



HAL
open science

Identification of critical effect factors for prediction of spatial and intra-annual variability of shallow groundwater nitrate in agricultural areas

Chunying Wang, Xinliang Wang, Gengchen Zhang, Feifei Zhang, Junfeng Li, Shuai Chen, Sabine Sauvage, J.M. Sánchez-Pérez, Yuping Han, Junguo Liu

► To cite this version:

Chunying Wang, Xinliang Wang, Gengchen Zhang, Feifei Zhang, Junfeng Li, et al.. Identification of critical effect factors for prediction of spatial and intra-annual variability of shallow groundwater nitrate in agricultural areas. *Science of the Total Environment*, 2023, 891, pp.164342. 10.1016/j.scitotenv.2023.164342 . hal-04278549

HAL Id: hal-04278549

<https://hal.science/hal-04278549v1>

Submitted on 13 Nov 2023

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

1 **Prediction of spatial and intra-annual variability of shallow**
2 **groundwater nitrate in agricultural areas**

3 Chunying Wang^{a*}, Xinliang Wang^a, Gengchen Zhang^a, Feifei Zhang^a, Junfeng Li^b, Shuai
4 Chen^a, Sabine Sauvage^c, José-Miguel Sánchez-Pérez^c, Yuping Han^a, Junguo Liu^{a,d}

5 ^a *School of Water Resources, North China University of Water Resources and Electric*
6 *Power, Zhengzhou 450045, PR China;*

7 ^b *College of Water Conservancy and Architecture Engineering, Shihezi University, Shihezi*
8 *832000, PR China;*

9 ^c *Laboratory of Functional Ecology and Environment, University of Toulouse, CNRS, UPS,*
10 *Toulouse INP, ENSAT campus, F-31326 Toulouse, France*

11 ^d *School of Environmental Science and Engineering, Southern University of Science and*
12 *Technology, Shenzhen 518055, PR China;*

13 **ABSTRACT**

14 Shallow groundwater nitrate nitrogen (NO₃⁻-N) concentrations in agricultural areas usually
15 show high spatial and intra-annual variability. It is hard to predict such concentrations due to
16 the complexity of influencing factors (e.g., different forms of N in soil, vadose zone
17 characteristics, and groundwater environmental conditions). Here, a large number of
18 groundwater and soil samples were collected monthly over two years at 14 sites to analyze
19 the soil and groundwater physiochemical properties and the stable isotopes of δ¹⁵N and δ¹⁸O
20 of groundwater NO₃⁻-N in agricultural areas. Based on field observations, a random forest
21 (RF) model was used to predict the groundwater NO₃⁻-N concentrations and reveal the
22 importance of effect factors. The results show that there are large spatiotemporal variations
23 in NO₃⁻-N, δ¹⁵N-NO₃⁻, and δ¹⁸O-NO₃⁻ in groundwater. NO₃⁻-N is the major dominant specie

24 of inorganic N in groundwater, and the groundwater NO_3^- -N concentration in 24% of the
25 samples failed to meet the drinking water standard of the WHO (10 mg L^{-1}). The RF model
26 satisfactorily predicted groundwater NO_3^- -N concentrations with R^2 of 0.92–0.93, RMSE of
27 3.87–4.94, and MAE of 2.10–2.89. Groundwater nitrite and ammonium are the most
28 important factors related to NO_3^- -N removal and production in groundwater. Denitrification
29 and nitrification were further identified by the relationships among $\delta^{15}\text{N-NO}_3^-$, $\delta^{18}\text{O-NO}_3^-$,
30 and NO_3^- -N, and by the ranges of $\delta^{15}\text{N-NO}_3^-$, $\delta^{18}\text{O-NO}_3^-$, temperature, pH, DO, and ORP in
31 groundwater. Soil-soluble organic nitrogen (S-SON) and the depth of groundwater table
32 were identified as vital factors related to N sourcing and leaching. Overall, as a first
33 approach to adopting a RF model for high spatiotemporal-resolution prediction of
34 groundwater NO_3^- -N variations, the findings of this study enable a better understanding of
35 groundwater N pollution in agricultural areas. Optimizing management of irrigation and N
36 inputs is anticipated to reduce S-SON accumulation and mitigate the threat to groundwater
37 quality in agricultural areas.

38

39

40

41 **Keywords:** Groundwater nitrate; Effect factor; Nitrogen transformation; Random forest;
42 Nitrogen and oxygen isotopes

43 **1. Introduction**

44 Groundwater is currently facing the severe challenges of depletion and deteriorating
45 quality due to natural and anthropogenic factors around the world (Liu et al., 2016;
46 Famiglietti and Ferguson, 2021). Among the many groundwater problems, nitrogen (N)
47 pollution has become a global issue due to the use of N fertilizers and manures and elevated
48 atmospheric deposition (Vystavna et al., 2017). Groundwater N pollution reduces N use
49 efficiency and threatens the safety of the water supply in China (Gan et al., 2022; Gao et al.,
50 2022). The nitrate nitrogen (NO_3^- -N) contamination of groundwater is a severe problem
51 threatening the environment and human health, especially in intensively irrigated
52 agricultural areas. Accurately predicting the variations of shallow groundwater NO_3^- -N
53 concentrations at high spatiotemporal resolution is very difficult due to the complex effect
54 factors and processes (Hinkle and Tesoriero, 2014; Biddau et al., 2019; He et al., 2022).
55 Therefore, exploration of the spatiotemporal patterns and any associated effect factors and
56 processes is urgently needed to prevent widespread water-quality issues around the world.

57 Climate variables, soil texture, land use, and human activities are commonly
58 investigated and recognized as the main factors affecting the inter-annual variations of
59 groundwater NO_3^- -N pollution on the large or macroscopic scale (Pennino et al., 2020; El
60 Amri et al., 2022; He et al., 2022). Nonetheless, comprehensive identification of the effect
61 factors and evaluation of their impact on dynamics of groundwater NO_3^- -N has seldomly
62 been done at a high spatial and monthly temporal resolution, especially in irrigated
63 agricultural areas. Climate variables and agricultural management activities, such as

64 precipitation, irrigation, and fertilizer input, cause the high level of variability in the
65 different forms of N content in soil and subsequent leaching to groundwater. Vadose zone
66 thickness (i.e., the depth of the groundwater table) contributes significantly to the variations
67 of groundwater NO_3^- -N concentrations by affecting the amounts of different forms of N
68 leaching and groundwater environmental conditions (Li et al., 2021; Weitzman et al., 2022).
69 Meanwhile, groundwater environmental parameters influence the different forms of N
70 content by impacting transformation processes. Environmental factors related to N
71 transformation in groundwater include temperature, dissolved oxygen (DO),
72 oxidation-reduction potential (ORP), and dissolved organic carbon (DOC). The difficulty in
73 understanding groundwater NO_3^- -N variations stems from the ability to determine the most
74 critical explanatory variables among the many factors. This question is challenging when
75 using common traditional methods, such as correlation analysis, attribution analysis, cluster
76 analysis, principal component analysis, and numerical and distributed hydrological models.

77 Overall, groundwater NO_3^- -N concentrations are controlled by diverse factors that are
78 linked through the complex interaction of multiple processes. These processes involve
79 inorganic and organic N transformation and transport in soil and groundwater (Zhang et al.,
80 2019). Surface soil N leaching to groundwater and the groundwater environment are both
81 affected by characteristics of the vadose zone. The leaching of different forms of N content
82 from the surface soil, as the source of N in groundwater, indirectly affects groundwater
83 NO_3^- -N dynamics through N transformation in groundwater. The N transformation in
84 groundwater is affected by the groundwater environment, and it increases the importance

85 and difficulty of identifying the roles of N transformation in groundwater. In terms of the N
86 in groundwater, the widely reported mineralization, nitrification, and denitrification
87 processes in groundwater can modify different forms of N species balance (Liu et al., 2022;
88 Shen et al., 2023). For instance, the mineralization and nitrification processes in
89 groundwater can produce NO_3^- -N, while denitrification can reduce NO_3^- -N. These
90 processes in groundwater play a key influence in the spatiotemporal variations of NO_3^- -N
91 dynamics in addition to N leaching in soil. Therefore, identifying the main N
92 transformation processes in groundwater and evaluating their importance will further
93 elucidate the mechanisms of groundwater NO_3^- -N pollution in agricultural areas.

94 The complex factors and processes influencing groundwater NO_3^- -N concentrations
95 lead to great difficulty in making predictions at high spatial and temporal resolution. For
96 instance, physically based numerical and distributed hydrological models (e.g., HYDRUS,
97 AgriFlux, and SWAT [Soil & Water Assessment Tool]) need to be coupled with
98 groundwater flow models, such as MODFLOW-MT3D and the model of Lasserre et al.
99 (1999) linked to a GIS (geographic information system), to predict groundwater NO_3^- -N
100 dynamics (Wang et al., 2016; El Amri et al., 2022). The performance of these models
101 basically depends on an adequate understanding of hydrological behaviors and
102 biogeochemical processes and the availability of detailed data on the properties of the
103 vadose zone and groundwater system. These data are usually difficult to measure and
104 collect, resulting in unsatisfactory model performance (Coppola et al., 2005). Machine
105 learning methods, for example, artificial neural networks, multiple logistic regressions,

106 generalized additive models, generalized linear models, support vector regressions, and
107 random forest models, have been widely used to predict the N status in agricultural and
108 natural waste waters (Chlingaryan et al., 2018; Bagherzadeh et al., 2021; He et al., 2022).
109 They were gradually accepted by researchers due to their advantages, such as their high
110 generalization ability and low cost, but most of them have the problem of overfitting
111 (Castrillo and García, 2020). Among these machine learning methods, the random forest
112 model has the advantages of strong resistance to overfitting, no feature selection, and
113 automatic data filling. It has been proven to accurately predict inter-annual groundwater
114 NO_3^- -N concentrations on a large scale (Band et al., 2020; He et al., 2022). However, the
115 random forest model has not yet been applied to predict intra-annual variations of shallow
116 groundwater NO_3^- -N in irrigated agricultural areas.

117 The North China Plain (NCP) is the main crop production area in China, and
118 groundwater resources in the NCP are generally dealing with the issue of N pollution. This
119 study was conducted in a semiarid heavily irrigated agricultural area located in the NCP,
120 that is characterized by a shallow groundwater table depth and, thus, vulnerable to
121 groundwater pollution. Monthly meteorological data were obtained from the Xinxiang
122 weather station, which is located near the study site. Two years' worth of monthly soil
123 physiochemical data and groundwater quality parameters were measured in the field and
124 laboratory. The objectives of this study were to: (1) reveal the spatial and intra-annual
125 variations of inorganic N species and isotopic signature of NO_3^- -N in groundwater, (2)
126 construct a random forest model to predict spatiotemporal variations of groundwater

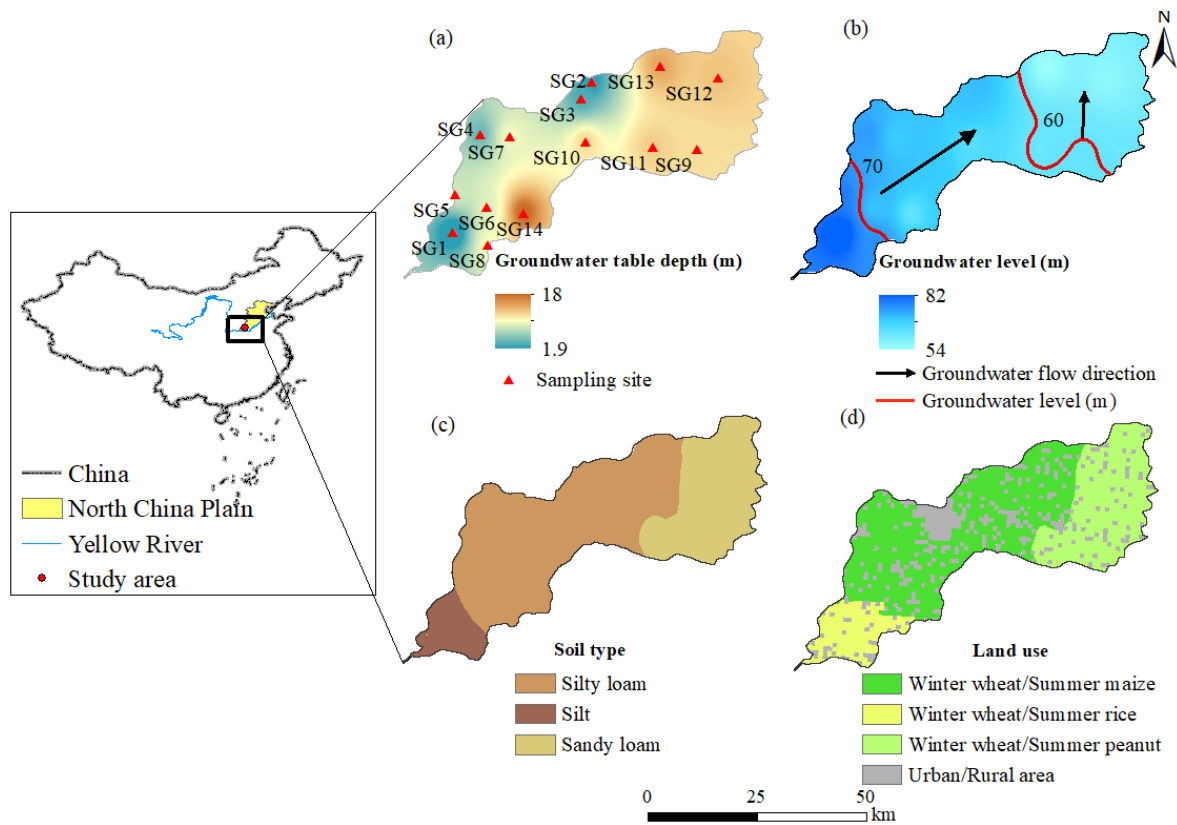
127 NO_3^- -N using a simplified approach to the model, and (3) identify the main effect factors
128 and processes, and evaluate their relationships with groundwater NO_3^- -N based on the
129 constructed random forest model and isotope approaches. The ability to prediction the
130 levels of groundwater NO_3^- -N pollution and uncover the mechanism that underlies its high
131 spatial and temporal variability is important for future management of agricultural N input
132 and water-quality protection in all developing and developed countries.

133 **2. Materials and Methods**

134 *2.1. Site description*

135 This study was conducted in an irrigated agricultural area (35°00'– 35°30' N, 113°31'–
136 114°25' E) located in the piedmont region of Taihang mountain, NCP, near the lower
137 reaches of the Yellow River (Fig. 1). The study site covers about 1,500 km² and belongs to
138 the temperate continental monsoon climate. Weather data from Xinxiang station, located
139 near our study site, were obtained from the China Meteorological Data Service Center
140 (<http://data.cma.cn/>). From 2010 to 2019, the average annual temperature was 15.6°C, with a
141 maximum temperature of 40°C and a minimum temperature of -13.1°C. The average annual
142 potential evaporation was 1,025 mm yr⁻¹, and the average annual rainfall was about 500 mm
143 yr⁻¹. The rainfall mainly occurred from June to September. According to the aquifer data
144 collected by field surveys, the average annual groundwater table depth ranged from 1.9 to
145 18 m, and the groundwater level ranged from 54.1 to 82.4 m (Fig. 1a–b). The particle size of
146 the soil was analyzed using a Malvern laser particle size analyzer as reported in our previous
147 study (Wang et al., 2021). According to the international soil texture classification standard,
148 the soil in the study area was classified into silt (8.3%), silty loam (61.1%), and sandy loam
149 (30.6%) (Fig. 1c). The main crops were winter wheat and summer maize (44%), winter
150 wheat and summer peanut (27%), and winter wheat and summer rice (11%) at the study site

151 (Fig. 1d). Other land uses, including rural residential areas, urban areas, and water body areas,
 152 constitute the remaining 18% of the study site (Fig. 1d). Irrigation and fertilization play
 153 important roles in ensuring stable crop production and the N applied on the agricultural
 154 land surface was mainly derived from chemical fertilizers, manure, and crop residues. The
 155 crop residues of winter wheat and summer corn/summer peanuts were all returned to the
 156 fields. The large amount and temporal variability in N input causes the groundwater N
 157 content to show high spatiotemporal variations and increase continuously, which poses a
 158 serious burden on groundwater N pollution and its management.



