

From the Department of Landscape Ecology and Resources Management
Professorship of Landscape, Water and Biogeochemical Cycles
of the Justus Liebig University Giessen

Nationwide estimation of groundwater nitrate concentrations using machine learning

Dissertation
for the degree Doctor of Natural Science (Dr. rer. nat.)
submitted by Lukas Knoll, M. Sc.

Referees from the Justus Liebig University Giessen:

Prof. Dr. Lutz Breuer (1st Supervisor)

Prof. Dr. Jan Siemens (2nd Supervisor)

Giessen, 2020

Abstract

Nitrate pollution of groundwater has been a well-known problem for decades, but today the debate about causes and possible solutions is very controversial. The EU Water Framework Directive has been implemented to restore the good status of water systems in Europe, which requires, among other things, a reduction of nitrate inputs to groundwater. In order to achieve this, harmonized and transparent approaches for evaluating the quality of groundwater bodies and clearly defined regulations for the implementation of mitigation measures need to be established. Important aspects to be taken into account when developing action plans are the estimation of nitrogen inputs to the subsurface, the spatial distribution of groundwater nitrate concentrations and the natural nitrate reduction capacity of the aquifer. For groundwater quality modelling on a country-wide scale, data-driven approaches, in particular machine learning techniques, have proven their worth. In this dissertation, an approach for the large-scale regionalisation of groundwater nitrate concentrations depending on spatial environmental parameters is developed. In a first step, several approaches for linking point information from monitoring sites with the spatial data from maps are investigated using the example of the federal state of Hesse, Germany. Four machine learning techniques based on different statistical model types are compared regarding their predictive performance. It can be shown that a 1,000 m circular buffer can describe the contribution area of a monitoring site in a simplified way and can be used for compiling the factors influencing the groundwater nitrate concentration. The random forest model outperforms classical multiple linear regression, simple classification and regression trees and boosted regression trees. In a second step, the approach will then be transferred to the entire federal republic of Germany. Based on the Water Framework Directive groundwater monitoring network, a random forest model is trained and the nitrate concentrations in groundwater is estimated for a 1 km x 1 km grid map. Good model predictive performance can be achieved with an R^2 of 0.52 where the redox conditions, the hydrogeological units and the percentage of arable land are identified as the most influential predictors for the estimation of groundwater nitrate concentration. By using quantile random forest for an uncertainty analysis, with a mean prediction interval of 53 mg NO_3^-/l large uncertainties are determined. Finally, the third part of this dissertation focuses on the estimation of nitrate reduction in groundwater. The estimated spatial distribution of groundwater nitrate concentrations together with data on nitrogen surplus are used in a simplified conceptual approach to estimate the integrated nitrate reduction across the unsaturated zone and the groundwater body. The determined nitrate reduction rates are on average 57% and strongly vary from no reduction up to a degradation of 100% with predominantly high reduction rates in northern Germany and lower reduction rates in the central and southern part of Germany. Nitrate reduction capacity is highly dependent on hydrogeological conditions, with reduction rates in porous aquifers and under anaerobic conditions, being significantly higher than in fractured consolidated aquifers and under aerobic conditions. With the overall approach presented here, spatial predictions can be made based on freely available geodata, which makes it an important contribution to the large-scale assessment of groundwater quality and can be used in the planning of mitigation measures

Table of Contents

Abstract	i
Table of Contents	ii
List of Tables.....	v
List of Figures	vi
1 Extended Summary	1
1.1 Introduction	1
1.1.1 Nitrate pollution – a brief overview	1
1.1.2 Nitrate pollution – a complex hydro-biogeochemical issue	1
1.1.3 Nitrate pollution – the current situation.....	2
1.1.4 Mapping and Analysis on a large scale – the capability of machine learning techniques	3
1.1.5 Objectives of the Dissertation	5
1.2 Machine learning techniques for the estimation of groundwater nitrate concentrations on a large scale	6
1.3 Groundwater nitrate concentration throughout Germany.....	9
1.4 Nitrate reduction throughout Germany	13
1.5 Conclusion and Outlook.....	14
2 Large scale prediction of groundwater nitrate concentrations from spatial data using machine learning.....	17
Abstract	17
2.1 Introduction	18
2.2 Materials and Methods	20
2.2.1 Study area	20
2.2.2 Basic modelling approach	20
2.2.3 Data pre-processing.....	20
2.2.4 Statistical model framework.....	25
2.2.5 Prediction.....	27
2.3 Results and Discussion.....	28

2.3.1	Model performance	28
2.3.2	Variable importance	31
2.3.3	Spatial prediction of groundwater nitrate concentrations	32
2.4	Conclusion.....	34
	Acknowledgements	34
	Supporting Information	35
3	Nation-wide estimation of groundwater redox conditions and nitrate concentrations through machine learning	43
	Abstract	43
3.1	Introduction	44
3.2	Materials and Methods	46
3.2.1	Database	47
3.2.2	Pre-processing	48
3.2.3	Predictive modelling	49
3.2.4	Uncertainty analysis	50
3.3	Results	50
3.3.1	Spatial distribution of groundwater redox conditions	50
3.3.2	Spatial distribution of groundwater nitrate concentration	54
3.4	Discussion	56
3.5	Conclusion.....	59
	Acknowledgements	59
	Data availability	59
	Supporting Information	60
4	Spatial distribution of integrated nitrate reduction across the unsaturated zone and the groundwater body in Germany.....	77
	Abstract	77
4.1	Introduction	78
4.2	Materials and Methods	80

Table of Contents

4.2.1	Nitrogen input load.....	82
4.2.2	Groundwater nitrate nitrogen load.....	83
4.2.3	Nitrate reduction.....	83
4.2.4	Hydrogeological conditions.....	84
4.3	Results	84
4.4	Discussion	90
4.5	Conclusions	93
	Author Contributions.....	93
	Acknowledgements	93
	Supporting Information	94
5	References	97
	Acknowledgements.....	116
	Declaration.....	117

List of Tables

Table 2-1: Summary of the spatial environmental predictor variables.	23
Table 3-1: Evaluation of model performance for the RF _{O₂} , RF _{Fe} and RF _{NO₃} models based on test data, training data and all data of the WFD data set.	52
Table 4-1: Mean NO ₃ -N load reduction rates and standard deviation for different hydrogeological attributes.....	89

List of Tables in Supporting Information

Table S 2-1: Criteria for data selection.	36
Table S 2-2: Summary statistics of nitrate concentrations in groundwater for training and test data set.....	36
Table S 2-3: Methods and tuning parameters of the models MLR, CART, RF and BRT used in the ‘caret’ package.	42
Table S 2-4: Predictive performance values MAE (Mean Absolute Error), RMSE (Root Mean Squared Error), R ² (coefficient of determination) for model training and testing for the models MLR (Multiple Linear Regression), CART (Classification and Regression Tree), RF (Random Forest) and BRT (Boosted Regression Trees) based on the 1000 m circular buffer zone.	42
Table S 3-1: Summary statistics of mean groundwater concentration (O ₂ , Fe and NO ₃) [mg/l] of the period 2009-2018 for the WFD data set.....	61
Table S 3-2: Spatial environmental predictor variables.	63
Table S 3-3: Classification point scheme for the characterisation of the redox conditions (acc. to LAWA (2018), modified).	69
Table S 4-1: Post-Hoc-Test results aquifer type.....	94
Table S 4-2: Post-Hoc-Test results consolidation.	94
Table S 4-3: Post-Hoc-Test results rock type.....	94
Table S 4-4: Post-Hoc-Test results geochemical rock type.....	95
Table S 4-5: Post-Hoc-Test results conductivity.....	95
Table S 4-6: Post-Hoc-Test results redox conditions.	96

List of Figures

Figure 1-1: Schematic workflow of the predictive modelling approach.	7
Figure 1-2: Map of Germany with the federal state Hesse (left) and a north to south profile line with elevations and the hydrogeological regions (right).	10
Figure 2-1: Workflow of the GIS-based statistical modelling approach, consisting of the three steps: data pre-processing, statistical model framework and prediction.	21
Figure 2-2: Five different buffer designs for the allocation of spatial data to the groundwater monitoring sites.	24
Figure 2-3: Histogram of nitrate concentration in groundwater (mean 2010 – 2017) in mg/l for training and test data set.	27
Figure 2-4: Comparison of the mean predictive performance MAE (Mean Absolute Error), RMSE (Root Mean Squared Error), R^2 (coefficient of determination) resulting from the 3 times repeated 10-fold cross-validation for MLR (Multiple Linear Regression), CART (Classification and Regression Tree), RF (Random Forest) and BRT (Boosted Regression Trees) based on the different contributing areas (circular buffer b1,500, b1,000, b500 and wedge buffer w45, w90), respectively.	29
Figure 2-5: Comparison of predictive performance MAE (Mean Absolute Error), RMSE (Root Mean Squared Error), R^2 (coefficient of determination) of the 30 models resulting from the 3 times repeated 10-fold cross-validation for MLR (Multiple Linear Regression), CART (Classification and Regression Tree), RF (Random Forest) and BRT (Boosted Regression Trees), based on the 1,000 m circular buffer zone, respectively.	30
Figure 2-6: Predicted vs. observed nitrate concentration for test and training data for the models MLR (Multiple Linear Regression), CART (Classification and Regression Trees), RF (Random Forest) and BRT (Boosted Regression Tree) based on the 1,000 m circular buffer zone.	30
Figure 2-7: Predictors and their relative importance (in %) for prediction of nitrate concentration in groundwater in Hesse by the models MLR (Multiple Linear Regression), CART (Classification and Regression Tree), RF (Random Forest) and BRT (Boosted Regression Trees) based on the 1,000 m circular buffer zone.	31
Figure 2-8: Predicted groundwater nitrate concentrations in Hesse with the models MLR (Multiple Linear Regression), CART (Classification and Regression Trees), RF (Random Forest) and BRT (Boosted Regression Tree) on a 1 x 1 km grid-map.	33
Figure 3-1: Scheme of the four-stage modelling process: set up of a database, data pre-processing, predictive modelling and uncertainty analysis. RF = Random Forest, QRF = Quantile Random Forest, RMSE = root mean squared error, R^2 = coefficient of determination, MAE = mean	

absolute error, PI = prediction interval, MPI = mean prediction interval, PICP = prediction interval coverage probability.	46
Figure 3-2: Boxplots of measured NO ₃ concentration for four classes of redox conditions based on the WFD data set.....	51
Figure 3-3: Importance ranking scaled from 0-100% of the six most important predictors for RF _{O₂} , RF _{Fe} , and RF _{NO₃} models based on the WFD training data. hyd_unit = hydrogeological units, GWR = groundwater recharge rate, L_KF = hydraulic conductivity, FC = field capacity, RTGW = groundwater residence time, HUMUS = humus content, Redox = redox conditions, arable_land = percentage of arable land, SWR_conc = potential nitrate concentration in seepage water, Nsurp_cm = calculated mean nitrogen surplus, forest = percentage of forest.....	51
Figure 3-4: Predicted vs. observed groundwater concentrations for the RF _{O₂} (a), RF _{Fe} (b) and RF _{NO₃} (c) model based on test data of the WFD data set. The blue lines show the regression lines between the observations and the predictions along with their 95% confidence intervals.	53
Figure 3-5: Spatial distribution of four redox classes for a 1 km x 1 km grid map of Germany derived from predicted oxygen and iron concentration in the period 2009-2018 based on the WFD data set (federal state borders: © GeoBasis-DE / BKG, 2017, Data license Germany - attribution - version 2.0 (www.govdata.de/dl-de/by-2-0)).....	54
Figure 3-6: Spatial prediction (1 km x 1 km) of the mean groundwater nitrate (NO ₃) concentration in the period 2009-2018 for Germany and the 90% prediction interval (PI) based on the WFD data set (federal state borders: © GeoBasis-DE / BKG, 2017, Data license Germany - attribution - version 2.0 (www.govdata.de/dl-de/by-2-0)).....	56
Figure 4-1: Flow paths and nitrate nitrogen loads (NO ₃ -N_loads) through the unsaturated zone and the groundwater body with: N_surplus _{a,f,u} = hydrospheric N surplus for different land use types (a = agriculture [including arable land, grassland and special crops], f = forest, u = urban) [kg N ha ⁻¹ a ⁻¹], Q _{sw} = seepage water [mm a ⁻¹], Q _{gw} = groundwater recharge [mm a ⁻¹], Q _{i,d} = interflow, drainage [mm a ⁻¹], c(NO ₃) _{gw} = NO ₃ concentration in groundwater [mg l ⁻¹], NO ₃ -N_load _{input_uz} = NO ₃ -N input load unsaturated zone [kg N ha ⁻¹ a ⁻¹], NO ₃ -N_load _{input_gw} = NO ₃ -N input load groundwater [kg N ha ⁻¹ a ⁻¹], NO ₃ -N_load _{gw} = groundwater NO ₃ -N load [kg N ha ⁻¹ a ⁻¹], NO ₃ -N_red = nitrate reduction rate [%], calculated values in italic, blue arrows represents water fluxes and orange arrows represents NO ₃ -N_loads.....	81
Figure 4-2: (a) Combined regression and density plot of seepage water (BGR, 2003a) and groundwater recharge rate (BGR, 2003b), (b) boxplot of the ratio of groundwater recharge and seepage water rate grouped by consolidated and unconsolidated aquifers.	85

Figure 4-3: (a) Nitrate nitrogen input load to the groundwater ($\text{NO}_3\text{-N_load}_{\text{input_gw}}$), (b) Groundwater nitrate nitrogen load ($\text{NO}_3\text{-N_load}_{\text{gw}}$), based on a 1 x 1 km grid map of Germany (federal state borders: © GeoBasis-DE / BKG, 2017). 86

Figure 4-4: Map of mean nitrate reduction ($\text{NO}_3\text{-N_red}$) across the unsaturated zone and the groundwater based on a 1 x 1 km grid map of Germany (federal state borders: © GeoBasis-DE / BKG, 2017). 87

Figure 4-5: Boxplots of $\text{NO}_3\text{-N}$ load reduction rates [%] for (a) aquifer type and (b) redox conditions. 88

List of Figures in Supporting Information

Figure S 2-1: Topographical map of Hesse, Germany (data source: © GeoBasis-DE / BKG, 2017); localization of 1,890 groundwater monitoring sites. 35

Figure S 2-2: Spatial predictors for 1 x 1 km grid map of Hesse; land use: percentage of urban land, grassland, forest and special crops. 38

Figure S 2-3: Spatial predictors for 1 x 1 km grid map of Hesse; land use: percentage of arable land and N surplus on agricultural land; hydrogeology: hydrogeological units. 39

Figure S 2-4: Spatial predictors for 1 x 1 km grid map of Hesse; hydrogeology & hydrology: seepage water rate, groundwater recharge rate, nitrate concentration in seepage water and nitrate concentration in groundwater recharge. 40

Figure S 2-5: Spatial predictors for 1 x 1 km grid map of Hesse; soil conditions: field capacity, humus content in top soil, soil groups. 41

Figure S 3-1: Density plots of mean groundwater concentrations of a) O_2 , b) Fe and c) NO_3) of the period 2009-2018 for the WFD data set. 61

Figure S 3-2: Localisation of selected groundwater monitoring sites of the WFD data set (blue) and sites used for the RF_{NO_3} model (red) (coordinates of the monitoring sites were provided by the LAWA-AG (German Working Group on water issues of the Federal States and the Federal Government – Committee for Groundwater and Water supply); state borders: © GeoBasis-DE / BKG, 2017, Data license Germany - attribution - version 2.0 (www.govdata.de/dl-de/by-2-0)). 62

Figure S 3-3: Spatial distribution of percentages of main land use types arable land (left) and forest (right) based on BKG (2016); (state borders: © GeoBasis-DE / BKG, 2017, Data license Germany - attribution - version 2.0 (www.govdata.de/dl-de/by-2-0)). 65

Figure S 3-4: Hydrogeological units for Germany (BGR & SGD, 2015). 66

Figure S 3-5: Spatial distribution of groundwater recharge rate (BGR, 2003b) (left) and groundwater residence times (Fuchs et al., 2010; Kunkel et al., 2007) (right); (state borders: © GeoBasis-DE / BKG, 2017, Data license Germany - attribution - version 2.0 (www.govdata.de/dl-de/by-2-0)).	67
Figure S 3-6: Mean area-weighted and land use specific hydrospheric nitrogen surplus (2007 – 2016) based on Häußermann et al (2019) (left); calculated potential nitrate concentration in seepage water (right) based on Häußermann et al (2019) and (BGR, 2003a); (state borders: © GeoBasis-DE / BKG, 2017, Data license Germany - attribution - version 2.0 (www.govdata.de/dl-de/by-2-0)).	68
Figure S 3-7: Boruta variable importance plots for a) O ₂ , b) Fe and c) NO ₃ .	70
Figure S 3-8: Spatial prediction of a) the mean and b) the median groundwater oxygen (O ₂) concentration in the period 2009-2018 based the WFD data set as well as c) the 90% prediction interval (PI); (state borders: © GeoBasis-DE / BKG, 2017, Data license Germany - attribution - version 2.0 (www.govdata.de/dl-de/by-2-0)).	72
Figure S 3-9: Spatial prediction of a) the mean and b) the median groundwater iron (Fe) concentration in the period 2009-2018 based the WFD data set as well as c) the 90% prediction interval (PI); (state borders: © GeoBasis-DE / BKG, 2017, Data license Germany - attribution - version 2.0 (www.govdata.de/dl-de/by-2-0)).	73
Figure S 3-10: Spatial prediction of a) the mean and b) the median groundwater nitrate (NO ₃) concentration in the period 2009-2018 based the WFD data set as well as c) the 90% prediction interval (PI); (state borders: © GeoBasis-DE / BKG, 2017, Data license Germany - attribution - version 2.0 (www.govdata.de/dl-de/by-2-0)).	74
Figure S 3-11: Density plots of predicted groundwater concentrations a) mean O ₂ , b) median Fe and c) mean NO ₃ based on the WFD data set.	75
Figure S 3-12: Sensitivity analysis of the effect of potentially misclassified redox classes on simulated mean NO ₃ concentration. (●) no variation of redox classes; (▽) redox classes (-1); (△) redox classes (+1).	76

1 Extended Summary

1.1 Introduction

1.1.1 Nitrate pollution – a brief overview

In the early twentieth century the solution to increase food production to handle the demand of the rapidly growing world's population, has already proven to be a major challenge to our water resources in the late twentieth century. The paradoxical reality that in order to meet our demand for food we are placing groundwater, as our most vital and valuable drinking water resource, under severe pressure, is mainly due to the intensive use of fertilizers in agriculture. With the development of the Haber-Bosch process, the industrial production of fertilizer was facilitated, leading to a dramatic increase in global agricultural production (Erisman et al., 2008). The interference in the nitrogen cycle has far-reaching negative environmental impacts such as water pollution and can also have serious consequences for human health (Erisman et al., 2013; van Grinsven et al., 2006; Vitousek et al., 1997). In the mid-1970s, water pollution became a public issue and a number of water-related directives were enacted, leading to the implementation of the EU Water Framework Directive 2000/60/EC (Kallis and Butler, 2001). The Water Framework Directive was introduced for an integrated protection with the objective of maintaining or restoring the 'good status' of water systems. As an integral part of the Water Framework Directive, the Nitrates Directive 91/676/EEC is an important instrument to protect water bodies against the agricultural pressures. In order to achieve the ambitious goals of the Water Framework Directive to meet the quality standards (WHO's maximum contamination level: 50 mg NO₃⁻/l or 10 mg NO₃-N/l in drinking water), effective mitigation measures are required to reduce nitrogen inputs to groundwater and surface water.

1.1.2 Nitrate pollution – a complex hydro-biogeochemical issue

For a better understanding of the problem, a closer look at the nitrogen (N) cycle is needed. With seven oxidation states, numerous possible conversion mechanisms and various environmental transport and storage processes the N cycle is the most complex of all major elements (Galloway et al., 2004). The atmosphere consists of about 78% of unreactive molecular nitrogen (N₂) which in this form is not available to organisms. Dry and wet deposition of nitrogen oxides (NO_x) produced by lightning and the transformation of molecular N into reactive N (N_r) by biological N fixation, result in large natural fluxes of N_r from the atmosphere into terrestrial and marine ecosystems (Fowler et al., 2013; Gruber and Galloway, 2008). N_r can be transformed either to inorganic reduced forms (e.g. ammonia (NH₃) and ammonium (NH₄⁺)) or inorganic oxidized forms like NO_x, nitric acid (HNO₃), nitrous oxide (N₂O), and nitrate (NO₃⁻), as well as organic compounds (e.g., urea, amines, proteins, and nucleic acids) (Galloway

et al., 2003). Nitrate is the result of the bacterial oxidation of ammonium called nitrification. Through denitrification nitrate is converted back to the unreactive form of gaseous N_2 (respectively other gaseous forms NO and N_2O) and is returned to the atmosphere (Rivett et al., 2008). Without extensive human interference in the natural N cycle through food and energy production, it would be balanced and no N_r would accumulate, however in the last decades, the anthropogenic N_r production was higher than those from all natural terrestrial systems (Ayres et al., 1994; Galloway et al., 2003, 2004).

As a consequence of the intensive use of fertilizers in agricultural systems respectively the application of more nitrogen than the crop can use lead to a nitrogen surplus, which accumulates in the subsurface (Ascott et al., 2016, 2017; Goulding, 2000; Worrall et al., 2015). The nitrogen surplus reaches the groundwater as nitrate via the seepage water, mostly with very long travel times. The long residence times cause a considerable amount of nitrate being stored in the unsaturated zone and a time lag from its input on agricultural systems to the arrival into the groundwater body (Ascott et al., 2017). While in the unsaturated zone a conservative transport occurs without a considerable reduction of nitrate loads, natural nitrate degradation takes place in the groundwater body (Ascott et al., 2017; Rivett et al., 2007).

Various processes, e.g. denitrification, assimilation, dissimilatory nitrate reduction to ammonium or anammox, cause nitrate reduction (Burgin and Hamilton, 2007; Francis et al., 2007; Knowles, 1982; Rivett et al., 2008). The predominant process is the denitrification, a chain of microbial reduction reactions to molecular nitrogen $NO_3^- \rightarrow NO_2^- \rightarrow NO \rightarrow N_2O \rightarrow N_2$ which strongly depends on the biogeochemical conditions of the aquifer (Knowles, 1982; Rivett et al., 2008). Denitrification processes occur under anaerobic conditions in the presence of denitrifying bacteria and electron donors (Rivett et al., 2008; Seitzinger et al., 2006). In general, there are two different types of denitrification, heterotrophic and autotrophic, depending on the type of denitrifier (Rivett et al., 2008). In heterotrophic denitrification, the bacteria use organic carbon as electron donor (Jørgensen et al., 2004). In the autotrophic denitrification the bacteria obtain the electrons from the oxidation of inorganic compounds, e.g. iron sulphide (pyrite, FeS_2) (Korom, 1992; Rivett et al., 2008). Both types of denitrification in groundwater bodies have been identified in several studies (Kludt et al., 2016; Schwientek et al., 2008; Wisotzky et al., 2018; Zhang et al., 2012). This clearly shows that the hydrogeological conditions of the aquifer, especially the prevailing redox conditions and the geochemical characteristics, can influence the intensity of nitrate pollution.

1.1.3 Nitrate pollution – the current situation

For the assessment of water and groundwater quality, nationwide monitoring programs, e.g. the Water Framework Directive groundwater monitoring network in Europe, are used to observe the occurrence

and distribution of groundwater pollution. All European Union Member States are obliged to submit the regular Evaluation Report on the implementation of the Nitrates Directive of the European Commission. According to the report for the period 2012-2015, approximately 13% of the groundwater monitoring stations of the monitoring network established under the Nitrates Directive exceed the threshold value of 50 mg NO₃⁻/l. Compared to the previous reporting period, this is a slight reduction of 1%, indicating that the mitigation measures are having only a very slow effect (EC, 2018). Furthermore, the report shows that the nitrate problem does not affect all European Member States equally. In Ireland, Finland and Sweden, for example, almost no groundwater monitoring stations exceed the threshold value of 50 mg NO₃⁻/l, whereas Malta (71%), Germany (28%) and Spain (21%) show exceedances well above the average (EC, 2018). The current nitrate report for Germany for the period of 2016 to 2018 shows a slight improvement of the situation with only 26.7% of the analysed groundwater monitoring sites with more than 50 mg NO₃⁻/l (BMU and BMEL, 2020).

The situation is similar in other major industrial nations. In a study of Burow et al. (2010) for example the groundwater nitrate concentrations measured in the period of 1991 to 2003 at the monitoring network of the U.S. Geological Survey National Water-Quality Assessment program were analysed. They found that the maximum contamination level was exceeded for 8% of the monitoring sites respectively for 20% of the monitoring sites in agricultural land-use setting. In a study for China by Gu et al. (2013) 28% of the investigated monitoring sites exceeded the WHO's maximum contamination level and more than 64% of those are located in agricultural regions.

1.1.4 Mapping and Analysis on a large scale – the capability of machine learning techniques

Monitoring networks cannot be designed to measure everywhere, and the density of monitoring networks can vary from region to region, limiting the availability of high-resolution groundwater quality data at the national level. However, the monitoring networks are designed in a way that they representatively cover the respective groundwater bodies, in order to evaluate their qualitative status. Regional and national groundwater-quality models that use detailed spatial data on environmental parameters and concentration of groundwater components can be used to complement the measured data to address data gaps and thus determine a detailed spatial distribution of the groundwater quality (Nolan and Hitt, 2006).

Besides classical geostatistical regionalisation methods such as kriging-based methods (Dalla Libera et al., 2017; Stein, 1999; Wriedt et al., 2019) or numerical modelling (Kim et al., 2008; Nguyen and Dietrich, 2018; Wriedt and Rode, 2006), statistical methods are widely used, especially for the

consideration of large scale problems like groundwater pollution (Boy-Roura et al., 2013; Gu et al., 2013; Nolan et al., 2002, 2015; Nolan and Hitt, 2006). More recently, machine learning approaches have been used in groundwater pollution estimation.

“Machine learning is a branch of artificial intelligence. Using computing, we design systems that can learn from data in a manner of being trained. The systems might learn and improve with experience, and with time, refine a model that can be used to predict outcomes of questions based on the previous learning.” (Bell, 2014). Machine learning techniques provide an effective method for the spatial regionalisation of dependent variables. A dependent variable e.g. groundwater concentration data can be predicted based on a data-driven statistical model by one or more independent or explanatory variables. By the use of nationwide data sets of spatial environmental predictors, it is possible to obtain spatial predictions of the model response variable (e.g. groundwater nitrate concentrations).

As a very effective tool, the algorithm random forest (RF), developed by (Breiman, 2001), is increasingly applied in the field of water research (Tyrallis et al., 2019). In recent studies RF was used to predict groundwater nitrate concentrations or redox conditions on large scales (Koch et al., 2019; Nolan et al., 2014; Ouedraogo et al., 2019; Rodriguez-Galiano et al., 2018; Stahl et al., 2020; Tesoriero et al., 2017; Wheeler et al., 2015; Wilson et al., 2020). In some of the studies, RF outperformed other traditional statistical methods (Ouedraogo et al., 2019; Tesoriero et al., 2017; Wheeler et al., 2015). The applications of RF and other machine learning techniques in many studies show their potential for spatial prediction on large scales and thus the opportunities in supporting the development of national strategies for mitigation measures to reduce nitrate inputs to groundwater (DeSimone et al., 2020; Ransom et al., 2017; Tesoriero et al., 2017).

In general groundwater quality modelling or the regionalisation of groundwater concentrations for areas where no measured values are available is associated with uncertainties (Grizzetti et al., 2015; Refsgaard et al., 2007). Especially when estimating parameters such as reactive nitrogen or nitrate, where complex hydro-biogeochemical processes are related to their occurrence and distribution, it is essential to consider the uncertainties. However, especially in large-scale studies on groundwater quality modelling, uncertainty analyses are not carried out regularly (Rahmati et al., 2019). Studies using machine learning techniques for estimating groundwater quality, which also addressed the uncertainties, have in some cases identified large uncertainties (Koch et al., 2019; Ransom et al., 2017).

1.1.5 Objectives of the Dissertation

The aim of this dissertation is to spatially predict nitrate concentrations in groundwater throughout Germany using a nationwide harmonised approach. The idea is first to generate a comprehensive data set of point measured concentration data at specific monitoring sites together with spatially available environmental data sets. By applying machine learning, patterns in the large data set can be identified in order to estimate the spatial distribution of the groundwater nitrate concentration. In combination with nitrogen surplus data for the entire country, the groundwater nitrate concentrations can be used to determine nitrate reduction rates. Three objectives are defined in this context:

1) Development of a nationwide harmonised approach for the estimation of groundwater nitrate concentration.

To achieve the first objective, different methods for an appropriate allocation of the spatial data to the monitoring sites are first examined in order to analyse several machine learning techniques with regard to their performance and the relevance of the respective predictors in estimating groundwater nitrate concentration. The development of the approach is based on the example of the federal state of Hesse, Germany (Chapter 2).

2) Estimating nitrate concentration in the groundwater throughout Germany.

To accomplish the second objective, the approach developed in the first part, based on the example of the federal state of Hesse, is adapted for the application to entire Germany. As an additional important predictor, the redox conditions in groundwater are taken into account and an uncertainty analysis is carried out (Chapter 3).

3) Determination of the nitrate reduction across the unsaturated zone and the groundwater body in Germany.

A simplified conceptual approach is used to achieve the third objective; determining an integrated nitrate reduction rate on the basis of nitrogen input into the unsaturated zone and the previously determined nitrate concentrations in groundwater. Furthermore, the hydrogeological conditions are analysed with regard to the nitrate reduction (Chapter 4).

1.2 Machine learning techniques for the estimation of groundwater nitrate concentrations on a large scale

Based on the federal state of Hesse in Germany (see Figure 1-1), an approach for the large scale prediction of groundwater nitrate concentrations is developed. The country provides a very heterogeneous landscape and thus represents very well several different characteristics of Germany. Different approaches for the allocation of the spatial environmental data to the respective monitoring sites are tested and a comparison of machine learning techniques based on different statistical model types is carried out to develop an optimal approach for predicting groundwater nitrate concentrations on a national scale:

Knoll, L., Breuer, L., Bach, M., 2019. Large scale prediction of groundwater nitrate concentrations from spatial data using machine learning. Science of The Total Environment 668, 1317–1327. <https://doi.org/10.1016/j.scitotenv.2019.03.045>

The general modelling process involves the three steps: database setup, data pre-processing and predictive modelling. In the first step all the information on groundwater quality data, provided as point data from monitoring sites is stored together with spatial environmental data in a geodatabase. In the pre-processing step, the environmental conditions (available as spatial data or maps) in the catchment area of the monitoring sites are compiled and linked to the groundwater quality data. This is performed simplified by using buffers, which approximately represent the contributing area. The main part is the predictive modelling. In this step, a machine learning algorithm is trained based on the dependent variable (e.g. groundwater nitrate concentration) and a set of independent explanatory predictors (environmental spatial data). The trained model can be applied to predict the dependent variable (model response variable) based on the same set of predictors underlying the model training process. In this way, estimations can be made for locations where no measured values of the target variable are available. Figure 1-1 gives a short summary of the workflow in the predictive modelling approach. A detailed description of the methodology of the approach is given in Chapter 2.

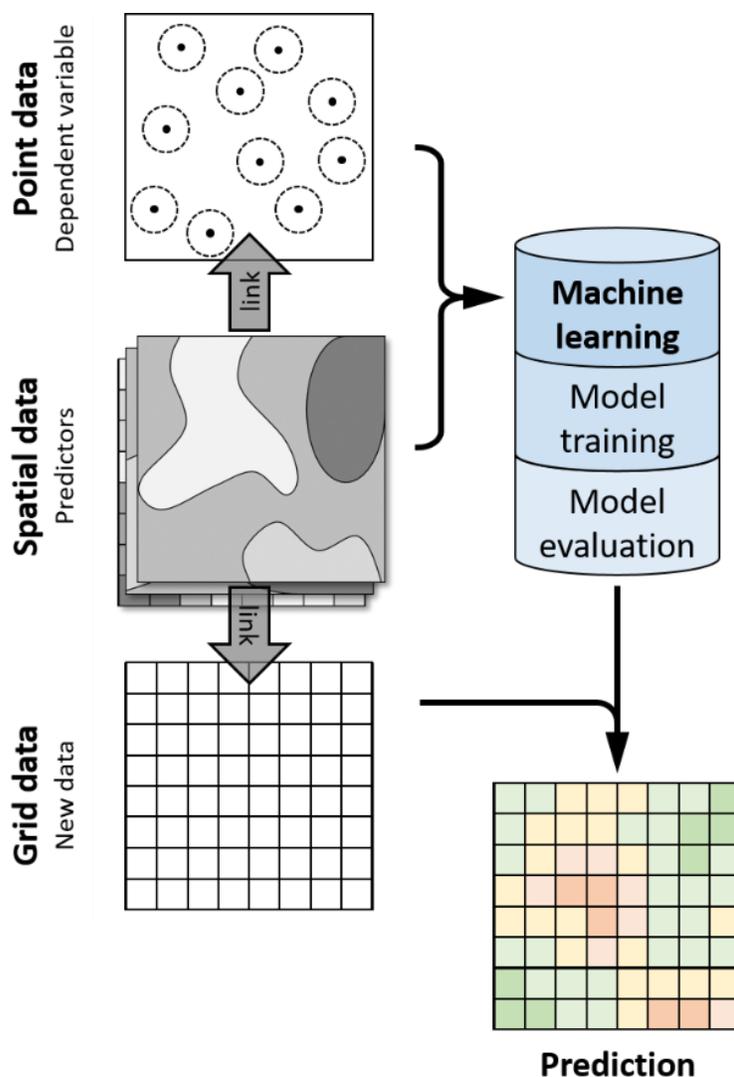


Figure 1-1: Schematic workflow of the predictive modelling approach.

The long-term time series of groundwater quality data and metadata of the groundwater monitoring service of the State of Hesse and further monitoring sites of the public water supply are collected. After applying a number of selection criteria, 1,890 monitoring sites can be used to perform a statistical modelling approach. The mean nitrate concentrations in the period 2010 to 2017 are used as a dependent variable (point data). As the explanatory independent variables, spatial environmental data (spatial data) are used such as land use, hydrogeology and soil conditions, which are available as maps throughout Germany. The spatial data has to be compiled for the contributing area of each monitoring site and thus linked to the point information. The predictive modelling is based on a grid map with a resolution of

1 km x 1 km. For each grid cell, the spatial environmental predictors also have to be compiled and linked to it.

In the study five different buffer designs representing the contributing area respectively the potential catchment area of a monitoring site are investigated for four statistical models regarding their predictive performance. Three circular buffers with a radius of 500 m, 1,000 m and 1,500 m and two wedge buffers with 1,000 m radius and apertures of 45° and 90° oriented in groundwater flow direction are analysed. The four investigated statistical model types are: multiple linear regression (MLR), classification and regression trees (CART), random forest (RF) and boosted regression trees (BRT). The models based on the circular buffer with 1,000 m radius outperform the models based on the other buffer designs and thus indicate an adequately description of the contributing area. The models based on wedge buffers all show a lower predictive performance, but a contributing area design based on precisely determined groundwater flow conditions would certainly lead to an improvement of the performance (Kihumba et al., 2016; Mattern et al., 2009). However, when considering even larger-scale studies (Germany-wide), the availability of such information cannot be expected. Furthermore, Johnson and Belitz (2009) and Wheeler et al. (2015) have also found that simplified circular buffers e.g. with a radius of 1,000 m can represent the contributing area sufficiently well.

The machine learning algorithms based on the four different statistical model types are compared in this study with respect to their performance in predicting groundwater nitrate concentrations. The evaluation of the model performance is based on the three objective functions: the coefficient of determination (R^2), the mean absolute error (MAE) and the root mean squared error (RMSE). After training each model, using three times repeated 10-fold cross-validation, the RF model (based on the 1,000 m circular buffer) outperforms the other model types. The RF model shows the highest R^2 value of 0.55 and the lowest MAE (5.90 mg/l) and RMSE (9.19 mg/l) after cross-validation indicating a good predictive performance. The comparison of the measured nitrate concentrations of the independent test data set underlines these results with similar values for the objective functions. The other model types also result in good to reasonable predictive performances. Compared to other studies dealing with large scale estimation of groundwater nitrate concentration, the model performance is in a similar or even better range (Boy-Roura et al., 2013; Nolan et al., 2014, 2015; Ransom et al., 2017; Tedd et al., 2014).

Within the best performing RF model, the hydrogeological units followed by the percentage of arable land and the nitrogen surplus are the most relevant predictors. A shift in the predictor ranking indicates that depending on the different model types, some predictors show a strong or less strong influence on nitrate prediction.

Looking at the predictions of the spatial distribution of nitrate concentrations in groundwater, all four models result in a similar pattern. High nitrate concentrations are predominantly found in the lowlands, which are intensively used for agricultural purposes. Overall, the threshold value of 50 mg NO₃⁻/l is exceeded for only 2% of the area of the federal state of Hesse.

The application of machine learning algorithms such as random forest allows a high extraction of the explanatory potential by the predictors used. However, reliable and representative input data is required for the respective predictors as well as for the dependent response variable. The approach presented in this first part of the dissertation shows the scope of application for the large-scale regionalisation of groundwater nitrate concentrations, in particular by the use of uniform nationwide data sets.

1.3 Groundwater nitrate concentration throughout Germany

The approach presented above is adapted, further developed, and applied to estimate the groundwater nitrate concentrations for the entire federal republic of Germany (see Figure 1-2). A data set of groundwater quality data covering the whole of Germany is compiled and the grid map is extended to cover the entire country. The main modification is (i) the assessment of redox conditions and their integration as additional predictor. Furthermore, (ii) the estimations are subjected to an uncertainty analysis:

Knoll, L., Breuer, L., Bach, M., 2020. Nation-wide estimation of groundwater redox conditions and nitrate concentrations through machine learning. Environ. Res. Lett. 15, 064004.

<https://doi.org/10.1088/1748-9326/ab7d5c>

The federal republic of Germany consists of 16 states and is located in Central Europe. The country covers an area of around 360,000 km² and has a diverse landscape (see Figure 1-2). From the coastal areas at the North and Baltic Sea, the country extends over the North German Plain and the Central German Uplands to the Alps at the southern border. About 50% of the area is used for agricultural purposes. The analyses performed in this dissertation are based on the grid map of Germany with a resolution of 1 km x 1 km (BKG, 2018). The federal state of Hesse, with its heterogeneous landscape of low mountain ranges and depressions (see Figure 1-2), provides a good representation of the different landscape characteristics of Germany. However, with the Germany-wide approach involving the North German Plain a large landscape element is added, which is not covered by the model set up for Hesse. Thus, for the Germany-wide approach an adaptation is necessary. For instance, groundwater redox conditions, which differ strongly between the North German Plain and the remaining part of the country, need to be implemented.

Since the assessment of uncertainties in environmental modelling like groundwater quality modelling is essential, an uncertainty analysis is added as a fourth step in the large-scale application of the approach for Germany. In the uncertainty analysis, the machine learning algorithm quantile random forest is used to derive prediction intervals of the estimates. These intervals can be used to verify the reliability of the estimates, as they describe the range in which the true value can be expected with a certain probability. A detailed description of the methodology of the approach is given in Chapter 3.

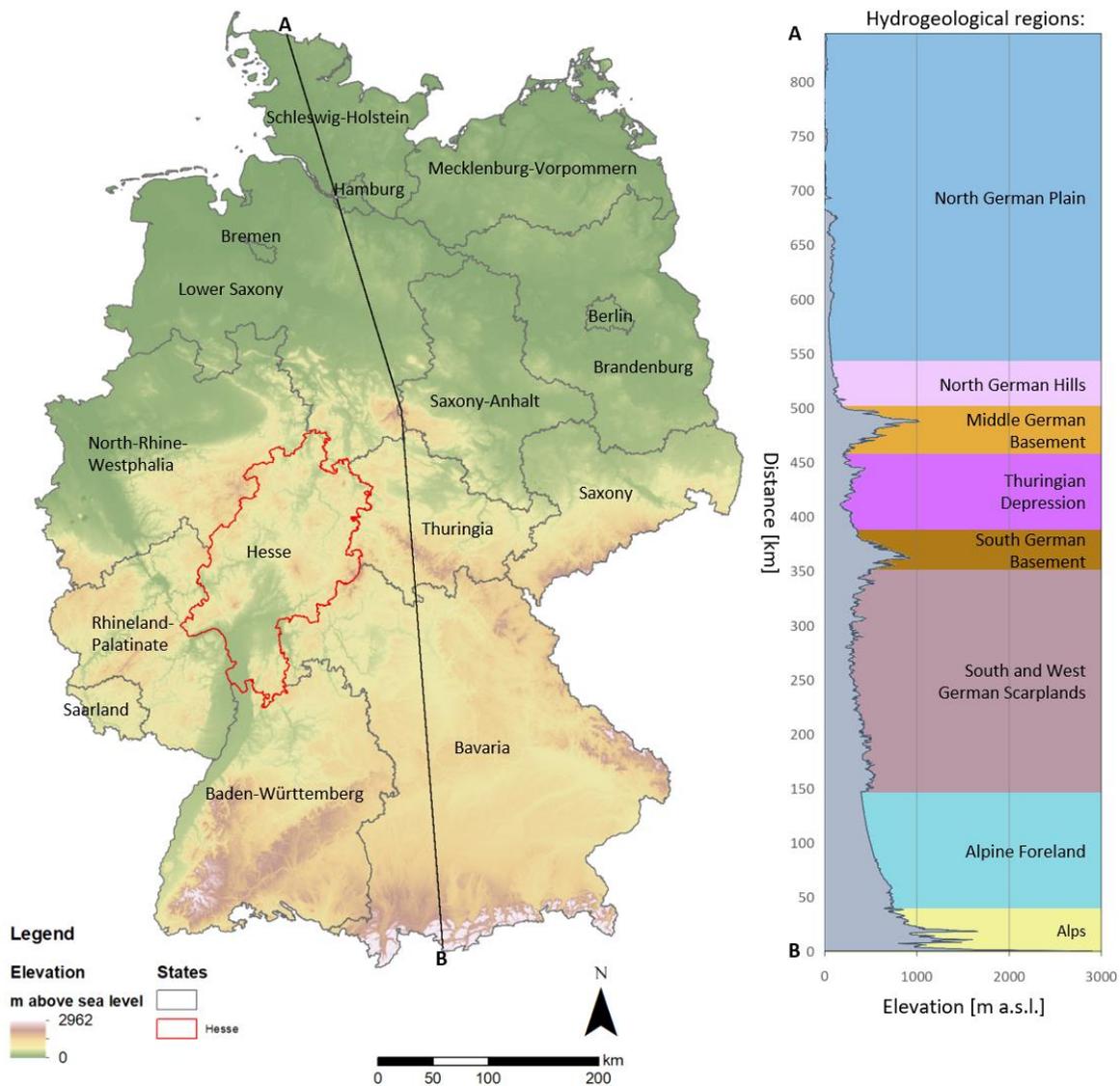


Figure 1-2: Map of Germany with the federal state Hesse (left) and a north to south profile line with elevations and the hydrogeological regions (right).

