Multivariate and long-term~~groundwater~~ time series analysis to assess the effect of nitrogen management policy on groundwater quality in Wallonia, BE

Elise Verstraeten[1], Alice Alonso[1], Louise Collier[2], Marnik Vanclooster[1]

[1]Earth and Life Institute - Environnemental Sciences, Université catholique de Louvain, Louvain-la-Neuve, 1348, Belgium
[2]Société Wallonne des Eaux, Verviers, 4800, Belgium

Correspondence to: Alice Alonso (alice.alonso@uclouvain.be)

Abstract. Groundwater nitrate contamination remains a significant environmental and health concern, as the effectiveness of nitrogen management programs has shown mixed results. This study leverages nearly two decades of groundwater monitoring data to assess and explain nitrate contamination trends in vulnerable zones in Wallonia, Belgium, following the implementation of the regional sustainable nitrogen management program.

Using nitrate concentration time series (2002–2020) from 36 monitoring points across four groundwater bodies, we calculated six nitrate pollution indicators that capture temporal dynamics. Spatially explicit variables describing potential contamination drivers were compiled, and their relationships with the nitrate pollution indicators were assessed using correlations and multivariate regressions.

A~~Results showed a decrease in the rate of nitrate concentration changes, though~~ lthough mean groundwater nitrate concentrations remained stable due to nitrogen legacy effects, there was an overall improvement as nitrate concentrations decreased over the study period. ~~mean concentrations remained stable, reflecting the influence of legacy nitrogen. Spatial patterns revealed~~Decreases~~ reductions~~ were observed in the Brusselian sands, where concentrations were initially higher, while~~and~~ increases were found in the Geer basin chalks, ~~which are~~ typically less contaminated. These diverging trends can be explained by differences in aquifer characteristics and nitrate transfer time lags. A~~This divergence likely reflects differences in aquifer characteristics and time lags in nitrate transfer, with shallow groundwater showing greater improvements than deeper zones. Agricultural land cover consistently exerted a significant influence on nitrate levels, underscoring the enduring impact~~gricultural land cover continues to have a negative impact on nitrate contamination~~of land use as a contamination driver~~, even after 20 years of PGDA implementation. The limited predictive power of the regression models highlights the multifaceted nature of groundwater nitrate contamination and the challenges in representing controlling factors, mainly due to lack of data.

Overall, this study emphasizes the need for sustained and adaptive nitrogen management policies, especially in vulnerable aquifers and cropland-dominated regions, alongside long-term monitoring to address time lags and nitrogen legacy effects. It also underscores the importance of detailed datasets to support advanced non-linear machine learning approaches that can capture the complex interactions involved.~~Tlimited predictive power of the multifaceted nature This study underscores the~~

~~importance of sustained and adaptive nitrogen management practices, particularly in vulnerable aquifers and cropland-dominated regions. It also suggests the need to sustain and intensify measures over extended periods to address time lags and nitrogen legacy effects, and the development of a more comprehensive and denser monitoring network to improve data resolution and track contamination trends effectively.~~

1 Introduction

For several decades, concerns have been~~being~~ raised about elevated nitrate concentrations in groundwaters. These high nitrate levels are harmful to the environment~~,~~ as they contribute to the biodiversity loss and eutrophication of the aquatic ecosystems (de Vries et al., 2024; Grizzetti et al., 2011; Hornung, 1999), and can pose serious health risks when the groundwater is used for drinking water (Bouchard et al., 1992; Comly, 1945; de Vries et al., 2024; Schullehner et al., 2018; M. Ward et al., 2018). Despite extensive research, u~~U~~nderstanding the factors underlying nitrate concentrations and trends is complicated by the diffuse, long-term and multi-causal nature of the contamination (Shukla & Saxena, 2018).

The primary anthropogenic sources of nitrates in groundwater are agricultural activities (Spalding & Exner, 1993; Wick et al., 2012). However, non-agricultural sources such as inadequate treatment and disposal of human waste and wastewater, landfill ~~and waste tip~~, and industrial waste can also be significant contributor to high concentrations in aquifers (Mattern et al., 2009; Vanclooster et al., 2020; Wakida & Lerner, 2005). While human activities ~~are the drivers of~~drive increased nitrogen input to the soils, the leaching of nitrates to~~levels in~~ groundwater ~~, the~~is ~~modulated~~governed by key natural processes, referred to as the nitrogen cycle (B. B. Ward, 2013). Specifically, ~~se increases are modulated by the natural properties such as climate, soil composition, geological formation, and depth of the groundwater tables~~n~~N~~itrate formation is ~~primarily~~ driven by ~~the~~ nitrification ~~process~~, a process in which ammonium, introduced through ~~human activities~~fertilizer and mineralization of organic matter~~waste~~, is oxidized to nitrate. In contrast, ~~N~~nitrate ~~degradation ,~~concentration levels decrease ~~on the other hand, occurs~~ through ~~its uptake by~~ plants and microbial uptake~~oorgani sms,~~ as well as through denitrification, ~~the~~ which transform~~sation~~ ~~of~~ nitrate into nitrogen gas~~, which takes place~~ under anaerobic conditions. These processes are ~~further~~ influenced by environmental factors ~~depend on natural properties~~ such as climate, soil composition, geological formation, and the depth of ~~the~~ groundwater tables.

To address the environmental and health risks associated with nitrate pollution, the European Union's Nitrate Directive (91/676/EEC) has played a central role in promoting best management practices (BMPs) that mitigate agricultural nitrate contamination. Established in 1991, the directive aims to protect water bodies by requiring member states to identify Nitrate Vulnerable Zones (NVZs) and implement targeted measures to reduce nitrate leaching, aiming to maintain surface and groundwater nitrate concentrations below 50 mg/L.

~~Recognizing the environmental and health hazards related to nitrate pollution, many regions have introduced best management practices (BMPs) to mitigate nitrate pollution. A notable example is the European Union's Nitrate Directive (91/676/EEC), established in 1991 to protect water bodies from agricultural nitrate pollution. Under this directive, member states must identify~~

~~Nitrate Vulnerable Zones (NVZs) and implement targeted measures to reduce nitrate leaching, with the aim of maintaining groundwater nitrate concentration levels below 50 mg/l.~~

Studies evaluating the effectiveness of such policies have shown mixed results. While many report encouraging signs of reduced nitrate pollution due to improved nitrogen management practices, others highlight persistent~~ing~~ challenges, with~~ a~~ continued increases or ~~minimal~~ limited improvement in concentrations in many locations (Ferguson, 2015; Hansen et al., 2012, 2017; Van Grinsven et al., 2012, 2016).

In Wallonia (Belgium), the European Directive was transposed into a program for sustainable nitrogen management (“Programme de Gestion Durable de l'Azote en Agriculture”, PGDA) in ~~, that was implemented at the end of~~ 2002. This program includes region-wide measures ~~that apply to the entire region of Wallonia, along~~and additional ~~measures only applicable~~requirements in the NVZs~~ in the Nitrate Vulnerable Zones~~, including restrictions on manure spreading, mandatory soil cover, and groundwater monitoring~~specific spreading periods and conditions, obligations regarding soil cover, and monitoring requirements~~ (Picron et al., 2017). In compliance with the Water Framework Directive (WFD, 2000/60/EC) and the Groundwater Directive (GWD, 2006/118/EC), the Walloon regional authority ~~assesses~~monitors groundwater nitrate concentrations, ~~and trends~~ to identify vulnerable areas and to evaluate the effectiveness of the PGDA (SPW - DEE - Direction des Eaux souterraines, 2024). ~~Since their most recent extension in 2013, the Nitrate Vulnerable Zones now cover 69% of the utilized agricultural land in Wallonia.~~ The observed effectiveness ~~of the program~~ so far ~~remains mitigated~~is debated without clear overall signs of improvement (Batlle Aguilar et al., 2007; SPW - DEE - Direction des Eaux souterraines, 2024). Many control sites keep exceeding the European guide level of 25 mg/l and most of the groundwater bodies in the Nitrate Vulnerable Zones (NVZ), partially or locally, have high levels and several exceedances of the standard of 50 mg/l~~, although~~. Despite this, ~~the~~ average nitrate ~~concentration values~~ levels in some aquifers have been shown to stabilize or~~and~~ even~~, in some aquifers, to~~ decrease in recent years.~~e.~~

To better understand the impacts of the PGDA on nitrogen dynamics, Sohier & Degré (2010) modeled soil nitrogen surpluses using a modified version of the process-oriented EPIC model (Williams et al., 1984). Their work demonstrated the positive effects of certain agricultural practices while~~but~~ also highlighting~~ed~~ significant variability driven by weather conditions, particularly precipitation. The regional authority uses this model to assess the agricultural soil nitrogen balance (SPW, 2022). However, ~~–. By using modelled data, tT~~the results are dependent on the hypotheses and assumptions about the physical processes underlying the model. When these processes are not well defined, which is challenging due to their complexity, the accuracy of the predictions is limited. ~~Although E~~efforts have been made to measure ~~in-situ~~ soil nitrate concentrations at the end of the agricultural crop season, known as potentially leachable nitrogen (APL). The in-situ APL measurements are correlated with agricultural practices such as fertilization and crop management, as well as with nitrate concentrations in soil water (Vandenberghe, 2016). However,~~–~~ these ~~such~~ measurements remain limited in both space and time, which constrains comprehensive assessments of nitrate contamination trends. Moreover, soil and soil water nitrate concentrations can differ from groundwater nitrate concentrations, as the latter are affected by nitrate transfer lags through soil matrices (Hansen et al., 2012; Mattern & Vanclooster, 2010; Visser et al., 2007) and the slow release of accumulated nitrogen (Ascott et al., 2017;

Kyte et al., 2023; Liu et al., 2024) — a phenomenon referred to as the “nitrogen legacy effect” (Basu et al., 2022; Van Meter et al., 2016).

~~. and the slow release of N that have accumulated over time in the landscape elements including soils and vadose zone , what certain author call “the nitrogen legacy effect” W~~n the other hand, ~~e~~hile ~~fforts have been made~~pursued to m~~measured soil n and analyse nitrate concentrations in agricultural soils at the end of the agricultural cropping season, – , known as ”potentially leachable nitrogen“ – *APL*. This indicator was shown to be correlated to agricultural practices such as fertilization and crop management, and to the nitrate concentrations of the water at the base of the root zone – control the processe, such as the time lag of nitrate transfer through the soil matrix Hence~~Given these complexities, ~~data-driven approaches~~ leveraging long-term groundwater quality data is a promising approach~~- and spanning long periods, are a promising path~~ (He et al., 2022; Rodriguez-Galiano et al., 2014)~~-~~. Such data encode all ~~the~~ relevant processes of the nitrate contamination process, they ~~and~~ are ~~, which offer the advantage of bein~~widely available and they ~~,~~ enabl~~inge~~ broader insights into nitrate pollution trends and drivers. For example, Vanclooster et al. (2020) used a data-driven approach to model groundwater nitrate pollution in the Brussels-Capital Region, Belgium, focussing on ~~, however they focused on~~ spatial variability~~, and did not address the drivers of~~ but not on the temporal evolution of nitrate concentrations.

~~(SPW - DEE - Direction des Eaux souterraines, 2024)~~This study aims to leverage groundwater monitoring data to ~~assess~~ interpret the effect ~~iveness~~ of the PGDA on groundwater quality in Wallonia over ~~an almost 20 year period~~nearly two decades. ~~In particular~~ Specifically, our objectives are:

(i) To assess the long-term (2002 – 2020) evolution of nitrate concentrations ~~since the~~following the PDGA implementation ~~of the PGDA~~;

(ii) To identify the factors driving ~~controlling~~ the nitrate concentration changes over time and across different locations.

2. Material and Methods

~~**To address the first objective, we captured the nitrate concentration dynamics between 2002 and 2020 with six indicators. To address the second objective, we computed a set of potential explanatory variables standing for environmental and anthropogenic factors which may impact the nitrate concentration dynamics. We then assessed the relationships between these explanatory variables and six nitrate pollution indicators.**~~

2.1 Study area

Our study area focuses on ~~the~~key -Nitrate Vulnerable Zones (NVZ) of the Region Wallonia, Belgium, where ~~because X~~nitrate contamination levels are the highest (SPW - DEE - Direction des Eaux souterraines, 2024). Specifically, ~~I~~it includes all monitored sites where levels currently exceed 50 mg/l. ~~since when?~~The PGDA defines ~~additional~~stricter -measures which are applicable only in these areas, such as specific spreading periods and conditions, obligations regarding soil cover, and monitoring requirements (Picron et al., 2017). ~~to mitigate the contamination~~We selected four groundwater bodies where ~~the~~a major drinking water company supplying the monitoring data operates several extraction sites:~~We selected four groundwater~~

~~bodies located in the Region of Wallonia, Belgium~~: the Geer basin chalks, the Brusselian sands, the Haine basin chalks and the Landenian sands (Figure 1~~Figure 1~~). ~~-~~The concept of "water body" was introduced within the Water Framework Directive to classify the various aquatic environments that characterize the European territory. A groundwater body consists of a distinct volume of groundwater within one or more aquifers. In Wallonia, the groundwater bodies were delineated by a group of experts based on hydrogeological criteria such as the extent of geological layers or the interaction with surface waters, as well as non-hydrogeological criteria such as the administrative limits.

~~The four selected groundwater bodies are part of Nitrate Vulnerable Zone (NVZ). The NVZ covers 69% of Wallonia's utilized agricultural land, and includes all monitored sites where levels exceeding 50 mg/l have been measured.~~ The Geer basin chalks and the Brusselian sands were part of the first nitrate vulnerable zones as defined in 1994, and have thus been subjected to the associated regulations since then. The Haine basin chalks and the Landenian sands were added in the NVZ in 2013.

