Risk assessment of unlined oil pits leaking into groundwater in the Ecuadorian Amazon: A modified GIS-DRASTIC approach

Juan Durango-Cordero, Mehdi Saqalli, Sylvain Ferrant, S. Bonilla, Laurence Maurice, P. Arellano, Arnaud Elger

Version of Record: https://www.sciencedirect.com/science/article/pii/S0143622821002447
Manuscript_0c90c381146dcb330ed24d8e69485015

Risk assessment of unlined oil pits leaking into groundwater in the Ecuadorian Amazon: A modified GIS-DRASTIC approach

Durango-Cordero, J.[a, b, c, d*], Saqalli, M.[c], Ferrant, S.[e], Bonilla, S.[a, f], Maurice, L.[g], Arellano P.[h] and Elger, A.[c]

[a] *Research Center for the Territory and Sustainable Habitat, Universidad Tecnológica Indoamérica, Machala y Sabanilla, 17030, Quito, Ecuador.*

[b] *Facultad de Arquitectura, Artes y Diseño, Universidad Tecnológica Indoamérica, Av. Manuela Sáenz y Agramonte, 180212, Ambato, Ecuador.*

[c] *Laboratoire d'Ecologie Fonctionnelle, Université de Toulouse, CNRS, INPT, UPS, Toulouse, France.*

[d] *GEODE, Université de Toulouse, CNRS, UT2J, Toulouse, France.*

[e] *CESBIO, Université de Toulouse, IRD, CNRS, CNES, UPS, INRA, Toulouse, France.*

[f] *Ingeniería en Biodiversidad y Recursos Genéticos. Facultad de Ciencias del Medio Ambiente. Universidad Tecnológica Indoamérica, Machala y Sabanilla, 170301, Quito, Ecuador.*

[g] *Géosciences Environnement Toulouse (GET), Université de Toulouse, CNRS, IRD, Toulouse, France.*

[h] *Michigan Technological University, College of Natural Resources and Environmental Sciences, Houghton, Michigan, USA.*

[*]**Corresponding author:** Juan Durango-Cordero

Email address; juandurango@uti.edu.ec

Telephone: + 593 984 140 671

Fax: +593 2 2 247 741

Research Center for the Territory and Sustainable Habitat

180150, Av. Manuela Sáenz y Agramonte

Ambato, Ecuador

Declarations of interest: none.

Risk assessment of unlined oil pits leaking into groundwater in the Ecuadorian Amazon: A modified GIS–DRASTIC approach

Highlights

- Aquifer vulnerability was assessed in forested and agricultural lands in an oil exploited region.
- Sensitivity analysis utilising the DRASTIC model was robust.
- Risk was mapped using oil pit hazard, a modified DRASTIC index, and cost-distance analysis.
- Micro-slopes and the distance of transported pollutants from the source were the main factors influencing the surface at risk.
- Maximum impacted surface may reach 13% of the Northern Ecuadorian Amazon.

Abstract

This study evaluates the risk of groundwater contamination from unlined oil pits, in the Northern Ecuadorian Amazon (NEA). Applying spatial analysis, several maps were provided for its integration in land use planning, public health improvement and future site-specific investigations. Two main maps were produced: (1) a vulnerability indexed map using a modified DRASTIC model and (2) a hazard map based on the past (1995-1997) and present (2018) contamination using a weighted density equation. The hazard was derived from hydrocarbon contained in oil pits associated with a cost-distance analysis to obtain different maximum distance ranges (MDR), to model the surface of potentially impacted groundwater. The results indicate a total calculated hydrocarbons of 39 052 tons. A MDR from 500-10 000 km was retained to map aquifers at risk, the maximum surface potentially at risk covers 13% of the NEA, while 83% of the area represents low to medium-low vulnerability. This study led to several recommendations, such as the level of suitability of the available information, and what gaps should be filled to improve future research. A surface of 271-766.5 km. in the 500-2000-m distance range should be prioritised for finer scale risk assessment.

Keywords: risk assessment, overlay index, Amazon, groundwater, oil pollution, land use planning.