Within the scope of the Germany-wide study, the monitoring sites of the Water Framework Directive groundwater monitoring network are used as the basis of predictive modelling. These represent a monitoring network designed according to uniform criteria with clearly defined requirements for a representative evaluation of the upper groundwater bodies. After applying specific selection criteria, the mean nitrate concentration in the period 2009 to 2018 of 5,414 groundwater monitoring sites is used as a dependent variable. Furthermore, the concentration data of oxygen and iron are used to describe the redox conditions prevailing in the groundwater. For the explanatory environmental spatial predictors, approximately the same data is used as for the example of Hesse, as these maps are already available throughout Germany. The pre-processing is done analogously to the procedure of the state of Hesse, however, the spatial predictors, which have to be linked to the monitoring sites as well as to the grid cells, are only compiled via the circular buffer with a radius of 1,000 m, since this contribution area provides the best prediction results in the first part.

With the previously presented predictive modelling approach (see Figure 1-1), it is also possible to predict groundwater oxygen and iron concentrations. For both RF models, for oxygen and for iron, the hydrogeological units prove to be the most relevant predictor, which is in line with the conclusions of Tesoriero et al. (2015, 2017). With the oxygen and iron groundwater concentrations, the redox conditions in the groundwater body are described and it can be shown that they have a considerable influence on groundwater nitrate concentrations. Four redox classes from aerobic, intermediate, anaerobic to strongly anaerobic are defined. Hence, these are added as predictors for the estimation of groundwater nitrate concentrations. Based on the predicted spatial distribution of oxygen and iron concentrations in groundwater, the redox classes are derived for the entire grid map and then used to estimate the spatial distribution of nitrate concentrations. Besides the hydrogeological units and the percentage of arable land, the redox conditions prove to be the most influential predictor in the nitrate RF model, which is in line with Ransom et al. (2017).

The predictive performance of the nitrate RF model for Germany is also evaluated after cross-validation resulting in good performance measures ($R^2=0.52$, MAE=12.71 mg/l and RMSE= 20.12 mg/l). The results show, that the nitrate RF model tends to underestimate high to extreme concentrations of >100 mg NO_3^-/l . Even if the model performance for the nationwide approach for Germany is somewhat lower than that for the model region of federal state of Hesse, it is still in a good range, particularly in view of the much larger scale. In addition, the Water Framework Directive groundwater monitoring network used for Germany has a considerably lower density of monitoring sites than the network used in the approach for the state of Hesse. Compared to other studies dealing with large scale prediction of nitrate concentration in groundwater the performance of the nitrate RF model is in a similar range (Nolan et al., 2015; Ransom et al., 2017; Wheeler et al., 2015).

Even if the model performance is within a good range, very high uncertainties respectively high prediction intervals are found in some areas by the uncertainty analysis (mean prediction interval of 53 mg/l). Particularly in nitrate polluted areas, the uncertainties are very high, which corresponds to the tendency of the nitrate RF model to underestimate high concentration ranges. It has also been found that with a denser monitoring network using additional monitoring sites operated by several federal states (around 13,000 groundwater monitoring sites), the uncertainties can be reduced and the model performance is increased. The use of additional monitoring sites in combination with the Water Framework Directive groundwater monitoring network, however, needs to ensure that the requirements for the representativeness of the monitoring network are met. Since these additional monitoring sites are not part of the WFD monitoring program. Therefore, it is not ensured that they meet the requirements of the WFD and hence they are not further considered in this dissertation.

Looking at the spatial distribution it can be shown that the redox conditions are clearly divided into the anaerobic North German Plain and the tendency of aerobic to intermediate conditions in the southern and Central German Uplands. This is similarly stated in a study of Kunkel et al. (2004) or Hannappel et al. (2018). The nitrate concentrations show a somewhat more heterogeneous distribution. On the one hand, regions with strongly anaerobic or nitrate-reducing conditions show no nitrate pollution, on the other hand, intensively agriculturally used regions clearly show high nitrate concentrations compared to forested mountainous regions with low nitrate concentrations. According to the estimations based on the nitrate RF model, around 10% of the regionalized groundwater nitrate concentration exceed the limit value of 50 mg NO₃⁻/l.

1.4 Nitrate reduction throughout Germany

The previously determined distribution of nitrate concentrations in groundwater is used in a third step together with data on nitrogen input into the subsurface to quantify an integrated nitrate reduction rate across the unsaturated zone and the groundwater body:

Knoll, L., Häußermann, U., Breuer, L., Bach, M., 2020. Spatial Distribution of Integrated Nitrate Reduction across the Unsaturated Zone and the Groundwater Body in Germany. Water 12, 2456. <https://doi.org/10.3390/w12092456>

For Germany, it can be assumed that under the soil-climatic conditions typical for the country, the nitrogen supplied in various reactive forms is displaced with seepage water only in the form of nitrate nitrogen ($\text{NO}_3\text{-N}$). To quantify the $\text{NO}_3\text{-N}$ load transported via the unsaturated zone into the groundwater body, the N input into the unsaturated zone, which is considered as area-weighted and land use specific hydrospheric N surplus, is calculated by multiplying with the ratio of groundwater recharge rate and seepage water rate. This ratio averages around 0.5 and shows that only about half of the seepage water reaches the groundwater body, the remaining amount enters the surface water via interflow or tile drainage systems. For the unsaturated zone a conservative transport with only a small decrease in the $\text{NO}_3\text{-N}$ loads can be expected (Ascott et al., 2017; Rivett et al., 2007). It has also been found that the ratio of the groundwater recharge rate to the seepage water rate in consolidated aquifers is much lower than in unconsolidated porous aquifers. This fact is usually based on morphological aspects, which favour higher interflow rates in consolidated units.

The mean $\text{NO}_3\text{-N}$ input load to the unsaturated zone is calculated with around $36 \text{ kg N ha}^{-1} \text{ a}^{-1}$ of which consequently only around half of the load ($17 \text{ kg N ha}^{-1} \text{ a}^{-1}$) enters the groundwater body. The spatial distribution shows the intensively agriculturally used areas in the north, north-west and south-east of Germany and as well as some lowland regions with high $\text{NO}_3\text{-N}$ loads up to $40 \text{ kg N ha}^{-1} \text{ a}^{-1}$. The overall $\text{NO}_3\text{-N}$ input into the unsaturated zone and into the groundwater body amounts to $1,196 \text{ kt N a}^{-1}$ and 580 kt N a^{-1} , respectively. According to the calculations, the $\text{NO}_3\text{-N}$ loads in the groundwater body with 236 kt N a^{-1} are substantially lower than the $\text{NO}_3\text{-N}$ input loads. Consequently, the difference of 344 kt N a^{-1} can be attributed to nitrate reduction. The mean nitrate reduction rate respectively the reduction of $\text{NO}_3\text{-N}$ loads is calculated with around 57%. The spatial distribution shows a clear north-south division with high reduction rates up to 100% in the north and very low reduction rates up to no reduction in central and southern part of Germany.

For specific hydrogeological features, considerable differences in the reduction rates have been analysed. It can be found that the nitrate reduction in porous aquifers is significantly higher than in fractured,

consolidated aquifers and groundwater bodies with anaerobic conditions tend to have significantly higher nitrate reduction rates than those with aerobic conditions.

The parsimonious conceptual approach presented here provides the possibility of estimating an integrated nitrate reduction rate in the unsaturated zone and in the groundwater body without the simulation of complex biogeochemical processes. In comparison to the results of this approach, Hirt et al. (2012), for example, calculated the same reduction rates for the Weser catchment with a more complex coupled model chain. It should be considered that the input data for the conceptual data-driven approach are partly subject to high uncertainties, so that this also applies to the determined reduction rates. However, the use of such a parsimonious conceptual approach is advantageous because knowledge about nitrogen turnover processes, especially in groundwater, is relatively limited and process-based models can fail due to a lack of knowledge. Due to the different types of input data, the different degree of uncertainty and the assumptions made, it is difficult to perform a precise uncertainty analysis. Another aspect that should be taken into account when evaluating reduction rates is that reduction capacity is not inexhaustible and, depending on local conditions and the degree of nitrate pollution, may decrease over time or reach capacity limits (Bergmann et al., 2013; Kludt et al., 2016; Wilde et al., 2017).

1.5 Conclusion and Outlook

The Water Framework Directive and the EU Nitrates Directive (91/676/EEC) require mitigation measures to achieve good chemical status. The first priority is therefore the identification of groundwater bodies with high nitrate pollution. In Germany, information from various sources on the share of groundwater above the limit value of 50 mg NO₃⁻/l varies between 10.4% and 49.4% which makes it difficult to argue for the implementation of measures (Bach et al., 2020b). This illustrates the requirements for a technically well-founded regionalisation method based on a representative data base. In view of the federal system in Germany, it is even more important to implement uniform monitoring programs and to specify clearly defined requirements for the monitoring networks.

With the application of machine learning respectively the random forest algorithm for the regionalisation of groundwater components like nitrate concentrations, new territory is entered, at least for Germany. In the international field, similar approaches have already been applied and accepted more widely, while in Germany, classical geostatistical regionalisation methods have been applied so far. The major advantage of machine learning methods compared to classical regionalisation methods is the possibility to use a variety of influencing parameters to predict a depending variable. However, this requires reliable and representative data sets of the relevant predictors and dependent response variables.

Within this dissertation, it is possible to carry out a first Germany-wide estimation of nitrate concentration in groundwater using machine learning. A random forest algorithm outperforms other statistical model types in predicting groundwater nitrate concentrations. The algorithm is trained by means of the Water Framework Directive groundwater monitoring network which is used for a representative assessment of the qualitative status of the German aquifers. Spatial data available throughout the country can be used as explanatory variables. Among these predictors, the redox conditions, followed by the hydrogeological units and the percentage of arable land are the most relevant for the prediction of nitrate concentrations in groundwater.

The spatial distribution of groundwater nitrate concentrations resulting from the use of the predictive modelling approach applied in this dissertation indicates that the limit value of 50 mg NO₃⁻/l is exceeded for around 10% of the area of Germany. This is not in line with the 27% of the aquifers identified by the Water Framework Directive as not meeting the good chemical status due to nitrate (UBA, 2017a). The classification of the chemical status according to Water Framework Directive is based on the precautionary principle according to GrwV (2010). According to this, the entire aquifer is classified as 'poor status' if only one monitoring site or a defined proportion within the groundwater body exceeds the corresponding threshold. In contrast, the approach taken here allows a more differentiated evaluation based on regionalized groundwater nitrate concentrations. Recent discussions in this context with regard to internal differentiation show even more the importance of a reliable spatial regionalisation on a higher resolution. Therefore, the findings of this dissertation can contribute to planning of the implementation of mitigation measures.

Not only the spatial distribution of nitrate concentrations in groundwater but also the distribution of the natural degradation capacity of aquifers is of vital importance in the development of strategies for action plans. In this dissertation, with the help of the predicted spatial distribution of the nitrate concentrations in groundwater and additional data on nitrogen input into the unsaturated zone, a simple conceptual approach is developed to enable the estimation of the spatial distribution of integrated nitrate reduction rates across the unsaturated zone and the groundwater body. It has been shown that the reduction rates strongly vary from no reduction up to 100% and that the hydrogeological conditions of the aquifer also have a clear influence on the nitrate reduction rates.

These results are already used in other studies dealing with nitrogen pollution. In Geupel et al. (submitted) an impact-based integrated national target for nitrogen is developed for Germany as part of the DESTINO project of the German Environment Agency (Bach et al., 2020a; Heldstab et al., 2020). This target value is used to derive a quantitative estimate of the annual permissible nitrogen loss rates to the environment in Germany as a target for political action. The integrated target value considers six

environmental sectors, each with its own quality targets. One of these sectors is groundwater, for which the necessary reduction of the N input is determined on the basis of groundwater nitrate concentrations $>50 \text{ mg NO}_3^-/\text{l}$, in order to comply with this limit value. For this purpose, the Germany-wide regionalized nitrate concentrations in groundwater bodies are considered.

In (Ebeling et al., 2020) riverine nutrient dynamics in 787 German catchments are investigated. A random forest model is applied to predict riverine nutrient concentrations and concentration-discharge ratios based on a variety of catchment characteristics. Using the map of the spatially distributed groundwater nitrate concentration together with calculated potential nitrate concentrations in seepage water (a ratio of N surplus (Häußermann et al., 2019) and the seepage water rate (BGR, 2003a)), the vertical concentration heterogeneity is calculated as one of the predictors. It has been found that natural attenuation reduces nitrate concentrations in the catchment and cause heterogeneity that controls the export dynamics into the surface waters. Furthermore, it can be identified that the vertical concentration heterogeneity predominantly influences the export dynamics.

In the conceptual river basin management system MoRE - Modelling of Regionalized Emissions (Fuchs et al., 2017a) nationwide nutrient emissions into surface waters are simulated. With the possibility of a pathway-specific quantification of emissions and river loads, MoRE is used as the reporting tool for the Federal Republic of Germany. The MoRE model originates mainly from the conceptual nutrient emission model MONERIS - MOdelling of Nutrient Emissions in RIver Systems (Venohr et al., 2011) and has largely adopted the simplified approaches for the representation of the inputs via the transport path: unsaturated zone \rightarrow groundwater \rightarrow surface water. Therein the N reduction is considered in a simplified way via an exponential function depending on the seepage water rate and hydrogeological conditions. The functions are based on nitrate concentrations from the 1990s in seepage water and groundwater of 217 groundwater monitoring sites (Venohr et al., 2011). With the nitrate reduction rates determined in this dissertation, a more comprehensive data base is now available based on nationwide and actual datasets. These nitrate reduction rates will be used to adapt the representation of the transport path across the unsaturated zone and groundwater in the model instrument MoRE (Knoll et al., under review).

This dissertation shows both the possibilities and the limitations of the application of machine learning techniques in the field of groundwater quality modelling respectively the large-scale regionalisation of groundwater components. The results presented here show the great potential of machine learning techniques, however, this work also shows the uncertainty of such large-scale data-driven approaches. The need for comprehensive uniform approaches to assess the qualitative status of water bodies underlines the need for research in this area.

2 Large scale prediction of groundwater nitrate concentrations from spatial data using machine learning

This chapter is published in the journal *Science of The Total Environment* 668, pages 1317–1327, 2019. <https://doi.org/10.1016/j.scitotenv.2019.03.045>

Knoll, L.^{1*}, Breuer, L.¹, Bach, M.¹

[1] Institute for Landscape Ecology and Resources Management, Research Centre for Biosystems, Land Use and Nutrition (iFZ), Justus Liebig University Giessen, Germany

* Correspondence to: Lukas Knoll (Lukas.Knoll@umwelt.uni-giessen.de)

Abstract

Reducing nitrogen inputs, in particular nitrate, to groundwater is becoming increasingly important to fulfil requirements of the European Water Framework Directive. When developing management plans for mitigation measures at larger scales, complex hydro-biogeochemical models reach their limits due to data availability and spatial discretization. To circumvent this problem, the spatial distribution of nitrate concentration in groundwater is estimated using a parsimonious GIS-based statistical approach. Point nitrate concentrations and spatial environmental data as predictors are used to train statistical models. In order to compile the spatial predictors with the respective monitoring sites, different designs of contributing areas (buffer zones) and their effects on the performance of different statistical models are investigated. Multiple Linear Regression (MLR), Classification and Regression Trees (CART), Random Forest (RF) and Boosted Regression Trees (BRT) are compared in terms of the predictive performance of each model according to various objective functions. We determine the most influential spatial predictors used in the respective models. After training the models with a subset of the data, we then predict the spatial nitrate distribution in groundwater for the entire federal state of Hesse, Germany on a 1 x 1 km grid by only the spatial environmental data. The Random Forest model outperforms the other models ($R^2=0.54$), relying on hydrogeological units, the percentage of arable land and the nitrogen balance as the three most influencing predictors based on a 1,000 m circular contributing area. The use of exclusively spatial available predictors is a big step forward in the prediction of nitrate in groundwater on regional scale.

2.1 Introduction

The protection of water resources is one of the biggest challenges of current environmental research. High nitrate concentrations deteriorate surface water and groundwater quality and pose a serious risk to human's health and the environment (WHO, 2017). In recent decades, the enhanced application of nitrogen (N) fertilizer and subsequently high diffuse nitrate losses from agricultural land are the main reasons for the poor state of water resources in Europe (EEA, 1999). Thus, improving water quality, in particular by reducing nitrate inputs, is an essential element of the EU Water Framework Directive 2000/60/EC (WFD) as well as the EU Nitrates Directive 91/676/EEC.

Nitrate concentrations in groundwater and surface waters depend on numerous factors such as land use, agricultural fertilizer application practice, hydrological and hydrogeological conditions, soil properties and weather conditions. A variety of model approaches have been developed in recent decades, to simulate the N input, turnover and transport in the hydrosphere in various temporal and spatial scales, and to develop improved management plans. The existing models range from simple empirical-conceptual approaches to complex process-based models (Bouraoui and Grizzetti, 2014; Breuer et al., 2008; Cherry et al., 2008; Wellen et al., 2015). Most of these models are limited in terms of considering spatially distributed groundwater flow and have limitations in reproducing spatial distributions of groundwater nitrate concentrations. Some studies provide approaches for a better representation of the groundwater component (Kim et al., 2008; Nguyen and Dietrich, 2018), which are however impractical on large scale.

For the development of river basin management plans required by the WFD, mainly conceptual models at national and regional scale are used (Grizzetti et al., 2008; Velthof et al., 2009). In Germany, the conceptual models MONERIS (Behrendt et al., 2000; Behrendt and Opitz, 1999; Venohr et al., 2011) and MoRE (Fuchs et al., 2017a, 2017b) are commonly used as the WFD reporting and management planning tools (Fuchs et al., 2010; Hirt et al., 2012). Both models estimate the potential nitrate concentration in leachate from the N surplus on agricultural land (i.e. the excess of N inputs over outputs on an agricultural field) and the seepage water rate for administrative units. The grid-based model chain GROWA-DENUZ-WEKU consists of a nutrient balance model, a water balance model (GROWA), a reactive nitrate transport model in soil (DENUZ) and reactive nitrate transport model in groundwater (WEKU) to predict nitrogen inputs into groundwater and surface water on high resolution (Kunkel et al., 2004, 2017, Kunkel and Wendland, 1997, 2002).

Maps with a high spatial resolution of the actual state of nitrate concentrations in groundwater have so far only been used sparsely for WFD action planning. Geostatistical approaches such as kriging-based methods are widely used to interpolate point measurements (Dalla Libera et al., 2017; Hu et al., 2005;

Nas and Berktaş, 2010; Wriedt et al., 2019). These techniques are point-based (i.e. depending on measured data) and consider the spatial covariance of the measurements. Some approaches consider additional information of environmental data to improve the spatial prediction (Bárdossy et al., 2003; Pebesma and de Kwaadsteniet, 1997). Other approaches like DRASTIC (Aller et al., 1985) systematically evaluate the groundwater vulnerability related to the hydrogeological setting (Babiker et al., 2005; Nixdorf et al., 2017; Ouedraogo et al., 2016; Panagopoulos et al., 2006). Nolan et al. (2002) and Nolan and Hitt (2006) create a groundwater vulnerability map derived from groundwater monitoring data linked with hydrogeological and land use characteristics.

In view of the data requirement, model complexity and runtime, a number of straightforward statistical models e.g. Multiple Linear Regression (MLR) are alternatively used to estimate nitrate pollution in groundwater at large scales (Boy-Roura et al., 2013; Liu et al., 2005; Mattern et al., 2009; Mfumu Kihumba et al., 2016; Ouedraogo and Vanclooster, 2016; Tedd et al., 2014; Wick et al., 2012). Mfumu Kihumba et al. (2016) and Mattern et al. (2009) compared MLR with the decision tree methodology Classification and Regression Tree (CART) for modelling nitrate pollution pressure. Machine learning techniques are of particular interest, as they can deal with complex data sets, a large number of predictor variables, and all scales of measures, including nominal scale. Ouedraogo et al. (2019) used Random Forest (RF) to model groundwater nitrate concentration at the African continent scale. Nolan et al. (2014, 2015) and Ransom et al. (2017) created a map of nitrate concentration in groundwater of the Central Valley, California using RF and Boosted Regression Trees (BRT). Wheeler et al. (2015) used CART and RF for predicting nitrate concentration in wells in Iowa. Tesoriero et al. (2017) and Rodriguez-Galiano et al. (2014, 2018) developed probability maps concerning nitrate contamination with RF in a region of Wisconsin and southern Spain. However, most of these studies use very comprehensive and regionally limited data sets.

The use of exclusively state-wide accessible spatial predictors and machine learning techniques is a promising way forward to estimating nitrate concentration in groundwater with high spatial resolution on a large scale. For the allocation of the spatial predictors to the respective monitoring sites, different designs of contributing areas (circular buffer and wedge buffer) and their effects on the predictive performance of different statistical models are investigated in the present study. The performance of four statistical models MLR, CART, RF and BRT to spatially depict the nitrate concentration in groundwater is compared by taking the example of Hesse, Germany. Furthermore, the most relevant hydrogeological and land use parameters are identified, which serve as predictors for the estimation of nitrate concentration. Last, MLR, CART, RF and BRT are used to establish 1 x 1 km grid maps of groundwater nitrate concentration for Hesse.

2.2 Materials and Methods

2.2.1 Study area

The study area covers the state of Hesse (Germany) with 21,115 km² (Figure S 2-1). The landscape of Hesse shows a relatively heterogeneous morphology with lowlands and low mountain ranges. Lowest elevations of around 70 m above sea level (NN) can be found in the Middle Rhine Valley and highest elevation up to 950 m NN in the eastern mountain ranges. The mean (2010 – 2017) annual air temperature is 9.2° C and the mean annual precipitation is 635 mm/a (DWD, 2018).

2.2.2 Basic modelling approach

The basic modelling approach consists of three main steps, a data pre-processing, a statistical model framework with different statistical models and a grid prediction (Figure 2-1). Pre-processing and statistical modelling is performed in a coupled environment of a Geographic Information System (ArcGIS v. 10.4) and a statistical computing software (R v. 3.4.1).

The nitrate concentration is available as point data and represents the response variable for the statistical models. The independent predictors consist of spatial geodata in the form of maps, which need to be linked to the point data. This is done by means of contributing area, which represent the potential catchment area of the respective monitoring sites. Based on the independent predictors, nitrate concentrations are finally calculated using the statistical models. On the basis of the statistical models, predictions can now be made. In the present study, the predictions rely on 1 x 1 km grid maps of the spatial predictors.

2.2.3 Data pre-processing

Point data

The Hessian Agency for Nature Conservation, Environment and Geology (HLNUG) maintains the groundwater database *GruWaH* to record groundwater quality throughout the state of Hesse. The *GruWaH* combines data from two monitoring networks, the groundwater monitoring service of the state of Hesse and the data from public water supply companies according to the water supply regulation (HLNUG, 2018).

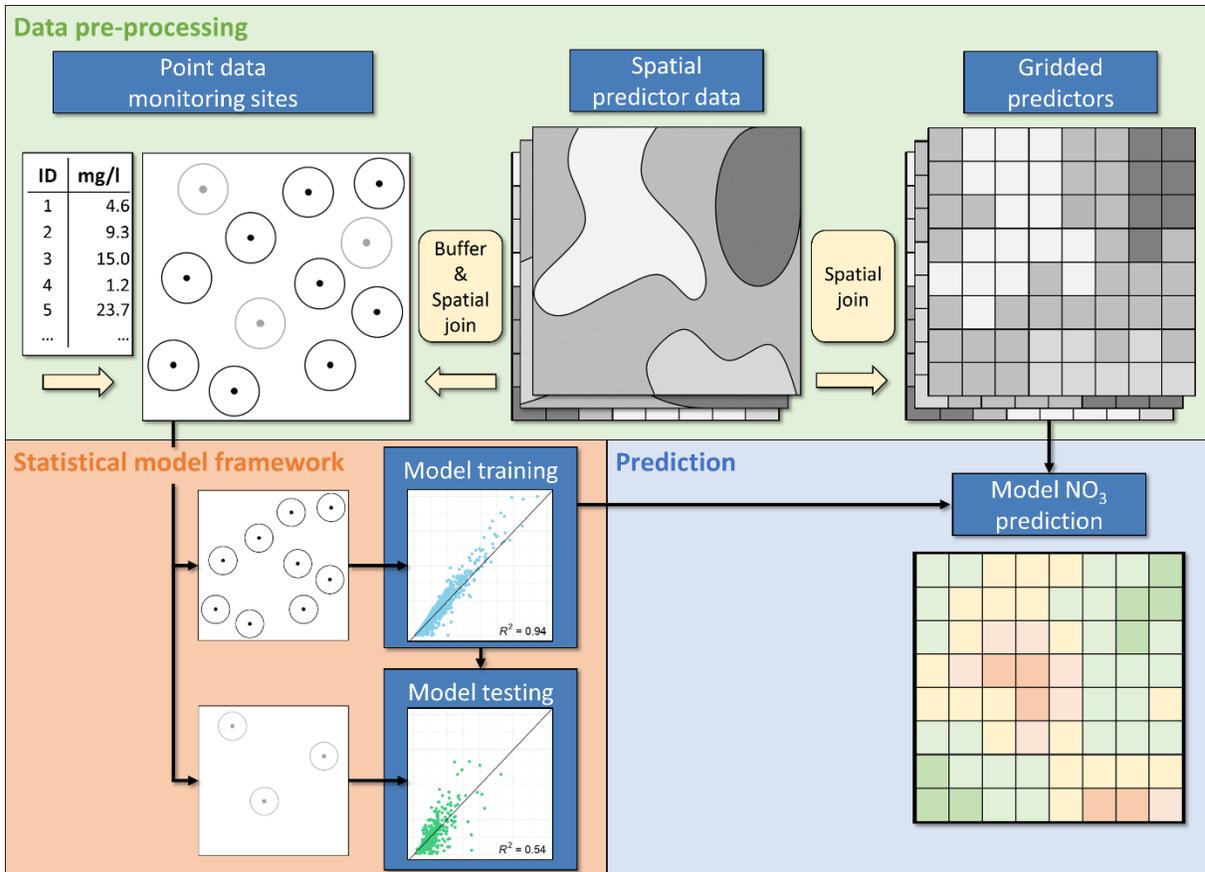


Figure 2-1: Workflow of the GIS-based statistical modelling approach, consisting of the three steps: data pre-processing, statistical model framework and prediction.

The groundwater monitoring service of the state of Hesse is a state-wide network operated through the HLNUG to map the groundwater quality of the various hydrogeological units of Hesse. Furthermore, the monitoring sites close spatial gaps between the monitoring sites of the public water supply companies (see below). The network comprises about 400 monitoring sites (86% groundwater monitoring stations and 14% springs). Sampling and analysis of the state groundwater monitoring sites is carried out by the State Laboratory of Hesse (Wiesbaden) according to DIN 38402-13.

Public water supply companies are obliged to measure raw water quality four times a year. The monitoring network consists of 2,630 wells and 1,670 springs. Sampling and the analysis is carried out by larger water supply companies in their own laboratories, smaller companies commission certified laboratories with the analysis. The groundwater samples are generally analysed in certified laboratories. The nitrate concentration is determined in accordance with the DIN EN ISO 10304-1:2009-07. Results

are compiled at the HLNUG, subjected to a plausibility check in advance, and entered into the *GruWaH* database.

The extensive *GruWaH* database provides time series of groundwater quality measurements dating back to the 1950s and the corresponding meta data of the particular monitoring sites. The sampling frequency varies from one-time to seasonal. The well depths varies from shallow to deep (>100 m) in different aquifer types. For the estimation of shallow groundwater nitrate concentrations using statistical models, a pre-selection of the relevant monitoring sites on which the models are based is required. We chose several criteria for data selection (Table S 2-1). After applying these criteria to the data set, 1,890 monitoring sites representing the recent mean nitrate concentration of shallow groundwater remained (Figure S 2-1). Descriptive statistics of the remaining nitrate concentrations can be found in Table S 2-2.

Spatial data

In previous studies related to the estimation of nitrate concentration in groundwater it was found that aquifer and soil properties, land use and N fertilizer are the most important predictors (Boy-Roura et al., 2013; Nolan et al., 2014; Ouedraogo and Vanclouster, 2016; Tedd et al., 2014; Wick et al., 2012). Since the approach of the present study should be applicable at the national level, an essential criterion for the selection of spatial predictors is their national availability in Germany. Hence, spatial predictors such as topography likely relevant on the small scale have not been considered for this analysis, also in the light of the coarse resolution (1 x 1 km) of the grid map. The Federal Institute for Geosciences and Natural Resources (BGR) and the Federal Agency for Cartography and Geodesy (BKG) provide a number of spatial thematic data (maps) on land use, hydrogeological and hydrological features as well as soil properties for Germany, which are used as predictors or independent variables (Table 2-1). These data are described in detail in Text S 2-1 of the Supporting Information.

Maps of the seepage water rate (BGR, 2003a) and groundwater recharge rate (BGR, 2003b) are provided by the BGR, the calculation is described in (Jankiewicz et al., 2005). The seepage water rate is calculated with the empirical TUB-BGR method and represents a land use differentiated balance between precipitation minus actual evapotranspiration and surface runoff. The mean annual groundwater recharge rate (BGR, 2003b) is based on a multi-stage regression method (Neumann, 2005). In a first step of this method, the baseflow index ($BFI = \text{base runoff} / \text{total runoff}$) is determined as the dependent variable according to slope, water network density, land cover, field capacity, groundwater distance and the proportion of direct runoff. Based on the results, two different model variants for low runoff and high runoff regions are developed. For low runoff regions the groundwater recharge rates correspond to the product of the BFI and the total runoff according to the water balance model BAGLUVA (Glugla et al., 2003). For high runoff regions, a second regression model estimates the groundwater recharge rate based

on the BFI calculated in the first step, the BAGLUVA total runoff and the groundwater distance as additional parameters.

The potential nitrate concentrations in seepage water and groundwater recharge are determined by a simple land use differentiated division of the N surplus (Bach et al., 2016) and the respective water quantities.

Table 2-1: Summary of the spatial environmental predictor variables.

Predictor variable	Units	Variable name	Variable type	Data source
Land use				
Urban Land	%	urban_land	numerical	LBM-DE2012 (BKG, 2016)
Arable Land	%	arable_land	numerical	
Grassland	%	grassland	numerical	
Forest	%	forest	numerical	
Special crops	%	special_crops	numerical	
N-surplus on agricultural land (mean 2011-2013)	kg ha ⁻¹ y ⁻¹	N_surplus	numerical	Bach et al. (2016)
Hydrogeology & Hydrology				
Hydrogeological units (9 variables)	(-)	hyd_unit	categorical	HYRAUM (BGR & SGD, 2015)
Seepage water rate	mm/a	SWR	numerical	SWR1000_250 (BGR, 2003a)
Groundwater recharge rate	mm/a	GWRR	numerical	HAD55_gwn1000_v1_raster (BGR, 2003b)
Nitrate concentration in seepage water	mg/l	SWR_conc	numerical	Calculated from seepage water rate and N-surplus
Nitrate concentration in groundwater recharge	mg/l	GWR_conc	numerical	Calculated from groundwater recharge rate and N-surplus
Soil conditions				
Soil groups (9 variables)	(-)	soil_unit	categorical	BGL5000 (BGR, 2005)
Field capacity (0-1 m soil depth)	mm	FC	numerical	FK10dm1000_250(BGR, 2015)
Humus content in top soil (0 -10 cm for grassland and forest; 0 - 30 cm for arable land)	%	HUMUS	numerical	HUMUS1000 OB(BGR, 2007)

Data pre-processing

On large scale, there is often only poor information on groundwater flow conditions, so that simplified methods have to be applied to determine potential catchment areas of, e.g. monitoring sites. Johnson and Belitz (2009) showed that a circular buffer can adequately represent the contributing area of a monitoring site where no detailed groundwater flow conditions are known. However, other studies indicate that contributing areas based on hydrogeological considerations and simplified hydrogeological flow models are superior and improve the predictive performance of statistical models (Mattern et al., 2009; Mfumu Kihumba et al., 2016). Therefore, we investigate the effect of five different buffer designs on the predictive performance of each statistical models. Each buffer design represents the contributing area of a given monitoring site. Three circular buffers with different radii (1,500 m, 1,000 m and 500 m) around the monitoring sites assume a homogenous influence of the surrounding area on the monitoring site (Figure 2-2). Furthermore, two wedge buffers with apertures of 45° and 90° oriented in groundwater flow direction consider a directed effect of the dominating flow condition (Figure 2-2). The groundwater flow direction is derived from a state-wide modelling of the groundwater surface for Hesse (Herrmann, 2010; Wendland et al., 2011). The spatial data is derived for each buffer design for all monitoring sites.

For subsequently prediction of groundwater nitrate concentration, we create a 1 x 1 km grid for the state of Hesse (21,754 grid cells) and perform the corresponding join processes with all predictors. The grid maps of all spatial predictors can be found in Figure S 2-2 to Figure S 2-5.

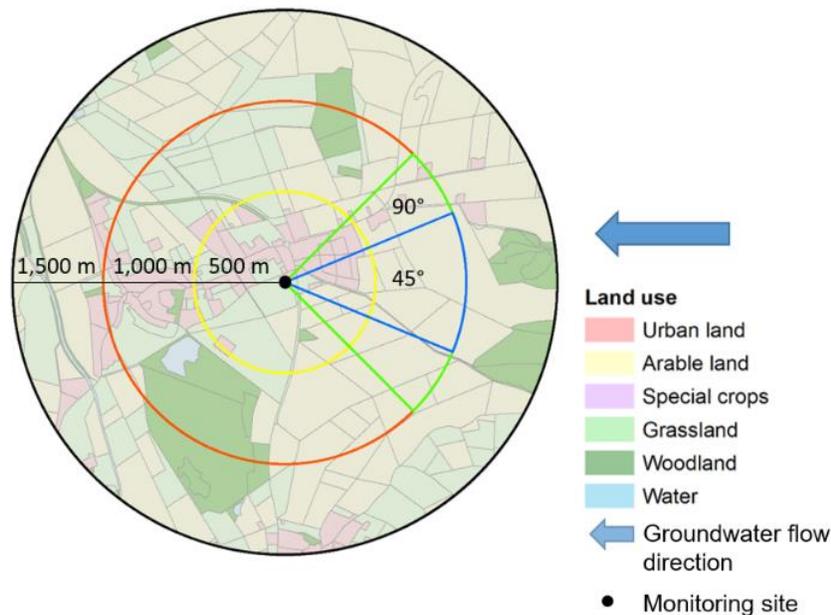


Figure 2-2: Five different buffer designs for the allocation of spatial data to the groundwater monitoring sites.

2.2.4 Statistical model framework

For the comparison of different statistical methods, we evaluate four statistical and machine learning methods for the spatial prediction of nitrate concentrations in groundwater: Multiple Linear Regression (MLR), Classification and Regression Tree (CART), Random Forest (RF), and Boosted Regression Trees (BRT). We use the R package ‘caret’ v. 6.0-78 (Kuhn, 2018) to train, compare and test the different statistical models. The model selection during the tuning process for CART, RF and BRT is made according to the lowest root mean squared error (RMSE). The individual models are described below, an overview of the methods and tuning parameters used can be found in Table S 2-3.

Multiple Linear Regression (MLR)

The MLR approach is an extension of a linear regression and uses multiple explanatory variables. We use nitrate concentration as dependent (or response) variable and all spatial variables (Table 2-1) as independent (or explanatory) variables. To optimize prediction performance of MLR and making the model interpretable it is necessary to identify the relevant explanatory variables to include in the model, which is also known as “subset selection” (Hastie et al., 2009). For this, we follow the stepwise forward variable selection according to the lowest Akaike information criterion (AIC) which is implemented in the ‘caret’ package.

Classification and Regression Tree (CART)

The methodology of Classification and Regression Tree (CART) (Breiman, 1984) is a flexible and nonparametric tool for data analysis. It is commonly used for modelling and exploring of complex datasets with either numeric or categorical predictor variables that can interact with each other and indicate nonlinear relationships (De’ath and Fabricius, 2000). CART uses recursive partitioning by split rules for single predictors, which creates homogeneous groups of the response variables (De’ath and Fabricius, 2000; Kuhn and Johnson, 2013). The tree structure of the methodology starts with the root node, which represents the entire data set. Each split creates two branches with nonterminal nodes, which end up in the terminal node or “leaves” giving a unique prediction for the response variable (Breiman, 1984; De’ath and Fabricius, 2000). With this approach, very large trees can arise with a tendency of over-fitting the data. Therefore, Breiman (1984) used a cost-complexity tuning method to prune the tree to smaller depth with the smallest error rate (Table S 2-3). The CART results allow a good interpretation of the dataset and provide a way to identify relevant predictors. Due to the simplicity of the methodology, the predictive performance is relatively low compared to other modelling techniques (De’ath, 2007; Kuhn and Johnson, 2013). Another disadvantage is the instability of a single tree that can occur with small changes in the data set affecting the split series (Hastie et al., 2009).

Random Forest (RF)

To overcome the disadvantage of CART, Breiman (2001) developed the ensemble method Random Forest (RF). The RF method uses the CART algorithm but in a way, that many trees are created. For each tree, a bootstrap or so-called out-of-bag (OOB) sample of randomly selected predictors is used. This technique is also known as bagging (bootstrap aggregation) (Breiman, 1996). The predictions of all trees of the “forest” are averaged to get an ensemble model prediction (Kuhn and Johnson, 2013). In the model training process, the number of trees and the number of predictors used at each split can be varied (Table S 2-3). This machine learning algorithm improves the model’s predictive performance by variance reduction and results in a more robust prediction. RF allows to quantify the impact of the predictors used in the ensemble. To do so, the differences of the predictive performances of non-permuted or permuted samples for each predictor of the OOB-samples are analysed (Breiman, 2001). The variable importance is measured in the mean increase in prediction performance.

Boosted Regression Trees (BRT)

Boosted Regression Trees (BRT) is another ensemble method, which strives to improve the model’s predictive performance. BRT is also based on the CART algorithm but additionally uses boosting algorithms, firstly developed by Freund and Schapire (Freund and Schapire, 1996). BRT differs mainly in its stage-wise forward process from RF. After an initial tree is fitted, further trees are adapted to the residuals of the preceding tree. Each iteration aims to minimize a loss function, a measure to represent a decrease in predictive performance. For the resulting BRT model (Table S 2-3), all trees are combined linearly (Elith et al., 2008; Hastie et al., 2009). We use stochastic gradient boosting, where at each iteration only a random subsample, similar to the bagging technique in RF, is used for fitting the new tree. This allows reducing the variance of the final model (Friedman, 2002).

Model training and evaluation of performance

The data set is divided into a training (80%, n=1,514) and a test data set (20%, n=376) for testing the model’s performances (Figure 2-3, Table S 2-2). The use of stratified random sampling, ensures that the distributions of the training and test data are approximately the same (Kuhn and Johnson, 2013). During model training three times repeated 10-fold cross-validation resampling technique is used for estimating model performance. In the ‘caret’ packages, the resampling technique can be defined consistently for all models. Training of the individual model parameters of the different models can then take place under the same basic conditions. At the end of the model training, we evaluate the model performance with the independent test data set to estimate the accuracy of the final prediction of the grid map.

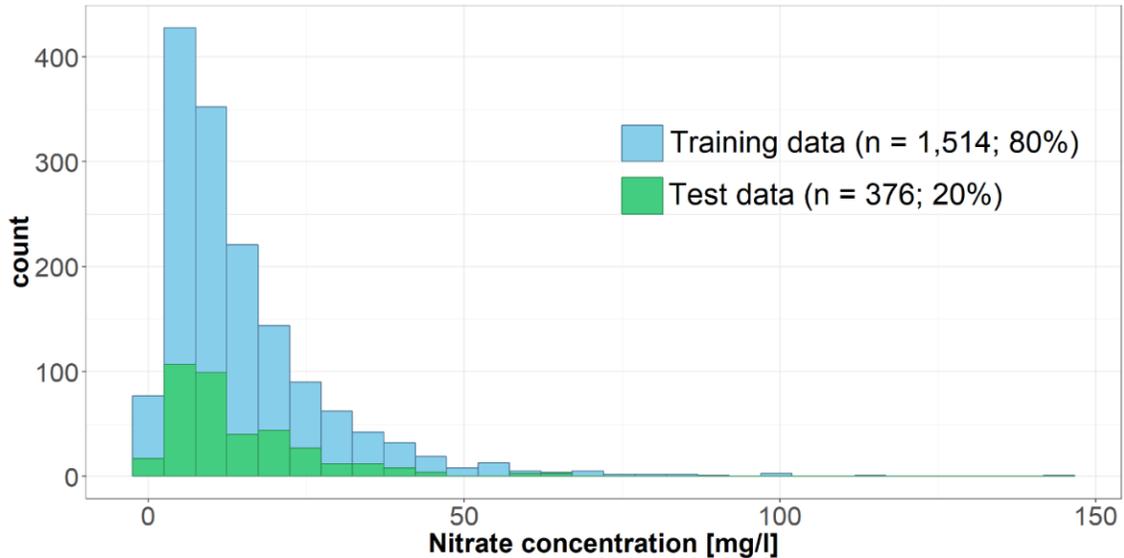


Figure 2-3: Histogram of nitrate concentration in groundwater (mean 2010 – 2017) in mg/l for training and test data set.

Because all models use a different way for assessing the variable importance, a normalized measure is used to compare the results of the four models. The relative variable importance measures are therefore scaled from 0 to 100, where the most relevant predictor is set to 100 (Kuhn, 2018).

In order to compare the different statistical models, a consistent way to evaluate the model performance is required. We calculate the objective functions mean absolute error (MAE) and the root mean squared error (RMSE) to express the average model prediction error in the unit of nitrate concentration. Furthermore, the commonly applied coefficient of determination R^2 is used, which is a measure for correlation and indicates the covariance in the model prediction.

2.2.5 Prediction

After all spatial predictors have been linked to the 1 x 1 km grid map of Hesse, a set of predictors identical to that used for model development is available for each grid cell. With the trained and tested models, a groundwater nitrate concentration can now be estimated according to the predictor values for each grid cell.

2.3 Results and Discussion

2.3.1 Model performance

The resulting predictive performance is measured by MAE, RMSE and R^2 of the particular statistical models. Each model type is trained based on the respective data sets resulting from the different contributing areas. The results for the mean model performances after training 30 models using three times repeated 10-fold cross-validation are presented in Figure 2-4. All model types based on the 1,000 m circular buffer zone show the highest R^2 , and except for CART, also the lowest RMSEs. CART shows the lowest RMSE based on the 500 m circular buffer zone. Furthermore, the MLR model shows the lowest MAE based on the 1,000 m circular buffer. For CART, RF and BRT the MAE based on the 1,500 m circular buffer slightly outperforms those considering the 1,000 m circular buffer. The results of the predictive performance based on the wedge buffers consistently show poorer goodness of fit values. Overall and in agreement with Wheeler et al. (2015), we find that the 1,000 m circular contributing area leads to the most favourable predictive performances. This indicates, that a circular buffer can adequately describe the contributing area of a monitoring site if no process-based information on the actual catchment area is available (Johnson and Belitz, 2009). A precise determination of the actual contributing area of each monitoring site, derived for example by a process-based representation of the groundwater flow conditions, would nevertheless be desirable, since in other studies this has led to an improvement of the predictive performance (Mattern et al., 2009; Mfumu Kihumba et al., 2016). In view of the application of our approach to the national scale of Germany, however, such information is likely not available in the near future.