159 Fig. 1. Location of the study area in the North China Plain and the sampling sites, and maps of (a) the
 160 depth of the groundwater table, (b) the groundwater level, (c) the soil type and (d) land use. The land use
 161 map was obtained from the Geospatial Data Cloud (<http://www.gscloud.cn/search>)
 162

163 2.2. Field monitoring, sampling and chemical analysis

164 Fourteen representative sampling sites were selected according to the geographical
 165 location, land use, soil types, and the groundwater table depth to conduct field monitoring
 166 and collect surface soil and groundwater samples (Fig. 1a and Table 1). The field

167 monitoring and sample collection were conducted monthly from May 2017 to April 2019 at
 168 the 14 observation sites (Fig. 1a). The field monitoring involved the use of a multiparameter
 169 water quality probe (HORIBA, Ltd., Japan) to determine the basic physicochemical
 170 **parameters** of groundwater, including water temperature (GW-Temp), pH (GW-pH), total
 171 dissolved solids concentration (GW-TDS), oxidation-reduction potential (GW-ORP), and
 172 dissolved oxygen (GW-DO). At the same time, the depth of the groundwater table
 173 (GW-Dep) was measured and groundwater water samples were collected using a bailer tube.
 174 The data on irrigation were collected through a survey of the local farmers.

175 **Table 1.** Soil texture, crops, depth of groundwater table, and groundwater level at sampling sites

Sampling site	Soil texture	Crop rotation	Groundwater table depth (m)	Groundwater level (m)
SG1	Silt	Winter wheat-summer rice	2	82.4
SG2	Silty loam	Winter wheat-summer corn	3.9	65.5
SG3	Silty loam	Winter wheat-summer corn	4.9	65.9
SG4	Silty loam	Winter wheat-summer corn	6	69.3
SG5	Silty loam	Winter wheat-summer corn	9.4	74.5
SG6	Silty loam	Winter wheat-summer corn	9.9	64.5
SG7	Silty loam	Winter wheat-summer corn	10.5	63.6
SG8	Silty loam	Winter wheat-summer corn	10.7	71.5
SG9	Sandy loam	Winter wheat-summer peanuts	12.2	60.1
SG10	Silty loam	Winter wheat-summer corn	12.3	60.8
SG11	Sandy loam	Winter wheat-summer peanuts	12.9	58.8
SG12	Sandy loam	Winter wheat-summer peanuts	13.3	55.2
SG13	Silty loam	Winter wheat-summer corn	14.8	54.1
SG14	Silty loam	Winter wheat-summer corn	18	60.1

176 Groundwater and soil samples were collected at a frequency of once a month during the
 177 two-year study period. The groundwater samples were collected from wells and filtered
 178 using a 0.45 μm filter. These groundwater samples were brought back to the laboratory, and
 179 1 L of each sample was stored in a refrigerator at a temperature of 2°C until physiochemical
 180 analysis. At the same time, 100 ml groundwater samples were immediately frozen until

181 isotope analysis. Moreover, soil samples were collected from farmland within 500 m of the
182 wells where the groundwater samples were collected. The soil samples were gathered at the
183 soil surface with a depth of 0–10 cm in the farmland. All the soil samples were randomly
184 collected from six sites and mixed thoroughly to obtain a representative soil sample (500 g)
185 at each sampling site. Meanwhile, additional soil samples were collected in aluminum boxes
186 that were sealed and brought back to the laboratory. The mixed soil samples were dried
187 naturally, crushed through a 2 mm sieve, and stored in a cool and dry place until
188 physiochemical analysis.

189 The soil samples stored in the aluminum boxes were further used to analyze the soil
190 water content (SWC) through the oven-drying method. The parameters measured in the
191 laboratory include groundwater dissolved organic carbon (GW-DOC), groundwater
192 dissolved organic N (GW-DON), groundwater ammonium (GW-NH₄⁺-N), groundwater
193 nitrate (GW-NO₃⁻-N), groundwater nitrite (GW-NO₂⁻-N), soil organic carbon (SOC),
194 soil-soluble organic N (S-SON), soil nitrate (S-NO₃⁻-N), soil nitrite (S-NO₂⁻-N), and soil
195 ammonium (S-NH₄⁺-N). All these chemical parameters of soil and groundwater were
196 determined by the colorimetric method using a spectrophotometer (Thermo Fisher Scientific,
197 Inc, USA) according to the procedures reported by Hood-Nowotny et al. (2010) and Wang
198 et al. (2021).

199 Since August 2018, the influence of N leaching on groundwater has been weaker in the
200 following eight-month dry period than in the wet period. Therefore, in this study the isotope
201 analysis was performed on groundwater samples to identify nitrification and denitrification
202 in groundwater from August 2018 to April 2019. The stable isotope (¹⁵N and ¹⁸O)

203 abundance of groundwater NO_3^- -N was determined using the denitrifying bacteria method.
204 A seed solution (of glycerol 500 μL + bacteria 500 μL) was shaken for 12–15 hrs, purged
205 with N_2 for 3 hrs, and then added to the sample and placed in a shaking table at 100 rpm
206 overnight, shaking and as determined by an IRMS-100 mass spectrometer. USG32, USG34,
207 and USG35 were used as standard samples, and the results were corrected based on a
208 two-point calibration method. The measured isotope values correspond to the international
209 standard substances, expressed as follows:

$$\delta_{\text{sample}}(\text{‰}) = \frac{R_{\text{sample}} - R_{\text{VSMOW}}}{R_{\text{VSMOW}}} \times 1000$$

211 [1]

212 where δ_{sample} is the isotope value of the corresponding sample, R_{sample} is the ratio of
213 heavy and light isotopic abundance of elements in the sample, and R_{VSMOW} is the ratio of
214 heavy and light isotopic abundance of Vienna Standard Mean Ocean Water (VSMOW).

215 2.3. Random forest model

216 A random forest model was used to predict the intra-annual variations of groundwater
217 NO_3^- -N at the study site. The random forest model is a machine learning algorithm based on
218 the combination of the bagging integrated learning theory and the random subspace
219 algorithm, and it overcomes the drawbacks of overfitting and instability (Breiman, 2001).
220 The model constructs several regression trees (ntree) by setting nodes (mtry) on a random
221 subset of the original training dataset, according to Amit and Geman (1997). The random
222 forest model divides the data into training and test sets by setting a certain ratio (P), and the P
223 ratio of 2:1 was chosen in this study. The random forest model was then calibrated using the
224 training subset and validated using the test subset. A package of *randomForest* in Rstudio
225 (version 4.2.2) was adopted to construct the random forest model.

226 The model accuracy was made to meet the research needs by adjusting the three
227 parameters n_{tree} , m_{try} , and P . Three error metrics, the root mean square error (RMSE), the
228 coefficient of determination (R^2), and the mean absolute error (MAE), were selected to
229 evaluate the random forest model's performance. The trained and validated random forest
230 model was used to analyze the importance of groundwater NO_3^- -N influencing factors. The
231 importance of an effect factor is defined as the increase in the predicted mean squared error
232 (MSE) after randomly permuting this factor (Breiman, 2001). It reflects the contribution of
233 each effect factor to a groundwater nitrate concentration. The normalized increased MSE,
234 that is, the relative importance, for each effect factor was between 0% and 100%. The
235 partial dependence shows the marginal effect of each explanatory variable for the response
236 after considering the average effects of the other variables. The partial dependence plots of
237 the important effect factors were used to analyze the relationship between a single
238 independent variable and the groundwater NO_3^- -N subject to the influence of the other
239 independent variables.

240 2.4. Data analysis

241 The spatial distribution of the measured data was interpolated using inverse distance
242 weighting in ArcGIS to analyze the spatial distribution characteristics of different N species
243 in groundwater from the study site. The Spearman coefficient (r), paired t-tests, and a
244 one-way analysis of variance (ANOVA) were used to calculate the correlation coefficient
245 and test the significance using SPSS software (version 25.0, IBM Corporation, Chicago,
246 Illinois). Denitrification was identified based on NO_3^- -N isotopes and their relationships,
247 that is, the enrichment of $\delta^{15}N$ - NO_3^- and $\delta^{18}O$ - NO_3^- with a slope of 0.5–1.0 ($\Delta\delta^{18}O/\delta^{15}N$),

248 and a decrease of NO_3^- -N concentrations. An increase of NO_3^- -N concentrations and
249 decrease of $\delta^{15}\text{N}-\text{NO}_3^-$ and $\delta^{18}\text{O}-\text{NO}_3^-$ were analyzed to identify nitrification.

250 **3. Results and Discussion**

251 *3.1. Statistical analysis of groundwater nitrate concentration and its effect factors*

252 Table 2 lists the variables analyzed in this study, that is, the precipitation, irrigation,
253 soil sand content (SSC), SWC, SOC, S-SON, S- NO_3^- -N, S- NH_4^+ -N, S- NO_2^- -N,
254 GW- NO_3^- -N, GW- NH_4^+ -N, GW- NO_2^- -N, GW-DON, GW-DOC, GW-Dep, GW-Temp,
255 GW-TDS, GW-pH, GW-DO, and GW-ORP. The 19 predictor variables in Table 2 were
256 considered as effect factors that could potentially influence groundwater GW- NO_3^- -N. The
257 maximum, minimum and mean values; standard deviation (SD); coefficient of variation
258 (CV); and quartile of these variables are shown in Table 2. Different forms of N content
259 from surface soil were considered instead of N input due to the complex sources and
260 transformation processes of N input. The statistical characteristics of N input were not
261 analyzed. As shown in Table 2, all the CV values, which can reflect the size of dispersion of
262 the measured data, ranged from 0.08 to 2.06, and most of the factors had small CVs and
263 low variability. However, the CV values of groundwater inorganic N were all greater than 1,
264 indicating that spatial and temporal patterns of groundwater inorganic N concentrations
265 were highly variable in the study area. The SD and CV of groundwater NO_3^- -N
266 concentrations were the largest compared with NH_4^+ -N and NO_2^- -N, indicating that the
267 spatiotemporal variability of groundwater NO_3^- -N concentrations was the highest among
268 the three forms of inorganic N.

269 Table 2 Summary statistics of the variables analyzed in this study (May 2017-April 2019)

	Variables	Max	Min	Mean	SD	CV	Quartiles		
							25th	50th	75th
Predictor variables, i.e., the effect factors that could potentially influence GW-NO ₃ ⁻ -N	Precipitation (mm)	98.20	0.20	35.47	33.07	0.93	4.78	26.65	67.95
	Irrigation (mm)	300.0	0.00	96.10	93.25	0.97	0.00	105.0 0	200.0 0
	GW-Dep (m)	19.00	0.10	10.03	4.48	0.45	5.82	10.90	13.03
	SSC (%)	71.26	12.45	31.10	16.60	0.53	21.4 0	25.70	28.2
	SWC (mg mg ⁻¹)	0.42	0.01	0.16	0.09	0.54	0.09	0.16	0.22
	SOC (g kg ⁻¹)	37.19	0.75	16.80	6.83	0.41	11.5 7	16.19	21.91
	S-SON (mg kg ⁻¹)	805.5 0	34.55	333.6 2	146.3 2	0.44	209. 80	335.0 6	440.3 3
	S-NO ₃ ⁻ -N (mg kg ⁻¹)	265.1 0	6.87	53.89	47.22	0.88	25.5 7	38.00	59.64
	S-NO ₂ ⁻ -N (mg kg ⁻¹)	11.40	0.01	2.74	1.80	0.66	1.68	2.28	3.15
	S-NH ₄ ⁺ -N (mg kg ⁻¹)	62.80	6.41	18.99	7.12	0.38	14.8 9	18.49	22.25
	GW-pH	8.70	3.42	7.14	0.58	0.08	6.84	7.13	7.52
	GW-Temp (°C)	31.20	6.95	18.40	4.07	0.22	15.5 6	17.85	20.80
	GW-DO (mg L ⁻¹)	50.00	2.65	11.78	7.28	0.62	7.31	9.91	13.37
	GW-ORP (mV)	308.0	-136. 0	137.2 4	120.6 7	0.88	46.2 5	172.5 0	235.0 0
	GW-DOC (mg L ⁻¹)	280.3 0	<LO D	42.84	43.27	1.01	13.3 5	25.83	64.60
	GW-TDS (g L ⁻¹)	3.17	0.22	1.15	0.65	0.56	0.73	0.95	1.30
	GW-DON (mg L ⁻¹)	808.3 8	<LO D	69.40	78.20	1.13	25.7 0	69.40	73.30
	GW-NO ₂ ⁻ -N (mg L ⁻¹)	0.52	<LO D	0.08	0.11	1.49	0.01	0.03	0.07
	GW-NH ₄ ⁺ -N (mg L ⁻¹)	22.80	0.04	1.52	1.61	1.05	0.77	1.23	1.93
	Response variable	GW-NO ₃ ⁻ -N (mg L ⁻¹)	132.2 0	0.09	8.31	14.77	1.78	1.60	3.60

270 Note: standard deviation (SD), coefficient of variation (CV), depth of groundwater table (GW-Dep),
271 soil sand content (SSC), soil water content (SWC), soil organic carbon content (SOC), soil soluble
272 organic N content (S-SON), soil nitrate content (S-NO₃⁻-N), soil nitrite content (S-NO₂⁻-N), soil
273 ammonium content (S-NH₄⁺-N), groundwater pH (GW-pH), groundwater temperature (GW-Temp),
274 groundwater dissolved oxygen content (GW-DO), groundwater oxidation-reduction potential
275 (GW-ORP), groundwater dissolved organic carbon concentration (GW-DOC), groundwater total
276 dissolved solids (GW-TDS), groundwater dissolved organic N concentration (GW-DON),
277 groundwater ammonium concentration (GW-NH₄⁺-N), groundwater nitrite concentration

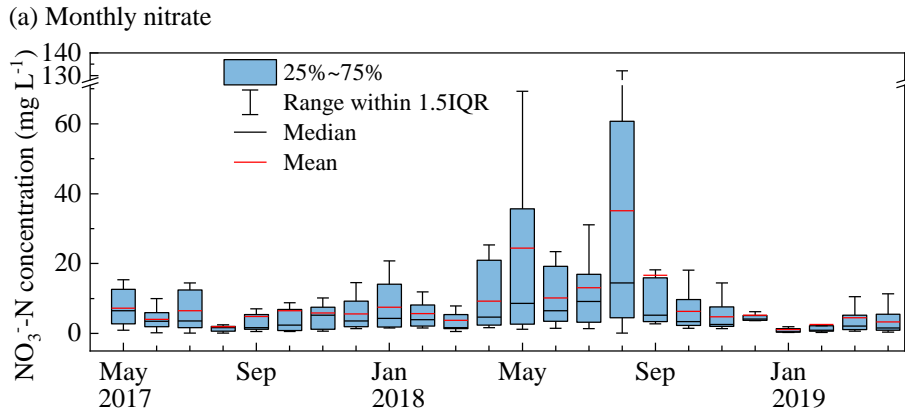
278 (GW-NO₂⁻-N), groundwater nitrate concentration (GW-NO₃⁻-N), Limit of detection (LOD).