The Geer basin chalks groundwater body covers an area of 440 km² and is located in the Meuse hydrographic basin (SPW, 2016). The groundwater body's aquifer, the Hesbaye aquifer, is said to have a substantial storage capacity and a high porosity. It is partly overlaid in its northeast portion by the Landenian sands groundwater body. Agricultural land covers approximately 68% of the land surface, with 14% of it being meadows and 86% crops. The region has a high population density with 340 inhabitants per square kilometre.

The Landenian sands groundwater body spans a surface of 206 km² and is located within the Scheldt hydrographic basin (SPW, 2010). The groundwater body's aquifer is the Landenian sands aquifer. Due to limited exploitation, the hydrogeological properties of this aquifer remain poorly defined. Agricultural land covers approximately 78% of the total land surface, with meadows accounting for 8% and crops making up the remaining 92%. The population density is relatively low, with 160 inhabitants per square kilometre.

The Brusselian sands groundwater body spans a surface of 964,5 km² and is situated in the Scheldt hydrographic basin (SPW, 2006b). Its aquifer is the Brusselian sands aquifer, which has a high storage capacity but a low hydraulic conductivity. Agriculture covers 71% of the land surface and another 10% is urban land.

The Haine basin chalks groundwater body covers an area of 644 km² and is situated in the Scheldt hydrographic basin (SPW, 2006a). The main aquifer of this water body is the Mons basin chalks aquifer. The aquifer's porosity has a permeability ranging from 10^{-5} to 10^{-7} m/s, while the fissures in the chalk formation entail a permeability of 2.10^{-3} to 5.10^{-5} m/s. In the northwest, the groundwater body is partially overlaid by the Haine valley sands groundwater body. The land surface area consists of 64% agricultural land and 23% urban land.

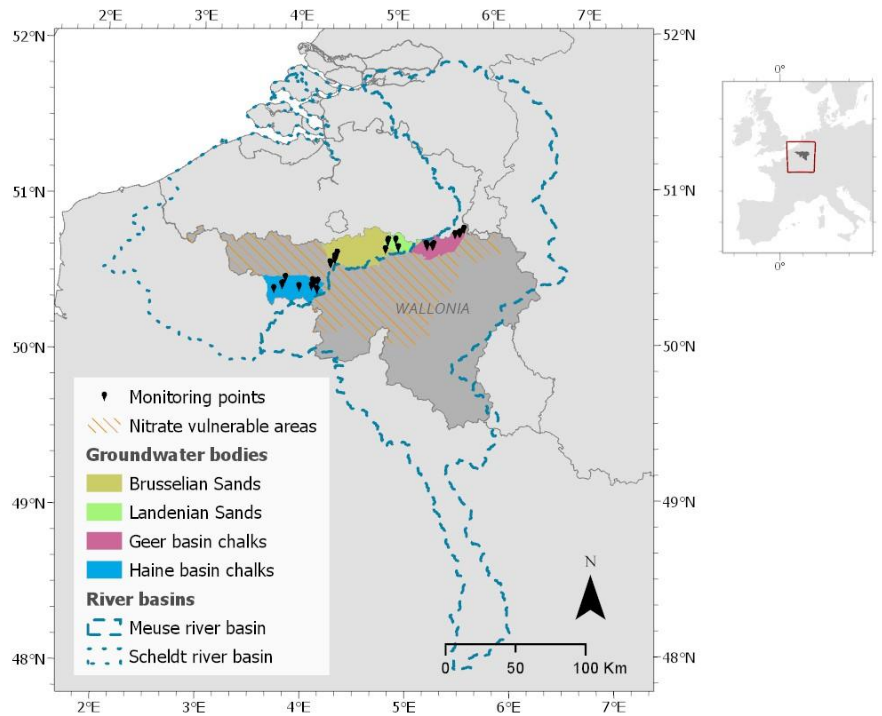


Figure 1. Location of the four groundwater bodies and the 36 monitoring points.

2.2 Groundwater nitrate concentration ~~dataResponse variables~~

2.2.1 Data collection~~Nitrate concentration data~~

The monitoring points are water intake structures exploited for drinking water production by the Société Wallonne des Eaux (SWDE). They are wells, galleries, springs or drains. ~~We selected points o~~Only points in unconfined ~~parts of the~~ aquifers, ~~and for which~~with a ~~the~~ water quality monitoring period ~~was longer than~~exceeding ten years were selected. Since our focus is on vulnerable groundwater with high nitrate concentrations, we excluded one monitoring point located in anoxic groundwater, where nitrate levels are below 3 mg/l, probably resulting from high denitrification rates ~~We removed points located in anoxic groundwater as they are prone to denitrification~~ (Rivett et al., 2008). ~~We defined the anoxic conditions as O_2 concentrations inferior to 0.5 mg/l and Mn concentrations superior to 0.05 mg/l according to~~ The final dataset included~~contained~~ 36 ~~monitoring~~ points from which 13 points are in the Haine basin chalks, nine points in the Geer basin chalks, nine points in the Brusselian sands and five points in the Landenian sands (Figure 1~~Figure 1~~).

Time series were available for some points from the sixties onwards, but ~~T~~the trend and causal analysis covers 2002-2020~~We focused our analysis on the period 2002 to 2020~~, starting at the onset ~~to start at the onset~~ of the sustainable nitrate management program (PGDA) and ending with the most recent year for which data were available. For nine out of the 36 points the first available data started after 2002: seven in 2003, one in 2006, and one in 2009. ~~-~~The temporal resolution of the nitrate measurements is variable, with the total number of measurements per point ranging from 53 to 948 over the study period. The water samples were analysed by the laboratory of the drinking water production company, the SWDE, under ISO 17025 accreditation.

The nitrate concentration time series contained some problematic values that were noticeably lower or higher than their neighbours due to reported human errors. We thus filtered the time series using a moving window of two years, with an upper and lower limit being the mean of the data within the window plus and minus three times the standard deviation. Figure 2~~Figure 2~~ shows the annual averages of the resulting time series.

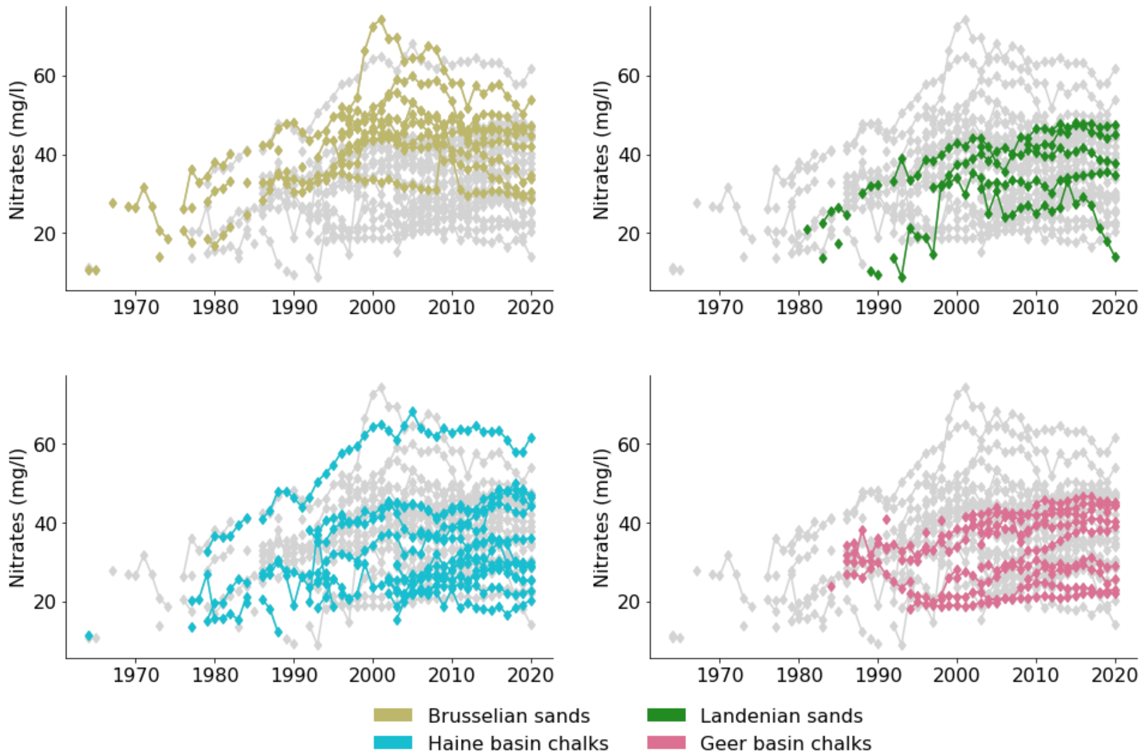


Figure 2. Time series of the yearly mean nitrate concentrations at the 36 monitoring points, sorted by groundwater body. In grey, the time series of all points. In colour, the time series of the points in each groundwater body.

2.2.2 Definition of nitrate pollution indicators~~Nitrate pollution indicators~~

We defined a set of six indicators that ~~collectively provide~~capture~~ a comprehensive view of~~ the state and rate of change~~s~~ in nitrate concentration in ~~over time, between~~ 2002~~, and~~ in 2020 and the evolution in between. The indicators are used to assess the spatial and temporal evolution of nitrate contamination at each monitoring point, and, they serve as dependent variables in our data analysis to identify factors influencing nitrate concentrations. The indicators and their interpretation are defined in Table 1~~Table 1~~ and illustrated in Figure 3~~Figure 3~~.

Table 1. Definition and interpretation of the nitrate pollution indicators.

Pollution indicator (I)		Unit	Usage and interpretation
I1	Average nitrate concentration in 2002	mg/l	Snapshot of nitrate concentration levels in the groundwater before the implementation of the PGDA. Serves as a baseline for comparison with future years.
I2	Average nitrate concentration in 2020	mg/l	Snapshot of nitrate concentration levels in the groundwater in a recent year. Allows for direct comparison with past data to assess changes over time.
I3	Concentration difference between 2020 and 2002 (I3 = I2-I1)	mg/l	Indicates how the concentration has changed since the implementation of the PGDA. A positive value indicates an increase in nitrate levels, while a negative value indicates a decrease.
I4	Slope in 2002	mg/l/year	Rate of change in nitrate concentration on 01-01-~~in~~ 2002. Provides insight into how rapidly nitrate levels were changing at the beginning of the period.
I5	Slope in 2020	mg/l/year	R~~Recent r~~ate of change in nitrate concentration on 01-01-2020. Useful for understanding recent dynamics and informing current policy decisions.
I6	Difference in slope between 2020 and 2002 (I6 = I5-I4)	mg/l/year	Indicates how the rate of change in nitrate concentration has changed over the study period. A positive value indicates an accelerating increase (or decelerating decrease) in nitrate levels, while a negative value suggests a decelerating increase (or accelerating decrease).

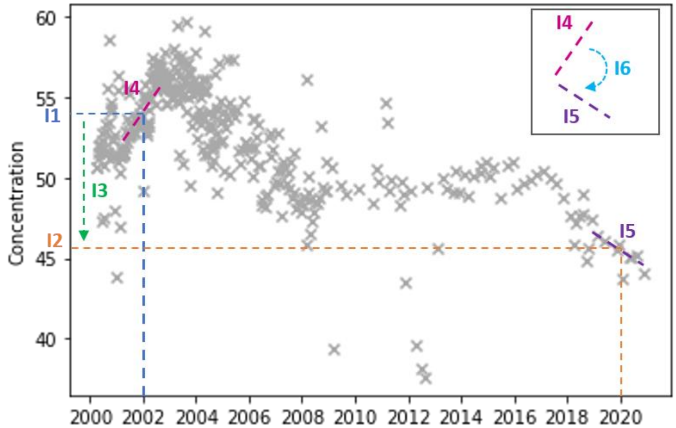


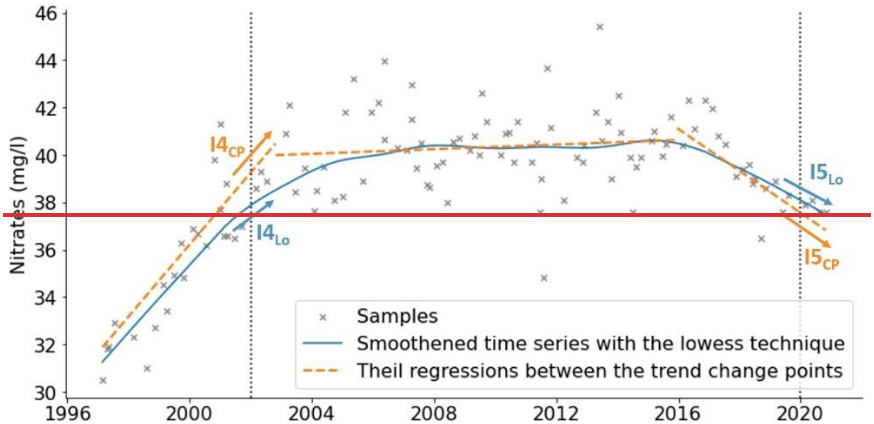
Figure 3. Illustration of the six indicators of the nitrate ~~concentration status and long-term change~~pollution indicators. ~~We use the~~T~~t~~imes series data from ~~of one~~ a ~~the~~ monitoring ~~points~~ station in the ~~X~~Brusselian sands ~~body~~is ~~are~~ used for illustrative purposes~~as example~~.

