm yr^{-1} (14 % of the mean 1 250 mm yr^{-1} precipitation) (Stigter, Ribeiro and Carvalho Dill 2005) (Teixeira, Chaminé, et al. 2008).

Because the terrain includes karstified limestone outcrops with higher infiltration capacity, a higher recharge fraction of 40 % was assigned to those zones, following the values reported by Silva (1988) for similar carbonate rocks (Stigter, Ribeiro and Carvalho Dill 2005).

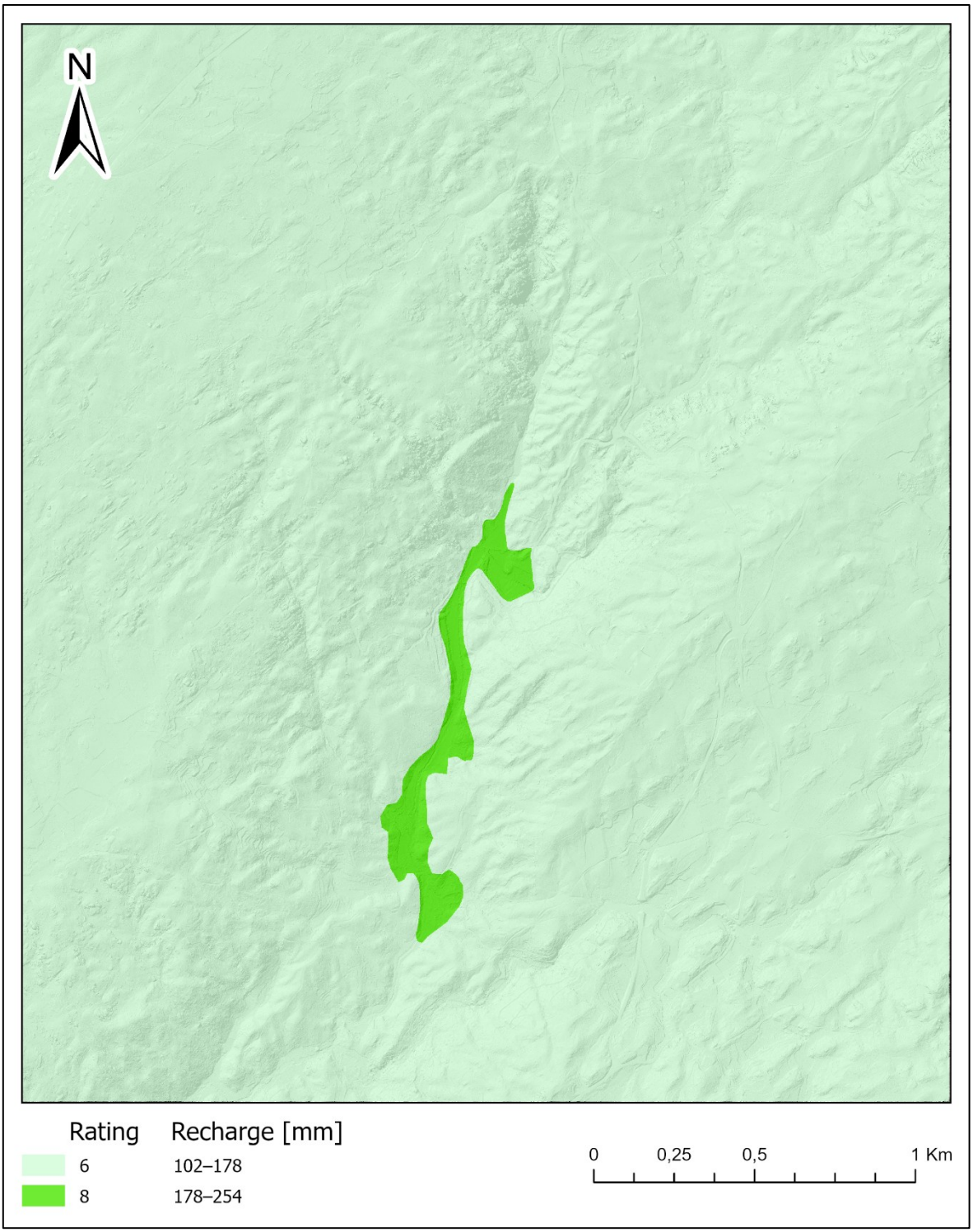

Figure 6: net recharge (R) classification over the study area

2.4.3.3 Aquifer media (A) ad hydraulic conductivity (C)

These two parameters will be discussed together as they are tightly connected to one another. The Caldas da Cavaca system is built on three distinct aquifer media that reflect its complex hard-rock setting.

1. Shallow unconfined aquifer (alluvial cover): this unit occupies the valley floor and consists mainly of sands and gravels derived from recent alluvium; water shows low mineralisation (EC < 20 µS cm^{-1}, pH 5-6.5) and the water-table exhibits a thickness of 3-5 meters. The transmissivity is of < 1 m^2 day^{-1}
2. Weathered to fractured granite aquifer: beneath the alluvium, groundwater is hosted in coarse-grained porphyritic granite that ranges from fresh to highly weathered (W1-5) and is intersected by a dense network of fractures and tectonic lineaments. This fractured granitic mass provides the primary hydraulic conductivity of the system, with transmissivities generally < 1 m^2 day^{-1}.
3. Deep confined hydromineral aquifer. The deepest water-bearing formation is a fresh granite body bounded by the Ribeira de Coja fault zone. It is largely unweathered (W1-2) and sealed by low-permeability overburden, producing a confined system with high temperature (≈ 29.8 °C), alkaline pH (8.3-8.6), silica ≈ 55 mg L^{-1} and fluoride up to 14 mg L^{-1} (Meerkhan, et al. 2016) (Teixeira, Chaminé, et al. 2014).

The aquifer media transition from unconsolidated alluvial sediments to progressively less weathered granitic rocks, with dolerite dykes interspersed in the granite body.