The presentation of the following modelling results are based on the 1,000 m circular buffer zone (Figure 2-5 and Table S 2-4). The RF model performs best with the highest mean R^2 value of 0.55 and the lowest MAE (5.90) and RMSE (9.19). The overall best model achieves an R^2 of even 0.67. BRT is only slightly worse with an R^2 of 0.51, whereas CART and MLR show lower predictive performances with an R^2 of 0.42 and 0.39, respectively. This tendency of similar performance of RF and BRT, clearly outperforming CART and MLR is obvious for all objective functions we estimate. The predictive performance of the observed nitrate concentration versus the predicted nitrate concentration is plotted for each model in Figure 2-6 and summarized in Table S 2-4. The predictive performance values for the training data set show significantly better results than for the test data for all models. The values for the test data are in a similar range than the results of the cross-validation measurements. The difference in the predictive performance measures for training (RF: $R^2 = 0.94$) and test data set (RF: $R^2 = 0.54$) indicate that particular statistical ensemble methods such as RF and BRT but also CART tend to an overfitting of the data, which result in optimistic model performances. The predictive performance measured by cross-validation or by the test data set represents more realistic predictive performance values. With RF as the best performing

model, we can achieve a R^2 value of 0.54, which indicates a strong correlation of the observed and predicted nitrate concentration in groundwater. Also the BRT, CART and MLR with R^2 values in the test data set of 0.53, 0.41 and 0.35 show good to reasonable results for the spatial nitrate prediction based on large-scale spatial input data as predictor variables.

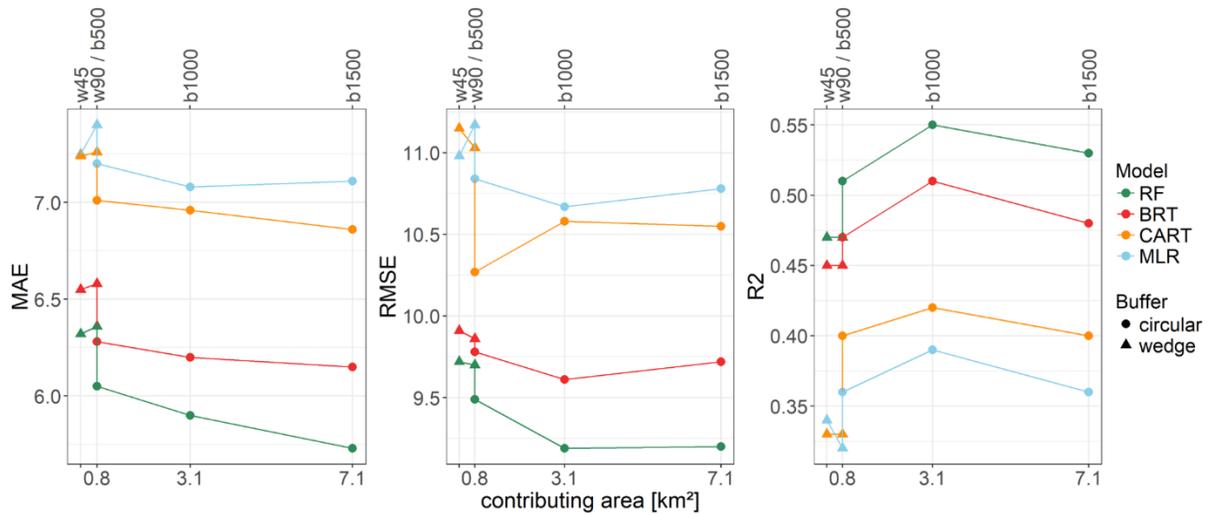


Figure 2-4: Comparison of the mean predictive performance MAE (Mean Absolute Error), RMSE (Root Mean Squared Error), R^2 (coefficient of determination) resulting from the 3 times repeated 10-fold cross-validation for MLR (Multiple Linear Regression), CART (Classification and Regression Tree), RF (Random Forest) and BRT (Boosted Regression Trees) based on the different contributing areas (circular buffer b1,500, b1,000, b500 and wedge buffer w45, w90), respectively.

In comparison with other studies (Boy-Roura et al., 2013; Nolan et al., 2014, 2015; Ransom et al., 2017; Tedd et al., 2014) we achieve a model performance of a similar or even better range. This is particular noteworthy, as we only use state-wide accessible mapped spatial predictors. Some of the aforementioned studies also considered point information like groundwater well depths, further hydrochemical measurements or previous model estimates of groundwater flow dynamics or “redox” conditions as an additional predictor.

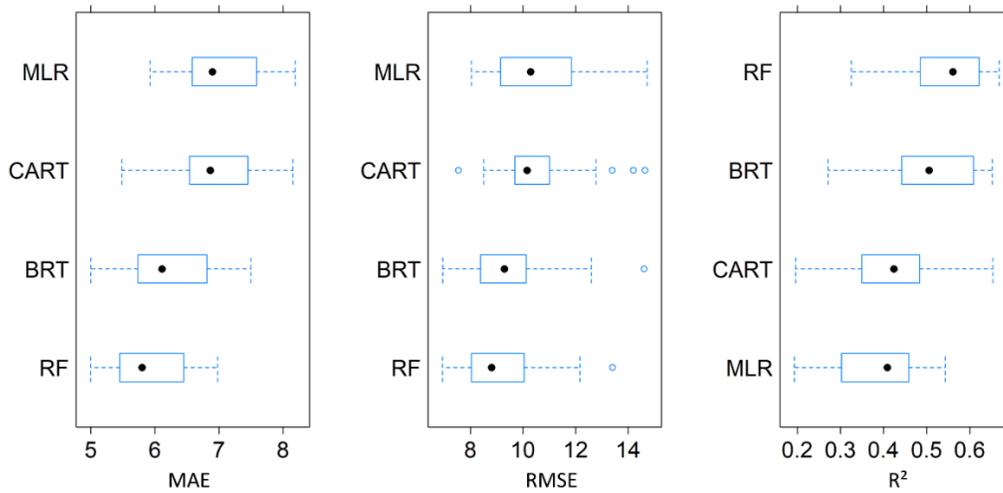


Figure 2-5: Comparison of predictive performance MAE (Mean Absolute Error), RMSE (Root Mean Squared Error), R² (coefficient of determination) of the 30 models resulting from the 3 times repeated 10-fold cross-validation for MLR (Multiple Linear Regression), CART (Classification and Regression Tree), RF (Random Forest) and BRT (Boosted Regression Trees), based on the 1,000 m circular buffer zone, respectively.

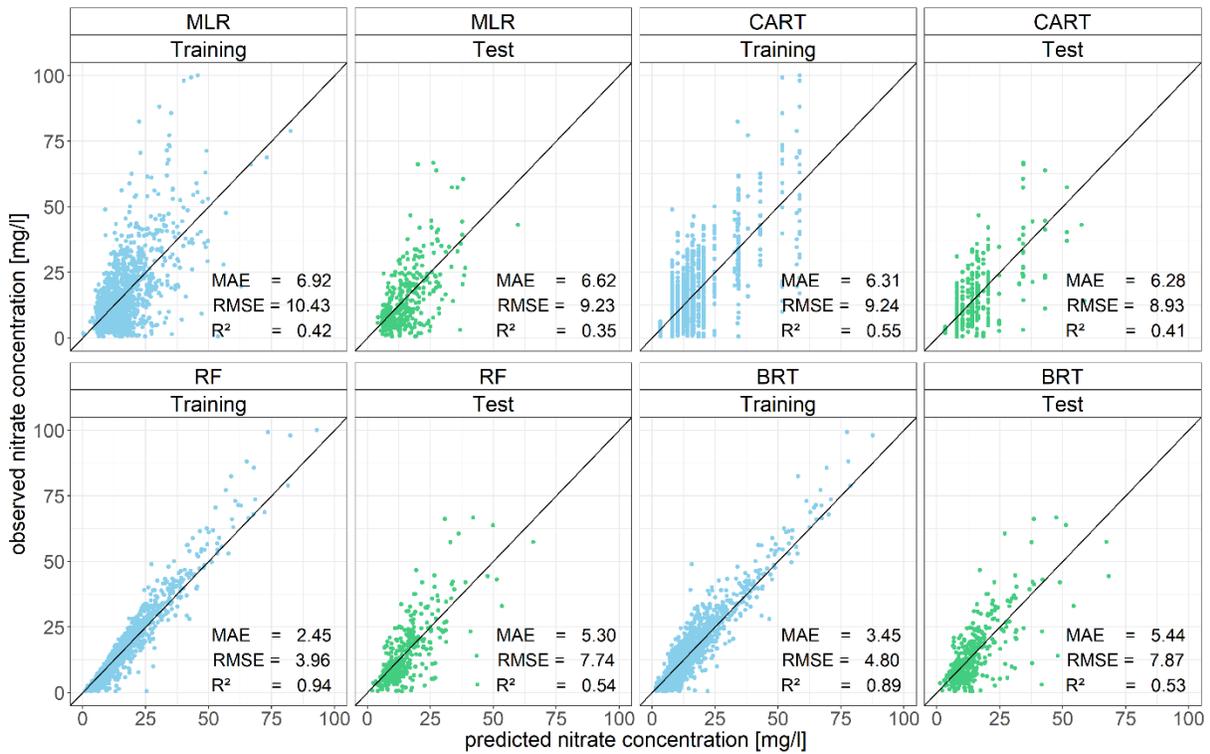


Figure 2-6: Predicted vs. observed nitrate concentration for test and training data for the models MLR (Multiple Linear Regression), CART (Classification and Regression Trees), RF (Random Forest) and BRT (Boosted Regression Tree) based on the 1,000 m circular buffer zone.

2.3.2 Variable importance

The spatial predictors from Table 2-1 are included in all four statistical models. The relative importance of the predictors are calculated for the respective models based on the 1,000 m circular buffer zone (Figure 2-7). For CART, the hydrogeological units, soil groups and groundwater recharge rate are the three most important predictors. The ranking for RF indicates that the hydrogeological units, the percentage of arable land and the N surplus are the three most influential predictors. For BRT and MLR also the percentage of arable land and the concentration in seepage water show the largest impact on nitrate prediction. The variety in the parameter rankings indicate that depending on the used statistical model type different predictors have stronger or less strong influence on nitrate prediction.

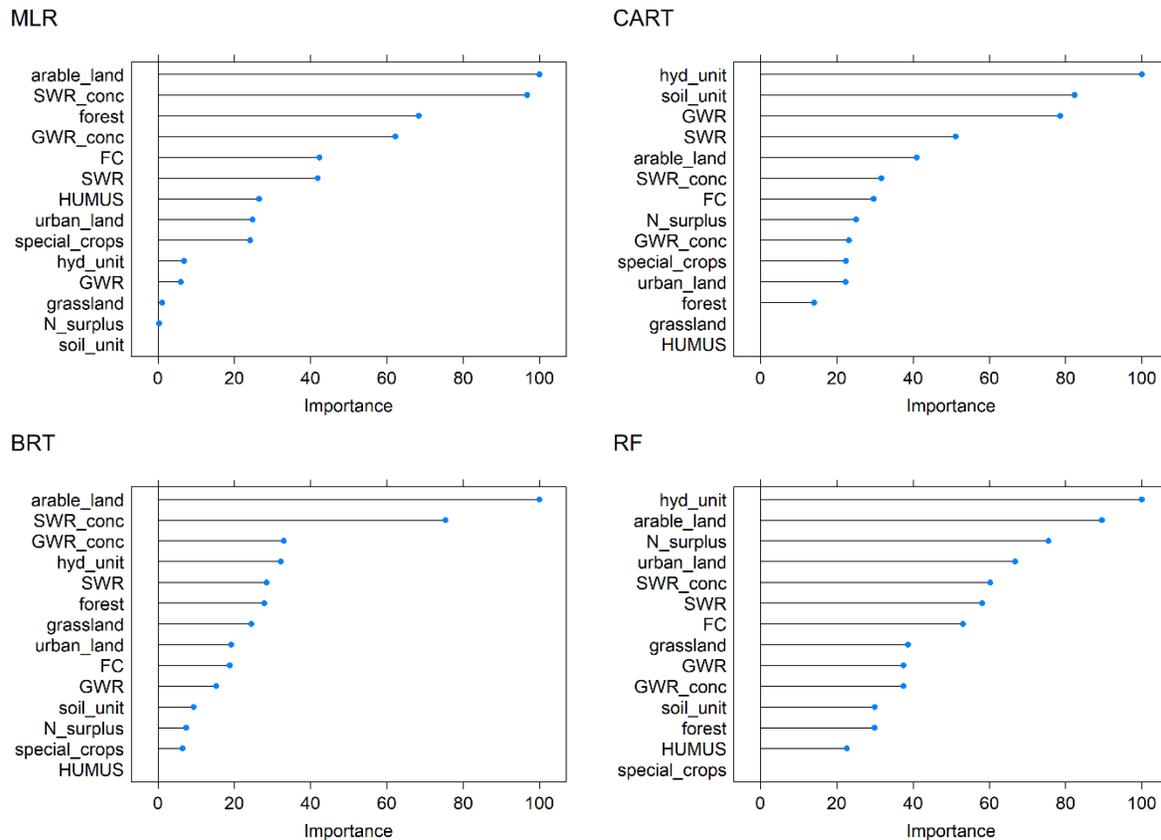


Figure 2-7: Predictors and their relative importance (in %) for prediction of nitrate concentration in groundwater in Hesse by the models MLR (Multiple Linear Regression), CART (Classification and Regression Tree), RF (Random Forest) and BRT (Boosted Regression Trees) based on the 1,000 m circular buffer zone.

The importance of specific predictors such as the hydrogeological conditions, soil types, land use or N surplus on nitrate prediction has also been found in other studies (Boy-Roura et al., 2013; Nolan et al., 2014; Ouedraogo and Vanclooster, 2016; Tedd et al., 2014; Wick et al., 2012). Frequently, the availability of spatial predictors is limited or differs from region to region, which may also result in a shift in parameter rankings. However, determining the most important predictor variables should also consider the general model performances and not only the ranking within each model.

2.3.3 Spatial prediction of groundwater nitrate concentrations

We predict the spatial distribution of nitrate concentrations in groundwater with all four models for Hesse (Figure 2-8). In general, all four models show similar patterns for the spatial nitrate distribution that only differ locally from each other. Areas with predominant agricultural land use like in the lowlands clearly show elevated nitrate concentrations. The spatial nitrate distribution mapped with the CART model show a less disperse pattern due to the constrained number of predicted nitrate concentrations resulting from one single decision tree. The proportion of the area in Hesse with nitrate concentrations exceeding the threshold of 50 mg/l nitrate (and the critical value of 37.5 mg/l nitrate) is 2.0% (8.3%), 2.4% (9.8%) 3.7% (11.4%) and 2.8% (10.3%) for the models RF, BRT, CART, and MLR, respectively. With regard to the impact of agricultural land use on nitrate concentrations, we have furthermore analysed only grid cells with more than 50% arable land. For this subset, the models RF, BRT, CART and MLR predict an exceedance of the threshold of 50 mg/l nitrate (and the critical value of 37.5 mg/l nitrate) on 8.0% (39.3%), 8.1% (34.5%), 11.6% (38.6%), and 13.1% (49.4%) of the area, respectively.

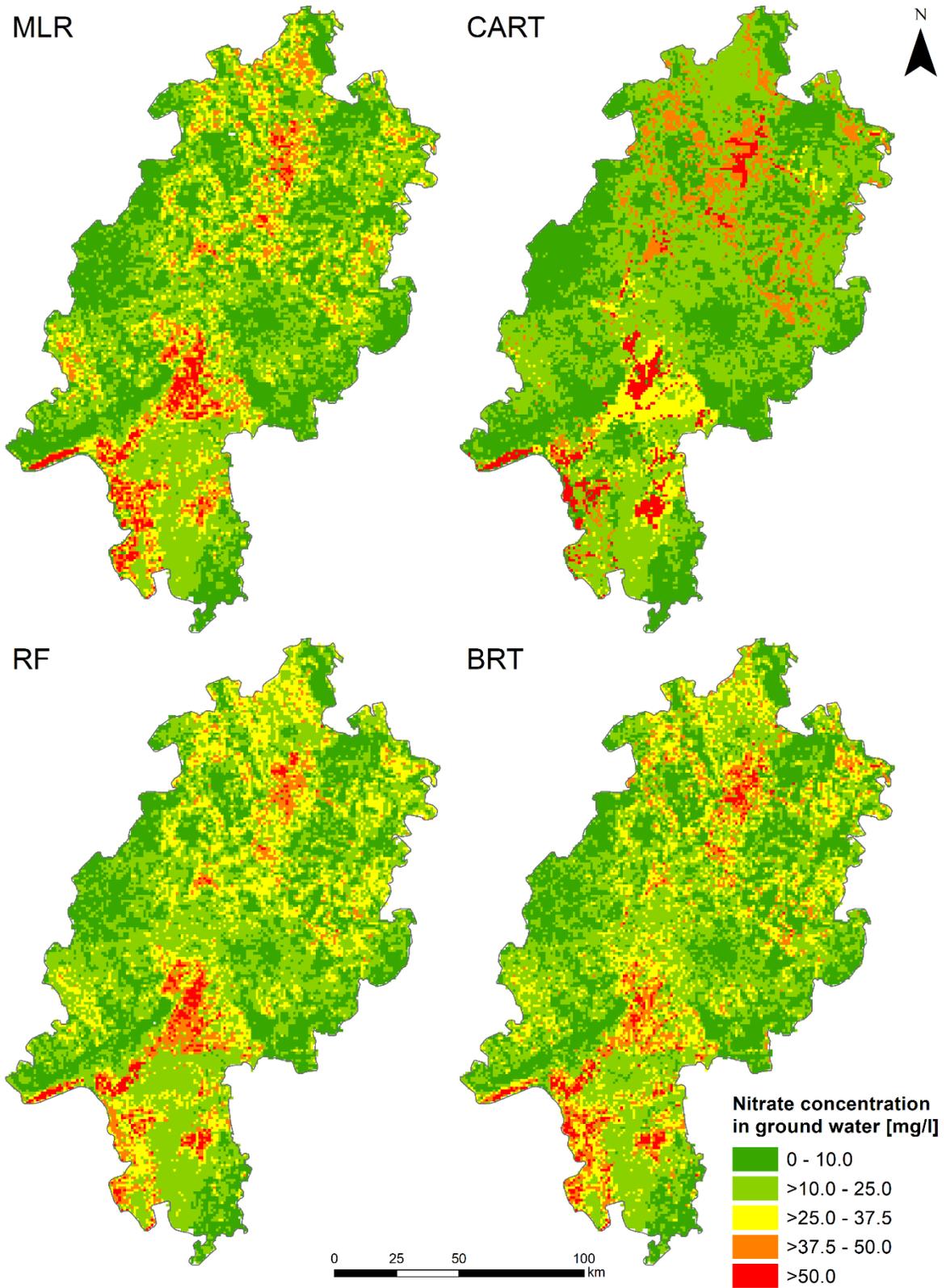


Figure 2-8: Predicted groundwater nitrate concentrations in Hesse with the models MLR (Multiple Linear Regression), CART (Classification and Regression Trees), RF (Random Forest) and BRT (Boosted Regression Tree) on a 1 x 1 km grid-map.

2.4 Conclusion

The GIS-based statistical approach allows the estimation of shallow groundwater nitrate concentration derived only from spatially mapped environmental parameters. The use of decision tree statistical models (CART, RF, BRT) enables the integration of numerical as well as categorical variables into one prediction approach. The models are trained with a point data set of groundwater nitrate concentrations with exclusively spatial environmental predictors. The models based on data sets of predictor variables derived by a circular buffer with 1,000 m radius leads to the highest predictive performance and thus best represents the contributing area around the nitrate concentration monitoring site.

All four model types result in reasonable to good predictions for the spatial distribution of nitrate concentration in groundwater. However, the RF model achieved the highest predictive performance with an RMSE of 7.74, MAE of 5.30 and R^2 of 0.54 for the test data set. For the RF model we identify the hydrogeological units, the percentage of arable land and the N surplus on agricultural used area as the most relevant predictors for estimating the spatial distribution of nitrate concentration in groundwater. Moreover, hydrological parameters and the occurrence of other land use types affect the prediction.

The determination of the actual contributing area of the monitoring sites through a process-based representation of the groundwater flow conditions could probably improve the predictive performance of the models. However, the inclusion of locally specific groundwater flow conditions is challenging on even larger scales due a lack of information. Furthermore, the present study does not take into account the hydrochemical environmental conditions in the groundwater body, which in any case influences redox-sensitive parameter such as nitrate. This information is often only available as point information on small scale.

The use of exclusively spatial available predictors as in our approach makes prediction of nitrate in groundwater on state level independently of further point data, which is a notable big advantage regarding the model's transferability and applicability in other regions where detailed point measurements are often not at hand.

Acknowledgements

The authors thank the Hessian Agency for Nature Conservation, Environment and Geology (HLNUG) for cooperation and providing the groundwater monitoring data of Hesse. The Federal Institute for Geosciences and Natural Resources (BGR) and the Federal Agency for Cartography and Geodesy (BKG) kindly provided the digital maps on land use, soils and hydrogeology. The project was funded by the German Federal Environment Agency (Umweltbundesamt, no. 37 1522 2200).

Supporting Information

Study Area

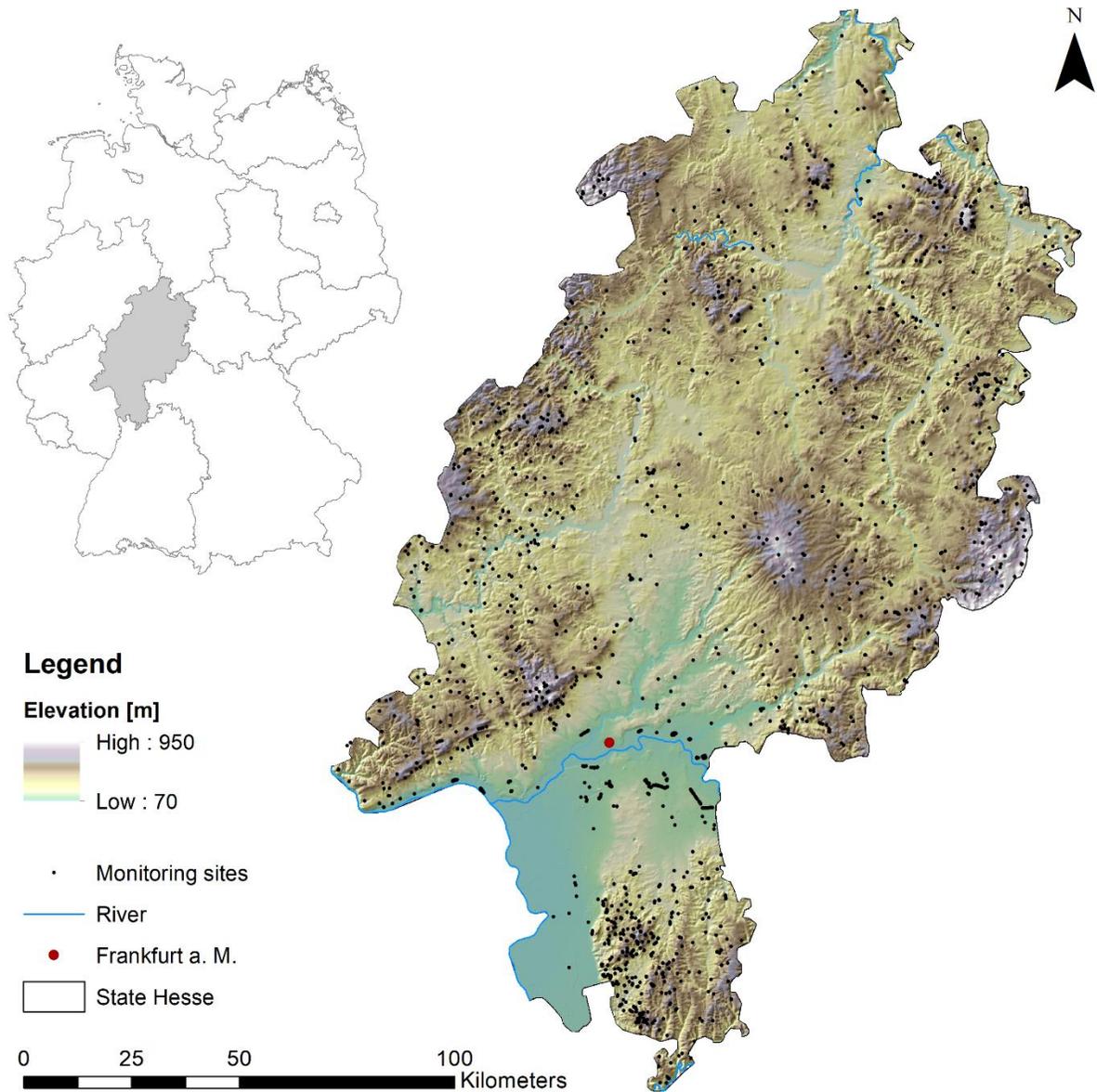


Figure S 2-1: Topographical map of Hesse, Germany (data source: © GeoBasis-DE / BKG, 2017); localization of 1,890 groundwater monitoring sites.

Data pre-processing

Point data

Table S 2-1: Criteria for data selection.

Criteria for data selection	Statement of reasons
Monitoring site: - welldepth < 100 m	- deep aquifers commonly integrate nitrate input from much greater area than covered by the 1 km radius buffer used for determination of predictor values
Measurements: - mean nitrate concentration from the period 2010 to 2017 - at least two measurements per site - mean nitrate concentration > 0.5 mg/l - if mean nitrate concentration exceeds 10 mg/l the coefficient of variation in the time series must be < 1	- restrict the time series to a near-term period of time (current decade) - ensures repeated sampling - different laboratories with various limits of detection (LoD) for nitrate; highest LoD is chosen for harmonisation of data used for statistical analysis - exclude highly variable nitrate concentration time series with potential outliers and measurement errors

Table S 2-2: Summary statistics of nitrate concentrations in groundwater for training and test data set.

	Training data nitrate concentration [mg/l]	Test data nitrate concentration [mg/l]
Count	1,514	376
Minimum	0.52	0.61
1 st Quantile	6.22	6.23
Median	10.55	10.62
Mean	14.90	14.23
3 rd Quantile	19.34	19.36
Maximum	144.71	66.66
Standard Deviation	13.74	11.39

Spatial predictors

Text S 2-1:

Land use and management

The land use map consists of 36 classes (BKG, 2016), which we aggregate to six dominant land use classes: forest (45%), arable land (24%), grassland (20%), urban land (10%), special crops (predominantly winegrowing) (0.3%) and water (0.7%). The agricultural land is mainly located in the low plains of the Upper Rhine Valley and the Lower Main Valley. The mountain ranges are predominately covered by forests. Densely populated areas are concentrated in the Rhine-Main Metropolitan Region around Frankfurt a. M. while the winegrowing areas occur in the south western parts of Hesse. Furthermore, we consider spatial information of the N surplus on utilized agricultural area (UAA) for administrative districts in Hesse from 2011 to 2013 (Bach et al., 2016). The average N surplus amounts to $57.6 \text{ kg N ha}^{-1} \text{ UAA a}^{-1}$.

Hydrogeology and soils

In Hesse, nine different hydrogeological units can be found (Figure S 2-3) (BGR & SGD, 2015). We used the hydrogeological units as categorical predictor variables. Equivalent to the main landscape units, there are predominately fractured aquifers with medium to very low hydraulic conductivities in the mountainous parts of Hesse and pore aquifers with high hydraulic conductivities in the lowlands.

The amount of seepage water in Hesse ranges from nearly zero mm/a in the floodplains of the Upper Rhine Valley up to 960 mm/a in some parts of the Vogelsberg (BGR, 2003a). The groundwater recharge rate shows a very similar pattern with lowest values of nearly zero mm/a in some parts of the Upper Rhine Valley and up to 340 mm/a in the Vogelsberg (BGR, 2003b). The higher values for the seepage water than the groundwater result from losses through interflow, which is much higher in the fractured mountainous areas than in the lowlands. Based on the N surplus data we use calculated nitrate concentrations (in mg/l) of the seepage water (SWR_conc) and the groundwater recharge (GWR_conc).

There are nine main soil groups in Hesse (BGR, 2005), which are classified according the soil substrate, water conditions, relief and climatic conditions. We use these as categorical predictors. Furthermore, we use information on field capacity of up to 1 m soil depth, with values of up to 500 mm in some alluvial zones or even 750 mm in some fens (BGR, 2015). Lower values of the field capacity of around 100 mm to 200 mm are predominately found in mountainous areas. We further use the top soil's humus content (%) as an indicator for possible nitrate degradation in soils (BGR, 2007). In the lowlands of Hesse the top soil humus content ranges from 0 to 3 %, in the mountainous areas from 3 to 6 %. Only in small areas of alluvial zones, in fens or areas with brown earth, which occur in the Odenwald or the Rhenish Massive, higher levels of 6 to 8 % occur.

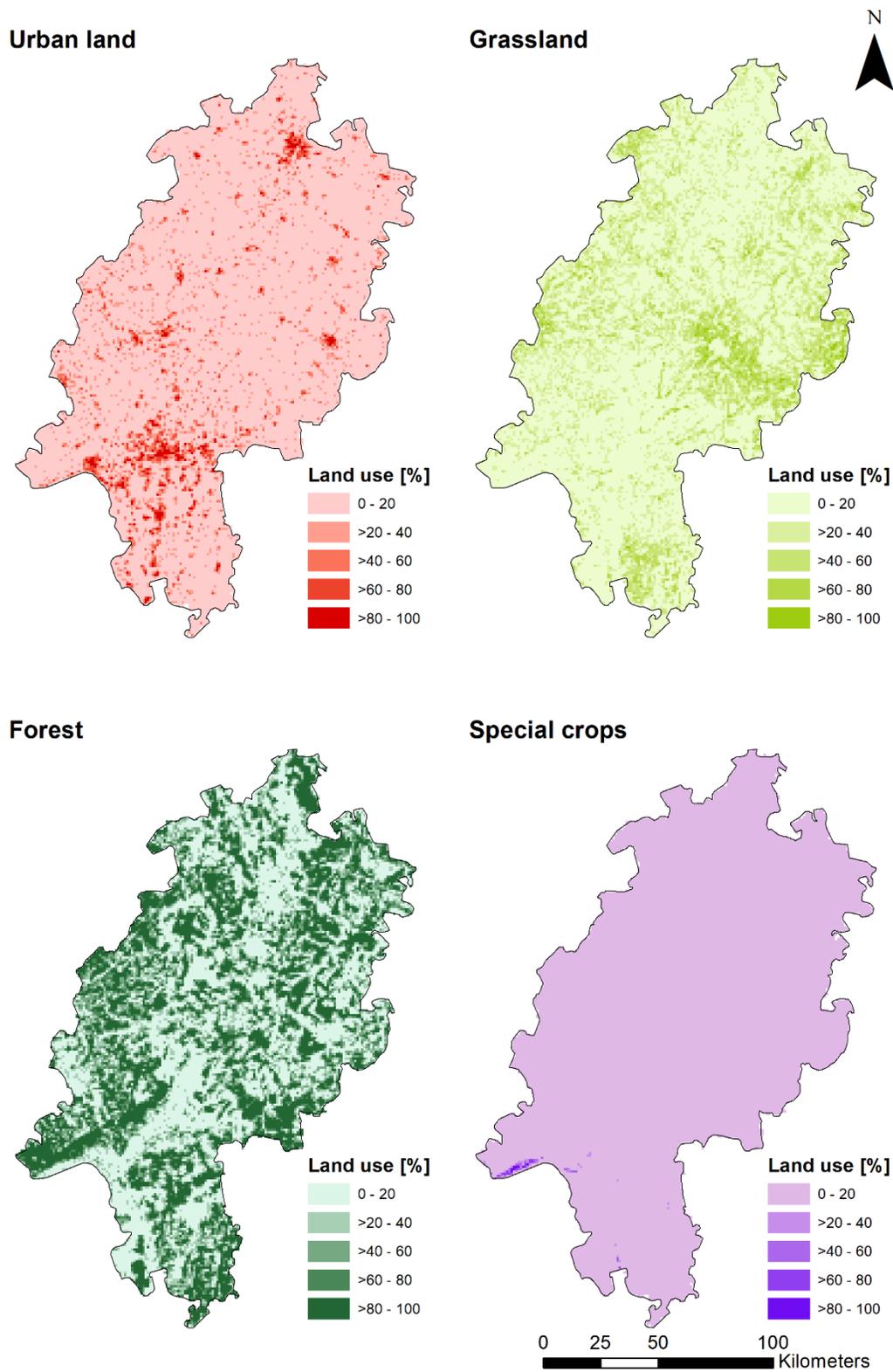


Figure S 2-2: Spatial predictors for 1 x 1 km grid map of Hesse; land use: percentage of urban land, grassland, forest and special crops.

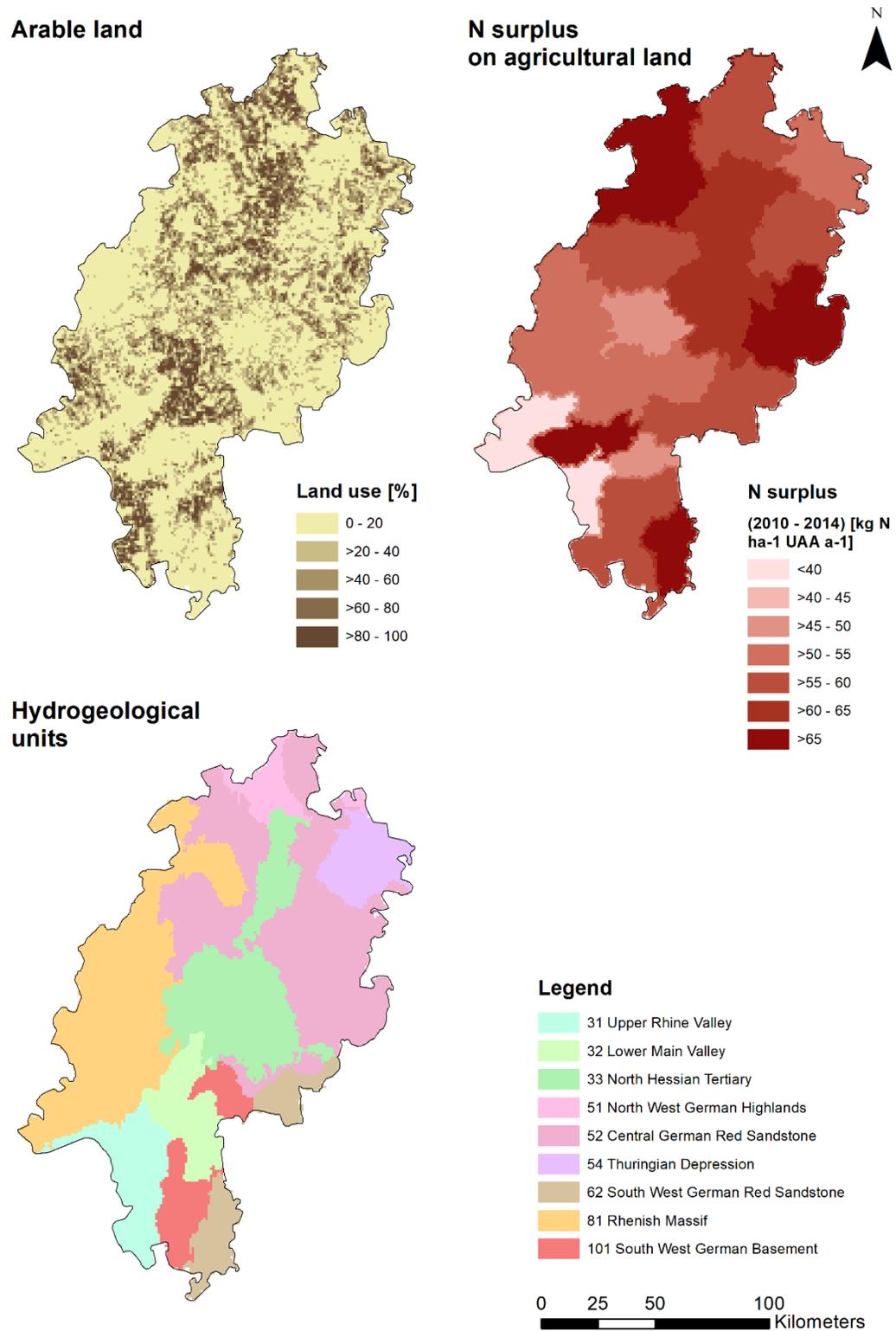


Figure S 2-3: Spatial predictors for 1 x 1 km grid map of Hesse; land use: percentage of arable land and N surplus on agricultural land; hydrogeology: hydrogeological units.

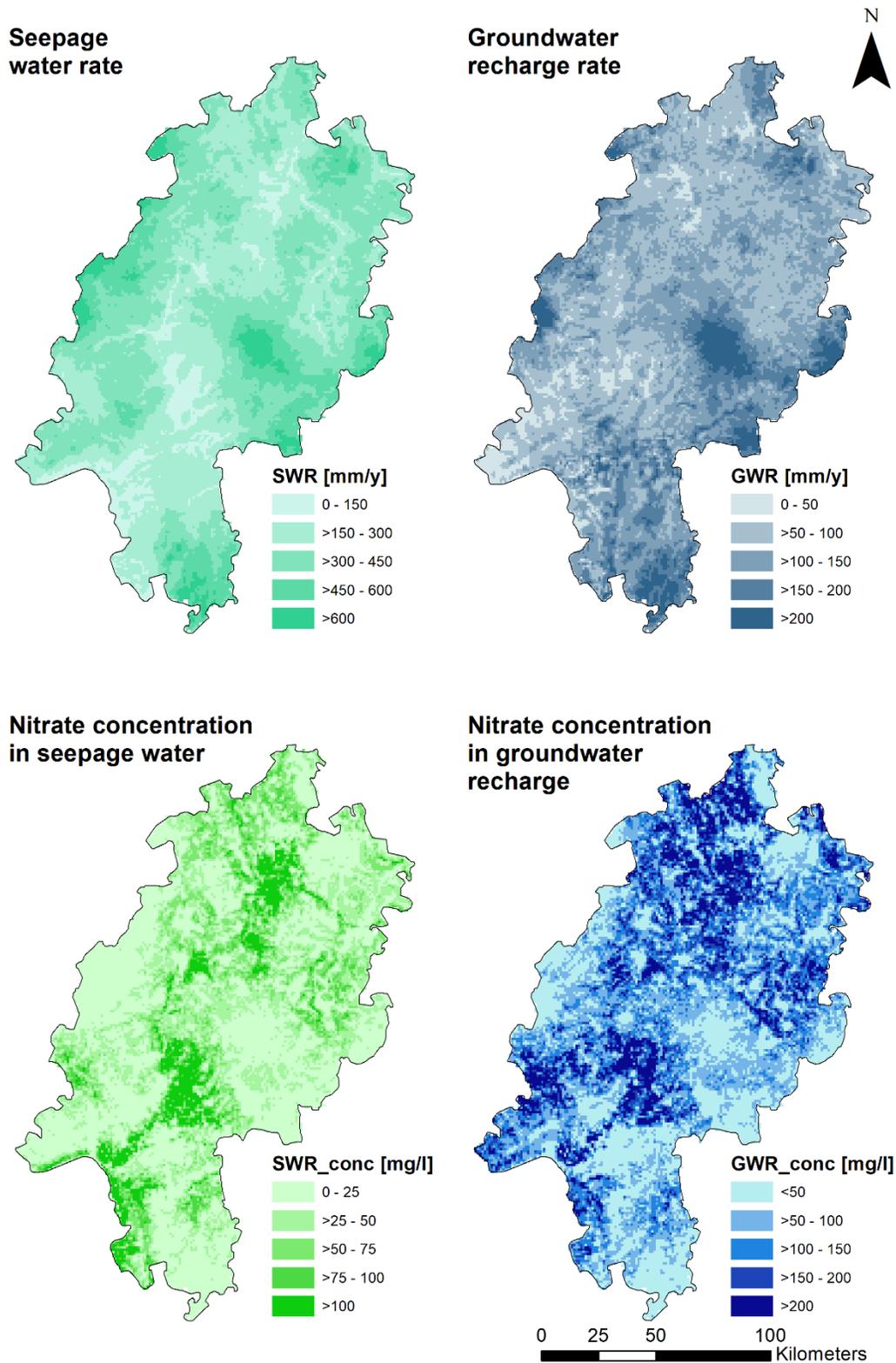


Figure S 2-4: Spatial predictors for 1 x 1 km grid map of Hesse; hydrogeology & hydrology: seepage water rate, groundwater recharge rate, nitrate concentration in seepage water and nitrate concentration in groundwater recharge.

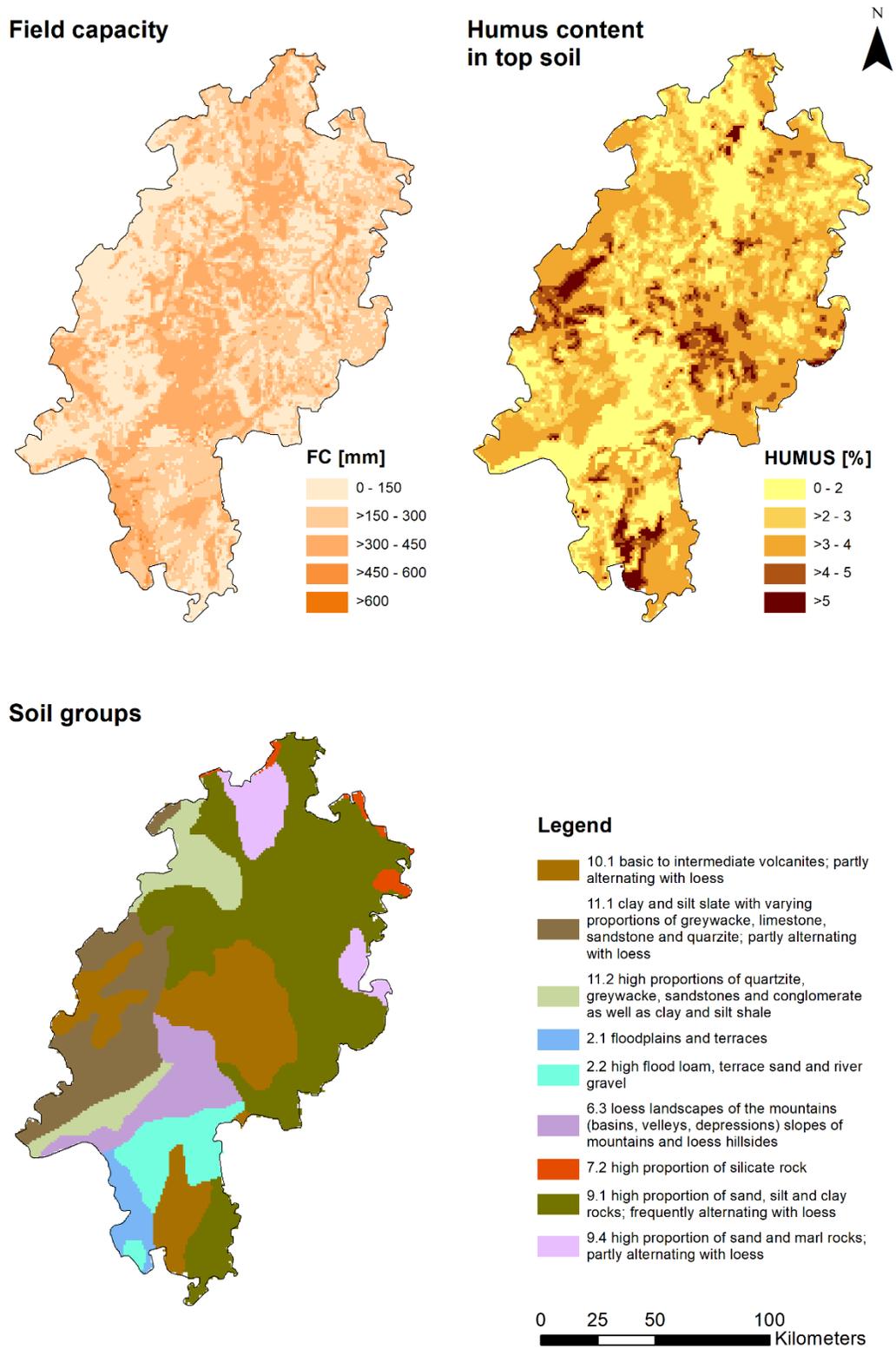


Figure S 2-5: Spatial predictors for 1 x 1 km grid map of Hesse; soil conditions: field capacity, humus content in top soil, soil groups.