1. Introduction

In the Northern Ecuadorian Amazon (NEA), oil production is alleged to have released hydrocarbons into ecosystems, causing human health issues (San Sebastián & Hurtig, 2005). Studies on hydrocarbon contamination have indicated high hydrocarbon concentrations in water and soil used by communities for daily life purposes (Barraza et al., 2018; Maddela et al., 2016; Merchán-Rivera et al., 2017; Wasserstrom, 2013). Maurice et al., (2019) and Merchán-Rivera et al., (2017) recently reported that drinking water and groundwater were contaminated by hydrocarbons. This contamination has previously been linked to a wide variety of health effects (Coronel Vargas et al., 2020; Ramirez et al., 2017; San Sebastián et al., 2001).

Previous studies have evaluated oil hazards, such as oil spills (Durango-Cordero et al., 2018) and gas flaring (Barraza et al., 2020; Durango-Cordero et al., 2019), as well as exposure patterns (Jaderne Houssou et al., 2019) and human health risk assessments (Barraza et al., 2018). Non-evaluated contaminant sources include unlined earthen reserves, mud drilling, and waste oil pits (hereafter referred to as 'pits'). Up until 1993, while extracting oil and gas in the NEA, oil producers (the CEPE–Texaco consortium being the most prominent) built and later abandoned approximately 900 pits (Buccina et al., 2013), into which 48 183 t of crude oil was poured (MAE-PRAS, 2016). As part of the 1995 Remedial Action Plan (RAP), only 158 pits have been covered with a barren soil coating (MAE-PRAS, 2016).

The purpose of a pit is to hold drilling cuts, drilling mud, and crude hydrocarbons that are not collected by the pipeline (MAE-PRAS, 2016). Technically, their size should be sufficient to ensure adequate storage until closure, and bottom

depth should not reach the groundwater table to ensure that toxic compounds do not adversely impact groundwater. A site chosen for the location of pits should be investigated in terms of vulnerable aquifers, and a specific action plan should be developed to protect them (EPA, 2014). The Ecuadorian National Programme of Environmental and Social Remediation (PRAS) was initiated in 2008 and registered historical data (1972–2013) on 1561 pits (141 were eliminated before 2011), of which 1420 were found to have a high-volume variability (see Section 2.5).

To investigate the potential contamination of groundwater, researchers should identify vulnerable zones across the NEA, that is, zones that would be more impacted by occasional or chronic contamination from unlined oil pits. Aquifer vulnerability is defined as the susceptibility of an aquifer to be reached and affected by pollutants, for example, the physicochemical properties of a pollutant (Andreo et al., 2006) and its diffusion into groundwater (Vrba & Zaporozec, 1994). Subsequently, the contamination risk posed to groundwater can be spatially defined as a function of hazard (i.e. the probability of contaminant leaching) and aquifer vulnerability (Kazakis & Voudouris, 2015; Metzger et al., 2008).

Aller et al. (1987) developed the DRASTIC model to assess groundwater vulnerability which conveys more information regarding groundwater contamination potential than other geographic information system (GIS)-based methods, such as the susceptibility index (SI); groundwater, overall lithology, and depth to water index (GOD); aquifer vulnerability index (AVI); and SINTACS index (acronym based on soil and hydrogeological parameters) (Gogu & Dassargues, 2000). The datasets were organised into seven parameters to assess groundwater vulnerability with different

79 rating scores and weights (Aller et al., 1987). The DRASTIC model has been modified
80 since its development, depending on available spatial data (Al-Adamat et al., 2003;
81 Merchant, 1994; Mimi & Assi, 2009; Stigter et al., 2006). For example, in the Daule
82 aquifer in Guayas, Ecuador, groundwater vulnerability was assessed using a reduced
83 set of parameters of the DRASTIC index model because of several intrinsic
84 shortcomings of the model (Merchant, 1994) or site-specific limitations (Ribeiro et
85 al., 2017). Some studies have incorporated land use types (Mimi et al., 2012),
86 including industry-related contamination (e.g. oil and gas exploitation) (Kazakis &
87 Voudouris, 2015; Mimi & Assi, 2009; Secunda et al., 1998), with the specific
88 physicochemical characteristics of pollutants. Other studies have replaced some of the
89 original parameters with other more specific parameters with respect to context and
90 data availability, for instance, replacing hydraulic conductivity with groundwater
91 velocity (Kazakis & Voudouris, 2015). More recently, the 'specific vulnerability to
92 anthropogenic pollution' (SVAP) was determined using multivariate and regression
93 analyses (Busico et al., 2020).
94 This study attempts to assess the risk of groundwater contamination from
95 hydrocarbons or associated pollutants, and to provide maps for integrating this risk
96 into land use planning, public health improvement, and future detailed site-specific
97 investigations when contamination data are only available for specific areas (MAE-
98 PRAS, 2016; Maurice et al., 2019). The innovation of the present study lies in the use
99 of information about past and present oil waste accumulated in oil pits to evaluate the
100 corresponding risk posed to aquifers in a very sensitive ecosystem. The
101 methodological approach enables researchers to (1) define vulnerability levels for

