Prediction model for cyanide environmental pollution in artisanal gold mining area by using logistic regression

Lovasoa Christine RAZANAMAHANDRY[a*], Harinaivo Anderson ANDRIANISA[a], Hela KAROUI[a], Joel PODGORSKI[b], and Hamma YACOUBA[c]

[a]International Institute for Water and Environmental Engineering (2iE), Department of Water and Sanitary Engineering, Laboratory of Water, Decontamination, Ecosystem and Health (LEDES), 01 PO Box 594, 01, Ouagadougou, Burkina Faso

[b]Swiss Federal Institute of Aquatic Science and Technology (EAWAG), Department of Water Resources and Drinking Water.Postfach 611 8600 Dübendorf. Switzerland.

[c]International Institute for Water and Environmental Engineering (2iE), Department of Civil and Hydraulic Engineering, Laboratory of Water Resources and Hydrology (LEAH), 01 PO Box 594, 01, Ouagadougou, Burkina Faso

*Corresponding author: Lovasoa Christine RAZANAMAHANDRY (c.razanamahandry@2ie-edu.org or tantely1989@gmail.com) Phone: + 226 66 69 16 57

Abstract

It has been reported that persistent cyanide pollution occurs in Artisanal small-scale gold mining (ASGM)-affected catchment areas in Burkina Faso. In the present study, the logistic regression method was employed to identify the factors that influence the spatial distribution of cyanide pollution as well as to predict the cyanide pollution map risk at catchment level. Soil samples were collected from two ASGM sites in the northern Zougnazagmiline(“North”) site and southern Galgouli(“South”) site parts of Burkina Faso, covering areas of 22 km² and 20 km², respectively. Free cyanide (FCN) concentration in each sample was measured. It was shown that the spatial distribution of cyanide was solely controlled by the soil type in Zougnazagmiline and both the soil type and conductivity in Galgouli. On the other hand, the cyanidation zones within the two catchments were the places where the highest risk of cyanide pollution occurs, with probabilities of 0.8 and 1 in Zougnazagmiline and Galgouli, respectively. More than 20% of the settled area in the Zougnazagmiline and 5% of that in Galgouli were exposed to cyanide pollution. Logistic regression was able to reliably predict cyanide contamination in areas affected by ASGM. The model could be useful for decision-makers to plan ASGM-site decontamination.

Key Words: Hazardous chemicals, Catchment area, diffuse pollution, Soil contamination, Risk assessment, Burkina Faso

1 Introduction

Artisanal and small-scale gold mining (ASGM) has been widespread throughout the world for over 2000 years (Hilson, 2002a; Weng et al., 2014), and more developed since the mid-1980s in West Africa, including Burkina Faso(Butaré & Keita, 2009; Grätz, 2009). In 2004, between 10% and15% of the gold mined in the world has been provided from ASGM (Adler et al., 2013; Grimaldi et al., 2015; Street et al., 2013; Telmer & Veiga, 2008). In Burkina Faso, small-scale artisanal miners produced approximately 12 tons of gold compared to an output of 14 tons from large-scale mines between 1986 and 1997(Gajigo et al., 2012; Guèye, 2001).

The ASGM sector provides a livelihood for millions of people throughout the world (Siegel & Veiga, 2009; Weng et al., 2014). In the case of sub-Saharan Africa, at least two million people are directly employed in ASGM, and an additional 10 million more people depend on the sector for their survival (Adler et al., 2013; Chupezi & et al., 2009; Hilson, 2009; Janneh & Ping, 2011; Schure et al., 2011; Weng et al., 2014). Nevertheless, several negative impacts are associated with ASGM such as an increase of infectious diseases, violence and crime, child labour and a lacking emphasis on education, loss of biodiversity and exposure of miners to strong hazardous chemicals (Adler et al., 2013). In the natural environment, ASGM induces changes to land use and landscapes, instability of the ground and landslide sand water, air and soil pollution (Adler et al., 2013; Guimaraes et al., 2011).Environmental pollution is primarily caused by the use of toxic chemicals products, including cyanide, which is widely used in post-processing to extract residual gold after mercury processing. The residual material is rich in cyanide ions, which can ultimately leach into the environment without treatment or control.(Adler et al., 2013; Bernstein, 2000; Sampat, 2003; Veiga et al.,2014; Velásquez-lópez et al., 2011).

Previous studies have shown that the main environmental parameters that control the distribution of pollutants are land cover, topography, geology, rainfall, temperature, soil type, and distance from the pollution source (Kheir et al.,2014;Venkataraman and Uddameri,2012). In addition, several chemical parameters, such as pH, soil conductivity and organic matter content, can also help explain the spatial distributions of pollutants (Kheir et al., 2014).

Cyanide is present in the environment as FCN, weak acid dissociable (WAD) and strong acid dissociable (SAD). The most toxic form is FCN. The subsurface behaviour of cyanide compounds in soil is governed by chemical and biological processes (Kjeldsen, 1999). The relevant chemical processes are adsorption, sorption, volatilization, complexation, sulfidation and dissolution-precipitation type reactions, which are sensitive to pH, temperature and redox conditions (Donato et al., 2007; Guo et al., 2014; Johnson, 2014; Klenk etal., 1996; Richards et al., 2000). On the other hand, microbial activity and plant uptake can affect the behaviour of cyanide in the environment (Kumar et al., 2016).

We have investigated the use, fate and behaviour of cyanide in two catchments areas affected by ASGM in Burkina Faso. It was found that up to 20 kg/week of cyanide could be illegally used in one catchment area for gold processing. Cyanide-containing leachate is then directly released into the environment without any treatment or control. FCN accumulates around the cyanide-processing zones, whereas some is also progressively transported to the catchment outlet through surface runoff and infiltration, which pollutes surface water, groundwater and soil within the catchment (Razanamahandry et al., 2016). However, the processes that control the transport of FCN are not sufficiently understood to allow the identification and targeting of pollution risk zones for the implementation of a remediation plan. In this regard, the Geographic Information Systems (GIS) could be useful in assessing cyanide pollution risk.

Related examples include modelling the adaptation of a mine-impacted community to landmine contamination (Benini et al. 2002; Schultz et al. 2016) and risk mapping of landmine hazard and its spatial distribution (Alegria et al. 2011; Chamberlayne, 2002; Lacroix et al. 2013; Schultz et al., 2016). In addition, logistic regression (LR) is one of the most important statistical techniques developed for analysing and classifying categorical variables (Agresti, 2002; Hair et al.,1998; Mokhtari, 2014; Pohar et al., 2004) A GIS LR approach has been applied to landslide susceptibility mapping (Guns & Vanacker, 2012; Schultz et al., 2016; Van Den Eeckhaut et al., 2006; Wang et al.,2015), disease mapping (Craig et al., 2007; Ekpo et al.,2008; Goovaerts et al., 2015; Schultz et al., 2016), vulnerability mapping (Ettinger et al., 2015; Schultz et al., 2016), wildfire distribution (Rodrigues et al., 2014; Schultz et al., 2016), crime mapping (Capla et al., 2011; Caplan, 2011; Schultz et al., 2016), post-fire soil erosion (Notario et al., 2014) and pollutants mapping (Venkataraman and Uddameri, 2012).The environmental and chemical parameters that influence the subsurface spatial distribution of FCN appear to have not yet been investigated using LR.

The aim of the present study is to(i) create a cyanide pollution risk map for ASGM sites by applying the LR method, (ii) identify possible risk factors that may explain the spatial distribution of cyanide contaminated areas, and (iii) identify areas of high risk so that appropriate remediation actions can be taken.

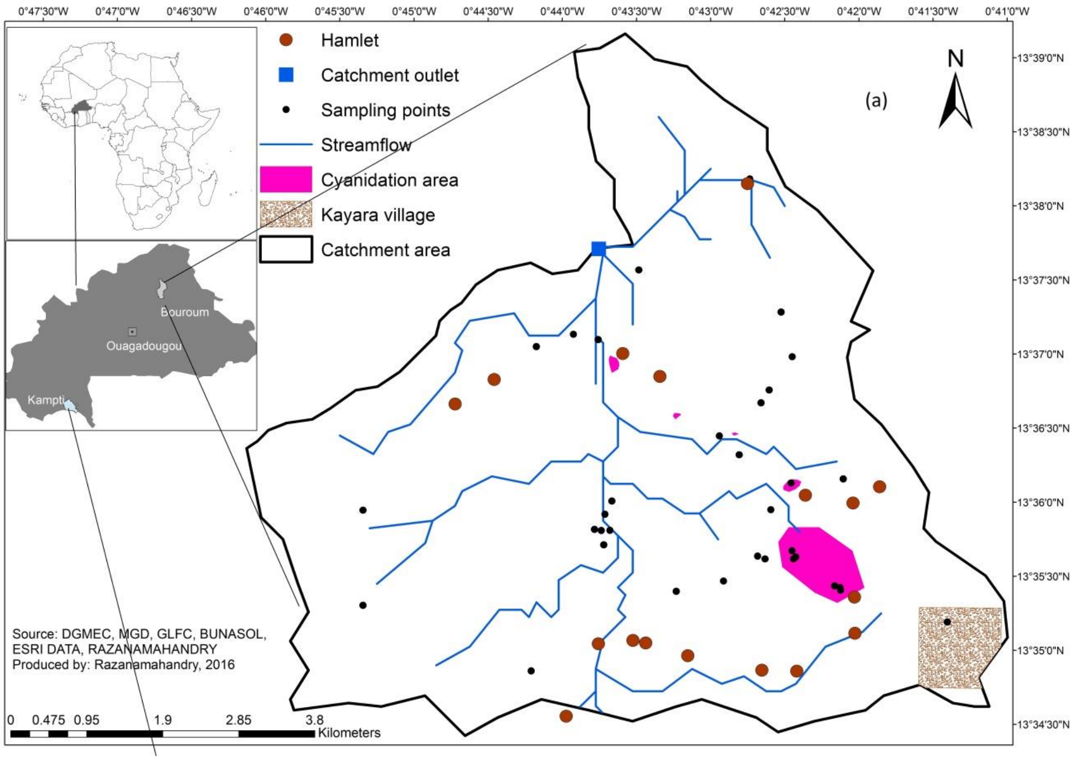
We first developed a conceptual model of the spatial distribution of cyanide pollution risk. Factors relevant to cyanide pollution transport were then identified and probability maps for cyanide pollution risk were created and analysed.