Model training

Table S 2-3: Methods and tuning parameters of the models MLR, CART, RF and BRT used in the ‘caret’ package.

Model	Method	Tuning parameter		
		Variable	Description	Value/Range
MLR	lmStepAIC	-	-	-
CART	rpart	cp	complexity Parameter	0.005-0.05
RF	rf	mtry	randomly selected predictors	3-10
		ntree	number of trees	1000
BRT	gbm	n.trees	number of boosting iterations	50 - 1000
		interaction.depths	max. tree depth	6-10
		shrinkage	shrinkage (learning rate)	0.01-0.05
		n.minobsinnode	min. terminal node size	10

Model performance

Table S 2-4: Predictive performance values MAE (Mean Absolute Error), RMSE (Root Mean Squared Error), R² (coefficient of determination) for model training and testing for the models MLR (Multiple Linear Regression), CART (Classification and Regression Tree), RF (Random Forest) and BRT (Boosted Regression Trees) based on the 1000 m circular buffer zone.

Model	3 x repeated 10-fold Cross-validation			Training data set			Test data set		
	MAE	RMSE	R ²	MAE	RMSE	R ²	MAE	RMSE	R ²
MLR	7.08	10.67	0.39	6.92	10.43	0.42	6.62	9.23	0.35
CART	6.96	10.58	0.42	6.31	9.24	0.55	6.28	8.93	0.41
RF	5.90	9.19	0.55	2.45	3.96	0.94	5.30	7.74	0.54
BRT	6.20	9.61	0.51	3.45	4.80	0.89	5.44	7.87	0.53

3 Nation-wide estimation of groundwater redox conditions and nitrate concentrations through machine learning

This chapter is published in the journal *Environmental Research Letters* 15 (064004), 2020. <https://doi.org/10.1088/1748-9326/ab7d5c>

Knoll, L.^{1*}, Breuer, L.¹, Bach, M.¹

[1] Institute for Landscape Ecology and Resources Management (ILR), Research Centre for BioSystems, Land Use and Nutrition (iFZ), Justus Liebig University Giessen, Germany

* Correspondence to: Lukas Knoll (Lukas.Knoll@umwelt.uni-giessen.de)

Abstract

The protection of water resources and development of mitigation strategies require large-scale information on water pollution such as nitrate. Machine learning techniques like random forest (RF) have proven their worth for estimating groundwater quality based on spatial environmental predictors. We investigate the potential of RF and quantile random forest (QRF) to estimate redox conditions and nitrate concentration in groundwater (1 km x 1 km resolution) using the European Water Framework Directive groundwater monitoring network as well as spatial environmental information available throughout Germany. The RF model for nitrate achieves a good predictive performance with an R^2 of 0.52. Dominant predictors are the redox conditions in the groundwater body, hydrogeological units and the percentage of arable land. An uncertainty assessment using QRF shows rather large uncertainties with a mean prediction interval (MPI) of 53.0 mg/l. This study represents the first nation-wide data-driven assessment of the spatial distribution of groundwater nitrate concentrations for Germany.

3.1 Introduction

The increasing use of groundwater as the world's most important drinking water resource underlines the importance of protecting it from pollution (Foster and Chilton, 2003). In Germany and many other countries, the quality of groundwater is affected by elevated nitrate (NO_3) concentrations attributed predominantly to diffuse nitrogen losses in agricultural systems (EEA, 2018; Galloway et al., 2008). Nitrate pollution of water systems has far-reaching environmental consequences and poses a serious risk to human health (Erisman et al., 2013; Reis et al., 2016; van Grinsven et al., 2006). The EU Water Framework Directive 2000/60/EC (WFD) was introduced for integrated water protection, aimed at maintaining or restoring the 'good status' of groundwater bodies. With regard to nitrate pollution, this requires measures to reduce nitrogen inputs to meet the WFD quality standard of $<50 \text{ mg/l NO}_3$.

The EU Member States are obliged to submit a regular inventory of the groundwater status. The latest report shows that 25% of groundwater bodies in the EU have a poor chemical status, 18% primarily due to nitrate pollution (EEA, 2018). The responsible authorities need to develop interdisciplinary, integrative and holistic river basin management plans for the implementation of the WFD, which requires comprehensive mapping of the pollution of water systems and identification of areas under significant pressure (Kallis and Butler, 2001; Voulvoulis et al., 2017).

Estimating the nitrate concentration in groundwater depends on several factors and processes. Land use patterns and the amount of the nitrogen surplus on agricultural land provide indications of input sources. In addition, the hydrogeological and biogeochemical conditions of the aquifer determine the occurrence of nitrate in the water body. Complex natural attenuation processes occur during the vertical displacement of nitrate in the unsaturated zone and the lateral transport in the aquifer in the direction of groundwater flow (Rivett et al., 2008). Denitrification, as the most significant process of nitrate removal in the groundwater body, depends strongly on the prevailing redox conditions (Korom, 1992; Rivett et al., 2008). Denitrification occurs under anaerobic conditions if denitrifying bacteria and electron donors are present (Rivett et al., 2008). Such anaerobic conditions are characterised by an oxygen (O_2) concentration of $\leq 1\text{-}2 \text{ mg/l}$ and iron (Fe) concentration of $\geq 0.1\text{-}0.2 \text{ mg/l}$ (Kunkel et al., 2004, 2017; McMahon and Chapelle, 2008; Rivett et al., 2008). Several studies dealt with the large-scale estimation of redox conditions in groundwater (Close et al., 2016; Koch et al., 2019; Rosecrans et al., 2017; Tesoriero et al., 2015, 2017) and Ransom et al. (2017) identified it as one of the most important parameters for the estimation of nitrate concentrations.

Data-driven statistical approaches have proven their worth in simulating large-scale groundwater quality. Nationwide groundwater nitrate concentration maps were developed with regression models for China (Gu et al., 2013) and for the United States (Nolan et al., 2002; Nolan and Hitt, 2006). Machine learning

(ML) techniques are receiving more and more attention in the field of water research. Khalil et al. (2005) showed the strong predictive capabilities in groundwater contamination modelling of different algorithms, e.g. artificial neural networks or support vector machines. As another effective tool, random forest (RF) developed by Breiman (2001) is frequently applied in water resources issues (Tyrallis et al., 2019). RF is a very powerful tool for modelling and evaluating large complex data sets with nonlinear relationships between both numerical and categorical explanatory variables. It is further characterised by its robustness and increased predictive performance (Tyrallis et al., 2019). Comparisons of different machine learning techniques in terms of prediction performance have shown that tree-based models such as RF outperform others (Knoll et al., 2019; Wang et al., 2016). Such tree-based ML techniques have already been used in several studies on large-scale prediction of groundwater nitrate concentrations (Knoll et al., 2019; Nolan et al., 2014, 2015; Ouedraogo et al., 2019; Ransom et al., 2017; Rodriguez-Galiano et al., 2014, 2018; Tesoriero et al., 2017; Wheeler et al., 2015). In terms of the estimation of groundwater redox conditions, Close et al. (2016) and Wilson et al. (2018) performed linear discriminant analyses. Whereas, other studies also made use of ML for predicting groundwater redox conditions or the redox interface (Friedel et al., 2020; Koch et al., 2019; Rosecrans et al., 2017; Tesoriero et al., 2015; Wilson et al., 2020).

Due to the complex processes influencing groundwater quality, estimations are accompanied by uncertainties that need to be quantified (Grizzetti et al., 2015; Refsgaard et al., 2007). As one of the few, Rahmati et al. (2019) investigated the uncertainties in machine learning approaches for estimating groundwater nitrate concentrations and showed that it is important to consider both the performance and the uncertainty for model evaluation. Ransom et al. (2017) used bootstrapping for a boosted regression tree model and Koch et al. (2019) extended RF with geostatistics to assess uncertainties. By modifying RF to quantile random forest (QRF), Meinshausen (2006) provided a method to consider a full conditional distribution for the predictions instead of the mean only. This enables not only the additional output of median predictions, but also to estimate uncertainties by determining prediction intervals. Even though this method is not yet common in water research, it has already been used in digital soil mapping (Szatmári and Pásztor, 2019; Vaysse and Lagacherie, 2017).

Here, we provide for the first time a data-driven estimation of groundwater nitrate concentrations for the federal state of Germany, which is characterized by a large share of agricultural land use of around 50% of its area (360,000 km²). The objectives of our study are to develop a RF model combined with QRF to: 1) characterise the spatial distribution of redox conditions in groundwater, 2) estimate groundwater nitrate concentrations considering the redox conditions, and 3) determine prediction intervals as an uncertainty assessment.

3.2 Materials and Methods

The general modelling process involves four steps, i.e., setting up a database, data pre-processing, predictive modelling and uncertainty analysis with RF combined with QRF (Figure 3-1). The predictive modelling as well as the uncertainty analysis is performed with R (v. 3.4.1). All spatial data are processed in the Geographical Information System ArcGIS (v. 10.4).

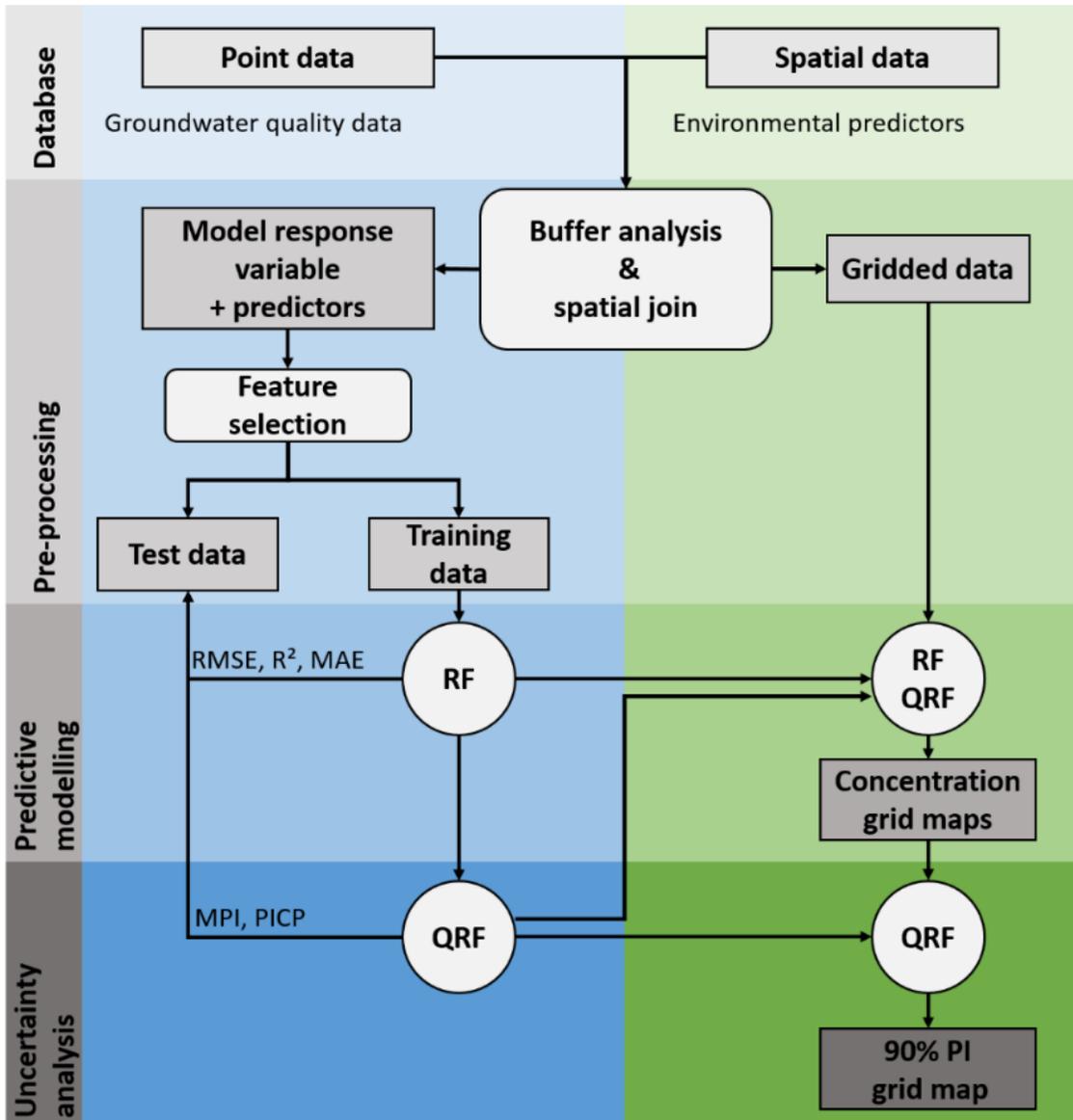


Figure 3-1: Scheme of the four-stage modelling process: set up of a database, data pre-processing, predictive modelling and uncertainty analysis. RF = Random Forest, QRF = Quantile Random Forest, RMSE = root mean squared error, R² = coefficient of determination, MAE = mean absolute error, PI = prediction interval, MPI = mean prediction interval, PICP = prediction interval coverage probability.

3.2.1 Database

According to WFD, the chemical status of groundwater is evaluated based on a representative monitoring network on which the EU Member States are required to report regularly. The WFD monitoring includes a surveillance network for the long-term and overall description of the groundwater status and an additional operational network. The monitoring sites selected in the WFD monitoring network represent the qualitative status of the respective groundwater body (EC, 2003). As the implementation of the WFD in Germany is the responsibility of the federal states, the design of the monitoring networks varies, which is reflected in particular in the numbers and distribution of the individual monitoring sites. In this study, we use a consolidated data set of groundwater concentration measurements of all WFD monitoring networks of the federal states (hereafter referred as WFD data set), which was provided by the responsible federal authorities. We considered a number of preconditions for the selection of the monitoring sites out of the WFD data set:

- Metadata and information on the sampling depth need to be available.
- Monitoring sites with a depth >100 m are excluded to avoid deep aquifers.
- Recent observation period (mean of 2009-2018).
- Monitoring sites with strong variations in the annual means are excluded (standard error of the mean (SEM) in the time series is >1 mg/l for O₂ and Fe and >10 mg/l for NO₃) to provide a representative concentration with robust means for the sampling period.
- Outliers in the O₂, Fe and NO₃ concentrations (>99% percentile) are removed from the data set.

Information on data acquisition and preparation is further described in Text S 3-1 of the Supporting Information. The groundwater quality data is statistically summarised in Table S 3-1 and Figure S 3-1, the spatial distribution of the monitoring sites is mapped in Figure S 3-2.

The groundwater concentration measurements (point data) for NO₃ (n=5,414), O₂ (n=5,837) and Fe (n=5,628) are used as the dependent variables. The redox condition classification is derived from O₂ and Fe concentrations and is taken into account in the nitrate model as predictor. To simulate groundwater nitrate concentrations, a set of 21 relevant spatial environmental predictors available throughout Germany is pre-selected according to expert knowledge, including 7 predictors on land use and management, 12 predictors on hydrogeology and hydrochemistry, and 2 predictors on soil conditions (Table S 3-2). A description of the data and information on data preparation can be found in Text S 3-2 and Figure S 3-3 to Figure S 3-6.

3.2.2 Pre-processing

Characterisation of redox conditions

The characterisation of the redox conditions is based on a four-point classification scheme according to LAWA (2018). O₂ concentration is the predominant determinant for the occurrence of denitrification, with concentrations <2 mg/l indicating anaerobic conditions and the occurrence of denitrification (Rivett et al., 2008). In this case, two points are assigned to the monitoring site. In the range of 2-5 mg/l O₂, there is only an average probability for the occurrence of denitrification, and one point is assigned. At O₂ concentrations >5 mg/l there are strong aerobic conditions, the probability of denitrification is low and no points are assigned to the monitoring site. Since denitrification preferably occurs under the presence of Fe concentrations ≥ 0.2 mg/l (Kunkel et al., 2004, 2017), we added an extra point. This results in four redox classes: (3) strongly anaerobic, (2) anaerobic, (1) intermediate and (0) aerobic (Table S 3-3).

Buffer analysis

The point data (O₂, Fe and NO₃ concentrations) are linked with the spatial predictors by means of contributing areas. The contributing area represents the potential catchment area of a monitoring site where the measured groundwater quality can be influenced by environmental factors. Since information on large-scale groundwater flow conditions is limited, we used a simplified procedure to determine contributing areas for each monitoring site, i.e., a circular buffer of 1000 m radius results in best model performances (Knoll et al., 2019). The spatial predictors are compiled within this buffer zone and are linked to the monitoring site. The spatial predictors are assigned to the individual cells of the grid map for Germany (358,171 grid cells) in the same way. The data pre-processing is further described in Text S 3-3.

Feature selection

First, a correlation matrix was used to test for strong correlations (pearson correlation coefficient $r > 0.75$) between the numerical predictors. This was the case for the leachate rate and the groundwater recharge rate, whereupon we removed the leachate rate from the data set. With the Boruta R package (Kursa and Rudnicki, 2010), an all-relevant feature selection algorithm, a final set of relevant predictors was selected, where rock type of the aquifer was removed as well (Table S 3-2). Predictor importance plots derived from the Boruta method can be found in Figure S 3-7. After feature selection, 11 predictors remain for the O₂ data set, 6 for the Fe data set, and 19 for the NO₃ data set.

3.2.3 Predictive modelling

For RF model training and evaluation, the data is first split into a training (80%) and test (20%) data set. In order to ensure an approximately equal distribution of the response variable in both data sets, stratified random sampling is applied (Kuhn and Johnson, 2013). For spatial prediction, the final model is trained with all data due to the relatively low density of monitoring sites. Details about the RF modelling approach is given in Text S 3-4. While RF considers only the conditional mean of a response prediction, QRF keeps all observations in this node and assesses the conditional distribution (Meinshausen, 2006). This continuous distribution function describes the range and variation of the response variable around the predicted mean and thus median values or an estimate for the uncertainty of the prediction can be obtained. We use the R package ‘caret’ v. 6.0-82 (Kuhn, 2018) for predictive modelling, in which the R package ‘randomForest’ (Liaw and Wiener, 2002) (method=‘rf’) and ‘quantregForest’ (Meinshausen, 2017) (method=‘qrf’) is implemented. In ‘caret’, we define the number of trees in the model (*n*tree=1000) and three times repeated 10-fold cross-validation (method=‘repeatedcv’, number=10, repeats=3) as resampling method in the model training function.

The evaluation of the model performance is based on three objective functions: the root mean squared error (RMSE), coefficient of determination (R^2) and mean absolute error (MAE). RMSE and MAE give an expression for the average model prediction error in the unit of the model response, where R^2 is a common measure indicating the variance in the prediction. We determine model performance for the training data set and validate it with the independent test data set. Because RF trees grow to a maximum size, it likely overfits the data in each tree, which results in essentially low prediction errors on the training data. Using cross-validation in the model training process avoids overfitting. Therefore the performance for the finally used model based on all data is determined through cross-validation. The relative predictor importance for each model is evaluated during the model training process by cross-validation based on the training data set with a normalised measure for importance from 0-100%, where the most important predictor is set to 100% (Kuhn, 2018).

After model training and setup of the grid map data set (1 km x 1 km) for Germany, the models are used to estimate the groundwater concentrations for each grid cell according to the assigned predictors. Before applying the model for NO_3 to the grid, the redox conditions must be determined for each grid cell according to the classification scheme based on the predicted O_2 and Fe concentrations (Table S 3-3).

3.2.4 Uncertainty analysis

To evaluate the reliability of the estimates, a prediction interval (PI) can be derived to determine the model uncertainty. The p -quantile of a distribution corresponds to the value not exceeded by the values of the response variable with a probability p . Based on the p -quantiles, the PI is defined as the range between the lower (PL^{lower}) and upper prediction limit (PL^{upper}). A 90%-PI indicates a range from the 5%-quantile to the 95%-quantile ($\frac{(1-p)}{2}$ to $\frac{(1+p)}{2}$), in which the true value is expected with a high probability ($p=0.9$). The mean prediction interval (MPI) and the prediction interval coverage probability (PICP) are suggested as statistical measures to evaluate the uncertainty of QRF (Rahmati et al., 2019; Shrestha and Solomatine, 2006). The MPI expresses the average width of all PIs, with lower values indicating lower uncertainties (Eq. [3-1]). With the PICP as a validation measure of the model uncertainty, the proportion of observed values (y_i) within the estimated PI is calculated for a given confidence level p (Eq. [3-2]) (Dogulu et al., 2015). Ideally, PICP corresponds approximately to the confidence interval p , otherwise the uncertainty is overestimated ($PICP > p$) or underestimated ($PICP < p$).

$$MPI = \frac{1}{n} \sum_{i=1}^n (PL_i^{upper} - PL_i^{lower}) \quad [3-1]$$

$$PICP = \frac{1}{n} \sum_{i=1}^n C, \quad \text{where } C = \begin{cases} 1, & PL_i^{lower} < y_i < PL_i^{upper} \\ 0, & \text{otherwise} \end{cases} \quad [3-2]$$

3.3 Results

3.3.1 Spatial distribution of groundwater redox conditions

In the pre-processing step, four classes characterising the prevailing redox conditions were determined based on the measured O_2 and Fe concentrations. Figure 3-2 shows a distinct influence of redox conditions on groundwater nitrate concentrations with a significant decrease with increasing anaerobic conditions. However, very high nitrate concentrations can still be seen in the anaerobic classes (2) and (3), despite potentially favourable conditions for denitrification. In these cases, either the denitrification potential is exceeded or the denitrification capacity is exhausted due to depleted electron donors (Green et al., 2008; Liao et al., 2012; Wilde et al., 2017).

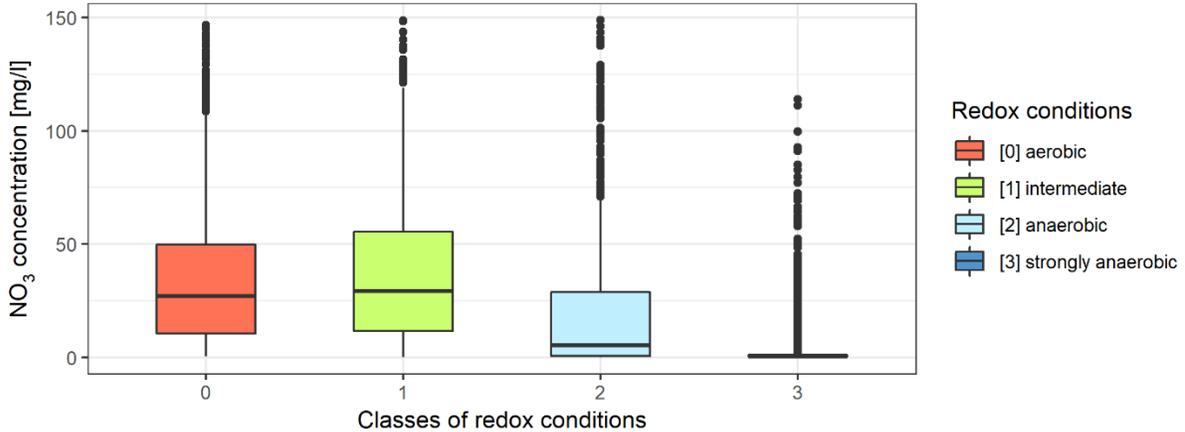


Figure 3-2: Boxplots of measured NO₃ concentration for four classes of redox conditions based on the WFD data set.

In order to map the redox conditions, the spatial distribution of O₂ and Fe groundwater concentrations must first be estimated. RF models for both parameters are trained with the pre-selected predictors. The importance of the predictors used in the RF model for O₂ (RF_{O2}) and Fe (RF_{Fe}) are ranked and depicted in Figure 3-3. For the RF_{O2} model, the hydrogeological units and the groundwater recharge rate are most relevant. In the RF_{Fe} model, the hydrogeological units dominate followed by field capacity. Tesoriero et al. (2015, 2017) also concluded that geology and parameters related to groundwater travel times, including groundwater recharge rate, are significant predictors for estimating aerobic conditions.

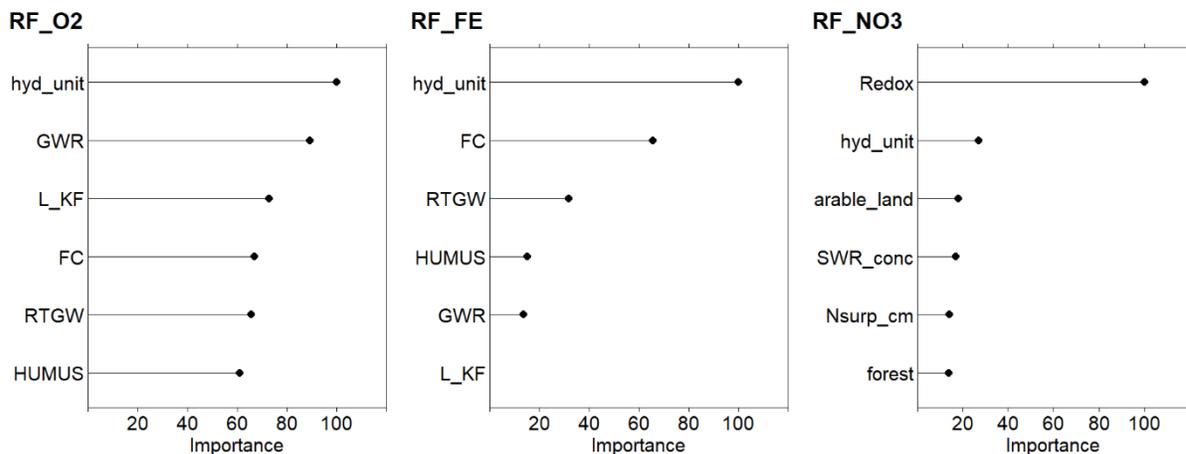


Figure 3-3: Importance ranking scaled from 0-100% of the six most important predictors for RF_{O2}, RF_{Fe}, and RF_{NO3} models based on the WFD training data. hyd_unit = hydrogeological units, GWR = groundwater recharge rate, L_KF = hydraulic conductivity, FC = field capacity, RTGW = groundwater residence time, HUMUS = humus content, Redox = redox conditions, arable_land = percentage of arable land, SWR_conc = potential nitrate concentration in seepage water, Nsurp_cm = calculated mean nitrogen surplus, forest = percentage of forest.

The results of the predictive performances of the RF models are summarised in Table 3-1. The RF_{O₂} model shows a very good predictive performance for the training data set. The RF_{O₂} model based on all data after cross-validation still achieves a good predictive performance (RMSE = 2.49 mg/l, R² = 0.53, MAE = 1.86 mg/l) similar to that of the test data set. The RF_{Fe} model also shows very good predictive performance for the training data set. However, for the independent test data, the RF_{Fe} model only has a reasonable predictive performance as has the RF_{Fe} model based on all data after cross-validation (RMSE = 3.00 mg/l, R² = 0.26, MAE = 1.66 mg/l). Similar studies, e.g. Tesoriero et al. (2017) and Rosecrans et al. (2017) can hardly be compared with our results, as they predict the probabilities of exceeding certain thresholds. Predictions versus observations for groundwater O₂ and Fe concentrations of the test data are shown in Figure 3-4. Despite its lower predictive performance compared to oxygen, iron provides an important additional contribution to the classification of the redox conditions.

The predicted grid maps of the groundwater concentrations of the redox parameters O₂ and Fe can be found in Figure S 3-8 and Figure S 3-9. For each map, the QRF model is used to determine the upper and the lower prediction limit (PL^{upper} = 95%-quantile, PL^{lower} = 5%-quantile) to derive the PI (Figure S 3-8 and Figure S 3-9). Density plots of the predicted redox parameters can be found in Figure S 3-11. The MPI of 7.4 mg/l for the RF_{O₂} model and 6.0 mg/l for the RF_{Fe} model indicate a very high uncertainty in the model prediction. The PICP values of 0.89 for RF_{O₂} and 0.90 for RF_{Fe} correspond to *p* (0.90), thus indicating that the uncertainties of the predictions are correctly assessed with QRF.

Table 3-1: Evaluation of model performance for the RF_{O₂}, RF_{Fe} and RF_{NO₃} models based on test data, training data and all data of the WFD data set.

Model	Test data			Training data			All data cross-validation (10-fold 3 times repeated)		
	RMSE [mg/l]	R ²	MAE [mg/l]	RMSE [mg/l]	R ²	MAE [mg/l]	RMSE (±sd) [mg/l]	R ² (±sd)	MAE (±sd) [mg/l]
RF _{O₂}	2.50	0.53	1.87	1.32	0.89	0.97	2.49±0.09	0.53±0.03	1.86±0.06
RF _{Fe}	3.09	0.24	1.65	1.62	0.84	0.88	3.00±0.25	0.26±0.05	1.66±0.11
RF _{NO₃}	20.90	0.51	12.97	8.66	0.93	5.39	20.12±0.97	0.52±0.04	12.71±0.51

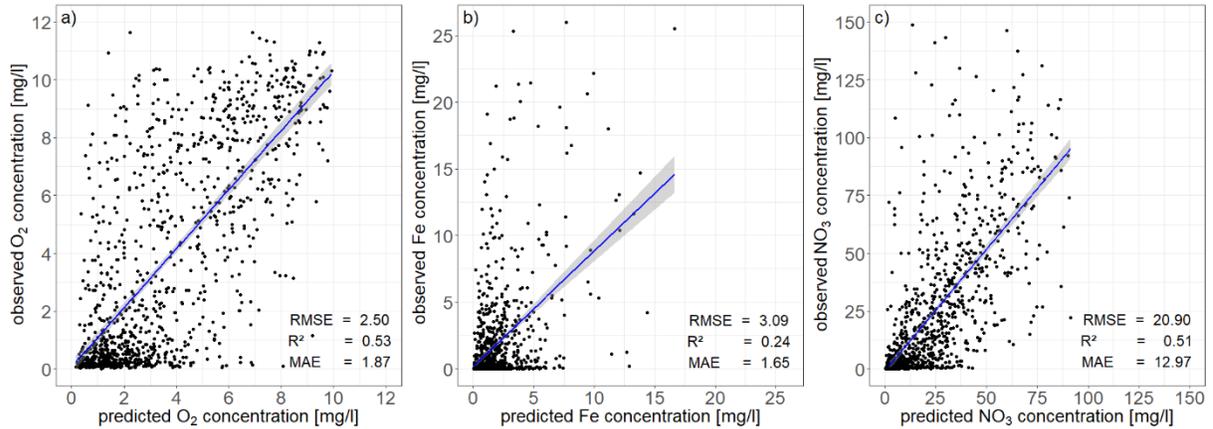


Figure 3-4: Predicted vs. observed groundwater concentrations for the RF_{O₂} (a), RF_{Fe} (b) and RF_{NO₃} (c) model based on test data of the WFD data set. The blue lines show the regression lines between the observations and the predictions along with their 95% confidence intervals.

According to the classification scheme in Table S 3-3, the spatial distribution of the redox condition is determined with the estimated O₂ and Fe concentrations in groundwater. Due to the low model performance of the RF_{Fe} model and its rather large uncertainties this could lead to misclassifications, especially in line with the low Fe concentration threshold for anaerobic conditions. As Fe concentrations show a strongly right skewed distribution, we use the median in favour of the average to describe the expected response. The resulting map of the derived redox conditions is shown in Figure 3-5. In the northern part of Germany as well as in some lowlands in the central and southern part predominantly strongly anaerobic to intermediate redox conditions occur. The southern part of Germany is dominated by intermediate to mostly aerobic redox conditions. This distribution indicates a higher denitrification potential in the unconsolidated aquifers of the North German Plain and local lowlands than in the consolidated units of central and southern Germany.

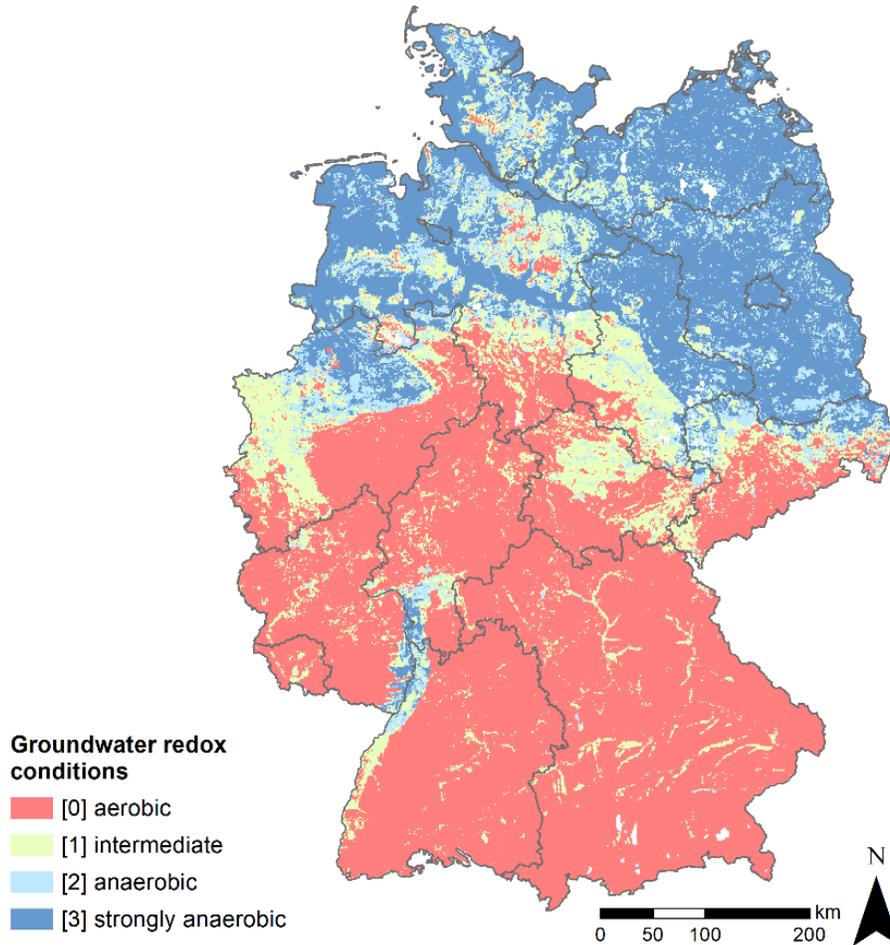


Figure 3-5: Spatial distribution of four redox classes for a 1 km x 1 km grid map of Germany derived from predicted oxygen and iron concentration in the period 2009-2018 based on the WFD data set (federal state borders: © GeoBasis-DE / BKG, 2017, Data license Germany - attribution - version 2.0 (www.govdata.de/dl-de/by-2-0)).

3.3.2 Spatial distribution of groundwater nitrate concentration

The RF model training for the NO_3 model (RF_{NO_3}) is based on a more comprehensive set of predictors compared to RF_{O_2} and RF_{Fe} (Table S 3-2). In addition to available spatial predictors, information on redox conditions derived from monitoring data of O_2 and Fe concentration (Figure 3-5) is used. The variable importance of the predictors used in the RF_{NO_3} model is ranked in Figure 3-3. The redox condition is by far the most relevant predictor, followed by the hydrogeological units and the percentage of arable land. The high relevance of the hydrogeological units and the percentage of arable land has also been found in Knoll et al. (2019). Ransom et al. (2017) used the redox parameters manganese and oxygen for nitrate prediction, which in that case turned out to be the most relevant predictors. Contrary to our expectations, the groundwater residence time within the RF_{NO_3} model is less relevant. Ransom et al.

(2017) added the groundwater age as a predictor and significantly improved the predictive performance of their groundwater nitrate model. However, the rather coarse resolution of groundwater age might be the reason for the low impact in our study.

The RF_{NO_3} model results in a good predictive performance for the training data set (Table 3-1). For the independent test data set, the model still provides good predictive performance as well as for the model based on all data after cross-validation (RMSE = 20.12 mg/l, $R^2 = 0.52$, MAE = 12.71 mg/l). Predictions versus observations for groundwater nitrate concentrations of the test data are shown in Figure 3-4. The RF_{NO_3} model tends to underestimate high concentrations (>50 mg/l) and has problems in the prediction of extreme values >100 mg/l. However, the performance of the model can be regarded as good, especially with respect to the large scale of the application and the relatively low density of monitoring sites. Other studies on large-scale prediction of groundwater nitrate concentrations result in comparable model performances (Knoll et al., 2019; Nolan et al., 2015; Ransom et al., 2017; Wheeler et al., 2015). The predicted grid map of the NO_3 concentration in groundwater is shown in Figure 3-6, a corresponding density plot can be found in Figure S 3-11. As shown before, results of the RF_{Fe} and the RF_{O_2} models contain partly large uncertainties. This could lead to misclassifications of the redox classes used for the NO_3 prediction. Since the redox condition is the most relevant predictor, we analysed its spatial sensitivity. In short, we added a random error of 5 to 30% to all classes in Figure 3-5, see details in Text S 3-5 and Figure S 3-12. The mean predicted NO_3 concentration varies more with an increase in misclassification. But even an assumed misclassification as large as 30% does not lead to a general questioning of our results. The predicted mean NO_3 concentration across Germany of 22.7 mg/l ranges within 21.0 to 24.9 mg/l under this assumption.

The QRF model is also used to determine the upper and the lower prediction limits for the RF_{NO_3} model. The derived 90%-PI is shown in Figure 3-6. Based on the test data, the MPI of 53.0 mg/l for the RF_{NO_3} model indicates a very high uncertainty in the model prediction. The PICP values for the RF_{NO_3} model of 0.91 corresponds to p (0.90), indicating correctly assessed prediction uncertainties with QRF. Other studies on prediction of groundwater nitrate concentration determined uncertainties in a similar range (Rahmati et al., 2019; Ransom et al., 2017). Koch et al. (2019) pointed out that the uncertainty can be significantly reduced with a more comprehensive data set. In line with this, we also tested a data set with additional monitoring sites ($n=13,038$ for NO_3) operated by several federal states in Germany, which resulted in a reduced, but still high uncertainty (MPI = 41.9 mg/l). Anyway, as these sites are not part of the WFD monitoring program, they were not further considered in this study.

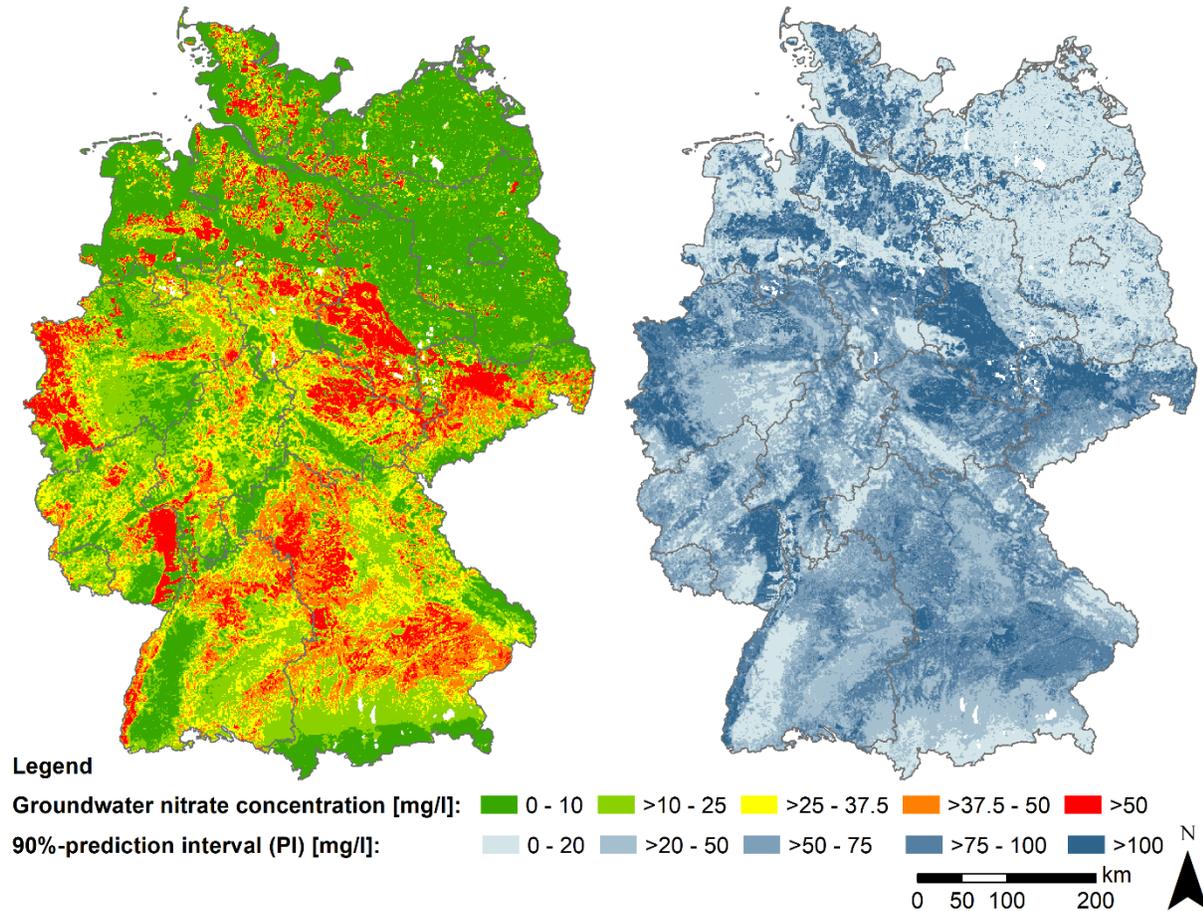


Figure 3-6: Spatial prediction (1 km x 1 km) of the mean groundwater nitrate (NO_3) concentration in the period 2009-2018 for Germany and the 90% prediction interval (PI) based on the WFD data set (federal state borders: © GeoBasis-DE / BKG, 2017, Data license Germany - attribution - version 2.0 (www.govdata.de/dl-de/by-2-0)).

3.4 Discussion

The high relevance of the redox conditions in the estimation of nitrate concentrations underlines the importance of a precise estimation of the hydrogeochemical conditions in the aquifers. Since hydrogeological units (Figure S 3-4) provide a decisive predictor for the estimation of the redox parameters, it is obvious that they are also of high relevance for the estimation of groundwater nitrate concentrations. Nevertheless, the anthropogenic impact on nitrate loads can only be represented properly if the inputs of nitrogen to the groundwater are correct. Thus, knowledge about percentage of agricultural land, the nitrogen surplus and the seepage rate are essential. Although natural turnover processes in the groundwater body are the most important determinant with regard to the groundwater pollution, the starting point for mitigation measures is to reduce the input, especially with regard to a finite reduction potential in the groundwater body.

It is apparent that the redox conditions are roughly divided into two regions (Figure 3-5). In the North German Plain and in some lowlands, anaerobic to strongly anaerobic conditions and thus preferably denitrifying conditions predominate. Conversely, aerobic to intermediate redox conditions tend to dominate in the southern and central German mountain regions. Similar findings are described in Kunkel et al. (2004) who conclude that aquifers with reduced nitrate degrading capacity occur in the North German Plain, while non-nitrate-degrading conditions prevail in aquifers of consolidated units. Hannappel et al. (2018) also found that the denitrification potential in quaternary and tertiary unconsolidated rocks is clearly higher than in the consolidated units.