279 *3.2. Spatiotemporal variations of inorganic N species in groundwater*

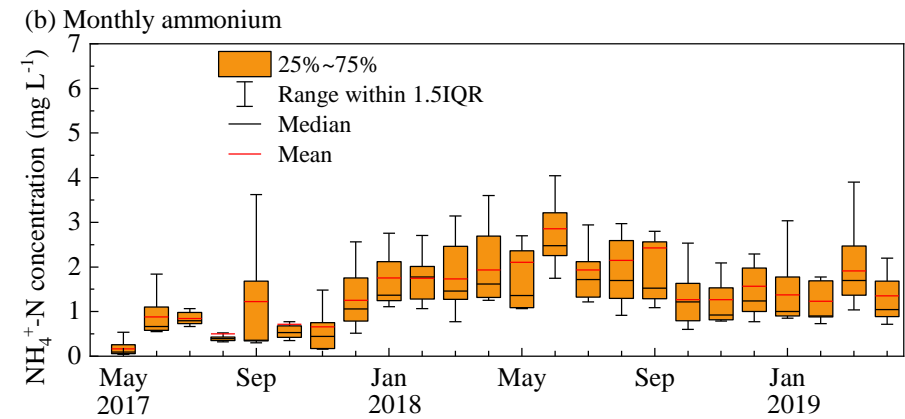
280 The temporal and spatial variations of groundwater inorganic N concentrations in the
281 study area are shown in Figs. 2–3. There is a large difference in the temporal variations of
282 groundwater inorganic N concentrations (Fig. 2a–c). The groundwater NO₃⁻-N
283 concentrations ranged from 0.09 to 132.16 mg L⁻¹. The groundwater NO₃⁻-N concentrations
284 were higher from May to August and usually decreased from September to April of the
285 following year. In this study, the peaks of the groundwater NO₃⁻-N concentrations were
286 caused by high level of NO₃⁻-N leaching after fertilization, irrigation, and heavy rain in the
287 rainy season (May–August). This result is consistent with the previous findings of Biddau
288 et al. (2019), who reported the high levels of NO₃⁻-N concentrations (up to 162 mg L⁻¹) in
289 shallow groundwater also occur when NO₃⁻-N leaching is high, particularly during
290 fertilization and irrigation periods. The groundwater NH₄⁺-N concentrations ranged from
291 0.04 to 22.8 mg L⁻¹. They showed a fluctuating increasing trend from May 2017 to May
292 2018, and the values were relatively high and stable from September 2018 to April 2019.
293 The concentrations of groundwater NO₂⁻-N ranged from 0.01 to 0.52 mg L⁻¹, and they were
294 higher from May to October than in other months for both 2017 and 2018. Comparing the
295 three inorganic forms of N, the concentrations of groundwater NO₃⁻-N were the highest,
296 while NO₂⁻-N was the lowest among them. Overall, all three forms of groundwater
297 inorganic N in the second year were generally higher than in the first year (Fig. 2d–f). This
298 increase in annual NO₃⁻-N indicates that the shallow groundwater system might not be able
299 to rapidly recover from N pollution.

300 The spatial variations of inorganic N concentrations in groundwater at the 14 sampling
301 sites are shown in Fig. 3 a–f. The GW-NO₃⁻-N concentrations were high and followed the
302 order of SG3, SG4, SG6, SG7, and SG2, while they were lower at the other sites.
303 Groundwater NH₄⁺-N concentrations were the highest at SG3, followed by SG2, SG7.
304 Groundwater NO₂⁻-N concentrations were the highest at SG6, followed by SG4, SG5, SG7.
305 Overall, all the forms of groundwater inorganic N concentrations varied spatially in the
306 study area. Among them, the spatial variations of GW-NO₃⁻-N concentrations were highly
307 variable in the study area. Site SG1 showed the lowest GW-Dep and low GW-NO₃⁻-N
308 concentrations (Table 1). This is likely because the soil NO₃⁻-N source was lost by
309 denitrification under intermittent ponding irrigation during the summer rice season. Thus,
310 the NO₃⁻-N leaching was reduced at SG1. Compared with sites SG8–SG14, the sites
311 SG2–SG4 and SG6–SG7 had lower GW-Dep and showed higher GW-NO₃⁻-N
312 concentrations (Table 1). This could be because NO₃⁻-N leaching decreased with the
313 increase of vadose zone thickness (Weitzman et al., 2022). Sites SG5 and SG8-SG14
314 showed similar GW-NO₃⁻-N concentrations but were characterized by different GW-Dep. It
315 indicates that in addition to N source and leaching, other groundwater environmental
316 factors might affect GW-NO₃⁻-N. Overall, for the GW-NO₃⁻-N concentrations, about 24%
317 of samples exceeded the World Health Organization (WHO)’s water quality standard of 10
318 mg L⁻¹.

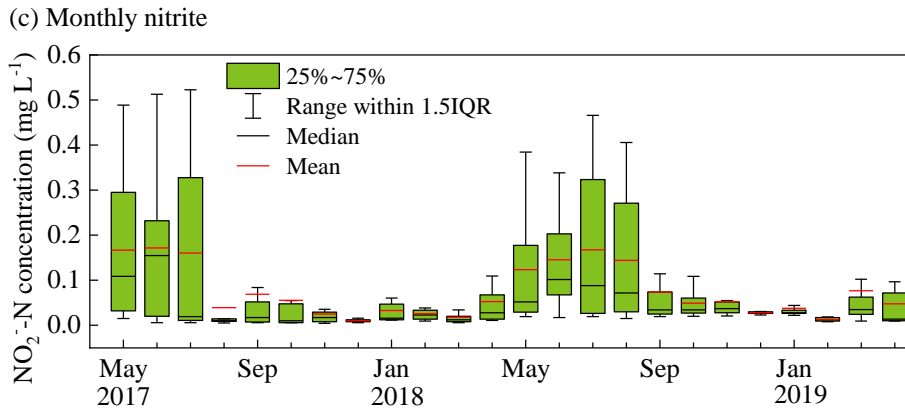
319



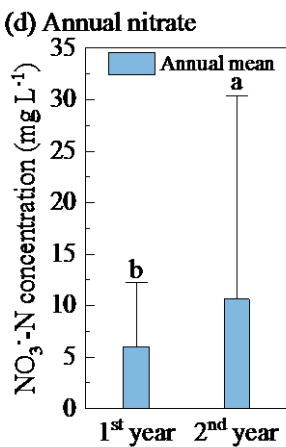
320



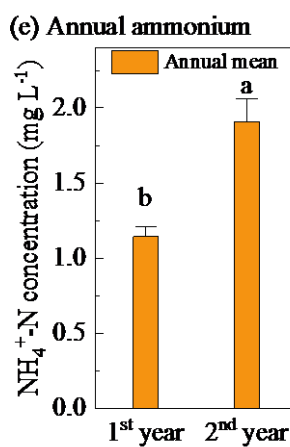
321



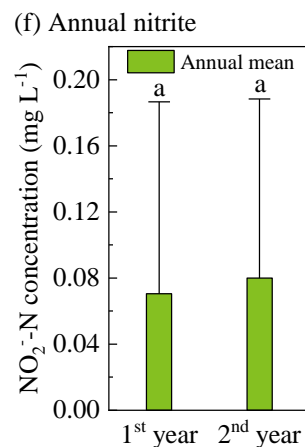
322



323



324



325

326

Fig. 2 Temporal variations of concentrations of (a) monthly nitrate, (b) monthly ammonium, (c) monthly nitrite, (d) annual nitrate, (e) annual ammonium, and (f) annual nitrite in groundwater. The different letters above the error bars in (d), (e), and (f) indicate significant difference ($p < 0.05$) between the different years.

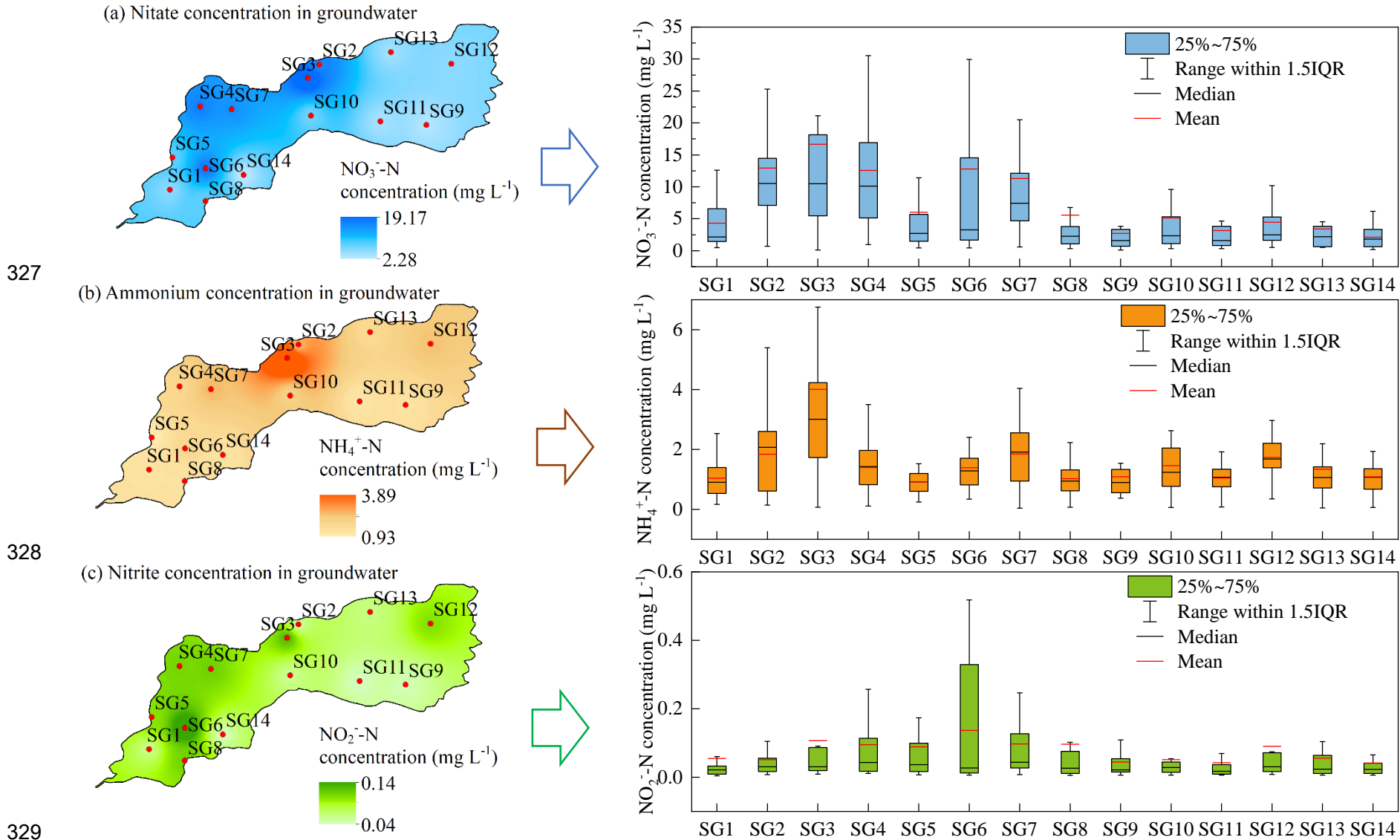


Fig. 3 Spatial distribution of concentrations of (a) nitrate, (b) ammonium, and (c) nitrite in groundwater

332

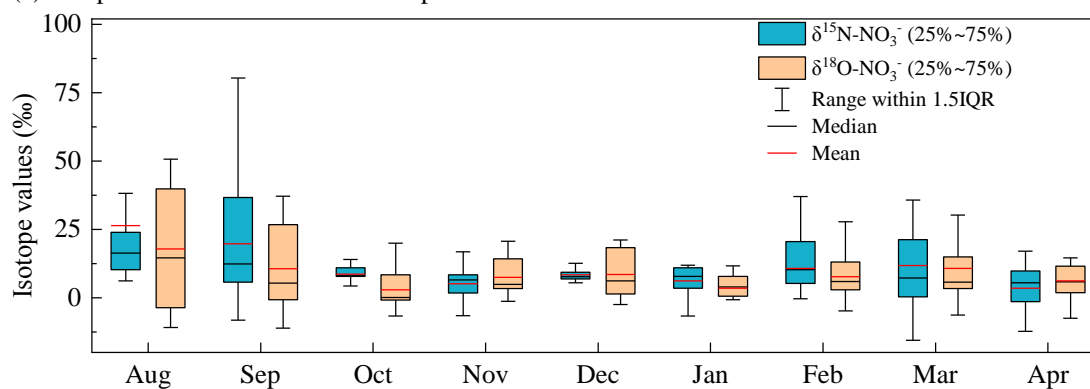
333 3.3. *Spatiotemporal variations of groundwater nitrate isotopes*

334 The $\delta^{15}\text{N-NO}_3^-$ values ranged from -20.16‰ to 107.88‰ with a mean value of
335 $11.16\pm 16.30\%$. The $\delta^{18}\text{O-NO}_3^-$ values ranged from -15.69‰ to 56.56‰ with a mean
336 value of $8.39\pm 12.40\%$. The temporal variations of the $\delta^{15}\text{N-NO}_3^-$ and $\delta^{18}\text{O-NO}_3^-$
337 values in groundwater from August 2018 to April 2019 are shown in Fig. 4a. The
338 $\delta^{15}\text{N-NO}_3^-$ and $\delta^{18}\text{O-NO}_3^-$ values showed similar temporal variation trends.
339 Groundwater NO_3^- -N in August was heavily affected by leaching because of the
340 application of fertilizer and manure and wheat residue return to the field in the wet
341 season when heavy and high-frequency precipitation happens. Meanwhile, in the
342 following months (September to April of the following year), the leaching amount was
343 reduced due to less precipitation in the dry seasons. The measured mean $\delta^{15}\text{N-NO}_3^-$
344 value was 26.36‰ in August 2018, and the mean $\delta^{18}\text{O-NO}_3^-$ value was 17.83‰ in the
345 same month. While the mean $\delta^{15}\text{N-NO}_3^-$ values ranged from 3.54‰ to 19.73‰ and the
346 mean $\delta^{18}\text{O-NO}_3^-$ values ranged 2.91‰ to 10.79‰ from September 2018 to April 2019,
347 both are lower than the measured values obtained in August 2018. During the periods
348 August–October 2018, December 2018–January 2019, and March–April 2019, the
349 $\delta^{15}\text{N-NO}_3^-$ and $\delta^{18}\text{O-NO}_3^-$ showed decreasing trends. The slight increase in the
350 $\delta^{15}\text{N-NO}_3^-$ and $\delta^{18}\text{O-NO}_3^-$ values after October 2018 (during wheat sowing) and
351 February 2019 (during wheat regeneration) can be attributed to the application of
352 fertilizer, crop residue returning to the field, and irrigation. These management practices
353 caused an increase in NO_3^- -N denitrification rates due to the increase of soil NO_3^- -N,

354 soil water, and organic carbon in the vadose zone.

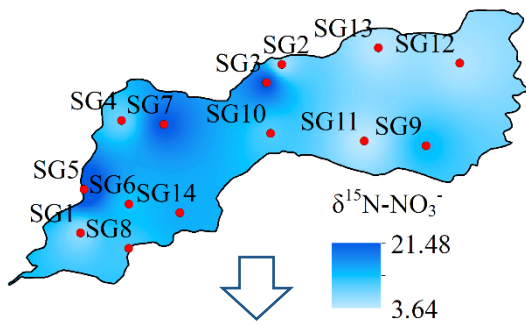
355 The spatial distribution of the $\delta^{15}\text{N-NO}_3^-$ and $\delta^{18}\text{O-NO}_3^-$ values in groundwater
356 are shown in Fig. 4b–c. The $\delta^{15}\text{N-NO}_3^-$ values from the SG5 (mean value of 29.55‰)
357 were obviously higher than those of the other sites (mean values ranged from 0.92‰ to
358 14.35‰) (Fig. 4d). The maximum $\delta^{15}\text{N-NO}_3^-$ value of 107.88 ‰ was observed in
359 August 2018 at SG5. This could be because the silt loam soil could have preserved soil
360 water, and the manure application could have provided a carbon source for
361 denitrification during wet season. Thus, the NO_3^- -N denitrification rates could be higher
362 during NO_3^- -N leaching in the vadose zone at this silt loam site during the late wet
363 season of August. Meanwhile, the $\delta^{18}\text{O-NO}_3^-$ values from the SG5, SG6, SG9, and
364 SG14 (mean values ranged from 13.97‰ to 18.03‰) were higher than those of the
365 other sites (mean values ranged from 2.05‰ to 8.67‰). At these sites, the $\delta^{18}\text{O-NO}_3^-$
366 values were as high as 33.93‰ to 50.69‰ in August 2018, indicating that NO_3^- -N in
367 precipitation was a direct source of recharge to the groundwater. This agrees well with
368 the measured $\delta^{18}\text{O-NO}_3^-$ values in precipitation which ranged from 35‰–59‰
369 (VSMOW) as reported by Spoelstra et al. (2001).