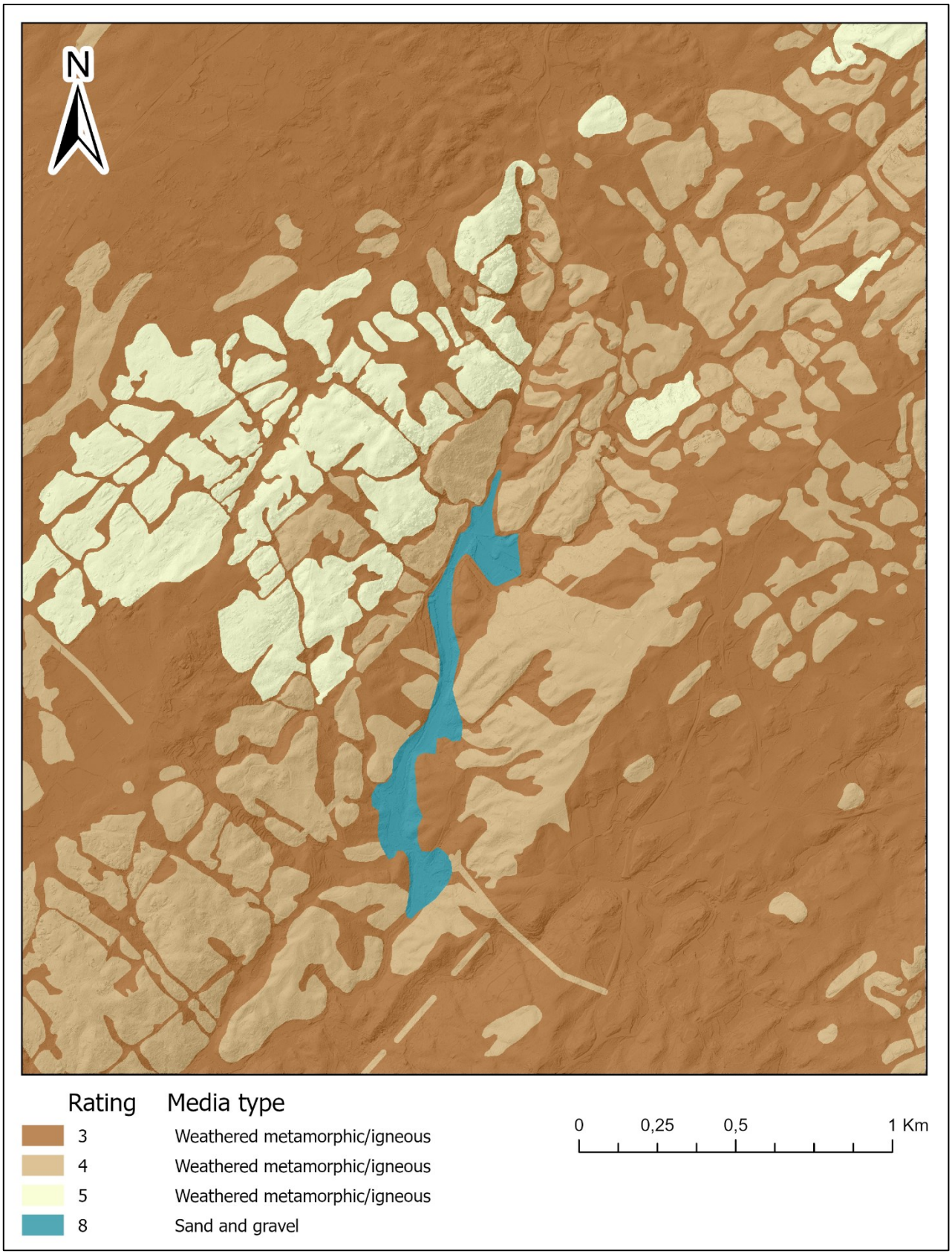

Figure 7: aquifer media (A) classification over the study area

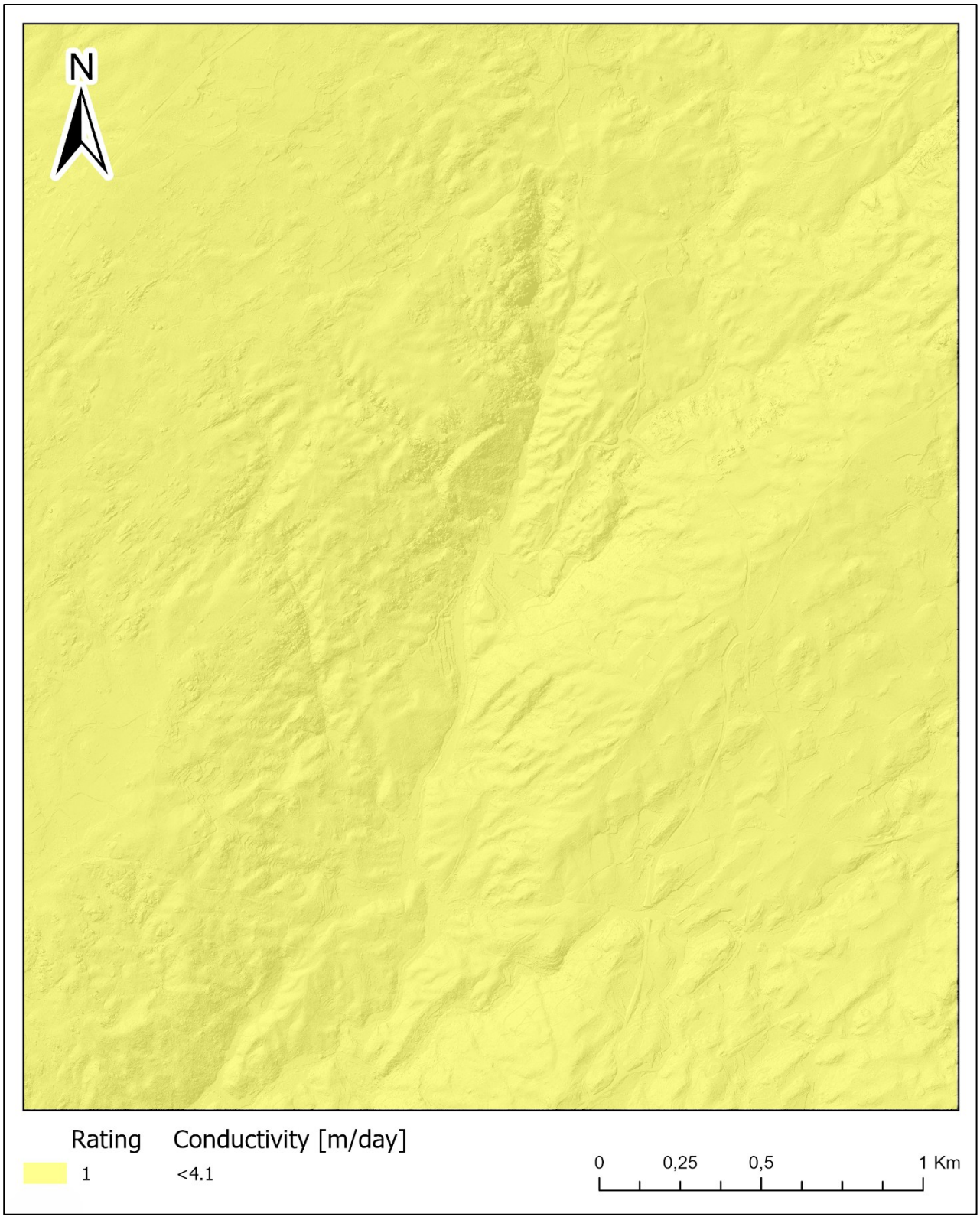

Figure 8: hydraulic conductivity (C) classification over the study area

2.4.3.4 Soil media (S)

The soil media within the Caldas da Cavaca study area is linked to the weathering processes of the underlying parent rock and the depositional history of the local hydrological system.

In the intrinsic vulnerability indexes the soil media (S) receives a moderate weight because it influences the attenuation of pollutants, especially where the soil is less developed (sand-rich) and the unsaturated zone is thin (Meerkhan, et al. 2016).

Thus, the Caldas da Cavaca area is dominated by a mosaic of sandy to silty-loam soils that, together with the degree of granite weathering, control infiltration potential and the vulnerability of the underlying fractured-rock aquifers.

Sandy Loam (High Vulnerability) is predominant in the region, accounting for around 60% of the entire area, specifically covering the areas of Highly Weathered Coarse-Grained Granite (W_{4-5}) and Fresh to Slightly Weathered Coarse Grained Granite (W_{1-2}). This texture contains a high percentage of sand particles and a low percentage of clay and organic matter. This coarse texture creates larger pore spaces, resulting in higher permeability and a lower capacity to adsorb contaminants. Consequently, pollutants can infiltrate rapidly through this soil layer with minimal attenuation or retardation, making the underlying groundwater highly vulnerable to contamination. In the study area, this texture is primarily associated with weathered granite (W_{4-5}) (Meerkhan, et al. 2016).

Silty Loam is associated with the moderately weathered coarse-grained granite (W_3) and accounts for around 25% of the study area. It possesses a texture intermediate between sand and clay. While it allows for water recharge, the presence of silt particles provides a better filtering capacity and a slower percolation rate compared to sandy loam, thereby offering moderate natural protection against the migration of contaminants into the aquifer.

Clay loam characterizes the soil cover developed over the alluvial deposits, fresh to weathered granite (W_{1-2}) and dolerite dykes: it accounts for about 15% of the study area. It contains a higher proportion of fine clay particles, which significantly decreases the soil's permeability and increases its specific retention. The clay minerals possess a high cation exchange capacity, which facilitates the sorption (binding) of contaminants, effectively retarding their movement into the subsurface. Therefore, these areas are assigned the lowest vulnerability rating of the study area.

Additionally, the land cover, specifically vegetation, plays a functional role in the hydraulic behaviour of the soil media. The decomposition of root systems in forest and agricultural areas contributes to organic matter content, which facilitates the formation of structural aggregates, thereby loosening the soil and enhancing hydraulic conductivity and infiltration potential (Teixeira, Chaminé, et al. 2008).

Figure 9: soil media (S) classification over the study area

2.4.3.5 Topography (T)

To derive the topographical gradient T required for the DRASTIC model, the “Slope” tool was utilized within ArcGIS Pro.

The tool calculates the steepness of the terrain by identifying the maximum rate of change in elevation (z-value) from each cell to its neighbours. This calculation is performed using a 3-by-3 cell moving window. As the algorithm iterates through the input raster, it examines the centre processing cell and its eight immediate neighbours. The tool then fits a plane to the z-values of the nine cells defined by the 3-by-3 neighbourhood. The slope for the processing cell is determined by the steepest gradient of this fitted plane. This approach ensures that local anomalies are smoothed relative to the immediate neighbourhood in order to a consistent gradient value (ESRI s.d.).

For this study, the output measurement parameter was set to percent rise, as required by the DRASTIC protocol (Stigter, Ribeiro and Carvalho Dill 2005).

