# Challenges in groundwater pollution management in transboundary basins: Groundwater pollution mapping at the African scale

Issoufou Ouedraogo[1], and Marnik Vanclooster[2]

[1]Ecole Supérieure d'Ingénieurs, Université de Fada N'Gourma, BP 46 Fada N'Gourma, Burkina Faso.

[2]Earth and Life Institute, Université catholique de Louvain, Croix du Sud 2, Box 2, B-1348 Louvain-la-Neuve, Belgium.

*Corresponding author*: ouedraogo.issoufou03@gmail.com

## Abstract

Groundwater is a major component of available freshwater and it is an essential natural resource for overall development. In the hydrological system, groundwater is closely linked to surface water. Groundwater and surface water should, therefore, be managed coherently together in an integrated way. However, there are some inherent structural challenges related to such an approach. First, the areal extent of surface water basins and the underlying groundwater systems often differ radically (BGR and UNESCO, 2012), making groundwater management by river basin organizations complex. Second, given the subsurface nature of groundwater, data on groundwater status is often scarce and of poor quality, in particular in many parts of the developing world like in Africa. There is, therefore, a need to improve the capacity of monitoring groundwater bodies as a support for integrated water management of large river basins.

We address this significant knowledge gap for groundwater pollution in Africa methods for assessing groundwater pollution risk at the African scale. To do so, we compile the most recent continental-scale information on soil, land use, geology, hydrogeology and climate in a Geographical Information System at the resolution of 15 kmx15 km and the 1:60,000,000 scale. We produced a vulnerability map by using the generic DRASTIC vulnerability indicator (Ouedraogo et al., 2016). This map revealed that groundwater is highly vulnerable in Central and West Africa groundwater basins, where the water table is very low. In addition, very low vulnerability classes are found in the large sedimentary basins of Africa deserts where groundwater is situated in very deep aquifers. The generic groundwater pollution risk map is obtained by overlaying the DRASTIC vulnerability indicator with current land use. The northern, central and western parts of the African continent are dominated by high vulnerability classes and very strongly related to water table depths and development of agricultural activities. The generic methods based on the DRASTIC model are further refined using statistical models. With these approaches, monitoring data are integrated into the vulnerability mapping. Given the availability of data, we concentrate first on nitrate vulnerability mapping. To this end, groundwater nitrate contamination data are compiled (Ouedraogo and Vanclooster, 2016) and used to calibrate as well linear and nonlinear statistical models (Ouedraogo et al.2018); the latter performing much better as compared to simple linear statistical models. This study will help to raise awareness of the manager's International Basin Authorities or Transboundary Basin Organizations in Africa and in particular on transboundary groundwater pollution issues.

***Keywords***: Groundwater basins, vulnerability, DRASTIC model, statistical modelling, African scale.

## 1. Introduction

Groundwater is a crucial natural resource supporting the development of the African continent, but it is subjected to many pressures. To define sustainable water resources management plans at the continental scale, assessments of groundwater resources and associated pressures are strongly needed (Hasiniaina et al., 2010). Several studies have already been conducted to improve our understanding of African groundwater systems. Indeed, the review by Xu and Usher (2006) was one of the most important qualitative works providing insights into common issues with regards to groundwater pollution in Africa at a different local scale. Notwithstanding the compilation book and the recent progress of technology such as remote sensing, no study has intended to assess the pan-African scale groundwater vulnerability for pollution. Assessing groundwater quality at a large scale is particularly important for monitoring progress in sustainable development goals, such as the target on water issues. In this context, assessing the groundwater vulnerability for pollution is important for designing efficient regional-scale groundwater management and protection strategies.