2 Materials and Methods

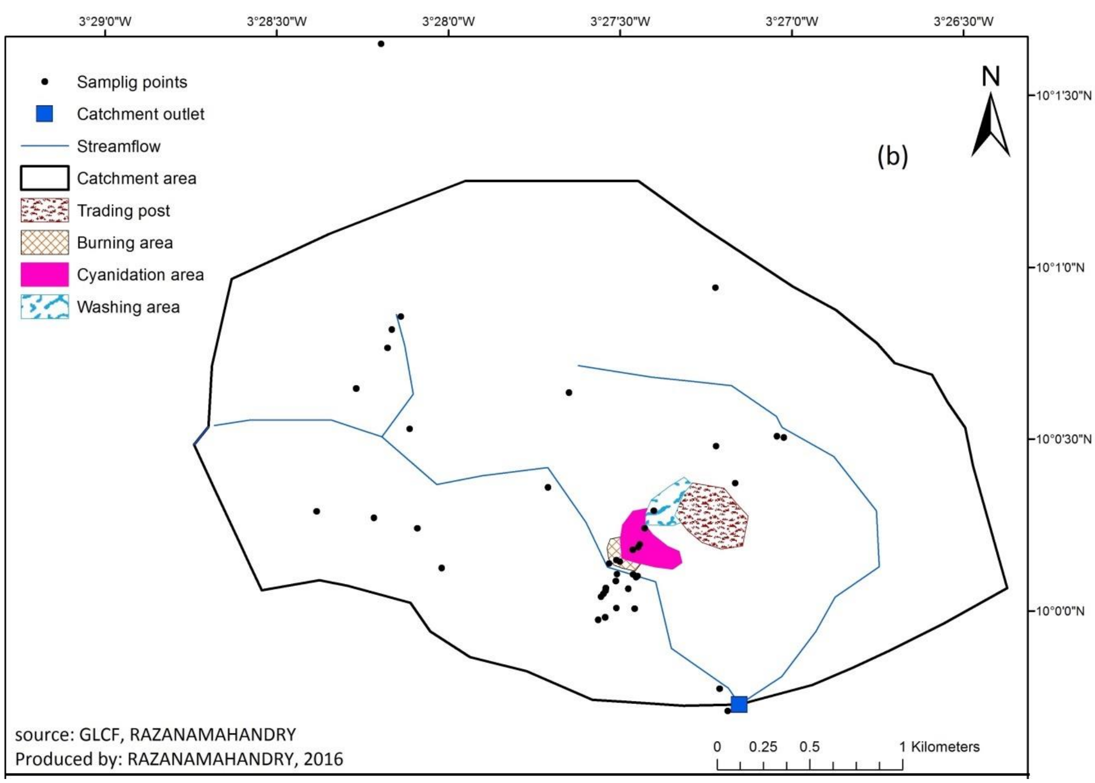
2.1 Samples and study areas

Two ASGM sites were selected for modelling cyanide contamination using LR based on their climatic and environmental conditions and mining activity. The first site, Zougnazagmiline, is located in the northern part of Burkina Faso (Fig.1a.). It is in the arid Sahelian climate zone with an average annual rainfall of less than 600 mm and contains primarily lixisols (FAO, 1998, 2001). All mining activities are illegal, and cyanidation takes place at several locations. The second area, Galgouli, is located in southern part of Burkina Faso (Fig.2a). It is a forest zone characterised by the Soudanese climate with an annual rainfall of 1200 mm and contains primarily arenosols (FAO, 1998, 2001). The site is controlled by a single private operator who holds a mining permit. Although cyanidation processing is illegal in Burkina Faso, it is still conducted in the area. However, as opposed to the Zougnazagmiline site, cyanidation takes place in only one location.

Two samples were collected from each site, in March 2015 and in April 2016.More than thirty points covering the cyanidation zones, catchment areas boundaries and outlets and the mining villages were selected for soil sampling (Fig. 1)., Five soil samples were collected at each sampling point in 20 cm intervals from the surface to 1 m depth. Immediately following collection, samples were wrapped in black plastic bags and kept in a cooler until arrival in the laboratory, where they were refrigerated until analysis, which was usually performed after 24 hours.



Hamlet
Catchment outlet
Sampling points
Streamflow
Cyanidation area
Kayara village
Catchment area
(a)
N
Bouroum
Ouagadougou
Kampti
Source: DGMEC, MGD, GLFC, BUNASOL, ESRI DATA, RAZANAMAHANDRY
Produced by: Razanamahandry, 2016
0 0.475 0.95 1.9 2.85 3.8
Kilometers
0°47'30"W 0°47'0"W 0°46'30"W 0°46'0"W 0°45'30"W 0°45'0"W 0°44'30"W 0°44'0"W 0°43'30"W 0°43'0"W 0°42'30"W 0°42'0"W 0°41'30"W 0°41'0"W
13°39'0"N 13°38'30"N 13°38'0"N 13°37'30"N 13°37'0"N 13°36'30"N 13°36'0"N 13°35'30"N 13°35'0"N 13°34'30"N



Samplig points
Catchment outlet
Streamflow
Catchment area
Trading post
Burning area
Cyanidation area
Washing area
(b)
N
3°29'0"W 3°28'30"W 3°28'0"W 3°27'30"W 3°27'0"W 3°26'30"W
10°1'30"N 10°1'0"N 10°0'30"N 10°0'0"N
source: GLCF, RAZANAMAHANDRY
Produced by: RAZANAMAHANDRY, 2016
0 0.25 0.5 1 Kilometers

Figure 1: Site locations and sampling points: (a) Zougnazagmiline, (b) Galgouli

2.2 Chemical reagents and analytical methods

The chemical reagents used, FCN extraction protocol and FCN analytical methods are described by Razanamahandry et al. (2016).

2.2 Logistic regression

2.2.1 Principle

Logistic regression (LR) isused to explain an observed or dependant variable through one or more independent predictor, or explanatory variables. In the present study, FCN concentrations represent the observed variable. Dependant and explanatory variables could be quantitative data. LR seeks and describes a relationship between the dependent variable and the explanatory variable (Shlutz et al., 2016) as shown in the equation (1) below:

$$p = \frac{1}{1 + e^{-y}} \qquad (1)$$

Where:

$p$denotes the probability of occurrence of an event, which is cyanide contamination in this case of this study.

$y$is a linearised regression equation (2) below:

$$y = \alpha + \beta_1 X_1 + \beta_2 X_2 + \cdots + \beta_n X_n \qquad (2)$$

Where:

α : theintercept

β_1, β_2,…, β_n : the coefficients of the n explanatory variables estimated by maximum likelihood (Real et al., 2006; Schultz et al., 2016)

X_1, X_2,…, X_n: the explanatory variables

The logit form of the model is:

$$logit\ (p) = log_e[\frac{p}{1-p}] \qquad (3)$$

where $logit\ (p)$ denotes the log (to base e) (Guns & Vanacker, 2012; Schultz et al., 2016). Equation (1) gives the probability of occurrence. A hazard map can then be established by considering the cyanide guide value of 0.5 mg FCN kg^{-1}applied by World Health Organisation for agricultural soil in Burkina Faso. Values greater than0.5 mg kg^{-1}pose a health risk. In this work, LR was used for finding independent variables related to cyanidation pollution and to produce a cyanide contamination hazard map.

The Receiver Operating Characteristics (ROC) curve assesses the predictive capability of a model by considering its trade-off between rates of true positive and true negative predictions (Van Den Eeckhaut et al., 2006; Wang et al., 2011; Schultz et al., 2016). A probability cut-off between high and low cases or events is ultimately selected to be used with a model to maximize its accuracy.