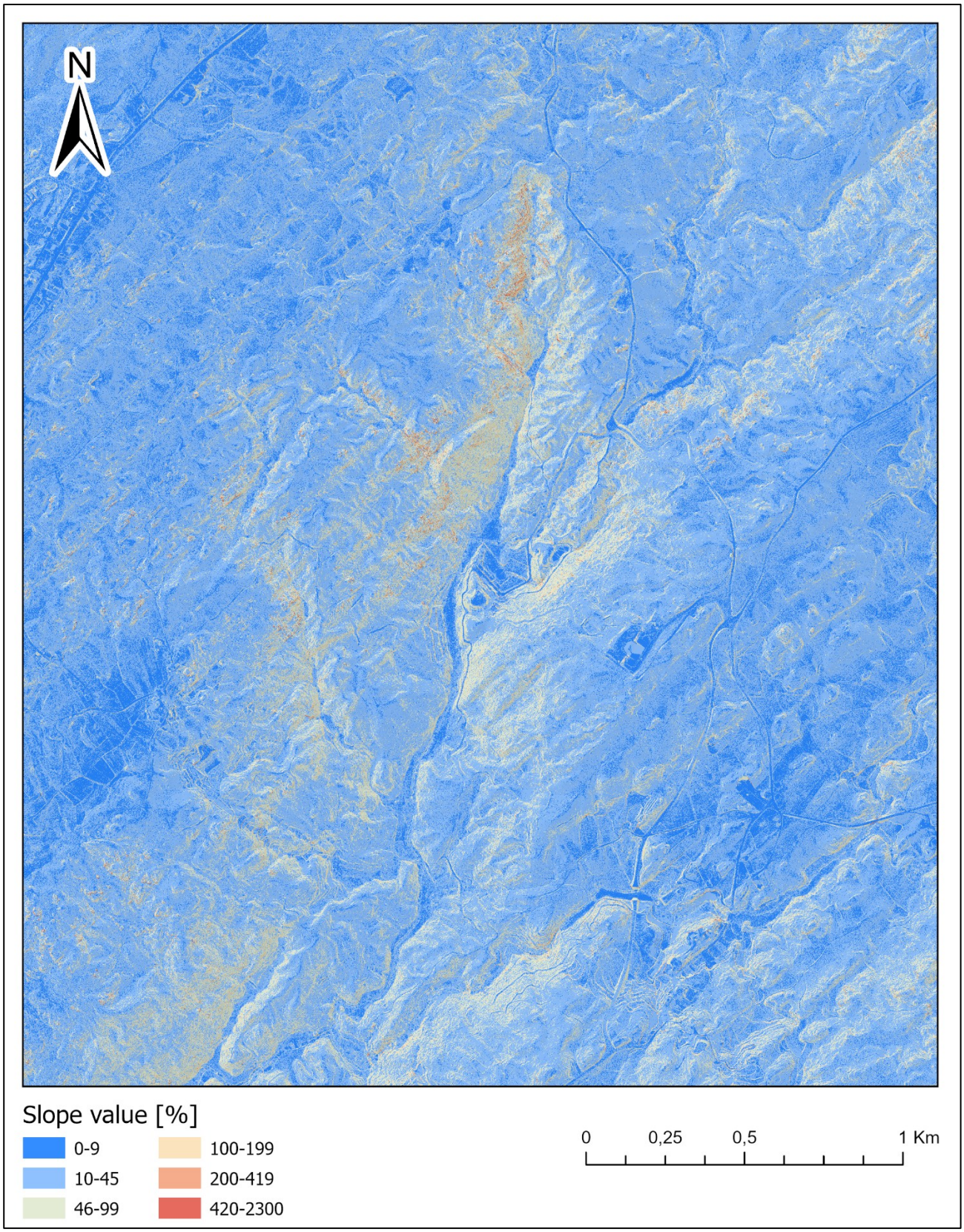

Figure 10: slope values, expressed in percentage rise, over the study area

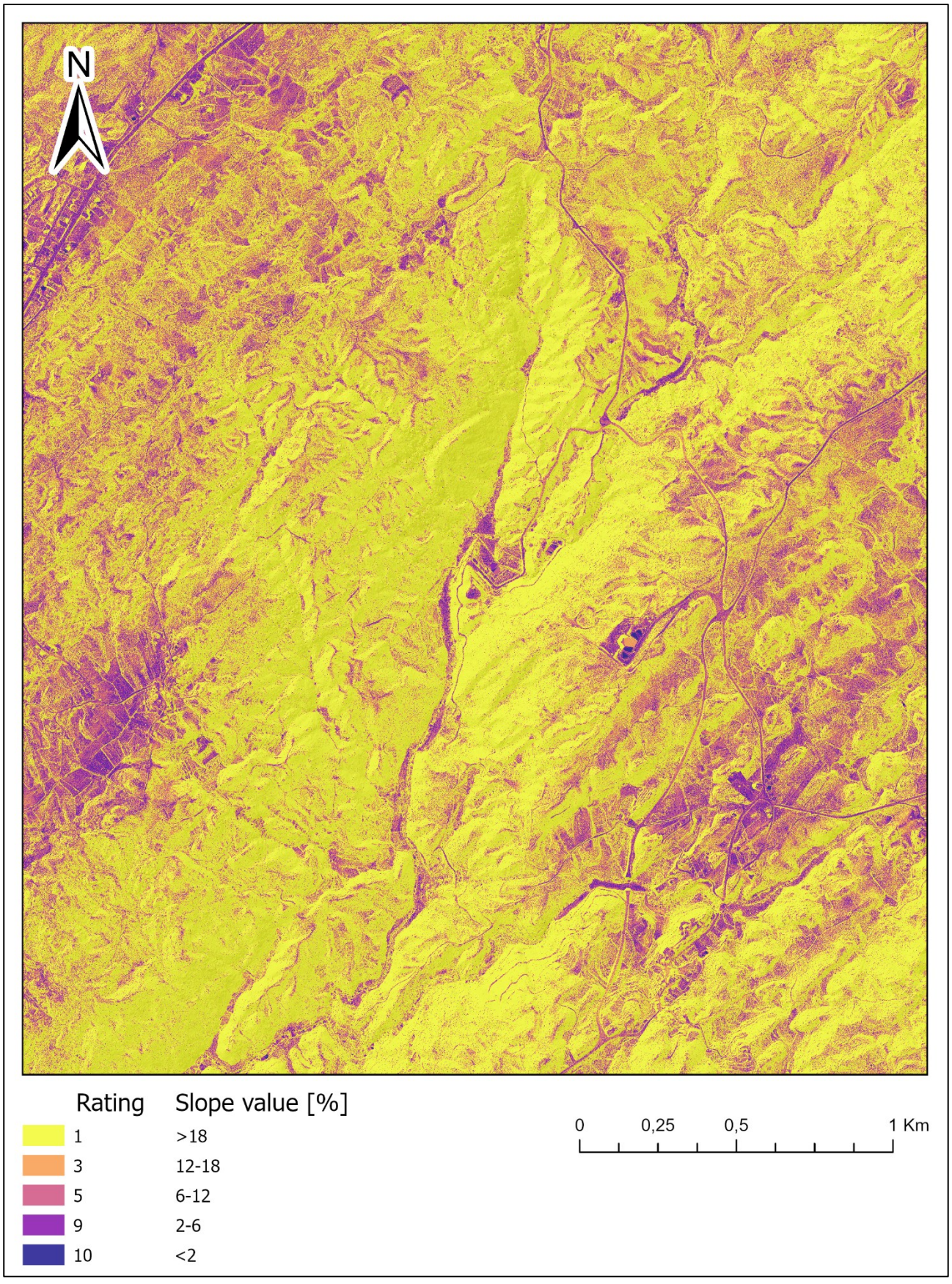

Figure 11: slope (T) classification over the study area

2.4.3.6 Impact of the vadose zone media (I)

In the DRASTIC index applied to the site, it is recognized that the thickness and hydraulic properties of this unsaturated layer affect contaminant attenuation and the travel time of pollutants to the water table (Teixeira, Chaminé, et al. 2014).

Where the vadose zone consists of thin, highly permeable sand, recharge is rapid and the attenuation capacity is low, so contaminants can reach the aquifer quickly.
In areas under moderate permeability loam or thicker unsaturated thickness, the vadose zone provides greater retardation, reducing vulnerability.

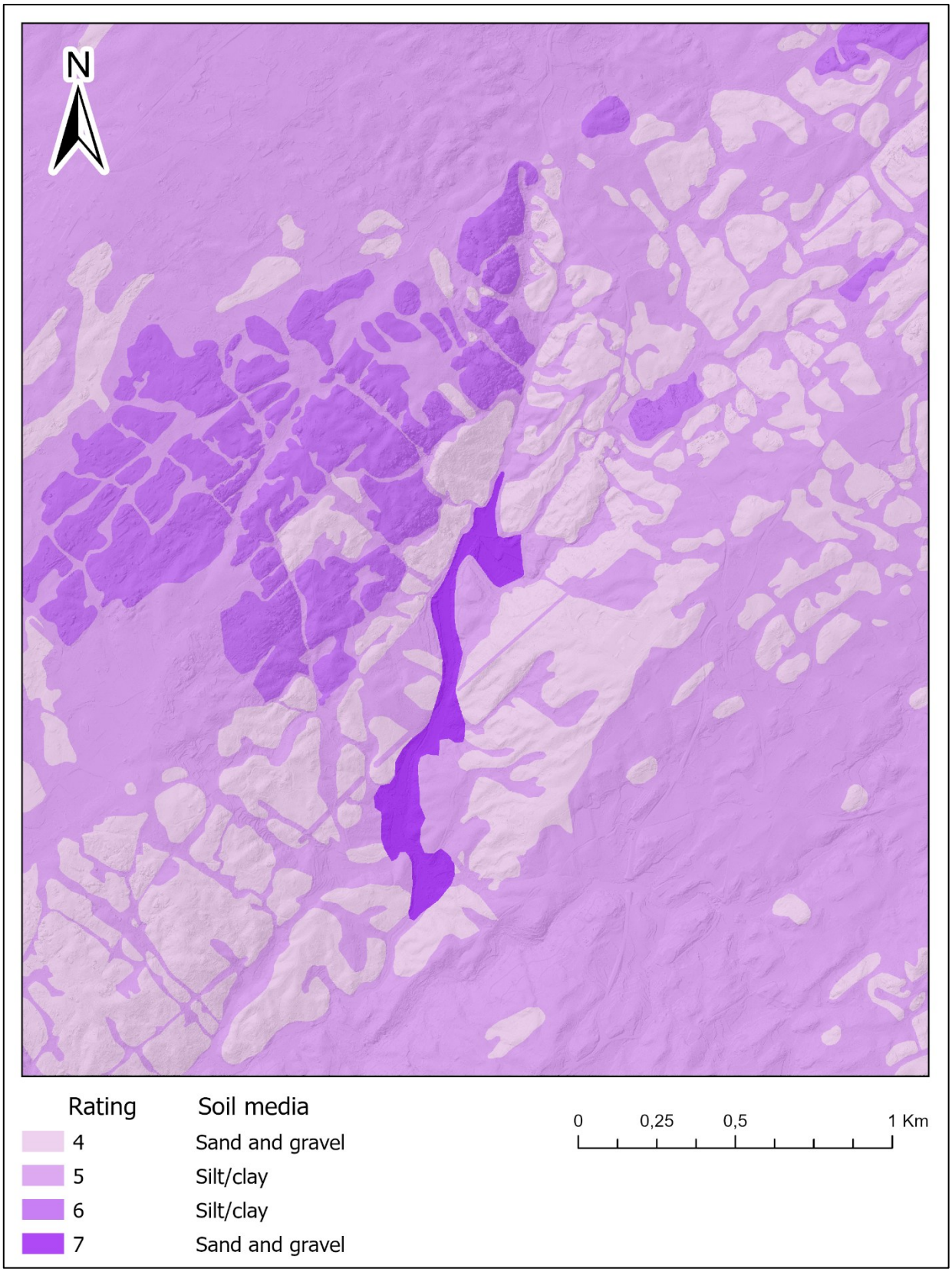

Figure 12: impact of the vadose media (I) classification over the study area

2.4.3.7 Land use (LU)

The integration of the Land Use (LU) parameter constitutes the fundamental adaptation of the Susceptibility Index (SI) relative to the intrinsic DRASTIC model. To spatially define this parameter, official vector data regarding land cover were acquired from the Direção-Geral do Território.

However, a direct application of this dataset was not feasible due to a semantic discrepancy between the standard national nomenclature (Direção-Geral do Território 2018) and the specific weighting scheme required by the SI method (Stigter, Ribeiro and Carvalho Dill 2005). The DGT dataset utilizes a hierarchical classification system with a high level of detail, whereas the SI framework, as defined by Ribeiro (2000), relies on a concise set of agricultural and industrial categories with assigned pollution vulnerability ratings (Table 4).

Table 4: Land use (LU) categories in the SI protocol

Land use	**Rating**
Agricultural areas	
Irrigation perimeters (annual crops), paddy fields	90
Permanent crops (orchards, vine yards)	70
Heterogeneous agricultural areas	50
Pastures and agro-forested areas	50
Artificial areas	
Industrial waste discharges, landfills	100
Quarries, shipyards, open-air mines	80
Continuous urban areas, airports, harbours, (rail)roads, areas with industrial or commercial	75
Discontinuous urban areas	70
Natural areas	
Aquatic environments (salt marshes, salinas, intertidal zones)	50
Forests and semi-natural zones	0
Water bodies	0

To resolve this incompatibility, a reclassification process was carried out. A correspondence system (Table 5) was utilized to map the detailed DGT land use codes to the "most fitting" broad categories defined in the SI methodology.

Table 5: reclassification process between DGT definitions and SI categories

Land use	**Assigned DGT classifications**	
SI categories	*DGT code(s)*	*DGT description*
Agricultural areas		
Irrigation perimeters (annual crops), paddy fields	2.1	Temporary crops
Permanent crops (orchards, vine yards)	2.2	Permanent crops
Heterogeneous agricultural areas	2.3	Heterogeneous agricultural areas
Pastures and agro-forested areas	3.1; 1.8.1.1	Pastures; Green spaces
Artificial areas		
Industrial waste discharges, landfills	1.5.2.1	Landfills
Quarries, shipyards, open-air mines	1.5; 1.6	Aggregate extraction areas, waste disposal areas, and construction yards; Equipment
Continuous urban areas, airports, harbours, (rail)roads, areas with industrial or commercial	1.1.1; 1.3.4.1	Continuous built fabric; Other tourist equipment and facilities
Discontinuous urban areas	1.1.2; 1.2	Discontinuous built fabric; Industry, commerce, and agricultural facilities
Natural areas		
Aquatic environments (salt marshes, salt flats, intertidal zones)	8.1	Wetlands
Forests and semi-natural zones	5.1	Forests
Water bodies	9.1	Inland bodies of water