(a) Temporal variations of stable isotopes of nitrate

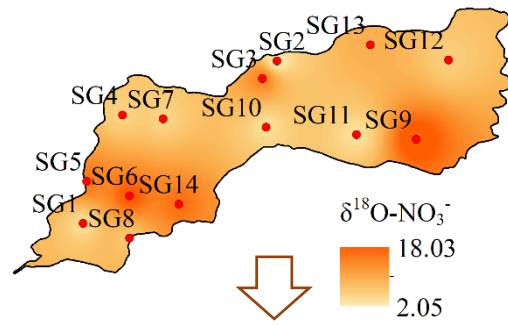


370

(b) Spatial variations of $\delta^{15}\text{N-NO}_3^-$ values

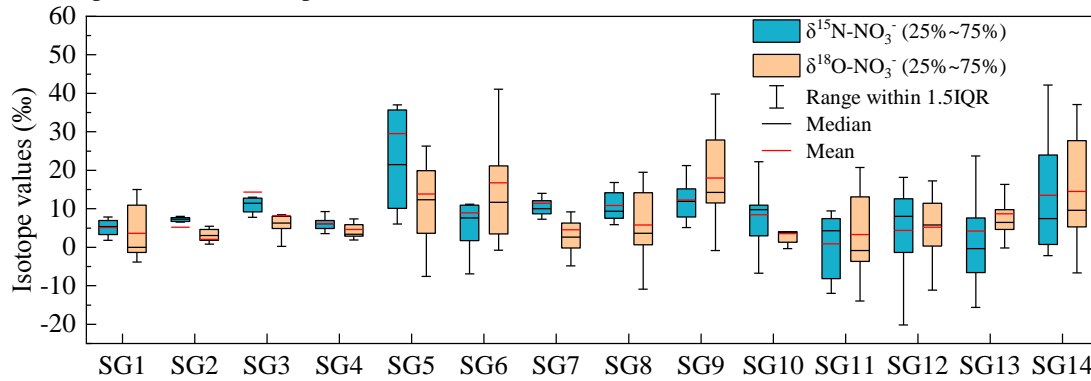


(c) Spatial variations of $\delta^{18}\text{O-NO}_3^-$ values



371

(d) Boxplots of stable isotopes of nitrate



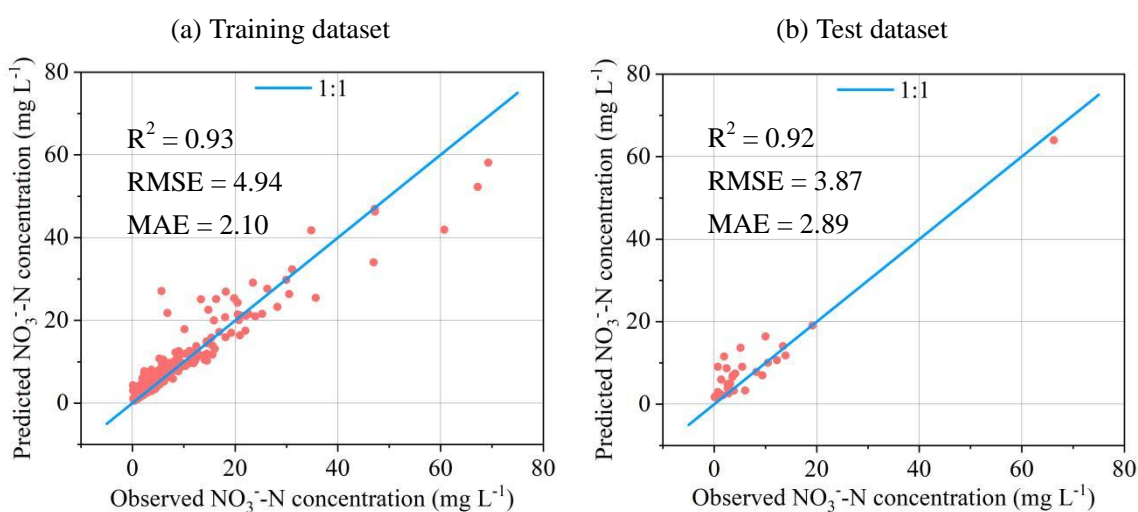
372

373 Fig. 4 (a) temporal variations of $\delta^{15}\text{N-NO}_3^-$ and $\delta^{18}\text{O-NO}_3^-$ values, spatial variations of (b) $\delta^{15}\text{N-NO}_3^-$
374 and (c) $\delta^{18}\text{O-NO}_3^-$ values, and (d) box plots of $\delta^{15}\text{N-NO}_3^-$ and $\delta^{18}\text{O-NO}_3^-$ values at 14 sampling sites

375 3.4. Prediction of spatiotemporal variations of groundwater nitrate

376 In this study, the 19 effect factors in Table 2 were set as predictor variables, and
377 GW- NO_3^- -N was set as a response variable when applying random forest model. Fig. 5
378 shows the comparison between the measured and random forest model predicted
379 GW- NO_3^- -N for the training dataset and test dataset. For the training dataset, the
380 simulation results yielded an R^2 of 0.93, an RMSE of 4.94, and an MAE of 2.10. After
381 training, the model was validated using the test dataset, which yielded an R^2 of 0.92, an
382 RMSE of 3.87, and an MAE of 2.89, indicating that the model simulation results were
383 good. Compared with previous results from random forest models reported by
384 Ouedraogo et al. (2019), Pennino et al. (2020) and He et al. (2022), the model

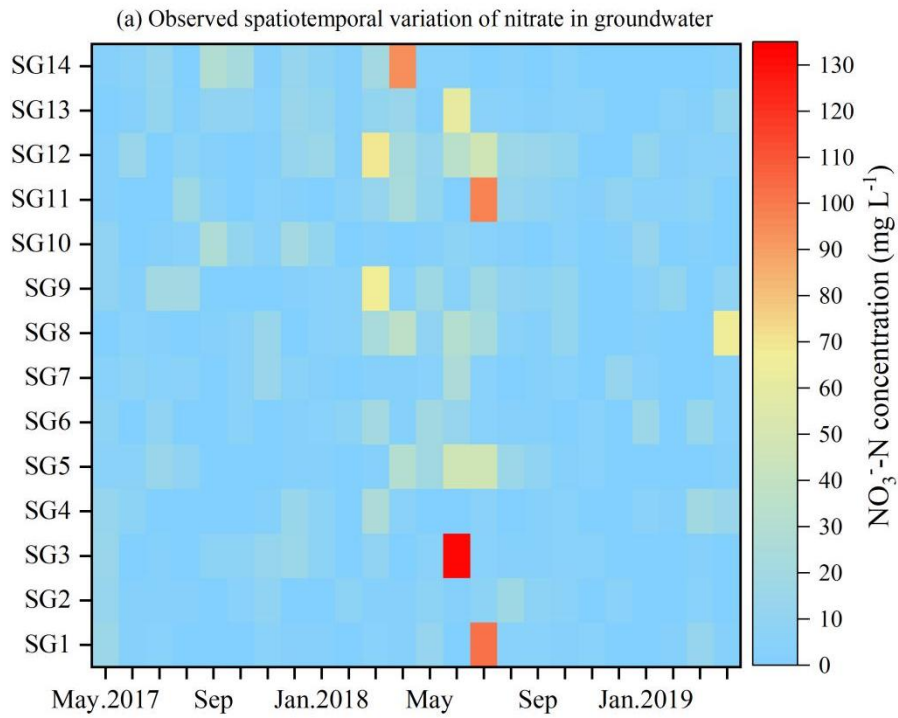
385 constructed in this study achieved better performance in predicting NO_3^- -N
 386 concentrations in groundwater. In addition, the result of this study showed better model
 387 accuracy (R^2 range of 0.92–0.93) compared to the studies of Knoll et al. (2019) and El
 388 Amri et al. (2022), which used other machine learning models, e.g., classification and
 389 regression trees and artificial neural network, to assess groundwater NO_3^- -N and yielded
 390 a model accuracy (R^2) range of 0.39–0.90.



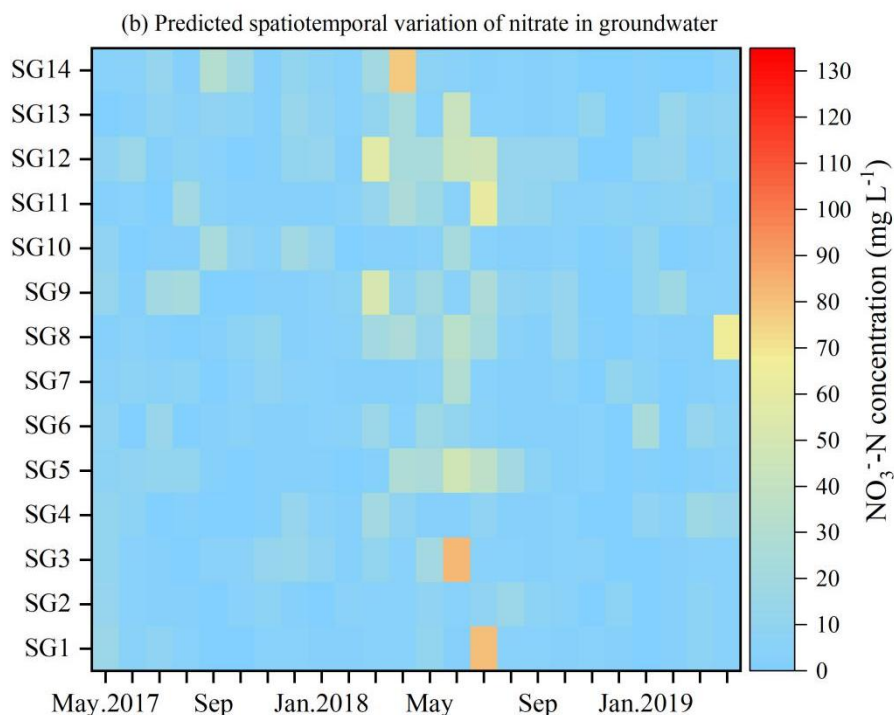
391 **Fig. 5** Comparison between the observed and random forest model predicted groundwater nitrate
 392 concentrations in groundwater for the (a) training and (b) test datasets

393 The spatial and temporal distribution of the observed and random forest model
 394 predicted groundwater NO_3^- -N concentrations are shown in Fig. 6. Comparing the
 395 measured and predicted spatial and temporal distributions of groundwater NO_3^- -N
 396 concentrations, the results showed that the predicted values were generally close to the
 397 measured values. Despite its remarkable performance, the random forest model may
 398 result in a loss when predicting the extreme ends or responses beyond the boundaries of
 399 the training data (Smarra et al., 2018). Therefore, predicting values beyond the range in
 400 the training data is not recommended. A representative training dataset is important for

401 assuring model performance when constructing a random forest model. In addition, it
402 was found that the performance of the model may be influenced by other factors, such
403 as the choice of a dependent variable, and independent variables, and the size of the
404 dataset.



405



406

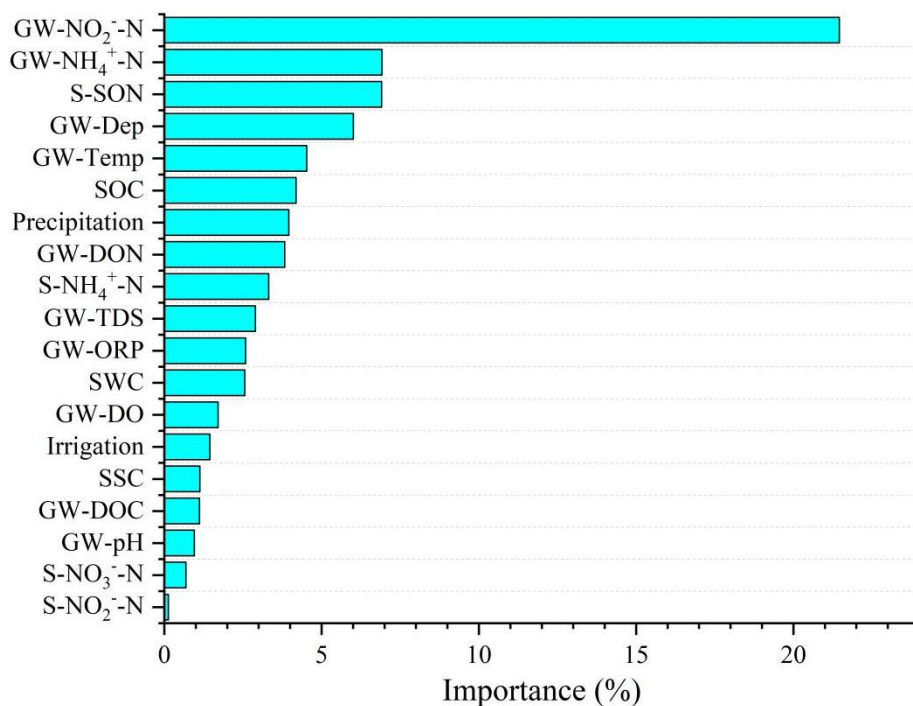
407 Fig. 6 Comparison between (a) observed and (b) the random forest model predicted spatiotemporal
 408 variation of nitrate in groundwater

409

410 3.5. Importance of effect factors on spatiotemporal variations of groundwater nitrate

411 The random forest importance and partial dependency analysis were used to
 412 calculate the importance of each influential factor on the spatiotemporal variations of
 413 GW-NO₃⁻-N, and to comprehensively analyze the relationship between environmental
 414 factors and GW-NO₃⁻-N (Figs. 7–8). As shown in Fig. 7, the key factors affecting
 415 GW-NO₃⁻-N concentrations were GW-NO₂⁻-N, GW-NH₄⁺-N, S-SON, and GW-Dep,
 416 with relative importance of 21.46%, 6.92%, 6.91%, and 6.01%, respectively, which are
 417 all higher than 5.0%. Meanwhile the relative importance of the other factors to
 418 GW-NO₃⁻-N concentrations were all less than 5.0%. The GW-DON as the substrate of
 419 mineralization contributed an importance of 3.83%, which was less important than
 420 GW-NO₂⁻-N and GW-NH₄⁺-N. Among the different surface soil N species, SON was
 421 the main influential factor on GW-NO₃⁻-N, followed by the S-NH₄⁺-N (3.32%) and
 422 S-NO₃⁻-N (0.69%), whereas the importance of S-NO₂⁻-N (0.13%) was the smallest. The

423 GW-Dep impacted the leaching of N out of the vadose zone. Moreover, the SOC, SWC,
 424 and SSC also influenced GW-NO₃⁻-N with a relative importance of 4.19%, 2.56%, and
 425 1.13%, respectively. Precipitation and irrigation were identified with a contribution of
 426 3.96% and 1.45%, respectively. As for groundwater environmental factors influencing
 427 N transformation by changing microbial survival, GW-Temp was identified with a
 428 relative importance of 4.53%. Finally, the relative importance of the other factors
 429 yielded the following order GW-TDS (2.90%), GW-ORP (2.59%), GW-DO (1.71%),
 430 GW-DOC (1.12%), and GW-pH (0.96%). The relationships between the effect factors
 431 and groundwater NO₃⁻-N were further analyzed and are shown in Fig. 8 according to their
 432 order of importance.