aquifers throughout the NEA, (2) estimate and map the hydrocarbon contents of pits and the associated hazard level, and (3) evaluate the risk of pit leakage to groundwater along the runoff stream network using a vulnerability index map overlaying a hazard map based on the potential propagation of a contaminant plume at the provincial/local scale.

2. Materials and Methods

2.1. Study area

This study, conducted in the NEA (Fig. 1A), was restricted to 35 051 km^2 in the Amazon lowlands (~144–900 m above sea level (a.s.l.)) (Fig. 1B). The area includes upstream oil/gas production activities that potentially impact forests, rivers, human settlements, and croplands. Rivers with high flow rates were excluded (Fig. 1B) because surface waters and their interactions with groundwater are highly dynamic and therefore difficult to characterise based on scarce data. The NEA has a highly warm and humid climate with an average rainfall of 2900 mm y^{-1} (Institute of Meteorology and Hydrology). The hydrological regime discharges 1000–5000 m^3 s^{-1}, generally in the Ecuadorian Amazon Basin, resulting in flash floods due to extreme rainfall events (Laraque et al., 2007). Evapotranspiration ranges from 960 mm y^{-1} to 1560 mm y^{-1} with annual average of 1300 mm· (Paca et al., 2019). The aquifers are unconfined or confined with medium to high permeability, producing small to large quantities of freshwater (US Army Corps, 1998) (Fig. A.3). The depth of the water table varies from 0.2 m on the eastern side of the NEA (near the Cuyabeno Reserve and Napo River) to more than 10 m on the northern and southwestern sides, and

reaches a depth of 1 m in the Joya de los Sachas and Shushufindi regions (Figs. 1B and A.1). To date, only four aquifers have been rigorously identified (Fig. 1B). More information on the aquifer characteristics can be found in the Supplementary Material (Table A.1 and Figs. A.1–A.6).

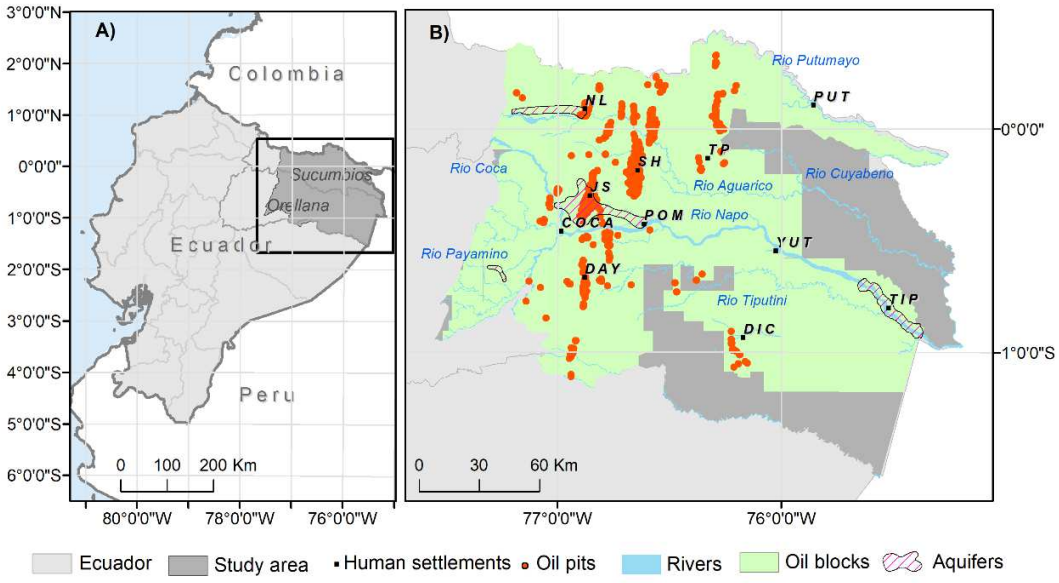


Fig. 1. Locations of oil pits. A: NEA region; B: study area within the concerned provinces of Sucumbíos and Orellana, including oilfields. Abbreviations of human settlements: DAY: Dayuma; NL: Nueva Loja; SH: Shushufindi; TP: Tarapoa; POM: Pompeya; PUT: Putumayo; COCA: Puerto Francisco de Orellana; JS: Joya de Los Sachas; YUT: Yuturi; DI: Dícaro; TIP: Tiputini.