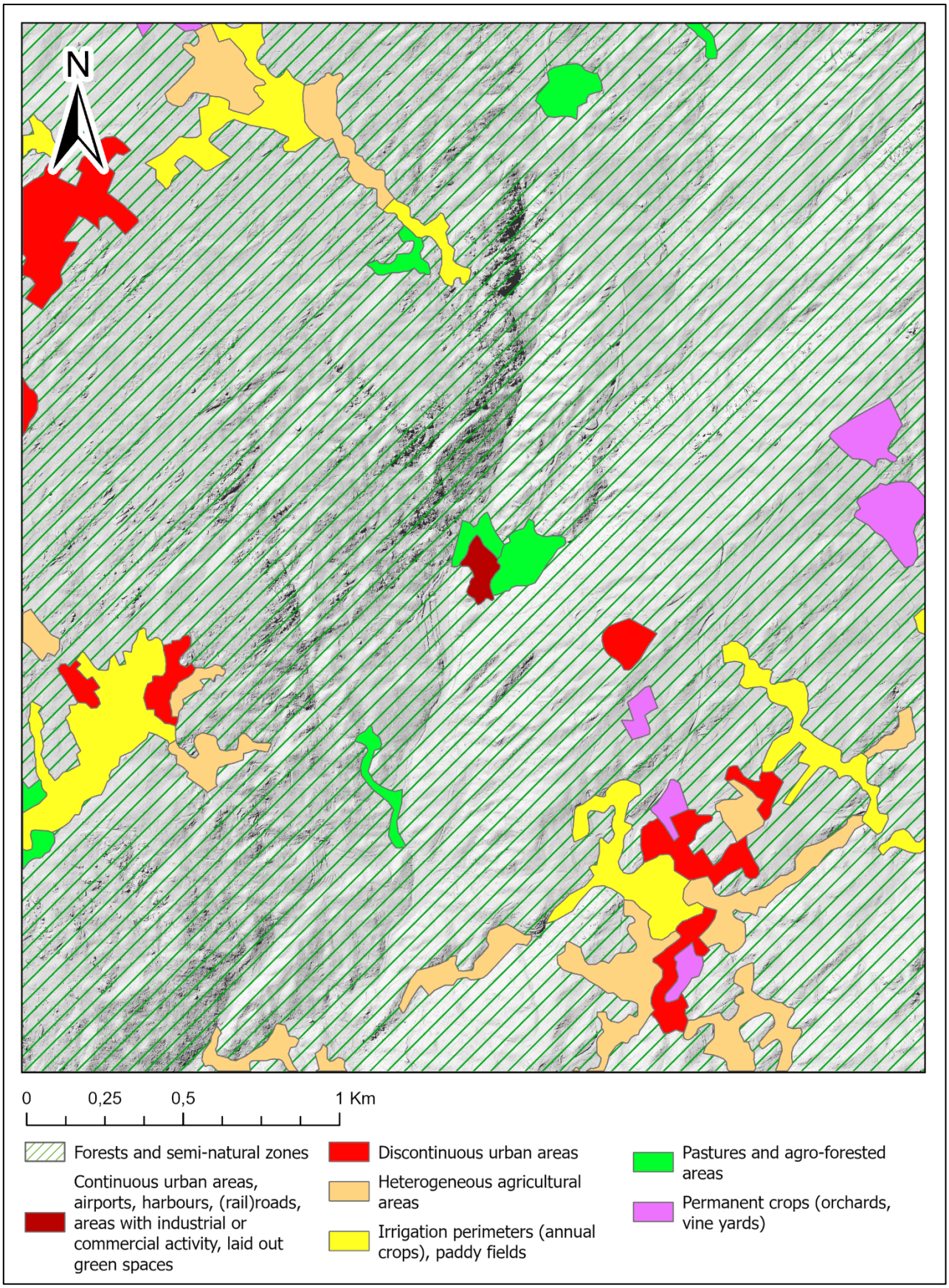

Figure 13: land use (LU) classification over the study area

3 Results

The following chapter presents the results of the hydrogeological vulnerability indices and the spatial distribution and enumeration of the arboreal vegetation.

3.1 Arboreal vegetation

The nDSM allows for a qualitative evaluation of the arboreal vegetation, especially after filtering out the lower height vegetation. The processing of the LiDAR point cloud and the subsequent generation of the nDSM allowed for the isolation of the arboreal strata from the terrain. As shown in Figure 4, the vegetation distribution is heterogeneous. Regarding the tree enumeration, the results obtained with each neighbourhood were visually reviewed and cross-referenced with satellite imagery (Figure 15 and Figure 16). This comparative validation process determined that the 30 pixels neighbourhood yielded the most accurate alignment with visible tree crowns and was consequently selected for the final enumeration. The final count amounted to 74791 treetops.

The application of the 5-meter height threshold (Figure 14) identified distinct clusters of high biomass. These clusters seem to be not uniformly distributed but appear to follow specific spatial trends, which will be analysed in the Discussion chapter in relation to the geological substrate.

As explained, the vegetation data present in the LiDAR database was filtered, in order to remove vegetation lower than 2 meters. By doing the opposite, the same technique has potential for the evaluation of shrub and bush population.

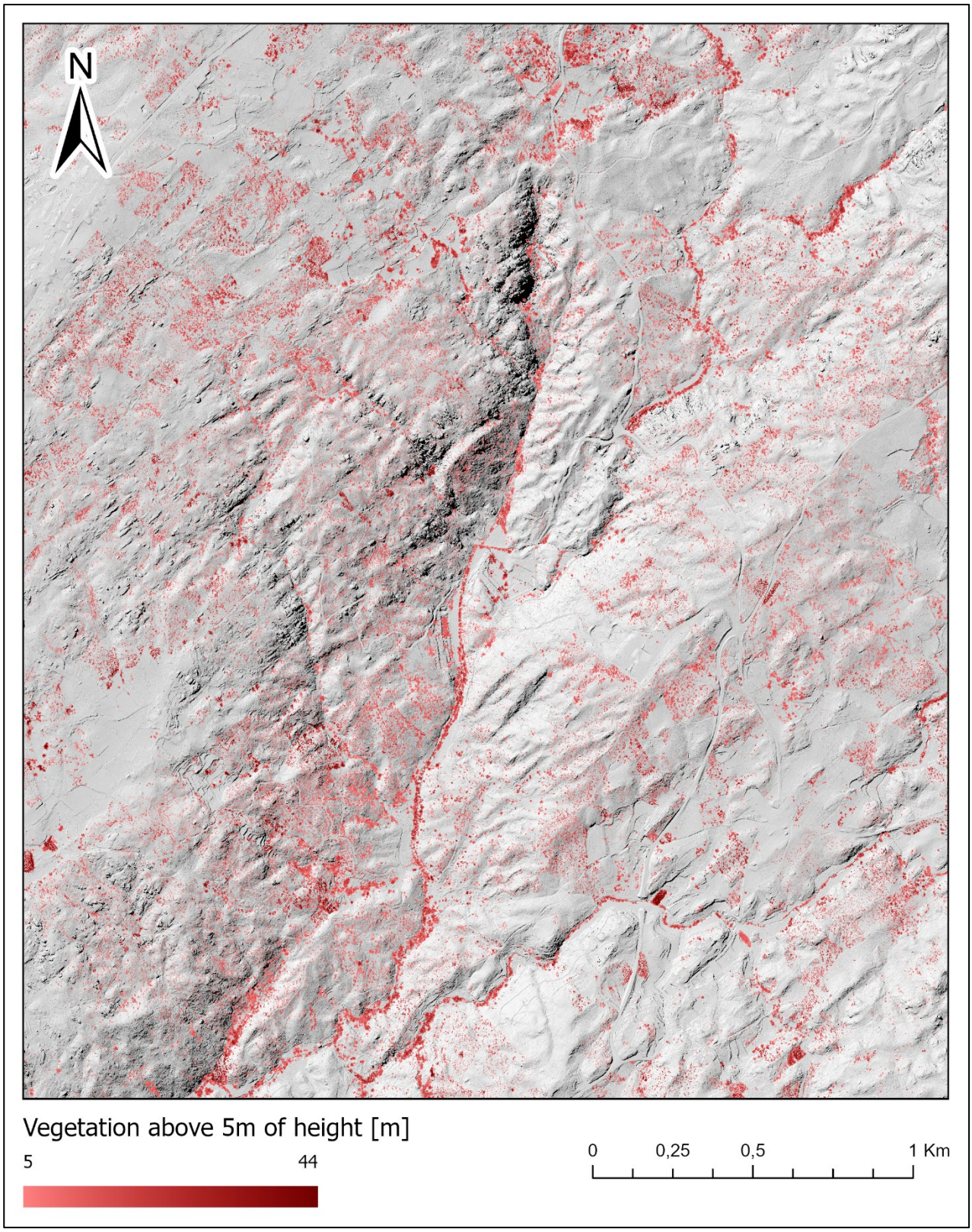

Figure 14: disposition of vegetation lidar data of more than 5m of height above ground

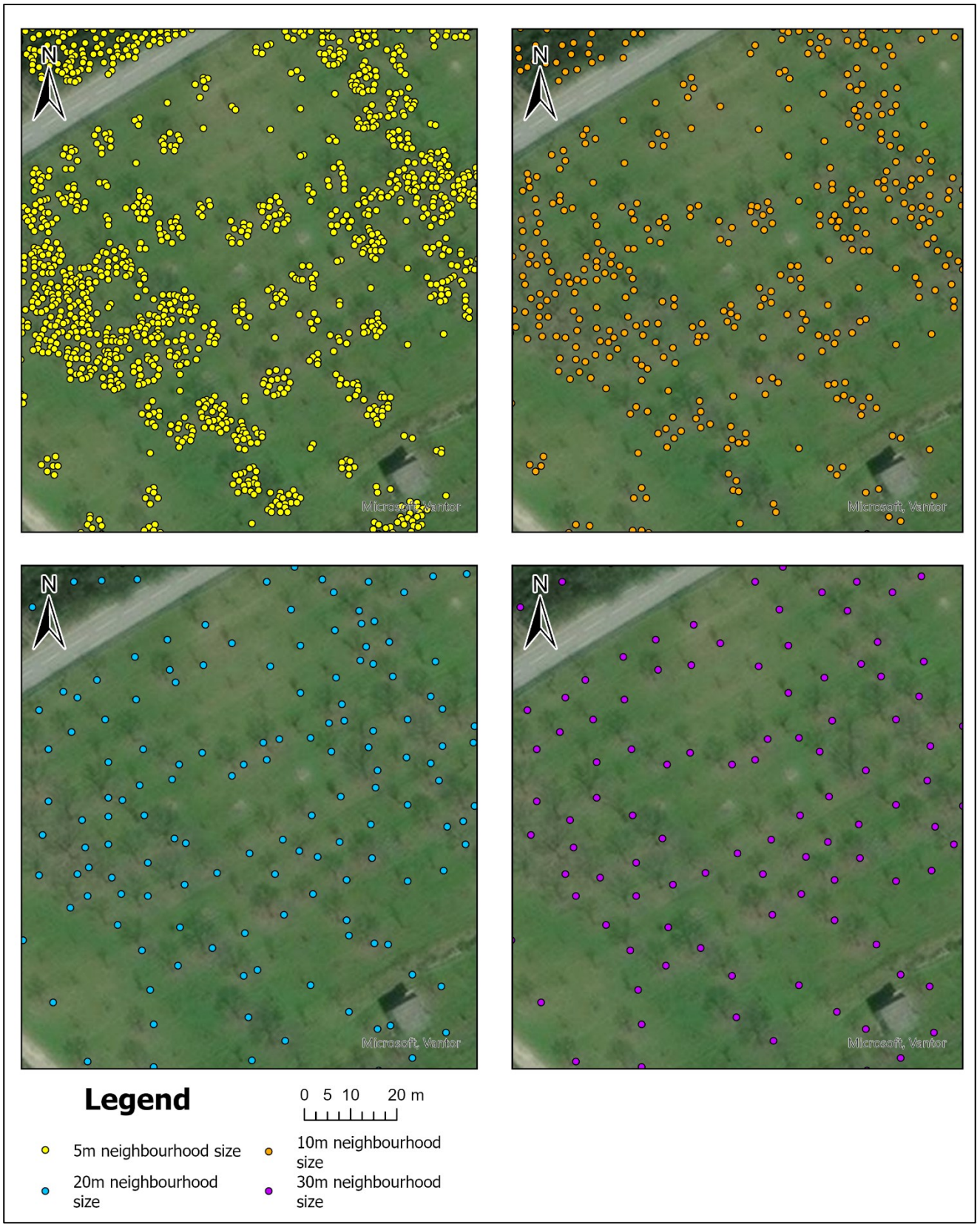

Figure 15: individual tree detection in an orchard using varying neighbourhood sizes