2.2.3 Model validation

Models were evaluated based on a number of statistical measurements, for example, the statistical coefficient as the Akaike Information Criterion (AIC) and the negative of twice the likelihood (-2 log L or LogLik-ratio), which is lower for a better fitting model (Allison, 2011). AIC is a relative goodness-of-fit statistic for comparing logistic regression models (Saefuddin et al., 2012). Lower AIC and - 2 log L values generally correspond to a better fitted model (Allison, 2011). Since AIC penalizes a model for using more parameters, minimizing the AIC optimizes the trade-off between goodness-of-fit and the number of parameters (Mcnew et al., 2013;Schultz et al., 2016). The Wald value of a variable coefficient is used to calculate its p-value, or significance, which should be equal to or less than 0.05. The Area Under the ROC Curve (AUC) typically ranges in value between 0.5 and 1, with values higher than 0.7 generally indicating stronger associations between the predicted and observed values (Van Den Eeckhaut et al., 2006). An ROC curve is a technique for visualizing, organizing and selecting classifiers based on their performance (Fawcett, 2003). It has been extended for use in visualizing and analyzing the behavior of diagnostic systems (Fawcett, 2003; Swets, 1988). ROC curves have long been used in signal detection theory to depict the trade-off between true positive rates (y axis called sensitivity axis) and false alarm rates (x axis called specificity) of classifiers (Egan, 1975; Fawcett, 2003; Swets et al., 2000a). The ROC curve of a perfectly accurate logistic regression model would run vertically from (0,0) to (0,1) (y axis) and then horizontally to (1,1) (x axis) (Brenning, 2005; Van Den Eeckhaut et al., 2006;Schultz et al., 2016).The diagonal line y = x represents the strategy of randomly guessing a class in which it can be expected to get half the positives and half the negatives correct (Fawcett, 2003). In order to get away from this diagonal into the upper triangular region, the classifier must exploit some information in the data (Fawcett, 2003). Any classifier that appears in the lower right triangle performs worse than random guessing (Fawcett, 2003).

2.2.4 LR tool

The LR model was created and downloaded from the freely available "Groundwater Assessment Platform" web site, provided by the Swiss Federal Institute of Aquatic Science and Technology (EAWAG)[Groundwater Assessment Platform (GAP), 2015]and funded by the Swiss Agency for Development and Cooperation (SDC).

2.2.5 Methodological approach

The procedure for modelling cyanide pollution was done in three steps as shown in Figure 2. The first step was to apply LR with all of the explanatory variables, which were evaluated according to their p-values. Variables with p-values less than or equal to 0.05 were retained. In the second step, the same explanatory variables from the first step were used in a stepwise LR (SLR), whereby variables were discarded individually in consecutive steps according to their effect on the AIC. The produced Model 2, which was again evaluated based on the selected explanatory variable coefficients (β_i). Explanatory variables with β_i lower than the respective Wald value were retained for use in producing the final model, Model 3.

A map of cyanide pollution hazard was created for each model. This map has a raster surface with a cell size X = 0.000171 and Y = 0.000171 and an angular units degree that contains continuous probability values ranging from 0 to 1,obtained by interpolating on a grid the predicted coefficient value points under ArcGIS software 10.1 version. The Inverse Weighted Distance (IDW) method was used for interpolation. This method determines the grid's cell values by using a linearly weighted combination of a set of the evaluation points (Schlutz et al., 2016). The cyanide pollution risk maps have been classified into five risk levels (very low, low, moderate, high, very high), based on the following cut-off values: 0.20, 0.40, 0.60, 0.80, and 1.

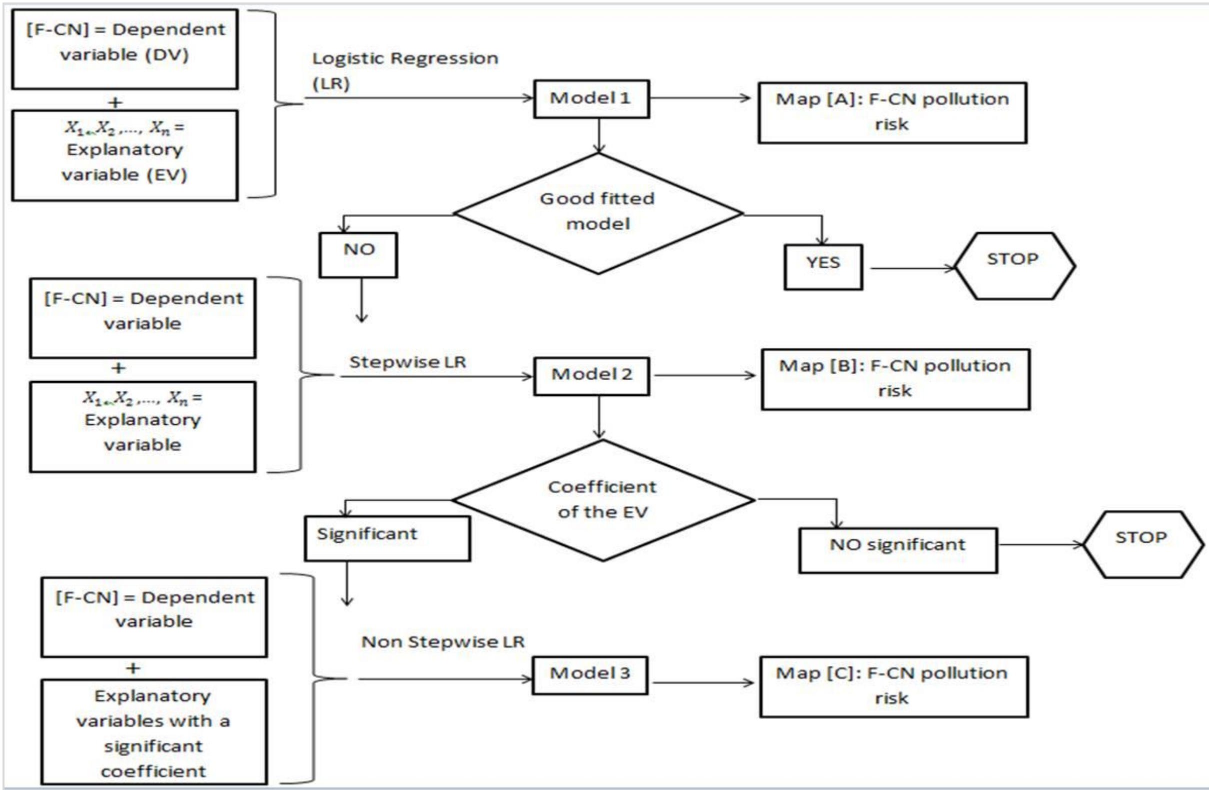


Figure 2: Methodological approach for building the model of cyanide pollution

3 Data

The model input data comprised the dependent variable and the explanatory variables as shown in Table 1. Explanatory variables were chosen for use in the model based on their relationship to the spatial distribution of cyanide pollution according to previous studies (Kheir et al.,2014;Venkataraman and Uddameri,2012). The FCN concentration data used are the mean FCN concentrations obtained from each soil sample during the 2015 and 2016 sampling campaigns. The FCN concentration data were uploaded in .CSV format, whereas. All explanatory variables have a raster image in .tif format.

The boundaries of the two catchment areas were also uploaded for delineating the predicted map of the model.

Table 1: Model data

Data	**Type**	**Variable**	**Coverage**	**Source**	**Cell size(X,Y),angular unit degree**
Geology	Continuous	Independent	2013	DGMEC [a]	0.022689,0.022689
Rainfall (mm)	Continuous	Independent	1970 - 2012	MGD[b]	0.024955,0.024955
Temperature (°C)	Continuous	Independent	1970 - 2012	MGD	0.024955,0.024955
Topographic elevation (m)	Continuous	Independent	2008	GLCF[c]	0.000833,0.000833
Land use	Continuous	Independent	2010	BUNASOL[d]	0.000280,0.000280
Soil type	Continuous	Independent	2010	BUNASOL	0.000289,0.000289
Soil pH	Categorical	Independent	2015-2016	Present study	0.00022,0.00022
Soil Conductivity (mS cm^{-1})	Categorical	Independent	2015-2016	Present study	9e-005,9e-005
Distance to cyanidation Ponds (m)	Categorical	Independent	2015-2016	Present study	8.6e-005,7.99e-005
F-CN Concentration (mg L^{-1})	Categorical	Dependent	2015-2016	Present study	7.99e-005,7.99e-0.005

[a]: Direction Général des Mines, Energie et Carrières au Burkina Faso www.mines.gov.bf

[b]: Direction Générale de la Météorologie au Burkina Faso www.meteoburkina.bf

[c]: Global Land Cover Facility glcfapp.glcf.umd.edu

[d]: Bureau National de Sol au Burkina Faso www.erails.net/BF/bunasols

4 Results and discussion

4.1 Factors influencing the spatial distribution of cyanide

Table 2 summarizes the LR model of each area.

In Zougnazagmiline, the three models have an AUC around 0.8 and p-values of 0.42, 0.05 and 0.01 for Model 1, Model 2 and Model 3, respectively. The Loglik-ratio of Model 3 was lower than that of Model 1 and Model 2 (-18.96 < 0.00). All models have good AUC values greater than 0.7. Only the p-value of Model 2 and Model 3 are equal to or less than 0.05.

