Science of the Total Environment 1043 (2026) 181948

Contents lists available at ScienceDirect

Science of the Total Environment

journal homepage: www.elsevier.com/locate/scitotenv

Disentangling agricultural and non-agricultural drivers of groundwater nitrate contamination using an integrated footprint–statistical approach

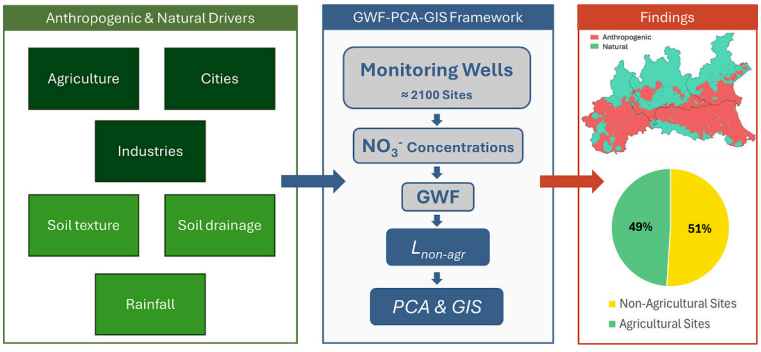
Diego Voccia [a,*], Giacomo Mortella [b], Marco Trevisan [a], Bartolo Forestieri [a], Lucrezia Lamastra [a]

[a] Department of Sustainable Food Process, Università Cattolica del Sacro Cuore, 29122, Piacenza, Italy
[b] Department of Sustainable Crop Production, Università Cattolica del Sacro Cuore, 29122, Piacenza, Italy

HIGHLIGHTS

- Integrated GWF, PCA and GIS to assess groundwater nitrate pollution
- Non-agricultural N loads exceed 10 kg/ha yr^{-1} in >50% of sites
- Nitrate contamination not solely driven by agricultural inputs
- Natural factors (rainfall, soil) strongly control nitrate transport
- Framework improves source apportionment in complex agro-urban systems

GRAPHICAL ABSTRACT



ARTICLE INFO

Keywords:
Grey water footprint
Nitrates
Groundwater pollution
Principal component analysis

ABSTRACT

Nitrate (NO_3^-) contamination of groundwater is a widespread environmental and public health concern, particularly in intensively managed agricultural regions. This study investigates the spatial and temporal dynamics of groundwater NO_3^- pollution across the Po Valley (northern Italy), integrating the Grey Water Footprint (GWF) framework with multivariate statistical and geospatial analyses. Using data from approximately 2100 monitoring sites between 2013 and 2023, nitrogen loads to groundwater were estimated and partitioned into agricultural and non-agricultural contributions. Results indicate that while agriculture remains an important source of NO_3^- pollution—especially in Emilia-Romagna and Piedmont—it does not fully explain observed contamination patterns. Non-agricultural sources, including urban and industrial activities, contribute significantly, particularly in Lombardy and, to a lesser extent, Veneto. The estimated non-agricultural nitrogen load ($L_{non\text{-}agr}$) exceeded 10 kg N/ha/yr in over 50% of sites across the study area, highlighting the relevance of diffuse anthropogenic pressures beyond farming. Principal Component Analysis (PCA) reveals that both anthropogenic factors (population density, industrial activity, and agricultural intensity) and natural drivers (precipitation, soil texture, and drainage) jointly control NO_3^- distribution. Natural factors, especially precipitation and soil properties, often play a dominant role in regulating nitrate transport and accumulation, explaining regional differences that cannot be attributed solely to nitrogen application rates. Temporal analysis further shows that the relative importance of these drivers varies over time, influenced by climatic variability. Uncertainty analysis underscores the influence of poorly constrained parameters, particularly groundwater recharge and natural

* Corresponding author.
E-mail address: diego.voccia@unicatt.it (D. Voccia).

https://doi.org/10.1016/j.scitotenv.2026.181948
Received 15 April 2026; Received in revised form 27 May 2026; Accepted 5 June 2026
Available online 10 June 2026
0048-9697/© 2026 The Authors. Published by Elsevier B.V. This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/).

background NO_3^- concentrations, on model outputs. Overall, the study demonstrates that groundwater nitrate pollution in the Po Valley arises from a complex interplay of sources and processes. These findings highlight the need for integrated management strategies that address both agricultural and non-agricultural inputs while accounting for hydrogeological and climatic conditions.

1. Introduction

Nitrate (NO_3^-) contamination of groundwater represents one of the most widespread threats to drinking water quality worldwide, posing significant risks to human health and aquatic ecosystems (Zhang et al., 2018). Although NO_3^- is not inherently toxic, its reduction to nitrite and subsequent formation of nitroso compounds in the human body has been associated with adverse health outcomes, including methemoglobinemia and increased cancer risk (Chamandoost et al., 2016). For this reason, strict regulatory limits have been established in many countries, including the European Union threshold of 50 mg/L set by Directive 91/676/EEC.

Nitrogen (N) is an essential nutrient for agricultural production, and the widespread use of synthetic fertilisers and manure has played a central role in sustaining global food security (Rütting et al., 2018). However, inefficient N uses and excessive application rates can lead to substantial environmental losses, particularly through NO_3^- leaching to groundwater. These losses are strongly influenced not only by fertilizer inputs but also by soil texture, drainage conditions, irrigation practices, and climatic factors such as precipitation intensity and distribution (Hina, 2024). As a result, similar agricultural N inputs can lead to markedly different groundwater contamination patterns across regions. While agriculture is often identified as the dominant source of groundwater NO_3^- pollution, numerous studies have demonstrated that urban and industrial activities can also contribute significantly (Rahman et al., 2021; Šrajbek et al., 2022). Leakage from sewage systems, inadequately treated wastewater, fertilization of urban green areas, industrial effluents, and landfill leachates may all increase NO_3^- concentrations in aquifers (Rahman et al., 2021). Disentangling the relative contributions of these multiple sources remains challenging, particularly in densely populated and intensively managed landscapes where point and diffuse sources coexist and interact with complex hydrogeochemical processes.

The Po Valley in northern Italy represents a paradigmatic case of these challenges. This region hosts one of Europe's most productive agricultural systems due to a relatively high water presence (Monteleone and Borzí, 2024) and relies heavily on groundwater as a primary source of drinking water, posing a potential risk (Ducci et al., 2026; Monteleone and Borzí, 2024). At the same time, it is characterised by high population density, extensive industrial development, and heterogeneous hydrogeological settings (Mattsson et al., 2025). Previous studies conducted in the Po Valley have reported widespread NO_3^- contamination and have primarily attributed it to agricultural practices (Severini et al., 2020). However, such interpretations often rely on simplified assumptions that equate NO_3^- presence with agricultural origin, potentially overlooking the role of non-agricultural and natural drivers.

Robust and communicable indicators are therefore required to assess groundwater pollution and to support evidence-based management and policy decisions. Among the available tools, the Water Footprint framework has gained increasing attention for its ability to link human activities to water quantity and quality impacts in an intuitive manner (Maranan et al., 2025). In particular, the Grey Water Footprint (GWF) quantifies the volume of freshwater required to dilute pollutants to acceptable concentrations and, when combined with the Water Pollution Level (WPL), provides insight into whether natural water recharge is sufficient to assimilate pollutant loads (Hoekstra et al., 2012). Despite its potential, the application of GWF and WPL to groundwater NO_3^- contamination remains limited, especially at large spatial scales and with explicit separation of agricultural and non-agricultural sources. In parallel, multivariate statistical techniques such as Principal Component Analysis (PCA) have been increasingly used to investigate groundwater contamination patterns and to infer potential pollution sources (Li et al., 2020, 2025; Liu et al., 2025; Ruan et al., 2024; Zhang et al., 2020). While effective in identifying dominant gradients and associations among variables, PCA-based studies often assume a priori that NO_3^- is an agricultural tracer, which may bias source attribution in complex socio-environmental contexts.

Against this background, the present study aims to improve the understanding of groundwater NO_3^- contamination in the Po Valley by integrating GWF accounting with multivariate and geospatial analyses. While specialized hydro chemical tracers or isotopic analyses represent the traditional 'gold standard' for source attribution (Blarasin et al., 2020), their application is frequently constrained at the regional scale by high costs and data fragmentation. Furthermore, many secondary ions (such as chloride, ammonium or sulfate) are often influenced by multiple, overlapping sources (e.g., road salts, geogenic salinity, or varying industrial effluents), which can confound the statistical 'disentangling' of specific agricultural and urban signals. Therefore, this work proposes a parsimonious, data-driven framework. This approach is designed not only as a scalable alternative for large-scale assessments where tracers are unavailable but also as a targeted methodology to avoid the interpretive uncertainties associated with multi-source chemical parameters. Specifically, this work (i) estimates spatially explicit N loads to groundwater across four regions of the Po Valley using the GWF–WPL framework, (ii) separates agricultural and non-agricultural contributions to NO_3^- leaching by leveraging high-resolution land-use indicators as physical proxies, and (iii) investigates the relative influence of anthropogenic and natural factors through Principal Component Analysis combined with GIS-based spatial interpolation. By moving beyond single-source assumptions and complex tracer requirements, this study seeks to provide a more comprehensive, transferable, and robust assessment of groundwater NO_3^- pollution in one of Europe's most critical agricultural and hydrological regions.

2. Materials and methods

2.1. Data collection

Groundwater NO_3^- levels, soil textures, cultivated area and regional N application rate for the regions of the Piedmont, Lombardy, Veneto and Emilia-Romagna were taken from Agenzia Regionale Prevenzione Ambiente (ARPA) database from 2013 to 2023 (Arpa Emilia Romagna, 2026; ARPA Lombardia, 2026; ARPA Piemonte, 2026; Arpa Veneto, 2026). All ARPA available sites (approximately 2100) were selected and reported in Fig. 1. Annual rainfall was obtained from Isprambiente database (Isprambiente, 2026) from 2013 to 2023 as total annual accumulated precipitation interpolation for each single year. Agricultural statistics like fertilisers N content total and crop-specific cultivated land were all collected from the Italian Institute of Statistics (ISTAT) (Istat, 2026). Municipality population densities were collected from the European Institute of Statistics (EUROSTAT) (Eurostat, 2026). Agricultural and industrial municipality relative area were extrapolated from Corine Land Cover 2018 dataset (European Environment Agency, 2020), and reported in Fig. 1. A natural soil drainage 1:250,000 resolution map was obtained by (Rivieccio, 2019) and used as a georeferenced raster layer in QGIS Geographic Information System, Version 4.0.0 (QGIS). Groundwater recharge coefficients were obtained from a 1 km spatial resolution map by (Martinsen et al., 2022) and used as a raster layer in QGIS for site specific value extraction.