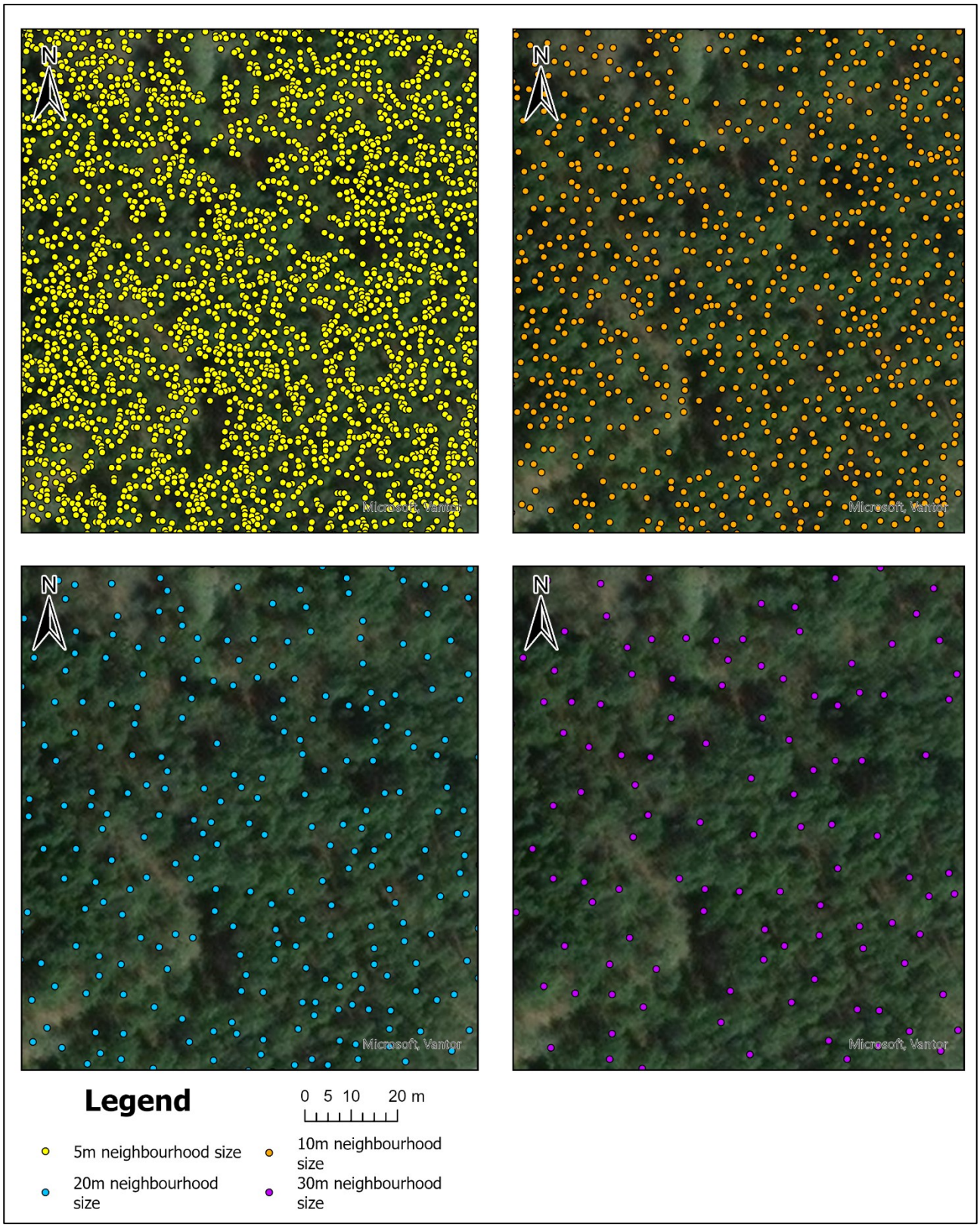

Figure 16: individual tree detection in a forested area using varying neighbourhood sizes

3.2 DRASTIC index

The results of the DRASTIC index evaluation are shown in Figure 17. The application of the DRASTIC model to the study area reveals a distinct spatial dichotomy in groundwater vulnerability, governed primarily by the contrasting hydrogeological properties of the alluvial depositional environment and the surrounding granitic bedrock. The resulting vulnerability maps highlight two primary zones: a central high-vulnerability channel and a surrounding region of moderate-to-low vulnerability.

Overall, most of the area is categorised as low risk with hotspots of "Moderate to low", as shown in Table 6:

Table 6: extent, in kilometres squared and percentage of the entire study area, of each DRASTIC categorization

Categorization	Fraction of the study area [km^2]	% of total area
Extremely low	*not present*	0%
Very low	*not present*	0%
Low	6,80	70,4%
Moderate to low	2,74	28,2%
Moderate to high	0,11	1,2%
High	0,01	0,2%
Very high	*not present*	0%
Extremely high	*not present*	0%

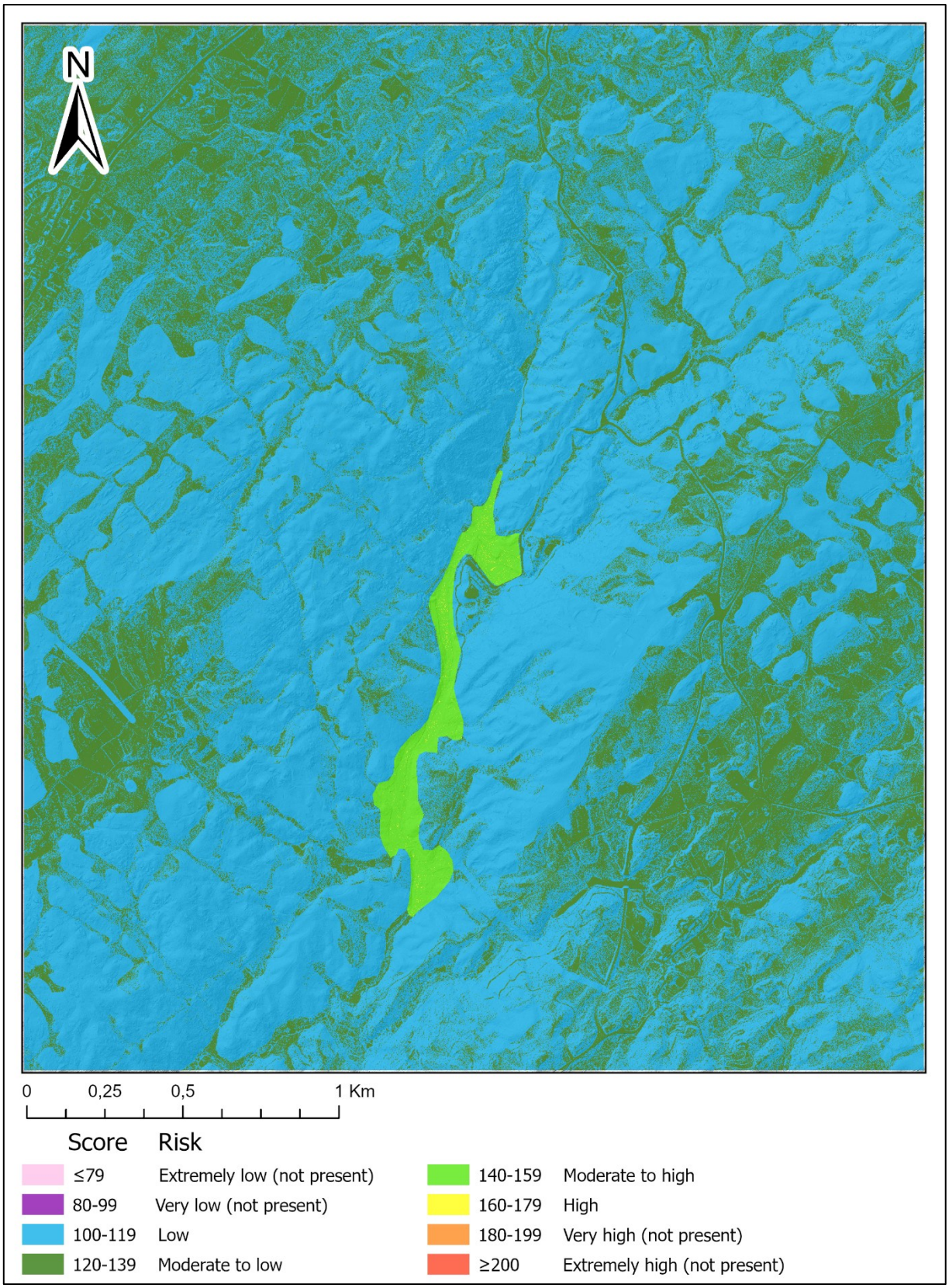

Figure 17: DRASTIC index results over the study area

3.3 SI index

While the DRASTIC model defines the hydrogeologic potential for contamination, the Susceptibility Index integrates anthropogenic pressure by incorporating land use data. This specific vulnerability assessment reveals a more nuanced risk landscape (Figure 18), highlighting where human activity spatially coincides with sensitive geological zones.

Overall, most of the area is categorised as low to very low risk with hotspots of moderate risk, as shown in Table 7:

Table 7: extent, in kilometres squared and percentage of the entire study area, of each SI categorization

Categorization	Fraction of the study area [km^2]	% of total area
Extremely low	*not present*	0%
Very low	5.77	59,7%
Low	2.78	28,8%
Moderate to low	0.69	7,10%
Moderate to high	0.40	4,10%
High	0.02	0,20%
Very high	0.01	0,10%
Extremely high	*not present*	0%

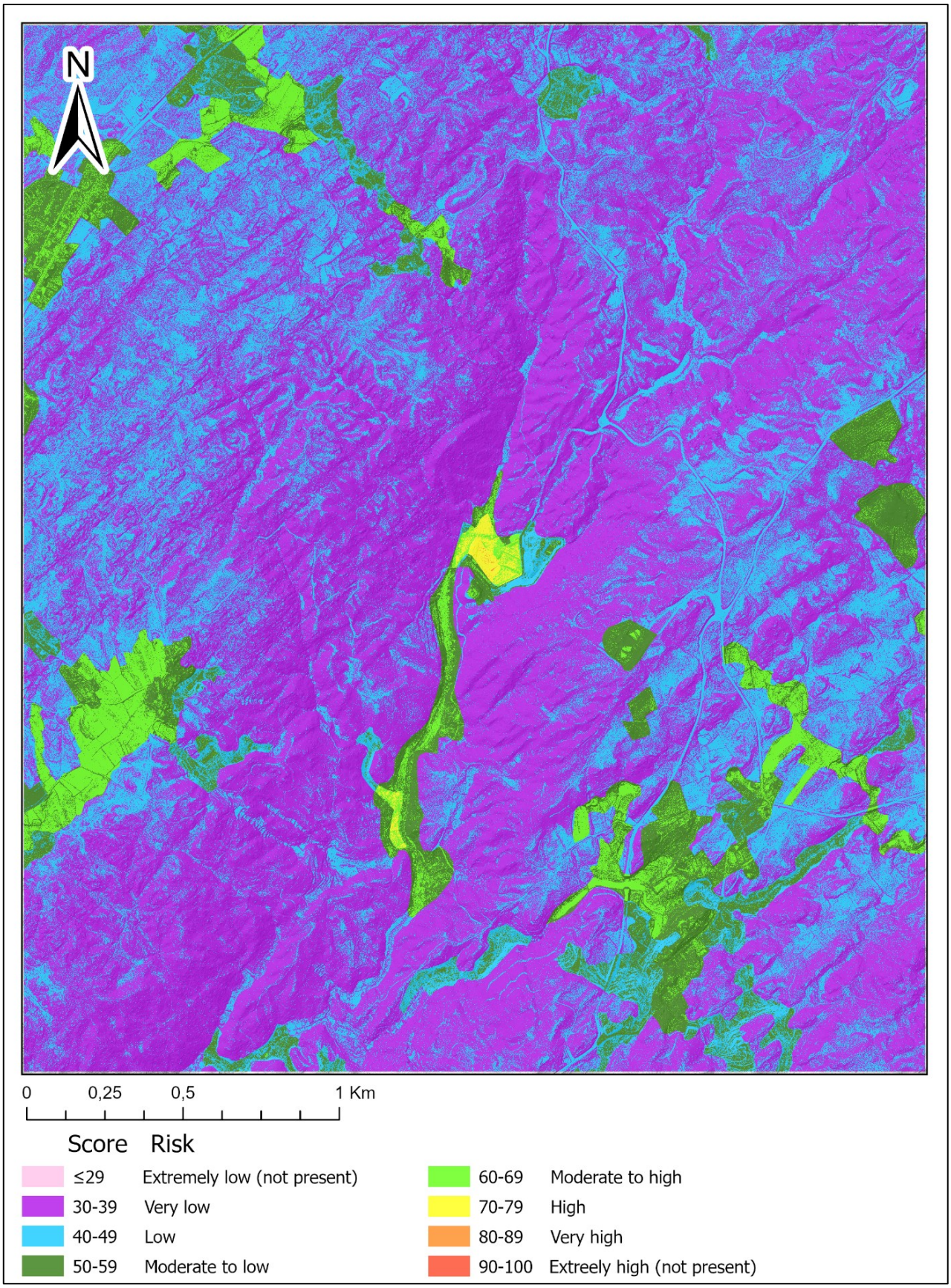

Figure 18: SI results over the study area

4 Discussion

While the DRASTIC and SI models provide a standardized assessment of pollution potential , the integration of LiDAR-derived vegetation metrics offers an independent line of evidence to validate the underlying hydrogeological conceptual model.