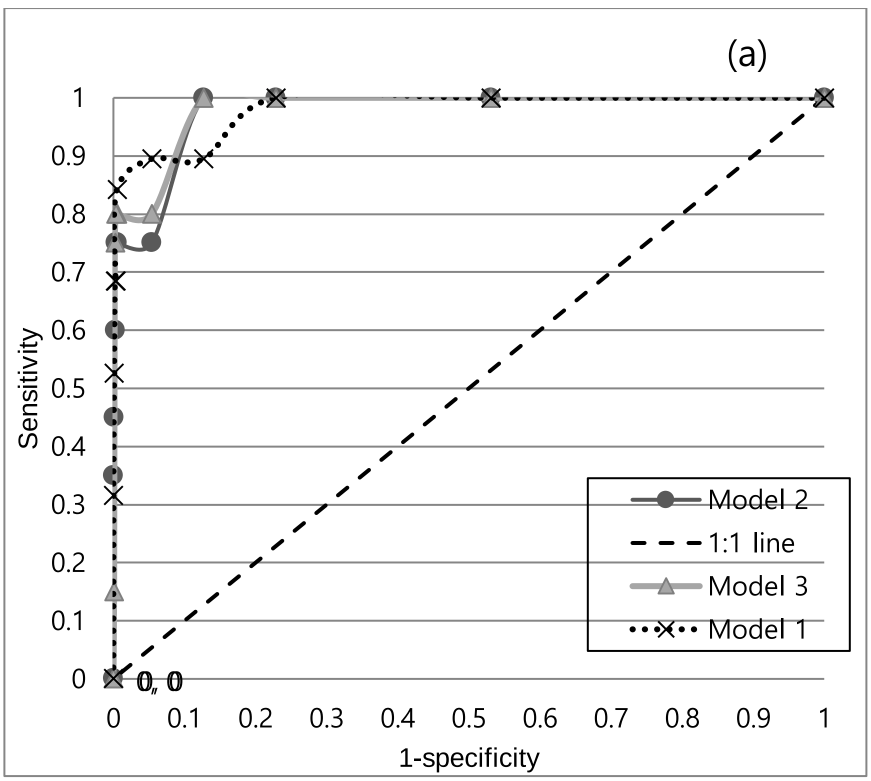
In Galgouli, the AUC of Model 1, Model 2 and Model 3 are 0.76, 0.74 and 0.69, respectively, whereas their respective p-values are 0.61, 0.05 and 0.02. In regard to the AUC value, good model fit has been established except for Model 3 in Galgouli., which indicates that the explanatory variables are related to the spatial distribution of cyanide contamination [*Lin et al.*, 2011].

Table 2: The p-value and the AUC values for each model

Site	Model name	Loglik-ratio	Deviance	X^2	Explanatory variable number	p-Value	AUC
Zougnazagmiline	**Model 1**	0.00	32.21	9.18	9	0.42	**0.79**
	Model 2	0.00	34.85	9.40	4	**0.05**	**0.78**
	Model 3	-18.96	37.92	6.33	1	**0.01**	**0.75**
Galgouli	**Model 1**	0.00	46.63	7.20	9	0.61	**0.76**
	Model 2	0.00	50.26	3.58	2	**0.05**	**0.74**
	Model 3	0.00	52.25	1.59	1	**0.02**	**0.69**

Figure 3 shows the ROC curves for the three models at each site. In terms of the p-value in both sites, the level of confidence increases from Model 1 to Models 2 and 3 (from 58 and 31 % to more than 95 %). Model 1 of both sites contains more explanatory variables than the other models. The p-values of the models improved as insignificant explanatory variables number were removed. In both sites, Model 1 has a p-value greater than 0.05, althoughModels2 and 3 have p-values less than 0.05. Model 1 is not significant and was therefore removed from consideration.

The Log-likelihood ratios (Loglik-ratio) for all models are almost the same, except for Model 3at Zougnazagmiline site, which was lower.



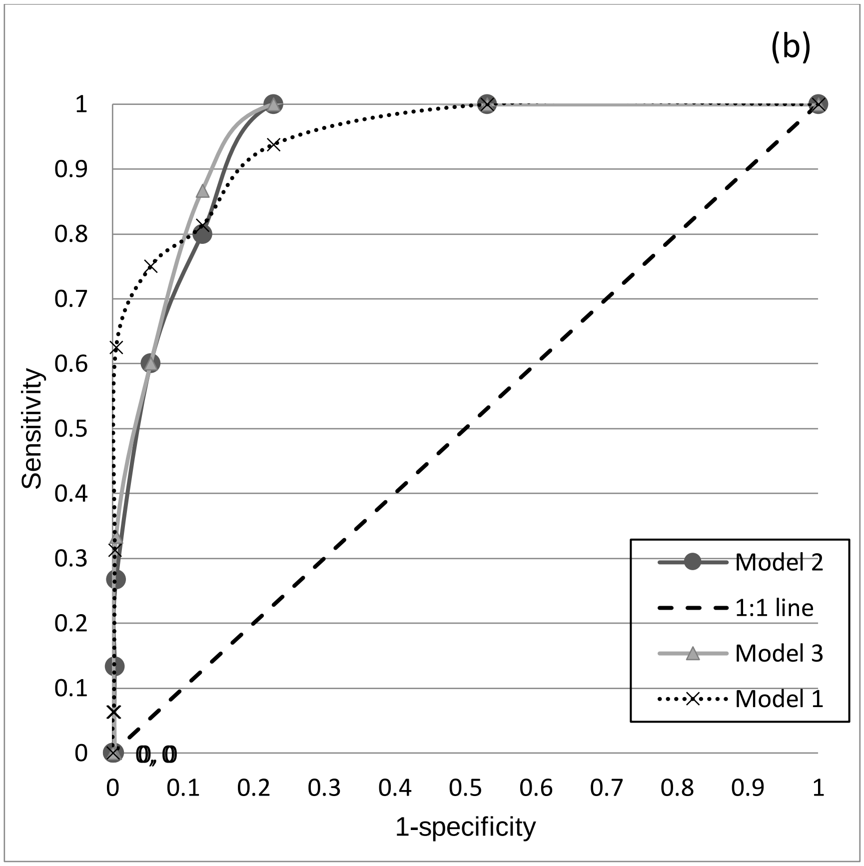


Figure 3: ROC curves for the (a) Zougnazagmilinesite and (b) Galgouli site

Table 3 shows the different explanatory variables and variable coefficient for each model. For the Zougnazagmiline site, the significant explanatory variable is the “soil type”. However, only the “soil type”and the“soil conductivity”are significant for Galgouli site. The explanatory variables that are best correlated (p-value = 0.05) with the spatial distribution of the cyanide are “soil type” and the “soil conductivity” for Model 2 in Zougnazagmiline and in Galgouli, respectively. Venkataraman and Uddameri (2012)found similar results in modelling arsenic and nitrate pollutants in drinking water with a multinomial logistic regression, whereby soil and aquifer properties were significant.

The soil types at the northern and southern parts of Burkina Faso are dominated by lixisols and arenosols, respectively (Pallo and Sawadogo, 2011). FAO (1998,2001) defines lixisols as soils with an argic layer from the soil surface to 100-200 cm depth. According to ATSDR(1997), the clay that is abundantly present in the argic layer inhibits FCN complexation with other metals. In addition, lixisols have a low level of plant nutrients and a high erodibility (FAO, 1998, 2001) and is a major inhibitor for plant growth and biomass production (Ehlers et al., 2010; Vitousek, 1984). Consequently, the biological activity of microorganisms and plants is not sufficient to degrade or complex FCN.FCN is therefore easily released, resulting in a positive correlation between soil type and FCN at Zougnazagmiline.

On the other hand, arenosols are characterised by sandy loam soils with low conductivity[FAO, 1998, 2001], which means that several cations and minerals are available in a low quantity [FAO,1999].Most of the FCN is released because the CN^- ligand does not form a complex compounds with the cations[Dai et al., 2012; Ghosh et al., 2006; Kjeldsen, 1999; Theis and West, 1986]. For this reason, soil conductivity mainly influences the distribution of FCN in Galgouli.

Table 3: Explanatory variables coefficients for Zougnazagmiline and Galgouli

Model name	**Variable**	**Coefficient**	**Std error**	**Wald**	**Significance**	**Odds**	**Lower**	**Upper**
ZOUGNAZAGMILINE SITE								
Model 1								
	Intercept	0.0249	0.1631	0.1527	0.8786	1.0252	0.7447	1.4113
	Geology	1.6255	2.2138	0.7343	0.4628	5.0808	0.0663	389.241
	Rainfall (mm)	-0.0476	0.1583	-0.3005	0.7638	0.9535	0.6991	1.3005
	Temperature (°C)	0.7543	4.9391	0.1527	0.8786	2.1261	0.0001	34010.2
	Soil pH	2.182	5.3559	0.4074	0.6837	8.8642	0.0002	320932
	Distance to cyanidation ponds (m)	0.0006	0.0008	0.7225	0.47	1.0006	0.999	1.0022