2.2. Area of study and site selection

The study area comprises four regions of northern Italy—Piedmont, Lombardy, Veneto, and Emilia-Romagna—covering the main extent of the Po Valley (≅ 89,077 km^2), one of the most important agricultural and hydrogeological systems in Europe (Monteleone and Borzí, 2024). Piedmont is in the western sector of the study area and is bounded by the Alps to the west and north and by the Apennines to the south. Its lowland area forms the western portion of the Po Plain and hosts one of the most significant groundwater reserves of northern Italy, characterised by extensive alluvial aquifers supplying drinking water and irrigation demands (De Luca et al., 2020). Lombardy occupies the central part of the Po Valley, extending between the Alpine foothills and the Po River. This region includes a highly permeable alluvial plain that has been widely recognized as particularly vulnerable to NO_3^- leaching due to intensive agricultural practices, high nitrogen inputs, and favourable hydrogeological conditions for solute transport (Sacchi et al., 2013). Veneto is situated in the eastern part of the study area, extending from the Adriatic coast to the Alpine chain. The region exhibits marked geomorphological heterogeneity, with approximately half of its territory characterised by mountainous and hilly landscapes and the remainder consisting of intensively cultivated plains (Gioia et al., 2021). Emilia-Romagna displays a pronounced geographical contrast, encompassing a northern alluvial plain, southern Apennine uplands, and an eastern coastal zone along the Adriatic Sea. The region has a long-standing agricultural tradition and represents one of Italy's most productive farming areasn (Ungaro et al., 2025). Monitoring sites were selected based on sampling frequency to ensure seasonal representation; specifically, only wells with biannual data (corresponding to spring and fall campaigns) were included. The sampling network primarily targeted shallow aquifers, which are most vulnerable to anthropogenic nitrogen inputs. Monitoring site densities were relatively uniform across the study area, with values of 2.3, 2.5, 1.7, and 2.7 per 100 km^2 for Piedmont, Lombardy, Veneto, and Emilia-Romagna, respectively. This distribution ensures balanced spatial coverage and minimizes sampling bias across the four regions. The entire area is characterised by cereal cultivation (mainly corn, wheat and rice) and sugarbeet. The wigrape vineyards are consistentand e approximately 30% of the total vineyards cultivated area in Italy (Istat, 2026). Swine and bovine are the major livestockin the area, with swine being approximately 60% of the total head, followed by bovine the 30% (Istat, 2026).

2.3. Grey Water Footprint use for Nitrogen Load estimation

N Load (L_N expressed as kg N/ha/yr) was estimated starting from NO_3^- concentration in groundwater sites as reported by (Voccia et al., 2022), assuming a starting groundwater thickness (R_0) of 500 mm as reported in Eq. 1. This assumption was made considering a top layer of the aquifer varying between 0 and 1000 mm, with an average value of 500 m, in contrast to an empty basin assumed by (Voccia et al., 2022). This approach assumes that the amount of NO_3^- before the recharge is negligible.

$$L_N = \frac{R_0 + R}{100} \cdot 14 \cdot \frac{C - C_{nat}}{62} \quad (1)$$

Where R is the annual groundwater aquifer recharge R (mm) and calculated from Eq. 2, C is the groundwater NO_3^- concentration in each sampling site (mg/L) and C_{nat} is the natural concentration in groundwater. For C_{nat} a 5 mg/L value was assumed as an average value of the natural concentration range usually reported in literarture (0.4–10 mg/L) (Buškulić et al., 2023; Franke et al., 2013).

$$R = P \cdot RC \quad (2)$$

Where P is the annual total precipitation (mm), RC is the recharge coefficient obtained from (Martinsen et al., 2022). To evaluate the non-agricultural contribution to the contamination level $L_{non\text{-}agr}$ (kg N/ha/yr), Eq. 3 was considered:

$$L_{non-agr} = L_N - L_{agr} \quad (3)$$

where L_{agr} (kg N/ha/yr) is the agricultural N leached in groundwater, that was estimated according to Eq. 4 as reported in (Franke et al., 2013):

$$L_{agr} = \alpha \cdot Dose \quad (4)$$

where $Dose$ is the amount of applied N (kg N/ha/yr) estimated from provincial nitrogen average application (that include synthetic and sludge N) and cultivated land for each province district of the studied area and α is the fraction of applied N reaching groundwater. Fraction α has been evaluated according to (Franke et al., 2013) through Eq. 5, with s_i being the score ranging from 0 to 1 and w_i being the weight associated with atmospheric input (N-deposition), soil proprieties (textures and drainage), climate (precipitation) and agricultural practice (N-fixation, application rate, plant uptake and management), all summarised in Table 1.

$$\alpha = 0.01 + 0.24 \cdot \frac{\sum_i s_i \bullet w_i}{\sum_i w_i} \quad (5)$$

The N-deposition score was set to 1 for all sites, based on values reported for the study area by (Leip et al., 2011) (> 1.5 g N/m^2/yr). Due to uncertainties in N-fixation, application rates, and plant uptake, an intermediate score of 0.5 was adopted, following the approach of (Franke et al., 2013) where no site-specific data are available.

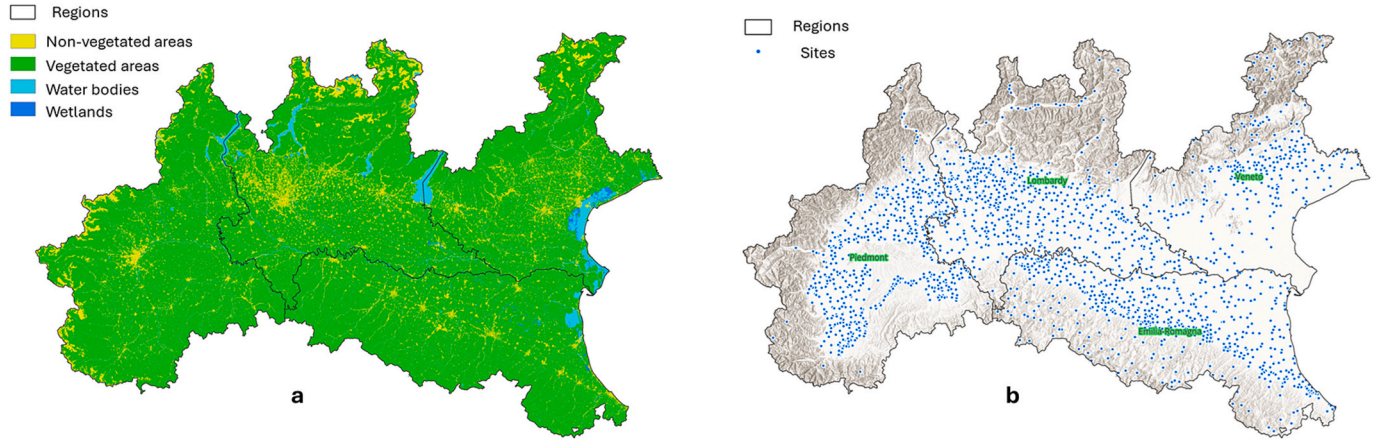


Fig. 1. – Land cover (a) and topography and sites distribution among the four regions (b).

Table 1
– Factor *a* weights and scores [34].

		Score	0	0.33	0.67	1
		Weight	Very low	Low	High	Very high
Atmospheric Input	N-deposition ($gN/m^2/yr$)	10	< 0.5	> 0.5	< 1.5	> 1.5
Soil	Texture for leaching	15	Clay	Silt	Loam	Sand
	Texture for runoff	10	Sand	Loam	Silt	Clay
	Drainage for leaching	10	Poorly to very poorly drained	Moderately to imperfectly drained	Well drained	Excessively to extremely drained
	Drainage for runoff	5	Excessively to extremely drained	Well drained	Moderately to imperfectly drained	Poorly to very poorly drained
Climate	Precipitation (mm)	15	0–600	600–1200	1200-1800	> 1800
Agricultural practice	N-fixation (kg/ha)	10	0	> 10	< 60	> 60
	Application rate	10	Very low	Low	High	Very high
	Plant uptake	5	Very low	Low	High	Very high
	Management practice	10	Best	Good	Average	Worst

Management practice was assigned a score of 0.33, representing a conservative estimate between 'best' and 'good' practices for an industrialized country such as Italy (Franke et al., 2013). All other scores were derived from site-specific parameters—soil texture, precipitation, and soil drainage—extracted from QGIS maps.

The locations of non-agricultural sites were assessed based on municipal population density, the proportion of cultivated land, and the proportion of industrial land, to evaluate the anthropogenic contribution to non-agricultural groundwater contamination. Additionally, to account for potential non-anthropogenic influences, factors such as precipitation, soil texture, and soil drainage were included in a Principal Component Analysis (PCA).

2.4. Statistical analysis

Statistical analyses were performed using RStudio (version 2022.02.0 + 443; R Core Team, 2024) to investigate correlations between natural variables (soil texture, soil drainage, and precipitation) and anthropogenic variables, including municipal population density and the proportions of cultivated and industrial land. Pearson's correlation coefficients were computed to quantify the strength and direction of these relationships, with statistical significance assessed at $p < 0.05$.

The suitability of the dataset for factor analysis was evaluated using the Kaiser–Meyer–Olkin (KMO) measure and Bartlett's test of sphericity (Li et al., 2020). KMO values closer to 1 indicate stronger intervariable correlations and greater reliability of factor analysis results. Values above 0.5, combined with a significant Bartlett's test ($p < 0.05$), are generally considered acceptable for proceeding with factor and principal component analyses (Haji Gholizadeh et al., 2016).

Two clustering approaches were applied: regional and temporal. In the regional approach, Piedmont, Lombardy, Veneto, and Emilia-Romagna were analyzed based on average $L_{non\text{-}agr}$ values for the period 2013–2023. These were classified into three categories: negligible (< 10 kg N/ha, NL), medium (< 50 kg N/ha, ML), and high (> 50 kg N/ha, HL). The temporal approach evaluated each year individually from 2013 to 2023. In both approaches, cluster centroids were calculated as the mean of all observations within each cluster.

Uncertainty in $L_{non\text{-}agr}$ was assessed through 10,000 Monte Carlo simulations considering random input values for NO_3^- concentration C, recharge R, L_{agr}, C_{nat} and groundwater thickness R_0. As the true statistical distributions of the input variables were unknown, the following assumptions were adopted:

- NO3− concentration C in wells, recharge R from precipitations and Lagr, varied with a triangular distribution centered on the mean value (μ) between a min value of $\mu - 2.45 \times \sigma$ and a max value of $\mu + 2.45 \times \sigma$, where σ is the standard deviations based on the 2013–2023 averages;
- Cnat was considered varying uniformly from 0.4 to 10 mg/L;
- R0 was considered varying from 0 to 1000 mm.

Sensitivity analysis was performed in R by calculating partial rank correlation coefficients (PRCC) between input variables and model outputs.

3. Results and discussion

3.1. Regional groundwater NO_3^- concentrations and $L_{non\text{-}agr}$

Comprehensive data on NO_3^- concentrations across the four regions are presented in Table 2 and via spatial distribution in Fig. 2. The table reports the number of monitoring sites where NO_3^- concentrations exceeded the regulatory limit of 50 mg/L between 2013 and 2023, expressed as an annual average, as well as the 95th percentile of NO_3^- concentrations for each region within the study area.

Emilia-Romagna exhibits the highest proportion of sites exceeding the legal threshold, with an average of 10% over the 2013–2023 period. Piedmont follows with an average of 6%, while Lombardy and Veneto show lower percentages of 3% and 2%, respectively. This marked difference may initially suggest that Emilia-Romagna experiences higher levels of groundwater contamination compared to the other regions.

Further analysis supports this interpretation: the 95th percentile of NO_3^- concentrations for the 2013–2023 period reaches 64.69 mg/L in Emilia-Romagna, the highest among the regions considered. This finding reinforces the conclusion that groundwater in Emilia-Romagna is more affected by NO_3^- pollution relative to other study areas.