2.2. *Data description*

Table 1.

Summary of data used for vulnerability and hazard analyses.

Vulnerability parameters	**Data**	**Scale**	**Sources**
Depth to the water table	Geopedological maps	1:25 000	MAGAP[1] (1984) SIGTIERRAS[2] (2015)
Final **R**echarge	Annual mean rainfall map DEM Hypsometry	1:250 000 90 m × 90 m pixel resolution	MAGAP (2002) IGM[3] (2010)
Aquifer media, lithology, and formations	Hydrogeology map	1:250 000	MAGAP (2005)
Soil media	Geopedological map	1:25 000	MAGAP (2002)
Topography, slope	DEM	90 m × 90 m pixel resolution	IGM[3] (2010)
Impact of zone vadose	DEM Soil types	90 m × 90 m pixel resolution 1:25 000	IGM[3] (2010) MAGAP (1984) SIGTIERRAS (2015)
Hydraulic **C**onductivity, replaced by the infiltration rate	Geopedological map	1:25 000	MAGAP (2002)
Hazard parameters			
Contaminant content	Total Petroleum Hydrocarbons	EPA 1311/418.1 modified method MONOIL method	ADF (2008) MONOIL-ANR project (2018)
	American petroleum index Gravity	Proof samples in several oil wells	National Board of Hydrocarbons (2014) ADF (2008)
	Soil bulk density	Generic defaults	Environmental Protection Agency-OHIO (2005)
	Soil texture	Derived from generic	MAGAP-SIGTIERRAS (1984-2019)
Pit volume	Oil pits	Point shapefile	MAE-PRAS[4] (2014)
Cost–distance	DEM	90 m × 90 m pixel resolution	IGM (2010)

[1] Ministry of Agriculture, Livestock, Aquaculture and Fisheries
[2] National System for Rural Land Information and Management and Technology Infrastructure
[3] Geographic Military Institute
[4] Programme of Social and Environmental Reparation by the Ministry of the Environment
[5] Amazon Defence Front

Several spatial sources were classified into two categories: (1) vulnerability and (2) hazard from pits. Both categories were evaluated using the available spatial of

139 the National Information System (http://sni.gob.ec/coberturas) and Geographical
140 Military Institute databases (http://www.geoportaligm.gob.ec), as well as local and
141 national government technical reports and institutional data (Burbano et al., 2015;
142 Cabrera, 2008; MAE-PRAS, 2016; MAGAP-SIGTIERRAS, 2015). The data are
143 summarised in Table 1.

144 *2.3. Vulnerability assessment*

145 Groundwater vulnerability was calculated using the modified DRASTIC index
146 (DI) using a 90 m × 90 m pixel size, as recommended by Jaderne Houssou et al.,
147 (2019) for cropland. The DI is a standardised model built upon predefined weights
148 derived from expert knowledge and corresponding rating scores according to the
149 contamination potential in a given area (Aller et al., 1987). Rating scores, depending
150 on the degree of vulnerability, may be categorised from low to high on a numeric scale
151 of 1–10. The DI was calculated using Equation (1):

152 $$DI = Dr \times Dw + Rr \times Rw + Ar \times Aw + Sr \times Sw + Tr \times Tw + Ir \times Iw + Cr \times Cw \quad (1)$$

153 where the capital letters indicate the respective parameter and subscripts ‘r’ and ‘w’
154 refer to its rating (based on a range of local information for the area being evaluated)
155 and weight (importance of the parameter), respectively. *D:* depth to the water table;
156 *R:* final recharge; *A:* aquifer media; *S:* soil type; *T:* topography; *I:* impact on the
157 unsaturated zone; *C:* normally hydraulic conductivity, but as it was not available, the
158 infiltration rate was used.