	Topographic elevation (m)	-0.1166	0.8138	-0.1432	0.8861	0.89	0.1806	4.3862
	Soil conductivity (mS cm^{-1})	0.0026	0.0089	0.2883	0.7731	1.0026	0.9853	1.0202
	Soil type	0.0093	0.0111	0.8351	0.4036	1.0093	0.9876	1.0315
	Land use	0.064	0.6424	0.0996	0.9207	1.0661	0.3027	3.7549
Model 2								
	Intercept	-3.2791	2.6349	-1.2445	0.2133	0.0377	0.0002	6.5869
	Distance to cyanidation ponds (m)	1.008	1.1631	0.8666	0.3861	2.7402	0.2804	26.7788
	Soil conductivity	-0.0054	0.0052	-1.0223	0.3066	0.9947	0.9845	1.0049
	Soil type	0.0139	0.0071	1.9462	**0.0516**	1.0139	0.9999	1.0282
	Land use	-0.0818	0.486	-0.1683	0.8663	0.9215	0.3555	2.3887
Model 3								
	Intercept	-1.2978	0.8099	-1.6025	0.109043	0.2731	0.0559	1.3357
	Soil type	0.0092	**0.0039**	2.3839	**0.017131**	1.0093	1.0016	1.017
GALGOULI SITE								
Model 1								
	Intercept	0.02	0.978	0.0205	0.9837	1.0202	0.1501	6.9358
	Geology	0.02	0.978	0.0205	0.9837	1.0202	0.1501	6.9358
	Rainfall (mm)	-0.0167	0.7233	-0.0231	0.9816	0.9835	0.2383	4.0588
	Temperature (°C)	0.5488	26.8382	0.0204	0.9837	1.7312	0	1E+23
	Soil pH	-0.1987	0.2716	-0.7316	0.4644	0.8198	0.4814	1.3961
	Distance to cyanidation ponds	0.0004	0.0009	0.4595	0.6459	1.0004	0.9987	1.0021

	(m)							
	Topographic elevation (m)	0.0029	0.0206	0.1387	0.8897	1.0029	0.9632	1.0442
	Soil conductivity (mS cm-1)	0.0005	0.0003	1.5717	0.116	1.0005	0.9999	1.0011
	Soil type	0.0729	0.4133	0.1765	0.8599	1.0757	0.4785	2.4183
	Land use	0.4548	0.7563	0.6014	0.5476	1.5759	0.3579	6.9393
Model 2								
	Intercept	-1.3562	0.5675	-2.3899	0.0169	0.2576	0.0847	0.7835
	Soil type	0.3126	0.2282	1.37	0.0707	1.367	0.8741	2.138
	Soil conductivity (mS cm-1)	0.0004	0.0003	1.4247	**0.054**	1.0004	0.9999	1.0009
Model 3								
	Intercept	-0.8933	0.4342	-2.0575	0.0396	0.4093	0.1748	0.9585
	Soil conductivity (mS cm-1)	0.0003	0.0003	1.22	**0.023**	1.0003	0.9998	1.0008

4.2 Cyanide contamination hazard

4.2.1 Cyanide hazard map

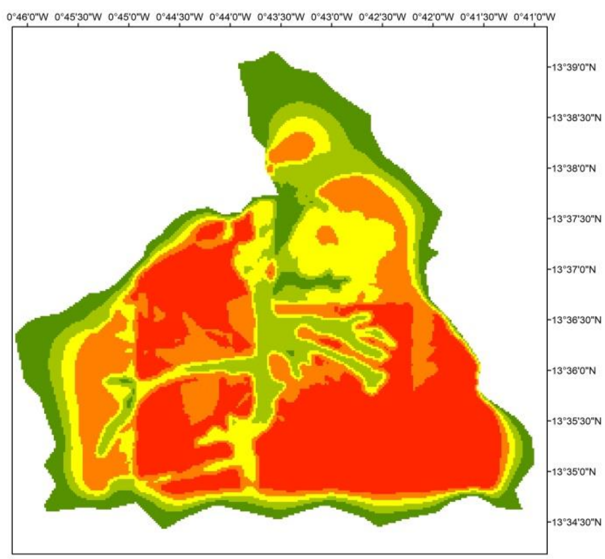
Figures 4 and 5 show the probability maps of the spatial distribution of cyanide contamination in Zougnazagmiline and Galgouli.

In Zougnazagmiline, cyanide contamination probability varies from very low to moderate near the catchment boundary, irrespective of the model. Figure 4 [Model 1] displays an over-predicted model for cyanide contamination hazard. In fact, the very high probability (p = 0.80 - 1) for cyanide contamination covers over 50% of the catchment area. Figure 4 [Model 2] and [Model 3], however, more precisely defines the zones that present a very high risk of cyanide contamination, which represent 30% and 20% of the total surface area, respectively.

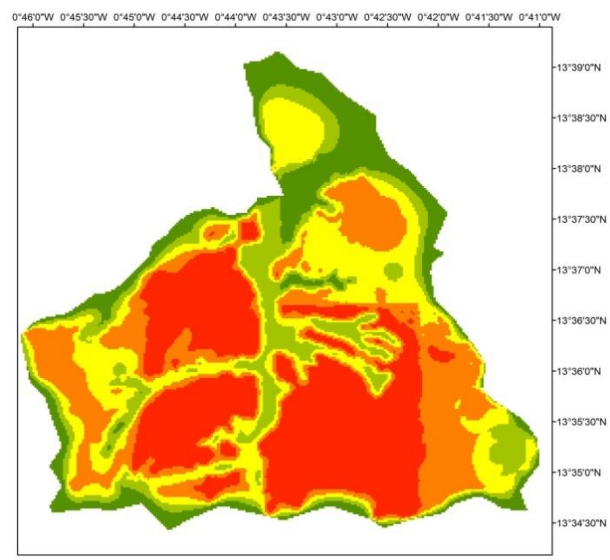
The zones with the greatest probability are near the cyanidation zones, which was expected. Part of the river bank also has a high chance of cyanide contamination, which is due to the river bank often containing clay, in which FCN is likely to be present.

On the other hand, there is a low probability of cyanide pollution around the catchment outlet as shown in Figure 4 [Model 3]. This is likely due to FCN volatilisation and dilution into the main stream flow because FCN is very soluble in surface water (Dash et al., 2009; Lötter, 2005) .

[Model 1]



[Model 2]



[Model 3]

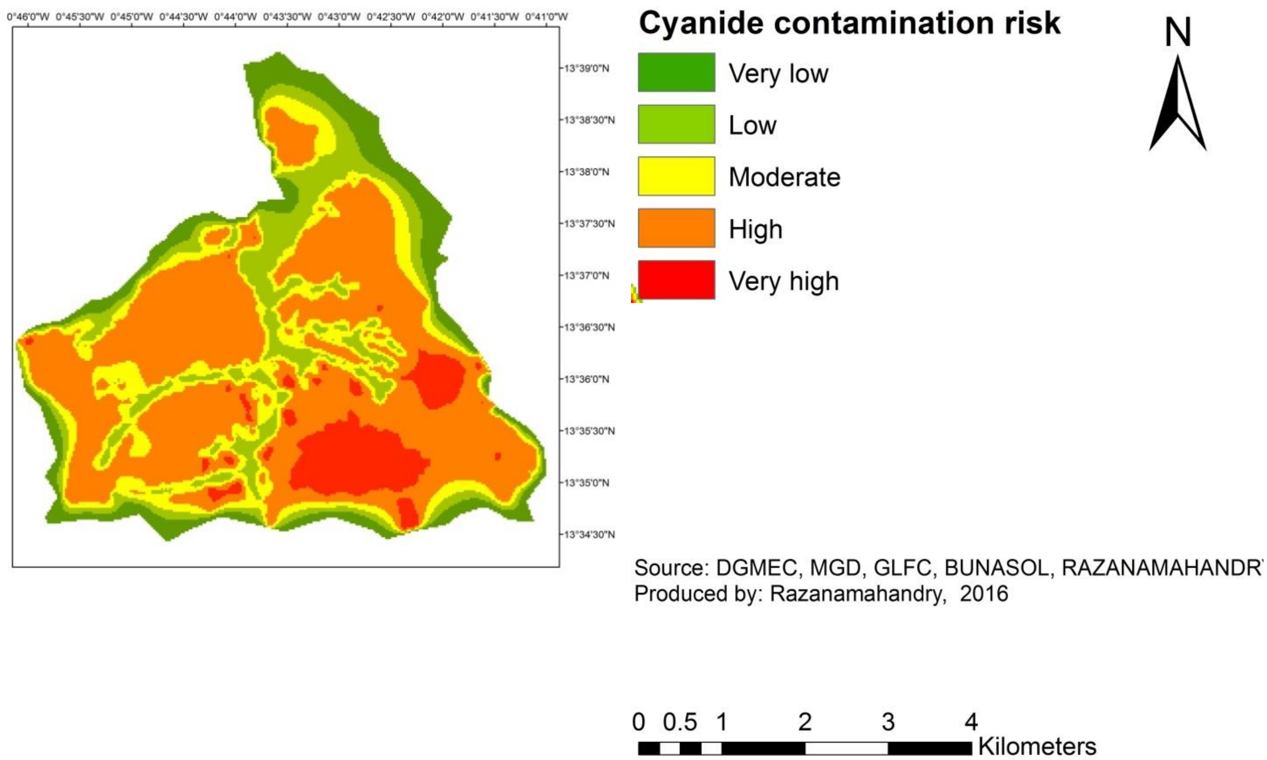


Figure 4 : Probability map for cyanide contamination risk at the North.