433

434 Fig. 7 Relative importance of potential effect factors to groundwater nitrate concentrations

435 *3.6. Relationship between effect factors and groundwater nitrate*

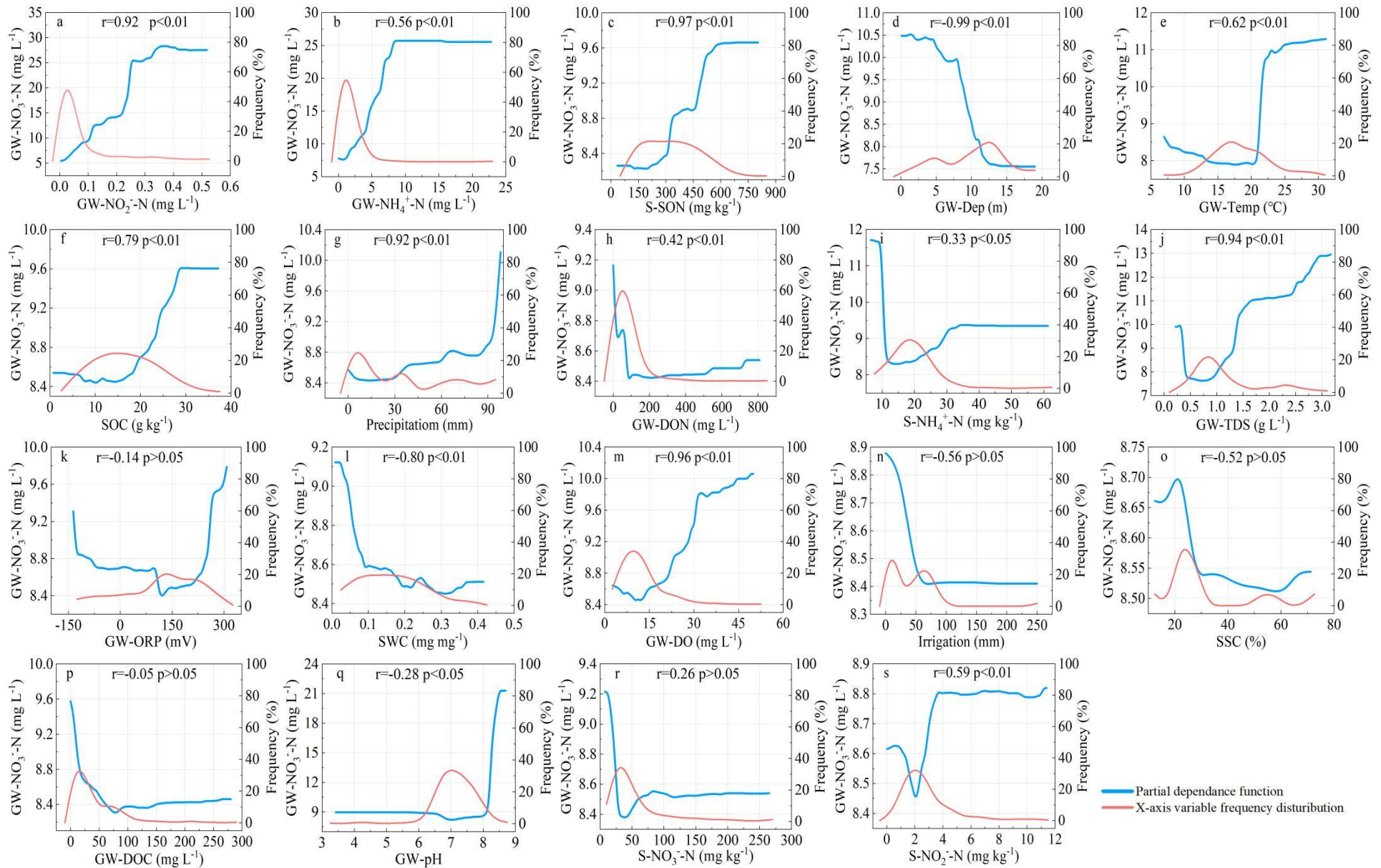
436 The relationships between the observed mean values of each explanatory variable,
 437 i.e., each effect factor (Table 2) and GW-NO₃⁻-N at the 14 sites are shown in Fig. S1 in the
 438 supporting information. Significant correlations were found between GW-NO₂⁻-N and

439 GW-NO₃⁻-N (r=0.84, p<0.01), between GW-NH₄⁺-N and GW-NO₃⁻-N (r=0.59, p<0.05),
440 between GW-Dep and GW-NO₃⁻-N (r=-0.68, p<0.01), between GW-TDS and
441 GW-NO₃⁻-N (r=0.56, p<0.05), and between GW-ORP and GW-NO₃⁻-N (r=0.57,
442 p<0.05). However, the interactions of all the effect factors resulted in nonsignificant
443 correlations between other factors and GW-NO₃⁻-N, and led to difficulty in reflecting
444 their impacts on groundwater NO₃⁻-N concentrations. Therefore, the partial dependence
445 plots based on the random forest model were used to show the complex nonlinear
446 relationships between GW-NO₃⁻-N and each effect factor, along with the frequency
447 distribution (Fig. 8a–s).

448 *3.6.1 Different forms of N in groundwater*

449 The GW-NO₂⁻-N and GW-NH₄⁺-N were significantly positively correlated with the
450 GW-NO₃⁻-N, with an r of 0.92 (p<0.01) and 0.56 (p<0.01), respectively (Fig. 8a–b). A
451 positive correlation (r = 0.42, p<0.01) was identified between GW-DON and
452 GW-NO₃⁻-N (Fig. 5h). However, a negative correlation was found when DON was
453 below 100 mg L⁻¹. The GW-NO₂⁻-N, an intermediate product of nitrification and
454 denitrification, is of great importance in the N transformation process and is the most
455 crucial factor as an indicator of the NO₃⁻-N transformation rate. In this study, the
456 presence or accumulation of NO₂⁻-N showed the highest impacts on groundwater
457 NO₃⁻-N concentrations compared to the other factors (Fig. 7). This could be due to
458 denitrification as a heterotrophic process. The previous results of Du et al. (2016)
459 indicated that denitrifying bacteria in the system preferred using NO₃⁻-N as an electron
460 acceptor rather than NO₂⁻-N. Thus, the accumulation and production of NO₂⁻-N can
461 occur during denitrification, and NO₂⁻-N could be regarded as an index of the activity of
462 denitrifying bacteria. The high-throughput sequencing analysis of Du et al. (2016)

463 revealed that the genus of *Thauera* bacteria was dominant in the denitrifying community
464 with high NO_2^- -N accumulation. According to the previous study, the DON retained by
465 the soil accounts for 25%-35% of the total DON, and the remaining amount enters the
466 groundwater with leaching (Zhou et al., 2003). The leached DON undergoes
467 mineralization (ammonification and nitrification) and subsequently causes an increase of
468 both NH_4^+ -N and NO_3^- -N in groundwater (Liu et al., 2022). The GW-DON can produce
469 GW- NH_4^+ -N through ammonification and then indirectly influence GW- NO_3^- -N
470 through nitrification of GW- NH_4^+ -N (Wang et al., 2020; Liu et al., 2022). In this study,
471 the GW- NH_4^+ -N was the second important factor influencing GW- NO_3^- -N. This could
472 be due to the nitrification processes, where the GW- NH_4^+ -N are the substrates for
473 NO_3^- -N production by the nitrification process. During the nitrification processes,
474 because the oxidation of NO_2^- -N to NO_3^- -N is rapid in natural systems, the slower
475 oxidation of NH_4^+ -N to NO_2^- -N is the main process that controls NO_2^- -N production
476 (Nikolenko et al., 2018). Both NO_3^- -N and NO_2^- -N appeared when organic N
477 mineralization occurred in groundwater. The GW- NH_4^+ -N showed higher importance
478 and a greater correlation coefficient than GW-DON, which might indicate that NH_4^+ -N
479 can be rapidly converted into NO_3^- -N. Thus, it is likely that the GW- NO_2^- -N is an
480 intermediate product of denitrification in this study. Overall, the availability of
481 GW- NO_2^- -N and GW- NH_4^+ -N are indicated them as the primary factors reflecting the
482 denitrification and nitrification rates. Their highly ranked importance showed clear
483 evidence that N transformation played a critical role in controlling groundwater NO_3^- -N
484 variations.



485

486

487

488

489

Fig. 8 Partial dependence plots of constructed random forest model. Blue lines show the partial dependence function, and red lines show the x-axis factor frequency

490 3.6.2 Different forms of N in surface soil and vadose zone characteristics

491 The spatiotemporal variations of GW-NO₃⁻-N depended on different forms of N
492 content in the soil and leaching. In terms of N forms in the soil, there are positive
493 correlations between SON and GW-NO₃⁻-N ($r = 0.97$, $p < 0.01$, Fig. 8c), between
494 S-NO₂⁻-N and GW-NO₃⁻-N ($r = 0.59$, $p < 0.01$, Fig. 8s), and between S-NH₄⁺-N and
495 GW-NO₃⁻-N ($r = 0.33$, $p < 0.05$, Fig. 8i). This is mainly because mineralization and
496 nitrification play critical roles in the production of NO₃⁻-N and the subsequent NO₃⁻-N
497 leaching. Moreover, a higher S-SON might indicate more accumulated organic and
498 inorganic N in the vadose zone. On the other hand, another possible explanation is
499 that the soil-leached SON, NH₄⁺-N, and NO₂⁻-N has undergone a transformation and
500 produced NO₃⁻-N in the vadose zone before entering into the groundwater. Despite
501 the fact that S-NO₃⁻-N has a direct impact on GW-NO₃⁻-N, no significant correlation
502 was found between them ($r = 0.26$, $p > 0.01$) (Fig. 8r). The S-NO₃⁻-N could be divided
503 into two groups by the threshold around 30 mg kg⁻¹. A negative correlation was found
504 when S-NO₃⁻-N was below 30 mg kg⁻¹, indicating that the lower S-NO₃⁻-N could be
505 attributed to more NO₃⁻-N leaching loss. In contrast, their relationship changed to be
506 positive when S-NO₃⁻-N was above 30 mg kg⁻¹, indicating that more NO₃⁻-N in
507 surface soil resulted in more NO₃⁻-N leaching loss into the groundwater.

508 The depth from the soil surface to the groundwater table plays an essential role
509 in the accumulation and reduction of N leaching loss in the vadose zone. The
510 GW-Dep was significantly negatively correlated with the GW-NO₃⁻-N ($r = -0.99$,
511 $p < 0.01$) (Fig. 8d). Our results showed that a shallower groundwater table leads to an
512 increasing amount of NO₃⁻-N in groundwater, consistent with many previous studies
513 (Awais et al., 2021; Li et al., 2021; El Amri et al., 2022; He et al., 2022). Other
514 characteristics of the vadose zone also affected N transformation and subsequent

515 leaching of NO_3^- -N, including SWC, SSC, and SOC. A significant negative
516 correlation ($r = -0.80$, $p < 0.01$) was found between SWC and GW- NO_3^- -N (Fig. 8l).
517 The increase of SWC caused the decrease of NO_3^- -N production by nitrification and
518 enhanced denitrification removal of NO_3^- -N in the soil (Sexstone et al., 1985).
519 However, it is notable that a positive correlation was identified when the SWC was
520 above 0.30. This could be because higher SWC increases NO_3^- -N leaching into the
521 groundwater. No significant negative correlation ($r = -0.56$, $p > 0.05$) was found
522 between irrigation and GW- NO_3^- -N (Fig. 8n). Meanwhile, a significant positive
523 correlation ($r = 0.92$, $p < 0.01$) was identified between precipitation and GW- NO_3^- -N
524 (Fig. 8g). Irrigation and precipitation influenced both soil water content and soil water
525 percolation. Irrigation mainly results in the increase of SWC but might not increase N
526 leaching loss because the higher SWC could enhance the denitrification rate, which
527 reduces NO_3^- -N content in root zone soil (Sexstone et al., 1985). However,
528 precipitation is more likely to increase soil-soluble N leaching into the groundwater
529 due to the increase of antecedent SWC by irrigation (Razzaghi et al., 2012). Thus,
530 precipitation exhibited higher importance than irrigation and SWC, as shown in Fig. 7.
531 No significant correlation ($p > 0.05$) was observed between SSC and GW- NO_3^- -N (Fig.
532 8o). However, it should be noted that when SSC is around 20%, the GW- NO_3^- -N was
533 obviously higher; that is, the GW- NO_3^- -N decreased with SSC when the SSC was less
534 than 60%. In contrast, the GW- NO_3^- -N showed a slight increase when the SSC was
535 higher than 60%. The SSC can influence SWC and the subsequent N leaching in soil.
536 Less sand content might be associated with higher antecedent SWC and more
537 preferential flow (Van Es et al., 2004; Razzaghi et al., 2012). Thus, the N leaching
538 might be increase with the decrease of SSC. On the other hand, the nitrification rate
539 increased when the SSC was higher than 60% in sandy loam soil. Thus, the soil

540 NO_3^- -N content and leaching correspondingly increased. A significant positive
541 correlation ($r = 0.79$, $p < 0.01$) was identified between SOC and GW- NO_3^- -N (Fig. 8f).
542 This could be due to the higher SOC, which usually associated with higher SON in
543 soil and corresponds to more N leaching. Nevertheless, when SOC was below 15 g
544 kg^{-1} , it had a minor impact on GW- NO_3^- -N.

545 3.6.3 Groundwater environmental conditions

546 The DO and ORP are indicators of redox status in the groundwater environment.
547 In this study, DO usually ranged from 4 to 20 mg L^{-1} . It is generally believed that the
548 appropriate DO concentration for denitrification is 2 mg L^{-1} (Peng et al., 2020).
549 Denitrification still exists when the DO concentration of groundwater is 2~6 mg L^{-1} ,
550 but the rate is reduced with higher DO (Peng et al., 2020). The average monthly DO
551 concentration in the study area from May 2017 to April 2019 was 11.78 mg L^{-1} , and
552 there was no record of below 2 mg L^{-1} values in any of the samples. Instead, the DO
553 values in 14.6% of the 336 samples were between about 2 and 6 mg L^{-1} , mainly in
554 April and August–October. Overall, a significant positive correlation ($r = 0.96$,
555 $p < 0.01$) was identified between GW-DO and GW- NO_3^- -N (Fig. 8m), which could be
556 due to nitrification. Meanwhile a negative correlation was identified when DO was
557 less than 10 mg L^{-1} (Fig. 8m). This could be due to stronger denitrification than
558 nitrification. An increase of GW- NO_3^- -N with ORP is observed when ORP is above
559 200 mV. This could be due to the positive correlation between ORP and actual
560 nitrification (Bohrerova et al., 2004). No significant correlation ($p > 0.05$) was found
561 between GW-ORP and GW- NO_3^- -N (Fig. 8k). When ORP was less than -100 mV, it
562 is notable that there was a negative correlation, while a positive correlation was
563 identified when ORP was great than 100 mV. Nevertheless, when the ORP was
564 between -100 mV and 100 mV, it had a minor impact on GW- NO_3^- -N. According to

565 the relationships between GW-NO₃⁻-N and DO and between GW-NO₃⁻-N and ORP
566 based on the partial dependence plots of the random forest model (Fig. 8), the
567 environment of groundwater can be classified as being under nitrate-reducing
568 conditions if the DO is less than 10 mg L⁻¹ and the ORP is less than -100 mV. On the
569 other hand, the groundwater environment can be described as being and under
570 nonreducing conditions if the DO is greater than 10 mg L⁻¹ and the ORP is greater
571 than 100 mV. This is in accordance with previous studies (Rivett et al., 2008; Jahangir
572 et al., 2017; Thayalakumaran et al., 2008). Groundwater DO and ORP are shown to
573 be pH dependent. Thus, the nitrification and denitrification rates could vary with
574 changes in pH (Bohrerova et al., 2004). The GW-pH showed a significantly negative
575 correlation with GW-NO₃⁻-N ($r = -0.28$, $p < 0.05$) (Fig. 8q). The groundwater NO₃⁻-N
576 decreased with an increasing trend of pH from 6.0 to 7.0, which could be attributed to
577 denitrification. Meanwhile groundwater NO₃⁻-N increased with pH from 7.0 to 8.0,
578 which could be attributed to nitrification. The results of this study are consistent with
579 previous studies, which reported that the highest denitrification removal of NO₃⁻-N
580 was at a pH of 6.0–7.5 and the most adapted pH range for nitrifying bacteria is 7.0 to
581 9.0 (Ghafari et al., 2009; Hu et al., 2018; Bergamasco et al., 2019). A drastic increase
582 of NO₃⁻-N is observed when pH was above 8.0. This could be attributed to soil water
583 percolation to shallow groundwater following fertilization and irrigation during the
584 rainy season (Jendia et al., 2020).