159 To define the rating scores in the NEA, the primary modification methods were:
160 (1) using the four-parameter sub-model of Seabra et al. (2009) for the *R* parameter
161 (Fig. A2.); and (2) replacing hydraulic conductivity with the infiltration rate. The
162 second modification was plausible because the depth to the water table is > 10 m in
163 most areas of the NEA, and pits have a maximum depth of 6 m. It was assumed that
164 water flows in a vertical direction at shallow depths, with discontinuities in areas of
165 subsurface and surface water (Miguez-Macho & Fan, 2012) (Fig. A.5). An extended
166 explanation of each parameter (as defined in the original model), along with its
167 ranking and assigned weight, is provided in Table A1 and Figs. A1–A6 (Aller et al.,
168 1987; Babiker et al., 2005; Merchant, 1994).

169 A sensitivity analysis was performed using map-removal (Lodwick et al., 1990)
170 and single-parameter analyses (Napolitano & Fabbri, 1996) to i) determine the
171 influence of parameter ratings and weights in the final index (Appendix A, Table A2–
172 A3), ii) confirm whether a similar output could be attained using fewer parameters,
173 and iii) depict variations among parameters (Babiker et al., 2005; Kazakis &
174 Voudouris, 2015).

175 *2.4. Hazard assessment and risk mapping*

176 In the present study, hazard refers to the probability of an event that can harm
177 a given environmental asset (Gleyze, 2002). The hazard potential was defined using
178 two datasets on the Total Petroleum Hydrocarbon (TPH) concentration as a surrogate
179 for the TPH mass content of a pit that is eventually able to flow into groundwater: (1)
180 a dataset on 87 sampled pits of the Amazon Defence Front (ADF) (Table 2) (field
181 campaign during 1995–1997) (Cabrera, 2008); and (2) a dataset on the chemical

182 composition of waste deposited in the AUCA-08 pit (field campaign in 2016). The
183 procedure and chemical compositions are described in Appendix A (Table A.4).
184 Despite being older, the ADF data were integrated for potential contaminant
185 propagation and risk mapping because the sampling effort was higher than in this
186 study (Cabrera, 2008). The median TPH concentration value was retained for
187 extrapolation and discussion. Descriptive statistics, e.g., median, quartiles and
188 coefficient of variation, were used to compare the two datasets.

189 Historically, the volumetric dimensions of pits have experienced physical
190 alterations, as summarised in Table 2. In addition, several remediation processes have
191 been implemented in the NEA (MAE-PRAS, 2016); however, a large number of pits
192 have been insufficiently remediated as they were only covered by a thin layer of clean
193 soil (1 m thick) without eliminating the residues, while others remained unlined
194 (MAE-PRAS, 2016). Consequently, the study area is covered with different soil
195 substrates that represent a gradient from solid soil to oil-like media (Amazon Defence
196 Front, 2008), which need to be addressed by first calculating the TPH content of each
197 pit based on its characteristics. To this end, a procedure should be implemented by
198 incorporating the density of the substrate (Sarah et al., 2015). The contents of a pit
199 gradually fluctuate between oil and soil muddy mixtures; thus, the TPH content is
200 difficult to measure for a single pit and is useless at a provincial scale. Following the
201 ADF report (Cabrera, 2008), it was assumed that each pit consists of 95% solid soil-
202 like material and 5% liquid oil-like material.

203 A density-weighted equation was used to estimate the TPH mass in each pit
204 from the pit volume (assuming a volume-to-volume oil proportion of 5% in the pit):

$$TPH\ (kg) = TPH_{Q2} \times \left((Sg\rho \times 0.05 + P\text{b} \times 0.95) \times V \times 10^{-6}\right) \quad (2)$$

where TPH_{Q2} (mg kg^{-1}) is the median TPH content (7519 mg kg^{-1}) stored in pits (Cabrera, 2008); *P*b (kg m^{-3}) is the average soil bulk density from the literature (Saxton & Rawls, 2006), which was assigned to pits using the soil type layer that overlaps the pit; $Sg\rho$ (kg m^{-3}) is the specific oil density, which was calculated using the corresponding American petroleum index (API), which is a measure of the petroleum density in each pit and enables the calculation of specific gravity (Sg) through the expression Sg = 141.5/(API + 131.5) (National Board of Hydrocarbons, 2016); and V (m^3) is the volume of the pit.

Table 2.

Descriptive statistics for the volumetric dimensions and TPH content of pits. Statistics were used to extrapolate data for pits with missing data (shown in bold).