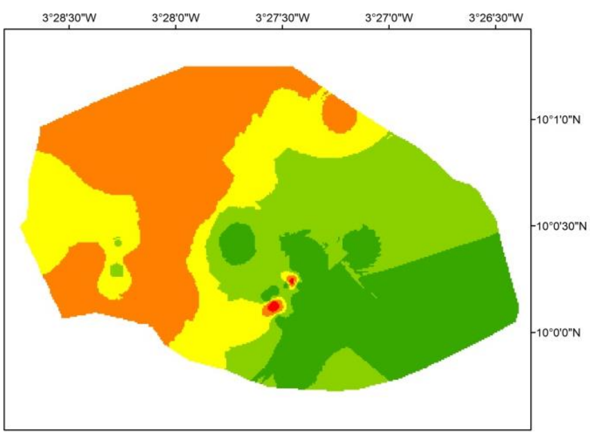
Model 1 of Galgouli(Fig. 5)is also over-predicted. The areas of high probability of cyanide contamination cover about40% of the total surface area. Model 2 and Model 3 in Figure 5 and give almost the same spatial distribution of cyanide in soil, with only 5% of the catchment area having a very high probability of cyanide contamination, which corresponds to the cyanidation zone. The explanatory factors selected in Model 2 and Model 3 are significantly correlated with and explain the spatial distribution of cyanide.

The area of predicted FCN contamination represents less of the catchment at Galgouli (5%) than at Zougnazagmiline (20%).

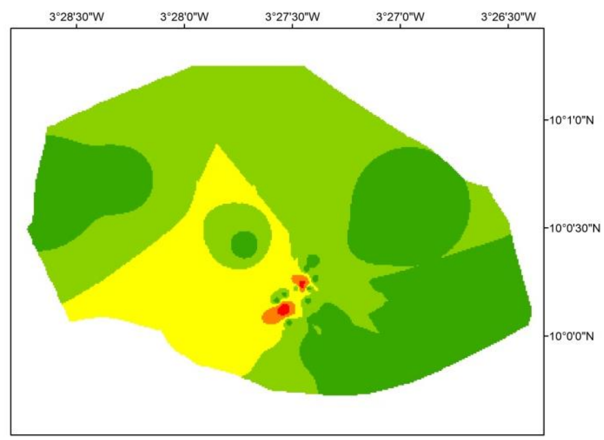
Botz et al. (2015),Bureau et al. (2011) and Kjeldsen (1999) have reported that FCN could take the anion cyanide form (CN^-) that reacts with metal cations under high pH conditions. Furthermore, Nsimba (2009) and Wong-chong et al.(2006) found that FCN takes the gaseous form HCN and could easily volatilize under acidic conditions. The climate of Zougnazagmiline is arid with a basic soil, whereas Galgouli is humid with acidic soil. Therefore, CN^- is the dominant component of FCN in Zougnazagmiline whereas HCN is more available in Galgouli. Our results suggest that FCN accumulation in arid Zougnazagmiline and volatilisation in humid Galgouli. The soil characteristics could influence the cyanide spatial distribution in the soil.

Since ASGM activities are more organized in Galgouli, the cyanidation zone is concentrated in one place. This is not the case in Zougnazagmiline, where the cyanidation zones were distributed throughout the catchment area, which increases the area subject to contamination. It could be a reason explaining the probability difference between the two sites.

[model 1]



[model 2]



[model 3]

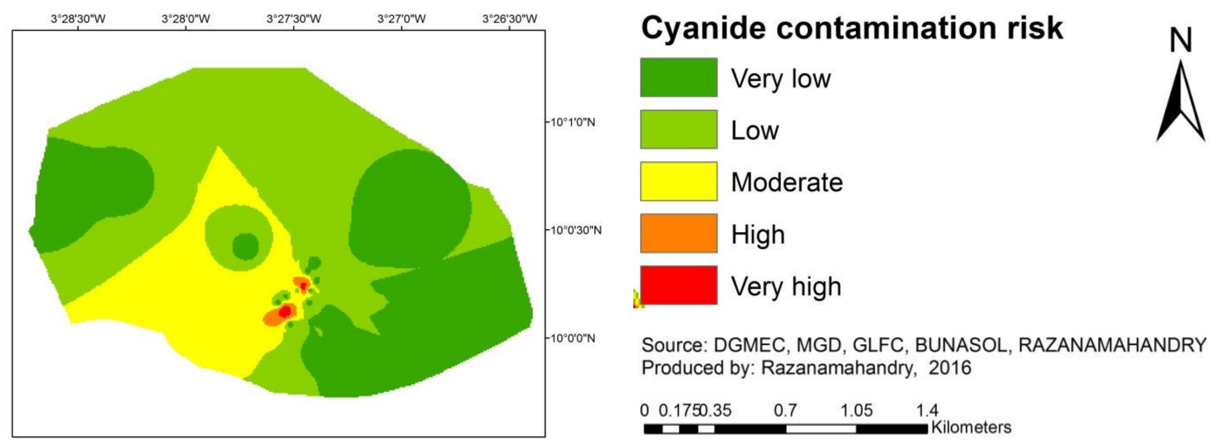


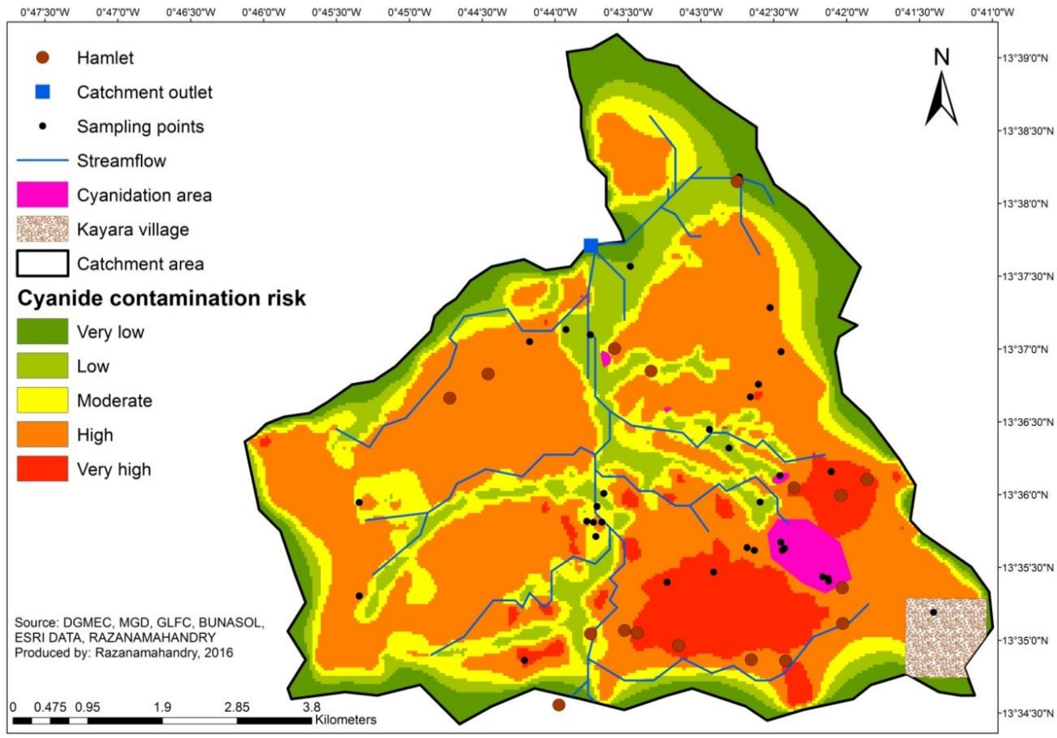
Figure 5 : Probability map for cyanide contamination risk at the South.

As said in previous paragraph, FCN was more fixed by the Zougnazagmiline soil than by that of Galgouli. It was reported that greater biological activity is observed in regions dominated by arenosols, which is characterised by sand particles, clay and sediment in which some structure and fertility are provided (FAO, 1998, 2001). Therefore, most of the FCN present in Galgouli was degraded by microbial activity, or reduced by vegetation uptake and complexation with metals, explaining the small area covered by FCN contamination (Figure 5).

4.2.2 Cyanide contamination exposure

The exposure of the populated areas (hamlets) to cyanide contamination is presented with Model 3 in Figures 4 and 5. More than 20% of the hamlets at Zougnazagmiline are exposed to cyanide contamination, whereas only 5% of the hamlets are exposed in Galgouli.