According to Xu and Usher (2006), the degradation of groundwater is the most serious water resource problem in Africa. The two main threats are overexploitation and contamination (MacDonald et al. 2013). Nitrate is a common chemical contaminant of groundwater and the level of contamination has also increased in many African aquifers (Spalding and Exner, 1993; Puckett et al. 2011). Nitrate ingestion has been linked to methemoglobinemia, adverse reproductive outcomes, and specific cancers (Ward et al. 2005). In addition, nitrate is often a proxy of other possible pollutants of groundwater. Nitrate contamination is therefore very informative for overall groundwater quality. Nitrate contamination of groundwater is, however, a space-time variable property and the level of contamination depends on many space-time variable environmental and anthropogenic attributes.

The prediction and mapping of nitrate contamination pose, therefore, formidable technical and scientific challenges. Statistical models can be used to explain the spatial distribution of observed nitrate concentration in terms of available environmental and anthropogenic attributes, or to discriminate sources of contamination (Nolan and Hitt, 2006). For example, Bauder et al. (1993) investigated the major controlling factors for nitrate contamination of groundwater in agricultural areas using multiple linear regression analysis with land use, climate, soil characteristics, and cultivation types as independent variables. Furthermore, nonlinear random forest techniques have been applied to model nitrate in an unconsolidated aquifer in southern Spain (Rodriguez-Galiano et al. 2014). Norouz et al. (2016) used also a random forest method to determine the Malekan Aquifer vulnerability with several variables such as variables related to the DRASTIC method. Most statistical models used in such a context uses linear regression, nonlinear regression, logistic regression or Bayesian approaches to extract the variables that explain nitrate contamination (Mair and El-kadi, 2013; Mattern et al. 2012). If properly identified and validated, the statistical models can be used to predict nitrate contamination, and map possible contamination at the regional scale. Such maps are extremely valuable for designing transboundary water resources management and protection programs.

The purposes of the study were to improve our understanding of groundwater quality issues at the African scale to support the monitoring of the UN Sustainable Development Goals and in particular Goal 6 Clean water and Sanitation. In this regard, we mapped firstly the vulnerability of groundwater to pollution at the continental scale. Secondly, we tried to identify the environmental factors contributing to the nitrate contamination of groundwater. We used the power of meta-analysis, to build pan-African database groundwater pollution by nitrates. Then we relied on the meta-database to develop an exploratory linear statistical model of nitrate degradation of groundwater in terms of driving attributes available in the best available database. Finally, we used the meta-database to develop a nonlinear statistical method and outputs were compared to those derived from a statistical linear model.

The mapping of groundwater vulnerability and statistical modelling of nitrate contamination at this scale can provide guidance for the planning and implementation of groundwater monitoring programs; in particular, the design of transboundary basins water management strategies adapted to the conditions of different regions of Africa.

## 2. Data and methods

To implement the DRASTIC index methodology at the pan-African scale, a best available database of environmental factors for the continent was established at the resolution of 15 kmx15 km and the 1:60,000,000 scale. This GIS database was related to land use, soil type, hydrogeology, topography, climatology, etc (Ouedraogo et al.2016). Spatial data were processed in ArcGIS10.3$^{TM}$, QGIS™ 2.2 and Matlab™. Data came in various spatial resolutions. We retained a 15 km x 15 km resolution in this study. We considered that this resolution is a reasonable compromise between different resolutions of the different datasets, computing constraints and regional extent. We obtained the vulnerability and risk maps, after classifying and assigning relative ratings and weights, then overlaying the individual maps in a GIS.

In addition to the pan-African hydrogeological geodatabase, we have analysed the distribution of nitrate concentration at the continental scale. This analysis was based on studies reported in the literature because of the scarcity of groundwater monitoring data. Thus, 250 publications dealing with groundwater pollution were retrieved from the literature, mainly from books and the internet web of Sciences science (mainly Scopus$^{TM}$). The data were filtered using the following criteria: (i) the publication should explicitly report on nitrate concentrations in groundwater; (ii) the article should be published after 1999 (Ouedraogo and Vanclooster, 2016). The dependent variable is the natural log of sampled groundwater nitrate concentration. The average mean nitrate concentration is 54.85 mg/L, the standard deviation is 89.91 mg/L, and the median concentration is 27.58 mg /L. The log transformation reduces the influence of very high nitrate values (up to 648 mg/l) on model predictions. The nitrate studies were separated into two groups: 80 % of the dataset for building the model (training dataset) and 20 % of the data set for the validation of the model (tested dataset) (Ouedraogo et al., 2018). Furthermore, we used “Random Forest Regression (RFR)” to explain and predict groundwater nitrate contamination at the continental scale and its performance was compared with multiple linear regression (MLR) techniques. The R statistical software was used for the modelling.