~~Note that we calculated the rate of change (slope) at the start and the end of the considered time period instead of conducting a single trend test. This choice is because the change in concentration values in most monitored locations was not gradual (Figure 2) but showed irregular or non-linear patterns within the considered time period, which is not suitable for a simple trend test.~~

The method used for slope computation (I4 and I5)~~can~~ affects~~lead~~ ~~to different conclusions in~~the trend diagnostics. The non-parametric Mann-Kendall test has been widely applied and recommended for groundwater pollution trend assessments (Frollini et al., 2021; Grath et al., 2001; Hirsch et al., 1991; Urresti-Estala et al., 2016). This method only captures a single linear trend, making it not directly applicable to detect trend changes in long, non-linear time series. To address this limitation, Lee et al. (2010) computed trends using local regression. In our analysis, ~~W~~we applied ~~used~~ two different approaches: ~~, based on~~ (i) computing tangent lines to a l~~l~~ocal~~owess~~ regression~~(Lo)~~ and ~~on~~ (ii) detecting change point and applying the Mann-Kendall test ~~(CP).~~, Since both methods produced similar slope indicators, we focus on the local regression approach in the main text, with details of the change-point detection approach provided in the supplementary materials.~~and tested whether the results from both approaches corroborated with each other.~~ ~~We referred to the indicators as $I4_{Lo}$ and $I5_{Lo}$, and $I4_{CP}$ and $I5_{CP}$, respectively.~~ The local regression~~For the Lo~~ approach consisted in~~we~~ ~~first smoothed~~smoothing the time series using locally weighted scatterplot smoothing (lowess) (Cleveland, 1979) and ~~we then determined $I4_{Lo}$ and $I5_{Lo}$ by~~ computing the slope of the tangent line ~~at the points corresponding to~~to the smoothed time series on January 1, 2002 (=I4), and on January 1, 2020 (=I5)~~, respectively~~. The lowess window length ~~for lowess~~ was set to eight years for time series longer than twelve years, allowing to capture the long-term fluctuations, and to 2/3 of the time series length for shorter series. Within a local window of 90 days, a linear interpolation was applied instead of a weighted regression to increase stability.

~~In the CP approach, we determined $I4_{CP}$ and $I5_{CP}$ as the slopes in 2002 and 2020 of the segments between the change points of the time series. The segments' slopes were calculated using the Theil slope estimator, robust to outliers~~

For the nine monitoring points whose first measurements were taken after 2002, the slope of the trend in 2002 ($I4_{Lo}$ ~~and $I4_{CP}$~~) was taken as the slope at the first data point, and the mean absolute nitrate concentration in 2002 (I1) was obtained by hindcasting that trend~~, using $I4_{CP}$~~. All time series cover a period longer than eight years, the minimum length recommended by Grath et al. (2001) for groundwater pollution trend assessments.



~~**Figure 4. Illustration of the two alternative methods, the lowess (Lo) and change point (CP) methods, used to model the time series and identify the trends in 2002 (I4 indicators) and in 2020 (I5 indicators).**~~

2.3 ~~Defining~~ ~~Nnitrate pollution~~~~contamination~~Candidate ~~potential~~explanatory variables ~~Potential explanatory variables~~

We computed a set of potential explanatory variables standing for environmental and anthropogenic factors which may impact the nitrate concentration dynamics. We used these variables as independent variables in our data analysis to identify factors influencing nitrate concentrations. ~~We tested the ability of a set of variables to explain the spatial and temporal variability of the nitrate concentration indicators.~~ The variables encompass both the inherent vulnerability to pollution and the anthropogenic influence, which include human activities that could cause or affect punctual and/or diffuse pollution. The definitions of these indicators, computation means and data sources are given in Table 2~~Table 2~~. The descriptive statistics of the values of these variables for the monitored locations are given in

Table 3

~~Table 3~~.

2.3.1 Delineation of the influence zones

The risk of groundwater contamination at a specific location is influenced by the traits of the land surface area that can potentially transport pollutants to it. Therefore, delineating this land surface area, here referred to as the 'influence zone', is crucial in the analysis of groundwater nitrate concentrations (Mattern et al., 2009). We defined the influence zones as the topographic surface watersheds of the legal~~ly~~ protect zones of~~ing~~ each water intake structure (SPW, 2020). These zones correspond either to the groundwater table area with a maximum transfer time of 50 days to the water intake structure as estimated through geological modelling, or they are defined as circular areas centred around the structure location, whose radius depends on the aquifer substrate: 100 meters for sandy aquifers, 500 meters for gravel aquifers and 1000 meters for karstic aquifers. The boundaries of these zones are available on the regional institution's geographical data portal (SPW, n.d.). We delineated the watersheds with the ArcGIS watershed toolbox, using a 2-meter resolution raster of flow direction and flow accumulation generated by the LIDAX project (SPW, 2019).

2.3.2 Inherent vulnerability

To quantify the natural vulnerability, we considered the seven factors of groundwater natural vulnerability to pollution as defined in the DRASTIC model of the U.S. Environmental Protection Agency (Aller et al., 1985): aquifer depth, recharge, aquifer media, soil type, the topography, impact of the vadose zone, and hydraulic conductivity. We used the depth of the water intake structures as a proxy of the depth to the groundwater table since piezometric measurements were not available for all structures. We used the mean annual rainfall as a proxy for the net recharge. We used a single categorical variable, namely the groundwater body of the water intake structure, as a proxy for the three DRASTIC vulnerability factors aquifer media, impact of the vadose zone and hydraulic conductivity. The variable standing for the topography was the mean slope in the influence zone calculated using a 2 m resolution digital slope product derived from a 1 m digital elevation model. We did not include the soil in the set of explanatory variables since the main soil type of all influence zones was identical, namely loam. We considered all the vulnerability variables to be time-invariant over the studied period.

To be able to include the categorical variable aquifer media (GWbody) in our analysis, we replaced it by four binary variables using one-hot encoding. We called the new variables GWbodyLS, GWbodyBS, GWbodyHBC and GWbodyGBC, they indicate respectively the monitoring points in the Landenian sands, the Brusselian sands, the Haine basin chalks and the Geer basin chalks. We do not expect any change of these variables over the study period.

2.3.3 Land ~~cover and land~~ use characteristics

Agricultural land is a significant source of nitrate leaching to groundwater (Cameron et al., 2013; Strebel et al., 1989), with nitrates originating from nitrogen fertilizers applied on cropland, and from grazing livestock on meadows. Consequently, both crop land and meadow area were considered as potential drivers in this study. We included specifically ~~the~~ potato crop cover, as it is known to leave a high concentration of potentially leachable nitrates in the upper soil layer after the growing season (Bah et al., 2015). In contrast, forested and green areas are generally less prone to nitrate leaching (Cameron et al., 2013; Zhang et al., 2013), hence we also considered it as driver, but expecting them to have a mitigating impact. Urban wastewater losses represent an additional nitrate source (Torres-Martínez et al., 2020). We used built infrastructure area as a proxy to estimate wastewater production. ~~We included specifically the potato crop cover as it is known to leave a high concentration of potentially leachable nitrates in the upper soil layer after the growing season .~~ We visually examined the trends in ~~land use type covers~~these variables over the study period~~ through a visual analysis~~, ~~which revealed~~finding that only the meadow area~~s~~

exhibited a trend. Therefore, ~~W~~we ~~thus~~ included a sixth variable that accounts for ~~,~~ the change in meadow cover over the study period.~~. This is particel~~

2.3.4 ~~Potential punctual~~Point pollution sources

~~Finally, we~~We considered three ~~types of~~ potential point~~unctual~~ pollution sources. Graveyards can contribute to nitrate pollution through the decomposition of organic materials (Mattern et al., 2009). Farm buildings are included because they are often associated with manure storage and handling, which can be a direct source of nitrates if not managed properly. Finally, buildings not connected to the collective sewage system indicate potential nitrate sources from septic tanks or dry wells, which can leach untreated wastewater. ~~, i.e. features related to the presence of graveyards, farms and buildings not connected to the collective sewage system. The latter can indicate the presence of septic tanks or dry wells.~~ The datasets used capture the situation in 2020 (Table 2~~Table 2~~). We expect little change for these variables over the studied period.

Table 2. Definition of the potential explanatory variables used in the statistical models. IZ: influence zone. SPW: Service Public de Wallonie. SPF: Service Public Fédéral. SPGE : Société Publique de Gestion de l'Eau. SWDE : Société Wallonne de Gestion de l'Eau. IRM : Institut Royal Météorologique.

Variable	ID	Definition and unit	Source dataset	Provider
Water extraction depth	Depth	Depth of the bottom part of the water intake structures (m)	Documentation of water intake structures	SWDE
Groundwater body	GWBody	Groundwater body (Brusselian sands, Landenian sands, Haine basin chalks and Geer basin chalks)	Documentation of water intake structures	SWDE
Mean annual ~~R~~rainfall	Rainfall	Interannual (1961-2019) average of annual rainfall (mm/year)	1961-2019, 5000 m resolution climate dataset for Belgium	IRM
Topography	TerrainSlope	Average of the terrain slope in the IZ (%)	1 m resolution digital slope model 2013-2014	SPW
Crop cover	CropLU	Interannual (1998-2019) mean percentage of IZ area with crop cover (%)	Anonymous agricultural land registry (annual data from 1998-2019)	SPW
Potato crop cover	PCropLU			
Meadow cover	MeadowLU			
Change in meadow cover	MeadowReg	Trend in meadow area calculated as the slope of the linear regression of the yearly meadow		

		area percentage between 2002 and 2020 (%/year)		
Built area	BuiltLU	Interannual (1998-2019) mean percentage of IZ area with built infrastructures (%)	Walloon land registry (annual data from 1998-2019)	SPF
Forested and green areas	GreenLU	Percentage of IZ area with forests and green spaces in 2003 (%)	Land cover map 2003	SPF
Presence of farm(s)	Farms	Number of farms in the IZ in 2020 divided by the surface of the IZ (nb/km^2)	Continuous cartographic mapping project	SPW
Presence of graveyard(s)	Graveyards	Number of graveyards in the IZ in 2020 divided by the surface of the IZ (nb/km^2)	Continuous cartographic mapping project	SPW
Buildings with automonous sewage regime	NoSewage	Number of buildings not connected to the sewage system in 2020 divided by the surface of the IZ (nb/km^2)	Walloon land registry (2020) Wastewater management plan	SPF SPGE

Table 3. Descriptive statistics of the independent variables for the 36 monitoring points. NA: not applicable.

Variable	Unit	Mean	Min	Median	Max
CropLU	%	57	6	53	91
PCropLU	%	5	0,1	5	14
MeadowLU	%	8	0,4	7	19
MeadowReg	%/year	-0,02	-0,66	-0,09	0,92
BuiltLU	%	3	0,2	2	8
GreenLU	%	6	0,1	3	63
Farms	nb/km^2	0,3	0	0,4	1,2
Graveyards	nb/km^2	0,06	0	0	0,3
NoSewage	nb/km^2	0,2	0,03	0,1	0,5
Depth	m	29	0,6	19	120

Rainfall	mm/year	814	760	807	867
TerrainSlope	%	4,6	2,6	4,4	9,0
Groundwater body	-	NA	NA	NA	NA

2.4 Data analysis

2.4.1 ~~Status and temporal e~~Evolution of the nitrate concentrations

We computed descriptive statistics and visuals to depict the past and present status and temporal changes of the nitrate concentrations in the studied groundwater bodies, using the nitrate pollution indicators ~~of~~ (Table 1).~~.~~

2.4.2 Bivariate analysis to identify the controlling factors

We tested the strength and direction of the association between the nitrate ~~concentration~~ pollution indicators (Table 1) and each independent variable (Table 2) separately by computing the Kendall rank correlation (also known as Kendall's tau coefficient, Kendall, 1938). The value of Kendall's tau ranges from -1 to 1 and the closer the coefficient is to either -1 or 1, the stronger the association. A higher positive value indicates a strong positive association, while a higher negative value indicates a strong negative association.

2.4.3 Multivariate linear regressions to identify the controlling factors

We used multiple linear regression to assess the individual contribution of each independent~~explanatory~~ variable (Table 2) ~~while considering the influence of the other variables, and thus, accounting~~to account for potential confounding effects.
Before applying the regression models, we ~~identified and removed variables that demonstrated high~~addressed multicollinearity by removing highly collinear variables, ~~. This approach was essential to~~hence ensur~~e~~ing that the remaining variables ~~in the model could~~ provide ~~clearer, more distinct~~reliable contributions to the analysis.~~s, enhancing the reliability and accuracy of our model's results.~~
First, we replaced the four binary variables representing the groundwater bodies, which were highly collinear, with one single binary variable 'Aquifer', distinguishing ~~representing~~ the aquifer media of the groundwater bodies: a value of 1 for ~~. A value of 1 indicates monitoring points in~~ the Brusselian and Landenian sands and a value of 0 ~~indicates monitoring points in~~for the Geer and Haine basin chalks.
We then removed one variable at a time until the variance inflation factor (VIF) values ~~(~~(Mansfield & Helms, 1982) of all remaining variables were below a threshold of 5 (James et al., 2013). The variable to remove at each iteration was selected based on its VIF value, correlation with other variables and importance as explanatory variable, assessed by the authors' expert judgment.

~~Using the remaining variables, we~~We built nine multiple linear regression models with the remaining independent variables, one for each indicator and slope calculation method.~~,~~ We used~~using~~ the ordinary least square (OLS) function of the Python *statsmodels* library (Seabold et al., 2010). We standardized the independent variables to ~~have~~ a mean of zero and a standard deviation of one,~~ which~~ facilitat~~es~~ing the comparison of their respective impacts on the nitrate concentration indicators. We applied a stepwise multiple linear regression to identify the most important predictor variables, eliminating at each iteration the independent variable with the highest p-value until the p-values of all remaining variables were below 0.05.