585 The DOC is considered as the source of electron donors for denitrification
586 (Thayalakumaran et al., 2008). Overall, there was no significant correlation ($r = -0.05$,
587 $p > 0.05$) observed between GW-DOC and GW-NO₃⁻-N (Fig. 8p). A negative
588 correlation was found when DOC was below 75 mg L⁻¹, while a positive correlation
589 was found when DOC was above 75 mg L⁻¹. A concentration value of DOC > 1.0 mg

590 L⁻¹ can indicate the presence of electron donors significant enough to support
591 denitrification (Rivett et al., 2008; Thayalakumaran et al., 2008). The DOC
592 concentrations of the 336 samples observed in this study ranged from 0.42 to 280.26
593 mg L⁻¹, with more than 99% of samples exceeding 1.0 mg L⁻¹. The negative
594 relationship between GW-NO₃⁻-N and DOC indicates the role of DOC as a carbon
595 source in the reduction processes (Rivas et al., 2017). The other electron donors, such
596 as Fe²⁺, Mn²⁺, and S²⁻, might also exist in groundwater and influence N
597 transformation, but their roles are not dependent on temperature, DO, and ORP (Pang
598 and Wang, 2021). A significant positive correlation (r = 0.94, p<0.01) was found
599 between GW-TDS and GW-NO₃⁻-N (Fig. 8j). The GW-TDS was usually higher than
600 0.5 g L⁻¹; however, when it was lower than this value, there was a negative correlation
601 between them. The other electron donors, rather than DOC, might have played critical
602 roles when TDS was below 0.50 g L⁻¹, causing the negative correlation between
603 GW-NO₃⁻-N and TDS. However, when it was higher than 0.5 g L⁻¹, a positive
604 correlation was found between NO₃⁻-N and TDS. This could be because the growth of
605 denitrification bacteria is reduced with the increase of salinity. A previous study
606 reported that denitrification was more sensitive to salt compared to nitrification
607 (Dinçer and Kargi, 1999). Besides, it could also be because the groundwater TDS was
608 increased by a large amount of NO₃⁻-N leaching.

609 The groundwater temperature controls microbial activity. Overall, a positive
610 correlation (r = 0.62, p<0.01) was found between GW-Temp and GW-NO₃⁻-N (Fig.
611 8e), which could be due to higher N leaching to groundwater in the summer period.
612 When the GW-Temp was below 20°C, a negative correlation was found between
613 GW-Temp and GW-NO₃⁻-N. This could be because the increased temperature

614 enhances the growth and activity of denitrifiers, thus, reducing NO_3^- -N in
615 groundwater (Singh et al., 2010). The groundwater temperature also impacts other
616 environmental factors, such as DO and DOC availability (Thayalakumaran et al.,
617 2008; Nikolenko et al., 2018). The DOC can be stimulated, and the DO can be
618 depleted by increased groundwater temperature (Guo et al., 2017). The suitable
619 temperature for the nitrification rate is 10°C – 35°C , with the optimum temperature
620 ranging between 25°C – 30°C (Hayatsu and Kosuge, 1993). In contrast, the suitable
621 temperature for denitrification is 20°C – 43°C (Strous et al., 1999). In general, higher
622 temperatures result in stronger denitrification and nitrification activity in the
623 groundwater environment with temperatures less than 25°C . The optimum
624 temperature for denitrification is relatively higher than that for nitrification
625 (Nikolenko et al., 2018).

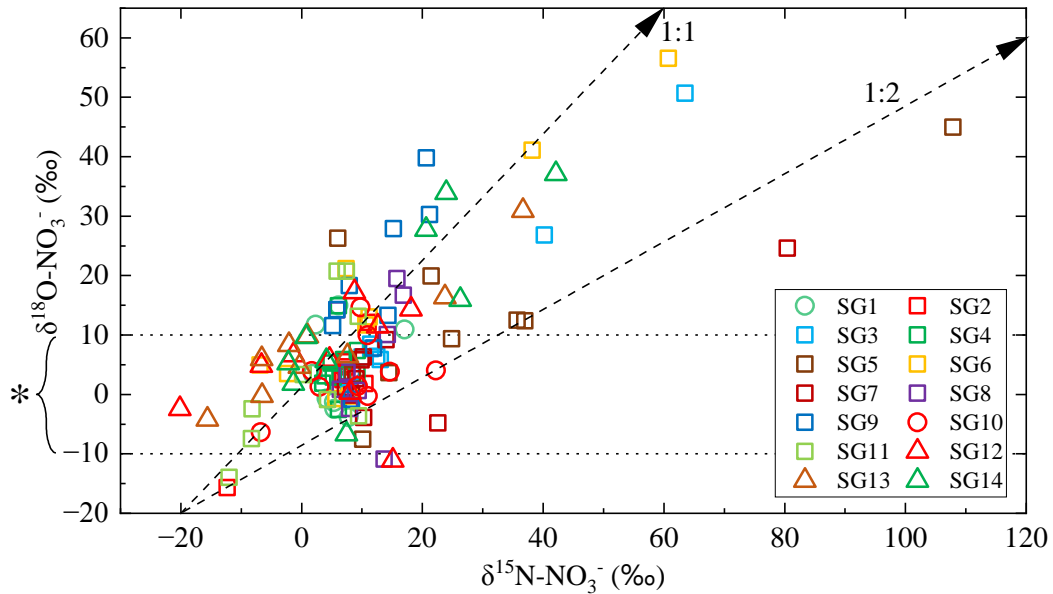
626 Although the relative importance of groundwater environmental factors (i.e.,
627 temperature, DO, ORP, DOC, TDS, and pH) to groundwater NO_3^- -N were less than
628 5.0% (Fig. 7), the partial dependence plots generated by the random forest model
629 suggested that they play important roles in influencing nitrification and/or
630 denitrification in groundwater (Fig. 8). The groundwater NO_3^- -N, $\delta^{15}\text{N}$ - NO_3^- and
631 $\delta^{18}\text{O}$ - NO_3^- exhibited tremendous changes in the summer periods in this study,
632 probably due to simultaneous leaching and denitrification (Figs. 2–4). The
633 nitrification rate also increases with an increase in temperature. In November, January,
634 and February, when the temperature is low, the DO and ORP increase, and less DOC

635 is consumed, indicating a more oxidizing condition. During these months, denitrifying
636 bacteria growth might be limited by the low temperature (Nikolenko et al., 2018).
637 Thus, the denitrification is low, and nitrification might superimpose on denitrification
638 during these low-temperature periods.

639 *3.7. Identification of key N transformation processes in groundwater using isotopes*

640 The simultaneous increase of $\delta^{15}\text{N-NO}_3^-$ and $\delta^{18}\text{O-NO}_3^-$ with a slope of 0.5-1.0
641 ($\delta^{18}\text{O}/\delta^{15}\text{N}$) is attributed to denitrification (Osaka et al., 2010; Wells et al., 2016) (Fig.
642 9). If strong denitrification occurs in groundwater, the slope of the linear relationship
643 (k) between $\delta^{15}\text{N-NO}_3^-$ and $\delta^{18}\text{O-NO}_3^-$ is close to 0.5, and $\delta^{15}\text{N-NO}_3^-$ in groundwater is
644 significantly negatively correlated with $\ln[\text{NO}_3^-]$ (Zhang et al., 2019). Among the 14
645 observation sites in the study area, the linear relationship slope (k) of $\delta^{15}\text{N-NO}_3^-$ and
646 $\delta^{18}\text{O-NO}_3^-$ in 8 sites, that is, SG1 (k = 0.98, $R^2 = 0.85$), SG3 (k = 0.83, $R^2 = 0.98$),
647 SG4 (k = 0.50, $R^2 = 0.047$), SG6 (k = 0.87, $R^2 = 0.87$), SG8 (k = 0.81, $R^2 = 0.12$),
648 SG13 (k = 0.59, $R^2 = 90$), SG14 (k = 0.82, $R^2 = 0.70$), and SG11 (k = 1.05, $R^2 = 0.50$),
649 ranged from 0.5 to 1.0 and with high $\delta^{15}\text{N-NO}_3^-$ composition, strongly supporting the
650 occurrence of denitrification in the NO_3^- -N source and/or transformation in
651 groundwater (Fig. 9). The k value at SG7 (0.30, $R^2 = 0.64$), SG5 (0.44, $R^2 = 0.84$),
652 and SG10 (0.34, $R^2 = 0.22$) did not fall into the denitrification line. However,
653 relatively higher $\delta^{15}\text{N-NO}_3^-$ and $\delta^{18}\text{O-NO}_3^-$ than the potential NO_3^- -N sources
654 indicated that weak denitrification still existed. The $\delta^{15}\text{N-NO}_3^-$ and $\delta^{18}\text{O-NO}_3^-$ at SG1,
655 SG4, SG10, SG11, and SG12 were relatively smaller, indicating weaker
656 denitrification. In this study, the groundwater $\delta^{15}\text{N-NO}_3^-$ was barely found to be
657 negatively correlated with $\ln[\text{NO}_3^-]$. It is likely that the isotopic fingerprint of
658 denitrification was simply masked by the pulse inputs of leached NO_3^- -N from the

659 vadose zone. In addition, it might also be because shallower groundwater has a short
 660 residence time. Thus, transport of NO_3^- -N affects the isotopic fingerprint of
 661 denitrification.



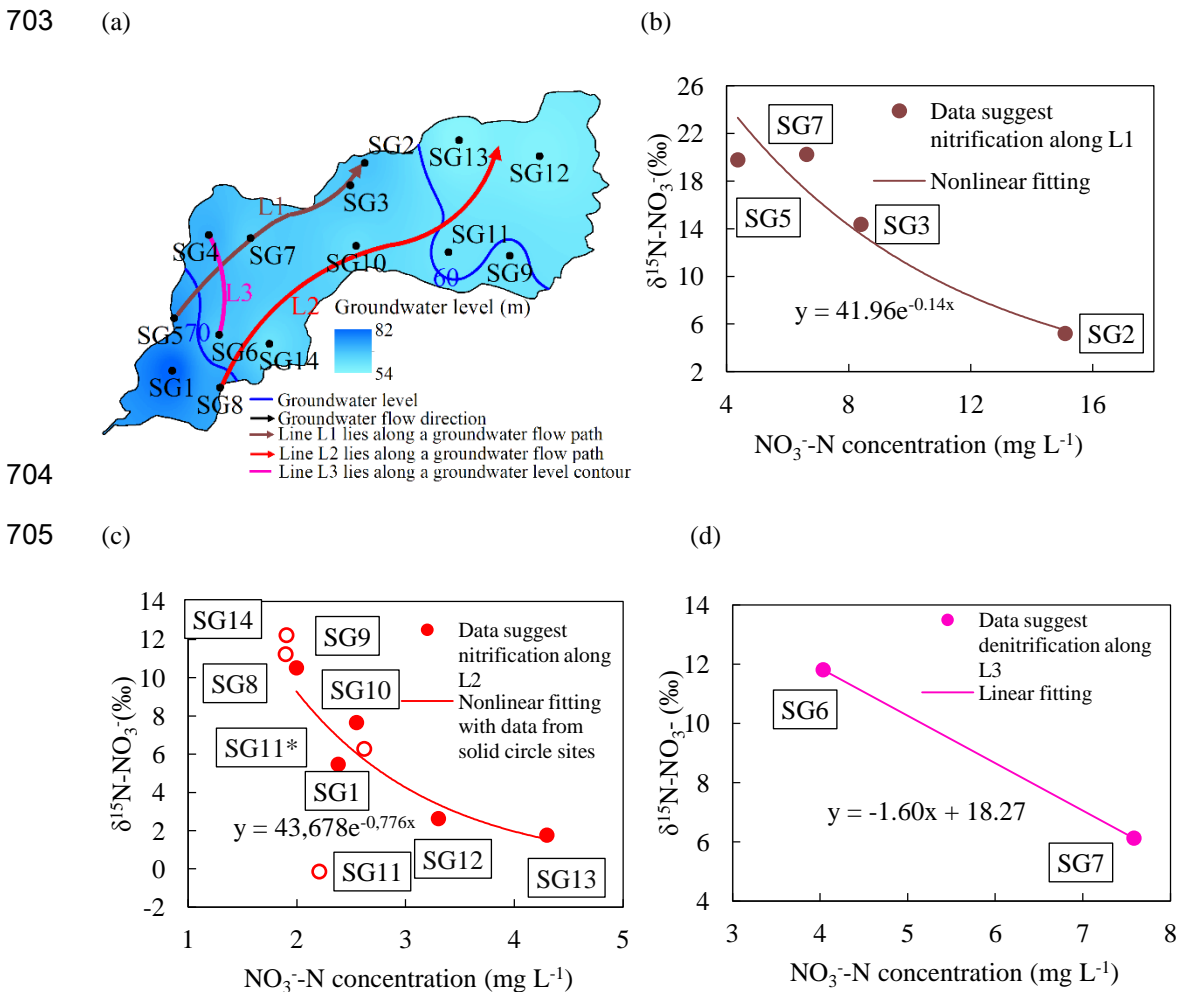
662 Fig. 9. Relationship between $\delta^{15}\text{N-NO}_3^-$ and $\delta^{18}\text{O-NO}_3^-$ in groundwater (the * indicates the
 663 range of nitrified $\delta^{18}\text{O-NO}_3^-$ within the dotted line, the 1:2 broken line indicates strong
 664 denitrification, and the 1:1 broken line indicates nitrification and denitrification)
 665

666 Nitrification will simultaneously increase NO_3^- -N concentrations and reduce the
 667 abundance of $\delta^{15}\text{N-NO}_3^-$ due to the isotope fractionation. The $\delta^{15}\text{N-NO}_3^-$ and
 668 $\delta^{18}\text{O-NO}_3^-$ vary along the 1:1 line and, with an R^2 closer to 1, indicates that the
 669 nitrification occurred concurrently with denitrification (Fig. 9). Moreover,
 670 nitrification in the groundwater environment resulted in the final $\delta^{18}\text{O-NO}_3^-$ values
 671 being mostly between -10‰ and 10‰ (Xue et al., 2009; Ji et al., 2021). In this study,
 672 the $\delta^{18}\text{O-NO}_3^-$ values of 64.28% of the 182 measured data were between -10‰ and
 673 10‰, indicating that nitrification also existed in the shallow groundwater of the study
 674 area. The NO_3^- -N is expected to be mainly from nitrifying processes since $\delta^{18}\text{O-NO}_3^-$
 675 values in 80% of the samples the ranged from -5‰ to 5‰, which is fits within the
 676 range of -6.1‰–5.2‰ (Osaka et al., 2010). The relationship between the monthly

677 $\delta^{15}\text{N-NO}_3^-$ and NO_3^- -N (from September 2018 to April 2019) in the 14 sites is shown
678 in Fig. 10. Nitrification along the groundwater flow path was identified, including in
679 the flow path along L1 surrounded by sites SG5, SG7, SG3, and SG1 and in the flow
680 path L2 surrounded by sites SG8, SG10, SG11, SG12, and SG13 (Fig. 10). Data with
681 negative $\delta^{15}\text{N-NO}_3^-$ abundance in September 2018, January 2019, and April 2019
682 were excluded from SG11 because the site was irrigated and fertilized in those
683 months. The more negative $\delta^{15}\text{N-NO}_3^-$ and $\delta^{18}\text{O-NO}_3^-$ ranges measured fell within the
684 expected range for NO_3^- -N produced from urea/urine (Wells et al., 2016). Wang et al.
685 (2020) has reported that nitrification occurred in the groundwater, which is consistent
686 with results of this study.

687 Along the groundwater flow path L1, nitrification increased NO_3^- -N
688 concentrations by about 10 mg L^{-1} . Meanwhile along the groundwater flow path L2,
689 nitrification increased NO_3^- -N concentrations by about 2 mg L^{-1} . Higher nitrification
690 rates along flow path L1 were observed. This is due to the shallower and decreased
691 depth of the groundwater table along the groundwater flow direction (from 9.37 to
692 3.86 m). On the other hand, along flow path L2, the depth of the groundwater table
693 increased from 2.0 to 14.8 m. The denitrification rates along flow path L1 were also
694 stronger, as evidenced by the higher $^{15}\text{N-NO}_3^-$ at sites around flow path L1. This is
695 because the temperature and contents of electron donors (DOC and TDS) at sites
696 around flow path L1 were higher than those around flow path L2. Therefore, both
697 denitrification and nitrification were higher in shallower groundwater around flow
698 path L1 than flow path L2. Overall, based on the DO, ORP, different forms of N
699 concentrations, and nitrate isotopic composition, it can be concluded that nitrification
700 occurred concurrently with denitrification in this shallow groundwater system. The

701 co-occurrence of denitrification and nitrification found in this study is consistent with
 702 the findings from a previous study by Osaka et al. (2010).