## 3. Results and Discussion

### 3.1 Mapping output

We produce the first groundwater vulnerability map at the African scale by using the generic DRASTIC index indicator (see Figure 1). We classified this index into 5 classes, ranging from very low to very high index. This map reveals that groundwater is highly vulnerable in Central and West Africa groundwater basins, where the water table is very low. In addition, very low vulnerability classes are found in the large sedimentary basins of the African deserts where groundwater is situated in very deep aquifers. The transboundary basins in northern, central and western parts of the African continent are dominated by high vulnerability classes and are very strongly related to water table depths and development of agricultural activities. Also, Figure 1 gives an overlaying of Transboundary aquifers (TBAs) defined by Altchenko and Villholth (2013). This figure could be used in example for transboundary aquifers management (International Network of Basin Organizations (INBO), Global Water Partnership (GWP)) in Africa.

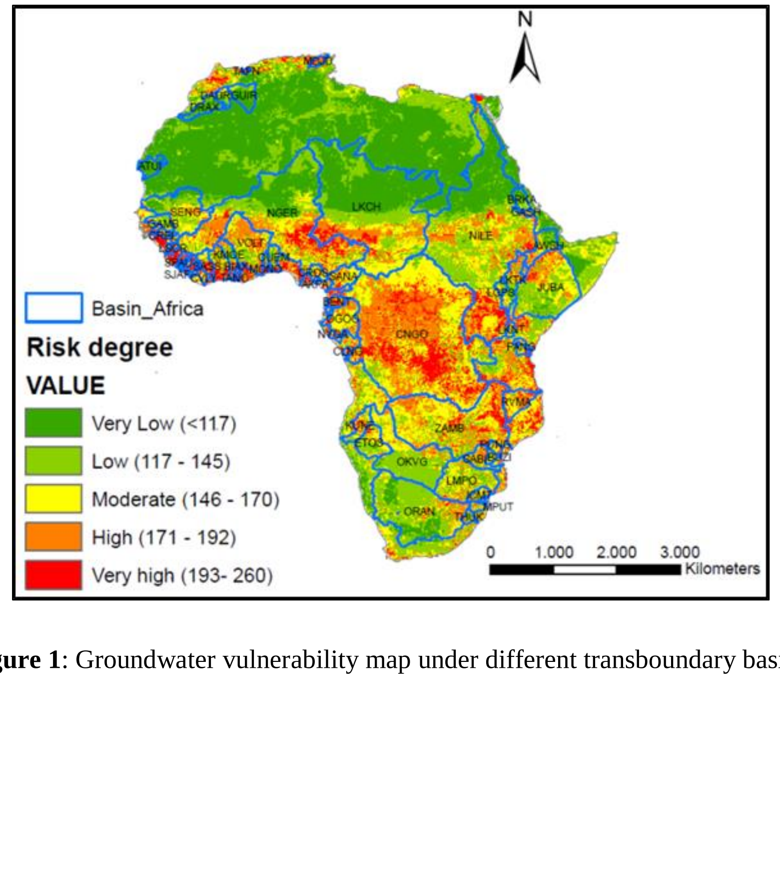


**Figure 1**: Groundwater vulnerability map under different transboundary basins.

### 3.2 Statistical modelling results

The results of statistical modelling are:

i. Using the stepwise MLR method we were able to select only those parameters which have a significant impact on the nitrate concentration values. The MLR model revealed that the four final explanatory variables yielded an $R^2$ of 0.64, indicating that 64 % of the variation of the observed logtransformed means nitrate concentration at the pan African scale is explained by the model. The four variables are: (i) depth to groundwater, (ii) recharge, (iii) aquifer type, and (iv) population density. The residual standard error is 0.95 (ln (mg/l)). A plot of predicted versus the observed log nitrate (ln NO3) concentration in Figure 2a for the training data, indicates the MLR model shows an acceptable fit with the observed data.