The north-eastern part of Germany shows predominantly unpolluted areas in glacially formed Late Pleistocene units. These largely confined aquifers are characterised by low flow velocities, long groundwater residence times and strongly anaerobic conditions due to thick covering till layers and are therefore well protected against contamination (Merz et al., 2009; Wendland et al., 2008). Merz et al. (2009) also describe aerobic areas in groundwater recharge regions with high NO_3 concentrations, which only in places are identified by the predictions of the RF_{NO_3} model. However, predictions are accompanied by rather high uncertainties, so that higher concentrations are at least within the confidence limits of the model. In the north-western part of Germany, high NO_3 concentrations occur in the Geest regions formed in the Middle Pleistocene, whereas in the lowlands the groundwater is almost nitrate-free, which is in line with Wriedt et al. (2019). The redox conditions tend to be less anaerobic to intermediate in the Geest areas in contrast to the lowlands where strong anaerobic conditions prevail, which was also reported by Eschenbach et al. (2018). Conversely, lowlands with sand and gravel deposits in the middle and southern part of Germany often show nitrate-polluted groundwater, due to a lack of cover layer, weakly anaerobic to intermediate redox conditions and intensive agricultural land use (Grimm-Strele et al., 2008; Knoll et al., 2019; Kuhr et al., 2013; Wendland et al., 2008). For the pore aquifers of the northern Alpine foothills, the RF_{NO_3} model predicts large areas of nitrate-polluted groundwater, despite generally well-developed cover layers and the associated high protective function of the aquifers in this region. The tendency to overestimate groundwater nitrate concentration in this region can be due to the very low density of monitoring sites. The uncertainties in this area are also high. In the consolidated units of central and southern Germany, fractured and karst aquifers are generally more vulnerable to contamination and predominantly indicate higher groundwater nitrate concentrations, which is also in line with a lower nitrate reduction potential due to aerobic to intermediate redox conditions (Hannappel et al., 2018). In these consolidated regions, the high nitrate concentrations in the groundwater correspond to areas with high nitrogen surpluses caused by intensively agriculturally used depressions. Only the mountain regions with extensive forest cover usually show no groundwater nitrate

pollution. Based on the described approach, about 10% of the area of Germany exceed the threshold of 50 mg/l NO₃.

In general, the uncertainties are large (MPI = 53.0 mg/l), especially in the regions where nitrate concentrations are high or where strong variations between very high and low concentrations exist. Ransom et al. (2017) also determined the highest uncertainties in the nitrate-polluted areas.

A possible source for prediction errors and high uncertainties is linked to the available predictors. A more precise assessment of groundwater residence times could lead to better results and lower uncertainty. However, there is no nation-wide detailed database on the depth to the groundwater surface and flow conditions or the thickness of cover layers, which would be important indicators for residence times and the protective function of the aquifer. In addition, the integration of further predictors, such as a detailed description of the geochemical character of the aquifer or detailed information on input sources (that are currently not available on large scales) could improve the model performance. Uncertainties and prediction errors may also arise from the underlying dependent variable. Investigations of the spatial distribution of redox conditions in aquifers have shown a decrease in aerobic conditions with depth (Close et al., 2016; Rosecrans et al., 2017). Due to this hydrogeochemical zoning of aquifers, the concentrations of the groundwater samples strongly depend on the horizon in which the wells are screened and generally the NO₃ concentrations decrease with increasing well depth (Ransom et al., 2017; Wheeler et al., 2015). Tesoriero et al. (2017) also showed a higher probability for high nitrate concentrations near the groundwater surface. Wriedt et al. (2019) reported a decrease in nitrate concentrations with depth, but they concluded that a vertical gradation of depth horizons cannot adequately describe the hydrogeochemical zoning of the aquifers. Long screened sections can also lead to mixing of groundwater from the aerobic and anaerobic zone during sampling. Adjacent measuring points, which supply samples from different horizons, thus lead to variations in the data set, likely causing high prediction uncertainties.

Finally, it should be pointed out that the results of the regionalized groundwater nitrate concentration, and thus the identified areas with NO₃ >50 mg/l (10%), presented in this study cannot be compared one-to-one with studies focusing on other research questions. The classification of the chemical status of the groundwater in Germany according to the WFD showed that even 27.1 % of the groundwater bodies did not meet the WFD quality standard for nitrate (UBA, 2017a). It should be taken into account that the WFD-assessment is based on the precautionary principle, i.e. the results consider extreme concentrations which are potentially harmful. Accordingly a groundwater body (average size around 316 km²) is entirely classified as 'poor status' if even one of the monitoring sites or a defined proportion of the area within the groundwater body does not meet the quality standard (GrwV, 2010). Thus, a much higher proportion

of the area is designated as nitrate-polluted compared to our analysis, where the spatial assessment is based on a much higher resolution and mean values excluding outliers. In the authors' view, which methodology most adequately characterises groundwater quality (with regard to nitrate) depends on the scientific context and the research objective.

3.5 Conclusion

This study shows the potential of machine learning applications such as RF and QRF in the field of water research on a national scale. We presented the first Germany-wide assessment of groundwater redox conditions and nitrate concentrations with a resolution of 1 km x 1 km using a uniform data-driven approach based on spatial environmental predictors. In addition to a large influence of redox conditions, hydrogeological units and the percentage of arable land on nitrate estimation, high prediction uncertainties were determined. A more detailed and comprehensive database stratified according to well depths would likely enhance the prediction of NO₃ concentrations in groundwater. Adaptations of national groundwater monitoring networks with extensive monitoring in different depth horizons would be a major step forward, not only for Germany. However, the results of this study can contribute to further research on nation-wide scales, e.g. the calculation of national nitrogen targets or quantifying nitrogen flows from groundwater into surface waters and for the identification of vulnerable regions in which detailed investigations regarding the implementation of mitigation measures are necessary.

Acknowledgements

The project underlying this publication was coordinated and funded by the German Environment Agency in the framework of the environment research plan of the Federal Ministry for the Environment, Nature Conservation and Nuclear Safety (project no. 3715 22 2200). We are very grateful to German Working Group on water issues of the Federal States and the Federal Government - Committee for Groundwater and Water supply (LAWA-AG) and the Water Management Authorities of the federal states for providing the groundwater data. We also declare that no conflicts of interest exist in the submission of this manuscript.

Data availability

The data that support the findings of this study are available from the corresponding author upon reasonable request.

Supporting Information

Data Base

Text S 3-1:

The administration, maintenance and the quality management of the groundwater data is the responsibility of the respective 16 federal states of Germany. The data were requested and provided centrally via the LAWA-AG (German Working Group on water issues of the Federal States and the Federal Government – Committee for Groundwater and Water supply). In the first step, the data sets of the individual states are transferred into a uniform structure in a database. Monitoring sites with missing metadata are removed from the data set. While some states have already provided annual averages, for others they have to be calculated before calculating the average for the past decade 2009-2018. Outliers and time series are removed according to the criteria set out in the Chapter 3.2.

Since the planning, design and implementation of monitoring is the responsibility of the federal states, the monitoring networks differ in their density and/or representativeness of the various land use types. For example, the density of monitoring sites per 1,000 km² varies (except for the federal city-states Bremen, Hamburg and Berlin) from 5.6 in Baden-Württemberg to 40.7 Rhineland-Palatinate. The mean density is 20.8 sites/1,000 km².

The monitoring sites of the WFD data set refer to the hydrogeological conditions of the upper main groundwater body. We exclude monitoring sites with a depth >100 m to avoid an influence from deep aquifers. A further vertical zoning by depth classification of the measuring sites is not applied in view of the small number of measuring sites. We attempt following a parsimonious modelling approach in our nationwide assessment. However, this simplification leads to uncertainties because, for example, different depth horizons are taken into account. Ultimately, this leads to a more average estimate, with less extremes. This should be considered when interpreting our results.

Table S 3-1: Summary statistics of mean groundwater concentration (O_2 , Fe and NO_3) [mg/l] of the period 2009-2018 for the WFD data set.

Statistics	O_2	Fe	NO_3
n	5,837	5,628	5,414
min.	0.00	0.002	0.01
1 st quantile	0.27	0.010	0.63
median	1.72	0.090	7.80
mean	3.52	1.614	21.46
3 rd quantile	6.81	1.530	33.45
max.	12.27	27.42	148.78

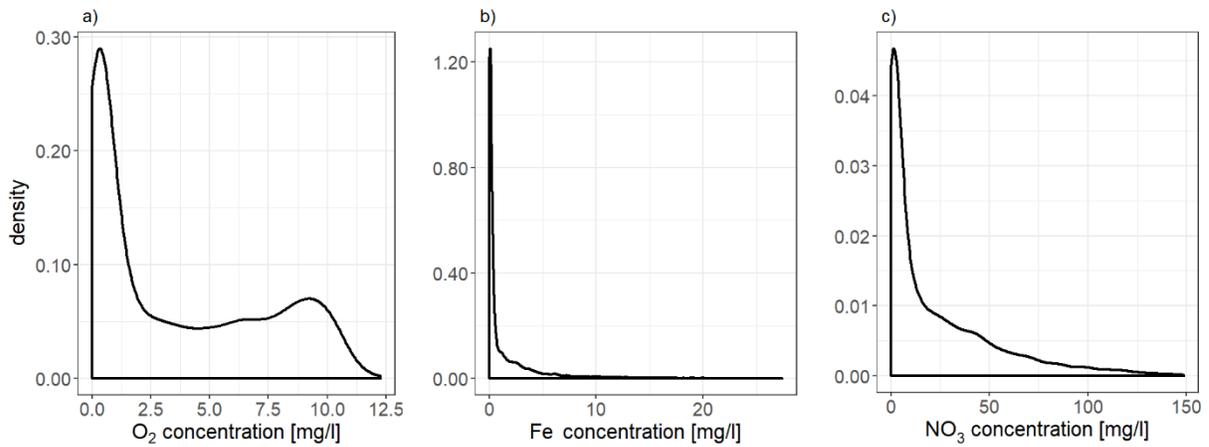


Figure S 3-1: Density plots of mean groundwater concentrations of a) O_2 , b) Fe and c) NO_3) of the period 2009-2018 for the WFD data set.

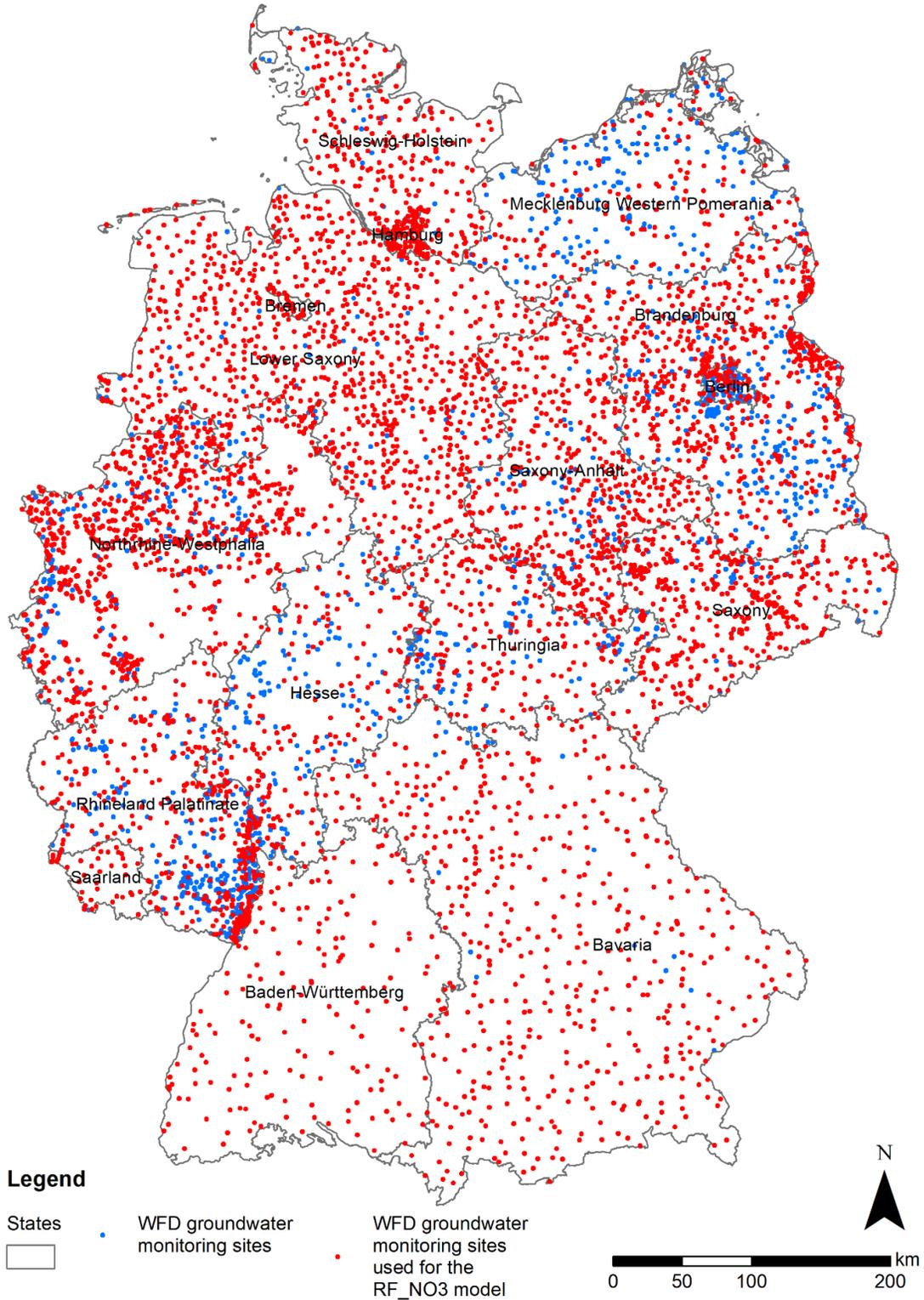


Figure S 3-2: Localisation of selected groundwater monitoring sites of the WFD data set (blue) and sites used for the RF_{NO_3} model (red) (coordinates of the monitoring sites were provided by the LAWA-AG (German Working Group on water issues of the Federal States and the Federal Government – Committee for Groundwater and Water supply); state borders: © GeoBasis-DE / BKG, 2017, Data license Germany - attribution - version 2.0 (www.govdata.de/dl-de/by-2-0)).

Table S 3-2: Spatial environmental predictor variables.

Predictor variable	Units	Variable name	Variable type	Model use	Data source
Land use					
Urban Land	%	urban_land	numerical	RF _{NO3}	LBM-DE2012 (BKG, 2016)
Arable Land	%	arable_land	numerical		
Grassland	%	grassland	numerical		
Forest	%	forest	numerical		
Special crops	%	special_crops	numerical		
Water	%	water	numerical		
Hydrospheric N surplus (mean 2007-2016)	kg ha ⁻¹ a ⁻¹	Nsurp_cm	numerical	RF _{NO3}	Häußermann et al. (2019), modified
Hydrogeology & Hydrology					
Hydrogeological units (32 classes)	(-)	hyd_unit	categorical	RF _{O2} , RF _{Fe} , RF _{NO3}	HYRAUM (BGR & SGD, 2015)
Rocktype of aquifer (3 classes)	(-)	L_GA	categorical	RF _{O2}	HÜK200 OGWL (BGR & SGD, 2016)
Consolidation of aquifer (3 classes)	(-)	L_VF	categorical	RF _{O2} , RF _{NO3}	
Type of aquifer (4 classes)	(-)	L_HA	categorical	RF _{O2} , RF _{NO3}	
Geochemical rock type of aquifer (7 classes)	(-)	L_GC	categorical	RF _{O2} , RF _{NO3}	
Hydraulic conductivity (11 classes)	(-)	L_KF	categorical	RF _{O2} , RF _{Fe} , RF _{NO3}	
Aquifer character (3 classes)	(-)	L_CH	categorical	RF _{O2} , RF _{NO3}	
Seepage water rate	mm/a	SWR	numerical	-	
Groundwater recharge rate	mm/a	GWR	numerical	RF _{O2} , RF _{Fe} , RF _{NO3}	HAD55_gwn1000_v1_raster (BGR, 2003b)
Nitrate concentration in seepage water	mg/l	SWR_conc	numerical	RF _{NO3}	Calculated from seepage water rate and N-surplus
Groundwater residence time	a	RTGW	numerical	RF _{O2} , RF _{Fe} , RF _{NO3}	Kunkel et al. (2007), cited after Fuchs et al. (2010)
Redox conditions (4 classes)	(-)	Redox	categorical	RF _{NO3}	Calculated from spatial predictions of groundwater O ₂ and Fe concentrations
Soil conditions					
Field capacity (0-1 m soil depth)	mm	FC	numerical	RF _{O2} , RF _{Fe} , RF _{NO3}	FK10dm1000_250 (BGR, 2015)
Humus content in top soil (0 -10 cm for grassland and forest; 0 - 30 cm for arable land)	%	HUMUS	numerical	RF _{O2} , RF _{Fe} , RF _{NO3}	HUMUS1000 (BGR, 2007)

Text S 3-2:

Land use

The land use map consists of 36 classes (BKG, 2016), which were aggregated to six dominant land use classes: forest (33%), arable land (32%), grassland (18%), urban land (8.5%), water (8%) and special crops (0.5%). The spatial distribution of the percentages of the two main types of land use, arable land and forest, is mapped in Figure S 3-3. Agricultural land is mainly located in the North German Plain and in Depressions and Lowlands in the central and southern parts of Germany. The mountain regions are predominantly forested.

The mean hydrospheric nitrogen (N) soil surface budget surplus, reduced by ammonia losses during application of organic fertilisers on agricultural land in Germany in the period 2007 to 2017 is about 55 kg N ha⁻¹y⁻¹ (district means: max. 119 kg ha⁻¹y⁻¹) (Häußermann et al., 2019). As a predictor for the groundwater nitrate estimation, we use an area-weighted land use specific calculation of the N surplus ($Nsurp_{cm}$) for each grid cell. In addition to the N surplus for arable land, forest is assigned 10 kg N ha⁻¹y⁻¹, urban land with 15 kg N ha⁻¹y⁻¹, grassland with the N surplus of arable land multiplied by 0.43 and special crops with N surplus of arable land multiplied by 1.5 (Eq. [S 4-1]). The highest N surpluses are recorded in western Lower Saxony and south-eastern Bavaria. The spatial distribution of the calculated mean N-surplus is mapped in Figure S 3-6.

$$Nsurp_{cm} = Nsurp_a w_a + Nsurp_u w_u + Nsurp_g w_g + Nsurp_f w_f + Nsurp_s w_s \quad [S 4-1]$$

$Nsurp_a$ = N-surplus arable land [kg N ha⁻¹ a⁻¹] (Häußermann et al., 2019)

$Nsurp_u$ = N-surplus urban land = 15 [kg N ha⁻¹ a⁻¹]

$Nsurp_g$ = N-surplus grassland = $Nsurp_a * 0.43$

$Nsurp_f$ = N-surplus forest = 10 [kg N ha⁻¹ a⁻¹]

$Nsurp_s$ = N-surplus special crops = $Nsurp_a * 1.5$

$w_{a,u,g,f,s}$ = area weighting factor [-]

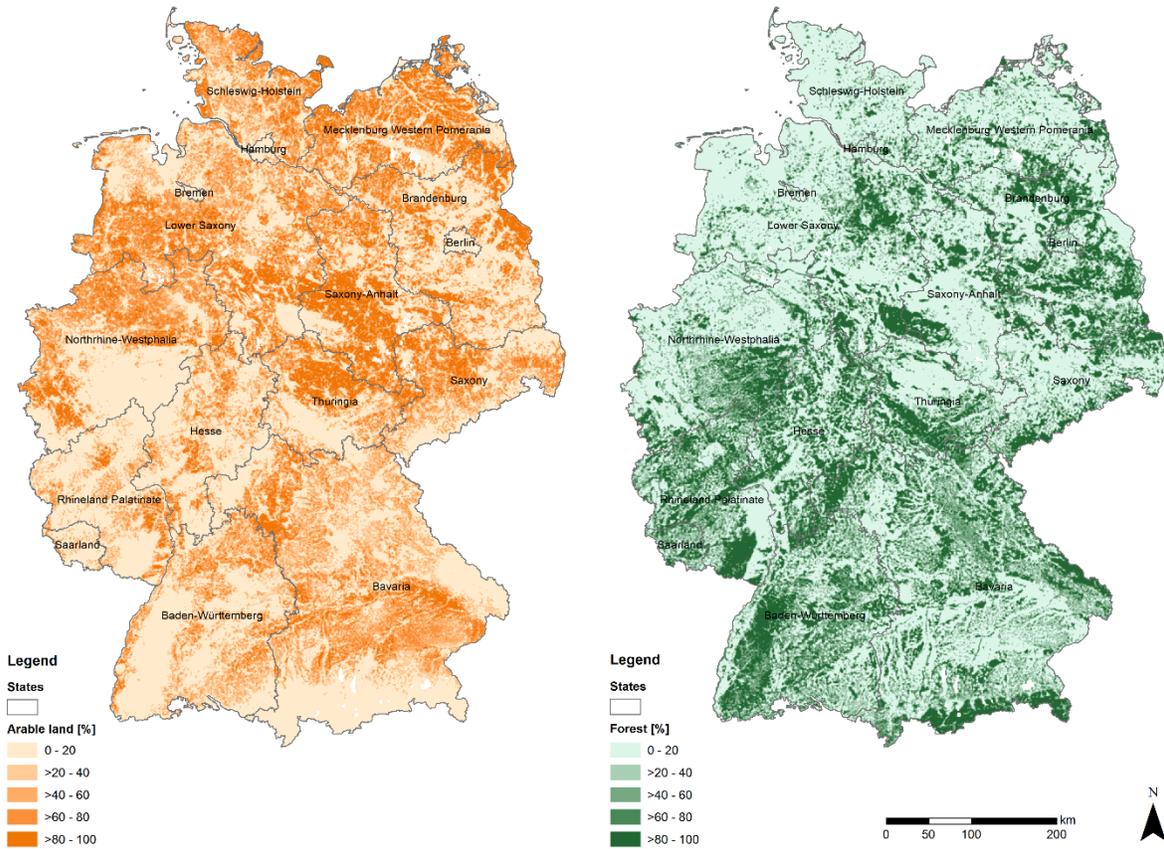


Figure S 3-3: Spatial distribution of percentages of main land use types arable land (left) and forest (right) based on BKG (2016); (state borders: © GeoBasis-DE / BKG, 2017, Data license Germany - attribution - version 2.0 (www.govdata.de/dl-de/by-2-0)).

Hydrogeology

In Germany, 36 different hydrogeological units are delineated (Figure S 3-4) (BGR & SGD, 2015). We used the hydrogeological units as categorical predictor variables. Due to model limitations, the number of categories must be reduced to 32. Units 11 and 12, 65 and 66, 91 and 94 and 96 and 97 were therefore aggregated. The 36 units and their designations are shown in Figure S 3-4. A detailed description of the aquifers of the hydrogeological units is given by the six categorical features rock type, consolidation, aquifer type, geochemical rock type, hydraulic conductivity, and the character of the aquifer derived from the hydrogeological map (BGR & SGD, 2016).

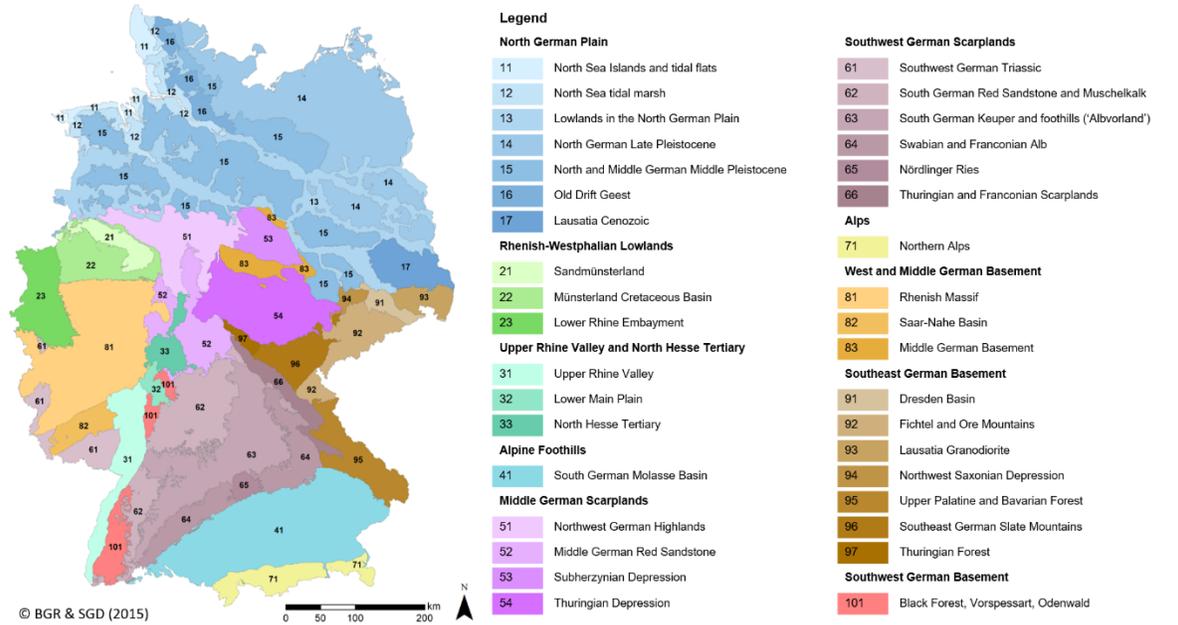


Figure S 3-4: Hydrogeological units for Germany (BGR & SGD, 2015).

The groundwater recharge rate is shown in Figure S 3-5 and reaches from very low values (<50 mm/a) in the north-eastern part of Germany up to high values (>400 mm/a) especially in the mountainous regions (BGR, 2003b). The groundwater residence times were made available to us on the basis of sub catchment areas and are shown in Figure S 3-5. The groundwater residence times reaches from < 10 years in the mountainous parts of central and southern Germany up to >200 years in the North German Plain and some lowlands (Fuchs et al., 2010; Kunkel et al., 2007).

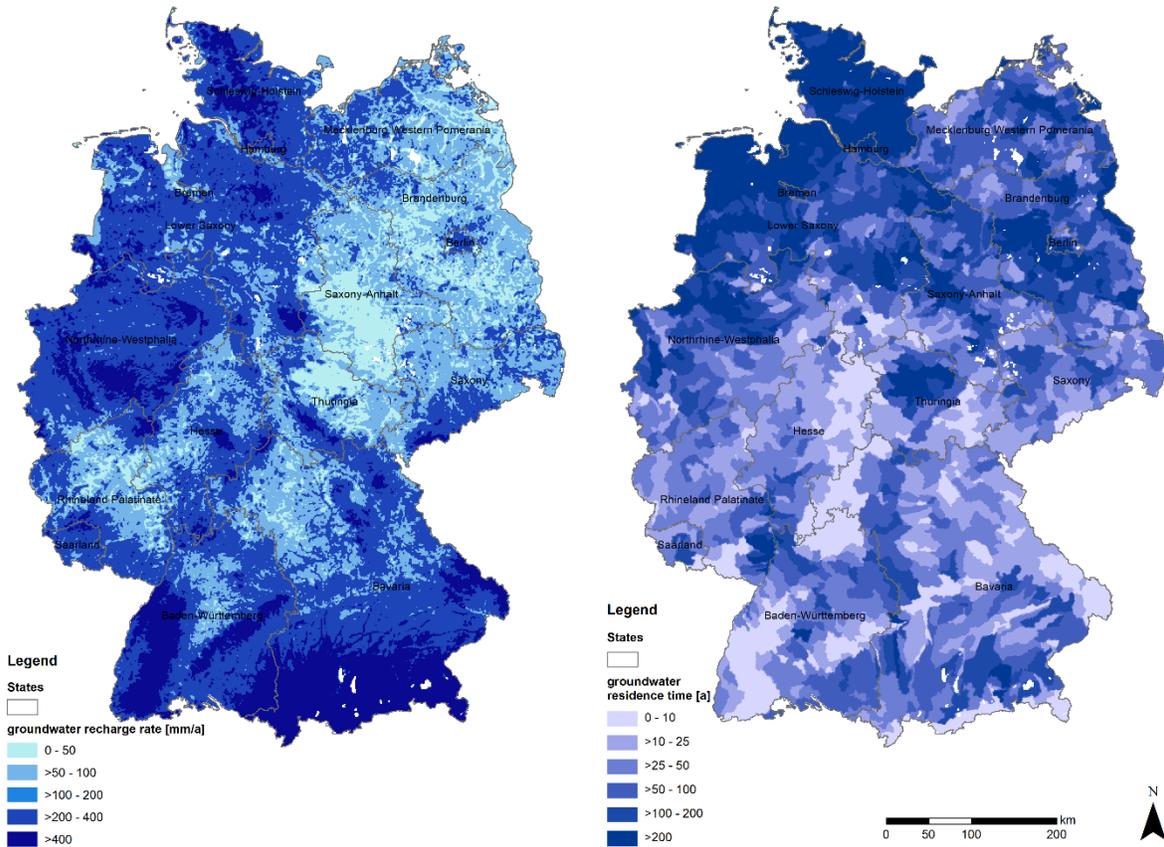


Figure S 3-5: Spatial distribution of groundwater recharge rate (BGR, 2003b) (left) and groundwater residence times (Fuchs et al., 2010; Kunkel et al., 2007) (right); (state borders: © GeoBasis-DE / BKG, 2017, Data license Germany - attribution - version 2.0 (www.govdata.de/dl-de/by-2-0)).

Potential nitrate concentration in seepage water

The amount of seepage water rate in Germany ranges from nearly zero mm/a in the floodplains in northeast Germany up to >1000 mm/a in high mountain ranges like the Black Forest and the Alps (BGR, 2003a). The higher values for the seepage water than for the groundwater result from losses through interflow, which is much higher in the fractured mountainous areas than in the lowlands.

Based on the calculated mean N surplus data we calculate the potential nitrate concentration (in mg/l) of the seepage water (SWR_conc):

$$SWR_{conc} = \frac{N_{surp_{cm}} * 100}{SWR} * 4.43 \text{ [mg/l]} \quad [S\ 4-2]$$

The highest potential nitrate concentration in seepage water occur in regions of north-eastern Germany with high N surplus (Figure S 3-3) in combination with moderate to low seepage water rates. The spatial distribution of the calculated nitrate concentration of the seepage water is shown in Figure S 3-6.

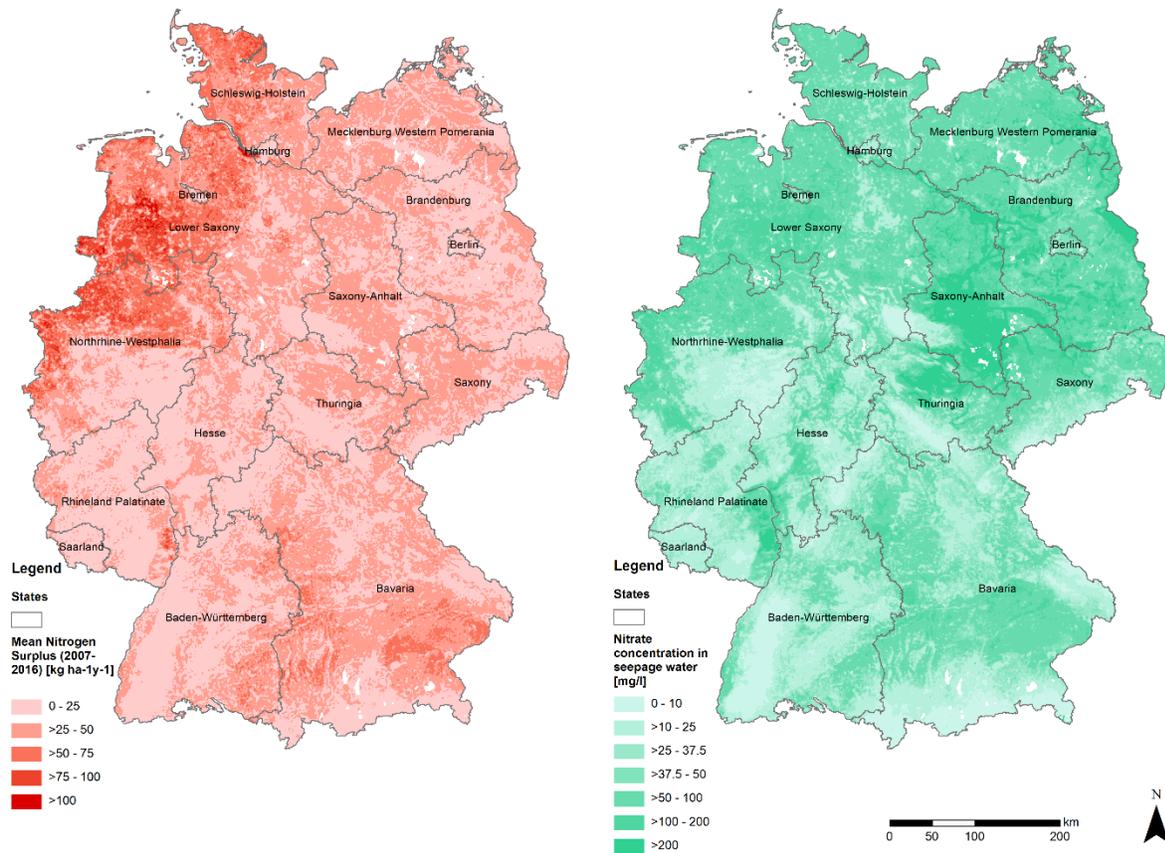


Figure S 3-6: Mean area-weighted and land use specific hydrospheric nitrogen surplus (2007 – 2016) based on Häußermann et al (2019) (left); calculated potential nitrate concentration in seepage water (right) based on Häußermann et al (2019) and (BGR, 2003a); (state borders: © GeoBasis-DE / BKG, 2017, Data license Germany - attribution - version 2.0 (www.govdata.de/dl-de/by-2-0)).

Soil conditions

We use information on field capacity of up to 1 m soil depth, with values of up to 500 mm in some alluvial zones especially in the North German Plain (BGR, 2015). In the mountainous areas lower values around 100 mm to 200 mm predominately occur. As a possible indicator for nitrate degradation in soils, we use the percentage of humus content in the topsoil (BGR, 2007).

Pre-processing

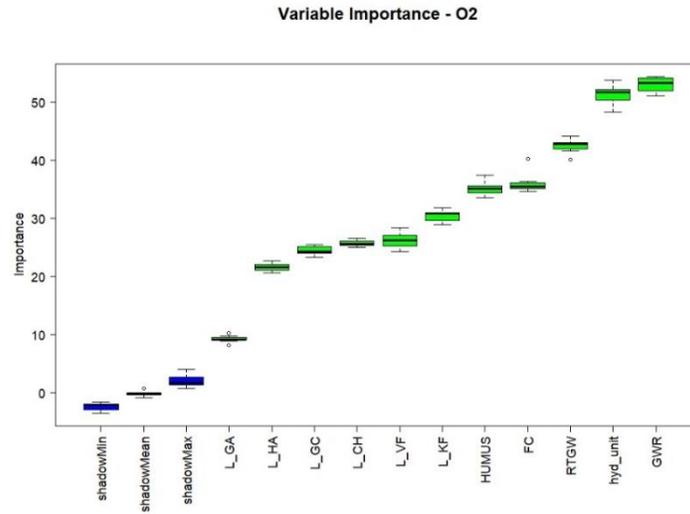
Text S 3-3:

The WFD data set is first transferred to a Geographical Information System ArcGIS (v. 10.4) into a point data set. The point data is linked to the spatial data (spatial predictors) described in Text S 3-2 by means of a buffer analysis. This is done using a circular buffer with a radius of 1000 m. With the tool 'spatial join' the means of numerical predictors are determined within the buffer. For categorical predictors the tool 'largest overlap' is used to assign the class with the dominant overlap. For the land use types, the percentages of the area within the buffer are calculated. These are also used as the area weighting factors to calculate the mean hydrospheric nitrogen (N) soil surface budget surplus. An analogous procedure of buffer analysis and spatial joining is used to link the spatial data with the grid map to which the predictions refer. We use the German Geo-Grid 1 km ('DE_Grid_ETRS89-UTM32_1km' © GeoBasis-DE / BKG, 2018) for the spatial prediction.

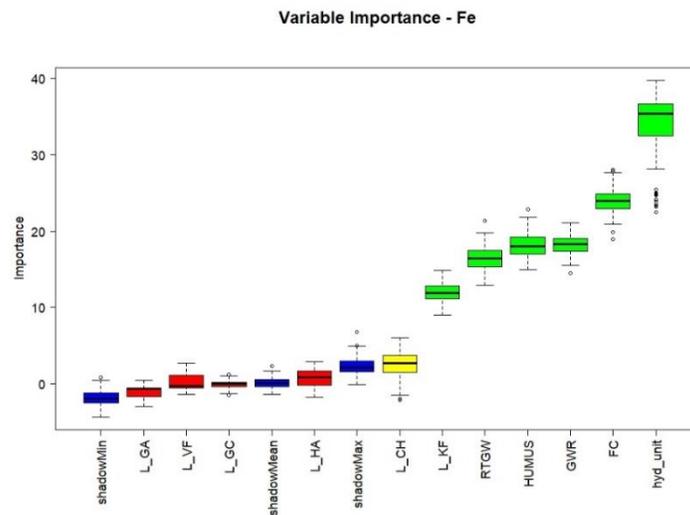
Table S 3-3: Classification point scheme for the characterisation of the redox conditions (acc. to LAWA (2018), modified).

		Oxygen (O ₂) thresholds		
		<2 [mg/l] (2 Points)	2-5 [mg/l] (1 Point)	>5 [mg/l] (0 Point)
Iron (Fe) thresholds	≥0.2 [mg/l] (1 Point)	(3) strong anaerobic	(2) anaerobic	(1) intermediate
	<0.2 [mg/l] (0 Point)	(2) anaerobic	(1) intermediate	(0) aerobic

a)



b)



c)

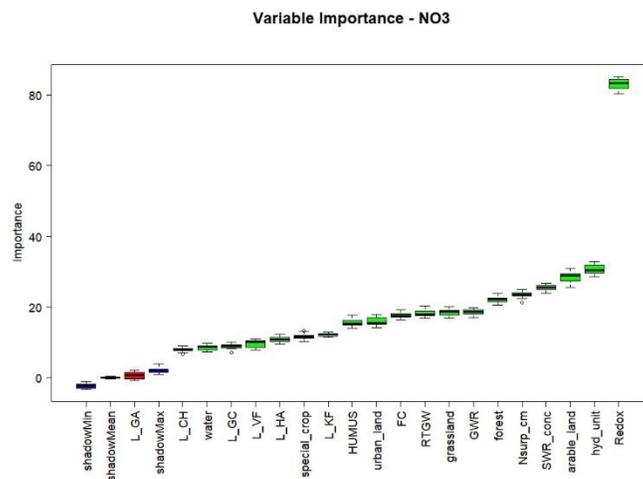


Figure S 3-7: Boruta variable importance plots for a) O₂, b) Fe and c) NO₃.

Predictive modelling

Text S 3-4:

Random Forest (RF)

The RF machine learning technique was developed by Breiman (2001) and is based on the classification and regression tree according to Breiman (1984). This nonparametric multivariate algorithm allows the modelling of nonlinear dependencies between both categorical and numerical parameters (De'ath and Fabricius, 2000). As an ensemble method, RF averages the results from many single decision trees and provides robust prediction with high performance (Kuhn and Johnson, 2013). In each decision tree, recursive partitioning is performed to split the data set of the response variable according to the predictors into homogeneous subsets described by the conditional mean (Hastie et al., 2009). Starting from the root node, the entire data set is divided into further non-terminal nodes until the tree has grown to a minimum number of observations per “leaf” or terminal node, respectively. At each node, the most important predictor for the further splitting of the data set is selected from the given predictors. RF contains two elements of randomness that increase variability among individual decision trees. First, the “bagging” technique is applied, where random bootstrap samples of the entire data set with replacement are created for each tree and then the predictions are aggregated by averaging over the full tree ensemble (Breiman, 1996; Hastie et al., 2009). In addition, only a random subset of predictor variables is used at each splitting node. Both steps lead to less accurate predictions of an individual tree and the correlation between the trees is reduced, which in turn leads to an increase in predictive performance for the ensemble average (Breiman, 2001). While model training is based on the bootstrap sample, the evaluation of model performance is carried out on the so-called out-of-bag sample, which is about 36% of the entire data set (Liaw and Wiener, 2002).

The RF model training requires two parameters: the number of trees in the ensemble (*ntree*) and the number of predictors that are randomly selected at each node (*mtry*). We set *ntree* to 1000 and tune the parameter *mtry* within a given range for each model with regard to the predictive performance (O₂ and Fe model: *mtry* = 2-4 and NO₃ model: *mtry* = 2-7). For O₂, Fe and NO₃ the model tuning gives best *mtry* values of 3, 2 and 7, respectively.

With the RF algorithm, it is also possible to estimate the importance of the predictors used in the ensemble by estimating the difference in the predictive performance of permuted and non-permuted predictors in the out-of-bag samples.

Results

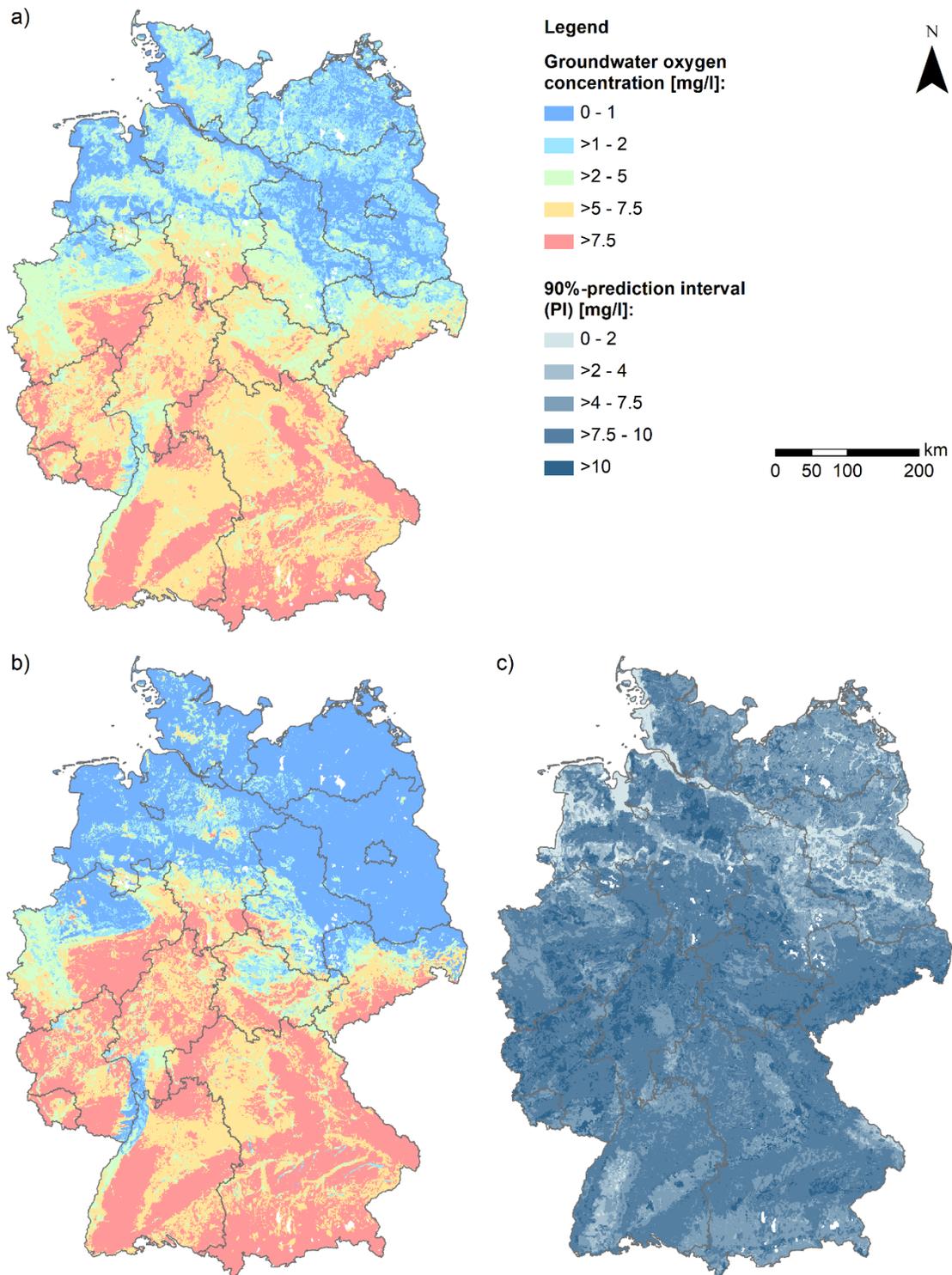


Figure S 3-8: Spatial prediction of a) the mean and b) the median groundwater oxygen (O₂) concentration in the period 2009-2018 based the WFD data set as well as c) the 90% prediction interval (PI); (state borders: © GeoBasis-DE / BKG, 2017, Data license Germany - attribution - version 2.0 (www.govdata.de/dl-de/by-2-0)).