3. Results

3.1 ~~Status and temporal evolution~~Evolution of the nitrate concentrations

Table 4~~Table 4~~ shows statistical summaries for the six ~~different~~ pollution indicators. It shows that the average concentration of nitrates in 2002 was 37.7 mg/l with a standard deviation of 12.2 mg/l and hardly decreased in 2020 with an average of 36.5 mg/l and a standard deviation of 10.8 mg/l. The mean change in concentration levels between 2002 and 2020 (I3) exhibits a slight decrease of 1.2 mg/l on average, but with a wide variation (standard deviation of 8.8 mg/l), ranging from a decrease of 21.3 mg/l to an increase of 12.5 mg/l. Forty seven percent ~~–~~of the monitored locations have witnessed a decrease in concentration, while the other 53% have seen an increase (Figure 4~~Figure 5~~). The average rate of change in nitrate concentrations (I4 and I5) are slightly negative whichever the method, but with variations ranging from a negative to a positive rate. They are slightly more negative in 2020 than in 2002. The maximum rate of change has decreased from ~~+2.2 mg/l/year ($I4_{CP}$) or~~ +2.7 mg/l/year~~ ($I4_{Lo}$)~~ in 2002 to 0.6 mg/l/year in 2020, which indicates an overall deceleration of the rate of change over the study period. This is confirmed by the negative values of the averaged I6.

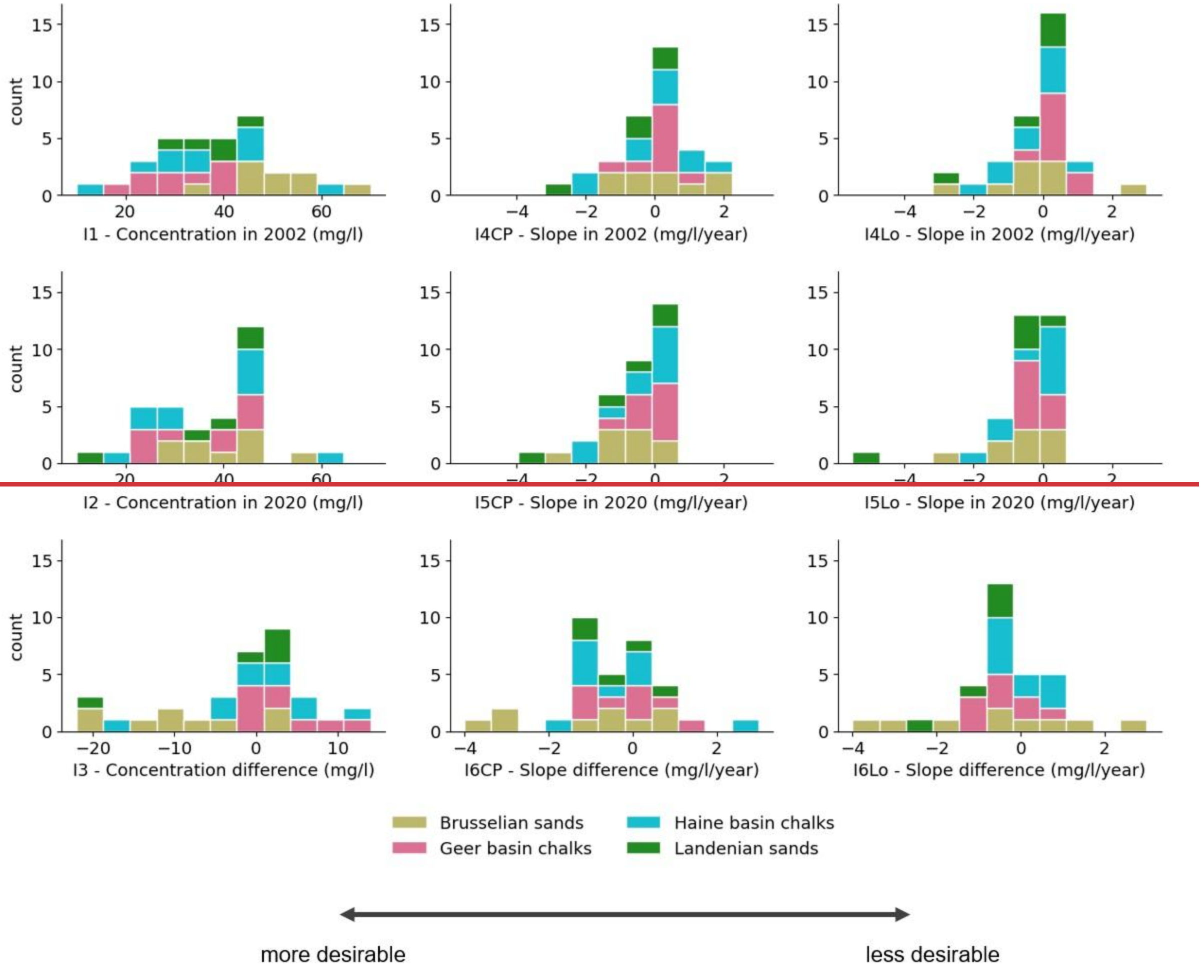
While these statistics indicate a slightly mild decrease in nitrate concentrations since 2002, the distributions of the indicators color-coded by the type of aquifer in Figure 4~~Figure 5~~ suggest that the decrease has mainly been significant in the Brusselian sands, while the pink stacks of the histogram for I3 indicate an increase in the Geer basin chalks.

~~**The statistics for the rate of change (I4 and I5) and the difference in rate (I6) are comparable regardless of the slope calculation method employed (CP or Lo).**~~

Table 4. Descriptive statistics of the six nitrate indicators. ~~**CP and Lo: change point and Lowess method for slope computation.**~~

Pollution indicator (I)	Unit	Mean ± standard deviation	Minimum	Percentile 25 (Q1)	Median Q2	Percentile 75 (Q3)	Maximum	IQR (Q3-Q1)
I1 – Concentration in 2002	mg/l	37,7 ± 12,2	15,4	28,3	38,1	44,3	69,2	16,0

I2 - Concentration in 2020	mg/l	36,5 ± 10,8	14,1	28,6	37,0	44,7	61,6	16,1
I3 - Concentration difference	mg/l	-1,2 ± 8,8	-21,3	-4,9	0,1	3,6	12,5	8,5
~~I4CP – Slope in 2002 – CP~~	~~mg/l/year~~	~~0,0 ± 1,1~~	~~-2,6~~	~~-0,4~~	~~0,1~~	~~0,6~~	~~2,2~~	~~0,9~~
I4~~Lo~~ – Slope in 2002~~ – Lo~~	mg/l/year	-0,1 ± 1,1	-3,1	-0,5	0,1	0,5	2,7	1,1
~~I5CP – Slope in 2020 – CP~~	~~mg/l/year~~	~~-0,5 ± 1,0~~	~~-3,9~~	~~-0,9~~	~~-0,2~~	~~0,2~~	~~0,6~~	~~1,1~~
I5~~Lo~~ – Slope in 2020~~ – Lo~~	mg/l/year	-0,5 ± 1,1	-5,2	-0,7	-0,2	0,1	0,6	0,8
~~I6CP – Slope difference – CP~~	~~mg/l/year~~	~~-0,5 ± 1,2~~	~~-3,6~~	~~-1,1~~	~~-0,6~~	~~0,0~~	~~2,4~~	~~1,1~~
I6~~Lo~~ – Slope difference~~ – Lo~~	mg/l/year	-0,4 ± 1,2	-3,6	-0,8	-0,4	-0,0	2,7	0,7


count
I1 - Concentration in 2002 (mg/l)
I4CP - Slope in 2002 (mg/l/year)
I4Lo - Slope in 2002 (mg/l/year)
I2 - Concentration in 2020 (mg/l)
I5CP - Slope in 2020 (mg/l/year)
I5Lo - Slope in 2020 (mg/l/year)
I3 - Concentration difference (mg/l)
I6CP - Slope difference (mg/l/year)
I6Lo - Slope difference (mg/l/year)
Brusselian sands
Geer basin chalks
Haine basin chalks
Landenian sands
more desirable
less desirable

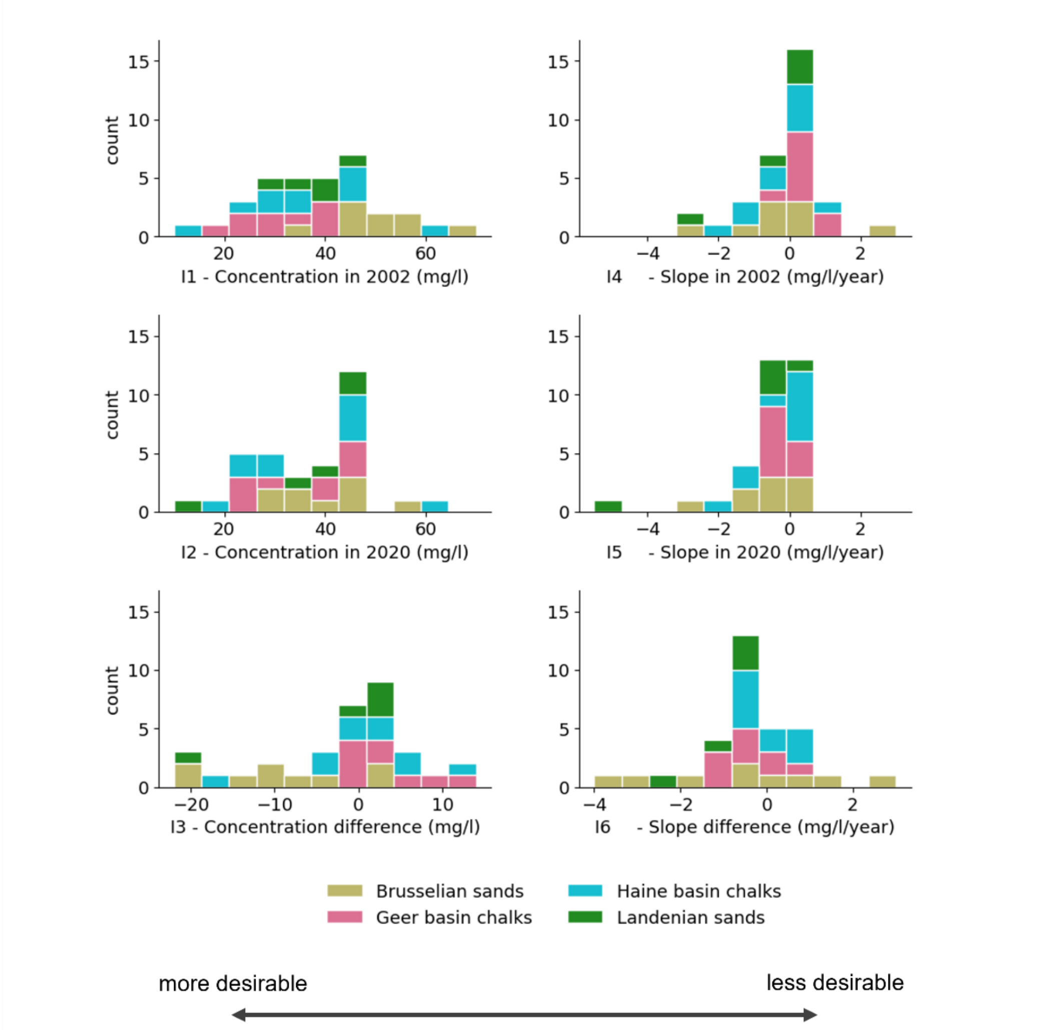


Figure 4~~5~~. Histograms of the six pollution indicators. The black arrow at the bottom indicates thate, for each indicator, lower values are more desirable than higher values, as this entails lower nitrate concentrations, decreasing trends and a decrease in concentration from 2002 to 2020. ~~CP: change point method for slope. Lo: Lowess method for slope.~~