ii. With machine learning technique, the percentage of variance explained in the final model of random forest regression (RFR) is 97.9 % (mean of squared residuals = 0.0407 ln (mg/l)). Figure 2b shows a plot of the predicted versus the observed ln-transformed nitrate concentration values for the training data, based on only 80 % of observations. The variable importance plot is a critical output of the random forest algorithm. Figure 3 shows the ranking of the relative importance of environmental attributes on groundwater nitrate occurrence. Higher values of percent increase in mean squared error (MSE) indicate higher importance.

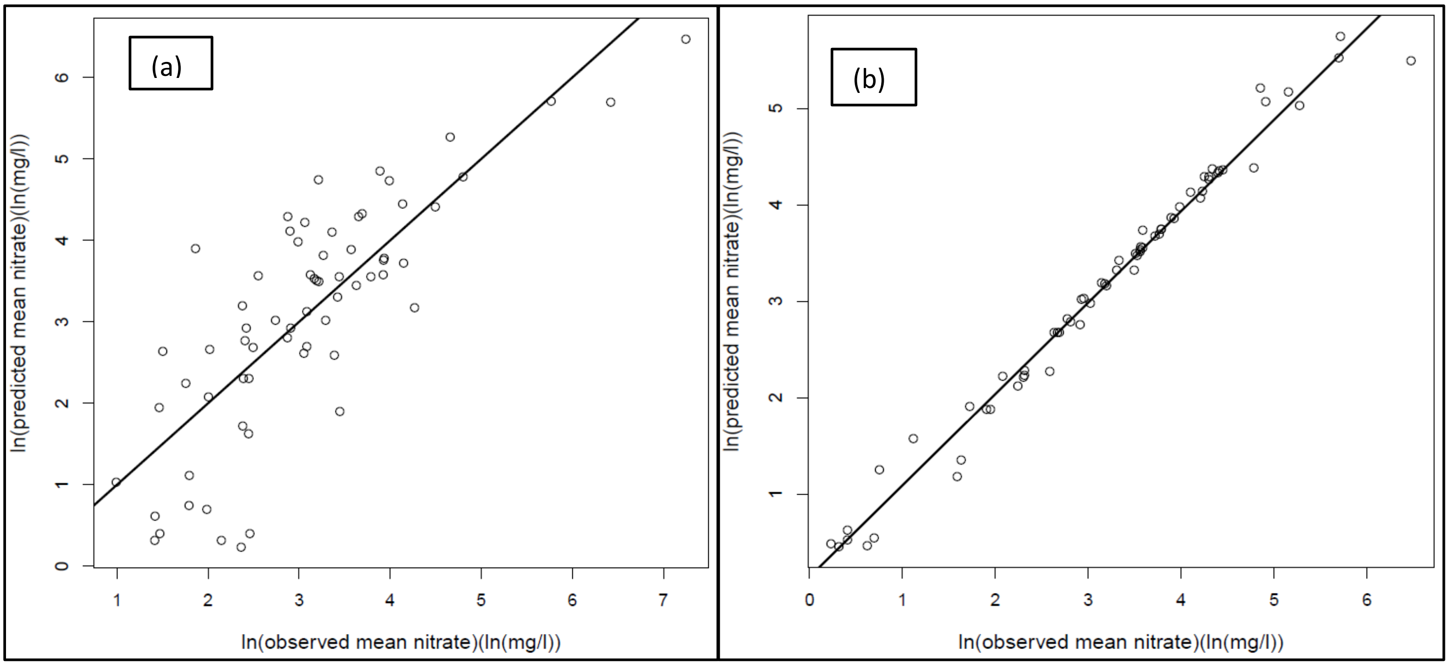


**Figure 2.** Comparison of observed and predicted log (ln) nitrate concentration on training dataset: (a) Linear regression and (b) RF regression.