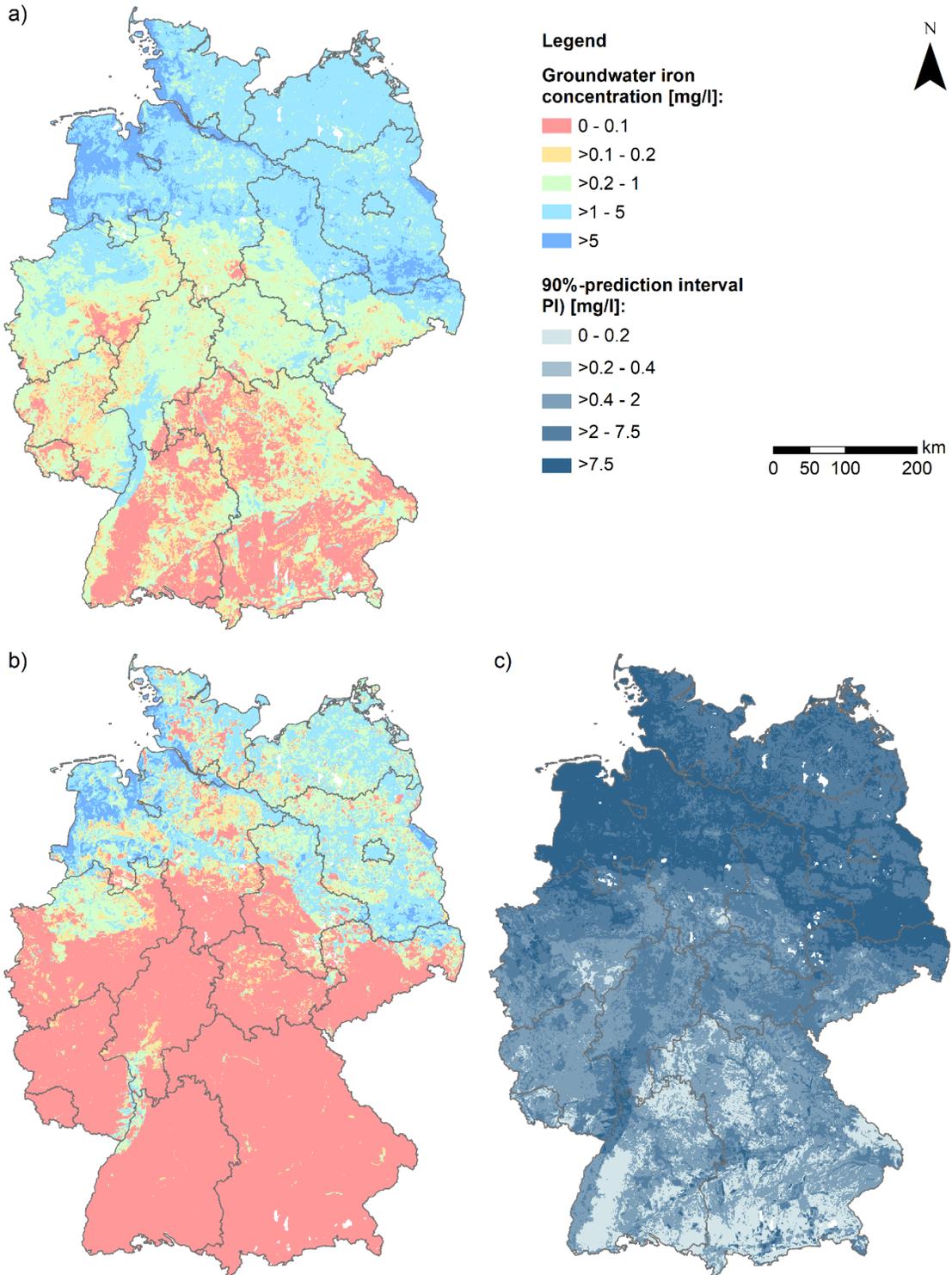


Figure S 3-9: Spatial prediction of a) the mean and b) the median groundwater iron (Fe) concentration in the period 2009-2018 based on the WFD data set as well as c) the 90% prediction interval (PI); (state borders: © GeoBasis-DE / BKG, 2017, Data license Germany - attribution - version 2.0 (www.govdata.de/dl-de/by-2-0)).

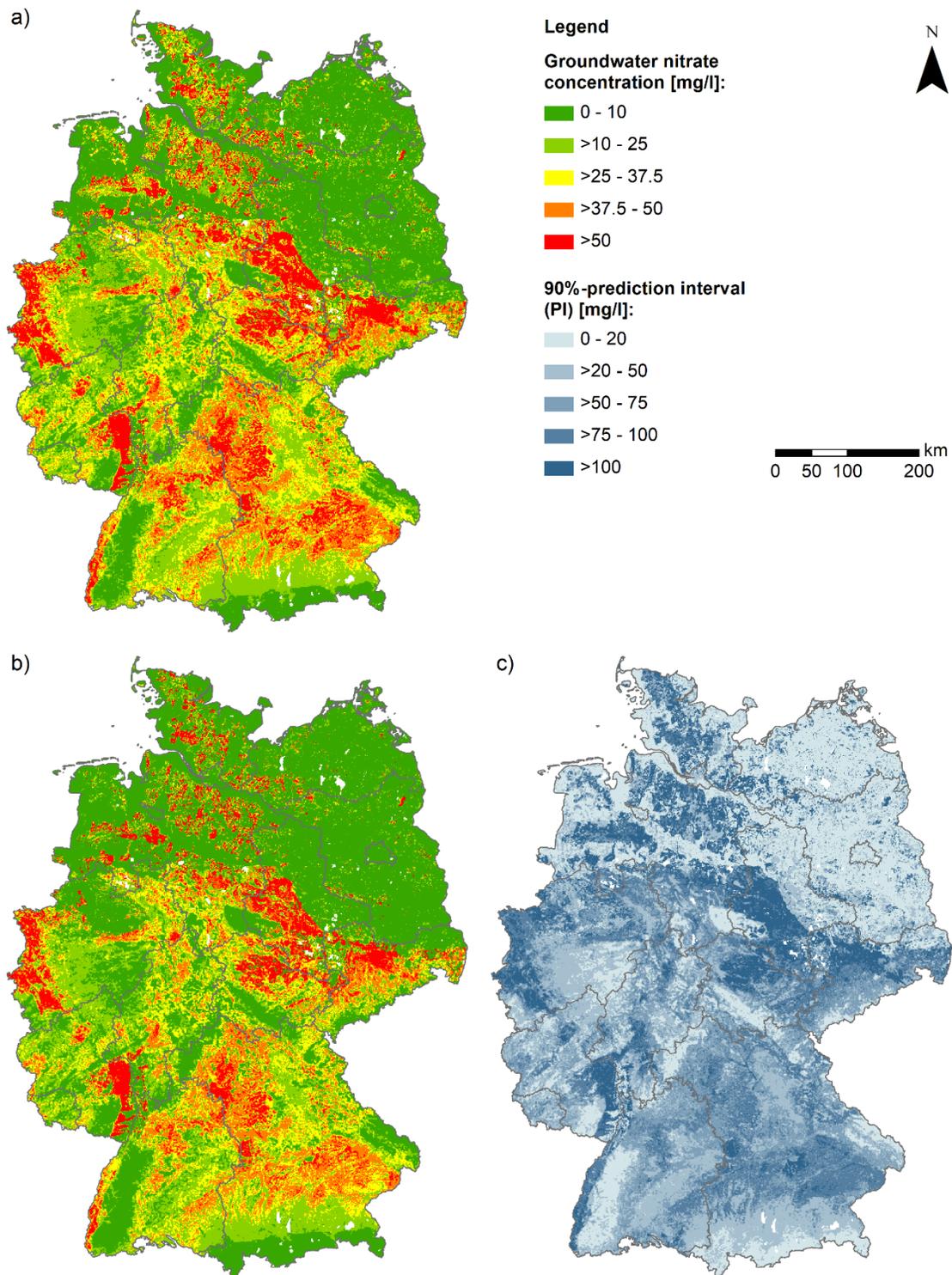


Figure S 3-10: Spatial prediction of a) the mean and b) the median groundwater nitrate (NO_3) concentration in the period 2009-2018 based the WFD data set as well as c) the 90% prediction interval (PI); (state borders: © GeoBasis-DE / BKG, 2017, Data license Germany - attribution - version 2.0 (www.govdata.de/dl-de/by-2-0)).

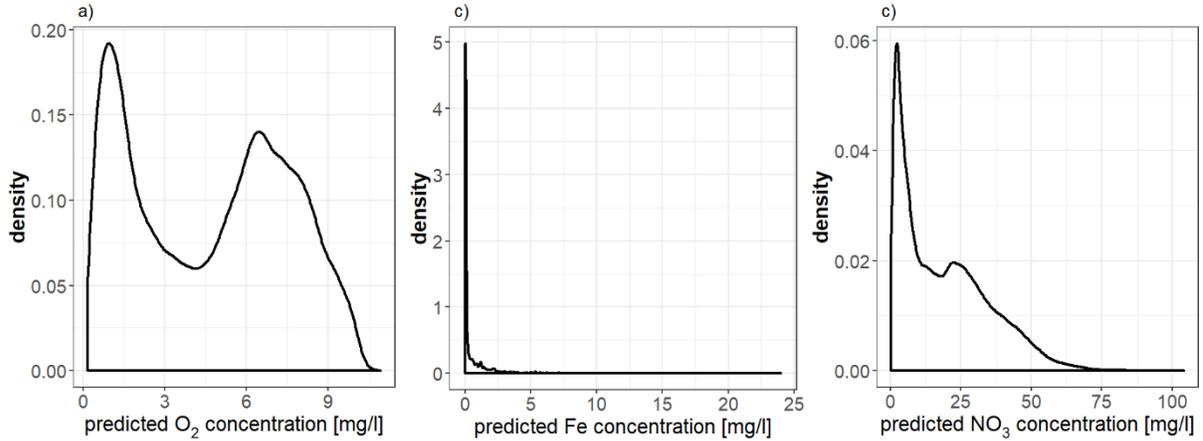


Figure S 3-11: Density plots of predicted groundwater concentrations a) mean O₂, b) median Fe and c) mean NO₃ based on the WFD data set.

Text S 3-5:

The redox condition is the most relevant spatial predictor for nitrate concentrations in groundwater. In order to analyse the effect of potential misclassifications of the redox conditions, a spatial sensitivity analysis is conducted. For this, we randomly vary for 5, 10, 15, 20, 25 and 30 % of the grid cells with regard to their redox class. A potential misclassification is represented in the sensitivity analysis by randomly allocating a grid cell to the next lower (-1) or the next higher (+1) class. Limitations are set by thresholds for the lower [0] and upper class [3]. To determine the effects of this variation on nitrate prediction, the mean predicted nitrate concentration for all grid cells without variation of the redox classes is plotted together with mean predicted nitrate concentration with varied redox classes (+/-1). Figure S 3-12 shows the results of this sensitivity analysis. The mean predicted NO₃ concentration varies more with increasing percentage of misclassification of the redox conditions. For instance, a random misclassification error of 20% of all grid cells results in a variation of the predicted overall mean by approximately 2.5 mg/l. Even a misclassification as large as 30% would not lead to a general questioning of our results with a predicted mean NO₃ concentration of 22.7 mg/l within a range of 21.0 to 24.9 mg/l.

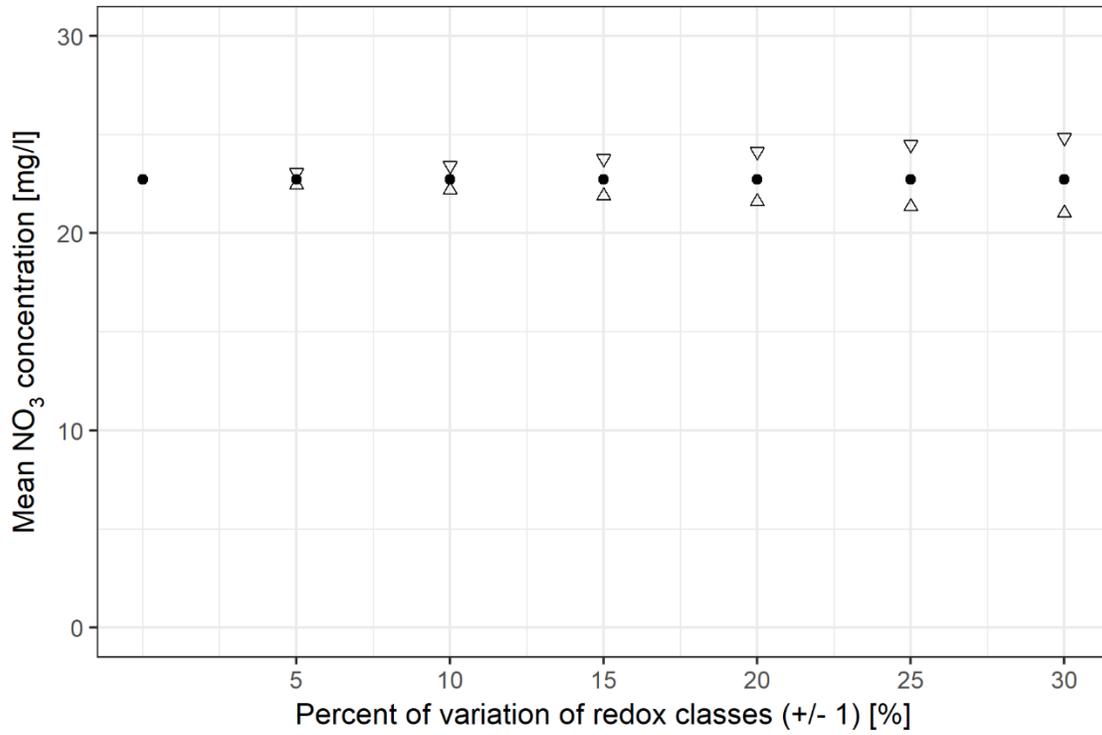


Figure S 3-12: Sensitivity analysis of the effect of potentially misclassified redox classes on simulated mean NO₃ concentration. (●) no variation of redox classes; (▽) redox classes (-1); (△) redox classes (+1).

4 Spatial distribution of integrated nitrate reduction across the unsaturated zone and the groundwater body in Germany

This chapter is published in the journal *Water* 12 (9), 2456, 2020.
<https://doi.org/10.3390/w12092456>

Knoll, L.^{1*}, Häußermann, U.¹, Breuer, L.^{1,2}, Bach, M.¹

[1] Institute for Landscape Ecology and Resources Management (ILR), Research Centre for BioSystems, Land Use and Nutrition (iFZ), Justus Liebig University Giessen, Heinrich-Buff-Ring 26, 35390 Giessen, Germany

[2] Centre for International Development and Environmental Research (ZEU), Justus Liebig University Giessen, Giessen, Germany

* Correspondence to: Lukas Knoll (Lukas.Knoll@umwelt.uni-giessen.de)

Abstract

Nitrate pollution in groundwater and mitigation strategies is currently a topic of controversial debate in Germany, and the demand for harmonised approaches for the implementation of regulations is increasing. Important factors that needs to be considered when planning mitigation measures are the nitrogen inputs into water bodies and the natural nitrate reduction capacity. The present study introduces a nationwide, harmonised and simplified approach for estimating nitrate reduction as an integral quantity across the unsaturated zone and the groundwater body. The nitrate reduction rates vary from 0% to 100%, and are on average 57%, with high values in the north of Germany and low values in the south. Hydrogeological characteristics are associated with the estimated nitrate reduction rates, whereby the influence of aquifer type and redox conditions are particularly relevant. The nitrate reduction rates are substantially higher in porous aquifers and under anaerobic conditions than in fractured, consolidated aquifers and under aerobic conditions. This contribution presents a harmonised conceptual approach to derive the nitrate reduction rate at a 1 x 1 km resolution. This information can be used when planning and designing mitigation measures to meet the groundwater nitrate limits.

4.1 Introduction

Human interference in the natural nitrogen (N) cycle has serious environmental consequences. Reactive nitrogen is a macronutrient essential for plant growth, but it also has a negative impact on human health, is involved in global warming and negatively affects terrestrial and aquatic ecosystems (Erisman et al., 2013). Therefore, there is an urgent need for sustainable integrated nitrogen management at all levels, from field site up to the global scale (Reis et al., 2016; Sutton et al., 2013). The pollution of water systems is seen as a major challenge by the European Commission. 25% of groundwater bodies in the EU have a poor chemical status, with 18% mainly affected by nitrate pollution (EEA, 2018). The EU Water Framework Directive 2000/60/EC (WFD) was implemented with the aim of maintaining or restoring the ‘good status’ of water bodies. Key to achieving this goal are effective strategies to reduce N inputs into groundwater and surface water. According to Fuchs et al. (Fuchs et al., 2010), 80% of the N load in surface waters in Germany originates from anthropogenic sources, with agriculture as the largest factor. The agricultural sector accounts for nearly 90% of the nitrate leaching to water systems (Bach et al., 2020a). Since the inputs largely result from diffuse N surpluses, initial measures are needed to reduce or optimise fertiliser management.

The pollution of groundwater bodies with nitrate depends considerably on the properties of the subsurface. These properties have to be taken into account in order to be able to decide where and to what extent mitigation measures are required to achieve good water status. Seitzinger et al. (2006) states that about 46% of nitrogen newly fixed by technical ammonia synthesis is removed by biogeochemical turnover processes in terrestrial soils on a global scale. This corresponds to a similar removal rate of 40% estimated for Europe (Van Egmond et al., 2002). In a study for the German federal state of North Rhine-Westphalia, an average denitrification rate of around 45% for the mobile N surplus in soils is deduced by Wendland et al. (2020) and Kuhr et al. (2013) have determined a range for N load reduction from 20% up to 80%. Moreover, groundwater bodies can have a substantial capacity for nitrate reduction, but this can vary widely in space and between 0% to 100% of the N loads can be removed in aquifers (McAleer et al., 2017; Seitzinger et al., 2006). For Denmark mean reduction rates in the groundwater of 63% were stated (Højberg et al., 2017). For the federal states of Mecklenburg-Vorpommern and a catchment in the north of Saxony-Anhalt (both in Germany), it was reported that 80% of the N input load was removed in the aquifers (Kunkel et al., 2017; Wriedt and Rode, 2006). A mean reduction rate of the N loads of 46% and maximum values above 80% in some marsh regions were modelled for the large river Weser catchment in Germany (Hirt et al., 2012). Nitrate reduction rates in the groundwater from 65%-83% were measured for different arable land systems in eastern China (Zhou et al., 2018). Højberg et al. (2017) reviewed studies on different scales (local to national) dealing with groundwater nitrate reduction in the Baltic Sea basin, concluding that the spatial variation can be significant in most countries. The studies

showed that information of the heterogeneous distribution of nitrate reduction in soils and aquifers based on harmonised approaches on high spatial resolution is needed to substantially improve the estimation of the fate of N loads in groundwater bodies on national scales.

Nitrate reduction in groundwater is primarily governed by denitrification processes. Knowles (1982) describes denitrification as a microbial reduction process of N oxides to molecular nitrogen as end product ($\text{NO}_3^- \rightarrow \text{NO}_2^- \rightarrow \text{NO} \rightarrow \text{N}_2\text{O} \rightarrow \text{N}_2$). Denitrification depends on the prevailing biogeochemical aquifer conditions and only occurs under anaerobic conditions when electron donors are available (Rivett et al., 2008; Seitzinger et al., 2006). Further, the groundwater residence time influences denitrification (Seitzinger et al., 2006). In general, a distinction is made between heterotrophic and autotrophic denitrification, depending on the type of denitrifying bacteria (Durand et al., 2011; Rivett et al., 2008). In heterotrophic denitrification, the bacteria use organic carbon as an electron donor, whereas in autotrophic denitrification they obtain the electrons from the oxidation of inorganic species, e.g. iron sulphide (pyrite, FeS_2) (Korom, 1992; Rivett et al., 2008). Both processes have been identified in several studies on nitrate degradation in groundwater (Kludt et al., 2016; Schwientek et al., 2008; Wisotzky et al., 2018; Zhang et al., 2012). Other studies have found that the nitrate concentration in groundwater strongly depends on the hydrogeological conditions and the prevailing redox conditions (Ransom et al., 2017; Knoll et al., 2020).

Due to the complexity of the environmental conditions that favour degradation processes and the difficulty of measuring these in the field, models are useful tools to estimate the magnitude of nitrate reduction at different scales (Boyer et al., 2006). Various studies have modelled and analysed denitrification on catchment scale (Durand et al., 2015; Green et al., 2016; Hirt et al., 2012; Oehler et al., 2009; Rivas et al., 2017; Whelan and Gandolfi, 2002). Wriedt and Rode (2006) successfully simulated nitrate transport and turnover processes in a small lowland catchment using a process based model which considers detailed geochemical reactions to describe denitrification. In studies on larger scales e.g. on a regional or national level, a simplified representation of nitrate reduction processes is needed due to limited data availability and the spatial resolution. Wendland et al. (2020) and Kunkel et al. (2017) coupled several models (GROWA-DENUZ-WEKU) to assess the denitrification in soils at a federal state level depending on soil type and residence time of percolation water according to a Michaelis-Menten kinetics, as well as the denitrification during groundwater transport assuming first order. Hirt et al. (2012) used the same approach and coupled their model chain with a nutrient emission model (MONERIS) for modelling the nitrogen loads into surface water on. In nationwide studies for Germany, simplified exponential functions depending on the seepage water rate and hydrogeological conditions were used to describe the reduction of N loads (Fuchs et al., 2017a; Venohr et al., 2011). These functions were derived by comparing the nitrate concentration in seepage water and groundwater

for different rock types (Venohr et al., 2011). The concentration data from 217 groundwater monitoring sites were taken into account and four hydrogeological rock types were distinguished. With regard to the large-scale problem and the heterogeneous landscapes in Germany, an improved approximation of the actual nitrate reduction rates could be achieved by a more comprehensive representation of the distribution of nitrate concentration in the groundwater and a stronger differentiation of the hydrogeological conditions.

Through the implementation of the EU Nitrates Directive and the EU Water Framework Directive, EU Member States have undertaken to report regularly on the implementation of the guidelines, which requires nationwide harmonised assessment of the water quality status. National assessments of groundwater quality and of nitrate reduction, can contribute to further studies, such as national assessments of nutrient inputs into surface waters (Fuchs et al., 2017a). Due to the limitations of modelling complex nitrate reduction processes on a large scale (national level), this study combines nationwide data sets with a simplified conceptual approach to derive nitrate reduction rates from nitrogen inputs to groundwater. The aims of the present study are: 1) to determine the nitrate-nitrogen input loads to the unsaturated zone and in the groundwater, 2) to quantify the reduction rates describing the integrated nitrate reduction across the unsaturated zone and the groundwater body, and 3) to analyse the relationships between the hydrogeological conditions and the resulting reduction rates.

4.2 Materials and Methods

This study focuses on a regionalisation for Germany, with a spatial resolution of 1 x 1 km grid size (BKG, 2018), in total approximately 360,000 grid cells. All data used in the study are handled in the geographical information system ArcGIS (v. 10.4). All analyses of the present study concerning possible relationships between the nitrate reduction and the hydrogeological parameters are made using R (v. 3.4.1).

Figure 4-1 shows the simulated flow paths of water and nitrogen loads in our model. Nitrogen (N) is supplied to the soil in different reactive forms of N: as mineral nitrogen (nitrate, ammonium) and organic nitrogen. Under the typical soil and climate conditions in Germany, only nitrate nitrogen ($\text{NO}_3\text{-N}$) in quantifiable quantities is displaced by the seepage water. The concentrations of ammonium and organic N in the soil solution are usually very low below the root zone. For simplification, it is therefore justified to assume $\text{NO}_3\text{-N}$ for the entire N load in the seepage water ($\text{NO}_3\text{-N}$ load). On the basis of the hydrospheric N surplus of different land uses (agricultural and urban areas as well as forests; $\text{N}_{\text{surplus}_{a,f,u}}$), the potential $\text{NO}_3\text{-N}$ input load into the unsaturated zone ($\text{NO}_3\text{-N}_{\text{load}_{\text{input}_{uz}}}$) is derived from a previous study (Häußermann et al., 2019). This is transported through the unsaturated zone with

seepage water (Q_{sw}), depending on the seepage water rate (BGR, 2003a). Due to interflow and tile drains ($Q_{i,d}$), some of the seepage water and thus also some of the NO_3-N load ($NO_3-N_{load_{i,d}}$) does not reach the groundwater. The remaining share represents the given groundwater recharge (Q_{gw}) (BGR, 2003b) and the estimated NO_3-N input load transported across the unsaturated zone to the groundwater by Q_{gw} ($NO_3-N_{load_{input_gw}}$). For the groundwater body, a groundwater nitrate concentration map ($c(NO_3)_{gw}$) is taken from a previous study (Knoll et al., 2020). Given the groundwater recharge Q_{gw} , the NO_3-N load in the groundwater ($NO_3-N_{load_{gw}}$) can be calculated. In order to describe the reduction processes in the unsaturated zone and in the groundwater body, the integrated nitrate reduction respectively NO_3-N load reduction rate (NO_3-N_{red}) is defined as the ratio of $NO_3-N_{load_{gw}}$ and $NO_3-N_{load_{input_uz}}$. The parameters and calculations are further described in the following sections.

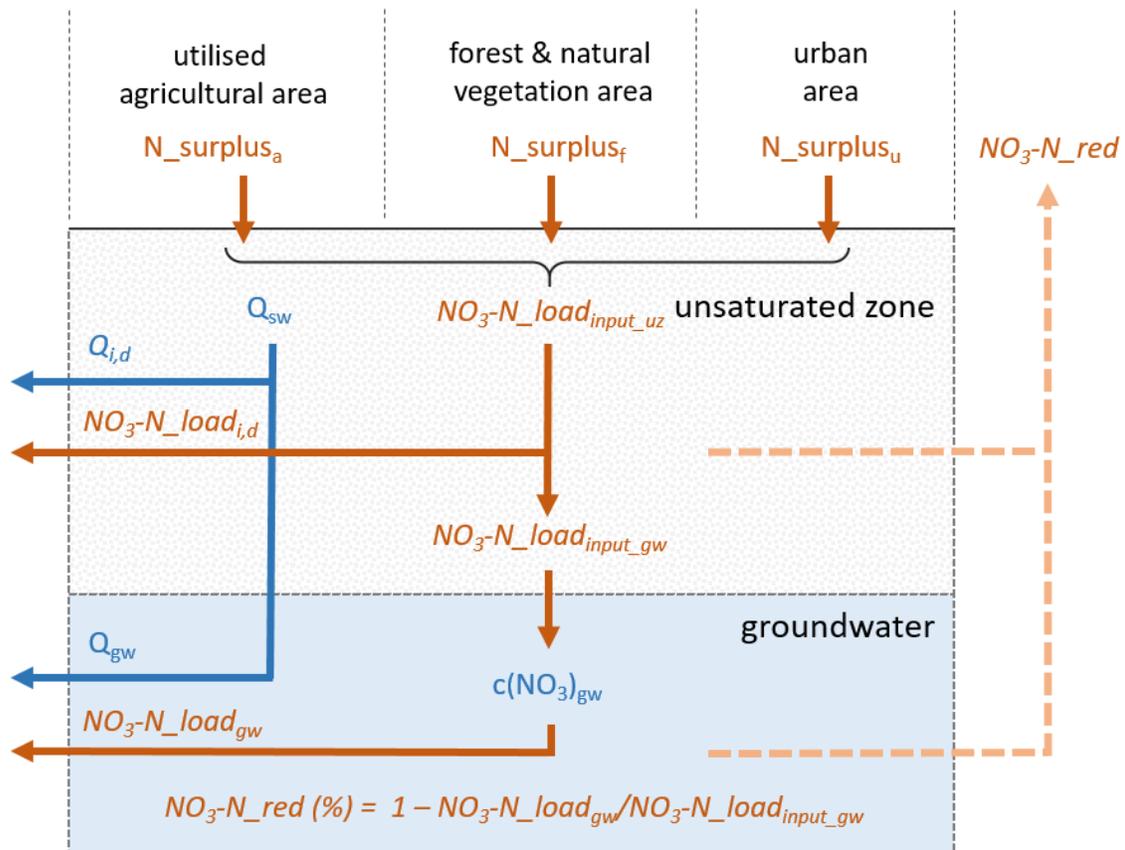


Figure 4-1: Flow paths and nitrate nitrogen loads (NO_3-N_{loads}) through the unsaturated zone and the groundwater body with: $N_{surplus_{a,f,u}}$ = hydrospheric N surplus for different land use types (a = agriculture [including arable land, grassland and special crops], f = forest, u = urban) [$kg\ N\ ha^{-1}\ a^{-1}$], Q_{sw} = seepage water [$mm\ a^{-1}$], Q_{gw} = groundwater recharge [$mm\ a^{-1}$], $Q_{i,d}$ = interflow, drainage [$mm\ a^{-1}$], $c(NO_3)_{gw}$ = NO_3 concentration in groundwater [$mg\ l^{-1}$], $NO_3-N_{load_{input_uz}}$ = NO_3-N input load unsaturated zone [$kg\ N\ ha^{-1}\ a^{-1}$], $NO_3-N_{load_{input_gw}}$ = NO_3-N input load groundwater [$kg\ N\ ha^{-1}\ a^{-1}$], $NO_3-N_{load_{gw}}$ = groundwater NO_3-N load [$kg\ N\ ha^{-1}\ a^{-1}$], NO_3-N_{red} = nitrate reduction rate [%], calculated values in italic, blue arrows represents water fluxes and orange arrows represents NO_3-N_{loads} .

4.2.1 Nitrogen input load

The hydrospheric N surplus is defined as the N surface budget surplus reduced by the gaseous N losses during and after application of organic and mineral fertilizer. The hydrospheric N surplus describes the amount of nitrogen potentially translocated to the hydrosphere; it is used as a key indicator characterising possible water pollution with nitrate from agricultural systems. The annual N budgets and the hydrospheric N surplus for the utilised agricultural area (UAA) were calculated on the basis of administrative units (district regions) (Häußermann et al., 2019). The average hydrospheric N surplus in Germany in the period 2007 - 2016 is calculated as 56 kg N ha⁻¹ a⁻¹ UAA, ranging from 13 to 146 kg N ha⁻¹ a⁻¹ UAA for the districts. As these values refer only to the inputs from agricultural land (arable land, grassland and special crops) constant values for the N surplus have been assumed for the remaining areas. Forest and natural areas have been assigned with 5 kg N ha⁻¹ a⁻¹, urban land with 18 kg N ha⁻¹ a⁻¹; for special crops on agricultural land the N surplus is multiplied by the factor 1.5. The spatial distribution of the five land use types is given by the land cover model (BKG, 2016) and their area proportions were assigned to the grid cells. An area-weighted and land-use specific hydrospheric N surplus is calculated according to Knoll et al. (2020) and is considered to be equivalent to the NO₃-N input load into the unsaturated zone ($NO_3-N_{load_{input_{uz}}}$):

$$NO_3-N_{load_{input_{uz}}} = N_{surplus_a} w_a + N_{surplus_u} w_u + N_{surplus_f} w_f + N_{surplus_s} w_s \quad [4-1]$$

$NO_3-N_{load_{input_{uz}}}$ = NO₃-N input load into the unsaturated zone [kg N ha⁻¹ a⁻¹]

$N_{surplus_a}$ = hydrospheric N-surplus arable land [kg N ha⁻¹ a⁻¹] (Häußermann et al., 2019)

$N_{surplus_u}$ = N-surplus urban land = 18 [kg N ha⁻¹ a⁻¹]

$N_{surplus_f}$ = N-surplus forest = 5 [kg N ha⁻¹ a⁻¹]

$N_{surplus_s}$ = N-surplus special crops = $N_{surplus_a} \times 1.5$ [kg N ha⁻¹ a⁻¹]

$w_{a,u,f,s}$ = area weighting factor [-] with $\sum w_i = 1$

Much of the NO₃-N load is transported with seepage water via interflow or drains ($Q_{i,d}$) into surface waters ($NO_3-N_{load_{i,d}}$), so that only the remaining load is transported across the unsaturated zone to the groundwater ($NO_3-N_{load_{input_{gw}}}$). For example, Kunkel et al. (2017) reported that for the federal state of Mecklenburg-Vorpommern around 35 % of the N load is transported via drainage systems into surface waters without any substantial nitrate reduction. The ratio of groundwater recharge Q_{gw} (BGR, 2003b) and seepage water rate Q_{sw} (BGR, 2003a) are used here as a proxy to estimate the interflow proportion and hence the NO₃-N input load to the groundwater via the unsaturated zone:

$$NO_3-N_{load_{input_gw}} = (Q_{gw}/Q_{sw}) \times NO_3-N_{load_{input_uz}} \quad [4-2]$$

$NO_3-N_{load_{input_gw}}$ = NO₃-N input load to the groundwater [kg N ha⁻¹ a⁻¹]

$NO_3-N_{load_{input_uz}}$ = NO₃-N input load into the unsaturated zone [kg N ha⁻¹ a⁻¹]

Q_{gw} = mean annual groundwater recharge rate [mm a⁻¹] (BGR, 2003b)

Q_{sw} = mean annual seepage water rate [mm a⁻¹] (BGR, 2003a)

Due to differences in the ways of estimation of seepage water (BGR, 2003a) and groundwater recharge rates (BGR, 2003b), it can happen that that lower seepage water rates than groundwater recharge rates occur, particularly in some lowland areas. In this case, the rates are considered equivalent ($Q_{gw}/Q_{sw} = 1$), since the groundwater recharge rate can only be as high as the seepage water rate.

4.2.2 Groundwater nitrate nitrogen load

In a study by Knoll et al. (2020), a 1 x 1 km grid map of groundwater nitrate concentrations in Germany was estimated using a data driven approach based on measured concentrations from the period 2009 to 2018. Applying the ‘random forest’ machine learning technique, the groundwater nitrate concentration map was modelled using several spatial predictors such as land use, hydrogeology and redox conditions throughout Germany. These predictions are applied in this study, since they provide a good assumption for nationwide estimates of groundwater NO₃-N loads. In order to calculate the groundwater loads, the groundwater recharge rate as an approximated value for the groundwater volume flow is used. Taking into account the conversion from nitrate to nitrate nitrogen per hectare, the groundwater NO₃-N load is calculated as follows:

$$NO_3-N_{load_{gw}} = (c_{gw} \times Q_{gw}) / 443 \quad [4-3]$$

$NO_3-N_{load_{gw}}$ = groundwater NO₃-N load [kg N ha⁻¹ a⁻¹]

c_{gw} = groundwater nitrate concentration [mg NO₃ l⁻¹] (Knoll et al., 2020)

Q_{gw} = mean annual groundwater recharge rate [mm a⁻¹] (BGR, 2003b)

443 = unit conversion factor from nitrate to nitrate nitrogen per hectare
(atomic weights: N=14, O=16)

4.2.3 Nitrate reduction

Considering the spatial resolution and the size of the study area, a straightforward approach is developed to estimate the integrated nitrate reduction across the unsaturated zone and the groundwater body. Such an approach is appropriate because data for the validation of reduction processes in the unsaturated zone are very limited. Since both, the initial value (NO₃-N input load) and the resulting value (groundwater

$\text{NO}_3\text{-N}$ load) are available from previous studies (Häußermann et al., 2019; Knoll et al., 2020), all possible natural degradation processes are taken into account when determining the nitrate reduction. The integrated nitrate reduction respectively the reduction of $\text{NO}_3\text{-N}$ loads is quantified as the ratio of groundwater $\text{NO}_3\text{-N}$ load to the $\text{NO}_3\text{-N}$ input load to the groundwater:

$$\text{NO}_3\text{-N}_{red} = (1 - (\text{NO}_3\text{-N}_{load_{gw}}/\text{NO}_3\text{-N}_{load_{input_{gw}}})) \times 100 \quad [4-4]$$

$\text{NO}_3\text{-N}_{red}$ = nitrate reduction [%]

$\text{NO}_3\text{-N}_{load_{gw}}$ = groundwater $\text{NO}_3\text{-N}$ load [$\text{kg N ha}^{-1} \text{ a}^{-1}$]

$\text{NO}_3\text{-N}_{load_{input_{gw}}}$ = $\text{NO}_3\text{-N}$ input load to the groundwater [$\text{kg N ha}^{-1} \text{ a}^{-1}$]

4.2.4 Hydrogeological conditions

In the hydrogeological shape map of Germany (HUEK200; (BGR & SGD, 2016)) the hydrogeology of the upper aquifer is described by the attributes ‘aquifer type’, ‘consolidation’, ‘rock type’, ‘geochemical rock type’ and ‘conductivity’. These five attributes are linked to the grid cells by the ArcGIS add-in tool ‘Spatial join – largest overlap’, which assigns the attributes with the dominant area that overlaps with the respective grid cell. As an additional hydrogeological parameter, the redox conditions are used as modelled by Knoll et al. (2020) with four redox classes, i.e. ‘strong anaerobic’, ‘anaerobic’, ‘intermediate’ and ‘aerobic’, based on groundwater oxygen and iron concentrations (specific limits are shown in Table 4-1).

4.3 Results

To calculate the $\text{NO}_3\text{-N}$ input loads to the groundwater, the ratio of Q_{gw} to Q_{sw} is estimated (Figure 4-2). The seepage water rate is generally higher than the groundwater recharge rate and the difference between both is greater in consolidated aquifers compared to unconsolidated aquifers. Therefore, the ratio Q_{gw}/Q_{sw} for consolidated aquifers is also considerably lower than for unconsolidated aquifers. This can probably be attributed to the morphological aspects of the mountain regions and the associated higher depth to the groundwater table thus causing a greater susceptibility for interflow. Agriculturally used areas are often tile drained, which also leads to lower Q_{gw}/Q_{sw} values; however, this is generally the case in less mountainous areas and lowlands. The average seepage water rate for Germany is about 287 mm a^{-1} , (BGR, 2003a) the average groundwater recharge rate is 125 mm a^{-1} (BGR, 2003b). Therefore, considering the assumptions for Eq. [4-2] more than 50% of the seepage water reaches the surface waters via interflow or drainage.

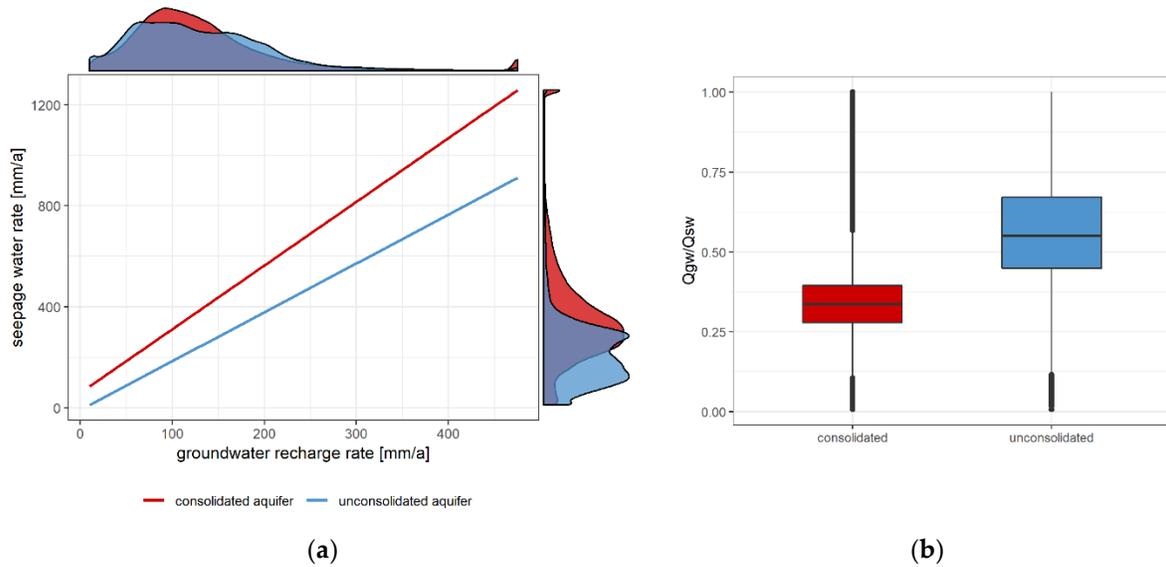


Figure 4-2: (a) Combined regression and density plot of seepage water (BGR, 2003a) and groundwater recharge rate (BGR, 2003b), (b) boxplot of the ratio of groundwater recharge and seepage water rate grouped by consolidated and unconsolidated aquifers.

The $\text{NO}_3\text{-N}$ input loads calculated according to Eq. [4-2] are shown in Figure 4-3a. The mean $\text{NO}_3\text{-N}$ input load into to unsaturated zone ($N_{load_{input_{uz}}}$) is around $36 \text{ kg N ha}^{-1} \text{ a}^{-1}$. Considering the proportional reduction by interflow and drainage, the mean $\text{NO}_3\text{-N}_{load_{input_{gw}}}$ is only about $17 \text{ kg N ha}^{-1} \text{ a}^{-1}$. The intensively managed agricultural systems are apparent in the north, north-west and south-east Germany, as well as in some preferential regions mainly in lowlands (Figure 4-3a). The highest loads in these regions are greater than $40 \text{ kg N ha}^{-1} \text{ a}^{-1}$.

In Figure 4-3b, the $\text{NO}_3\text{-N}$ loads in groundwater derived according to Eq. [4-3] are shown. It should be noted again, that these are $\text{NO}_3\text{-N}$ loads and not groundwater nitrate concentrations and therefore cannot be linked to the limit value of $50 \text{ mg NO}_3 \text{ l}^{-1}$. It is already evident that the $\text{NO}_3\text{-N}_{load_{gw}}$ (mean $6 \text{ kg N ha}^{-1} \text{ a}^{-1}$) is in most cases substantially lower than the $\text{NO}_3\text{-N}_{load_{input_{gw}}}$ (mean $17 \text{ kg N ha}^{-1} \text{ a}^{-1}$). The spatial distribution of high $\text{NO}_3\text{-N}_{load_{gw}}$ is not spatially correlated with high $\text{NO}_3\text{-N}_{load_{input_{gw}}}$.