Based on ISTAT data for the 2013–2023 period, average agricultural nitrogen loads of 81, 124, 134, and 138 kg N/ha/yr were calculated for Piedmont, Emilia-Romagna, Veneto, and Lombardy, respectively. These values do not fully correspond to the spatial patterns reported in Table 2, suggesting that agriculture alone may not be the dominant driver of NO_3^- contamination in the study area. Northern Italy is widely recognized for its intensive agricultural activity, largely supported by the abundant water resources of the Po River basin (Xue et al., 2024). Nevertheless, while agriculture remains a major source of NO_3^- inputs to groundwater, urban and industrial areas may also contribute significantly through processes such as leakage from sewage and water supply systems, fertilization of urban green spaces, and the presence of landfills (Lerner, 2003; Wakida and Lerner, 2005). Spatial analysis (Fig. 2) reveals a clear correlation between elevated NO_3^- levels and proximity to urban and industrial artificial areas.

Although NO_3^- concentrations provide a useful indicator of pollution levels, the groundwater footprint (GWF) and/or water pollution level (WPL) offer a more comprehensive assessment of groundwater quality. Furthermore, distinguishing between agricultural and non-agricultural potential nitrogen loads provides deeper insight into the relative contributions of different anthropogenic sources of groundwater contamination. For this reason, $L_{non\text{-}agr}$ was estimated for each monitoring site across all regions for the 2013–2023 period, following an approach like

Table 2
– Percentage of ARPA sites coupled according to NO_3^- concentration level ranges.

	2013	2014	2015	2016	2017	2018	2019	2020	2021	2022	2023	Average
	Piedmont											
Nr. Sites	507	490	507	549	509	532	518	326	455	285	284	451
< 10 mg/L	36%	34%	37%	43%	38%	42%	44%	32%	43%	37%	36%	38%
10–25 mg/L	31%	34%	33%	33%	34%	30%	30%	35%	33%	38%	35%	33%
25–50 mg/L	27%	23%	24%	20%	23%	23%	20%	26%	20%	21%	25%	23%
> 50 mg/L	7%	9%	6%	4%	5%	5%	6%	8%	4%	5%	4%	6%
Av. (mg/L)	20.52	20.90	20.09	16.98	18.86	18.48	17.73	21.73	16.73	17.73	18.03	18.89
95th (mg/L)	56.55	54.24	53.85	46.15	49.18	50.05	51.46	57.05	46.73	46.94	49.27	51.04
	Lombardy											
Nr. Sites	398	497	418	458	456	436	387	460	452	378	372	428
< 10 mg/L	47%	44%	44%	49%	48%	49%	43%	47%	48%	50%	51%	47%
10–25 mg/L	22%	23%	23%	23%	24%	24%	27%	26%	27%	24%	24%	24%
25–50 mg/L	27%	29%	29%	26%	24%	24%	26%	24%	22%	22%	23%	25%
> 50 mg/L	4%	4%	3%	3%	4%	3%	5%	3%	3%	4%	2%	3%
Av. (mg/L)	17.16	18.13	17.55	15.89	16.15	15.96	17.42	15.89	15.61	15.14	14.88	16.34
95th (mg/L)	46.86	48.48	46.75	46.12	46.29	47.25	47.59	44.29	44.22	44.99	44.29	46.10
	Veneto											
Nr. Sites	272	275	278	285	289	289	286	285	282	276	277	281
< 10 mg/L	56%	59%	61%	58%	60%	59%	59%	59%	61%	63%	60%	59%
10–25 mg/L	24%	22%	23%	27%	25%	27%	26%	25%	24%	23%	24%	24%
25–50 mg/L	17%	16%	14%	13%	15%	12%	14%	14%	12%	12%	14%	14%
> 50 mg/L	3%	3%	2%	2%	1%	2%	2%	2%	2%	2%	2%	2%
Av. (mg/L)	12.93	12.92	11.97	12.03	11.45	11.39	11.78	12.20	12.34	11.53	11.87	12.04
95th (mg/L)	43.22	45.17	44.02	40.38	37.80	37.16	39.10	40.00	41.63	39.55	40.06	40.73
	Emilia-Romagna											
Nr. Sites	353	497	436	437	505	460	306	121	500	447	474	412
< 10 mg/L	58%	60%	57%	58%	64%	58%	58%	39%	65%	62%	64%	58%
10–25 mg/L	10%	11%	12%	12%	11%	13%	10%	23%	11%	12%	13%	13%
25–50 mg/L	20%	17%	18%	19%	16%	19%	21%	23%	17%	18%	16%	19%
> 50 mg/L	13%	12%	13%	11%	9%	10%	10%	15%	8%	8%	7%	10%
Av. (mg/L)	19.30	18.20	19.61	18.09	15.59	17.72	17.18	24.02	14.97	14.88	13.80	17.58
95th (mg/L)	75.72	74.59	73.20	70.90	62.80	67.05	59.63	62.50	58.50	53.50	53.26	64.69
	Total Area											
Nr. Sites	1530	1759	1639	1729	1759	1717	1497	1192	1689	1386	1407	1573
< 10 mg/L	47%	48%	48%	51%	52%	51%	50%	45%	54%	54%	54%	50%
10–25 mg/L	23%	22%	23%	24%	23%	23%	24%	28%	23%	23%	22%	24%
25–50 mg/L	23%	22%	22%	20%	20%	20%	21%	22%	19%	18%	19%	21%
> 50 mg/L	7%	8%	7%	5%	5%	5%	6%	5%	5%	5%	4%	6%
Av. (mg/L)	18.02	18.11	17.94	16.15	16.00	16.44	16.40	17.43	15.18	14.87	14.56	16.46
95th (mg/L)	56.55	58.04	55.17	50.48	50.04	51.17	51.50	51.17	48.89	49.48	48.71	51.93

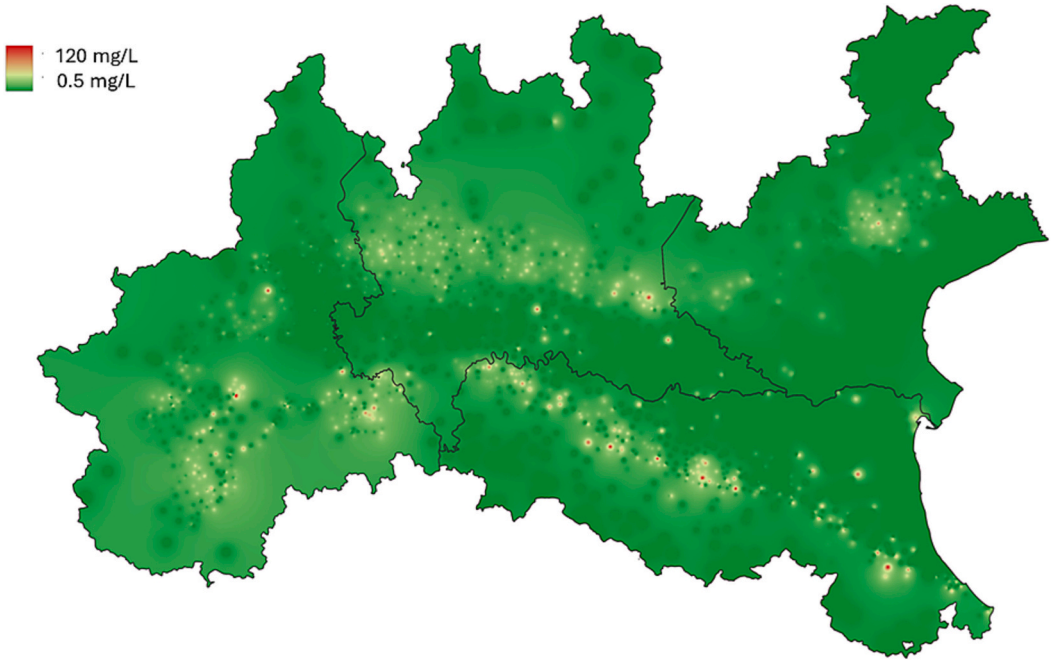


Fig. 2. – Map of the 2013–2023 average annual NO_3^- concentration.

that proposed by (Voccia et al., 2022). Table 3 summarizes the average results for this period, including site-specific mean nitrogen application rates (N-dose, kg N/ha), the estimated α fractions, average total nitrogen load (L_N), and the proportion of sites with $L_{non\text{-}agr}$ values exceeding 10 kg N/ha/yr. In addition, the table reports average annual precipitation (P) and recharge (R) for each site.

In (Voccia et al., 2022), a constant weight of 0.5 was assigned to soil texture and drainage in the calculation of the α coefficient, resulting in an estimated value of approximately 10%. In contrast, the present study incorporates geospatially distributed data for these parameters, yielding a higher average α value of approximately 13%. Furthermore, nitrogen application rates (N-dose) were estimated at the provincial rather than the regional scale, thereby improving the spatial resolution and overall accuracy of the assessment. These methodological enhancements increase both the precision and spatial sensitivity of the analysis, allowing for a more detailed evaluation of groundwater pollution patterns.

The results presented in Table 3 are consistent with the trends observed in NO_3^- concentrations, with Veneto consistently exhibiting the lowest 95th percentile $L_{non\text{-}agr}$ values and Emilia-Romagna the highest. Analysis of inter-regional averages reveals a strong positive Pearson correlation between precipitation, and both mean and 95th percentile $L_{non\text{-}agr}$ values, with correlation coefficients exceeding 0.8 across all regions. While previous studies have suggested that higher precipitation intensity enhances NO_3^- leaching and thus groundwater contamination (Hess et al., 2020; Jabloun et al., 2015; Martinez-Feria et al., 2019), the results of this study indicate a more complex relationship. For example, Veneto, with an average annual precipitation of 1126 mm, shows lower contamination levels (in terms of average L_N) compared to Emilia-Romagna, where mean annual precipitation is 725 mm. While this study successfully identifies regional drivers using a decadal dataset, it is important to note that the reliance on annual averages may overlook high-frequency seasonal dynamics. NO_3^- transport is sensitive to the precise timing of recharge and fertilization; therefore, while the current approach captures long-term structural patterns, future research focusing on intra-annual variability could further refine the understanding of transient leaching pulses in these aquifers.

These findings suggest that precipitation alone does not control groundwater contamination. Instead, groundwater quality is influenced by a combination of factors, including soil texture, drainage conditions, agricultural and industrial practices, nitrogen application rates, and the degree of urbanization. This interpretation is further supported by the relatively weak correlation between N-dose and average NO_3^- (or L_N) values, with Pearson correlation coefficients ranging from −0.67 to +0.47, indicating that agricultural inputs alone cannot fully explain the observed patterns.

$L_{non\text{-}agr}$ represents the nitrogen present in groundwater after

Table 3
– Average Annual Precipitation (AAP), Average Groundwater Recharge (AGR), Average N-DOSE (AND), Average Fraction α (AFα), Average N Leached (AL_N) and Average N non-agricultural leached ($AL_{non\text{-}agr}$) statistical analysis.