This chapter discusses the spatial correlations observed between surface biomass and subsurface conditions, specifically evaluating how canopy height mirrors the degree of granite weathering and fracture density. Furthermore, it examines the resulting DRASTIC and SI vulnerability maps, analyzing the "hydrogeological window" created by the central alluvial corridor. Finally, these findings are situated within the context of previous research in the Beiras Variscan granitic belt.

4.1 Vegetation disposition patterns

After visually reviewing the imagery, it is possible to take notice of some patterns in the disposition of the trees that do not appear to be random. In particular, areas with a more weathered geological composition tend to present trees of greater height, as shown in Figure 19 and Figure 21. The presence of water bodies also seems to influence the growth potential of the surrounding vegetation: Figure 20**Errore. L'origine riferimento non è stata trovata.** shows how the presence of a stream creates a contiguous row of higher-than-average trees.

However, to determine if the spatial distribution of vegetation within the study area is dependent on geological substrate, a statistical analysis was performed correlating the geological data with tree occurrence. The objective was to test the hypothesis that specific geological substrates (or their state of degradation) exert a significant influence on tree density and distribution.

Before proceeding with the statistical tests, a pre-processing step was applied to remove confounding anthropogenic variables. Since tree presence in urbanized or intensively cultivated zones is primarily driven by human management rather than environmental suitability, these areas would introduce bias into the geological assessment. To mitigate this, the Land Use vector layer was used as a mask. All polygons not classified as "Forests and semi natural zones" (Figure 13) were spatially subtracted from the geological layer. This reduced the

total study area from A=9.66km^2 to around A=8.47km^2, and the tree population from N=74791 to N=69682

Consequently, the subsequent statistical analyses were restricted exclusively to these zones. This ensures that the results reflect, to the greatest extent possible, the natural ecological preference of the vegetation for the underlying substrate rather than land-use history.

4.1.1 Preliminary assessment: total vegetation distribution (>2m)

The initial phase of the investigation quantified the total abundance of woody vegetation to establish a baseline for geological preference. The primary dataset consisted of the tree enumeration obtained previously, as shown in chapter 3.1, where a vertical threshold of 2 meters had already been applied to exclude low-lying herbaceous vegetation.

To test the null hypothesis that vegetation is distributed randomly across the landscape irrespective of substrate, a Chi-Square Goodness of Fit test was performed. The analysis yielded a statistically significant result (χ^2= 3.1x10^{-28}, $p < 0.001$, Table 8), formally rejecting the null hypothesis. However, while the distribution was technically non-random, the Preference Index for the dominant granitic formations showed little variation and remained close to neutrality (1.0), ranging only slightly from 0.97 (Weathered) to 1.06 (Fresh).

This lack of strong differentiation indicated that, at least for the 2-meter threshold, no strong correlation was in place between tree population and underlying geological strata. It was determined that a more refined analysis targeting established, mature trees would be necessary to further investigate the hypothesis.

4.1.2 Refined Analysis: Mature Canopy Distribution (>5m)

Following the inconclusive results of the total vegetation assessment, the analysis was refined to isolate established, mature trees. The vertical height threshold for the tree population was increased to 5 meters (Figure 14). This filtration reduced the sample size to N = 57516 treetops, representing the mature canopy stratum.

The Chi-Square Goodness of Fit test was repeated using this high-canopy dataset. The results showed an even stronger statistical significance (χ^2= 7.0x10^{-53}, $p < 0.001$, Table 8). Crucially,

the Preference Index revealed a distinct ecological gradient compared to the preliminary analysis. Mature trees exhibited a clear positive association with "Fresh" granitic substrates (Index = 1.13), while "Weathered" substrates showed a pattern of avoidance (Index = 0.95). This shift confirms that as the canopy height threshold increases, the influence of the geological substrate becomes more pronounced, with mature stands showing a clearer preference for less degraded lithologies. This is also in direct contrast with the initial hypothesis that weathered substrate would be linked to higher treetops.

Table 8: preliminary and refined analysis results for correlation between tree count and underlying geology

		Number of trees		Preference Index	
Substrate type	Total Area [km^2]	All trees	Trees >5m	All trees	Trees >5
Alluvium	0.10	722	640	0.8516	0.9145
Dolerite dykes	0.04	313	279	1.0063	1.0868
Fresh to slighlty weathered granite (W_{1-2})	1.00	8807	7736	1.0652	1.1335
Medium weathered granite (W_3)	2.52	21616	17896	1.0430	1.0461
Weathered granite (W_{4-5})	4.80	38224	30965	0.9670	0.9490
Total of above	8.47	69682	57516		

X^2	
All trees	Trees >5
3.1E-28	7.0E-53

4.1.3 Gradient analysis: geological degradation and tree density

To further investigate the observed preference for "Fresh" over "Weathered" substrates, a correlation analysis was performed to quantify the strength of this relationship. The granitic categories were reclassified into an ordinal scale representing the degree of weathering degradation (1=Fresh to 3=Weathered). Tree density (n/km^2) was calculated for each individual geological polygon within the natural zone mask. The degradation data is ordinal, so a Spearman's Rank Correlation (ρ) was selected, revealed a weak negative correlation (ρ = -0.18). The negative coefficient confirms the trend identified in the Chi-Square test: as the degree of geological weathering increases, tree density decreases. However, the low magnitude of the coefficient indicates a weak relationship, suggesting that while geological substrate is a contributing factor, tree distribution in the study area is likely primarily driven by other environmental variables such as slope or soil depth, or competition with human activity

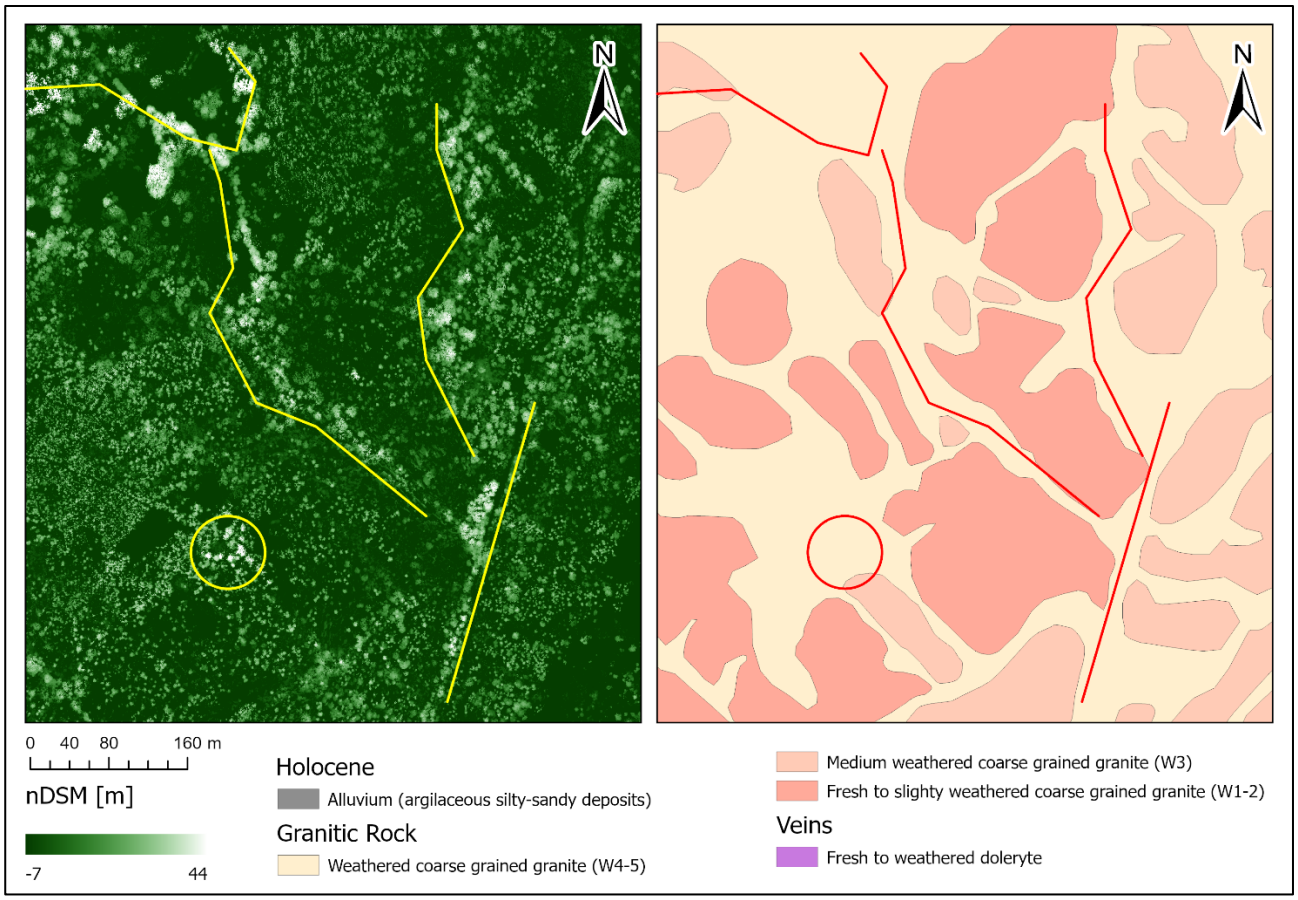

Figure 19: nDSM and geology: higher vegetation often overlaps with more weathered geology:

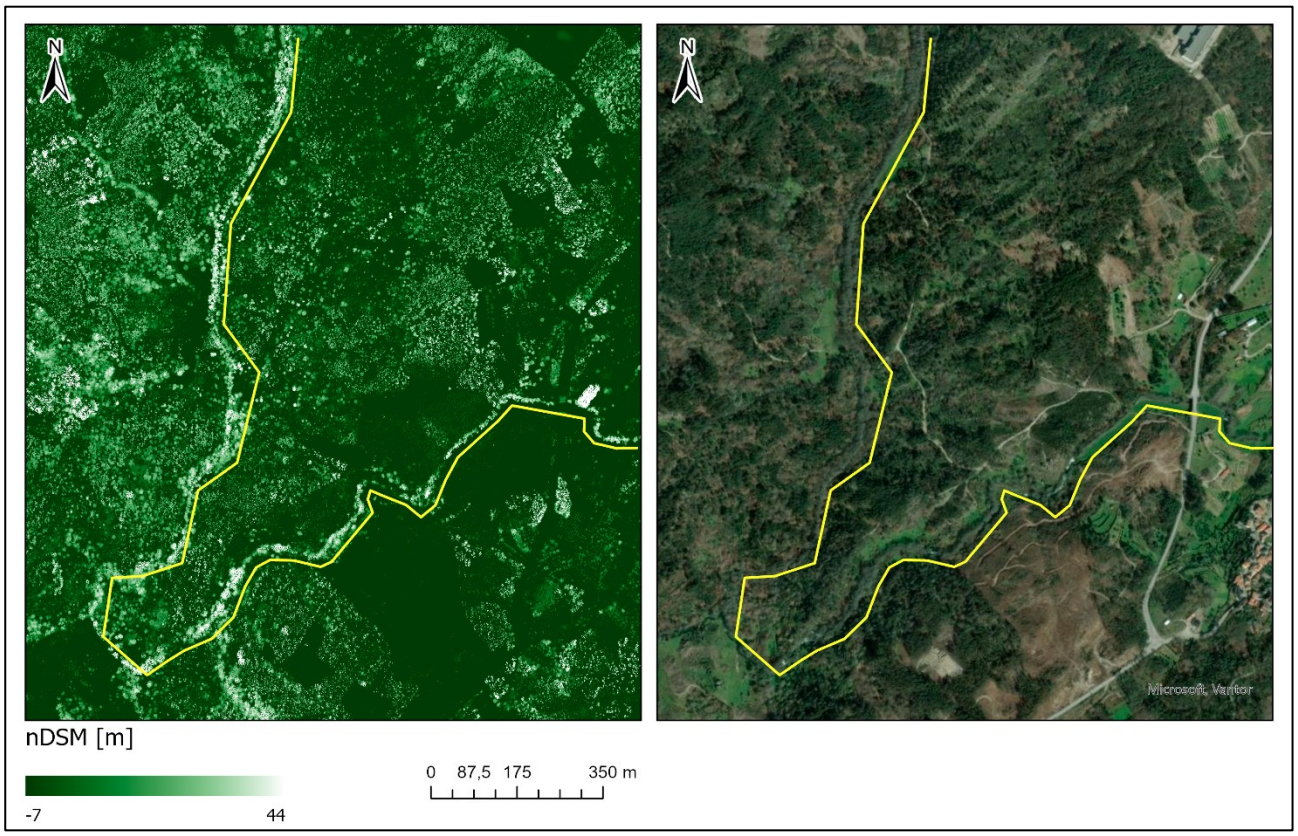

Figure 20: nDSM and hydrology: higher vegetation in correspondence with a creek

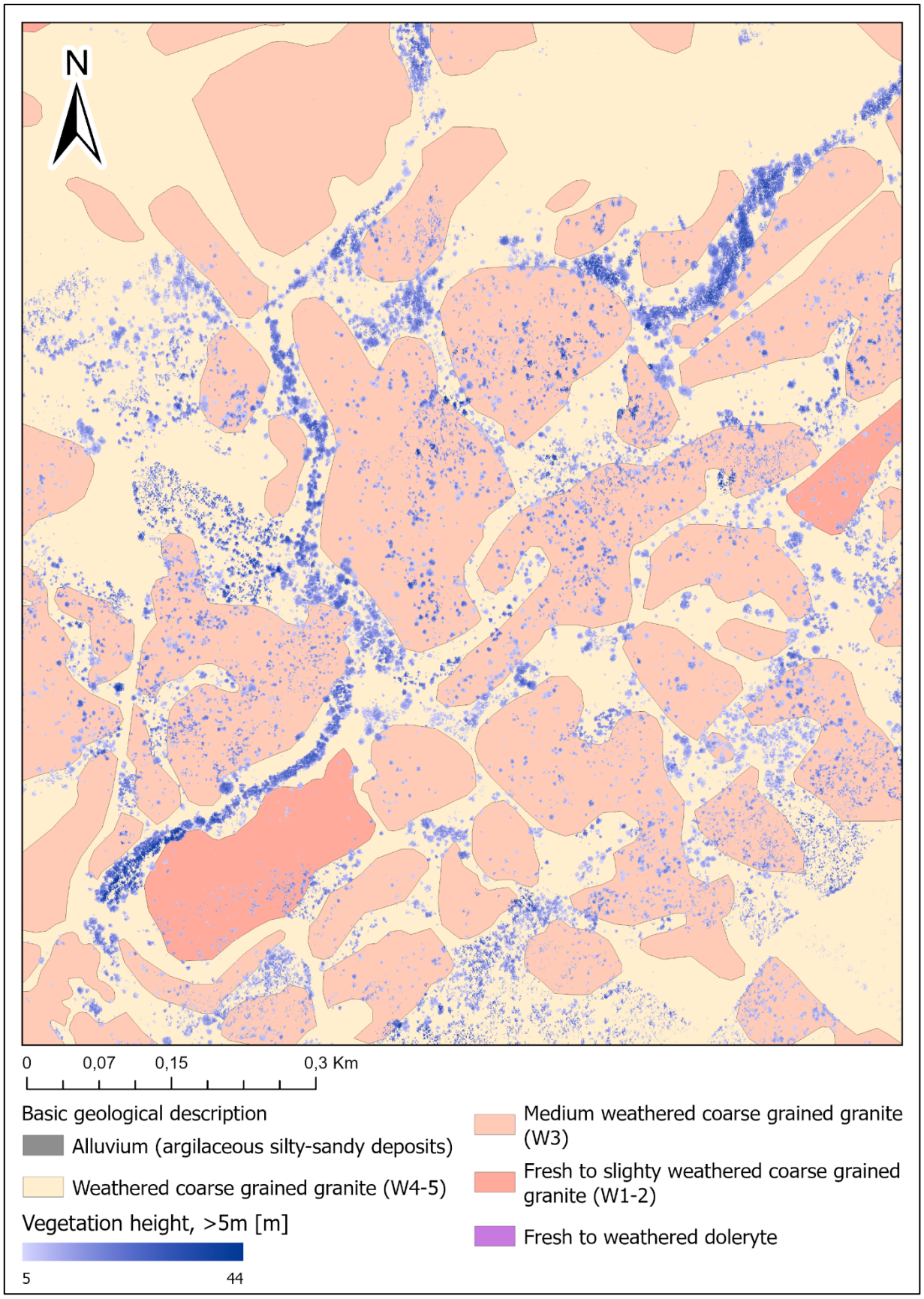

Figure 21: nDSM and geology: taller vegetation often overlaps with more weathered geology

4.2 Indices analysis

The results obtained through the application of the DRASTIC and SI indices provide a comprehensive view of the hydrogeological risk landscape in Caldas da Cavaca. This section analyses the performance of these models, examining how they differentiate between the natural protective barriers of the granitic highlands and the higher-risk hydrogeological windows found in the alluvial valleys.

By contrasting the intrinsic results of DRASTIC with the land-use-weighted scores of the SI, we can identify where anthropogenic pressures, such as urban centres and intensive agriculture, spatially coincide with the most fragile geological units.

4.2.1 DRASTIC

The results obtained via the DRASTIC index are consistent with what is expected: more permeable rock substrate, associated with flatter slope, shallower water table and higher recharge, result in higher risk.

4.2.1.1 The granitic bedrock zone

Surrounding the central alluvial channel, the DRASTIC index values decrease significantly, falling primarily into the "Moderate to Low" category with scores between 120 and 139 (indicated in blue), which covers 70,4% of the study area. This extensive zone correlates directly with the presence of Granitic Rock (specifically fresh to weathered coarse-grained granite, grades W1-W5) and Dolerite veins.

The lower vulnerability scores in this domain are attributable to the more robust natural barriers provided by the crystalline bedrock environment:

- Geological control: the aquifer media (A) in this region transitions to weathered metamorphic and igneous rock, receiving significantly lower ratings of 3, 4, and 5 (Figure 7). Unlike the porous alluvium, the hydraulic conductivity in this fractured media is generally lower and more heterogeneous, restricting the rapid spread of pollutant plumes.

- Hydraulic proximity: the water table in these upland areas is generally deeper, recorded between 4.9 and 9.1 meters, receiving a rating of 7 (Figure 5). This increased vertical distance provides a thicker buffer zone for natural attenuation processes to occur.
- Vadose zone attenuation: crucially, the vadose zone (I) in the granitic areas is dominated by silt and clay matrices (Ratings 5 and 6, Figure 12). The presence of fine-grained clays increases the cation exchange capacity (CEC) of the soil profile, enhancing the sub surface's ability to retard the migration of pollutants through adsorption.

Although the central alluvial corridor exhibits increased susceptibility to contamination, the study area as a whole does not reach 'High' or 'Extreme' vulnerability thresholds (Index >160). The absence of scores in the upper quartiles suggests that, while the central valley requires careful environmental management, the region benefits generally from significant natural attenuation provided by the surrounding aquitards.

4.2.1.2 The central alluvial zone

The most critical feature identified in the vulnerability assessment is a north-south trending corridor characterized by "Moderate to High" DRASTIC scores ranging from 140 to 159 (indicated in light green). This zone exhibits a near-perfect spatial coincidence with the Holocene geological deposits, specifically the Alluvium (argillaceous silty-sandy deposits) identified in the geological description.

The elevated vulnerability index in this sector is driven by the cumulative effect of high-risk ratings across the primary hydrogeological parameters:

- Geological control: the underlying lithology consists of unconsolidated sand and gravel, which was assigned a high aquifer media (A) rating of 8 (Figure 7). This coarse-grained composition facilitates rapid groundwater flow and offers minimal retardation to potential contaminants.
- Hydraulic proximity: this zone corresponds to the area of shallowest groundwater occurrence. The depth to water (D) parameter (Figure 5) here ranges between 1.5 and 4.6 meters, resulting in a maximum rating of 9. The proximity of the water table to the surface drastically reduces the travel time for surface pollutants to reach the saturated zone, which in turn further decreases the attenuation potential by the soil.

- Vadose zone attenuation: the impact of the vadose zone (I) further exacerbates vulnerability in this corridor. The unsaturated zone is composed primarily of sand and gravel (Rating 7, Figure 12), lacking the clay content necessary to effectively filter or adsorb contaminants before they enter the aquifer.

Consequently, the central alluvial valley represents a "hydrogeological window" where the natural protective capacity of the subsurface is minimal, making it the primary hotspot for potential groundwater contamination.

4.2.2 SI

This index shows different results everywhere the land cover differs from natural or semi-natural. Overall, the SI shows higher risk classifications.

4.2.2.1 The urban hotspots

The most significant deviation from the intrinsic DRASTIC model is the emergence of a localized "High" vulnerability zone, with spots of “Very high” vulnerability in the northern-central sector of the alluvial valley. This classification, absent in the DRASTIC results, represents the area of maximum specific risk within the study domain.