3.2 Identification of the controlling factors

3.2.1 Bivariate analysis

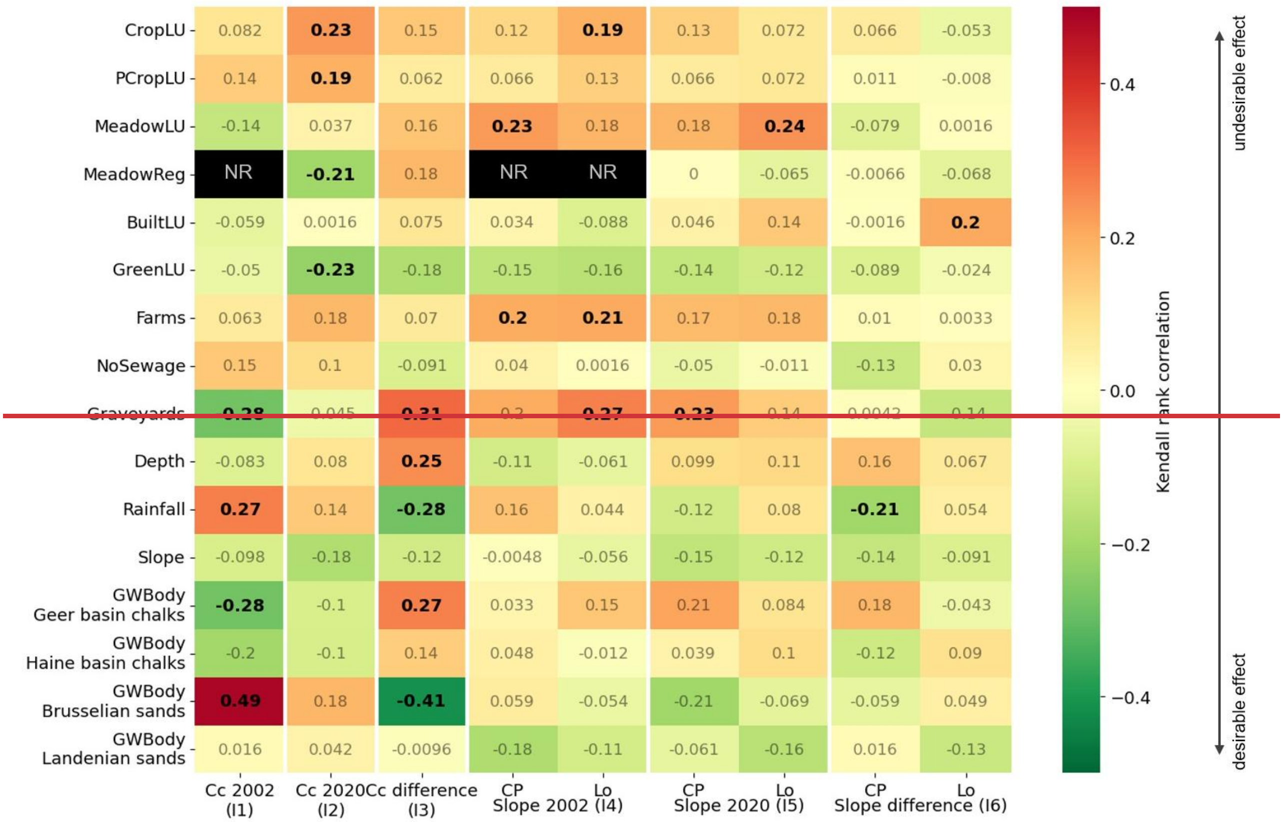
The analysis reveals a consistent positive correlation between the crop and potato crop area and all the pollution indicators, and hence an undesirable effect of these variables (Figure 5~~Figure 6~~). This effect is only significative on the nitrate concentration in 2020 (I2) ~~-~~and rate of change in 2002 (I4).~~with the Lo method (I4$_{Lo}$).~~ On the other hand, the analysis reveals a consistent negative correlation between the forest and green space area and all the pollution indicators, and hence a desirable effect of this variable. This effect is only significative on the nitrate concentration in 2020 (I2). Results also indicate a significant positive relationship, or undesirable effect, of the meadow area on the rate of change in ~~both 2002 and~~ 2020 (I5~~I4CP and I5Lo~~). Temporal trend in the meadow area (MeadowReg) shows a significant negative correlation with concentrations in 2020 (I2), suggesting a desirable effect of an increase in meadow area on I2, but shows a positive correlation with the change in concentrations (I3), suggesting an undesirable effect of an increase in meadow area on I3.

The results show a positive relationship and hence undesirable effect of the number of farms on all indicators, and of the number of graveyards on the change in concentrations (I3) and the rate of changes in 2002 and 2020 (I4 and I5). Conversely, graveyards displayed a significant negative relationship with concentrations in 2002 (I1). Finally, there is no detected influence of the presence of building area, a proxy for population density, and buildings not connected to wastewater treatment plants.

The depth of the water intake structure, serving as a surrogate for groundwater table depth, shows a significant correlation with the indicator of change in concentration (I3). However, it does not exhibit any correlation with the other indicators. There is a significant negative relationship between the annual rainfall, a proxy for recharge, and the indicator of change in concentrations (I3), ~~and change in rate of change (I6$_{CP}$),~~ while there is a significant positive relationship with the concentration in 2002 (I1).

A weak negative relationship (positive effects) was found between the terrain slope and all the indicators.

Results also confirm a clear influence of the aquifer media. While concentrations in 2002 and 2020 were higher in Brusselian sands and lower in the chalks aquifers, the decrease and rate of decrease has been more prominent in the sands.


CropLU 0.082 0.23 0.15 0.12 0.19 0.13 0.072 0.066 -0.053
PCropLU 0.14 0.19 0.062 0.066 0.13 0.066 0.072 0.011 -0.008
MeadowLU -0.14 0.037 0.16 0.23 0.18 0.18 0.24 -0.079 0.0016
MeadowReg NR -0.21 0.18 NR NR 0 -0.065 -0.0066 -0.068
BuiltLU -0.059 0.0016 0.075 0.034 -0.088 0.046 0.14 -0.0016 0.2
GreenLU -0.05 -0.23 -0.18 -0.15 -0.16 -0.14 -0.12 -0.089 -0.024
Farms 0.063 0.18 0.07 0.2 0.21 0.17 0.18 0.01 0.0033
NoSewage 0.15 0.1 -0.091 0.04 0.0016 -0.05 -0.011 -0.13 0.03
Graveyards -0.28 0.045 0.31 0.2 0.27 0.23 0.14 0.0042 -0.14
Depth -0.083 0.08 0.25 -0.11 -0.061 0.099 0.11 0.16 0.067
Rainfall 0.27 0.14 -0.28 0.16 0.044 -0.12 0.08 -0.21 0.054
Slope -0.098 -0.18 -0.12 -0.0048 -0.056 -0.15 -0.12 -0.14 -0.091
GWBody Geer basin chalks -0.28 -0.1 0.27 0.033 0.15 0.21 0.084 0.18 -0.043
GWBody Haine basin chalks -0.2 -0.1 0.14 0.048 -0.012 0.039 0.1 -0.12 0.09
GWBody Brusselian sands 0.49 0.18 -0.41 0.059 -0.054 -0.21 -0.069 -0.059 0.049
GWBody Landenian sands 0.016 0.042 -0.0096 -0.18 -0.11 -0.061 -0.16 0.016 -0.13
Cc 2002 (I1)
Cc 2020 (I2)
Cc difference (I3)
CP Lo Slope 2002 (I4)
CP Lo Slope 2020 (I5)
CP Lo Slope difference (I6)
0.4
0.2
0.0
−0.2
−0.4
Kendall rank correlation
undesirable effect
desirable effect

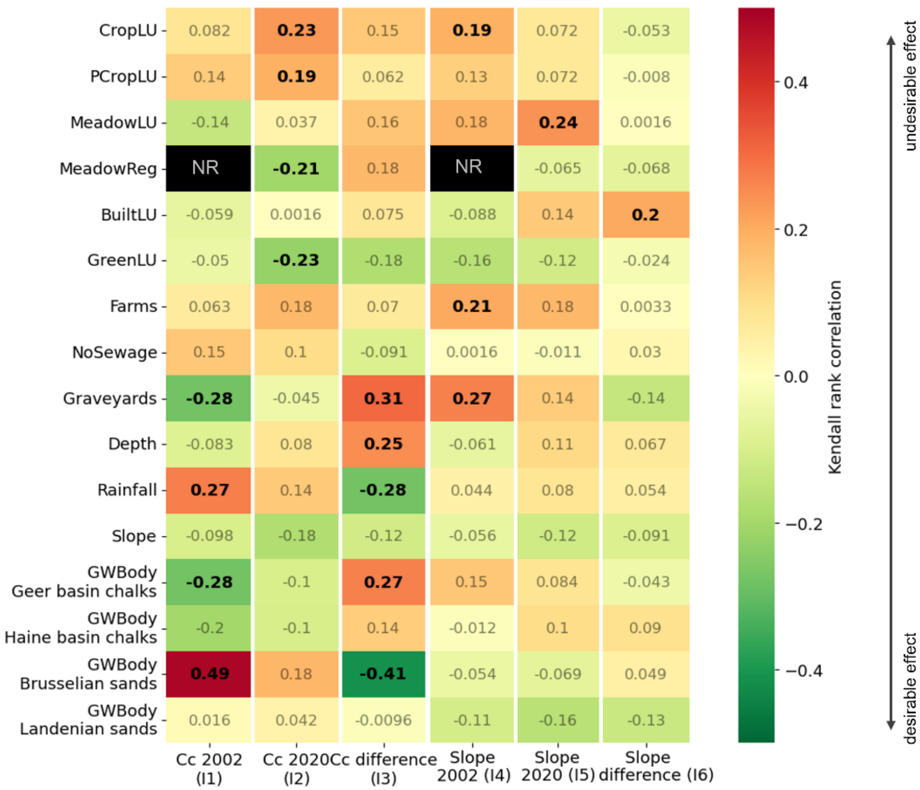


Figure 5~~6~~ Heatmap of the Kendall rank correlation coefficients between the explanatory variables and the six pollution indicators. Coefficients in bold indicate a significant relationship (p-value<0.1) NR: not relevant. The black arrow on the right indicates that low correlation is more desirable than high correlation, as low correlation entails that high values of the explanatory variables are correlated to low values for the six pollution indicators and vice versa.

3.2.2 Multivariate linear regression analyses

In the process of refining our multivariate regression models, we made several adjustments to address issues of multicollinearity among the variables. We removed the variable ‘Rainfall’ for its high VIF of 242. The variables ‘CropLU’ and ‘PCropLU’ had a VIF of respectively 20 and 12, and were highly correlated (R^2=0.77). Considering the importance of 'CropLU' as explanatory variable, we chose to retain it and remove 'PCropLU.' The next variables we removed were ‘Slope’ and ‘NoSewage’ as they each had the highest VIF among the remaining variables (9 and 7 respectively). Finally, we removed the variables ‘Farms’ and ‘Graveyards’, allowing to satisfy the condition of all remaining variables having a VIF < 5. The decision to exclude 'Farms' was based on its redundancy with ‘CropLU’ and ‘MeadowLU’, which already represent agricultural activity. As for ‘Graveyards’, their very sparse distribution in the considered areas led us to expect a limited effect.

The results of the multivariate regression models with the selected independent variables are presented in Table 5~~Table 5~~. The table shows the coefficients of the independent variables for each model, after stepwise removal of all non-significant variables (p-value of coefficient < 0.05). Note that the interpretation of the coefficient value is difficult because the independent variables have been normalized. In this normalized context, the coefficient indicates the expected change in the dependent variable per standard deviation change in the independent variable. However, it allows to interpret the relative importance of the variables, as a higher coefficient indicates a higher change of the pollution indicator per standard deviation change.

The regression models highlight the significant role of the aquifer media in explaining the variability of multiple indicators. Sandy aquifers tend to have higher nitrate concentrations but have also shown more desirable concentration changes and rates of change over the study period. The land use variables (CropLU, MeadowLU, GreenLU) exhibit varying influences across the response indicators. Larger crop areas correlate with higher concentrations and less desirable rate changes in 2020 (I2 and I5~~Lo~~), while more forested and green areas are associated with more favorable concentration changes (I3)~~ and rates of change (I4CP and I5CP)~~.

The models explain only 18 to 46% of the variance in the indicator values, as indicated by the R^2 coefficients.

Table 5~~5~~. Coefficients of the variables used in the multiple linear regressions. Only statistically significant coefficients (p-value < 0.05) are shown. ~~CP: change point method for slope. Lo: Lowess method for slope.~~ NR : not relevant.

	Cc 2002 (I1)	**Cc 2020 (I2)**	**Difference cc (I3)**	**Rate of change 2002 (I4)**	**Rate of change 2020 (I5)**	**Difference rate of change (I6)**
Constant	+ 37,7	+ 36,5	-	-	- 0,49	-
CropLU	-	+ 5,3	-	-	+ 0,66	-
MeadowLU	-	+ 4,0	-	-	-	-
MeadowReg	NR	-	-	-	-	-
BuiltLU	-	-	-	-	+ 0,62	-
GreenLU	-	-	- 3,9	-	-	-
Aquifer	+ 6,2	+ 5,4	- 4,8	-	- 0,38	-
Depth	-	+ 5,6	-	-	-	-
R^2	0,26	0,40	0,46	-	0,31	-

4. Discussion

4.1 Evolution of the nitrate concentrations ~~Status and temporal evolution of the nitrate concentrations~~

Since the implementation of the PDGA, the average nitrate concentration across ~~the various~~monitoring sites ~~studied~~ has shown ~~remained~~ relative~~ly~~ stability~~le~~ (Table 4). However, this ~~stability masks~~overall trend conceals ~~underlying~~ variations: 53% of the sites ~~have experienced~~ recorded an increase in nitrate levels, while~~and~~ 47% ~~have seen~~ experienced a decrease. ~~Notably~~ ~~Notably, the~~The most significant reduction ~~in nitrate concentrations between 2002 and 2020~~ was observed in the Brusselian sands, while concentrations increased at most sites in the Geer basin chalks (~~as detailed in Table 3 and~~ Figure 4~~Figure 5~~). These findings are consistent with those reported by ~~, which aligns with~~~~confirms results from~~ (SPW - DEE - Direction des Eaux souterraines, 2024)~~SPW, (2020), supporting the reliability of our dataset.~~. The observed disparity in nitrate trends between these regions reflects differences in the hydrogeological characteristics of the aquifers. The Brusselian sands, characterized by higher permeability and lower nitrogen storage capacity, may exhibit shorter response times to changes in surface nitrogen loading, resulting in more immediate declines in nitrate concentrations. Conversely, the Geer basin chalks, with greater capacity for nitrogen storage and slower groundwater flow, may exhibit a delayed response to reduced nitrogen inputs. This lag effect suggests that, despite reductions in nitrogen loading (SPW, 2022), improvements in nitrate levels in the Geer basin chalks may only become evident in the coming years, provided that nitrogen management measures continue to be implemented~~This lag effect suggests that improvements in nitrate levels in the Geer basin chalks might only become apparent in subsequent years, contingent upon sustained reductions in nitrogen loading and continued implementation of nitrogen management measures.~~ As Liu et al. (2024) emphasize, addressing the delayed response caused by lag effects and nitrogen legacy requires the implementation of sustained, long-term strategies.