	2013	2014	2015	2016	2017	2018	2019	2020	2021	2022	2023	Average
	Piedmont											
AAP (mm/yr)	996	1309	771	933	516	1137	1075	788	713	464	725	857
AGR (mm/yr)	597	776	463	556	307	679	636	471	421	282	444	512
AND (kg/ha/yr)	99.05	98.71	90.55	92.69	89.80	82.94	74.16	89.60	98.21	53.83	71.47	85.55
AFα(-)	0.13	0.14	0.13	0.13	0.12	0.14	0.13	0.13	0.13	0.12	0.13	0.13
AL_N (kg/ha/yr)	49.59	58.46	42.90	39.95	33.79	48.82	44.98	46.77	33.92	31.02	38.63	42.62
$AL_{non\text{-}agr}$ (kg/ha/yr)	38.83	45.63	33.35	29.79	25.84	37.39	34.42	36.85	25.18	23.36	29.45	32.74
$\%L_{non\text{-}agr} > 10$ kg/ha/yr	66%	69%	64%	60%	60%	62%	59%	70%	58%	62%	64%	63%
95th $L_{non\text{-}agr}$ (kg/ha/yr)	122.13	134.23	106.24	97.21	75.31	123.24	112.41	102.18	81.92	71.99	86.38	101.20
	Lombardy											
AAP (mm/yr)	1134	1486	707	925	675	943	993	948	806	514	831	906
AGR (mm/yr)	658	856	405	533	388	551	579	546	463	303	482	524
AND (kg/ha/yr)	111.27	116.42	98.16	124.84	135.25	134.74	147.10	191.29	185.21	135.76	219.86	145.45
AFα(-)	0.14	0.15	0.13	0.14	0.13	0.14	0.14	0.14	0.13	0.12	0.13	0.13
AL_N (kg/ha/yr)	45.49	58.13	36.19	38.37	33.16	38.54	43.12	38.42	34.98	27.32	34.63	38.94
$AL_{non\text{-}agr}$ (kg/ha/yr)	36.17	46.43	28.66	30.01	25.92	30.20	34.38	29.93	27.10	21.28	26.96	30.64
$\%L_{non\text{-}agr} > 10$ kg/ha/yr	55%	59%	56%	53%	52%	52%	58%	54%	53%	49%	50%	54%
95th $L_{non\text{-}agr}$ (kg/ha/yr)	110.47	142.93	84.00	99.27	87.11	101.90	104.55	94.74	84.70	75.84	92.98	98.04
	Veneto											
AAP (mm/yr)	1309	1731	815	1170	916	1147	1376	1125	976	737	1089	1126
AGR (mm/yr)	747	984	465	669	519	652	783	637	554	419	619	641
AND (kg/ha/yr)	126.20	120.46	134.83	93.24	127.48	117.61	108.67	116.92	118.42	66.41	87.65	110.72
AFα(-)	0.14	0.15	0.13	0.14	0.13	0.14	0.14	0.14	0.13	0.13	0.14	0.14
AL_N (kg/ha/yr)	36.34	43.78	25.78	31.64	26.42	29.70	34.21	31.58	29.63	23.80	30.12	31.18
$AL_{non\text{-}agr}$ (kg/ha/yr)	26.53	31.78	18.34	22.39	18.36	20.61	24.02	22.47	21.20	16.61	21.28	22.15
$\%L_{non\text{-}agr} > 10$ kg/ha/yr	47%	48%	41%	46%	44%	46%	46%	46%	42%	40%	44%	44%
95th $L_{non\text{-}agr}$ (kg/ha/yr)	105.36	127.40	80.03	88.39	75.86	75.66	96.17	89.26	87.12	68.38	87.92	89.23
	Emilia-Romagna											
AAP (mm/yr)	897	1034	691	729	584	748	846	626	547	555	723	725
AGR (mm/yr)	584	655	454	482	371	488	551	403	345	366	461	469
AND (kg/ha/yr)	111.58	93.42	106.15	133.66	130.38	111.62	107.78	168.65	150.95	135.70	100.67	122.78
AFα(-)	0.13	0.13	0.13	0.13	0.12	0.13	0.13	0.13	0.12	0.12	0.13	0.13
AL_N (kg/ha/yr)	47.19	46.30	42.28	39.52	29.73	39.62	41.25	49.30	27.78	28.89	28.43	38.21
$AL_{non\text{-}agr}$ (kg/ha/yr)	40.37	38.87	36.01	33.33	24.39	33.08	34.28	40.52	22.54	23.38	22.53	31.75
$\%L_{non\text{-}agr} > 10$ kg/ha/yr	43%	40%	43%	43%	36%	43%	42%	61%	35%	38%	36%	42%
95th $L_{non\text{-}agr}$ (kg/ha/yr)	174.20	174.82	148.01	141.82	108.92	138.44	131.54	114.73	99.10	94.95	97.38	129.45
	Total Area											
AAP (mm/yr)	1065	1347	741	919	642	985	1065	914	733	561	824	890
AGR (mm/yr)	637	797	446	550	381	591	632	533	432	342	494	530
AND (kg/ha/yr)	109.95	105.62	104.15	111.66	119.42	109.61	106.48	143.40	140.48	105.08	123.73	116.32
AFα(-)	0.13	0.14	0.13	0.13	0.13	0.13	0.14	0.13	0.13	0.12	0.13	0.13
AL_N (kg/ha/yr)	45.61	52.64	38.12	38.05	31.25	40.53	41.68	40.17	31.67	27.89	32.46	38.19
$AL_{non\text{-}agr}$ (kg/ha/yr)	36.31	41.78	30.31	29.52	24.22	31.59	32.39	31.11	24.25	21.46	24.85	29.80
$\%L_{non\text{-}agr} > 10$ kg/ha/yr	55%	55%	53%	51%	48%	52%	53%	57%	47%	47%	47%	51%
95th $L_{non\text{-}agr}$ (kg/ha/yr)	125.44	147.45	107.11	104.99	88.27	110.74	113.84	101.05	87.99	82.41	92.93	105.66

subtracting the estimated agricultural contribution under a worst-case scenario. As such, it can be interpreted as an indicator of non-agricultural anthropogenic nitrogen inputs, primarily associated with urban and industrial sources. Lombardy and Piedmont exhibit the highest proportions of sites with $L_{non\text{-}agr}$ values exceeding 10 kg N/ha/yr, at 54% and 63%, respectively. In contrast, Emilia-Romagna and Veneto show lower proportions, at 42% and 44%, respectively, suggesting a comparatively smaller influence of non-agricultural activities on groundwater pollution in these regions.

3.2. Uncertainty and sensitivity analysis

Uncertainty analysis was conducted using Monte Carlo simulations (10,000 iterations) for each monitoring site, based on average input values for the 2013–2023 period. For each site, the median $L_{non\text{-}agr}$ value and its corresponding 95% confidence interval were computed from the simulation outputs and subsequently interpolated using QGIS. Fig. 3 illustrates the spatial distribution of the interpolated median $L_{non\text{-}agr}$ values. Elevated non-agricultural contributions are predominantly observed in areas surrounding urban centers, supporting and reinforcing findings from previous study in Emilia-Romagna and highlighting the

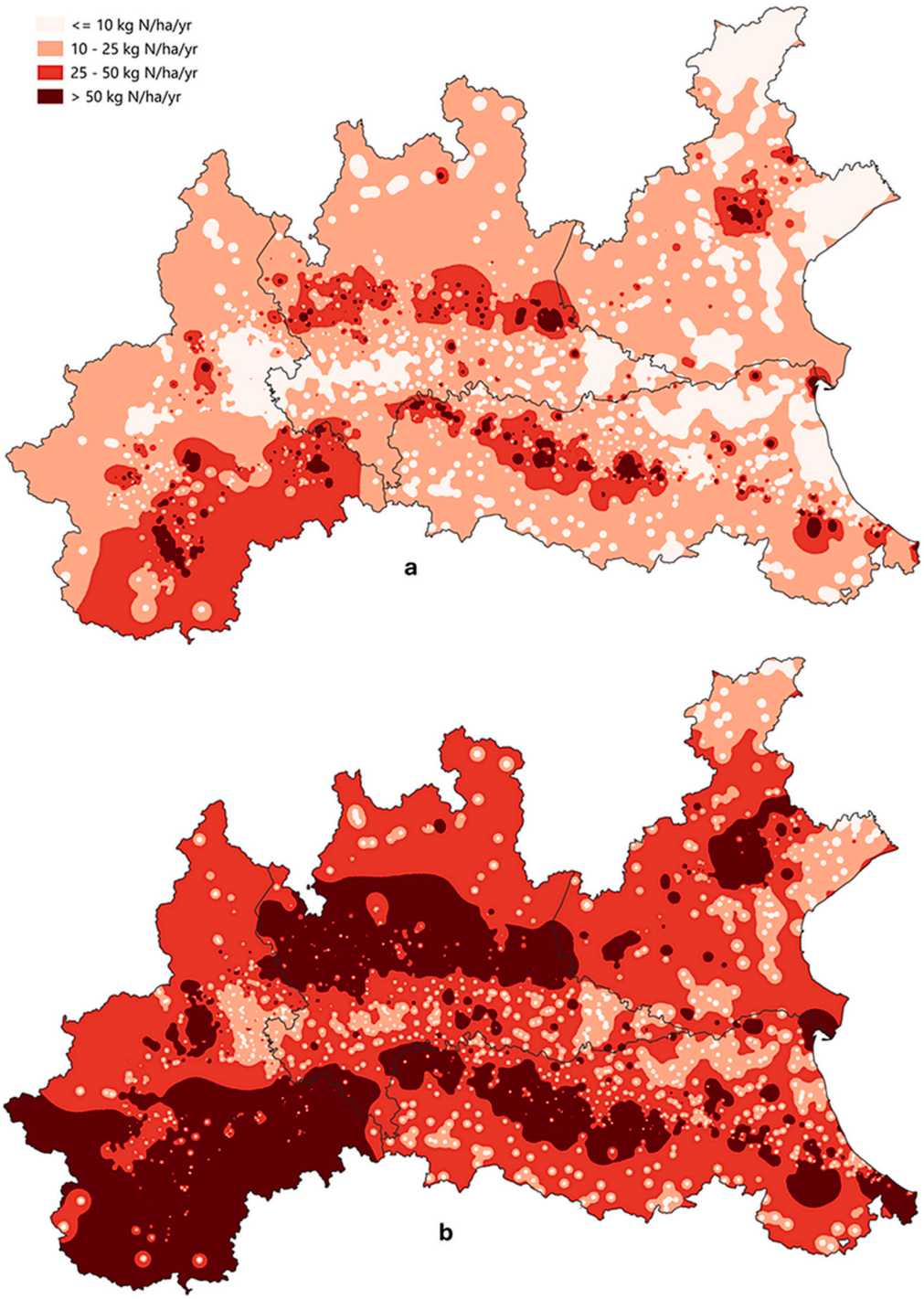


Fig. 3. – Map of the median $L_{non\text{-}agr}$ obtained in 10,000 Montecarlo simulations starting from average annual values from 2013 to 2023 (a) and correspondent 95% confidence interval (b).

relevance of additional anthropogenic sources beyond agriculture (Voccia et al., 2022). While $L_{non\text{-}agr}$ provides a spatially explicit indicator of nitrate enrichment beyond current agricultural activities, its interpretation requires caution. In this study, we interpret $L_{non\text{-}agr}$ as an integrated non-agricultural anthropogenic signal rather than a direct measurement of specific industrial or urban effluents. This distinction is necessary because, as a residual indicator, it may encompass legacy nitrogen and be modulated by local hydrogeological processes such as denitrification. The stability of the $L_{non\text{-}agr}$ signal under these varying parameters suggests it is a robust proxy for non-crop related pressures. Nevertheless, we acknowledge that without independent validation via isotopic tracers, the specific partitioning between legacy inputs and active urban-industrial discharge remains a subject for future investigation.

Overall, the analysis reveals a certain degree of uncertainty, reflecting the variability and limited knowledge of key input parameters, as described in the Materials and Methods section. Sensitivity analysis indicates that the largest contribution to output uncertainty arises from R_o (residual groundwater recharge), followed by natural background NO_3^- concentrations and measured NO_3^- levels in groundwater, as shown in Fig. 4. The spatial distribution of uncertainty further indicates that the highest variability is associated with areas characterized by elevated $L_{non\text{-}agr}$ values.