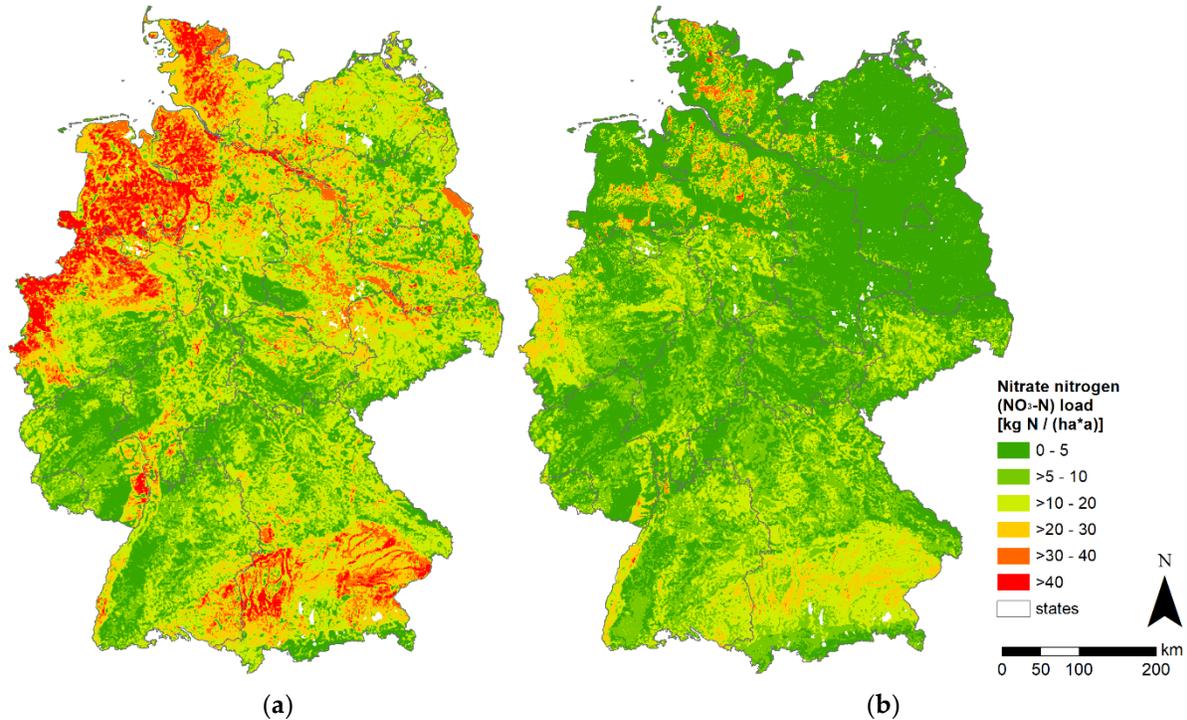


Figure 4-3: (a) Nitrate nitrogen input load to the groundwater ($NO_3-N_{load_{input_{gw}}}$), (b) Groundwater nitrate nitrogen load ($NO_3-N_{load_{gw}}$), based on a 1 x 1 km grid map of Germany (federal state borders: © GeoBasis-DE / BKG, 2017).

In summary for Germany, the annual NO_3-N input into the unsaturated zone ($NO_3-N_{load_{input_{uz}}}$) amounts to $1,196\ kt\ N\ a^{-1}$. The $NO_3-N_{load_{input_{gw}}}$ accounts for $580\ kt\ N\ a^{-1}$, of which $236\ kt\ N\ a^{-1}$ are discharged from the groundwater ($NO_3-N_{load_{gw}}$) into the surface waters. Thus the difference of $344\ kt\ N\ a^{-1}$ is due to NO_3-N load reduction (NO_3-N_{red}) in the unsaturated zone and the groundwater body. Figure 4-3 shows that the NO_3-N loads in the groundwater body are substantially reduced compared to the NO_3-N input loads. The reduction of NO_3-N loads, is quantified based on Eq. [4-4], where the percentage of removed NO_3-N load is calculated. The spatial distribution of the reduction rate [%] is shown in Figure 4-4. There is a marked north-south contrast with high reduction rates of up to 100% in the North German Plain and low to zero reduction in central and southern Germany. In some regions, predominantly in mountainous areas but also for example in the Geest regions in northern Germany, reduction rates are at or below zero. Excluding the areas with no NO_3-N reduction, the average reduction rate from NO_3-N input across the unsaturated zone and the groundwater body is 57%.

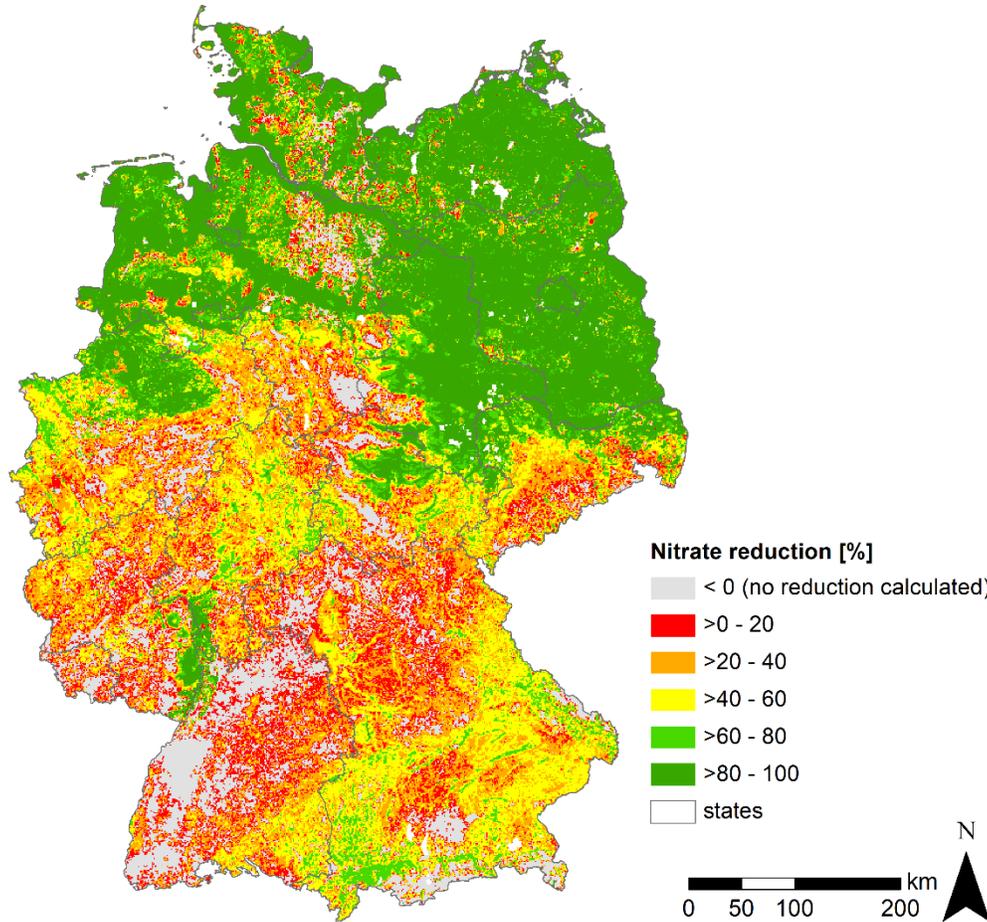


Figure 4-4: Map of mean nitrate reduction ($\text{NO}_3\text{-N}_{\text{red}}$) across the unsaturated zone and the groundwater based on a 1 x 1 km grid map of Germany (federal state borders: © GeoBasis-DE / BKG, 2017).

The nitrate reduction rates are further analysed regarding their dominant hydrogeological features. In Table 4-1, the mean values and the standard deviations for the $\text{NO}_3\text{-N}$ load reduction are summarised for the attributes 'aquifer types', 'consolidation', 'rock type', 'geochemical rock type', 'conductivity' and 'redox conditions'. The differences between the characteristics within the hydrogeological features were statistically analysed using a Post-Hoc-Test (TukeyHSD) (see Table S 4-1 to Table S 4-6). It is obvious, that porous unconsolidated aquifers tend to have significantly higher reduction rates (>70%) than consolidated fractured and karstified aquifers with below average reduction rates (32.8 – 36.1%) (Figure 4-5a). Since the porous aquifers predominantly have sedimentary origins, this difference in $\text{NO}_3\text{-N}$ load reduction is also apparent in the rock type, with significantly higher reduction rates being observed for sedimentary rocks than for metamorphic and magmatic rocks. Looking at the geochemical rock type, 'silicatic' rocks tend to have a significantly higher $\text{NO}_3\text{-N}$ load reduction potential; 'anthropogenic' and

‘sulphatic/halitic’ rocks show the highest and the lowest reductions rates, respectively. However, both classes are rare, occurring with a share of <0.5%. No obvious trend can be deduced for ‘conductivity’. The highest mean $\text{NO}_3\text{-N}$ load reduction rates of 77.6% are found for high conductivities, as well as for the overlapping class medium to moderate with 73.9%. However, medium conductivities only show an $\text{NO}_3\text{-N}$ load reduction rate below the average with 42.6%, and for moderate conductivities even lower. The lowest mean $\text{NO}_3\text{-N}$ load reduction values (28.5%) occur for the class ‘very low’, but both the class ‘very low’ and also ‘very high to high’ are rare, with a share of <0.2% of the overall area. With regard to redox conditions, there is a clear trend with increasing reduction rates from aerobic (32.9%) to strongly anaerobic aquifer conditions (95.8%) (Figure 4-5b).

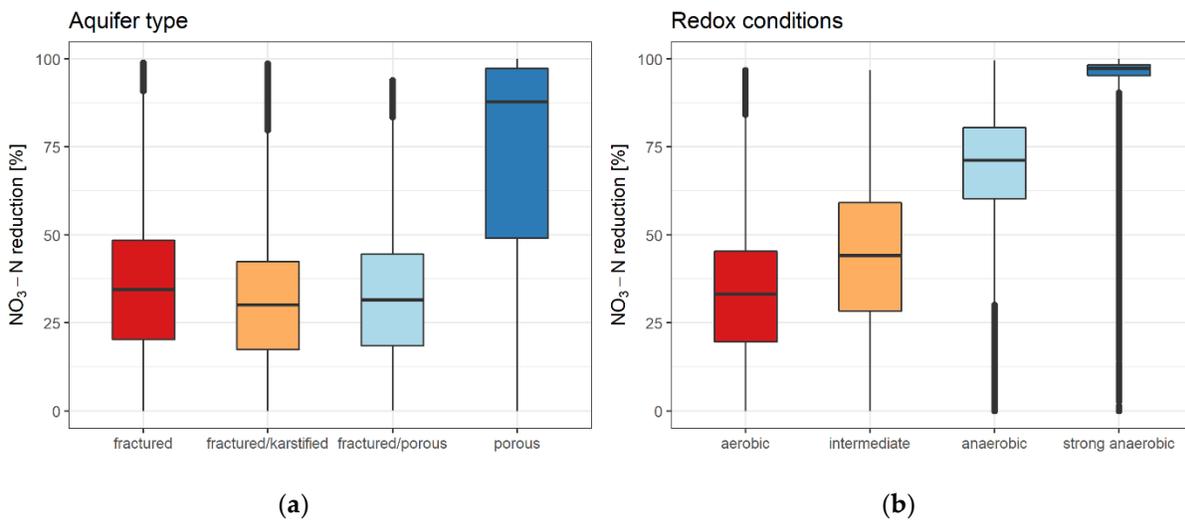


Figure 4-5: Boxplots of $\text{NO}_3\text{-N}$ load reduction rates [%] for (a) aquifer type and (b) redox conditions.

Table 4-1: Mean NO₃-N load reduction rates and standard deviation for different hydrogeological attributes.

Hydrogeology	NO ₃ -N load reduction [%]	
	mean	standard deviation
aquifer type¹		
fractured	36.1	21.3
fractured/karstified	32.9	21.7
fractured/porous	32.8	18.9
porous	72.5	28.4
consolidation¹		
consolidated	35.4	21.3
unconsolidated	72.6	28.4
rock type¹		
magmatic	37.6	16.9
metamorphic	36.1	17.8
sedimentary	58.9	31.9
geochemical rock type¹		
anthropogenic	67.1	24.0
sulphatic	37.9	23.5
sulphatic/halitic	20.0	11.2
carbonatic	35.1	20.2
silicatic/carbonatic	42.6	23.2
silicatic	64.6	32.1
silicatic/organic	55.3	25.1
conductivity¹		
very high (>10 ⁻² m s ⁻¹)	46.5	22.2
high (10 ⁻² – 10 ⁻³ m s ⁻¹)	77.6	27.1
medium (10 ⁻³ – 10 ⁻⁴ m s ⁻¹)	42.6	27.1
moderate (10 ⁻⁴ – 10 ⁻⁵ m s ⁻¹)	39.2	22.9
low (10 ⁻⁵ – 10 ⁻⁷ m s ⁻¹)	40.3	23.3
very low (10 ⁻⁷ – 10 ⁻⁹ m s ⁻¹)	28.5	14.7
<i>overlapping classes</i>		
very high to high (>10 ⁻³ m s ⁻¹)	27.4	14.8
medium to moderate (10 ⁻³ – 10 ⁻⁵ m s ⁻¹)	73.9	31.6
low to very low (<10 ⁻⁵ m s ⁻¹)	37.4	21.0
moderate to low (10 ⁻⁴ – 10 ⁻⁷ m s ⁻¹)	38.1	21.5
variable	56.0	27.6
redox conditions²		
aerobic (O ₂ >5 mg l ⁻¹ and Fe <0.2 mg l ⁻¹)	32.9	17.2
intermediate (O ₂ >5 mg l ⁻¹ and Fe ≥0.2 mg l ⁻¹ or O ₂ 2-5 mg l ⁻¹ and Fe <0.2 mg l ⁻¹)	44.8	22.7
anaerobic (O ₂ 2-5 mg l ⁻¹ and Fe ≥0.2 mg l ⁻¹ or O ₂ <2 mg l ⁻¹ and Fe <0.2 mg l ⁻¹)	69.1	16.6
strongly anaerobic (O ₂ <2 mg l ⁻¹ and Fe ≥0.2 mg l ⁻¹)	95.8	5.7

¹ attribute classes derived from HUEK200 (BGR & SGD, 2016); ² (Knoll et al., 2020).

4.4 Discussion

To the best of our knowledge, this study provides for the first time a conceptual, parsimonious modelling approach for estimating the spatial distribution at a high resolution (1 km) of the integrated nitrate reduction in the unsaturated zone and within the groundwater body on a large scale, applied here for Germany. Previous studies on N fluxes in Germany have shown that a significant fraction of the N loads in surface waters originate from the groundwater. However, the order of magnitude of the N load reduction rates during the groundwater transfer is a major unknown (Fuchs et al., 2017b). Although the approach presented here is simplified, the results provide a valuable approximation for estimating the potential losses of N loads via the groundwater path. It is not possible to estimate the uncertainties involved in the approach due to the assumptions made. For the hydrospheric N surplus on agriculturally utilised areas, uncertainties of $\pm 10 \text{ kg N ha}^{-1} \text{ a}^{-1}$ are stated (Häußermann et al., 2019). Large uncertainties are also found for groundwater nitrate concentrations, with prediction intervals of 53 mg l^{-1} and a mean absolute error of 12.7 mg l^{-1} (Knoll et al., 2020). No uncertainty analyses are available for the estimates of the N surplus on non-agricultural areas, the seepage water rate, or groundwater recharge rate. It can be argued, that both input variables $NO_3\text{-N}_{load_{input_gw}}$ and $NO_3\text{-N}_{load_{gw}}$ are technically well established and documented. They therefore provide the best available and state-of-the-art data sets for Germany. If it is accepted that the estimates of $NO_3\text{-N}_{load_{input_gw}}$ and $NO_3\text{-N}_{load_{gw}}$ are reliable, it could be concluded that the results of Eq. [4-4] are also reliable within the scope of the overall uncertainties of the approach, which however cannot be determined quantitatively. In order to examine the reliability of the results, the N_2/Ar -method could be applied. The N_2/Ar ratio is used to calculate the concentration of excess- N_2 in the groundwater, which can be used to derive the denitrification rate. Since corresponding nationwide measurement programs have not yet been established, a validation of the approach presented here can currently not be carried out with the N_2/Ar -method, but offers great potential for further research.

The conceptual approach for estimating the nitrate reduction along its transport paths from the source through the groundwater body includes many processes. Studies dealing with this topic (Hirt et al., 2012; Kunkel et al., 2017) usually distinguish between reduction in the unsaturated zone and the groundwater. However, for the unsaturated zone, a small decrease in nitrate concentration or a conservative transport can be assumed (Ascott et al., 2017; Rivett et al., 2007). The capability of validating the nitrate reduction in the unsaturated zone is limited to a very small number of measured nitrate concentrations in seepage water (Kreins et al., 2010). Since groundwater nitrate concentration data are available in a much higher density than measured values for nitrate concentration in the seepage water, Wendland et al. (2020) compared modelled nitrate concentrations in seepage water with observed nitrate concentrations in groundwater, and found good agreement. However, they note that in regions with increased denitrification capacity, measured groundwater concentrations are considerably lower than the simulated

seepage water concentrations. In the present study, it was done the other way round and the nitrate concentration in the groundwater was used as input data to derive the overall nitrate reduction. This approach provides a simple, uniform and robust method for the quantification of nitrate reduction, depending on existing nationwide data sets.

The nitrate reduction for Germany in the present study ranges from 0% to 100%, which was also the case in the study by Seitzinger et al. (2006). Comparing the overall mean integrated nitrate reduction of 57% calculated for Germany with other studies indicates similar dimensions, however, throughout Germany there is a very heterogeneous spatial distribution of nitrate reduction. For example, looking at the federal state of Mecklenburg-Vorpommern with its widespread glacially formed Late Pleistocene units, a mean integrated nitrate reduction of around 89% was calculated, while Kunkel et al. (2017) estimated around 40% of the N input being denitrified in the soil, and 85% of the N load transported via groundwater being removed within the aquifer. From this, one can infer an integrated reduction rate for the transport path from soil to the groundwater and within the groundwater body of around 74%, which is in a similar range to the reduction rates determined in the present study. For the Weser catchment, a mean $\text{NO}_3\text{-N}$ load reduction of about 46% was calculated, which is exactly the same as stated in Hirt et al. (2012). The Weser catchment covers typical hydrogeological characteristics occurring in Germany, from the Central German uplands to the North German Plain. According to Fuchs et al. (2017b), the mean N input between 2006 to 2011 into surface waters in Germany via the groundwater path amounted to 283 kt N, which is in line with the 236 kt N estimated in this study. At this point, it has to be considered that these loads can be further reduced by the nitrate reduction capacity of the riparian zone when entering the surface waters. Hill (2019) reviewed research studies from recent decades dealing with nitrate removal in the riparian zone and found considerable variations in reduction capacity depending on hydrogeological properties.

While Häußermann et al. (2019) report maximum values of the N surplus of up to $162 \text{ kg N ha}^{-1} \text{ a}^{-1}$ (mean $77 \text{ kg N ha}^{-1} \text{ a}^{-1}$), the values of the area-weighted land-use specific hydrospheric N input loads are considerably lower. This reflects the impact on the $\text{NO}_3\text{-N}$ input loads of areas with less anthropogenic control, such as natural or forested land. For some 12% of the area of Germany no reduction was calculated, which is due to the straightforward nature of our approach. When the N inputs are underestimated or the groundwater loads are overestimated, a reduction of less than zero results. A closer look at the regions with 'no reduction' shows that this mainly concerns forested areas or natural vegetation, e.g. the Alps or the Black Forest. In any case these areas are not relevant with respect to anthropogenic water pollution.

The spatial distribution of the nitrate reduction rates reflects very well the spatial hydrogeological structure in Germany. Rivas et al. (2017) concluded, that the factors 'soil texture', 'drainage class',

‘aquifer material’ and ‘rock types’ influence the groundwater denitrification potential. They found that well drained soils and rocks are unsuitable for denitrification processes. However, in contrast to Rivas et al. (2017), the conductivities in this study do not show a clear effect on nitrate reduction. The present study further shows that for Germany the parameters ‘aquifer type’, ‘consolidation’, and in particular ‘redox conditions’ significantly affect the intensity of the nitrate reduction.

Groundwater bodies with anaerobic conditions show a distinct influence on the nitrate reduction, which is to be expected since anaerobic conditions favour denitrification processes (Rivett et al., 2008). The unconsolidated porous aquifers occurring in the North German Plain are characterised by largely confined aquifers, low flow velocities, long groundwater residence times and, due to the thick covers of till layers, strongly anaerobic conditions (Knoll et al., 2020; Merz et al., 2009; Wendland et al., 2008). As a result, the North German Plain is dominated by high nitrate reduction rates. The Central German uplands are predominantly formed by fractured or karstified consolidated aquifer with mainly aerobic to intermediate redox conditions, that are unsuitable for denitrification processes and thus, lead to low reduction rates compared to the Quaternary deposits (Hannappel et al., 2018). The Alpine foreland in the south of Germany show a rather heterogeneous distribution of nitrate reduction. The southern part, mainly characterised by Late Pleistocene glacial deposits, has high reduction rates. Further north, Quaternary fluvial gravel deposits with high conductivities and, in contrast to the pore aquifers in northern Germany, very low to no nitrate reduction occur, which however is in line with the findings for well drained rocks (Rivas et al., 2017). The Tertiary sedimentary deposits in the northern Alpine foreland have in places thick loess covers, which leads to again higher nitrate reduction rates in this area.

The spatial distribution of the natural nitrate reduction have to be taken into account when discussing management strategies for groundwater bodies that do not meet the ‘good status’ according to the Water Framework Directive with respect to nitrate concentration. Mitigation measures are principally intended to reduce N surplus, mainly by restrictions on fertiliser use. In accordance with the German Groundwater Ordinance (GrwV, 2010), groundwater bodies are to be evaluated entirely according to the precautionary principle. In the context of the current revision of the Fertiliser Ordinance, the issue of internal differentiation of nitrate-sensitive areas require to make a 20% reduction of N fertilising is gaining more and more relevance. For planning and designing mitigation measures, information on the $\text{NO}_3\text{-N}$ input load, the groundwater nitrate concentrations or the groundwater $\text{NO}_3\text{-N}$ load, and the $\text{NO}_3\text{-N}$ load reduction quantified in the present study could be applied to determine the N fertilising reduction necessary to meet the groundwater nitrate limit ($50 \text{ mg NO}_3 \text{ l}^{-1}$).

Looking at the potential nitrate reduction rates, finally one must also consider that denitrification processes are not inexhaustible (Kludt et al., 2016; Wilde et al., 2017). Water management experts in

Germany regularly draw attention to the lack of knowledge about the denitrification capacity in the unsaturated zone and in groundwater in the regions of Germany (Bergmann et al., 2013), and whether similar rates of subsurface nitrate reduction can be expected in the longer term (UBA, 2017b).

4.5 Conclusions

In the present study, a parsimonious modelling approach has been developed to estimate the nationwide nitrate reduction or NO₃-N load reduction rates from existing data sets on N surplus calculations, water fluxes and estimates of groundwater nitrate concentrations. The method follows a simplified conceptual approach without the need to simulate complex biogeochemical processes. The nitrate reduction is considered as an integrated measure across the unsaturated zone and the groundwater body.

Besides a clear spatial differentiation with high nitrate reduction rates in north Germany and low nitrate reduction rates in the south, it could also be shown that the hydrogeological characteristics have a substantial influence on the degree of reduction. The nitrate reduction quantified in the present study could be taken into account in the planning and design of mitigation measures. Particularly in the current debate on the designation of vulnerable areas, the demand for harmonised approaches is growing.

Author Contributions

Conceptualization, Lukas Knoll and Martin Bach; Formal analysis, Lukas Knoll; Funding acquisition, Lutz Breuer and Martin Bach; Investigation, Lukas Knoll; Methodology, Lukas Knoll; Project administration, Martin Bach; Resources, Lukas Knoll, Uwe Häußermann and Martin Bach; Validation, Lukas Knoll; Visualization, Lukas Knoll; Writing – original draft, Lukas Knoll; Writing – review & editing, Uwe Häußermann, Lutz Breuer and Martin Bach.

Acknowledgements

This research was funded by the German Environment Agency in the framework of the environment research plan of the Federal Ministry for the Environment, Nature Conservation and Nuclear Safety (project no. 3715 22 2200). The hydrogeological data used are available from the Federal Institute for Geosciences and Natural Resources (BGR) in Hanover (https://www.bgr.bund.de/EN/Themen/Wasser/Produkte/produkte_node_en.html).

Supporting Information

Table S 4-1: Post-Hoc-Test results aquifer type.

Hydrogeology	NO ₃ -N load reduction [%]		Post-Hoc-Test (TukeyHSD [3])			
	mean	standard deviation	p-value			
aquifer type¹			fractured	fractured/ karstified	fractured/ porous	porous
fractured	36.1	21.3				
fractured/karstified	32.9	21.7	0*			
fractured/porous	32.8	18.9	0*	0.9865067		
porous	72.5	28.4	0*	0*	0*	

*significant difference (p-value <0.05), ¹ (BGR & SGD, 2016).

Table S 4-2: Post-Hoc-Test results consolidation.

Hydrogeology	NO ₃ -N load reduction [%]		Post-Hoc-Test (TukeyHSD [3])	
	mean	standard deviation	p-value	
consolidation¹			consolidated	unconsolidated
consolidated	35.4	21.3		
unconsolidated	72.6	28.4	0*	

*significant difference (p-value <0.05), ¹ (BGR & SGD, 2016).

Table S 4-3: Post-Hoc-Test results rock type.

Hydrogeology	NO ₃ -N load reduction [%]		Post-Hoc-Test (TukeyHSD [3])		
	mean	standard deviation	p-value		
rock type¹			magmatic	metamorphic	sedimentary
magmatic	37.6	16.9			
metamorphic	36.1	17.8	0.0007181*		
sedimentary	58.9	31.9	0*	0*	

*significant difference (p-value <0.05), ¹ (BGR & SGD, 2016).

Table S 4-4: Post-Hoc-Test results geochemical rock type.

Hydrogeology	NO ₃ -N load reduction [%]		Post-Hoc-Test (TukeyHSD [3])						
	mean	standard deviation	p-value						
geochemical rock type¹			a	su	s/h	c	s/c	s	s/o
anthropogenic (a)	67.1	24.0							
sulphatic (su)	37.9	23.5	0*						
sulphatic/halitic (s/h)	20.0	11.2	0*	0.0000440*					
carbonatic (c)	35.1	20.2	0*	0*	0.0011193*				
silicatic/carbonatic (s/c)	42.6	23.2	0*	0*	0*	0*			
silicatic (s)	64.6	32.1	0.0193857*	0*	0*	0*	0*		
silicatic/organic (s/o)	55.3	25.1	0*	0*	0*	0*	0*	0*	

*significant difference (p-value <0.05), ¹ (BGR & SGD, 2016).

Table S 4-5: Post-Hoc-Test results conductivity.

Hydrogeology	NO ₃ -N load reduction [%]		Post-Hoc-Test (TukeyHSD [3])										
	mean	standard deviation	p-value										
conductivity¹			1	2	3	4	5	6	7	8	9	10	11
very high (1) (>10 ⁻² m s ⁻¹)	46.5	22.2											
high (2) (10 ⁻² – 10 ⁻³ m s ⁻¹)	77.6	27.1	0*										
medium (3) (10 ⁻³ – 10 ⁻⁴ m s ⁻¹)	42.6	27.1	0*	0*									
moderate (4) (10 ⁻⁴ – 10 ⁻⁵ m s ⁻¹)	39.2	22.9	0*	0*	0*								
low (5) (10 ⁻⁵ – 10 ⁻⁷ m s ⁻¹)	40.3	23.3	0*	0*	0*	0.0014464*							
very low (6) (10 ⁻⁷ – 10 ⁻⁹ m s ⁻¹)	28.5	14.7	0*	0*	0*	0*	0.0111273						
very high to high (7) (>10 ⁻³ m s ⁻¹)	27.4	14.8	0.0000032*	0*	0.0007664*	0.322525		0.9999999					
medium to moderate (8) (10 ⁻³ – 10 ⁻⁵ m s ⁻¹)	73.9	31.6	0*	0*	0*	0*	0*	0*	0*				
low to very low (9) (<10 ⁻⁵ m s ⁻¹)	37.4	21.0	0*	0*	0*	0*	0*	0*	0*	0.1449279			
moderate to low (10) (10 ⁻⁴ – 10 ⁻⁷ m s ⁻¹)	38.1	21.5	0*	0*	0*	0*	0*	0*	0*	0*	0*	0*	
variable (11)	56.0	27.6	0*	0*	0*	0.0000008*	0*	0*	0*	0.0803958	0*	0.0021666*	0*

*significant difference (p-value <0.05), ¹ (BGR & SGD, 2016).

Table S 4-6: Post-Hoc-Test results redox conditions.

Hydrogeology	NO ₃ -N load reduction [%]		Post-Hoc-Test (TukeyHSD (“TukeyHSD function R Documentation,” n.d.))			
	mean	standard deviation	p-value			
redox conditions²			aerobic	intermediate	anaerobic	strongly anaerobic
aerobic	32.9	17.2				
intermediate	44.8	22.7	0*			
anaerobic	69.1	16.6	0*	0*		
strongly anaerobic	95.8	5.7	0*	0*	0*	

* significant difference (p-value <0.05), ² (Knoll et al., 2020).

5 References

- Aller, L., Bennett, T., Lehr, J., Petty, R., 1985. DRASTIC: a standardized system for evaluating ground water pollution potential using hydrogeologic settings. U.S. Environmental Protection Agency, Washington, D.C., EPA/600/2-85/018.
- Ascott, M.J., Goody, D.C., Wang, L., Stuart, M.E., Lewis, M.A., Ward, R.S., Binley, A.M., 2017. Global patterns of nitrate storage in the vadose zone. *Nature Communications* 8, 1416. <https://doi.org/10.1038/s41467-017-01321-w>
- Ascott, M.J., Wang, L., Stuart, M.E., Ward, R.S., Hart, A., 2016. Quantification of nitrate storage in the vadose (unsaturated) zone: a missing component of terrestrial N budgets. *Hydrological Processes* 30, 1903–1915. <https://doi.org/10.1002/hyp.10748>
- Ayres, R.U., Schlesinger, W.H., Socolow, R.H., 1994. Human Impacts on the Carbon and Nitrogen Cycles, in: *Industrial Ecology and Global Change*. Cambridge University Press., Cambridge, pp. 121–156. <https://doi.org/10.1017/CBO9780511564550.011>
- Babiker, I.S., Mohamed, M.A.A., Hiyama, T., Kato, K., 2005. A GIS-based DRASTIC model for assessing aquifer vulnerability in Kakamigahara Heights, Gifu Prefecture, central Japan. *Science of The Total Environment* 345, 127–140. <https://doi.org/10.1016/j.scitotenv.2004.11.005>
- Bach, M., Häußermann, U., Klement, L., **Knoll, L.**, Breuer, L., Weber, T., Fuchs, S., Heldstab, J., Reutimann, J., Schäppi, B., 2020a. Reactive nitrogen flows in Germany 2010-2014 (DESTINO Report 2), UBA-Texte 65/2020. Umweltbundesamt, Dessau-Roßlau.
- Bach, M., Klement, L., Häußermann, U., 2016. Bewertung von Maßnahmen zur Verminderung von Nitrateinträgen in die Gewässer auf Basis regionalisierter Stickstoff-Überschüsse. Teil I: Beitrag zur Entwicklung einer ressortübergreifenden Stickstoff-Strategie, UBA-Texte 55/2016. German Environment Agency (UBA), Dessau-Roßlau.
- Bach, M., **Knoll, L.**, Häußermann, U., Breuer, L., 2020b. Nitratbelastung des Grundwassers in Deutschland - Ist das Messnetz schuld? *Wasserwirtschaft* 2020, 54.
- Bárdossy, A., Giese, H., Grimm-Strele, J., Barufke, K.-P., 2003. SIMIK+ – GIS-implementierte Interpolation von Grundwasserparametern mit Hilfe von Landnutzungs- und Geologiedaten. *Hydrologie und Wasserbewirtschaftung* 47, 13–20.
- Behrendt, H., Huber, P., Kornmilch, M., Opitz, D., Schmoll, O., Scholz, G., Uebe, R., 2000. Nutrient Emissions into River Basins of Germany, UBA-Texte 23/2000. German Environment Agency (UBA), Dessau-Roßlau.

- Behrendt, H., Opitz, D., 1999. Retention of nutrients in river systems: dependence on specific runoff and hydraulic load. *Hydrobiologia* 410, 111–122. <https://doi.org/10.1023/A:1003735225869>
- Bell, Ja., 2014. *Machine Learning*, 1st ed. John Wiley & Sons, Ltd. <https://doi.org/10.1002/9781119183464>
- Bergmann, A., Dietrich, P., van Straaten, L., Frabko, U., van Berk, W., Kiefer, J., 2013. Konsequenzen nachlassenden Nitratabbauvermögens in Grundwasserleitern (Abschlussbericht). Deutscher Verein Gas- u. Wasserfach (DVGW), Bonn.
- BGR, 2015. Field capacity of German soils to a depth of 1 m (FK10dm1000_250). Federal Institute for Geosciences and Natural Resources (BGR), Hannover. Soil Information System FISBo: <https://produktcenter.bgr.de/terraCatalog/OpenSearch.do?search=5a75335e-ba46-452f-8792-ac7e9b49da88&type=/Query/OpenSearch.do> (accessed Nov 2, 2018).
- BGR, 2007. Organic Matter Content of Top-Soils in Germany 1:1,000,000 (HUMUS1000OB). Federal Institute for Geosciences and Natural Resources (BGR), Hannover. Soil Information System FISBo: https://www.bgr.bund.de/DE/Themen/Boden/Projekte/Stoffeigenschaften-abgeschlossen/Flaechenrepraesentative_Auswertungen/Corg.html (accessed Nov 2, 2018).
- BGR, 2005. Soil Scapes in Germany 1:5,000,000 (BGL5000). Federal Institute for Geosciences and Natural Resources (BGR), Hannover. Soil Information System FISBo: https://www.bgr.bund.de/DE/Themen/Boden/Informationsgrundlagen/Bodenkundliche_Karten_Datenbanken/Themenkarten/BGL5000/bg15000_node.html (accessed Nov 2, 2018).
- BGR, 2003a. Mean Annual Rate of Percolation from the Soil in Germany (SWR1000), Hydrogeologischer Atlas von Deutschland. Federal Institute for Geosciences and Natural Resources (BGR), Hannover. Soil Information System FISBo: https://www.bgr.bund.de/DE/Themen/Boden/Bilder/Bod_Themenkarten_HAD_4-5_g.html (accessed Nov 2, 2018).
- BGR, 2003b. Mean Annual Groundwater Recharge of Germany at the scale of 1:1,000,000 (HAD 5.5), Hydrogeologischer Atlas von Deutschland. Federal Institute for Geosciences and Natural Resources (BGR), Hannover. <https://produktcenter.bgr.de/terraCatalog/OpenSearch.do?search=5a75335e-ba46-452f-8792-ac7e9b49da88&type=/Query/OpenSearch.do> (accessed Nov 2, 2018).
- BGR & SGD, 2016. Hydrogeologische Übersichtskarte von Deutschland 1:200.000, Oberer Grundwasserleiter (HÜK200 OGWL). Digitaler Datenbestand, Version 3.0. Hannover.

- BGR & SGD, 2015. Hydrogeological spatial structure of Germany (HYRAUM). Federal Institute for Geosciences and Natural Resources (BGR), Hannover. https://www.bgr.bund.de/DE/Themen/Wasser/Projekte/abgeschlossen/Beratung/Hyraum/hyraum_projektbeschr.html (accessed Nov 2, 2018).
- BKG, 2018. Geographische Gitter für Deutschland in UTM-Projektion (GeoGitter national) - DE_Grid_ETRS89-UTM32_1km, Frankfurt a.M. <https://gdz.bkg.bund.de/index.php/default/geographische-gitter-fur-deutschland-in-utm-projektion-geogitter-national.html>.
- BKG, 2017. Verwaltungsgebiete 1:250 000 (Ebenen), Stand 01.01. (VG250 01.01.) Bundesamt für Kartographie und Geodäsie, Frankfurt a.M. URL <https://gdz.bkg.bund.de/index.php/default/open-data/verwaltungsgebiete-1-250-000-ebenen-stand-01-01-vg250-ebenen-01-01.html>
- BKG, 2016. Digitales Landbedeckungsmodell für Deutschland (LBM-DE2012). Federal Agency for Cartography and Geodesy (BKG), Frankfurt a.M. <https://www.bkg.bund.de/DE/Ueber-das-BKG/Geoinformation/Fernerkundung/Landbedeckungsmodell/landbedeckungsmodell.html> (accessed Nov 2, 2018).
- BMU, BMEL, 2020. Nitratbericht 2020. Bundesministerium für Umwelt, Naturschutz und nukleare Sicherheit & Bundesministerium für Ernährung und Landwirtschaft.
- Bouraoui, F., Grizzetti, B., 2014. Modelling mitigation options to reduce diffuse nitrogen water pollution from agriculture. *Sci. Total Environ.* 468, 1267–1277. <https://doi.org/10.1016/j.scitotenv.2013.07.066>
- Boyer, E.W., Alexander, R.B., Parton, W.J., Li, C., Butterbach-Bahl, K., Donner, S.D., Skaggs, R.W., Grosso, S.J.D., 2006. Modeling Denitrification in Terrestrial and Aquatic Ecosystems at Regional Scales. *Ecological Applications* 16, 2123–2142. [https://doi.org/10.1890/1051-0761\(2006\)016\[2123:MDITAA\]2.0.CO;2](https://doi.org/10.1890/1051-0761(2006)016[2123:MDITAA]2.0.CO;2)
- Boy-Roura, M., Nolan, B.T., Menció, A., Mas-Pla, J., 2013. Regression model for aquifer vulnerability assessment of nitrate pollution in the Osona region (NE Spain). *Journal of Hydrology* 505, 150–162. <https://doi.org/10.1016/j.jhydrol.2013.09.048>
- Breiman, L., 2001. Random Forests. *Machine Learning* 45, 5–32. <https://doi.org/10.1023/A:1010933404324>
- Breiman, L., 1996. Bagging Predictors. *Machine Learning* 24, 123–140. <https://doi.org/10.1023/A:1018054314350>
- Breiman, L., 1984. *Classification and Regression Trees*. Routledge.

References

- Breuer, L., Vache, K.B., Julich, S., Frede, H.-G., 2008. Current concepts in nitrogen dynamics for mesoscale catchments. *Hydrol. Sci. J.-J. Sci. Hydrol.* 53, 1059–1074. <https://doi.org/10.1623/hysj.53.5.1059>
- Burgin, A.J., Hamilton, S.K., 2007. Have we overemphasized the role of denitrification in aquatic ecosystems? A review of nitrate removal pathways. *Front. Ecol. Environ.* 5, 89–96. [https://doi.org/10.1890/1540-9295\(2007\)5\[89:HWOTRO\]2.0.CO;2](https://doi.org/10.1890/1540-9295(2007)5[89:HWOTRO]2.0.CO;2)
- Burow, K.R., Nolan, B.T., Rupert, M.G., Dubrovsky, N.M., 2010. Nitrate in Groundwater of the United States, 1991–2003. *Environ. Sci. Technol.* 44, 4988–4997. <https://doi.org/10.1021/es100546y>
- Cherry, K.A., Shepherd, M., Withers, P.J.A., Mooney, S.J., 2008. Assessing the effectiveness of actions to mitigate nutrient loss from agriculture: A review of methods. *Science of The Total Environment* 406, 1–23. <https://doi.org/10.1016/j.scitotenv.2008.07.015>
- Close, M.E., Abraham, P., Humphries, B., Lilburne, L., Cuthill, T., Wilson, S., 2016. Predicting groundwater redox status on a regional scale using linear discriminant analysis. *Journal of Contaminant Hydrology* 191, 19–32. <https://doi.org/10.1016/j.jconhyd.2016.04.006>
- Dalla Libera, N., Fabbri, P., Mason, L., Piccinini, L., Pola, M., 2017. Geostatistics as a tool to improve the natural background level definition: An application in groundwater. *Science of The Total Environment* 598, 330–340. <https://doi.org/10.1016/j.scitotenv.2017.04.018>
- De'ath, G., 2007. Boosted trees for ecological modeling and prediction. *Ecology* 88, 243–251. [https://doi.org/10.1890/0012-9658\(2007\)88\[243:BTFEMA\]2.0.CO;2](https://doi.org/10.1890/0012-9658(2007)88[243:BTFEMA]2.0.CO;2)
- De'ath, G., Fabricius, K.E., 2000. Classification and Regression Trees: a powerful yet simple technique for ecological data analysis. *Ecology* 81, 3178–3192. [https://doi.org/10.1890/0012-9658\(2000\)081\[3178:CARTAP\]2.0.CO;2](https://doi.org/10.1890/0012-9658(2000)081[3178:CARTAP]2.0.CO;2)
- DeSimone, L.A., Pope, J.P., Ransom, K.M., 2020. Machine-learning models to map pH and redox conditions in groundwater in a layered aquifer system, Northern Atlantic Coastal Plain, eastern USA. *Journal of Hydrology: Regional Studies* 30, 100697. <https://doi.org/10.1016/j.ejrh.2020.100697>
- Dogulu, N., López López, P., Solomatine, D.P., Weerts, A.H., Shrestha, D.L., 2015. Estimation of predictive hydrologic uncertainty using the quantile regression and UNEEC methods and their comparison on contrasting catchments. *Hydrology and Earth System Sciences* 19, 3181–3201. <https://doi.org/https://doi.org/10.5194/hess-19-3181-2015>

- Durand, P., Breuer, L., Johnes, P.J., Billen, G., Klein, J.J.M. de, 2011. Nitrogen processes in aquatic ecosystems, in: Sutton, M. A., Howard, C. M., Erisman, J. W., Billen, G., et Al. (Eds.) European Nitrogen Assessment. Cambridge, pp. 126–146.
- Durand, P., Moreau, P., Salmon-Monviola, J., Ruiz, L., Vertes, F., Gascuel-Oudou, C., 2015. Modelling the interplay between nitrogen cycling processes and mitigation options in farming catchments. *The Journal of Agricultural Science* 153, 959–974. <https://doi.org/10.1017/S0021859615000258>
- DWD, 2018. Climate Data Center: regional averages DE. German Weather Service (DWD). Deutscher Wetterdienst. URL ftp://ftp-cdc.dwd.de/pub/CDC/regional_averages_DE/annual/ (accessed 11.2.18).
- Ebeling, P., Kumar, R., Weber, M., **Knoll, L.**, Fleckenstein, J., Musolff, A., 2020. Archetypes and Controls of Riverine Nutrient Export Across German Catchments. *Earth and Space Science Open Archive*. <https://doi.org/10.1002/essoar.10503375.1>
- EC, 2018. Report from the Commission to the Council and the European Parliament. European Commission.
- EC, 2003. Guidance No 7 - Monitoring under the Water Framework Directive. URL [https://circabc.europa.eu/sd/a/63f7715f-0f45-4955-b7cb-58ca305e42a8/Guidance%20No%207%20-%20Monitoring%20\(WG%202.7\).pdf](https://circabc.europa.eu/sd/a/63f7715f-0f45-4955-b7cb-58ca305e42a8/Guidance%20No%207%20-%20Monitoring%20(WG%202.7).pdf) (accessed 10.22.19).
- EEA, 2018. European waters -- Assessment of status and pressures 2018. European Environment Agency. URL <https://www.eea.europa.eu/publications/state-of-water> (accessed 6.21.19).
- EEA, 1999. Groundwater quality and quantity in Europe. European Environment Agency. URL <https://www.eea.europa.eu/publications/groundwater07012000> (accessed 11.2.18).
- Elith, J., Leathwick, J.R., Hastie, T., 2008. A working guide to boosted regression trees. *J. Anim. Ecol.* 77, 802–813. <https://doi.org/10.1111/j.1365-2656.2008.01390.x>
- Erisman, J.W., Galloway, J.N., Seitzinger, S., Bleeker, A., Dise, N.B., Petrescu, A.M.R., Leach, A.M., de Vries, W., 2013. Consequences of human modification of the global nitrogen cycle. *Philos. Trans. R. Soc. Lond., B, Biol. Sci.* 368, 20130116. <https://doi.org/10.1098/rstb.2013.0116>
- Erisman, J.W., Sutton, M.A., Galloway, J., Klimont, Z., Winiwarter, W., 2008. How a century of ammonia synthesis changed the world. *Nature Geoscience* 1, 636–639. <https://doi.org/10.1038/ngeo325>
- Eschenbach, W., Budziak, D., Elbracht, J., Hoepfer, H., Krienen, L., Kunkel, R., Meyer, K., Well, R., Wendland, F., 2018. Possibilities and limitations of validating modelled nitrate inputs into

- groundwater at the macroscale using the N-2/Ar-method. *Grundwasser* 23, 125–139. <https://doi.org/10.1007/s00767-018-0391-6>
- Foster, S.S.D., Chilton, P.J., 2003. Groundwater: the processes and global significance of aquifer degradation. *Philos Trans R Soc Lond B Biol Sci* 358, 1957–1972. <https://doi.org/10.1098/rstb.2003.1380>
- Fowler, D., Coyle, M., Skiba, U., Sutton, M.A., Cape, J.N., Reis, S., Sheppard, L.J., Jenkins, A., Grizzetti, B., Galloway, J.N., Vitousek, P., Leach, A., Bouwman, A.F., Butterbach-Bahl, K., Dentener, F., Stevenson, D., Amann, M., Voss, M., 2013. The global nitrogen cycle in the twenty-first century. *Philosophical Transactions of the Royal Society B: Biological Sciences* 368, 20130164. <https://doi.org/10.1098/rstb.2013.0164>
- Francis, C.A., Beman, J.M., Kuypers, M.M.M., 2007. New processes and players in the nitrogen cycle: the microbial ecology of anaerobic and archaeal ammonia oxidation. *The ISME Journal* 1, 19–27. <https://doi.org/10.1038/ismej.2007.8>
- Freund, Y., Schapire, R.E., 1996. Experiments with a New Boosting Algorithm, in: *Proceedings of the Thirteenth International Conference on International Conference on Machine Learning, ICML'96*. Morgan Kaufmann Publishers Inc., San Francisco, CA, USA, pp. 148–156.
- Friedel, M.J., Wilson, S.R., Close, M.E., Buscema, M., Abraham, P., Banasiak, L., 2020. Comparison of four learning-based methods for predicting groundwater redox status. *Journal of Hydrology* 580, 124200. <https://doi.org/10.1016/j.jhydrol.2019.124200>
- Friedman, J.H., 2002. Stochastic gradient boosting. *Computational Statistics & Data Analysis, Nonlinear Methods and Data Mining* 38, 367–378. [https://doi.org/10.1016/S0167-9473\(01\)00065-2](https://doi.org/10.1016/S0167-9473(01)00065-2)
- Fuchs, S., Kaiser, M., Kiemle, L., Kittlaus, S., Rothvoß, S., Toshovski, S., Wagner, A., Wander, R., Weber, T., Ziegler, S., 2017a. Modeling of Regionalized Emissions (MoRE) into Water Bodies: An Open-Source River Basin Management System. *Water* 9, 239. <https://doi.org/10.3390/w9040239>
- Fuchs, S., Scherer, U., Wander, R., Behrendt, H., Venohr, M., Opitz, D., Hillenbrand, T., Marscheider-Weidemann, F., Götz, T., 2010. Berechnung von Stoffeinträgen in die Fließgewässer Deutschlands mit dem Modell MONERIS, UBA-Texte 45/2010. German Environment Agency (UBA), Dessau-Roßlau.
- Fuchs, S., Weber, T., Wander, R., Toshovski, S., Kittlaus, S., Reid, L., Bach, M., Klement, L., Hillenbrand, T., Tettenborn, F., 2017b. Effizienz von Maßnahmen zur Reduktion von Stoffeinträgen, UBA-Texte 05/2017. German Environment Agency (UBA), Dessau-Roßlau.