~~On an encouraging note,~~Encouragingly, the nitrate concentration trends show improvement, with lower ~~the~~ rates of change ~~in nitrate concentrations is slightly lower~~ in 2020 (I5) ~~than it is in~~compared to 2002 (I4)~~.~~, and a drop in the ~~(Table 4), which is also confirmed by the multivariate regression models (Table 5). The~~ maximum rate of increase from 2,7 mg/l/year in 2002~~change~~ ~~in nitrate concentrations has shown a downward trend from 2002~~ to only 0,6 mg/l/year in 2020 (Table 4~~Table 4~~). This suggests~~indicates~~ that, although nitrate levels continue to rise at~~in~~ some sites, the pace has slowed~~the rate of increase is slower than it used to be. This trend suggests a gradual improvement~~, indicating a ~~and could be seen as a~~ positive outcome of the measures implemented under the PDGA.

4.2 Identification of the controlling factors

Our analysis confirms the anticipated relationship between agricultural land use and nitrate contamination, with~~, as~~ nitrate concentrations were positively correlated with cropland area (Figure 5 and Table 5, I1 and I2) (Gurdak & Qi, 2012; Wick et al., 2012). However, while this correlation was expected for 2002 (I1, prior to the full implementation of the PGDA), its persistence in 2020 (I2) is concerning. Nearly two decades after the introduction of the PGDA, which was designed to reduce nitrogen inputs and nitrate leaching, the correlation between cropland and high nitrate concentrations suggests limited effectiveness of the measures in altering the relationship between land use and groundwater quality. This lack of decoupling highlights the challenge of mitigating "legacy nitrogen" effects, whereby nitrate accumulated in soils and aquifers from past agricultural practices continues to leach into groundwater long after inputs have been reduced (Basu et al., 2022; Van Meter et al., 2016).

Additionally, one might have expected the croplands to be negatively related to the rate of nitrate concentration change (I5 in Table 5) due to the implementation of nitrogen management measures under the PGDA. Contrary to this expectation, the observed positive relationship ~~-~~suggests that nitrate accumulation in groundwater is ongoing. This may reflect the combined effects of legacy nitrogen and potentially insufficient compliance or enforcement of PGDA measures in some regions (Ascott et al., 2017; Hansen et al., 2012).

These findings emphasize the importance of considering both historical nitrogen loads and ongoing agricultural practices. Efforts to enhance the effectiveness of nitrate reduction policies should consider the incorporation of measures to accelerate the recovery of aquifers, such as the promotion of deep-rooted crops that reduce leaching. Enhanced monitoring and stricter enforcement of fertilizer application limits may also help to mitigate further contamination. Furthermore, given the evident lag effects, long-term policy evaluations should account for the temporal dynamics of nitrate transfer and accumulation within aquifers.

Forested and green areas exhibit a negative association with nitrate pollution, showing a lesser contribution to nitrate leaching. This aligns with findings by Zhang et al. (2013) and (Cameron et al., 2013), who noted that forests act as natural buffers by promoting nitrogen uptake and reducing runoff. Moreover, the data suggests that nitrate contamination evolution (I3) was slightly better in more forested zones (Table 5). This implies that PGDA measures might currently be more effective in areas with lower inherent vulnerability. Interestingly, while reductions in meadow area over time correlate with lower nitrate concentrations, meadow area itself does not show a significant impact on nitrate levels. This may result from the failure to differentiate between pasture—typically associated with high nitrate leaching—and other types of meadows (Sacchi et al., 2013). Addressing this distinction in future assessments could enhance the specificity of land-use-related policies.

~~Our analysis aligns with the expected impacts of agricultural and forested land use on nitrate contamination. Although potato cropland do not show a greater effect compared cropland overall, both variables are linked with higher nitrate concentrations as well as higher rates of change (I1, I2, I4 and I5 in Table 5). Forested and green areas show a negative association with nitrate pollution, hence less contributing to nitrate leaching as showed by other authors (Zhang 2012, Cameron 2013). Interestingly, while a reduction in meadow area over time correlates with lower nitrate concentrations, meadow area itself does not show any impact on nitrate levels, possibly because we did not differentiate between pasture, typically associated with high nitrate leaching, and other meadow types (Sacchi et al. 2013).~~

No significant negative effect was observed from built infrastructures, which~~t~~ contrasts with the findings of Mattern et al. (2009) who identified residential land as having negative influence on nitrate concentrations in the Brusselian sands. This discrepancy may be due to our study's focus on semi-rural areas, where built-up areas are limited (under 8% of total land use,

Table 3) and the associated sewage pressure is likely low. Expanding monitoring to include more urbanized areas could clarify the influence of residential land use on nitrate trends.~~This could be explained by our study's focus on semi-rural areas, with built area not exceeding 8% (Table 3), hence low sewage pressure is expected.~~

Aquifer type emerges as a critical predictor of nitrate vulnerability, with sandy aquifers showing higher contamination levels than chalk aquifers. This suggests that policies should prioritize sandy aquifer regions, particularly those under cropland, for targeted measures like stricter nitrogen application limits and buffer zones. However, aquifer type also predicts changes over time (I3), with a general decrease in nitrate concentrations in the Brusselian sands contrasting with increases in the Geer basin chalks (Figure 5). This finding underscores the importance of tailoring strategies to the hydrogeological characteristics of aquifers. For example, in chalk aquifers, where response times to management measures are longer due to nitrogen storage and slower groundwater flow, additional interventions may be needed to address legacy nitrogen. ~~Aquifer type emerges as a key predictor, revealing varying vulnerability among groundwater bodies, with higher contamination in the sandy aquifers compared to the chalk aquifers.~~

Depth also plays a significant role, with shallow groundwater points showing more improvement in nitrate concentrations compared to deeper ones. This delayed response in deeper groundwater bodies to the PGDA measures reflects a time lag in nitrate transfer through the vadose zone. Supporting evidence from (Masetti et al., 2008) in Northern Italy highlights how higher annual precipitation and recharge rates can shorten lag times and improve nitrate trends.~~Looking at changes over time, aquifer type also appears as a significant predictor of nitrate concentration changes (I3), with as already highlighted in section 4.1, a general decrease in the Brusselian sands, but a general increase in the Geer basin chalks (Figure 5). Depth is also significantly correlated to I3, with shallow groundwater points showing more improvement than deeper ones. This suggests a delayed response in deeper groundwater bodies to the PGDA measures. Higher annual precipitation, indicative of increased recharge, could shorten this lag (Figure 5). Such a relationship between groundwater recharge and nitrate concentrations was also found in Northern Italy's groundwater (Masetti et al., 2008)(Masetti et al., 2008)~~

The correlation between graveyards and less favorable nitrate trends warrants careful consideration, as sparse distribution of the graveyards may obscure localized effects. Targeted monitoring near such potential point sources could help identify specific mitigation needs.

~~There is slight indication that nitrate contamination evolution (I3) was better in more forested zones (Table 5), which suggests that PGDA measures might currently be more effective in lesser vulnerable areas. Graveyards correlate with less favorable nitrate trends, though results for the point pollution variables should be analysed with care, as their sparse distribution (Table 3) may obscure more localized effects.~~

~~(Wick et al., 2012)Gurdak & Qi (2012)~~The low predictive power of the multivariate models (R^2 = 18–46%) reflects the complexity of groundwater systems and the factors influencing nitrate concentrations Groundwater systems are influenced by numerous natural and anthropogenic elements, some of which have not been fully captured due to data limitations or to unaccounted sources of nitrate leachate (Masetti et al., 2008). This complexity is compounded by non-linear relationships and time lags between surface changes and groundwater response, which our linear regression models may not fully capture (Wick et al., 2012). Furthermore, the limited number of groundwater data points (36) restricts the model's ability to account for the full variability of the system (Ishwaran, 2007).

To mitigate multicollinearity, variables with high Variance Inflation Factor (VIF) values were removed from the multivariate regression, which likely improved the stability of the coefficient estimates but could also have led to the exclusion of significant predictors. Rainfall was removed as it exhibited highest collinearity, as expected, given its regional variability, which also applies to other variables such as aquifer type, crop land use, and depth. Potato cropland was quite obviously highly collinear with cropland and therefore excluded. Interestingly, land use variables were not as collinear as anticipated, which can be explained ~~-~~by the fact that built infrastructure only accounts for buildings rather than the total urban area, and that the land use variables (crop + forest + built) thus not encompass the entire area. The methodological trade-off of removing potential controlling factors to avoid collinearity further contributes to the relatively low model performance (Ishwaran, 2007).

~~**4.3 Indicators of nitrate pollution trend**~~

~~**We provide six nitrate pollution indicators useful to provide a detailed picture of nitrate pollution state and trends over a given period. However, when using these indicators, it is necessary to keep in mind certain limitations. The indicator on the difference in nitrate concentration (I3) may overlook nuances, such as temporary spikes, short-term fluctuations, or declines within the period. The indicators I4 and I5, the slopes at the beginning and end of the study period reflect only the trend at a specific point in time and may not be representative of longer-term patterns.**~~

~~**Finally, the indicator I6, the difference in slopes, is more abstract and might be harder to interpret. It also assumes that the slopes are linear and may not capture non-linear changes.**~~

~~**It is also clear that the value of these indicators depends on the quality and completeness of the underlying data.**~~

~~**The indicators of rate of change (I4, I5 and I6) are not very sensitive to the chosen method for defining local slope, as evidenced by the similar statistics (Table 4) and correlation values (Figure 6Error! Reference source not found.). This indicates a robustness of these indicators and increases confidence in their values.**~~

4.3~~4~~ ~~Database and i~~Challenges in defining~~nfluence zone~~ the pollution indicators and the independent variables

We provide six nitrate pollution indicators that collectively capture nitrate pollution state and trends over time. However, the use of these indicators, comes with certain limitations. The indicator on the difference in nitrate concentration (I3) may miss short-term fluctuations, while indicators I4 and I5, representing the slopes at the beginning and end of the study period, may not be representative of the longer-term trend. The indicator I6, the difference in slopes, is quite abstract and harder to interpret. The reliability of these indicators depends on the quality and completeness of the underlying data, as well as on the methods used to compute them. The indicators of rate of change (I4, I5 and I6) demonstrate robustness, as they are not very sensitive to the chosen method for defining local slope, as evidenced by the similar statistics (see supplementary materials - Table S1), which increases confidence in their values.

The precise delineation of the influence zones is essential for effectively characterizing and quantifying the independent variables representing the ~~factors~~ factors ~~that potentially impact inherent vulnerability to pollution and identifying possible pollution sources~~affecting nitrate pollution, as highlighted by Nobre et al. (2007) and Mattern et al. (2009). However, ~~the task of accurately~~ defining these zones is fraught with challenges, ~~including the intricate nature of~~due to complex subsurface geology, the ever-changing dynamics of groundwater flow, and constraints related to data availability. In our study, we ~~had to resort to~~used methodological simplifications to delineate these zones, acknowledging ~~that this approach, while the best feasible under the circumstances, does introduce a certain degree of limitation to our analysis.~~the resulting approximation. Another pitfall ~~resides in the fact that~~is the spatial overlap of the influence zones ~~among~~ of some monitoring points ~~were overlapping spatially~~. This overlap implies that the monitoring points are not entirely independent, leading to similar values of ~~ities in~~ potential explanatory variables for these points. This dependency in our dataset ~~Consequently, this~~ could have influenced the ~~identification of factors affecting nitrate pollution~~data analysis.

~~Our study's~~A key strength of our study lies in the comprehensive computation of a broad array of ~~potential explanatory~~independent variables ~~that could influence observed pollution levels and their changes~~representing potential controlling factors. However, ~~this strength is counterbalanced by certain~~ some~~data availability-driven~~ choices~~assumptions~~ made to characterize these variables, driven by data-availability limitations,~~,~~ may have introduced additional noise or uncertainty to the data analysis. ~~which also represent a potential weakness in our work. For instance, in considering~~For instance, we used the variable 'Depth Structure,' rather than the more precise variable ‘Groundwater table depth’, unavailable at the necessary spatial scale, to account for the travel time in the saturated zone. Similarly, ‘Rainfall’ was used as a proxy for ‘Precipitation surplus’, also unavailable at the necessary spatial scale, to approximate net recharge. These substitutions, while necessary, can affect that accuracy of our analysis, and mainly the lack of more significant relationships between the independent variable and the nitrate pollution indicators. Furthermore, some potentially important controlling factors, such as agricultural practices, mitigation measures, livestock density, and manure and fertilizer storage practices, could not be included in our analysis due to the lack of regionally available proxies for these variables.~~our analysis focuses on the bottom of the structure, thereby incorporating the travel time in the saturated zone whose thickness is changing. This approach contrasts with methods like DRASTIC, where the assessment of vulnerability is based on conditions at the top of the aquifer. Such a difference in methodology could lead to variations in interpreting the vulnerability.~~

~~**Regarding the 'Recharge' variable our analysis exclusively considered precipitation as a contributing factor, neglecting the potential effects of evapotranspiration and the complex processes that control the actual recharge. Including evapotranspiration might have offered more nuanced insights into the recharge process and its impact on pollution levels.**~~

4.5 Perspectives

The findings of our current study open several pathways for future research. A primary direction, contingent to the data-availability~~ of additional data~~, is the enhancement of independent variables representing controlling factors.~~, would involve conducting multivariate linear regressions with different subsets of variables to gain a more nuanced understanding of the factors driving nitrate pollution.~~ Future work should ~~R~~refine existing variables by distinguishing between more specific crop types and rotations, as well as differentiating grazed from non-grazed meadows. ~~ing the selection of predictors, such as a more detailed classification of land use that differentiates between meadow and pasture, could enhance the precision of our analysis.~~ Additionally, incorporating data on fertilizer~~ the~~ application rates~~ of nitrogen fertilizer~~, livestock densities, nitrogen surplus estimations, in-situ leachable nitrogen measurements, the state of sewage systems, ~~–~~actual groundwater table depth and precipitation surplus, and concentrations of other pollutants, representative of nitrate-generating activities, would help increase the representation of all factors potentially influencing nitrate contamination. However, these advancements depend heavily on the availability of comprehensive and open datasets.~~potentially leachable nitrates, or metrics derived from the gross nitrogen balance, as done by, would be valuable.~~

~~Exploring variables that represent the concentration and trends of other pollutants, like specific pesticides and pharmaceuticals, could provide insights into the environmental impact of agricultural and urban activities. Improving the estimation of water table depth is another aspect that warrants attention, given its relevance in understanding groundwater dynamics.~~

Our data-driven approach could be combined with complementary methodologies. ~~The integration of ~~~~G~~groundwater dating and~~ outputs from~~ chemical and isotopic analyses (Böhlke & Denver, 1995; Christiaens et al., 2023; Mattern et al., 2011; Vanclooster et al., 2020) could offer critical temporal perspectives on the source and evolution of groundwater contamination.