A major source of uncertainty is linked to the limited knowledge of groundwater system properties, particularly the effective volume of the aquifer and recharge dynamics. In many cases, these parameters are not directly measurable and must be approximated, introducing significant variability into model outputs (Healy and Scanlon, 2010; Scanlon et al., 2002). Similarly, the estimation of natural background NO_3^- concentrations (C_{nat}) remains challenging due to spatial heterogeneity and the influence of both geogenic and low-level anthropogenic sources. Reported background values can vary widely depending on hydrogeological conditions, often ranging from <1 mg/L up to 10 mg/L in some aquifers (Edmunds and Shand, 2008; Rivett et al., 2008). This wide range justifies the assumptions adopted in this study but also contributes substantially to overall uncertainty. These findings emphasize that uncertainty is inherently linked to both data limitations and conceptual model assumptions. Improving the characterization of aquifer properties and background NO_3^- levels would therefore be essential for reducing uncertainty and enhancing the reliability of groundwater pollution assessments.

3.3. Municipality statistical analysis

To better interpret the results, municipality-level indicators, population density, relative industrial area, and relative cultivated area, were considered proxies for anthropogenic pressures. These variables allow the spatial characterization of $L_{non\text{-}agr}$ values in relation to urbanization (population density, PD), industrialization (industrial land share, ID), and agricultural activity (cultivated land share, AD). To further investigate the contribution of different pollution sources and to support the assumptions outlined previously, a PCA was conducted. Monitoring sites were grouped according to $L_{non\text{-}agr}$ classes, while both natural variables like P, soil texture (T), and drainage (D) and anthropogenic variables (PD, ID, AD) were included in the analysis. The KMO values ranged from 0.55 to 0.63, indicating a moderate yet acceptable sampling adequacy for exploratory analysis. Bartlett's test of sphericity yielded $p < 0.05$ in all cases, confirming statistically significant correlations among variables and supporting the suitability of the PCA approach. The first two principal components explain a cumulative variance of 53.43% (PC1: 32.49%; PC2: 22.66%), while the inclusion of PC3 increases the total explained variance to 69.3%. Given these values, the PCA is utilized primarily as an exploratory tool to identify associations among variables rather than a standalone method for definitive source attribution. Table 4 summarizes the contributions and loadings of the variables for each principal component. Population density (PD) and industrial density (ID) show strong positive loadings on PC1, while agricultural density (AD) exhibits a negative loading. These results suggest a primary association between PC1 and anthropogenic variables related to urban and industrial development. In contrast, PC2 appears to be influenced by natural environmental factors, particularly precipitation and soil texture, whereas PC3 is largely associated with soil properties. This exploratory grouping highlights the different potential drivers of nitrate variability within the study area.

Fig. 5 illustrates the distribution of monitoring sites and regional

Table 4
– Variable contribution to PC1, PC2 and PC3.

	PC1	PC2	PC3
T	4%	12%	60%
D	0%	33%	4%
P	7%	27%	−23%
PD	35%	−7%	2%
AD	−27%	−5%	10%
ID	26%	−16%	0%

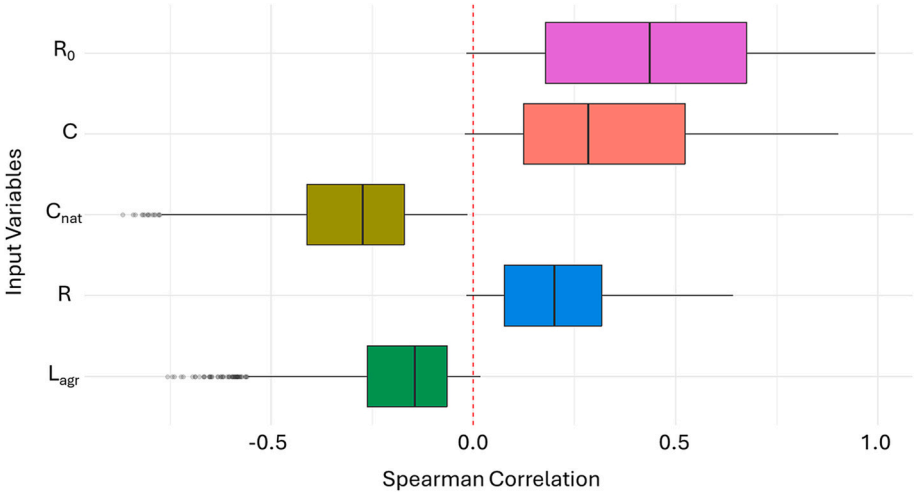


Fig. 4. – Boxplot of Spierman Correlations on the input variables. R_0: starting groundwater thickness; C: nitrates concentration; C_{nat}: background nitrates concentration; R: annual groundwater recharge; L_{agr}: agricultural load.

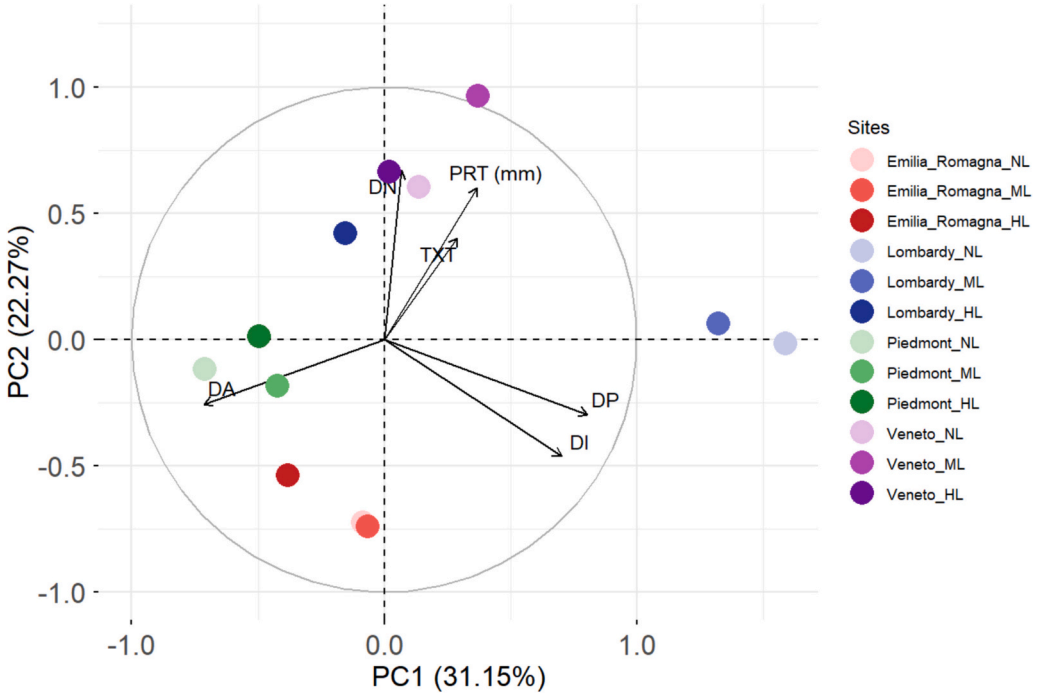


Fig. 5. - Principal Component Analysis performed on $L_{non\text{-}agr}$ values from 2013 to 2023 for all sites. Each group color represents a different region: red for Emilia-Romagna; blue for Lombardy; green for Piedmont; purple for Veneto. D: soil drainage; T: soil texture; P: annual precipitation; DA: Agricultural density; DP: Population density; DI: Industrial density; NL: $L_{non\text{-}agr} < 10$; ML: $L_{non\text{-}agr} = 10$–50 kg N/ha/yr; HL: $L_{non\text{-}agr} > 50$ kg N/ha/yr.

centroids within the PC1-PC2 space, overlaid with variable loading vectors. Sites are categorized into three classes (NL, ML, and HL) based on their $L_{non\text{-}agr}$ values. The centroids for Piedmont and Emilia-Romagna are positioned in close proximity to the agricultural density vector, consistent with the significant role of agricultural land use in these regions. Notably, the Emilia-Romagna_HL centroid shows the strongest alignment with the agricultural vector, suggesting a spatial association between intensive cultivation and elevated NO_3^- levels. Conversely, in Piedmont, the ML and NL centroids are closer to this vector, reflecting a comparatively different pattern of agricultural influence compared to the high-impact zones of Emilia-Romagna.

Conversely, the centroids for Veneto and Lombardy are positioned closer to vectors associated with natural factors, indicating a relatively lower association with the anthropogenic variables analyzed. To further

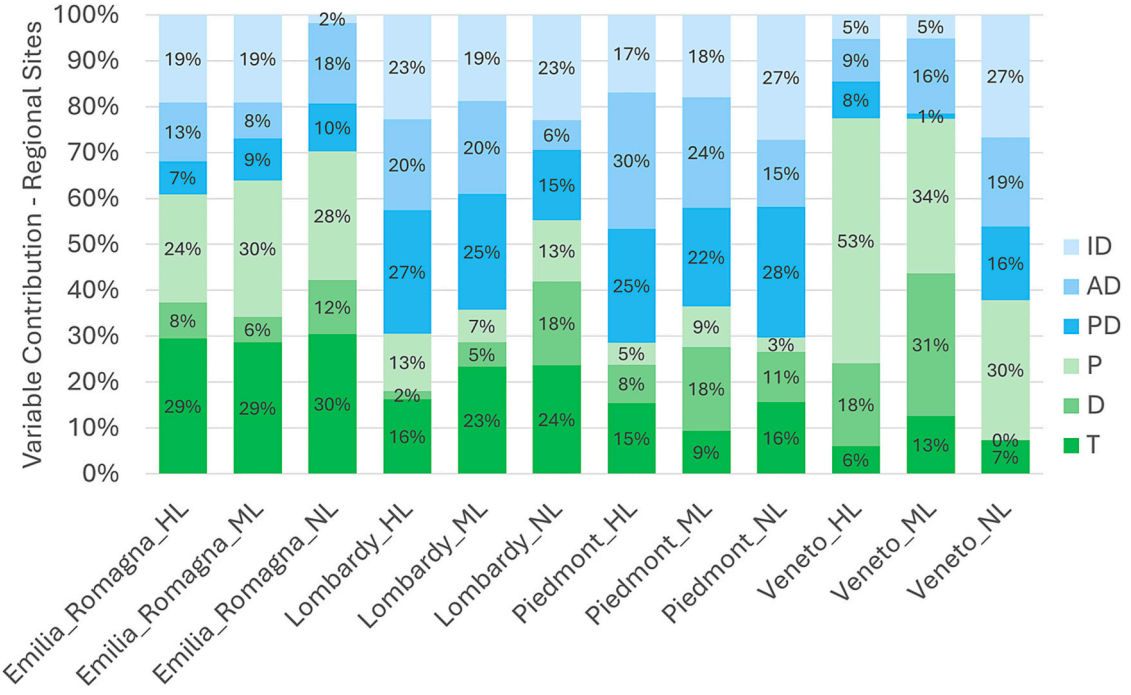


Fig. 6. – Variables contribution to regional centroids of Fig. 5. In shades of green, the natural contribution like precipitation (P), soil texture (T) and soil drainage (D); in shades of blue, the anthropogenic contribution like municipal industrial density (ID), municipal agricultural density (AD) and municipal population density (PD).

explore these relationships, the principal component scores of the centroids were decomposed into variable contributions (Fig. 6). Although Emilia-Romagna centroids are located in a region of the PCA space aligned with agricultural density, the decomposition of these contributions reveals that natural factors, particularly soil texture and precipitation, account for a substantial portion of the observed variance. Within this context, agriculture remains the primary anthropogenic variable associated with the NO_3^- patterns in this region. In Lombardy, the centroids are oriented towards the anthropogenic gradient (population and industrial density) in the PCA biplot. This pattern is further supported by the contribution analysis, where anthropogenic factors account for 64–70% in HL and ML sites, while natural factors explain the majority of the variance in NL sites. Although Piedmont centroids align the agricultural density vector, the decomposition of contributions reveals a more complex structure; anthropogenic contributions range from 64% to 72%, with agriculture identified as the primary anthropogenic variable associated with the HL and ML classes. Similarly, Veneto centroids are positioned closer to natural variables in the PCA space; the contribution analysis indicates that natural factors, especially precipitation, account fro a high proportion of overall variability, although agricultural and industrial activities contribute across all classes to varying extents.