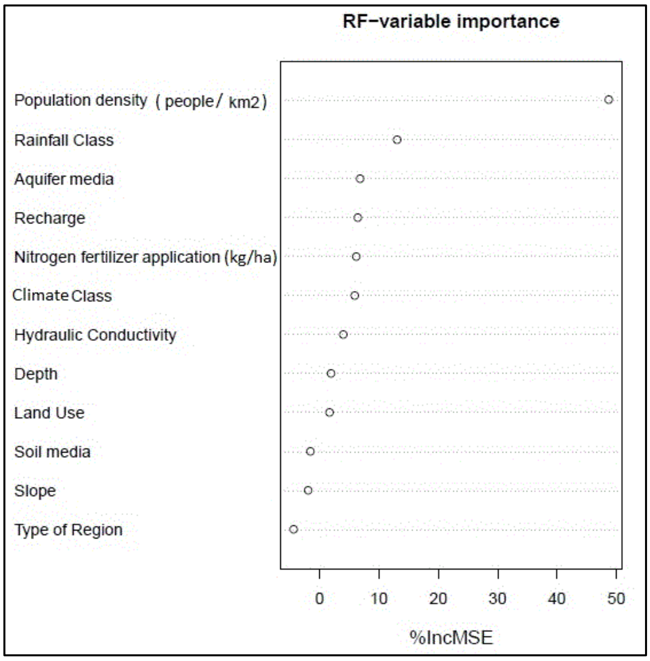


**Figure 3.** Variable importance according to the percent increase in mean squared error (%IncMSE).

The comparison between the results shows that RFR has a much higher predictive ability compared to MLR. The better performance of RFR is due to the ability to handle nonlinear relationships between nitrate pollution and explaining factors. The most important explanatory variable with the strongest influence on nitrate concentration at the pan-African scale identified in both models was the population density. This is most likely related to the lack of sanitation in major parts of the African continent and the development of small-scale farming systems in peri-urban environments. A strong influence of the population density on groundwater nitrate contamination is consistent with a previous UNEP study (e.g. Nolan et al. 2006; Mattern et al., 2009; UNEP/DEWA, 2014; Mfumu Kihumba et al. 2015; Ouedraogo et al.2016, and 2018). This study at the pan African scale could prompt national or international authorities to foster targeted local investigations because nowadays, many international agencies investigate on transboundary basins aquifers in Africa (OSS, IGAD, GWP, AMCOW, etc.).

## 4. Conclusion

The main lesson of the study is to remind is that shallow groundwater poses pollution problems in Africa. The maps that were designed in this study can increase awareness of citizens and leaders in areas where groundwater pollution is likely to be significant. In addition, they could prompt national or international authorities to foster targeted local investigations. In fact, environmental management needs to be operatively performed at regional and local scales, but investment policies can be addressed at continental (transboundary water management) or even global scales by international agencies and authorities (e.g., FAO, UNEP/UNESCO, INBO, and AMCOW). The map should serve as a general guideline for planners and decision-makers with land-use and water management development issues. Results of the statistical modelling have shown that groundwater nitrate contamination is strongly linked to population density.
This analysis represents an important step toward developing tools that will allow us to accurately predict the distribution of nitrate contamination in groundwater in the context of climate change.

**Acknowledgments:** GIS shapefiles to generate generic attributes were obtained from different sources through the World and also online. In this regard, special thanks go to Tom Gleeson, Petra Döll, Nils Moosdoorf, and P. Trambauer. This study has been supported by Université catholique de Louvain through the 'Fonds Spécial de Recherche' and IDB Merit Scholarship Program for PhD thesis.