(a)



0°47'30"W
0°47'0"W
0°46'30"W
0°46'0"W
0°45'30"W
0°45'0"W
0°44'30"W
0°44'0"W
0°43'30"W
0°43'0"W
0°42'30"W
0°42'0"W
0°41'30"W
0°41'0"W
13°39'0"N
13°38'30"N
13°38'0"N
13°37'30"N
13°37'0"N
13°36'30"N
13°36'0"N
13°35'30"N
13°35'0"N
13°34'30"N
N
Hamlet
Catchment outlet
Sampling points
Streamflow
Cyanidation area
Kayara village
Catchment area
Cyanide contamination risk
Very low
Low
Moderate
High
Very high
Source: DGMEC, MGD, GLFC, BUNASOL, ESRI DATA, RAZANAMAHANDRY
Produced by: Razanamahandry, 2016
0 0.475 0.95 1.9 2.85 3.8
Kilometers

(b)

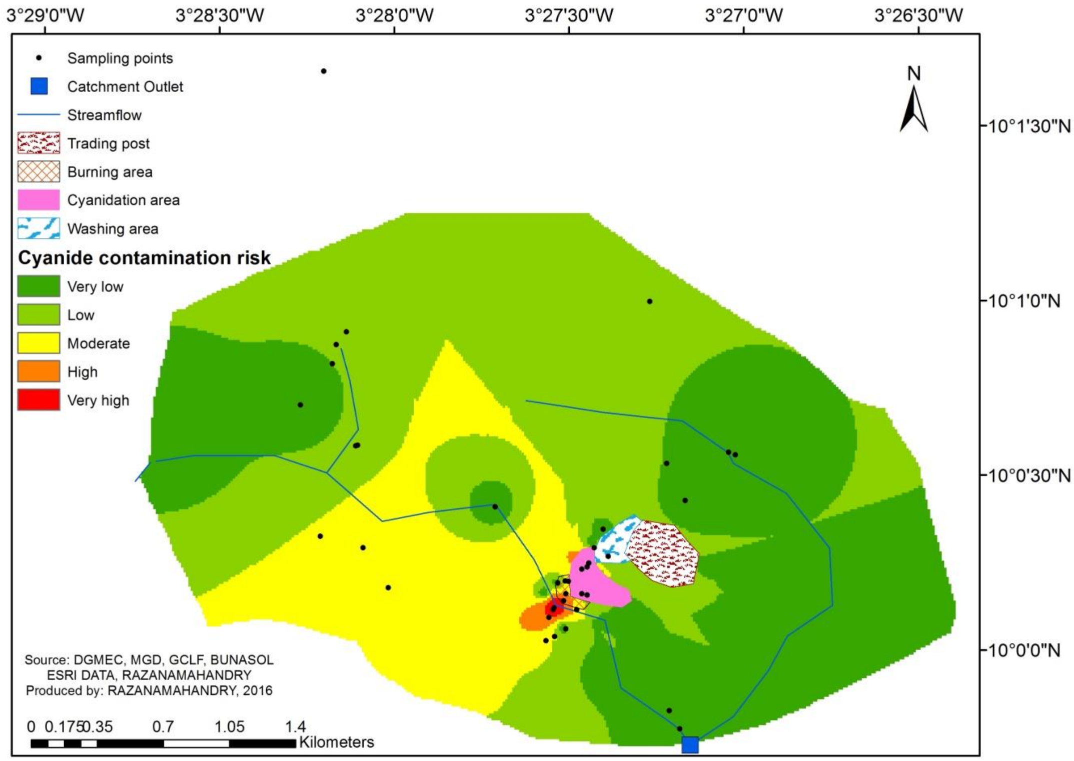


Figure 6: Cyanide contamination exposure (a) in Zougnazagmiline and (b) in Galgouli

5 Conclusions

ASGM is the primary economic activity of the people of Zougnazamiline and Galgouli. Since water and soil contamination by cyanide is widespread at these sites, it was deemed necessary to determine the most vulnerable areas in order to prioritize restoration of the degraded ecosystem. Environmental factors related to the spatial distribution of cyanide have been evaluated. Three predictive models using LR were created for each site, which have different climate conditions and soil characteristics. The most important factors influencing the FCN distribution are the soil type in Zougnazagmiline and the soil conductivity in Galgouli. Therefore, when the zone is arid, only the soil type would influence the FCN distribution but if it is humid, both the soil type and the soil conductivity would be the main influencing parameter. Environmental factors such as the distance from cyanidation zones, topographic elevation and land use are likely to increase cyanide contamination risk. Since the soil conductivity and soil type are dependent upon the soil composition, that aspect needs to be investigated in depth to fully understand FCN distribution in ASGM affected areas. Moreover, the LR model should also be tested in the zone under Soudanese-Sahelian climate to determine the main parameters that influence the FCN distribution in semi-arid areas. This would then allow for the prediction of FCN distribution for any ASGM area in Burkina Faso based on its climate and soil characteristics.

Acknowledgments

This present study was funded by the Swiss Agency for Development and Cooperation under the 3E Programme, Grant Number 81016359/1. Authors would like to particularly thank Dr Fowe TAZEN, and Mr Sina THIAM, who gave additional advices for the data analysis, and to the laboratory technicians who contributed in data collection and FCN analysis.

References

Adler, R., Bergquist, B.A., Adler, S.E., Remy, J., Guimarães, D., Lees, P.S.J., Niquen, W., Velasquez-lópez, P.C., Veiga, M.M., 2013. Challenges to measuring , monitoring , and addressing the cumulative impacts of artisanal and small-scale gold mining in Ecuador. Resour. Policy 1–10. doi:10.1016/j.resourpol.2013.03.007

Alegria, A.C., Sahli, H., Zimanyi, E., 2011. Application of density analysis for landmine risk mapping, in: IEEE International Conference on spatial data mining and geographical knowledge Services. bll 223 – 228.

Benini, A.A., Moulton, L.H., Conley, C.E., 2002. Landmines and local community adaptation. J. Contingencies Cris. Manag. 10, 82–94. doi:http:// dx.doi.org/10.1111/1468-5973.00184.

Bernstein, P., 2000. The power of gold: the history of an obsession, John Wiley. ed.

Butaré, I., & Keita, S., 2009. Aspects environnementaux liés au développement du secteur minier

[WWW Document]. URL http://www.africabusinessmarket.com/fichier/1290108564aspect.pdf

Caplan, J.M., 2011. Mapping the spatial influence of crime correlates : a comparison of operationalization schemes and implications for crime analysis and criminal justice practice. Cityscape 13, 57 – 83.

Caplan, J.M., Kennedy, L.W., Miller, J., 2011. Risk terrain modeling: brokering criminological theory and GIS methods for crime forecasting. Justice Q. 28, 360 – 381. doi:http://dx.doi.org/10.1080/07418825.2010.486037

Chamberlayne, E., 2002. A GIS model for minefield area prediction: The minefield likelihood procedure., Ph.D. thes. ed.

Chupezi, T.J., et al., 2009. Impacts of Artisanal Gold and Diamond Mining on Livelihoods and the Environment in the Sangha Tri-National Park Landscape, Center for. ed. Bogor.

Craig, M.H., Sharp, B.L., Mabaso, M.L.H., Kleinschmidt, I., 2007. Developing a spatial-statistical model and map of historical malaria prevalence in Botswana using a staged variable selection procedure. Int. J. Health Geogr. 6, 44. doi:http://dx.doi.org/10.1186/1476-072X-6-44

Dai, X., Simons, A., Breuer, P., 2012. A review of copper cyanide recovery technologies for the cyanidation of copper containing gold ores. Miner. Eng. 25, 1–13. doi:10.1016/j.mineng.2011.10.002

Dash, R.R., Gaur, A., Balomajumder, C., 2009. Cyanide in industrial wastewaters and its removal : A review on biotreatment. J. Hazard. Mater. 163, 1–11. doi:10.1016/j.jhazmat.2008.06.051

Egan, J.P., 1975. Signal Detection Theory and ROC Analysis. Ser. Cogn. Perception. Acad.

Press. New York.

Ehlers, K., Bakken, L.R., Frostegård, Å., Frossard, E., Bünemann, E.K., 2010. Phosphorus limitation in a Ferralsol: Impact on microbial activity and cell internal P pools. Soil Biol. Biochem. 42, 558–566. doi:10.1016/j.soilbio.2009.11.025

Ekpo, U.F., Mafiana, C.F., Adeofun, C.O., Solarin, A.R., Idowu, A.B., 2008. Geographical information system and predictive risk maps of urinary schistosomiasis in Ogun state, Nigeria. BMC Infect. Dis. 8, 1471 – 2334. doi:http:// dx.doi.org/10.1186/1471-2334-8-74

Ettinger, S., Mounaud, L., Magill, C., Yao-Lafourcade, A.-F., Thouret, J.-C., Manville, V., Al., E., 2015. Building vulnerability to hydro-geomorphic hazards: estimating damage probability from qualitative vulnerability assessment using logistic regression. J. Hydrol. doi:http://dx.doi.org/10.1016/j.jhydrol.2015.04.017

FAO, 2001. Lecture notes on the major soils of the World, Wordl Soil Resources Reports. doi:10.1136/gut.27.11.1400-b

FAO, 1999. Soil salinity assessment: Methods and interpretation of electrical conductivity measurements. Irrig. Drain. Pap. 57 165.

FAO, 1998. World Reference Base for Soils Resources. Rome, Italy.

Fawcett, T., 2003. ROC Graphs : Notes and Practical Considerations for Data Mining Researchers, Hewlett-Pa. ed.

Gajigo, O., Mutambatsere, E., Ndiaye, G., 2012. Gold Mining in Africa : Maximizing Economic Returns for Countries, African Development Bank Group.

Ghosh, R.S., Dzombak, D.A., Wong-Chong, G.M., 2006. Physical and forms of cyanide, in: CYanide in Water and Soil: Chemistry , Risk and Management. bll 15 – 23.

Goovaerts, P., Xiao, H., Adunlin, G., Ali, A., Tan, F., Gwede, C.K. et al., 2015. Geographically-weighted regression analysis of percentage of late-stage prostate cancer diagnosis in florida. Appl. Geogr. 62, 191 – 200. doi:http:// dx.doi.org/10.1016/j.apgeog.2015.04.018

Grätz, T., 2009. Moralities , risk and rules in West African artisanal gold mining communities : A case study of Northern Benin $ 34, 12–17. doi:10.1016/j.resourpol.2008.11.002

Grimaldi, M., Guédron, S., Grimaldi, C., 2015. Impact of gold mining on mercury contamination and soil degradation in Amazonian Ecosystems of French Guiana. Land-use Chang. Impacts Soil Process. Trop. Savannah Ecosyst. 1, 95–106. doi:10.1017/CBO9781107415324.004

Groundwater Assessment Platform (GAP), 2015. Mapping and information platform for geogenic groundwater contamination.