707 Fig. 10. (a) Locations of sampling sites along the groundwater flow path lines (L1 and L2) or the
 708 groundwater-level contour line (L3) and the relationship between monthly average (from
 709 September 2018–April 2019) groundwater $\delta^{15}\text{N-NO}_3^-$ and $\text{NO}_3^- \text{-N}$ concentrations at the 14
 710 sampling sites along (b) L1, (c) L2, and (d) L3 (SG11*: data with negative $\delta^{15}\text{N-NO}_3^-$ abundance
 711 in September 2019, January 2019, and April 2019 were excluded from SG11)

712 Site SG1 was located upstream of groundwater flow with the shallowest depth of
 713 the groundwater table. However, $\text{NO}_3^- \text{-N}$ concentration at SG1 ranged between SG8
 714 and SG12 as shown in Fig. 10. This is because the irrigation water source was from
 715 the Yellow River at SG1, while groundwater was abstracted for irrigation at other

716 sites. Moreover, the groundwater at SG1 might have received NO_3^- -N from vertical
717 and lateral recharge of the Yellow River. It is notable, though, that the irrigation
718 schedule during the summer rice growth period at SG1 was different from that of the
719 other sites. High $\delta^{15}\text{N}-\text{NO}_3^-$ and low NO_3^- concentration values were found at SG14.
720 The groundwater table depth of SG14 was the deepest and was caused by
721 groundwater abstraction for paper production in factories, according to the Xinxiang
722 City Water Resources Bulletin in 2017 (XCWRB, 2017). The deep groundwater table
723 meant that NO_3^- -N might have already been denitrified before arriving in the
724 groundwater (Vidon and Hill, 2004). As for SG9, there was a river nearby where
725 surface water could recharge the groundwater. The mean DO concentration was about
726 14.6 mg L^{-1} at SG9, while it ranged from 10.2 to 12.8 mg L^{-1} at the other sites. The
727 relatively higher DO at SG9 than the other sites could support the reasoning that
728 surface water existed recharge to the groundwater. The denitrification occurred
729 underneath the riverbed, and the denitrified NO_3^- -N continuously flowed to the
730 groundwater at SG9. In addition, the DOC was lower at SG9 (41.9 mg L^{-1}) than at the
731 neighboring sites of SG11 (45.8 mg L^{-1}) and SG 12 (45.2 mg L^{-1}) because DOC was
732 consumed by denitrification underneath the riverbed. Therefore, the isotopic
733 composition and proportion of NO_3^- -N sources at SG9 differed from that of the other
734 sites. The sites SG6 and SG4 near L3 in Fig. 10 lay along the groundwater water-level
735 contour line. The value of $\delta^{15}\text{N}-\text{NO}_3^-$ at SG6 was higher than that at SG4. The
736 groundwater table depth at SG6 was slightly higher than that of SG4. In addition, the
737 concentration of DOC in the groundwater of SG6 was higher than that at SG4.
738 Therefore, the stronger denitrification at SG6 could be attributed to the sufficient
739 carbon sources in the groundwater to support denitrification. Furthermore, the deeper

740 vadose zone thickness could provide enough time for denitrification before entering
741 the groundwater.

742 **4. Implications and Limitations**

743 *4.1. Groundwater nitrate pollution risk and attenuation in agricultural areas*

744 The spatiotemporal variations of groundwater NO_3^- -N concentrations can be
745 explained by inorganic and organic N leaching and subsequent N transformation in
746 groundwater. A shallower groundwater table is more vulnerable to N contamination.
747 The deeper the vadose zone, the less likely contaminants are coming from leaching.
748 However, continuous organic and inorganic fertilizer application under the combined
749 effects of irrigation and precipitation can result in massive N accumulation in the deep
750 vadose zone, and the peak of N content in soil contentiously moves downward to
751 groundwater (Weitzman et al., 2022). The inorganic and organic N in groundwater
752 located in a region with both a shallower groundwater table and high groundwater
753 level can be transported to a lower groundwater water-level region. In addition, the
754 groundwater is undergoing nitrification along the groundwater-level gradient. Thus,
755 slow DON mineralization and nitrification might continuously increase NO_3^- -N along
756 the groundwater flow path, threatening groundwater quality for a long time. Therefore,
757 the soil DON and NH_4^+ -N leaching and the subsequent transformation in groundwater
758 should not be overlooked when evaluating groundwater N pollution and making
759 nonpoint source control policies. In our research, the groundwater flow could export
760 the N out of the study area, thus, contributing to the temporal decrease of NO_3^- -N in
761 the local groundwater while also posing as a groundwater pollution threat to
762 neighboring areas.

763 The rapid decrease of NO_3^- -N during May–August indicated that the leached
764 NO_3^- -N was consumed immediately by denitrification. The temporal increases of

765 NO_3^- -N were observed from September 2017 to January 2018 and from January 2019
766 to March 2019. In addition, the increase of NO_3^- -N occurred along the groundwater
767 flow during September 2018–April 2019. These results suggested that the NO_3^- -N
768 accumulation might occur under low groundwater temperature conditions. This is
769 mainly due to the denitrification rate decreasing at lower temperatures in groundwater
770 (Singh et al., 2010). Therefore, sufficient DOC must be provided for the efficient
771 removal of groundwater NO_3^- -N when the groundwater temperature is suitable or
772 optimum for denitrification, particularly in late spring, summer, and early autumn. In
773 the future, quantifying nitrification and denitrification rates at different time scales
774 (e.g., daily, monthly, seasonally, or annually) is necessary to put forward N
775 attenuation measures. Agronomists recommend the application of organic fertilizer
776 and crop residue return with reduced chemical fertilizer in China and around the
777 world to increase soil quality (Zhao et al., 2016). Management of organic fertilizers
778 and crop residue is as crucial as chemical fertilizer application for reducing the risk of
779 SON accumulation and subsequent groundwater and surface water pollution while
780 providing enough DOC for denitrification. Furthermore, scientists and policy makers
781 should pay attention to the leaching risk of stocked N and evaluate its transport delay
782 time to protect groundwater quality in agricultural areas.

783 *4.2. Implications of groundwater nitrate prediction and limitations of this study*

784 The performance of the random forest model demonstrates that it can be well
785 applied for predicting spatial and intra-annual variations of groundwater nitrate
786 concentrations. The importance analysis for the influencing factors likewise provides
787 strong evidence for determining the main influencing factors and processes of NO_3^- -N
788 in groundwater. In the NCP irrigation area, the effect of N transformation in

789 groundwater is as important as soil N leaching on the variations of NO_3^- -N in shallow
790 groundwater. The findings of this study can provide technical support for the rapid
791 prediction and evaluation of N pollution in shallow groundwater through readily
792 available effect factors at higher spatial and temporal resolution. The simplified
793 approach of the random forest model can be applied in irrigation regions around the
794 world to predict the spatial and temporal variations of NO_3^- -N in groundwater. Unlike
795 NO_3^- -N, adsorption/desorption processes dominate DON and NH_4^+ -N transport in the
796 soil-groundwater systems (Vandenbruwane et al., 2007; Wang et al., 2021). Therefore,
797 other factors related to adsorption/desorption processes are required to predict
798 spatiotemporal variations of DON and NH_4^+ -N concentrations in groundwater using
799 machine learning models (Wang et al., 2016; Wang et al., 2021).

800 The N leaching and transformation in groundwater, especially denitrification and
801 nitrification, must be considered in other numerical and/or distributed groundwater
802 flow models to accurately predict NO_3^- -N dynamics. The NO_3^- -N in groundwater
803 tends to be the most sensitive to NO_2^- -N, indicating that biotic factors play critical
804 roles in controlling the N transformation (Yang et al., 2012). These biochemical
805 processes cause the difficulty and challenge of building and verifying groundwater
806 flow and solute transport models. It is essential to study microbial communities and
807 growth response to changes in N and C substrates and environmental factors to gain
808 better insights into N transformation processes and rates (Nikolenko et al., 2018). It
809 should be noted that except for the commonly considered nitrification and

810 denitrification processes identified in this study, other N transformation processes in
811 shallow groundwater are also possible. For instance, NO_2^- -N accumulation in
812 denitrification can provide the substrate for anammox (Smith et al., 2015), DNRA
813 (Jahangir et al., 2017), and shortcut nitrification-denitrification (Hu et al., 2018).
814 Therefore, it is necessary to measure the N transformation rates using the ^{15}N tracing
815 method, and analyze microbial communities based on in situ experiments and
816 laboratory simulation of the groundwater environment to precisely understand the fate
817 of N in groundwater in future studies.

818 **5. Conclusions**

819 Nitrate as the primary source of inorganic N pollutants in shallow groundwater,
820 showed highly variable spatiotemporal patterns in an irrigated agricultural area, the
821 NCP. Monthly soil physiochemical data, groundwater quality data, and groundwater
822 NO_3^- -N isotopic composition data were measured and analyzed from May 2017 to
823 April 2019 in the study area. This research focused on predicting spatial and
824 intra-annual variations of shallow groundwater NO_3^- -N and evaluating their main
825 effect factors and processes. The random forest model, as a simplified approach
826 compared to numerical and distributed hydrological models, performed well in
827 predicting the spatial and intra-annual variations of NO_3^- -N in groundwater with R^2
828 value of 0.93 and 0.92, RMSE value of 4.94 and 3.87, and MAE values of 2.10 and
829 2.89 for the training and test datasets, respectively. The application of the random
830 forest model may have important implications for groundwater NO_3^- -N prediction in
831 other regions worldwide. Evaluation of the importance of the effect factors
832 emphasized the roles of N transformation, including denitrification and nitrification,

833 in controlling spatiotemporal variations of NO_3^- -N in addition to N leaching. The
834 impacts of soil SON on groundwater NO_3^- -N was the greatest among different forms
835 of N in surface soil. The temporal increase of NO_3^- -N was mainly attributed to
836 NO_3^- -N leaching, while the temporal decrease of NO_3^- -N can be attributed to
837 denitrification. Although the N leaching usually decreases with the increase of
838 groundwater table depth, nitrification along the groundwater flow path will adversely
839 affect groundwater quality in deep vadose zone areas. Monitoring and management of
840 SON in soil and DON in groundwater will become increasingly important since
841 agronomists and policy makers widely recommend the use of manure instead of
842 chemical fertilizers in developing and developed countries. The groundwater NO_3^- -N
843 is significantly related to nitrite which is an indicator of biotic factors, but
844 groundwater NO_3^- -N is less sensitive to other environmental factors. The prediction of
845 groundwater NO_3^- -N dynamics using hydrological or biochemical models should
846 address all forms of N leaching and transformation in groundwater, which is still a
847 considerable challenge due to the involvement of biological factors. The ability to
848 identify the N transformation process and quantify the N transformation rates is
849 urgently needed and deserves further attention through the adoption of
850 microbiological techniques and ^{15}N tracing methods based on in situ and laboratory
851 groundwater experiments.

852 **Declaration of Competing Interest:** The authors declare no conflict of interest.

853 **Data availability:** The data analyzed and presented in this study are openly available.

854 **Acknowledgment:** This research was funded by the National Natural Science
855 Foundation of China (No. 51609084). The authors sincerely thank the anonymous
856 reviewers for their insightful comments.

857 **References**

- 858 [1] Amit Y, Geman D. 1997. Shape quantization and recognition with randomized
859 trees. *Neural Comput.*; 9: 1545-1588. <https://doi.org/10.1162/neco.1997.9.7.1545>
- 860 [2] Awais M, Aslam B, Maqsoom A, Khalil U, Ullah F, Azam S, et al. 2021.
861 Assessing nitrate contamination risks in groundwater: A machine learning
862 approach. *Appl. Sci.*; 11: 10034. <https://doi.org/10.3390/app112110034>
- 863 [3] Bagherzadeh F, Mehrani M-J, Basirifard M, Roostaei J. 2021. Comparative study
864 on total nitrogen prediction in wastewater treatment plant and effect of various
865 feature selection methods on machine learning algorithms performance. *J. Water*
866 *Process. Eng.*; 41: 102033. <https://doi.org/10.1016/j.jwpe.2021.102033>
- 867 [4] Band SS, Janizadeh S, Pal SC, Chowdhuri I, Siabi Z, Norouzi A, et al. 2020.
868 Comparative analysis of artificial intelligence models for accurate estimation of
869 groundwater nitrate concentration. *Sensors*; 20.
870 <https://doi.org/10.3390/s20205763>
- 871 [5] Bergamasco MAM, Braos LB, Guidini Lopes I, Cruz MCP. 2019. Nitrogen
872 mineralization and nitrification in two soils with different pH levels. *Commun.*
873 *Soil Sci. Plant Anal.*; 50: 2873-2880.
874 <https://doi.org/10.1080/00103624.2019.1689250>
- 875 [6] Biddau R, Cidu R, Da Pelo S, Carletti A, Ghiglieri G, Pittalis D. 2019. Source
876 and fate of nitrate in contaminated groundwater systems: Assessing spatial and
877 temporal variations by hydrogeochemistry and multiple stable isotope tools. *Sci.*
878 *Total Environ.*; 647: 1121-1136. <https://doi.org/10.1016/j.scitotenv.2018.08.007>
- 879 [7] Bohrerova Z, Stralkova R, Podesvova J, Bohrer G, Pokorny E. 2004. The
880 relationship between redox potential and nitrification under different sequences of
881 crop rotations. *Soil Tillage Res.*; 77: 25-33.
882 <https://doi.org/10.1016/j.still.2003.10.006>
- 883 [8] Breiman L. 2001. Random Forests. *Mach. Learn.*; 45: 5-32.
884 <https://doi.org/10.1023/A:1010933404324>
- 885 [9] Castrillo M, García ÁL. 2020. Estimation of high frequency nutrient
886 concentrations from water quality surrogates using machine learning methods.
887 *Water Res.*; 172: 115490. <https://doi.org/10.1016/j.watres.2020.115490>
- 888 [10] Chlingaryan A, Sukkarieh S, Whelan B. 2018. Machine learning approaches for
889 crop yield prediction and nitrogen status estimation in precision agriculture: A
890 review. *Comput. Electron. Agric.*; 151: 61-69.
891 <https://doi.org/10.1016/j.compag.2018.05.012>
- 892 [11] Coppola Jr EA, Rana AJ, Poulton MM, Szidarovszky F, Uhl VW. 2005. A neural
893 network model for predicting aquifer water level elevations. *GROUNDWATER*;
894 43: 231-241. <https://doi.org/10.1111/j.1745-6584.2005.0003.x>
- 895 [12] Dinçer AR, Kargi F. 1999. Salt inhibition of nitrification and denitrification in
896 saline wastewater. *Environ. Technol.*; 20: 1147-1153.
897 <https://doi.org/10.1080/09593332008616912>