Numerical process-oriented models could allow to better take into account the mechanisms known to affect nitrate leaching to groundwater.

Expanding the groundwater nitrate concentration dataset by incorporating additional monitoring points would strengthen the representativeness and predictive power of models.

With a larger dataset and better quality independent variables, more advanced, non-linear machine learning techniques could be employed to uncover new insights and capture the complexities of nitrate contamination. These approaches could provide a deeper understanding of the underlying processes and hence help guide future best management practices. However, expanding the dataset and collecting better-quality independent variables is a challenging task due to the limited availability of long-term and spatially explicit data, and the complexities associated with data collection across multiple stakeholders.

Addressing these challenges will be crucial for the development of more robust and advanced analytical approaches in future studies.

~~**Expanding the network of monitoring points, particularly in vulnerable regions, would significantly enhance the representativeness and reliability of our study.**~~

~~**Moreover, experimenting with alternative predictive models, allowing to capture non-linear effects could lead to more robust findings. However, it's important to recognize that more sophisticated models typically require larger datasets for effective training. Thus, expanding our dataset is a crucial step for such advanced modelling. This expansion is not a trivial task due to the lack of monitoring points with long records. Moreover, due to the multiplicity of actors owning the records, data collection and pretreatment is time consuming. Addressing these challenges is essential for the successful implementation of more complex analytical models in future studies.**~~

5. Conclusions

~~**Our study leverages two decades of groundwater monitoring data to evaluate nitrate contamination trends and identify controlling factors in nitrate-vulnerable zones following the implementation of the sustainable nitrogen management program (PGDA) in 2002 in Wallonia. The spatially extensive and long-term dataset enabled a comprehensive analysis of both spatial variability and temporal dynamics.**~~

This study contributes to global scientific knowledge by investigating the long-term spatial and temporal dynamics of nitrate contamination in groundwater, linking these trends to land use, hydrogeological conditions, and nitrogen management policies. By leveraging two decades of data, it provides insights into persistent contamination sources, such as cropland, and highlights the critical role of aquifer characteristics and depth in mediating nitrate responses. To enhance policy effectiveness, our results suggest the relevance of targeted interventions by prioritizing ~~sandy~~vulnerable aquifers and cropland-dominated zones ~~zones with stricter nitrogen application limits and conservation measures~~. They also highlight the need to address time lags and nitrogen legacy effects by sustaining and intensifying measures over extended periods. ~~Our study also underscores the need to sustain and perhaps expand the monitoring systems to better capture spatial variability and localized impacts of land use, and long-time changes.~~

The limited predictive power of the regression models reflects the inherent complexity of groundwater nitrate contamination and the difficulty of adequately representing its controlling factors. Constraints such as data availability and required simplifications in defining independent variables ~~further~~ limit model reliability. ~~These findings~~This underlines the need for more detailed and accessible datasets that better capture the controlling factors~~o deepen our understanding of contamination~~

~~dynamics~~, and the need to sustain and perhaps expand the monitoring systems to better capture spatial variability and localized impacts of land use, and long-term changes. This will f~~. Enhanced data availability would f~~acilitate the development of advanced machine-learning ~~modeling~~modelling approaches ~~, including non-linear machine learning techniques,~~ capable of better addressing the multifaceted and nonlinear processes governing nitrate pollution.

~~To enhance policy effectiveness, our results suggest the relevance of targeted interventions by prioritizing sandy aquifers and cropland-dominated zones with stricter nitrogen application limits and conservation measures. They also highlight the need to address time lags and nitrogen legacy effects by sustaining and intensifying measures over extended periods. Our study also underscores the need to sustain and perhaps expand the monitoring systems to better capture spatial variability and localized impacts of land use, and long-time changes.~~

Data availability

The data used in this study are not publicly available. For data access, please contact the corresponding author.~~Particularly, the temporal focus revealed a decrease in nitrate concentration rate of changes in 2020 compared to 2002, while mean concentrations remained almost stable overall. We foundshowed that all monitored sites had nitrate concentrations below the 50 mg/l threshold in 2020, but that the average concentration across the studied sites has remained relatively stable, although with significant variations across sites. We also showed that the average rate of change in nitrate concentration was decreasing in 2020, and a decelerating increase (or accelerating decrease) in the rate of change in nitrate concentrations compared to 2002.~~

~~Our findings aligns with previously observed patterns. A temporal decrease in nitrate levels in the Brusselian sands, where concentrations are typically higher, and an increase in the Geer basin chalks, where they are generally lower. Improvements were more pronounced in shallow groundwater points than in deeper ones, likely due to varying time lags between the implementation of regulatory measures and observable changes in groundwater quality. Locations with lower conductivity may require additional time to exhibit reductions in nitrate concentrations, underscoring the importance of long-term strategies and sustained efforts in managing and monitoring groundwater quality. The study also finds a delayed response of deeper groundwater bodies to nitrogen regulation measures. We also observed a persistent negative influence of agricultural land cover on nitrate concentrations over the entire study period, even on current nitrate concentration trends. This highlights the ongoing necessity of implementing and maintaining best management practices (BMPs) in agricultural areas to mitigate contamination risks effectively.~~

~~The limited predictive power of the regression models reflect the complexity of groundwater nitrate contamination and the challenges of adequately representing controlling factors. Limitations related to data availability and the simplifications made in defining independent variables further hinder the development of reliable models. These findings emphasize the need for more detailed and open-access datasets to improve our understanding of contamination dynamics. Such improvements would enable the development of more advanced models, including non-linear machine learning techniques, which could better account for the multifaceted nature of groundwater nitrate pollution.~~

~~Our study provides insights into the current state and temporal evolution of nitrate concentrations in groundwater since the implementation of the nitrogen regulation (PGDA) in the agricultural crop production sector in Wallonia, Belgium. We showed that all monitored sites had nitrate concentrations below the 50mg/l threshold in 2020, but that the average concentration across the studied sites has remained relatively stable, although with significant variations across sites. We also showed that the average rate of change in nitrate concentration was decreasing in 2020, and a decelerating increase (or accelerating decrease) in the rate of change in nitrate concentrations compared to 2002.~~

~~Overall, the results post-PGDA implementation confirm the complex interplay of factors influencing nitrate pollution and trend, with land use and aquifer characteristics emerging as significant determinants. Positive relationships were found between crop and potato cultivation, meadow areas, and nitrate levels and change, highlighting the significant impact of farming activities. The study also finds a delayed response of deeper groundwater bodies to nitrogen regulation measures. Notable reduction in nitrate levels was also observed, especially in the Brusselian sands likely due to the higher conductivity and faster renewal rate of that aquifer. These results show encouraging sign about the effectiveness of the PGDA and suggest that longer time lag between the implementation of regulatory measures and observable changes in groundwater quality might explain that the other locations with lower conductivity may only exhibit decreases in nitrate concentrations in future years. They underscore the importance of long-term approaches and sustained efforts in managing and monitoring groundwater quality.~~

~~The multivariate regression models were only able to explain between 18 to 40% of the variability of the indicators, which suggests that while the identified factors are influential, other unaccounted variables or inherent complexities in nitrate pollution dynamics are at play.~~

Author contribution

EV and AA conceptualized the study. EV performed the data curation and formal analysis. LC provided the data. MV did the funding acquisition and project administration. AA and MV carried out the supervision. AA and EV wrote the original manuscript.

Competing interests

One of the coauthors is a member of the editorial board of Hydrology and Earth System Sciences.

Ackowledgments

The authors thank Olivier Bastin, Romain Millecamps and Julien Minet with whom they collaborated during the EGERIEC project, in the context of which this research was conducted. The authors also thank the SPGE for financing the EGERIEC project. The authors thank the laboratory of the Société Wallonne des Eaux, and more specifically Sébastien Ronkart, for providing the water quality data. The authors thank OpenAI for providing ChatGPT, which was used to assist in the writing process.

References

Aller, L., Benett, T., Lehr, J. H., & Petty, R. J. (1985). *DRASTIC: A Standardized System for Evaluating Ground Water Pollution potential using hydrogeologic settings*.

Ascott, M. J., Gooddy, D. C., Wang, L., Stuart, M. E., Lewis, M. A., Ward, R. S., & Binley, A. M. (2017). Global patterns of nitrate storage in the vadose zone. *Nature Communications 2017 8:1*, *8*(1), 1–7. https://doi.org/10.1038/s41467-017-01321-w

Bah, B., Vandenberghe, C., & Colinet, G. (2015). *Méthodologie de diagnostic environnemental autour de captages d'eau potabilisable sensibles qualitativement du point de vue nitrate & application à six sites de la société wallonne des eaux. Rapport d'activités final. Partie GRENERA. Convention S.P.G.E. - SWDE*.

Basu, N., Van Meter, K., Byrnes, D., Cappellen, P. Van, Brouwer, R., Jacobsen, B., Jarsjö, J., Rudolph, D., Cunha, M., Nelson, N., Bhattacharya, R., Destouni, G., & Olsen, S. B. (2022). Managing nitrogen legacies to accelerate water quality improvement. *Nature Geoscience*, *15*, 97–105. https://doi.org/10.1038/s41561-021-00889-9

Batlle Aguilar, J., Orban, P., Dassargues, A., & Brouyère, S. (2007). Identification of groundwater quality trends in a chalk aquifer threatened by intensive agriculture in Belgium. *Hydrogeology Journal*, *15*(8), 1615–1627. https://doi.org/10.1007/s10040-007-0204-y

Böhlke, J. K., & Denver, J. M. (1995). Combined Use of Groundwater Dating, Chemical, and Isotopic Analyses to Resolve the History and Fate of Nitrate Contamination in Two Agricultural Watersheds, Atlantic Coastal Plain, Maryland. *Water Resources Research*, *31*(9), 2319–2339. https://doi.org/10.1029/95WR01584

Bouchard, D. C., Williams, M. K., & Surampalli, R. Y. (1992). Nitrate Contamination of Groundwater: Sources and Potential Health Effects. *Journal - American Water Works Association*, *84*(9), 85–90. https://doi.org/10.1002/J.1551-8833.1992.TB07430.X

Cameron, K. C., Di, H. J., & Moir, J. L. (2013). Nitrogen losses from the soil/plant system: a review. *Annals of Applied Biology*, *162*(2), 145–173. https://doi.org/10.1111/aab.12014

Christiaens, L., Orban, P., Brouyère, S., & Goderniaux, · Pascal. (2023). Tracking the sources and fate of nitrate pollution by combining hydrochemical and isotopic data with a statistical approach. *Hydrogeology Journal*, *31*, 1271–1289. https://doi.org/10.1007/s10040-023-02646-1

Cleveland, W. S. (1979). Robust locally weighted regression and smoothing scatterplots. *Journal of the American Statistical Association*, *74*(368), 829–836. https://doi.org/10.1080/01621459.1979.10481038

Comly, H. H. (1945). Cyanosis in Infants caused by Nitrates in Well Water. *JAMA*, *129*, 112–116. https://doi.org/10.1001/JAMA.1945.02860360014004

de Vries, W., Posch, M., Simpson, D., de Leeuw, F. A. A. M., van Grinsven, H. J. M., Schulte-Uebbing, L. F., Sutton, M. A., & Ros, G. H. (2024). Trends and geographic variation in adverse impacts of nitrogen use in Europe on human health, climate, and ecosystems: A review. *Earth-Science Reviews*, *253*, 104789. https://doi.org/10.1016/J.EARSCIREV.2024.104789

Ferguson, R. (2015). Groundwater Quality and Nitrogen Use Efficiency in Nebraska's Central Platte River Valley. *Journal of Environmental Quality*, *44 2*, 449–459. https://doi.org/10.2134/jeq2014.02.0085

Frollini, E., Preziosi, E., Calace, N., Guerra, M., Guyennon, N., Marcaccio, M., Menichetti, S., Romano, E., & Ghergo, S. (2021). Groundwater quality trend and trend reversal assessment in the European Water Framework Directive context: an example with nitrates in Italy. *Environmental Science and Pollution Research*, *28*(17), 22092–22104. https://doi.org/10.1007/S11356-020-11998-0/FIGURES/5

Grath, J., Scheidleder, A., Uhlig, S., Weber, K., Kralik, M., Keimel, T., & Gruber, D. (2001). *The EU Water Framework Directive: Statistical aspects of the identification of groundwater pollution trends, and aggregation of monitoring results. Final Report.*

Grizzetti, B., Bouraouri, F., Billen, G., Van Grinsven, H., Cardoso, A. C., Thieu, V., Garnier, J., Curtis, C., Howarth, R., & Johnes, P. (2011). Nitrogen as a threat to European water quality. In M. A. Sutton, C. M. Howard, J. W. Erisman, G. Billen, A. Bleeker, P. Grennfelt, H. Van Grinsven, & B. Grizzetti (Eds.), *The European Nitrogen Assessment* (1st ed., pp. 379–404). Cambridge University Press. https://centaur.reading.ac.uk/20869/

Gurdak, J., & Qi, S. L. (2012). Vulnerability of recently recharged groundwater in principal [corrected] aquifers of the United States to nitrate contamination. *Environmental Science & Technology*, *46 11*, 6004–6012. https://doi.org/10.1021/es300688b

Hansen, B., Dalgaard, T., Thorling, L., Sørensen, B., & Erlandsen, M. (2012). Regional analysis of groundwater nitrate concentrations and trends in Denmark in regard to agricultural influence. *Biogeosciences*, *9*(8), 3277–3286. https://doi.org/10.5194/BG-9-3277-2012

Hansen, B., Thorling, L., Schullehner, J., Termansen, M., & Dalgaard, T. (2017). Groundwater nitrate response to sustainable nitrogen management. *Scientific Reports 2017 7:1*, *7*(1), 1–12. https://doi.org/10.1038/s41598-017-07147-2

He, S., Wu, J., Wang, D., & He, X. (2022). Predictive modeling of groundwater nitrate pollution and evaluating its main impact factors using random forest. *Chemosphere*, *290*, 133388. https://doi.org/10.1016/J.CHEMOSPHERE.2021.133388

Hirsch, R. M., Alexander, R. B., & Smith, R. A. (1991). Selection of methods for the detection and estimation of trends in water quality. *Water Resources Research*, *27*(5), 803–813. https://doi.org/10.1029/91WR00259

Hornung, M. (1999). The Role of Nitrates in the Eutrophication and Acidification of Surface Waters. In W. S. Wilson, A. S. Ball, & R. H. Hinton (Eds.), *Managing risks of Nitrates to Humans and the Environment* (pp. 155–174). Woodhead Publishing.