These results highlight that agricultural practices are not the sole drivers of NO_3^- variability; instead, natural factors (particularly precipitation and soil properties) play a fundamental role in controlling NO_3^- transport and accumulation in groundwater. To enhance the spatial interpretation of these results, the first three principal components (PC1, PC2, and PC3) were interpolated geographically using average data from 2013 to 2023 (Fig. 7). Based on the variable loadings, the positive and negative scores of each component were interpreted to reflect different environmental and anthropogenic gradients. The spatial distribution of PC1 scores suggests that agricultural activity is a predominant association in Piedmont and Emilia-Romagna, whereas industrial and urban pressures appear more prominent in Lombardy. These observations are consistent with the findings from a 2019 study conducted in Turin-Cuneo area of the Piedmont region, based on NO_3^- different isotope monitoring, confirm the anthropogenic source of pollution is mainly due to both synthetic and organic nitrogen (Lasagna and De Luca, 2019). Veneto exhibits a more balanced contribution from diverse anthropogenic sources. Furthermore, the spatial patterns of PC2 and PC3 underscore the influence of natural factors, particularly soil texture and precipitation, in shaping groundwater contamination.

This analysis underscores the importance of evaluating the combined impact of both anthropogenic and natural sources to accurately assess the sources of groundwater NO_3^- pollution and implement effective mitigation strategies. Some studies attempted to evaluate the potential source of pollution of several contaminants in groundwater (including NO_3^-), adopting methods based on PCA. In a study conducted in China (Zhang et al., 2020), NO_3^- pollution was mainly attributed to agricultural practices, followed by urban and industrial pollution. It is important to note that in their approach, they assumed NO_3^- as a conventional pollutant from agricultural activities, while ammonia and sulfates as conventional pollutants from urban sewage leakage and industrial effluent, respectively. Based on those assumptions, they attributed each component obtained from the application of the method to a specific contamination source if the component had a strong positive loading on

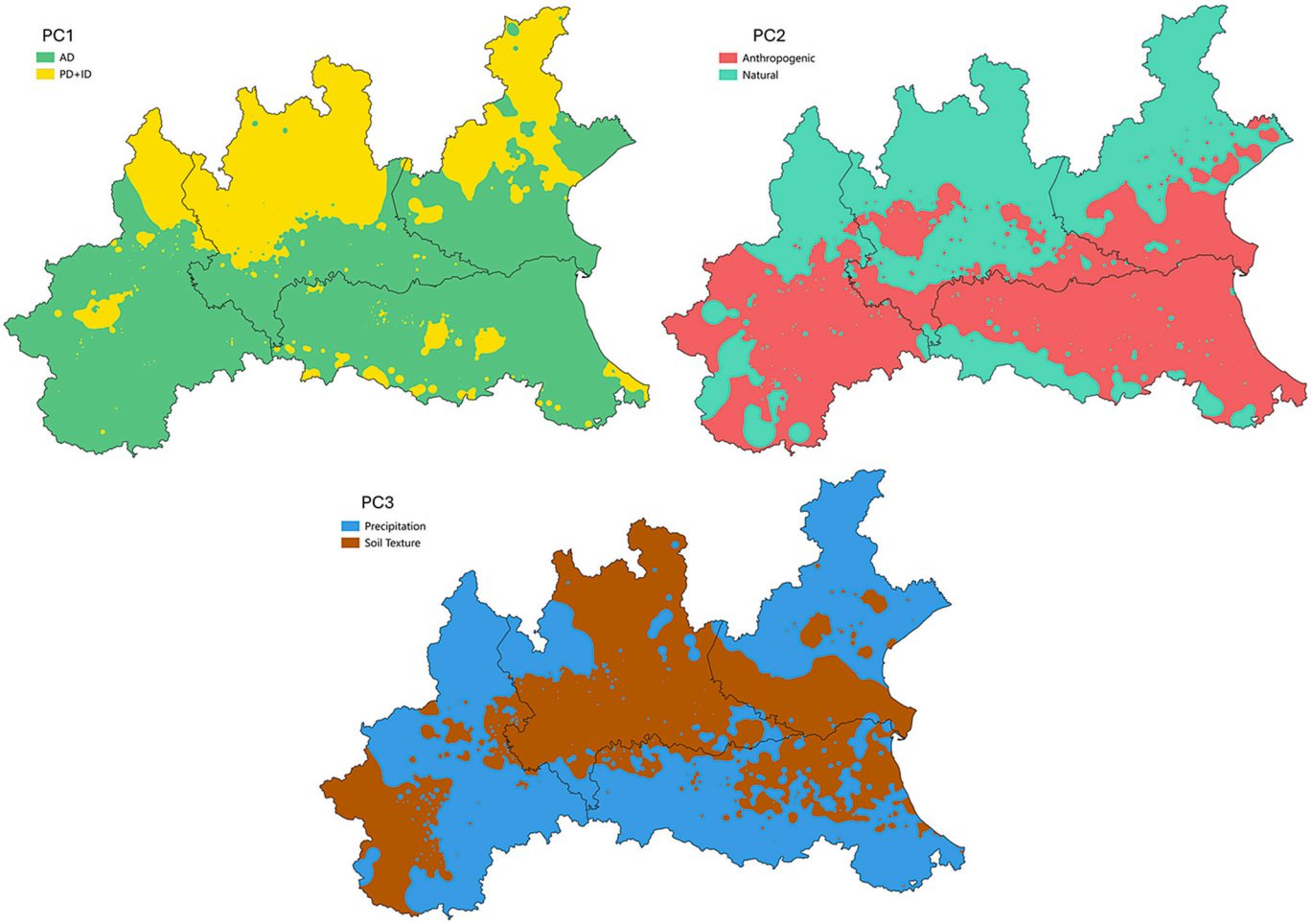


Fig. 7. - Principal Components PC1, PC2 and PC3 main positive and negative variables loading from average data interpolation from 2013 to 2023.

a characteristic pollutant. In a similar study conducted in China (Chen et al., 2023), results show that NO_3^- pollution was mainly attributed to industrial activities, followed by agricultural and urban pollution, further corroborating the concept that agriculture is not the only potential source of groundwater contamination. Component spatial distribution was also performed in another Chinese study based on the same assumption that NO_3^- pollution was mainly because of agricultural practices (Li et al., 2020), concluding fertilizer application is the main responsible for groundwater contamination.

3.4. Temporal PCA from 2013 to 2023

A temporal PCA was also performed by clustering sites according to the year and $L_{non\text{-}agr}$ class (HL, ML, NL). The results are shown in Fig. 8 and highlight that the relative importance of anthropogenic and natural drivers varies across years, indicating a dynamic system where climatic conditions and human activities interact in shaping groundwater NO_3^- contamination. For HL sites, anthropogenic factors generally account for the majority of the observed variance, except for 2014, when natural factors prevail. This shift in the statistical structure is consistent with the exceptionally high precipitation recorded in 2014, which likely enhanced natural leaching and dilution processes.

For ML sites, centroids are located in the positive side of PC1, suggesting a combined association of urban-industrial activities and natural drivers. In contrast, NL sites are predominantly located in the negative portion of PC1, where agricultural variable is more pronounced. This distribution further supports the premise that NO_3^- variability cannot be attributed solely to agriculture.

The temporal evolution of variable contributions (Fig. 9) indicates that natural factors account for the majority of the variance in most years, with precipitation often representing the primary driver. A notable exception occurred in 2020, when soil texture emerged as the dominant factor, accompanied by an increase in anthropogenic contributions. This anomaly cannot be explained solely by precipitation totals (which were lowest in 2022 (Xue et al., 2024)), but rather by the specific precipitation pattern in 2020, characterized by dry periods interspersed with intense rainfall events (https://www.arpae.it/it/notizie/il-clima-in-italia-nel-2020-e-trend-1961-2020). Under such conditions, soil properties play a crucial role in controlling infiltration and NO_3^- transport (Domnariu et al., 2020).

Anthropogenic contributions exhibit peaks in 2016 and 2020, particularly concerning agricultural variables, which represent the primary anthropogenic associated factor in 8 out of 11 years. These peaks align with period of increased fertilizer application, as reported by ISTAT. The 2016 peak may be associated with regulatory changes (e.g., Directive 2016/2284), while the 2020 peak may reflect increased