## References


Altchenko Y. and K.G. Villholth, (2013). Transboundary aquifer mapping and management in Africa: a harmonised approach Hydrogeological J.,DOI: 10.1007/s10040-013-1002-3.

Bauder, JW., Sinclair KN., and Lund, RE., (1993). Physiographic and land use characteristics associated with nitrate nitrogen-nitrogen in Montana groundwater. Journal of Environmental Quality 22: 255-262.

BGR and UNESCO, (2012). River and Groundwater Basins of the World. (www.whymap.org).

Hasiniaina, F., Zhou, J., Guoyi, Z.L., (2010). Regional assessment of groundwater vulnerability in tamtsag basin, Mongolia using drastic model. Journal of American Science 6 (11), 65–78.

MacDonald, A., M., R. Taylor, G. and H. Bonsor, C. (2013). (eds.) Groundwater in Africa – is there sufficient water to support the intensification of agriculture from “Land Grabs”." Hand book of land and water grabs in Africa. pp 376- 383.

Mair, A. and A. I. El-Kadi, (2013). Logistic regression modeling to assess groundwater vulnerability to contamination in Hawaii, USA. J Contam Hydrol. 153: 1-23.

Mattern S. , Fasbender, D., and Vanclooster, M. (2009). Discriminating sources of nitrate pollution in a sandy aquifer. Journal of hydrology 376:275-284

Mattern, S., et al.,(2012). Bayesian Data Fusion (BDF) of monitoring data with a statistical groundwater contamination model to map groundwater quality at the regional scale. Journal of Water Resource and Protection Vol.4 (No.11).

Mfumu, Kihumbo, A., Ndembo, J., and Vanclooster, M.. (2015). Modelling nitrate pollution pressure using a multivariate statistical approach: the case of Kinshasa groundwater body, Democratic Republic of Congo. Hydrogeology Journal. DOI 10.1007/s10040-015-1337-z

Nolan, BT., Hitt, KJ., (2006). Vulnerability of shallow groundwater and drinking-water wells to nitrate in the United States. Environmental Science & Technology 40: 7834-7840.

Norouz, H., Asghari moghaddam Negar, A.M., Attaallah, N. (2016). Determining vulnerable areas of malekan plain aquifer for nitrate, using random forest method. JOURNAL OF ENVIRONMENTAL STUDIES WINTER 2016 , Volume 41 , Number 4 (76); Page(s) 923 To 942. http://www.sid.ir/En/Journal/ViewPaper.aspx?ID=550917 (Accessed online August, 2rd 2018).

Ouedraogo, I., and Vanclooster, M. (2016). A meta-analysis and statistical modelling of nitrates in groundwater at the African scale. In: Hydrology and Earth System Sciences, Vol. 20, no.6, p. 2353-2381. DOI: 10.5194/hess-20-2353-2016.

Ouedraogo, I., Defourny, P., Vanclooster, M. (2016). Mapping the groundwater Vulnerability for pollution at the pan-African scale. Science of the Total Environment, Vol. 544, p. 939-953.DOI: 10.1016/j.scitotenv.2015.11.135.

Ouedraogo, I., Defourny, P., Vanclooster, M. (2018). Application of random forest regression and comparison of its performance to multiple linear regression in modeling groundwater nitrate concentration at the African continent scale. Hydrogeology Journal. https://doi.org/10.1007/s10040-018-1900-5

Rodriguez-Galiano, V., Mendes, M.P., Garcia-Soldado, M.J., Chica-Olmo, M., and 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). Sci Total Environ 476-477: 189-206. https://doi.org/10.1016/j.scitotenv.2014.01.001.

Ward, M. H., et al.,(2005). Workgroup Report: Drinking-Water Nitrate and Health—Recent Findings and Research Needs. Environmental Health Perspectives 113(11): 1607-1614.

Xu, Y., and Usher, B. (2006). Groundwater pollutionj in Africa, Taylor&Francis/Balkema, The Netherlands. 353pp.