Guèye, D., 2001. Small-Scale Mining in Burkina Faso.

Guimaraes, J.R.D., Betancourt, O., Rodrigues, M., Barriga, R., Cueva, E., Betancourt, S., 2011. Long-range effect of cyanide on mercury methylation in a gold mining area in southern Ecuador. Sci. Total Environ. 409, 5026–5033. doi:10.1016/j.scitotenv.2011.08.021

Guns, M., Vanacker, V., 2012. Logistic regression applied to natural hazards: rare event logistic regression with replications. Nat. Hazards Earth Syst. Sci. 12, 1937 – 1947. doi:http://dx.doi.org/10.5194/nhess-12-1937-2012.NHESS.

Guo, B., Peng, Y., Espinosa-gomez, R., 2014. Cyanide chemistry and its effect on mineral flotation. Miner. Eng. 66-68, 25–32. doi:10.1016/j.mineng.2014.06.010

Hilson, G., 2009. Small-scale mining, poverty and economic development in sub- Saharan Africa: an overview. Resour. Policy 34, 1–5.

Hilson, G., 2002. The future of small-scale mining: environmental and socioeco- nomic perspectives. Futures 34, 863 – 872.

Janneh, A., Ping, J., 2011. Minerals and Africa’s Development: The International Study Group

Report on Africa’s Mineral Regimes, Economic C. ed. Addis Ababa.

Johnson, C.A., 2014. Applied Geochemistry The fate of cyanide in leach wastes at gold mines : An environmental perspective. Appl. Geochemistry 1. doi:10.1016/j.apgeochem.2014.05.023

Kheir, R.B., Shomar, B., Greve, M.B., Greve, M.H., 2014. On the quantitative relationships between environmental parameters and heavy metals pollution in Mediterranean soils using GIS regression-trees: The case study of Lebanon. J. Geochemical Explor. doi:10.1016/j.gexplo.2014.05.015

Kjeldsen, P., 1999. Behaviour of cyanides in soil and groundwater: a review. Water, Air, Soil Pollut. 115, 279–307.

Klenk H, Griffiths A, Huthmacher K, Itzel H, Knorre H, Voight C, et al., 1996. Cyano, inorganics, in: Gerhartz W, Yamamoto YS, Kaudy L, Pfefferkorn R, Rounsaville JF, Ullmanns encyclopedia of industrial chemistry, A8. VCH. bll 159 – 190.

Kumar, R., Saha, S., Dhaka, S., Kurade, M.B., Kang, C.U., Baek, S.H., Jeon, B.-H., 2016. Remediation of cyanide-contaminated environments through microbes and plants: a review of current knowledge and future perspectives. Geosystem Eng. 0, 0. doi:doi=10.1080%2F12269328.2016.1218303

Lacroix, P., Herzog, J., Eriksson, D., Weibel, R., 2013. Methods for visualizing the explosive remnants of war. Appl. Geogr. 41, 179 – 194. doi:http://dx.doi.org/ 10.1016/j.apgeog.2013.04.007

Lin, Y.-P., Cheng, B.-Y., Chu, H.-J., Chang, T.-K., Yu, H.-L., 2011. Assessing how heavy metal pollution and human activity are related by using logistic regression and kriging methods. Geoderma 163, 275–282. doi:10.1016/j.geoderma.2011.05.004

Lötter, N., 2005. Cyanide volatilisation from gold leaching operations and tailing storage facilities submitted in fulfilment of the requirements of the degree Master of Metallurgical Engineering in the Faculty of Engineering , Built Environment and Informa, submitted. ed. Master of Metallurgical Engineering. Faculty of Engineering, Built Environment and information technology. University of Pretoria. South Africa.

Notario, del P., Jesús, S., Ruiz-Gallardo, José-Reys, 2014. Modelling post-fire soil erosion hazard using ordinal logistic regression: A case study in South-eastern Spain. Geomorphology. doi:10.1016/j.geomorph.2014.12.005

Pallo, F., Sawadogo, N., 2011. Essai de corrélation entre les caractéristiques du climat, de la végétation et des sols au Burkina Faso. Int. J. Biol. Chem. Sci. 4, 1839–1850. doi:10.4314/ijbcs.v4i5.65563

Razanamahandry, L.C., Andrianisa, H.A., Karoui, H., Kouakou, K.M., Yacouba, H., 2016. Biodegradation of free cyanide by bacterial species isolated from cyanide-contaminated artisanal gold mining catchment area in Burkina Faso. Chemosphere 157, 71–78. doi:10.1016/j.chemosphere.2016.05.020

Real, R., Barbosa, A.M., Vargas, J.M., 2006. Obtaining environmental favour- ability functions from logistic regression. Environ. Ecol. Stat. 13, 237 – 245.

Richards, B.K., Steenhuis, T.S., Peverly, J.H., Mcbride, M.B., 2000. E € ect of sludge-processing mode , soil texture and soil pH on metal mobility in undisturbed soil columns under accelerated loading 109.

Rodrigues, M., de la Riva, J., Fotheringham, S., 2014. Modeling the spatial variation of the explanatory factors of human-caused wildfires in spain using geographically weighted logistic regression. Appl. Geogr. 48, 52 – 63.

doi:http://dx.doi.org/10.1016/j.apgeog.2014.01.011.

Sampat, P., 2003. Scrapping mining dependence. Published in State of the World 2003. New York, NY: Norton.

Schultz, C., Cecilia, A., Cornelis, J., Sahli, H., 2016. Comparison of spatial and aspatial logistic regression models for landmine risk mapping. Appl. Geogr. 66, 52–63. doi:10.1016/j.apgeog.2015.11.005

Schure, J., et al., 2011. Is the god of diamonds alone? The role of institutions in artisanal mining in forest landscapes, CongoBasin. Resour. Policy 33, 363 – 371.

Siegel, S., Veiga, M.M., 2009. Artisanal and small-scale mining as an extralegal economy: De Soto and the redefinition of "'formalization'". Resour. Policy 34, 51–56.

Street, L., Palmberg, J., Artigas, J.C., Grubb, M., 2013. Gold demand trends: full year 2012, World Gold. ed. London, United Kingdom.

Swets, J., 1988. Measuring the accuracy of diagnostic systems. Science (80-.). 240, 1285 – 1293.

Swets, J.A., Dawes, R.M., Monahan, J., 2000. Better decisions through science. Sci. Am. 283, 82– 87. doi:http://www.psychologicalscience.org/newsresearch/publications/journals/%siam.pdf.

Telmer, K., Veiga, M., 2008. World Emissions of Mercury from Small-scale and Artisanal Gold Mining, in: Mercury Fate and Transport in the Global Atmosphere: Emissions, Measurements and Models. Springer, Germany.

Theis, T. L.;West, M.J., 1986. Effects of cyanide complexation on adsorption of trace metals at the surface of goethite. Environ. Technol. Lett. 7, 309 – 318.

Van Den Eeckhaut, M., Vanwalleghem, T., Poesen, J., Govers, G., Verstraeten, G., Vandekerckhove, L., 2006. Prediction of landslide susceptibility using rare events logistic regression: a case-study in the Flemish Ardennes (Belgium). Geomorphology 76, 392 – 410. doi:http://dx.doi.org/10.1016/j.geomorph. 2005.12.003

Veiga, M.M., Angeloci-santos, G., Meech, J.A., 2014. The Extractive Industries and Society Review of barriers to reduce mercury use in artisanal gold mining. Biochem. Pharmacol. doi:10.1016/j.exis.2014.03.004

Velásquez-lópez, P.C., Veiga, M.M., Klein, B., Shandro, J.A., Hall, K., 2011. Cyanidation of mercury-rich tailings in artisanal and small-scale gold mining : identifying strategies to manage environmental risks in Southern Ecuador. J. Clean. Prod. 19, 1125–1133. doi:10.1016/j.jclepro.2010.09.008

Venkataraman, K., Uddameri, V., 2012. Modeling simultaneous exceedance of drinking-water standards of arsenic and nitrate in the Southern Ogallala aquifer using multinomial logistic regression. J. Hydrol. 458-459, 16–27. doi:10.1016/j.jhydrol.2012.06.028

Vitousek, P.M., 1984. Litterfall, nutrient cycling, and nutrient limitation in tropical forests. Ecology 65, 285 – 298.

Wang, L.-J., Guo, M., Sawada, K., Lin, J., Zhang, J., 2015. Landslide susceptibility mapping in Mizunami city, Japan: a comparison between logistic regression, bivariate statistical analysis and multivariate adaptive regression spline models. Catena 135, 271 – 282. doi:http://dx.doi.org/10.1016/j.catena.2015.08.007

Weng, L., Endamana, D., Klintuni, A., Levang, P., Margules, C.R., Sayer, J.A., 2014. The Extractive Industries and Society Asian investment at artisanal and small-scale mines in rural Cameroon. Biochem. Pharmacol. doi:10.1016/j.exis.2014.07.011