Ishwaran, H. (2007). Variable importance in binary regression trees and forests. *Electronic Journal of Statistics*, *1*, 519–537. https://doi.org/10.1214/07-EJS039

J. R. Williams, C. A. Jones, & P. T. Dyke. (1984). A Modeling Approach to Determining the Relationship Between Erosion and Soil Productivity. *Transactions of the ASAE*, *27*(1), 0129–0144. https://doi.org/10.13031/2013.32748

James, G., Witten, D., Hastie, T., & Tibshirani, R. (2013). *An Introduction to Statistical Learning*. Springer. https://doi.org/10.1007/978-1-4614-7138-7

Kendall, M. G. (1938). A New Measure of Rank Correlation. *Biometrika*, *30*, 81–93.

Kyte, E., Cey, E., Hrapovic, L., & Hao, X. (2023). Nitrate in shallow groundwater after more than four decades of manure application. *Journal of Contaminant Hydrology*, *256*, 104200. https://doi.org/10.1016/j.jconhyd.2023.104200

Lee, H. W., Bhang, K. J., & Park, S. S. (2010). Effective visualization for the spatiotemporal trend analysis of the water quality in the Nakdong River of Korea. *Ecological Informatics*, *5*(4), 281–292. https://doi.org/10.1016/J.ECOINF.2010.05.004

Liu, X., Beusen, A. H. W., van Grinsven, H. J. M., Wang, J., van Hoek, W. J., Ran, X., Mogollón, J. M., & Bouwman, A. F. (2024). Impact of groundwater nitrogen legacy on water quality. *Nature Sustainability 2024 7:7*, *7*(7), 891–900. https://doi.org/10.1038/s41893-024-01369-9

Mansfield, E. R., & Helms, B. P. (1982). Detecting Multicollinearity. *The American Statistician*, *36*(3), 158–160.

Masetti, M., Poli, S., Sterlacchini, S., Beretta, G. P., & Facchi, A. (2008). Spatial and statistical assessment of factors influencing nitrate contamination in groundwater. *Journal of Environmental Management*, *86*(1), 272–281. https://doi.org/10.1016/J.JENVMAN.2006.12.023

Mattern, S., Fasbender, D., & Vanclooster, M. (2009). Discriminating sources of nitrate pollution in an unconfined sandy aquifer. *Journal of Hydrology*, *376*(1–2), 275–284. https://doi.org/10.1016/J.JHYDROL.2009.07.039

Mattern, S., Sebilob, M., & Vancloostera, M. (2011). Identification of the nitrate contamination sources of the Brusselian sands groundwater body (Belgium) using a dual-isotope approach. *Isotopes in Environmental and Health Studies*, *47*(3), 297–315. https://doi.org/10.1080/10256016.2011.604127

Mattern, S., & Vanclooster, M. (2010). Estimating travel time of recharge water through a deep vadose zone using a transfer function model. *Environmental Fluid Mechanics*, *10*(1), 121–135. https://doi.org/10.1007/S10652-009-9148-1/METRICS

Nobre, R. C. M., Rotunno Filho, O. C., Mansur, W. J., Nobre, M. M. M., & Cosenza, C. A. N. (2007). Groundwater vulnerability and risk mapping using GIS, modeling and a fuzzy logic tool. *Journal of Contaminant Hydrology*, *94*(3–4), 277–292. https://doi.org/10.1016/J.JCONHYD.2007.07.008

Picron, P., Copus, A., & Wouez, D. (2017). WATER PROTECTION AGAINST NITRATE AND PESTICIDES FROM AGRICULTURAL SOURCES – HISTORY OF ACTIONS TAKEN AT FARM SCALE IN WALLONIA (BELGIUM) FOR 15 YEARS. *Farm Machinery and Processes Management in Sustainable Agriculture, IX International Scientific Symposium*, 306–311. https://doi.org/10.24326/fmpmsa.2017.55

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 Research*, *42*(16), 4215–4232. https://doi.org/10.1016/J.WATRES.2008.07.020

Rodriguez-Galiano, V., Mendes, M. P., Garcia-Soldado, M. J., Chica-Olmo, M., & Ribeiro, L. (2014). Predictive modeling of groundwater nitrate pollution using Random Forest and multisource variables related to intrinsic and specific vulnerability: A case study in an agricultural setting (Southern Spain). *Science of The Total Environment*, *476–477*, 189–206. https://doi.org/10.1016/J.SCITOTENV.2014.01.001

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. *Applied Geochemistry*, *34*, 164–180. https://doi.org/10.1016/J.APGEOCHEM.2013.03.008

Schullehner, J., Hansen, B., Thygesen, M., Pedersen, C. B., & Sigsgaard, T. (2018). Nitrate in drinking water and colorectal cancer risk: A nationwide population-based cohort study. *International Journal of Cancer*, *143*(1), 73–79. https://doi.org/10.1002/IJC.31306

Seabold, Skipper, & Josef Perktold. (2010). statsmodels: Econometric and statistical modeling with python. *Proceedings of the 9th Python in Science Conference*.

Shukla, S., & Saxena, A. (2018). Global Status of Nitrate Contamination in Groundwater: Its Occurrence, Health Impacts, and Mitigation Measures. *Handbook of Environmental Materials Management*, 1–21. https://doi.org/10.1007/978-3-319-58538-3_20-1

Sohier, C., & Degré, A. (2010). Modelling the effects of the current policy measures in agriculture: An unique model from field to regional scale in Walloon region of Belgium. *Environmental Science & Policy*, *13*(8), 754–765. https://doi.org/10.1016/J.ENVSCI.2010.08.008

Spalding, R., & Exner, M. (1993). Occurrence of nitrate in groundwater-a review. *Journal of Environmental Quality*, *22*, 392–402. https://doi.org/10.2134/JEQ1993.00472425002200030002X

SPW. (n.d.). *Catalogue des données et services | Géoportail de la Wallonie*. Retrieved June 21, 2023, from https://geoportail.wallonie.be/catalogue-donnees-et-services

SPW. (2006a). *Masse d'eau souterraine RWE030 Craies de la Haine*. http://eau.wallonie.be/fme/RWE030.pdf

SPW. (2006b). *Masse d'eau souterraine RWE051 - Sables du Bruxellien*. http://eau.wallonie.be/fme/RWE051.pdf

SPW. (2010). *Etat des lieux de la masse d'eau RWE053 Sables du Landénien (Est)*. http://eau.wallonie.be/fme/RWE053.pdf

SPW. (2016). *Fiche de caractérisation de la masse d'eau RWM040 « Craies du Geer »*. http://eau.wallonie.be/fme/RWM040.pdf

SPW. (2019). *Amélioration des axes de concentration du ruissellement générés sur base du LIDAR 2015 à la résolution de 1m et des données cours d'eau dérivées - Rapport final.*

SPW. (2020). *Zones de protection des captages d'eau souterraine - État de l'environnement wallon*. http://etat.environnement.wallonie.be/contents/indicatorsheets/EAU%2016.html

SPW. (2022). *Bilan d'azote en agriculture et flux d'azote des sols vers les eaux - État de l'environnement wallon*. http://etat.environnement.wallonie.be/contents/indicatorsheets/SOLS%204.html

SPW - DEE - Direction des Eaux souterraines. (2024). *Etat des nappes et des masses d'eau souterraine de Wallonie. Février 2024. Vingt-deuxième année.*

Strebel, O., Duynisveld, W. H. M., & Böttcher, J. (1989). Nitrate pollution of groundwater in western Europe. *Agriculture, Ecosystems & Environment*, *26*(3–4), 189–214. https://doi.org/10.1016/0167-8809(89)90013-3

Torres-Martínez, J. A., Mora, A., Knappett, P. S. K., Ornelas-Soto, N., & Mahlknecht, J. (2020). Tracking nitrate and sulfate sources in groundwater of an urbanized valley using a multi-tracer approach combined with a Bayesian isotope mixing model. *Water Research*, *182*, 115962. https://doi.org/10.1016/j.watres.2020.115962

Urresti-Estala, B., Gavilán, P. J., Pérez, I. V., & Cantos, F. C. (2016). Assessment of hydrochemical trends in the highly anthropised Guadalhorce River basin (southern Spain) in terms of compliance with the European groundwater directive for 2015. *Environmental Science and Pollution Research*, *23*(16), 15990–16005. https://doi.org/10.1007/S11356-016-6662-9/FIGURES/6

Van Grinsven, H. J. M., Ten Berge, H. F. M., Dalgaard, T., Fraters, B., Durand, P., Hart, A., Hofman, G., Jacobsen, B. H., Lalor, S. T. J., Lesschen, J. P., Osterburg, B., Richards, K. G., Techen, A. K., Vertès, F., Webb, J., & Willems, W. J. (2012). Management, regulation and environmental impacts of nitrogen fertilization in northwestern Europe under the Nitrates Directive; A benchmark study. *Biogeosciences*, *9*(12), 5143–5160. https://doi.org/10.5194/BG-9-5143-2012

Van Grinsven, H. J. M., Tiktak, A., & Rougoor, C. W. (2016). Evaluation of the Dutch implementation of the nitrates directive, the water framework directive and the national emission ceilings directive. *NJAS - Wageningen Journal of Life Sciences*, *78*, 69–84. https://doi.org/10.1016/J.NJAS.2016.03.010

Van Meter, K., Basu, N., Veenstra, J., & Burras, C. (2016). The nitrogen legacy: emerging evidence of nitrogen accumulation in anthropogenic landscapes. *Environmental Research Letters*, *11*. https://doi.org/10.1088/1748-9326/11/3/035014

Vanclooster, M., Petit, S., Bogaert, P., Lietar, A., Vanclooster, M., Petit, S., Bogaert, P., & Lietar, A. (2020). Modelling Nitrate Pollution Vulnerability in the Brussel's Capital Region (Belgium) Using Data-Driven Modelling Approaches. *Journal of Water Resource and Protection*, *12*(5), 416–430. https://doi.org/10.4236/JWARP.2020.125025

Vandenberghe, C. (2016). *Performance de l’indicateur Azote Potentiellement Lessivable pour l’évaluation de la gestion du nitrate en agriculture et du risque pour la qualité de l’eau* [Thesis., ORBi-University of Liège]. https://orbi.uliege.be/handle/2268/201372

Visser, A., Broers, H. P., Van Der Grift, B., & Bierkens, M. F. P. (2007). Demonstrating trend reversal of groundwater quality in relation to time of recharge determined by 3H/3He. *Environmental Pollution*, *148*(3), 797–807. https://doi.org/10.1016/J.ENVPOL.2007.01.027

Wakida, F. T., & Lerner, D. N. (2005). Non-agricultural sources of groundwater nitrate: a review and case study. *Water Research*, *39*(1), 3–16. https://doi.org/10.1016/J.WATRES.2004.07.026

Ward, B. B. (2013). Nitrification. *Encyclopedia of Ecology: Volume 1-4, Second Edition*, *2*, 351–358. https://doi.org/10.1016/B978-0-12-409548-9.00697-7

Ward, M., Jones, R. R., Brender, J., de Kok, T. D., Weyer, P., Nolan, B., Villanueva, C., & van Breda, S. V. (2018). Drinking Water Nitrate and Human Health: An Updated Review. *International Journal of Environmental Research and Public Health*, *15*, null. https://doi.org/10.3390/ijerph15071557

Wick, K., Heumesser, C., & Schmid, E. (2012). Groundwater nitrate contamination: Factors and indicators. *Journal of Environmental Management*, *111*, 178–186. https://doi.org/10.1016/J.JENVMAN.2012.06.030

Zhang, X., Xu, Z., Sun, X., Dong, W., & Ballantine, D. (2013). Nitrate in shallow groundwater in typical agricultural and forest ecosystems in China, 2004–2010. *Journal of Environmental Sciences*, *25*(5), 1007–1014. https://doi.org/10.1016/S1001-0742(12)60139-9