Spatial correlation with the Land Use map identifies the driver of this elevated risk:

This high-risk node directly overlays an area classified as "Continuous urban areas" and "Industrial or commercial activity" (red, dark red and green polygons in Figure 13).

The placement of these high-load activities atop the highly permeable Holocene alluvium (Figure 18) creates a "worst-case" scenario. The intrinsic vulnerability of the shallow aquifer (Depth to Water < 4.6m) is compounded by the potential discharge of persistent urban, industrial and agricultural pollutants.

4.2.2.2 The forested surroundings

Most of the study area, particularly the eastern and western highlands, is classified as "Very Low to "Low" vulnerability.

This extensive low-risk domain is maintained by the protective synergy of geology and land cover: these areas are dominated by "Forests and semi-natural zones" (Hatched Green texture in Figure 13). The absence of significant pollutant sources (fertilizers, urban runoff) in these natural zones keeps the specific vulnerability score minimal.

Even where discontinuous urban settlements appear in the western uplands, the SI score often remains in the "Moderate to Low" range (score 50-59) rather than spiking to "High." This demonstrates the buffering capacity of the Granitic Bedrock and the deeper water table, which naturally mitigate the impact of localized human activity.

The SI analysis refines the vulnerability model by demonstrating that the highest risk is not merely geological but also anthropogenic. The study area's critical management zone is the central alluvial valley, specifically where urban and intensive agricultural land uses intersect with the shallow, permeable aquifer: this convergence transforms a region of moderate intrinsic vulnerability into a localized zone of high specific risk.

4.2.3 Comparison with previous research

The spatial dichotomy identified between intrinsic and specific vulnerability in this study, especially in the north-western section of the study area, which report vastly different risk classifications, is consistent with the findings of Teixeira et al. (2014) and Meerkhan et al. (2016).

- DRASTIC Index (Intrinsic): This study identified the alluvial valleys and highly weathered granites (W_{4-5}) as the most geologically fragile zones. This mirrors the work of Teixeira et al. (2014), who noted that the weathered granitic mantle and alluvial deposits create a shallow phreatic system that is naturally susceptible to surface processes. The same work shows how the northwestern section of the study area, in particular the area around Quinta das Lameiras, has two different and non-compatible risk classifications in DRASTIC and SI;
- The transition to the Susceptibility Index (SI) demonstrates that agricultural and urban-industrial hotspots significantly amplify groundwater risk in naturally fragile areas. This finding aligns with the conclusions of Stigter et al. (2006), who proved, in other Portuguese contexts, that intrinsic models tend to underestimate the impact of

intensive land use, particularly regarding nitrate contamination. While Stigter et al. observed higher SI scores across the vast majority of the Faro region, due to pervasive agriculture, the overall DRASTIC and SI scores in Caldas da Cavaca remain lower and more similar to each other. This is because 87,6% of the area is occupied by "Forests and semi-natural zones," which receive a Land Use rating of 0. However, the underlying behaviour of the indices is identical in both studies: in every instance where land use deviates from natural cover, such as in urban centres or irrigation perimeters, the SI scores deviate sharply from DRASTIC, reflecting a localized but significant increase in specific vulnerability.

5 Conclusion

5.1 Indexes results

The application of the DRASTIC and SI models to the Caldas da Cavaca region revealed two distinct but complementary perspectives on groundwater protection.

The DRASTIC index, representing the natural defence capacity of the aquifer system, identified the alluvial valleys and zones of highly weathered granite as the areas of highest vulnerability.

The spatial distribution of risk is primarily controlled by the Aquifer Media (A) and Topography (T) parameters. The flat, permeable alluvial sectors facilitate rapid infiltration, offering little geological barrier to contaminants. Conversely, the steep slopes and fresh crystalline bedrock (W1-2) acted as natural protective barriers, resulting in "Low" to "Very Low" vulnerability scores in the upland areas due to high runoff potential and low primary permeability.

The transition to the Susceptibility Index (SI) introduced the anthropogenic reality of the study area.

In the northern agricultural sectors, the SI significantly amplified the vulnerability scores compared to DRASTIC. This shift confirms that the most geologically fragile areas (alluvium) are coincidentally the sites of highest pollutant loading (intensive irrigation and fertilization). In the forested sectors, the SI reduced the effective vulnerability. Even where the geology allows for infiltration, the absence of contaminant sources (LU rating of 0) effectively "masks" the intrinsic risk, highlighting the protective role of the arboreal cover.

5.2 Statistical Verification of Vegetation Distribution

The statistical integration of the Normalized Digital Surface Model (nDSM) with the geological data provided a quantitative validation of the relationship between substrate and vegetation structure. While the initial assessment of total vegetation volume, showed only a weak correlation with geological categories, the refined analysis of the mature canopy stratum (>5m) revealed a somewhat statistically significant ecological signal.

Contrary to the expectation that deep weathering profiles would support higher biomass due to water retention, the analysis identified a preference for Fresh to Slightly Weathered

granites (Preference Index >1.1). Conversely, the zones of Deep Weathering and Alluvium were associated with lower-than-expected tree densities (Index < 1.0). This was further confirmed by a Spearman's Rank correlation, which yielded a negative coefficient ($\rho = -0.18$), indicating that tree density generally declines as the degree of weathering increases.

This pattern lets us hypothesize that in the Caldas da Cavaca region, the distribution of mature trees is driven less by hydrogeological suitability and more by anthropogenic land-use pressure. The deeper, fertile saprolite of the weathered zones is usually more useful for a variety of anthropic activities, leading to forest clearance. In contrast, the "Fresh" granitic zones, which are characterized by steeper slopes and thinner soils, might have acted as ecological refugia, preserving mature woodland simply because these areas were unsuitable for other uses.

5.3 Implications for land management

The combination of the DRASTIC and SI maps provides a roadmap for local spatial planning: the areas classified as "High" in both models represent the critical intersection of geological weakness and agricultural pressure. The "Low" vulnerability of the forested areas is conditional on the maintenance of the current land cover. Deforestation or conversion of these zones to agriculture would strip away the this protection, exposing the underlying aquifer to the intrinsic vulnerability predicted by the DRASTIC model.

5.4 Limitations of the Study

While the multi-scalar approach proved effective, certain limitations must be acknowledged:

- The LiDAR survey represents a single temporal snapshot. Multi-temporal data or radiometric analysis (NDVI, RE, narrowband etc.) would allow for the assessment of vegetative stress, which could further refine the identification of groundwater-dependent ecosystems.
- The generalization required for the regional Land Use map (1:25,000) may mask small-scale point sources of pollution that are visible in the high-resolution LiDAR data.
- Although the relationship between geology and tree density was statistically significant, the correlation strength was weak. This implies that while geological

substrate is a contributing factor, it is likely a secondary driver. Topographic variables (slope, aspect) and historical land management likely exert a stronger influence on the current distribution of the forest canopy.

5.5 Bibliography

Aller, Linda, Jay Lehr, and Truman Bennet. “Drastic: A Standardized System to Evaluate Groundwater Pollution Potential using Hydrogeologic Setting.” *Geological Society of India*, 1987: 23-37.

Barbulescu, Alina. “Assessing Groundwater Vulnerability: DRASTIC and DRASTIC-Like Methods: A Review.” *Water*, 2020.

Carter, Jamie, et al. *Lidar 101: An Introduction to Lidar Technology, Data, and Applications.* Charleston: NOAA Coastal Services Center, 2012.

Direção-Geral do Território. “Carta de Uso e Ocupação do Solo para 2018.” 2018.

—. *Dados LiDAR de Portugal continental.* n.d. https://snig.dgterritorio.gov.pt/rndg/srv/por/catalog.search#/metadata/077a8c94-8b46-4a8a-8796-0d7fc4662f0c (accessed 12 08, 2025).

Direção-Geral do Território. “Especificações técnicas da Carta de Uso e Ocupação do Solo (COS) de Portugal Continental para 2018.” 2019.

ESRI. *Con Function*. n.d. https://pro.arcgis.com/en/pro-app/latest/help/analysis/raster-functions/con-function.htm (accessed 12 10, 2025).

—. *How Focal Statistics work.* n.d. https://pro.arcgis.com/en/pro-app/latest/tool-reference/spatial-analyst/how-focal-statistics-works.htm (accessed 10 12, 2025).

—. *How Point to Raster works.* n.d. https://pro.arcgis.com/en/pro-app/latest/tool-reference/conversion/how-point-to-raster-works.htm (accessed 12 10, 2025).

—. *Slope (3D Analyst).* n.d. https://pro.arcgis.com/en/pro-app/latest/tool-reference/3d-analyst/slope.htm (accessed 12 18, 2025).

Francés, Paralta, Fernandes, and Ribeiro. “Development and application in the Alentejo region of a method to assess the vulnerability of groundwater to diffuse agricultural pollution: the susceptibility index.” 2002.

Giglierano, James. *Lidar Basics for Mapping Applications.* Manual, Iowa City: Iowa Geological Survey, n.d.

Lohani, Bharat. “Airborne Altimetric LiDAR: Principle, Data collection, processing and Applications.” University lecture material, Kanpur, n.d.

Meerkhan, Helen, José Teixeira, Jorge Espinha Marques, Maria José Afonso, and Helder Chaminé. “Delineating Groundwater Vulnerability and Protection Zone Mapping in Fractured Rock Masses: Focus on the DISCO Index.” *Water* (MDPI), 2016.

Melin, Markus, Aurélie Shapiro, and Paul Glover-Kapfer. *LiDAR for ecology and conservation*. Woking: WWF, 2017.

O'Leary, Friedman, and Pohn. “Lineament, linear, lineation: Some proposed new standards for old terms.” *Geological Society of America Bulletin: Vol. 87*, 1976: 1463-1469.

Ribeiro. “IS: um nuovo índice de susceptibilidade de aquíferos á contaminaçao agrícola.” Lisbon, 2000.

Scesi, Laura, and Paola Gattinoni. *La circolazione idrica negli ammassi rocciosi.* Casa Editrice Ambrosiana, 2007.

Stigter, Ribeiro, and Carvalho Dill. "Evaluation of an intrinsic and a specific vulnerability assessment method in comparison with groundwater salinisation and nitrate contamination levels in two agricultural regions in the south of Portugal." *Hyrogeology Journal*, 2005.

Teixeira, et al. "A comprehensive analysis of groundwater resources using GIS and multicriteria tools (Caldas da Cavaca, Central Portugal): environmental issues." *Environmental Earth Sciences* (Springer), 2014.

Teixeira, et al. "Integrated approach of hydrogeomorphology and GIS mapping to the evaluation of groundwater resources: An example from the hydromineral system of Caldas da Cavaca, NW Portugal." In *Global Groundwater Resources and Management*, 227-249. Oslo: B. S. Paliwal, 2008.

Ulusay, Resat. *The ISRM Suggested Methods for Rock Characterization, Testing and Monitoring: 2007-2014.* Springer Cham, 2015.

Vrba, and Zaporozec. In *Guidebook on Mapping Groundwater Vulnerability, vol 16*, 131. Hannover, 1994.

Wang, Cheng, Xuebo Yang, Xi Xiaohuan, Sheng Nie, and Pinliang Dong. *Introduction to LiDAR Remote Sensing.* Boca Raton: CRC Press, 2024.