- 898 [13]Du R, Peng Y, Cao S, Li B, Wang S, Niu M. 2016. Mechanisms and microbial
899 structure of partial denitrification with high nitrite accumulation. *Appl. Microbiol.*
900 *Biotechnol.*; 100: 2011-2021. <https://doi.org/10.1007/s00253-015-7052-9>
- 901 [14]El Amri A, M'Nassri S, Nasri N, Nsir H, Majdoub R. 2022. Nitrate concentration
902 analysis and prediction in a shallow aquifer in central-eastern Tunisia using
903 artificial neural network and time series modelling. *Environ. Sci. Pollut. Res. Int.*;
904 29: 43300-43318. <https://doi.org/10.1007/s11356-021-18174-y>
- 905 [15]Famiglietti JS, Ferguson G. 2021. The hidden crisis beneath our feet. *Science*;
906 372: 344-345. <https://doi.org/10.1126/science.abh2867>
- 907 [16]Foulquier A, Malard F, Mermillod-Blondin F, Montuelle B, Dolédec S, Volat B,
908 et al. 2011. Surface water linkages regulate trophic interactions in a groundwater
909 food web. *Ecosystems*; 14: 1339-1353. <http://doi.org/10.1007/s10021-011-9484-0>
- 910 [17]Gan L, Huang G, Pei L, Gan Y, Liu C, Yang M, et al. 2022. Distributions, origins,
911 and health-risk assessment of nitrate in groundwater in typical alluvial-pluvial
912 fans, North China Plain. *Environ. Sci. Pollut. Res. Int.*; 29: 17031-17048.
913 <https://doi.org/10.1007/s11356-021-17067-4>
- 914 [18]Gao Z, Han C, Yuan S, Liu J, Peng Y, Li C. 2022. Assessment of the
915 hydrochemistry, water quality, and human health risk of groundwater in the
916 northwest of Nansi Lake Catchment, north China. *Environ. Geochem. Health*; 44:
917 961-977. <https://doi.org/10.1007/s10653-021-01011-z>
- 918 [19]Guo H, Ye C, Zhang H, Pan S, Ji Y, Li Z, et al. 2017. Long-term nitrogen &
919 phosphorus additions reduce soil microbial respiration but increase its
920 temperature sensitivity in a Tibetan alpine meadow. *Soil Biol. Biochem.*; 113:
921 26-34. <https://doi.org/10.1016/j.soilbio.2017.05.024>
- 922 [20]Hayatsu M, Kosuge N. 1993. Autotrophic nitrification in acid tea soils. *Soil Sci.*
923 *Plant Nutr.*; 39: 209-217. <https://doi.org/10.1080/00380768.1993.10416992>
- 924 [21]He S, Wu J, Wang D, He X. 2022. Predictive modeling of groundwater nitrate
925 pollution and evaluating its main impact factors using random forest.
926 *Chemosphere*; 290: 133388. <https://doi.org/10.1016/j.chemosphere.2021.133388>
- 927 [22]Hinkle SR, Tesoriero AJ. 2014. Nitrogen speciation and trends, and prediction of
928 denitrification extent, in shallow US groundwater. *J. Hydrol.*; 509: 343-353.
929 <https://doi.org/10.1016/j.jhydrol.2013.11.048>
- 930 [23]Hood-Nowotny R, Umana NH-N, Inselbacher E, Oswald- Lachouani P, Wanek W.
931 2010. Alternative methods for measuring inorganic, organic, and total dissolved
932 nitrogen in soil. *Soil Sci. Soc. Am. J.*; 74: 1018-1027.
933 <https://doi.org/10.2136/sssaj2009.0389>
- 934 [24]Hu W, Zhou Y, Min X, Liu J, Li X, Luo L, et al. 2018. The study of a pilot-scale
935 aerobic/Fenton/anoxic/aerobic process system for the treatment of landfill
936 leachate. *Environ. Technol.*; 39: 1926-1936.
937 <https://doi.org/10.1080/09593330.2017.1344325>
- 938 [25]Jahangir MMR, Fenton O, Müller C, Harrington R, Johnston P, Richards KG.
939 2017. In situ denitrification and DNRA rates in groundwater beneath an
940 integrated constructed wetland. *Water Res.*; 111: 254-264.

- 941 <https://doi.org/10.1016/j.watres.2017.01.015>
- 942 [26]Jendia AH, Hamzah S, Abuhabib AA, El-Ashgar NM. 2020. Removal of nitrate
943 from groundwater by eggshell biowaste. *Water Supply*; 20: 2514-2529.
944 <http://doi.org/10.2166/ws.2020.151>
- 945 [27]Ji W, Xiao J, Toor GS, Li Z. 2021. Nitrate-nitrogen transport in streamwater and
946 groundwater in a loess covered region: Sources, drivers, and spatiotemporal
947 variation. *Sci. Total Environ.*; 761: 143278.
948 <https://doi.org/10.1016/j.scitotenv.2020.143278>
- 949 [28]Knoll L, Breuer L, Bach M. 2019. Large scale prediction of groundwater nitrate
950 concentrations from spatial data using machine learning. *Sci. Total Environ.*; 668:
951 1317-1327.<https://doi.org/10.1016/j.scitotenv.2019.03.045>
- 952 [29]Lasserre F, Razack M, Banton O. 1999. A GIS-linked model for the assessment of
953 nitrate contamination in groundwater. *J. Hydrol.*; 224: 81-90.
954 [https://doi.org/10.1016/S0022-1694\(99\)00130-4](https://doi.org/10.1016/S0022-1694(99)00130-4)
- 955 [30]Li Z, Zhang Q, Qiao Y, Leng P, Zhang Q, Du K, et al. 2021. Influence of the
956 shallow groundwater table on the groundwater N₂O and direct N₂O emissions in
957 summer maize field in the North China Plain. *Sci. Total Environ.*; 799: 149495.
958 <https://doi.org/10.1016/j.scitotenv.2021.149495>
- 959 [31]Liu J, Liu Q, Yang H. 2016. Assessing water scarcity by simultaneously
960 considering environmental flow requirements, water quantity, and water quality.
961 *Ecol. Indic.*; 60: 434-441. <https://doi.org/10.1016/j.ecolind.2015.07.019>
- 962 [32]Liu Y, Xin J, Wang Y, Yang Z, Liu S, Zheng X. 2022. Dual roles of dissolved
963 organic nitrogen in groundwater nitrogen cycling: Nitrate precursor and
964 denitrification promoter. *Sci. Total Environ.*; 811: 151375.
965 <https://doi.org/10.1016/j.scitotenv.2021.151375>
- 966 [33]Nikolenko O, Jurado A, Borges AV, Knöller K, Brouyère S. 2018. Isotopic
967 composition of nitrogen species in groundwater under agricultural areas: A review.
968 *Sci. Total Environ.*; 621: 1415-1432.
969 <https://doi.org/10.1016/j.scitotenv.2017.10.086>
- 970 [34]Osaka Ki, Ohte N, Koba K, Yoshimizu C, Katsuyama M, Tani M, et al. 2010.
971 Hydrological influences on spatiotemporal variations of $\delta^{15}\text{N}$ and $\delta^{18}\text{O}$ of nitrate
972 in a forested headwater catchment in central Japan: Denitrification plays a critical
973 role in groundwater. *Journal of Geophysical Research: Biogeosciences*; 115.
974 <https://doi.org/10.1029/2009JG000977>
- 975 [35]Ouedraogo I, Defourny P, Vanclooster M. 2019. Validating a continental-scale
976 groundwater diffuse pollution model using regional datasets. *Environ. Sci. Pollut.*
977 *Res. Int.*; 26: 2105-2119. <https://doi.org/10.1007/s11356-017-0899-9>
- 978 [36]Pang Y, Wang J. 2021. Various electron donors for biological nitrate removal: A
979 review. *Sci. Total Environ.*; 794: 148699.
980 <https://doi.org/10.1016/j.scitotenv.2021.148699>
- 981 [37]Peng B, Liang H, Wang S, Gao D. 2020. Effects of DO on N₂O emission during
982 biological nitrogen removal using aerobic granular sludge via shortcut
983 simultaneous nitrification and denitrification. *Environ. Technol.*; 41: 251-259.

- 984 <https://doi.org/10.1080/09593330.2018.1494757>
- 985 [38] Pennino MJ, Leibowitz SG, Compton JE, Hill RA, Sabo RD. 2020. Patterns and
986 predictions of drinking water nitrate violations across the conterminous United
987 States. *Sci. Total Environ.*; 722: 137661.
988 <https://doi.org/10.1016/j.scitotenv.2020.137661>
- 989 [39] Razzaghi F, Plauborg F, Jacobsen S-E, Jensen CR, Andersen MN. 2012. Effect of
990 nitrogen and water availability of three soil types on yield, radiation use
991 efficiency and evapotranspiration in field-grown quinoa. *Agric. Water Manag.*;
992 109: 20-29. <https://doi.org/10.1016/j.agwat.2012.02.002>
- 993 [40] Rivas A, Singh R, Horne D, Roygard J, Matthews A, Hedley MJ. 2017.
994 Denitrification potential in the subsurface environment in the Manawatu River
995 catchment, New Zealand: Indications from oxidation-reduction conditions,
996 hydrogeological factors, and implications for nutrient management. *J. Environ.*
997 *Manag.*; 197: 476-489. <https://doi.org/10.1016/j.jenvman.2017.04.015>
- 998 [41] Rivett MO, Buss SR, Morgan P, Smith JWN, Bemment CD. 2008. Nitrate
999 attenuation in groundwater: A review of biogeochemical controlling processes.
1000 *Water Res.*; 42: 4215-4232. <https://doi.org/10.1016/j.watres.2008.07.020>
- 1001 [42] Sexstone AJ, Parkin TB, Tiedje JM. 1985. Temporal response of soil
1002 denitrification rates to rainfall and irrigation. *Soil Sci. Soc. Am. J.*; 49: 99-103.
1003 <https://doi.org/10.2136/sssaj1985.03615995004900010020x>
- 1004 [43] Shen Z, Xin J, Wu H, Jiang Z, Peng H, Xu F, et al. 2023. Kinetic and molecular
1005 evidence for DON transformation in the deep vadose zone: Important
1006 implications for soil nitrogen budgeting and groundwater nitrate management. *J.*
1007 *Hydrol.*; 616: 128782. <https://doi.org/10.1016/j.jhydrol.2022.128782>
- 1008 [44] Singh BK, Bardgett RD, Smith P, Reay DS. 2010. Microorganisms and climate
1009 change: terrestrial feedbacks and mitigation options. *Nat. Rev. Microbiol.*; 8:
1010 779-790. <http://doi.org/10.1038/nrmicro2439>
- 1011 [45] Smarra F, Jain A, de Rubeis T, Ambrosini D, D'Innocenzo A, Mangharam R.
1012 2018. Data-driven model predictive control using random forests for building
1013 energy optimization and climate control. *Appl. Energy*; 226: 1252-1272.
1014 <https://doi.org/10.1016/j.apenergy.2018.02.126>
- 1015 [46] Smith RL, Böhlke JK, Song B, Tobias CR. 2015. Role of anaerobic ammonium
1016 oxidation (anammox) in nitrogen removal from a freshwater aquifer. *Environ. Sci.*
1017 *Technol.*; 49: 12169-12177. <https://doi.org/10.1021/acs.est.5b02488>
- 1018 [47] Spoelstra J, Schiff S, Elgood R, Semkin R, Jeffries D. 2001. Tracing the sources
1019 of exported nitrate in the Turkey Lakes Watershed using $^{15}\text{N}/^{14}\text{N}$ and $^{18}\text{O}/^{16}\text{O}$
1020 isotopic ratios. *Ecosystems*; 4: 536-544.
1021 <https://doi.org/10.1007/s10021-001-0027-y>
- 1022 [48] Strous M, Kuenen JG, Jetten MS. 1999. Key physiology of anaerobic ammonium
1023 oxidation. *Appl. Environ. Microbiol.*; 65: 3248-50.
1024 <https://doi.org/10.1128/aem.65.7.3248-3250.1999>
- 1025 [49] Thayalakumaran T, Bristow KL, Charlesworth PB, Fass T. 2008. Geochemical
1026 conditions in groundwater systems: Implications for the attenuation of

1027 agricultural nitrate. *Agric. Water Manag.*; 95: 103-115.
 1028 <https://doi.org/10.1016/j.agwat.2007.09.003>

1029 [50] van Es HM, Schindelbeck RR, Jokela WE. 2004. Effect of manure application
 1030 timing, crop, and soil type on phosphorus leaching. *J. Environ. Qual.*; 33:
 1031 1070-1080. <https://doi.org/10.2134/jeq2004.1070a>

1032 [51] Vandenbruwane J, De Neve S, Qualls R G, Sleutel S, Hofman G. 2007.
 1033 Comparison of different isotherm models for dissolved organic carbon (DOC)
 1034 and nitrogen (DON) sorption to mineral soil. *Geoderma*, 139(1-2), 144-153.
 1035 <https://doi.org/10.1016/j.geoderma.2007.01.012>

1036 [52] Vidon P, Hill AR. 2004. Denitrification and patterns of electron donors and
 1037 acceptors in eight riparian zones with contrasting hydrogeology. *Biogeochemistry*;
 1038 71: 259-283. <http://doi.org/10.1007/s10533-004-9684-1>

1039 [53] Vystavna Y, Diadin D, Valeriy Y, Hejzlar J, Vadillo I, Huneau F, et al. 2017.
 1040 Nitrate contamination in a shallow urban aquifer in East Ukraine: Evidence from
 1041 hydrochemical, stable nitrate isotope, and land use analysis. *Environ. Earth Sci.*;
 1042 76. <https://doi.org/10.1007/s12665-017-6796-1>

1043 [54] Wang C, Jiang R, Boithias L, Sauvage S, Sánchez-Pérez JM, Mao X, Han Y,
 1044 Hayakawa A, Kuramochi K, Hatano R. 2016. Assessing potassium environmental
 1045 losses from a dairy farming watershed with the modified SWAT model. *Agr.*
 1046 *Water Manage.*; 175, 91-104. <https://doi.org/10.1016/j.agwat.2016.02.007>

1047 [55] Wang C, Wu D, Mao X, Hou J, Wang L, Han Y. 2021. Estimating soil ammonium
 1048 adsorption using pedotransfer functions in an irrigation district of the North
 1049 China Plain. *Pedosphere*; 31: 157-171.
 1050 [https://doi.org/10.1016/S1002-0160\(20\)60054-6](https://doi.org/10.1016/S1002-0160(20)60054-6)

1051 [56] Wang Y, Peng J, Cao X, Xu Y, Yu H, Duan G, et al. 2020. Isotopic and chemical
 1052 evidence for nitrate sources and transformation processes in a plateau lake basin
 1053 in Southwest China. *Sci. Total Environ.*; 711: 134856.
 1054 <https://doi.org/10.1016/j.scitotenv.2019.134856>

1055 [57] Weitzman JN, Brooks JR, Compton JE, Faulkner BR, Mayer PM, Peachey RE, et
 1056 al. 2022. Deep soil nitrogen storage slows nitrate leaching through the vadose
 1057 zone. *Agr Ecosyst Environ*; 332: 107949.
 1058 <https://doi.org/10.1016/j.agee.2022.107949>

1059 [58] Wells NS, Baisden WT, Horton T, Clough TJ. 2016. Spatial and temporal
 1060 variations in nitrogen export from a New Zealand pastoral catchment revealed by
 1061 stream water nitrate isotopic composition. *Water Resour. Res.*; 52: 2840-2854.
 1062 <https://doi.org/10.1002/2015WR017642>

1063 [59] Xinxiang City Water Resources Bureau (XCWRB), 2017. Xinxiang City Water
 1064 Resources Bulletin. <http://slj.xinxiang.gov.cn/ggl/1983.html>

1065 [60] Xue D, Botte J, De Baets B, Accoe F, Nestler A, Taylor P, et al. 2009. Present
 1066 limitations and future prospects of stable isotope methods for nitrate source
 1067 identification in surface- and groundwater. *Water Res.*; 43: 1159-1170.
 1068 <https://doi.org/10.1016/j.watres.2008.12.048>

- 1069 [61] Yang X, Wang S, Zhou L. 2012. Effect of carbon source, C/N ratio, nitrate and
1070 dissolved oxygen concentration on nitrite and ammonium production from
1071 denitrification process by *Pseudomonas stutzeri* D6. *Bioresour. Technol.*; 104:
1072 65-72. <https://doi.org/10.1016/j.biortech.2011.10.026>
- 1073 [62] Zhang Y, Xu B, Guo Z, Han J, Li H, Jin L, et al. 2019. Human health risk
1074 assessment of groundwater arsenic contamination in Jinghui irrigation district,
1075 China. *J. Environ. Manag.*; 237: 163-169.
1076 <https://doi.org/10.1016/j.jenvman.2019.02.067>
- 1077 [63] Zhao B, Zhang J, Yu Y, Karlen DL, Hao X. 2016. Crop residue management and
1078 fertilization effects on soil organic matter and associated biological properties.
1079 *Environ. Sci. Pollut. Res.*; 23: 17581-17591.
1080 <http://doi.org/10.1007/s11356-016-6927-3>
- 1081 [64] Zhou J, Green M, Shaviv A. 2003. Mineralization of organic N originating in
1082 treated effluent used for irrigation. *Nutr. Cycl. Agroecosyst.*; 67: 205-213.
1083 <https://doi.org/10.1023/B:FRES.00000003599.60911.a2>
- 1084