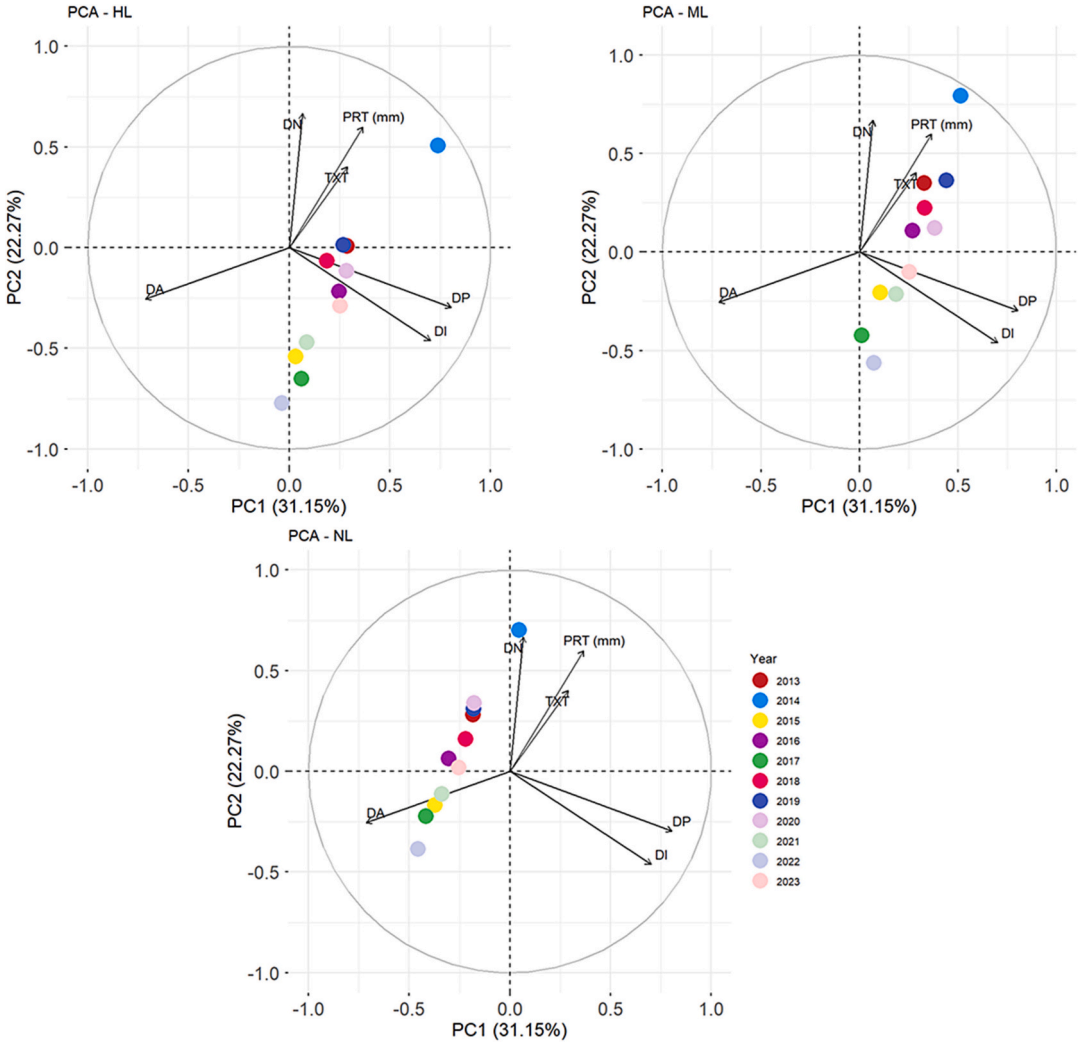


Fig. 8. - Principal Component Analysis performed on $L_{non\text{-}agr}$ values from 2013 to 2023 for all sites. Each color represents a year, divided in NL: $L_{non\text{-}agr} < 10$ kg N/ha/yr, ML: $L_{non\text{-}agr} = 10$–50 kg N/ha/yr and HL: $L_{non\text{-}agr} > 50$ kg N/ha/yr. D: soil drainage; T: soil texture; P: annual precipitation; DA: Agricultural density; DP: Population density; DI: Industrial density;

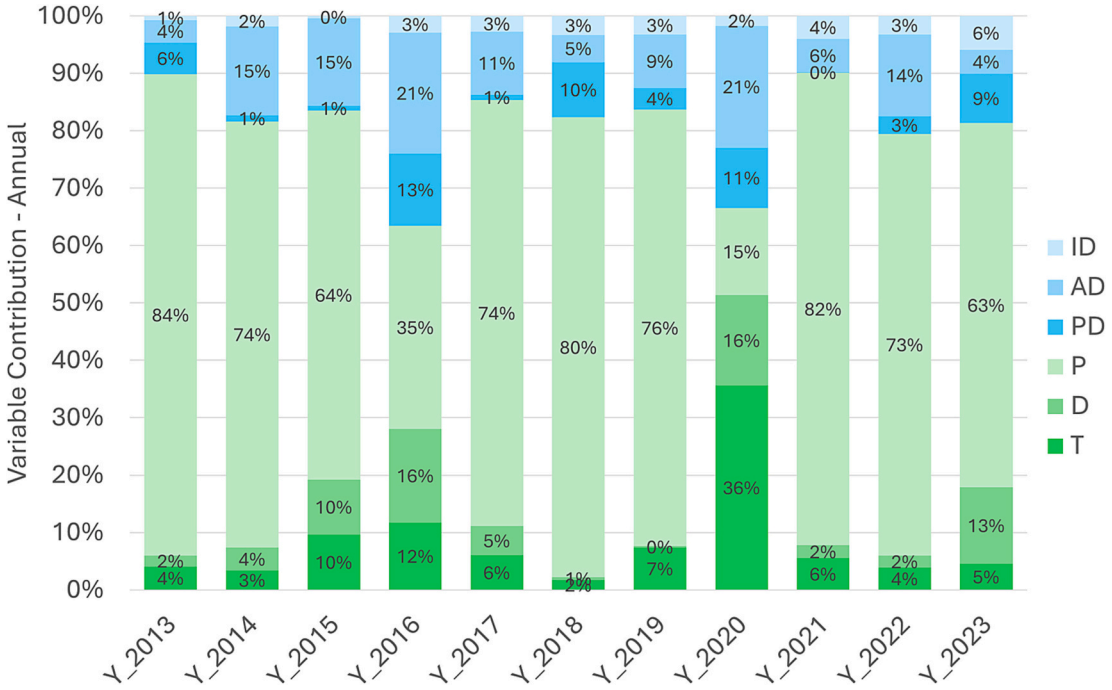


Fig. 9. – Variables contribution to annual centroids of Fig. 6. In shades of green, the natural contribution like precipitation (P), soil texture (T) and soil drainage (D); in shades of blue, the anthropogenic contribution like municipal industrial density (ID), municipal agricultural density (AD) and municipal population density (PD).

agricultural inputs combined with reduced industrial activity during the COVID-19 pandemic.

While the PCA provides valuable insights into the statistical associations among variables, the current approach to 'disentangling' nitrate sources does not rely on any single analytical tool. Instead, the identification of contamination drivers emerges from the convergence of multiple lines of evidence. The spatial interpolation maps (Figs. 2, 3 and 7) first identified regional hotspots; the land-use footprint analysis then quantified the specific anthropogenic pressures in those areas; and finally, the exploratory PCA confirmed that these spatial patterns align with distinct multivariate groupings. By integrating these independent methods, it is possible to move beyond simple correlations, providing a more robust and multifaceted assessment of how natural and anthropogenic factors collectively shape the hydrogeochemical landscape.

Despite the robustness of the integrated footprint-statistical approach presented here, certain limitations remain that warrant consideration in future research. This study focused on a parsimonious selection of variables, specifically NO_3^- and land-use footprints, to minimize interpretive ambiguity arising from the multi-source nature of common hydrochemical tracers such as chloride or sulfate. While this choice enhances the scalability and clarity of the regional 'disentangling' process, the absence of these secondary ions, as well as isotopic tracers, limits the ability to quantify specific transformation processes such as denitrification or to resolve overlapping signals in highly complex industrial-agricultural interfaces. Future investigations could build upon this regional framework by incorporating broader hydrochemical suites at a more localized scale to validate the source associations identified here. Integrating such tracers would provide a multi-dimensional validation of the footprint-based proxies, potentially refining the precision of source apportionment in areas with heterogeneous anthropogenic pressures.

4. Conclusions

This study provides a comprehensive assessment of groundwater NO_3^- pollution in northern Italy by integrating statistical, geospatial, and comprehensive uncertainty analyses. The results demonstrate that NO_3^- contamination arises from a complex interaction between anthropogenic signals and natural factors rather than a single dominant source. While agriculture remains a significant contributor, particularly in regions such as Emilia-Romagna and Piedmont, it does not fully explain the observed spatial and temporal variability in NO_3^- concentrations. In contrast, both natural and non-agricultural anthropogenic signals, associated mainly to urban and industrial footprints, play a more prominent role in regions such as Lombardy and Veneto. The use of $L_{non\text{-}agr}$ as an indicator of non-agricultural nitrogen inputs proved effective in disentangling these contributions. Crucially, the robustness of this indicator was validated through a sensitivity analysis that explicitly accounted for uncertainties in agricultural loads, natural background concentrations, and annual recharge variability. The PCA results highlight that the relative importance of these drivers varies across contexts, whereas natural factors, especially precipitation and soil texture, often regulate NO_3^- transport and accumulation. However, the interpretation of non-agricultural signals must remain cautious; while statistically robust, $L_{non\text{-}agr}$ acts as an integrated indicator that could potentially incorporates legacy nitrogen and is modulated by processes like denitrification. The temporal analysis confirms that this balance is not static but shifts over time in response to climatic variability and changes in land use practices. Uncertainty analysis also underscores the importance of key knowledge gaps, particularly regarding groundwater system properties and natural background NO_3^- concentrations, which remain critical limitations. Improving the characterization of these parameters would enhance the robustness of future assessments. Overall, this study emphasizes the need fonon-crop-relatedtegrated management strategies that consider both agricultural and non-agricultural sources, as well as the underlying hydrogeological context. Such an approach is essential for effectively mitigating groundwater NO_3^- pollution and ensuring sustainable water resource management. Overall, this study emphasizes the need for integrated management strategies that address both agricultural and non –crop-related anthropogenic pressures. By proposing a parsimonious and scalable framework that avoids the interpretive ambiguities of multi-source chemical tracers, this work provides a transferable tool for regional-scale water resource management in complex hydrological basins.

CRediT authorship contribution statement

Diego Voccia: Writing – review & editing, Writing – original draft, Visualization, Software, Resources, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **Giacomo Mortella:** Writing – review & editing, Data curation. **Marco Trevisan:** Writing – review & editing, Conceptualization. **Bartolo Forestieri:** Writing – review & editing, Data curation. **Lucrezia Lamastra:** Writing – original draft, Supervision, Methodology, Conceptualization.

Declaration of competing interest

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

Data availability

The data used in this work is all referenced in the manuscript and is publicly available.

References

Arpa Emilia Romagna, 2026. Agenzia Regionale per la Protezione Ambientale - Emilia Romagna [WWW Document]. URL. https://www.arpae.it/it. (Accessed 13 March 2026).

Blarasin, M., Cabrera, A., Matiatos, I., Becher Quinodóz, F., Giuliano Albo, J., Lutri, V., Matteoda, E., Panarello, H., 2020. Comparative evaluation of urban versus agricultural nitrate sources and sinks in an unconfined aquifer by isotopic and multivariate analyses. Sci. Total Environ. 741, 140374. https://doi.org/10.1016/j.scitotenv.2020.140374.

Buškulić, P., Parlov, J., Kovač, Z., Nakić, Z., 2023. Estimation of Nitrate Background Value in Groundwater under the Long-Term Human Impact. Hydrology 10, 63. https://doi.org/10.3390/hydrology10030063.

Chamandoost, S., Fateh Moradi, M., Hosseini, M.-J., 2016. A Review of Nitrate and Nitrite Toxicity in Foods. Journal of Human, Environment, and Health Promotion 1, 80–86. https://doi.org/10.29252/jhehp.1.2.80.

Chen, Z., Zhou, Q., Lv, J., Jiang, Y., Yang, Hai, Yang, Hui, Mei, S., Jia, Z., Zhang, H., Jin, Y., Liu, L., Shen, R., 2023. Assessment of groundwater quality using APCS-MLR model: a case study in the pilot promoter region of Yangtze River Delta Integration Demonstration Zone, China. Water (Basel). 15, 225. doi:https://doi.org/10.3390/w15020225.

De Luca, D.A., Lasagna, M., Debernardi, L., 2020. Hydrogeology of the western Po plain (Piedmont, NW Italy). J. Maps 16, 265–273. https://doi.org/10.1080/17445647.2020.1738280.

Domnariu, H., Paltineanu, C., Marica, D., Lacatsusu, A.-R., Rizea, N., Lazar, R., Popa, G. A., Vrinceanu, A., Balaceanu, C., 2020. Influence of Soil-Texture on nitrate leaching from small-scale lysimeters toward groundwater in various environments. Carpathian Journal of Earth and Environmental Sciences 15, 301–310. https://doi.org/10.26471/cjees/2020/015/130.

Ducci, L., Pinardi, R., Francesco, F. Di, Meo, C., Rizzo, P., Rezaei Kalvani, S., Segadelli, S., De Nardo, M.T., Celico, F., 2026. The influence of hydrogeological and anthropogenic factors on PFAS distribution in deep multilayer alluvial aquifer: the case study of Parma Plain, Northern Italy. Water (Basel). 18, 117. doi:https://doi.org/10.3390/w18010117.