- Galloway, J.N., Aber, J.D., Erisman, J.W., Seitzinger, S.P., Howarth, R.W., Cowling, E.B., Cosby, B.J., 2003. The Nitrogen Cascade. *BioScience* 53, 341–356. [https://doi.org/10.1641/0006-3568\(2003\)053\[0341:TNC\]2.0.CO;2](https://doi.org/10.1641/0006-3568(2003)053[0341:TNC]2.0.CO;2)
- Galloway, J.N., Dentener, F.J., Capone, D.G., Boyer, E.W., Howarth, R.W., Seitzinger, S.P., Asner, G.P., Cleveland, C.C., Green, P.A., Holland, E.A., Karl, D.M., Michaels, A.F., Porter, J.H., Townsend, A.R., Vöösmary, C.J., 2004. Nitrogen Cycles: Past, Present, and Future. *Biogeochemistry* 70, 153–226. <https://doi.org/10.1007/s10533-004-0370-0>
- Galloway, J.N., Townsend, A.R., Erisman, J.W., Bekunda, M., Cai, Z., Freney, J.R., Martinelli, L.A., Seitzinger, S.P., Sutton, M.A., 2008. Transformation of the Nitrogen Cycle: Recent Trends, Questions, and Potential Solutions. *Science* 320, 889–892. <https://doi.org/10.1126/science.1136674>
- Geupel M., Heldstab J., Schächli B., Reutiman J., Bach M., Häußermann U., **Knoll L.**, Klement L., Breuer L., Fuch S., Weber T., 2020. A national nitrogen target for Germany. (submitted)
- Glugla, G., Jankiewicz, P., Rachimow, C., Lojek, K., Richter, K., Fürtig, G., Krahe, P., 2003. Wasserhaushaltsverfahren zur Berechnung vieljähriger Mittelwerte der tatsächlichen Verdunstung und des Gesamtabflusses (BfG-Bericht No. 1342), BfG-Bericht. Bundesanstalt für Gewässerkunde, BfG, Koblenz, Berlin.
- Goulding, K., 2000. Nitrate leaching from arable and horticultural land. *Soil Use and Management* 16, 145–151. <https://doi.org/10.1111/j.1475-2743.2000.tb00218.x>
- Green, C.T., Jurgens, B.C., Zhang, Y., Starn, J.J., Singleton, M.J., Esser, B.K., 2016. Regional oxygen reduction and denitrification rates in groundwater from multi-model residence time distributions, San Joaquin Valley, USA. *Journal of Hydrology*, RESIDENCE TIMES IN SUBSURFACE HYDROLOGICAL SYSTEMS: Signature of hydrological processes and impact on environmental applications 543, Part A, 155–166. <https://doi.org/10.1016/j.jhydrol.2016.05.018>
- Green, C.T., Puckett, L.J., Böhlke, J.K., Bekins, B.A., Phillips, S.P., Kauffman, L.J., Denver, J.M., Johnson, H.M., 2008. Limited Occurrence of Denitrification in Four Shallow Aquifers in Agricultural Areas of the United States. *Journal of Environmental Quality* 37, 994–1009. <https://doi.org/10.2134/jeq2006.0419>
- Grimm-Strele, J., Casper, M., van Dijk, P., Finck, M., Gudera, T., Korte, S., 2008. Modelling system MoNit for simulation of groundwater pollution by nitrate of the Upper Rhine valley aquifer. *WasserWirtschaft* 98, 55–59.

- Grizzetti, B., Bouraoui, F., Marsily, G.D., 2008. Assessing nitrogen pressures on European surface water. *Global Biogeochemical Cycles* 22. <https://doi.org/10.1029/2007GB003085>
- Grizzetti, B., Passy, P., Billen, G., Bouraoui, F., Garnier, J., Lassaletta, L., 2015. The role of water nitrogen retention in integrated nutrient management: assessment in a large basin using different modelling approaches. *Environ. Res. Lett.* 10, 065008. <https://doi.org/10.1088/1748-9326/10/6/065008>
- Gruber, N., Galloway, J.N., 2008. An Earth-system perspective of the global nitrogen cycle. *Nature* 451, 293–296. <https://doi.org/10.1038/nature06592>
- GrwV, 2010. Grundwasserverordnung vom 9. November 2010 (BGBl. I S. 1513), zuletzt geändert durch Artikel 1 der Verordnung vom 4. Mai 2017, BGBl. I S. 1044.
- Gu, B., Ge, Y., Chang, S.X., Luo, W., Chang, J., 2013. Nitrate in groundwater of China: Sources and driving forces. *Global Environmental Change* 23, 1112–1121. <https://doi.org/10.1016/j.gloenvcha.2013.05.004>
- Hannappel, S., Koepf, C., Bach, T., 2018. Characterization of the denitrification potential of aquifers in Saxony-Anhalt. *Grundwasser* 23, 311–321. <https://doi.org/10.1007/s00767-018-0402-7>
- Hastie, T., Tibshirani, R., Friedman, J., 2009. *The Elements of Statistical Learning: Data Mining, Inference, and Prediction*, Second Edition, 2nd ed, Springer Series in Statistics. Springer-Verlag, New York.
- Häußermann, U., Bach, M., Klement, L., Breuer, L., 2019. Stickstoff-Flächenbilanzen für Deutschland mit Regionalgliederung Bundesländer und Kreise – Jahre 1995 bis 2017. Umweltbundesamt.
- Heldstab, J., Schäppi, B., Reutimann, J., Bach, M., Häußermann, U., Knoll, L., Klement, L., Breuer, L., Fuchs, S., Weber, T., 2020. Integrated nitrogen indicator, national nitrogen target and the current situation in Germany (DESTINO Report 1). Umweltbundesamt.
- Herrmann, F., 2010. Entwicklung einer Methodik zur großräumigen Modellierung von Grundwasserdruckflächen am Beispiel der Grundwasserleiter des Bundeslandes Hessen. BTU Cottbus, Cottbus.
- Hill, A.R., 2019. Groundwater nitrate removal in riparian buffer zones: a review of research progress in the past 20 years. *Biogeochemistry* 143, 347–369. <https://doi.org/10.1007/s10533-019-00566-5>
- Hirt, U., Kreins, P., Kuhn, U., Mahnkopf, J., Venohr, M., Wendland, F., 2012. Management options to reduce future nitrogen emissions into rivers: A case study of the Weser river basin, Germany. *Agricultural Water Management* 115, 118–131. <https://doi.org/10.1016/j.agwat.2012.08.005>

- HLNUG, 2018. Grundwasserbeschaffenheitsbericht 2017, Grundwasser in Hessen, Heft 3. Hessisches Landesamt für Naturschutz, Umwelt und Geologie (HLNUG), Wiesbaden.
- Højberg, A.L., Hansen, A.L., Wachniew, P., Żurek, A.J., Virtanen, S., Arustiene, J., Strömqvist, J., Rankinen, K., Refsgaard, J.C., 2017. Review and assessment of nitrate reduction in groundwater in the Baltic Sea Basin. *Journal of Hydrology: Regional Studies* 12, 50–68. <https://doi.org/10.1016/j.ejrh.2017.04.001>
- Hu, K., Huang, Y., Li, H., Li, B., Chen, D., White, R.E., 2005. Spatial variability of shallow groundwater level, electrical conductivity and nitrate concentration, and risk assessment of nitrate contamination in North China Plain. *Environment International, Soil Contamination and Environmental Health* 31, 896–903. <https://doi.org/10.1016/j.envint.2005.05.028>
- Jankiewicz, P., Neumann, J., Duijnsveld, W.H.M., Wessolek, G., Wycisk, P., Hennings, V., 2005. Abflusshöhe – Sickerwasserrate – Grundwasserneubildung – Drei Themen im Hydrologischen Atlas von Deutschland. *Hydrologie und Wasserbewirtschaftung* 49, 2–13.
- Johnson, T.D., Belitz, K., 2009. Assigning land use to supply wells for the statistical characterization of regional groundwater quality: Correlating urban land use and VOC occurrence. *Journal of Hydrology* 370, 100–108. <https://doi.org/10.1016/j.jhydrol.2009.02.056>
- Jørgensen, P.R., Urup, J., Helstrup, T., Jensen, M.B., Eiland, F., Vinther, F.P., 2004. Transport and reduction of nitrate in clayey till underneath forest and arable land. *Journal of Contaminant Hydrology* 73, 207–226. <https://doi.org/10.1016/j.jconhyd.2004.01.005>
- Kallis, G., Butler, D., 2001. The EU water framework directive: measures and implications. *Water Policy* 3, 125–142. [https://doi.org/10.1016/S1366-7017\(01\)00007-1](https://doi.org/10.1016/S1366-7017(01)00007-1)
- Khalil, A., Almasri, M.N., McKee, M., Kaluarachchi, J.J., 2005. Applicability of statistical learning algorithms in groundwater quality modeling. *Water Resources Research* 41. <https://doi.org/10.1029/2004WR003608>
- Kihumba, A.M., Longo, J.N., Vanclooster, M., 2016. Modelling nitrate pollution pressure using a multivariate statistical approach: the case of Kinshasa groundwater body, Democratic Republic of Congo. *Hydrogeol. J.* 24, 425–437. <https://doi.org/10.1007/s10040-015-1337-z>
- Kim, N.W., Chung, I.M., Won, Y.S., Arnold, J.G., 2008. Development and application of the integrated SWAT–MODFLOW model. *Journal of Hydrology* 356, 1–16. <https://doi.org/10.1016/j.jhydrol.2008.02.024>

- Kludt, C., Weber, F.-A., Bergmann, A., Knöller, K., Berthold, G., Schüth, C., 2016. Identifizierung der Nitratabbauprozesse und Prognose des Nitratabbaupotenzials in den Sedimenten des Hessischen Rieds. *Grundwasser* 21, 227–241. <https://doi.org/10.1007/s00767-015-0317-5>
- Knoll, L.**, Breuer, L., Bach, M., 2019. Large scale prediction of groundwater nitrate concentrations from spatial data using machine learning. *Science of The Total Environment* 668, 1317–1327. <https://doi.org/10.1016/j.scitotenv.2019.03.045>
- Knoll, L.**, Breuer, L., Bach, M., 2020. Nation-wide estimation of groundwater redox conditions and nitrate concentrations through machine learning. *Environ. Res. Lett.* 15, 064004. <https://doi.org/10.1088/1748-9326/ab7d5c>
- Knoll, L.**, Häußermann, U., Breuer, L., Bach, M., 2020. Spatial Distribution of Integrated Nitrate Reduction across the Unsaturated Zone and the Groundwater Body in Germany. *Water* 12, 2456. <https://doi.org/10.3390/w12092456>
- Knoll L.**, Bach M., Häußermann U., Breuer L., Fuch S., Morling K., Weber T., 2020. Abschätzung der Stickstoffreduktion in der ungesättigten Zone und im Grundwasser für die Modellierung des N-Eintrags in Oberflächengewässer mit dem Modellinstrument MoRE. Umweltbundesamt, Dessau-Roßlau, UBA-Texte 00/2020, 96 S. (under review)
- Knowles, R., 1982. Denitrification. *Microbiol Rev* 46, 43–70.
- Koch, J., Stisen, S., Refsgaard, J.C., Ernsten, V., Jakobsen, P.R., Højberg, A.L., 2019. Modeling Depth of the Redox Interface at High Resolution at National Scale Using Random Forest and Residual Gaussian Simulation. *Water Resources Research* 55, 1451–1469. <https://doi.org/10.1029/2018WR023939>
- Korom, S.F., 1992. Natural denitrification in the saturated zone: A review. *Water Resources Research* 28, 1657–1668. <https://doi.org/10.1029/92WR00252>
- Kreins, P., Behrendt, H., Gömann, H., Hirt, U., Kunkel, R., Seidel, K., Tetzlaff, B., Wendland, F., 2010. Analyse von Agrar- und Umweltmaßnahmen im Bereich des landwirtschaftlichen Gewässerschutzes vor dem Hintergrund der EG-Wasserrahmenrichtlinie in der Flussgebietseinheit Weser—AGRUM Weser (No. Sonderheft 336), *Landbauforschung—vTI agriculture and forestry research*. Braunschweig.
- Kuhn, M., 2018. Caret: Classification and Regression Training, R package version 6.0-82. <https://cran.r-project.org/web/packages/caret/index.html>.
- Kuhn, M., Johnson, K., 2013. *Applied Predictive Modeling*. Springer-Verlag, New York.

- Kuhr, P., Haider, J., Kreins, P., Kunkel, R., Tetzlaff, B., Vereecken, H., Wendland, F., 2013. Model Based Assessment of Nitrate Pollution of Water Resources on a Federal State Level for the Dimensioning of Agro-environmental Reduction Strategies. *Water Resour Manage* 27, 885–909. <https://doi.org/10.1007/s11269-012-0221-z>
- Kunkel, R., Bach, M., Behrendt, H., Wendland, F., 2004. Groundwater-borne nitrate intakes into surface waters in Germany. *Water Sci Technol* 49, 11–19. <https://doi.org/10.2166/wst.2004.0152>
- Kunkel, R., Herrmann, F., Kape, H.-E., Keller, L., Koch, F., Tetzlaff, B., Wendland, F., 2017. Simulation of terrestrial nitrogen fluxes in Mecklenburg-Vorpommern and scenario analyses how to reach N-quality targets for groundwater and the coastal waters. *Environ. Earth Sci.* 76, 146. <https://doi.org/10.1007/s12665-017-6437-8>
- Kunkel, R., Montzka, C., Wendland, F., 2007. Anwendung des im Forschungszentrum Jülich entwickelten Modells WEKU zur Ableitung flussgebietsbezogener Aufenthaltszeiten des Grundwassers in Deutschland (Endbericht). Forschungszentrum Jülich.
- Kunkel, R., Wendland, F., 2002. The GROWA98 model for water balance analysis in large river basins - the river Elbe case study. *J. Hydrol.* 259, 152–162. [https://doi.org/10.1016/S0022-1694\(01\)00579-0](https://doi.org/10.1016/S0022-1694(01)00579-0)
- Kunkel, R., Wendland, F., 1997. WEKU – a GIS-Supported stochastic model of groundwater residence times in upper aquifers for the supraregional groundwater management. *Environmental Geology* 30, 1–9. <https://doi.org/10.1007/s002540050126>
- Kursa, M.B., Rudnicki, W.R., 2010. Feature Selection with the Boruta Package. *Journal of Statistical Software* 36, 1–13. <https://doi.org/10.18637/jss.v036.i11>
- LAWA, 2018. Konzept zur Beurteilung der Grundwassergüte anhand weiterer Stoffgehalte unter Berücksichtigung eines möglichen Nitratabbaus. Bund/Länder-Arbeitsgemeinschaft Wasser (German Working Group on Water Issues), Ständiger Ausschuss „Grundwasser und Wasserversorgung (AG)“, Stand 06. Februar 2018.
- Liao, L., Green, C.T., Bekins, B.A., Böhlke, J.K., 2012. Factors controlling nitrate fluxes in groundwater in agricultural areas. *Water Resour. Res.* 48, W00L09. <https://doi.org/10.1029/2011WR011008>
- Liaw, A., Wiener, M., 2002. Classification and Regression by randomForest. *R News* 2, 18–22.
- Liu, A.G., Ming, J.H., Ankumah, R.O., 2005. Nitrate contamination in private wells in rural Alabama, United States. *Sci. Total Environ.* 346, 112–120. <https://doi.org/10.1016/j.scitotenv.2004.11.019>

References

- Mattern, S., Fasbender, D., Vanclooster, M., 2009. Discriminating sources of nitrate pollution in an unconfined sandy aquifer. *Journal of Hydrology* 376, 275–284. <https://doi.org/10.1016/j.jhydrol.2009.07.039>
- McAleer, E.B., Coxon, C.E., Richards, K.G., Jahangir, M.M.R., Grant, J., Mellander, P.E., 2017. Groundwater nitrate reduction versus dissolved gas production: A tale of two catchments. *Science of The Total Environment* 586, 372–389. <https://doi.org/10.1016/j.scitotenv.2016.11.083>
- McMahon, P.B., Chapelle, F.H., 2008. Redox Processes and Water Quality of Selected Principal Aquifer Systems. *Groundwater* 46, 259–271. <https://doi.org/10.1111/j.1745-6584.2007.00385.x>
- Meinshausen, N., 2017. quantregForest: Quantile Regression Forests.
- Meinshausen, N., 2006. Quantile Regression Forests. *Journal of Machine Learning Research* 7, 983–999.
- Merz, C., Steidl, J., Dannowski, R., 2009. Parameterization and regionalization of redox based denitrification for GIS-embedded nitrate transport modeling in Pleistocene aquifer systems. *Environ Geol* 58, 1587. <https://doi.org/10.1007/s00254-008-1665-6>
- Mfumu Kihumba, A., Ndembo Longo, J., Vanclooster, M., 2016. Modelling nitrate pollution pressure using a multivariate statistical approach: the case of Kinshasa groundwater body, Democratic Republic of Congo. *Hydrogeology Journal* 24, 425–437. <https://doi.org/10.1007/s10040-015-1337-z>
- Nas, B., Berktaş, A., 2010. Groundwater quality mapping in urban groundwater using GIS. *Environ Monit Assess* 160, 215–227. <https://doi.org/10.1007/s10661-008-0689-4>
- Neumann, J., 2005. Flächendifferenzierte Grundwasserneubildung von Deutschland - Entwicklung und Anwendung des makroskaligen Verfahrens HAD-GWNeu. Martin-Luther-Universität Halle-Wittenberg, Halle/Saale.
- Nguyen, V.T., Dietrich, J., 2018. Modification of the SWAT model to simulate regional groundwater flow using a multicell aquifer. *Hydrological Processes* 32, 939–953. <https://doi.org/10.1002/hyp.11466>
- Nixdorf, E., Sun, Y., Lin, M., Kolditz, O., 2017. Development and application of a novel method for regional assessment of groundwater contamination risk in the Songhua River Basin. *Science of The Total Environment* 605–606, 598–609. <https://doi.org/10.1016/j.scitotenv.2017.06.126>

- Nolan, B.T., Fienen, M.N., Lorenz, D.L., 2015. A statistical learning framework for groundwater nitrate models of the Central Valley, California, USA. *J. Hydrol.* 531, 902–911. <https://doi.org/10.1016/j.jhydrol.2015.10.025>
- Nolan, B.T., Gronberg, J.M., Faunt, C.C., Eberts, S.M., Belitz, K., 2014. Modeling Nitrate at Domestic and Public-Supply Well Depths in the Central Valley, California. *Environ. Sci. Technol.* 48, 5643–5651. <https://doi.org/10.1021/es405452q>
- Nolan, B.T., Hitt, K.J., 2006. Vulnerability of Shallow Groundwater and Drinking-Water Wells to Nitrate in the United States. *Environ. Sci. Technol.* 40, 7834–7840. <https://doi.org/10.1021/es060911u>
- Nolan, B.T., Hitt, K.J., Ruddy, B.C., 2002. Probability of Nitrate Contamination of Recently Recharged Groundwaters in the Conterminous United States. *Environ. Sci. Technol.* 36, 2138–2145. <https://doi.org/10.1021/es0113854>
- Oehler, F., Durand, P., Bordenave, P., Saadi, Z., Salmon-Monviola, J., 2009. Modelling denitrification at the catchment scale. *Science of The Total Environment* 407, 1726–1737. <https://doi.org/10.1016/j.scitotenv.2008.10.069>
- Ouedraogo, I., Defourny, P., Vanclooster, M., 2019. Application of random forest regression and comparison of its performance to multiple linear regression in modeling groundwater nitrate concentration at the African continent scale. *Hydrogeol J* 27, 1081–1098. <https://doi.org/10.1007/s10040-018-1900-5>
- Ouedraogo, I., Defourny, P., Vanclooster, M., 2016. Mapping the groundwater vulnerability for pollution at the pan African scale. *Science of The Total Environment* 544, 939–953. <https://doi.org/10.1016/j.scitotenv.2015.11.135>
- Ouedraogo, I., Vanclooster, M., 2016. A meta-analysis and statistical modelling of nitrates in groundwater at the African scale. *Hydrol. Earth Syst. Sci.* 20, 2353–2381. <https://doi.org/10.5194/hess-20-2353-2016>
- Panagopoulos, G.P., Antonakos, A.K., Lambrakis, N.J., 2006. Optimization of the DRASTIC method for groundwater vulnerability assessment via the use of simple statistical methods and GIS. *Hydrogeol J* 14, 894–911. <https://doi.org/10.1007/s10040-005-0008-x>
- Pebesma, E.J., de Kwaadsteniet, J.W., 1997. Mapping groundwater quality in the Netherlands. *Journal of Hydrology* 200, 364–386. [https://doi.org/10.1016/S0022-1694\(97\)00027-9](https://doi.org/10.1016/S0022-1694(97)00027-9)

References

- Rahmati, O., Choubin, B., Fathabadi, A., Coulon, F., Soltani, E., Shahabi, H., Mollaefar, E., Tiefenbacher, J., Cipullo, S., Ahmad, B.B., Tien Bui, D., 2019. Predicting uncertainty of machine learning models for modelling nitrate pollution of groundwater using quantile regression and UNEEC methods. *Science of The Total Environment* 688, 855–866. <https://doi.org/10.1016/j.scitotenv.2019.06.320>
- Ransom, K.M., Nolan, B.T., A. Traum, J., Faunt, C.C., Bell, A.M., Gronberg, J.A.M., Wheeler, D.C., Z. Rosecrans, C., Jurgens, B., Schwarz, G.E., Belitz, K., M. Eberts, S., Kourakos, G., Harter, T., 2017. A hybrid machine learning model to predict and visualize nitrate concentration throughout the Central Valley aquifer, California, USA. *Science of The Total Environment* 601–602, 1160–1172. <https://doi.org/10.1016/j.scitotenv.2017.05.192>
- Refsgaard, J.C., van der Sluijs, J.P., Højberg, A.L., Vanrolleghem, P.A., 2007. Uncertainty in the environmental modelling process – A framework and guidance. *Environmental Modelling & Software* 22, 1543–1556. <https://doi.org/10.1016/j.envsoft.2007.02.004>
- Reis, S., Bekunda, M., Howard, C.M., Karanja, N., Winiwarter, W., Yan, X., Bleeker, A., Sutton, M.A., 2016. Synthesis and review: Tackling the nitrogen management challenge: from global to local scales. *Environ. Res. Lett.* 11, 120205. <https://doi.org/10.1088/1748-9326/11/12/120205>
- Rivas, A., Singh, R., Horne, D., Roygard, J., Matthews, A., Hedley, M.J., 2017. Denitrification potential in the subsurface environment in the Manawatu River catchment, New Zealand: Indications from oxidation-reduction conditions, hydrogeological factors, and implications for nutrient management. *Journal of Environmental Management* 197, 476–489. <https://doi.org/10.1016/j.jenvman.2017.04.015>
- Rivett, M.O., Buss, S.R., Morgan, P., Smith, J.W.N., Bemment, C.D., 2008. Nitrate attenuation in groundwater: A review of biogeochemical controlling processes. *Water Research* 42, 4215–4232. <https://doi.org/10.1016/j.watres.2008.07.020>
- Rivett, M.O., Smith, J.W.N., Buss, S.R., Morgan, P., 2007. Nitrate occurrence and attenuation in the major aquifers of England and Wales. *Quarterly Journal of Engineering Geology and Hydrogeology* 40, 335–352. <https://doi.org/10.1144/1470-9236/07-032>
- Rodriguez-Galiano, V.F., Luque-Espinar, J.A., Chica-Olmo, M., Mendes, M.P., 2018. Feature selection approaches for predictive modelling of groundwater nitrate pollution: An evaluation of filters, embedded and wrapper methods. *Science of The Total Environment* 624, 661–672. <https://doi.org/10.1016/j.scitotenv.2017.12.152>

- Rodriguez-Galiano, V.F., Mendes, M.P., Garcia-Soldado, M.J., Chica-Olmo, M., Ribeiro, L., 2014. Predictive modeling of groundwater nitrate pollution using Random Forest and multisource variables related to intrinsic and specific vulnerability: A case study in an agricultural setting (Southern Spain). *Science of The Total Environment* 476–477, 189–206. <https://doi.org/10.1016/j.scitotenv.2014.01.001>
- Rosecrans, C.Z., Nolan, B.T., Gronberg, J.M., 2017. Prediction and visualization of redox conditions in the groundwater of Central Valley, California. *Journal of Hydrology* 546, 341–356. <https://doi.org/10.1016/j.jhydrol.2017.01.014>
- Schwientek, M., Einsiedl, F., Stichler, W., Stögbauer, A., Strauss, H., Maloszewski, P., 2008. Evidence for denitrification regulated by pyrite oxidation in a heterogeneous porous groundwater system. *Chemical Geology* 255, 60–67. <https://doi.org/10.1016/j.chemgeo.2008.06.005>
- Seitzinger, S., Harrison, J.A., Böhlke, J.K., Bouwman, A.F., Lowrance, R., Peterson, B., Tobias, C., Drecht, G.V., 2006. Denitrification Across Landscapes and Waterscapes: A Synthesis. *Ecological Applications* 16, 2064–2090. [https://doi.org/10.1890/1051-0761\(2006\)016\[2064:DALAWA\]2.0.CO;2](https://doi.org/10.1890/1051-0761(2006)016[2064:DALAWA]2.0.CO;2)
- Shrestha, D.L., Solomatine, D.P., 2006. Machine learning approaches for estimation of prediction interval for the model output. *Neural Networks, Earth Sciences and Environmental Applications of Computational Intelligence* 19, 225–235. <https://doi.org/10.1016/j.neunet.2006.01.012>
- Stahl, M.O., Gehring, J., Jameel, Y., 2020. Isotopic variation in groundwater across the conterminous United States – Insight into hydrologic processes. *Hydrological Processes* 34, 3506–3523. <https://doi.org/10.1002/hyp.13832>
- Stein, M.L., 1999. *Interpolation of spatial data: some theory for kriging*. Springer, New York.
- Sutton, M.A., Bleeker, A., Howard, C.M., Bekunda, M., Grizzetti, B., de Vries, W., van Grinsven, H.J.M., Abro, Y.P., Adhya, T.K., Billen, G., Davidson, E.A., Datta, A., Diaz, R., Erisman, J.W., Liu, X., J., Oenema, O., Palm, C., Raghuram, N., Reis, S., Scholz, R.W., Sims, T., Westhoek, H., Zhang, F.S., 2013. *Our Nutrient World - The challenge to produce more food and energy with less pollution*. Centre for Ecology and Hydrology (CEH), Edinburgh UK.
- Szatmári, G., Pásztor, L., 2019. Comparison of various uncertainty modelling approaches based on geostatistics and machine learning algorithms. *Geoderma* 337, 1329–1340. <https://doi.org/10.1016/j.geoderma.2018.09.008>
- Tedd, K.M., Coxon, C.E., Misstear, B.D.R., Daly, D., Craig, M., Mannix, A., Williams, N.H.H., 2014. An integrated pressure and pathway approach to the spatial analysis of groundwater nitrate: A

References

- case study from the southeast of Ireland. *Science of The Total Environment* 476–477, 460–476. <https://doi.org/10.1016/j.scitotenv.2013.12.085>
- Tesoriero, A.J., Gronberg, J.A., Juckem, P.F., Miller, M.P., Austin, B.P., 2017. Predicting redox-sensitive contaminant concentrations in groundwater using random forest classification. *Water Resources Research* 53, 7316–7331. <https://doi.org/10.1002/2016WR020197>
- Tesoriero, A.J., Terziotti, S., Abrams, D.B., 2015. Predicting Redox Conditions in Groundwater at a Regional Scale. *Environmental Science & Technology* 49, 9657–9664. <https://doi.org/10.1021/acs.est.5b01869>
- TukeyHSD function | R Documentation, n.d. URL <https://www.rdocumentation.org/packages/stats/versions/3.6.2/topics/TukeyHSD> (accessed 7.24.20).
- Tyralis, H., Papacharalampous, G., Langousis, A., 2019. A Brief Review of Random Forests for Water Scientists and Practitioners and Their Recent History in Water Resources. *Water* 11, 910. <https://doi.org/10.3390/w11050910>
- UBA, 2017a. *Gewässer in Deutschland: Zustand und Bewertung*. Umweltbundesamt, Dessau-Roßlau.
- UBA, 2017b. *Quantifizierung der landwirtschaftlich verursachten Kosten zur Sicherung der Trinkwasserbereitstellung*, UBA-Texte 43/2017. Umweltbundesamt, Dessau-Roßlau.
- Van Egmond, K., Bresser, T., Bouwman, L., 2002. The European Nitrogen Case. *ambi* 31, 72–78. <https://doi.org/10.1579/0044-7447-31.2.72>
- van Grinsven, H.J., Ward, M.H., Benjamin, N., de Kok, T.M., 2006. Does the evidence about health risks associated with nitrate ingestion warrant an increase of the nitrate standard for drinking water? *Environ Health* 5, 26. <https://doi.org/10.1186/1476-069X-5-26>
- Vaysse, K., Lagacherie, P., 2017. Using quantile regression forest to estimate uncertainty of digital soil mapping products. *Geoderma* 291, 55–64. <https://doi.org/10.1016/j.geoderma.2016.12.017>
- Velthof, G.L., Oudendag, D., Witzke, H.P., Asman, W. a. H., Klimont, Z., Oenema, O., 2009. Integrated Assessment of Nitrogen Losses from Agriculture in EU-27 using MITERRA-EUROPE. *Journal of Environmental Quality* 38, 402–417. <https://doi.org/10.2134/jeq2008.0108>
- Venohr, M., Hirt, U., Hofmann, J., Opitz, D., Gericke, A., Wetzig, A., Natho, S., Neumann, F., Hürdler, J., Matranga, M., Mahnkopf, J., Gadegast, M., Behrendt, H., 2011. Modelling of Nutrient Emissions in River Systems – MONERIS – Methods and Background. *International Review of Hydrobiology* 96, 435–483. <https://doi.org/10.1002/iroh.201111331>

- Vitousek, P.M., Aber, J.D., Howarth, R.W., Likens, G.E., Matson, P.A., Schindler, D.W., Schlesinger, W.H., Tilman, D.G., 1997. Human Alteration of the Global Nitrogen Cycle: Sources and Consequences. *Ecological Applications* 7, 737–750. [https://doi.org/10.1890/1051-0761\(1997\)007\[0737:HAOTGN\]2.0.CO;2](https://doi.org/10.1890/1051-0761(1997)007[0737:HAOTGN]2.0.CO;2)
- Voulvoulis, N., Arpon, K.D., Giakoumis, T., 2017. The EU Water Framework Directive: From great expectations to problems with implementation. *Science of The Total Environment* 575, 358–366. <https://doi.org/10.1016/j.scitotenv.2016.09.228>
- Wang, B., Oldham, C., Hipsey, M.R., 2016. Comparison of Machine Learning Techniques and Variables for Groundwater Dissolved Organic Nitrogen Prediction in an Urban Area. *Procedia Engineering*, 12th International Conference on Hydroinformatics (HIC 2016) - Smart Water for the Future 154, 1176–1184. <https://doi.org/10.1016/j.proeng.2016.07.527>
- Wellen, C., Kamran-Disfani, A.-R., Arhonditsis, G.B., 2015. Evaluation of the Current State of Distributed Watershed Nutrient Water Quality Modeling. *Environ. Sci. Technol.* 49, 3278–3290. <https://doi.org/10.1021/es5049557>
- Wendland, F., Bergmann, S., Eisele, M., Gömann, H., Herrmann, F., Kreins, P., Kunkel, R., 2020. Model-Based Analysis of Nitrate Concentration in the Leachate—The North Rhine-Westfalia Case Study, Germany. *Water* 12, 550. <https://doi.org/10.3390/w12020550>
- Wendland, F., Berthold, G., Fritsche, J.-G., Herrmann, F., Kunkel, R., Voigt, H.-J., Vereecken, H., 2011. Konzeptionelles hydrogeologisches Modell zur Analyse und Bewertung von Verweilzeiten in Hessen. *Grundwasser* 16, 163–176. <https://doi.org/10.1007/s00767-011-0169-6>
- Wendland, F., Blum, A., Coetsiers, M., Gorova, R., Griffioen, J., Grima, J., Hinsby, K., Kunkel, R., Marandi, A., Melo, T., Panagopoulos, A., Pauwels, H., Ruisi, M., Traversa, P., Vermooten, J.S.A., Walraevens, K., 2008. European aquifer typology: a practical framework for an overview of major groundwater composition at European scale. *Environ Geol* 55, 77–85. <https://doi.org/10.1007/s00254-007-0966-5>
- Wheeler, D.C., Nolan, B.T., Flory, A.R., DellaValle, C.T., Ward, M.H., 2015. Modeling groundwater nitrate concentrations in private wells in Iowa. *Science of The Total Environment* 536, 481–488. <https://doi.org/10.1016/j.scitotenv.2015.07.080>
- Whelan, M.J., Gandolfi, C., 2002. Modelling of spatial controls on denitrification at the landscape scale. *Hydrological Processes* 16, 1437–1450. <https://doi.org/10.1002/hyp.354>

- WHO, 2017. Guidelines for drinking-water quality, 4th edition, incorporating the 1st addendum. WHO. URL http://www.who.int/water_sanitation_health/publications/drinking-water-quality-guidelines-4-including-1st-addendum/en/ (accessed 11.2.18).
- Wick, K., Heumesser, C., Schmid, E., 2012. Groundwater nitrate contamination: Factors and indicators. *Journal of Environmental Management* 111, 178–186. <https://doi.org/10.1016/j.jenvman.2012.06.030>
- Wilde, S., Hansen, C., Bergmann, A., 2017. Decreasing denitrification capacity in aquifers: scaled model-based evaluation. *Grundwasser* 22, 293–308. <https://doi.org/10.1007/s00767-017-0373-0>
- Wilson, S.R., Close, M.E., Abraham, P., 2018. Applying linear discriminant analysis to predict groundwater redox conditions conducive to denitrification. *J. Hydrol.* 556, 611–624. <https://doi.org/10.1016/j.jhydrol.2017.11.045>
- Wilson, S.R., Close, M.E., Abraham, P., Sarris, T.S., Banasiak, L., Stenger, R., Hadfield, J., 2020. Achieving unbiased predictions of national-scale groundwater redox conditions via data oversampling and statistical learning. *Science of The Total Environment* 705, 135877. <https://doi.org/10.1016/j.scitotenv.2019.135877>
- Wisotzky, F., Wohnlich, S., Böddeker, M., 2018. Nitratreduktion in einem quartären Grundwasserleiter in Ostwestfalen, NRW. *Grundwasser* 23, 167–176. <https://doi.org/10.1007/s00767-017-0379-7>
- Worrall, F., Howden, N.J.K., Burt, T.P., 2015. Evidence for nitrogen accumulation: the total nitrogen budget of the terrestrial biosphere of a lowland agricultural catchment. *Biogeochemistry* 123, 411–428. <https://doi.org/10.1007/s10533-015-0074-7>
- Wriedt, G., de Vries, D., Eden, T., Federolf, C., 2019. Mapping groundwater nitrate concentrations in Lower Saxony. *Grundwasser* 24, 27–41. <https://doi.org/10.1007/s00767-019-00415-0>
- Wriedt, G., Rode, M., 2006. Modelling nitrate transport and turnover in a lowland catchment system. *Journal of Hydrology, Measurement and Parameterization of Rainfall Microstructure* 328, 157–176. <https://doi.org/10.1016/j.jhydrol.2005.12.017>
- Zhang, Y.-C., Slomp, C.P., Broers, H.P., Bostick, B., Passier, H.F., Böttcher, M.E., Omoregie, E.O., Lloyd, J.R., Polya, D.A., Van Cappellen, P., 2012. Isotopic and microbiological signatures of pyrite-driven denitrification in a sandy aquifer. *Chemical Geology* 300–301, 123–132. <https://doi.org/10.1016/j.chemgeo.2012.01.024>

Zhou, W., Ma, Y., Well, R., Wang, H., Yan, X., 2018. Denitrification in Shallow Groundwater Below Different Arable Land Systems in a High Nitrogen-Loading Region. *Journal of Geophysical Research: Biogeosciences* 123, 991–1004. <https://doi.org/10.1002/2017JG004199>

Acknowledgements

Finally, I would like to thank the people that have made this work possible:

First of all, I would like to thank my supervisor Lutz for the straightforward and open discussion and his always very constructive support during the time of my PhD. Special thanks also belong to Martin for the initiation and coordination of the project in which I was able to participate. Without Martin this dissertation would not have been possible, many thanks for your open ear and the input on any research ideas and questions that arose during the work and meetings.

Thanks goes to Antje Ullrich from the German Environment Agency which funded the project (no. 3715 22 2200) in the framework of the environment research plan of the Federal Ministry for the Environment, Nature Conservation and Nuclear Safety and to the working group of Prof. Dr. Fuchs from the Karlsruhe Institute of Technology which cooperates on the project.

To all colleagues and friends from our institute the Department of Landscape Ecology and Resources Management: Thank you for the great time and the enjoyable discussions during coffee, lunch, beer, field trips to Schwingbach or at conferences.

I want to thank all my friends from the time in Darmstadt, Freiburg and Munich for many gorgeous moments and the very important balance in my life. Special thanks to Jonas not only for the numerous technical discussions and the input from the agricultural field but also for the many entertaining after work hours.

A substantial thank goes to my family, especially, Regina, Stefan and Leoni for your valuable support, the continuous motivation and all the great time with you.

I thank Jule for her endless support and for making the PhD a very special and enjoyable time.

Declaration

I declare that I have completed this dissertation single-handedly without the unauthorized help of a second party and only with the assistance acknowledged therein. I have appropriately acknowledged and cited all text passages that are derived verbatim from or are based on the content of published work of others, and all information relating to verbal communications. I consent to the use of an anti-plagiarism software to check my thesis. I have abided by the principles of good scientific conduct laid down in the charter of the Justus Liebig University Giessen „Satzung der Justus-Liebig-Universität Gießen zur Sicherung guter wissenschaftlicher Praxis“ in carrying out the investigations described in the dissertation.”

Ich erkläre: Ich habe die vorgelegte Dissertation selbständig und ohne unerlaubte fremde Hilfe und nur mit den Hilfen angefertigt, die ich in der Dissertation angegeben habe. Alle Textstellen, die wörtlich oder sinngemäß aus veröffentlichten Schriften entnommen sind, und alle Angaben, die auf mündlichen Auskünften beruhen, sind als solche kenntlich gemacht. Bei den von mir durchgeführten und in der Dissertation erwähnten Untersuchungen habe ich die Grundsätze guter wissenschaftlicher Praxis, wie sie in der „Satzung der Justus-Liebig-Universität Gießen zur Sicherung guter wissenschaftlicher Praxis“ niedergelegt sind, eingehalten

Giessen, 6th November 2020

Lukas Knoll