Edmunds, W.M., Shand, P., 2008. Natural groundwater quality. Blackwell Pub.

European Environment Agency, 2020. CORINE Land Cover 2018. WWW Document. https://doi.org/10.2909/960998c1-1870-4e82-8051-6485205ebbac.

Eurostat, 2026. European Institute of Statistics. WWW Document. URL. https://ec.europa.eu/eurostat. (Accessed 13 March 2026).

Franke, N.A., Boyacioglu, H., Hoekstra, A.Y., 2013. Grey water footprint accounting: Tier 1 supporting guidelines, Value of Water Research Report Series vol. No. 65. Delft, the Netherlands.

Gioia, E., Casareale, C., Colocci, A., Zecchini, F., Marincioni, F., 2021. Citizens' perception of geohazards in Veneto Region (NE Italy) in the context of climate change. Geosciences (Basel). 11, 424. https://doi.org/10.3390/geosciences11100424.

Haji Gholizadeh, M., Melesse, A.M., Reddi, L., 2016. Water quality assessment and apportionment of pollution sources using APCS-MLR and PMF receptor modeling techniques in three major rivers of South Florida. Sci. Total Environ. 566–567, 1552–1567. https://doi.org/10.1016/j.scitotenv.2016.06.046.

Healy, R.W., Scanlon, B.R., 2010. Estimating Groundwater Recharge. Cambridge University Press. https://doi.org/10.1017/CBO9780511780745.

Hess, L.J.T., Hinckley, E.-L.S., Robertson, G.P., Matson, P.A., 2020. Rainfall intensification increases nitrate leaching from tilled but not no-till cropping systems in the U.S. Midwest. Agric. Ecosyst. Environ. 290, 106747. https://doi.org/10.1016/j.agee.2019.106747.

Hina, N.S., 2024. Global meta-analysis of nitrate leaching vulnerability in synthetic and organic fertilizers over the past four decades. Water (Basel). 16, 457. doi:https://doi.org/10.3390/w16030457.

Hoekstra, A.Y., Chapagain, A.K., Aldaya, M.M., Mekonnen, M.M., 2012. The Water Footprint Assessment Manual. Routledge. https://doi.org/10.4324/9781849775526.

Isprambiente, 2026. Isprambiente Database [WWW Document]. URL https://www.isprambiente.gov.it/ (accessed 3.13.26).

Istat, 2026. Italian Institute of Statistics [WWW Document]. URL. www.istat.it. (Accessed 13 March 2026).

Jabloun, M., Schelde, K., Tao, F., Olesen, J.E., 2015. Effect of temperature and precipitation on nitrate leaching from organic cereal cropping systems in Denmark. Eur. J. Agron. 62, 55–64. https://doi.org/10.1016/j.eja.2014.09.007.

Lasagna, M., De Luca, D.A., 2019. Evaluation of sources and fate of nitrates in the western Po plain groundwater (Italy) using nitrogen and boron isotopes. Environ. Sci. Pollut. Res. 26, 2089–2104. https://doi.org/10.1007/s11356-017-0792-6.

Leip, A., Achermann, B., Billen, G., Bleeker, A., Bouwman, A.F., de Vries, W., Dragosits, U., Döring, U., Fernall, D., Geupel, M., Herolstab, J., Johnes, P., Le Gall, A. C., Monni, S., Nevečeřal, R., Orlandini, L., Prud'homme, M., Reuter, H.I., Simpson, D., Seufert, G., Spranger, T., Sutton, M.A., van Aardenne, J., Voß, M., Winiwarter, W., 2011. Integrating nitrogen fluxes at the European scale. In: The European Nitrogen Assessment. Cambridge University Press, pp. 345–376. https://doi.org/10.1017/CBO9780511976988.019.

Lerner, D.N., 2003. ESTIMATING URBAN LOADS OF NITROGEN TO GROUNDWATER. Water Environ. J. 17, 239–244. https://doi.org/10.1111/j.1747-6593.2003.tb00475.x.

Li, Q., Zhang, H., Guo, S., Fu, K., Liao, L., Xu, Y., Cheng, S., 2020. Groundwater pollution source apportionment using principal component analysis in a multiple land-use area in southwestern China. Environ. Sci. Pollut. Res. 27, 9000–9011. https://doi.org/10.1007/s11356-019-06126-6.

Li, X., Zhang, Y., Xu, L., Jiang, J., Zhang, C., Wang, G., Huan, H., Tian, D., Guo, J., 2025. Groundwater pollution source identification based on a coupled PCA–PMF–Mantel framework: a case study of the Qujiang River Basin. Water (Basel). 17, 2881. doi:https://doi.org/10.3390/w17192881.

Liu, Y., Luo, J., Wu, Y., Zhang, Z., Zheng, X., Zheng, T., 2025. Impact of burial conditions on NO3- -N source apportionment in groundwater: Insights from PCA-APCS-MLR and MixSIAR methods. https://doi.org/10.5194/egusphere-2025-5482.

Lombardia, A.R.P.A., 2026. Agenzia Regionale per la Protezione Ambientale - Lombardia [WWW Document]. URL. https://www.arpalombardia.it/. (Accessed 13 March 2026).

Maranan, C., Cruz, G., Santos, F., 2025. Quantifying blue, green, and gray water footprints in a mixed land use urban catchment for sustainable urban water management. Frontiers in Water 7. https://doi.org/10.3389/frwa.2025.1655691.

Martinez-Feria, R., Nichols, V., Basso, B., Archontoulis, S., 2019. Can multi-strategy management stabilize nitrate leaching under increasing rainfall? Environ. Res. Lett. 14, 124079. https://doi.org/10.1088/1748-9326/ab5ca8.

Martinsen, G., Bessiere, H., Caballero, Y., Koch, J., Collados-Lara, A.J., Mansour, M., Sallasmaa, O., Pulido-Velazquez, D., Williams, N.H., Zaadnoordijk, W.J., Stisen, S., 2022. Developing a pan-European high-resolution groundwater recharge map – Combining satellite data and national survey data using machine learning. Sci. Total Environ. 822, 153464. https://doi.org/10.1016/j.scitotenv.2022.153464.

Mattsson, F., Neuberger, A., Heikkinen, L., Gramlich, Y., Paglione, M., Rinaldi, M., Decesari, S., Zieger, P., Riipinen, I., Mohr, C., 2025. Enrichment of organic nitrogen in fog residuals observed in the Italian Po Valley. Atmos. Chem. Phys. 25, 7973–7989. https://doi.org/10.5194/acp-25-7973-2025.

Monteleone, B., Borzí, I., 2024. Drought in the Po Valley: Identification, Impacts and Strategies to Manage the Events. Water (Basel). 16, 1187. https://doi.org/10.3390/w16081187.

Piemonte, A.R.P.A., 2026. Agenzia Regionale per la Protezione Ambientale - Piemonte [WWW Document]. URL. https://www.arpa.piemonte.it/. (Accessed 13 March 2026).

Rahman, A., Mondal, N.C., Tiwari, K.K., 2021. Anthropogenic nitrate in groundwater and its health risks in the view of background concentration in a semi arid area of Rajasthan, India. Sci. Rep. 11, 9279. https://doi.org/10.1038/s41598-021-88600-1.

Rivett, M.O., Buss, S.R., Morgan, P., Smith, J.W.N., Bemment, C.D., 2008. Nitrate attenuation in groundwater: A review of biogeochemical controlling processes. Water Res. 42, 4215–4232. https://doi.org/10.1016/j.watres.2008.07.020.

Rivieccio, R., 2019. SOIL DATABASE OF ITALY AT 1 : 250,000 SCALE AND THEMATIC APPLICATIONS FOR ASSESSING ENVIRONMENTAL SERVICES. Università degli Studi della Tuscia, Viterbo.

Ruan, D., Bian, J., Wang, Y., Wu, J., Gu, Z., 2024. Identification of groundwater pollution sources and health risk assessment in the Songnen Plain based on PCA-APCS-MLR and trapezoidal fuzzy number-Monte Carlo stochastic simulation model. J. Hydrol. (Amst). 632, 130897. https://doi.org/10.1016/j.jhydrol.2024.130897.

Rütting, T., Aronsson, H., Delin, S., 2018. Efficient use of nitrogen in agriculture. Nutr. Cycl. Agroecosyst. 110, 1–5. https://doi.org/10.1007/s10705-017-9900-8.

Sacchi, E., Acutis, M., Bartoli, M., Brenna, S., Delconte, C.A., Laini, A., Pennisi, M., 2013. Origin and fate of nitrates in groundwater from the central Po plain: Insights from isotopic investigations. Appl. Geochem. 34, 164–180. https://doi.org/10.1016/j.apgeochem.2013.03.008.

Scanlon, B.R., Healy, R.W., Cook, P.G., 2002. Choosing appropriate techniques for quantifying groundwater recharge. Hydrogeol. J. 10, 18–39. https://doi.org/10.1007/s10040-001-0176-2.

Severini, E., Bartoli, M., Pinardi, M., Celico, F., 2020. Reactive silica traces manure spreading in alluvial aquifers affected by nitrate contamination: a case study in a

High Plain of Northern Italy. Water (Basel). 12, 2511. doi:https://doi.org/10.3390/w12092511.

Šrajbek, M., Kranjčević, L., Kovač, I., Biondić, R., 2022. Groundwater nitrate pollution sources assessment for contaminated wellfield. Water (Basel). 14, 255. doi:https://doi.org/10.3390/w14020255.

Ungaro, F., Tarocco, P., Calzolari, C., 2025. Leveraging soil geography for land use planning: assessing and mapping soil ecosystem services indicators in Emilia-Romagna, NE Italy. Geographies 5, 39. https://doi.org/10.3390/geographies5030039.

Veneto, Arpa, 2026. Agenzia Regionale per la Protezione Ambientale - Veneto [WWW Document]. URL. https://www.arpa.veneto.it/. (Accessed 13 March 2026).

Voccia, D., Mortella, G., Ferrari, F., Fontanella, M.C., Trevisan, M., Lamastra, L., 2022. The Anthropic Pressure on the Grey Water Footprint: The Case of the Vulnerable Areas of the Emilia-Romagna Region in Italy. Sustainability 14, 16353. https://doi.org/10.3390/su142416353.

Wakida, F.T., Lerner, D.N., 2005. Non-agricultural sources of groundwater nitrate: a review and case study. Water Res. 39, 3–16. https://doi.org/10.1016/j.watres.2004.07.026.

Xue, C., Ghirardelli, A., Chen, J., Tarolli, P., 2024. Investigating agricultural drought in Northern Italy through explainable Machine Learning: Insights from the 2022 drought. Comput. Electron. Agric. 227, 109572. https://doi.org/10.1016/j.compag.2024.109572.

Zhang, H., Cheng, S., Li, H., Fu, K., Xu, Y., 2020. Groundwater pollution source identification and apportionment using PMF and PCA-APCA-MLR receptor models in a typical mixed land-use area in Southwestern China. Sci. Total Environ. 741, 140383. https://doi.org/10.1016/j.scitotenv.2020.140383.

Zhang, W., Ruan, X., Bai, Y., Yin, L., 2018. The characteristics and performance of sustainable-releasing compound carbon source material applied on groundwater nitrate in-situ remediation. Chemosphere 205, 635–642. https://doi.org/10.1016/j.chemosphere.2018.04.133.