Statistic	**TPH (mg kg^{-1})**	**Area (m^2)**	**Volume (m^3)**	**Depth (m)**
Minimum	4	11.55	30	0.16
1st quartile	2809	241.55	585.73	2
Median	**7519**	**500**	**2000**	**2.19**
3rd quartile	21 749	977.92	2000	3.3
Maximum	900 000	7114	18 432	10
Average	56 322	835.7	1979.48	2.79
S.D.	170 472	914.48	2163.1	2.66
Variation coefficient	3.03	1.09	1.09	0.95
Number of pits with missing data	1333	144	1,200	1298
Number of pits with data	87	1276	220	122
Total number of pits	1420	1420	1420	1420

*A linear regression was used to fill in missing values for area and depth (r^2 = 0.40, n = 220).

Pollutant transport depends on the physicochemical characteristics of hydrocarbons and contamination sources (Fried et al., 1979; Kaasschieter, 1999; Ngo et al., 2014). Generally, hydrocarbons are hydrophobic pollutants, in which low-molecular-weight polycyclic aromatic and volatile hydrocarbons (e.g. benzene, toluene, and xylene) are able to dissolve in water; however, pits can also contain heavy metals (e.g. Ba, Cd, Hg, Ni, Mn, and V), which may also be soluble in water (Maurice et al., 2019). Thus, we assumed that the concentrations of the above-mentioned soluble contaminants were derived from the TPH content of a pit, coupled with the volume of the pit (Table 2) and generic data of physicochemical characteristics (Table 1).

In the absence of detailed and exhaustive data on the fate of hydrocarbons in the NEA, the contaminating plume concept considers the pollutant flow rate to be similar to that of water at a provincial scale; therefore, it propagates along a downward gradient flow direction. Likewise, the contamination plume derived from these characteristics was mapped using 'maximum distance ranges' (MDRs), that is, semi-quantitative scenarios (from 500 m to 10 km) indicating the surface potentially reached by water-transported pollutants until the pollutants are immobilised or cease to cause harm. Based on the importance of initial sources (i.e. TPH contents of pits), this method assesses and maps the relative concentrations of potential pollutants at any point in a given area, without the need to define the proportion of contaminants mobilised (Khan et al., 2018; Zhang et al., 2018). To determine the spatial extent of a hazard, a cost-distance analysis was conducted (Kaffa et al., 2021; Udoh & Ekanem, 2011; Verutes et al., 2020). This method calculates the least accumulative cost-

241 distance for each pixel to the nearest source over a cost surface, given a maximum
242 distance (ESRI, 2013). The contamination originating from a pit depends on the
243 pollutant content of the pit and the distance downstream that the pollutant plume can
244 ultimately reach. Indeed, soil water leakage and runoff propagate along a downward
245 gradient (the cost surface), whereby the intensity decreases with increased distance
246 from the source. The MDRs were reclassified into classes from 'low' (the most
247 distant) to 'high' (the closest) by considering that the hazard increases with increased
248 proximity to the pollutant source, as water leakage and runoff are assumed to be the
249 major dispersion processes.

250 Risk was calculated as a function of the probability of a hazardous source
251 being present and the threat it represents (Gleyze, 2002; Metzger et al., 2008), by
252 multiplying the hazard propagation at a user-defined distance by the vulnerability
253 value. The two maps were reclassified from 'low' to 'high' for each pixel following
254 the Jenks Natural Breaks Optimisation method. All treatments were implemented
255 using ArcGIS®. Four MDRs (from 500 m to 10 km) with contrasting spatial patterns
256 of risk were selected for analysis after several MDRs were explored.

257 ## *3. Results*

258 ### *3.1 Influence of groundwater parameters on vulnerability*

259 The rating scores represent the influence of individual parameters on the
260 DRASTIC vulnerability index across the NEA (Fig. 2). Rating scores are one factor
261 determining vulnerability that considers the NEA-specific conditions; the other factor
262 is composed of weights from the empirical/adapted model. The parameters

263 contributing to the vulnerability are shown in Fig. 2. The overall vulnerability index
264 depends on the multiplication of the rating scores and weights defined by the model.
265 Thus, medium-high, and high vulnerability ratings are strongly influenced by soil-
266 media parameters. After the weights are included, the relative importance of the
267 medium ratings is given by water-related parameters (Fig. 3). Detailed results on
268 weights and the sensitivity analysis can be found in Appendix A (Tables A2 and A3).

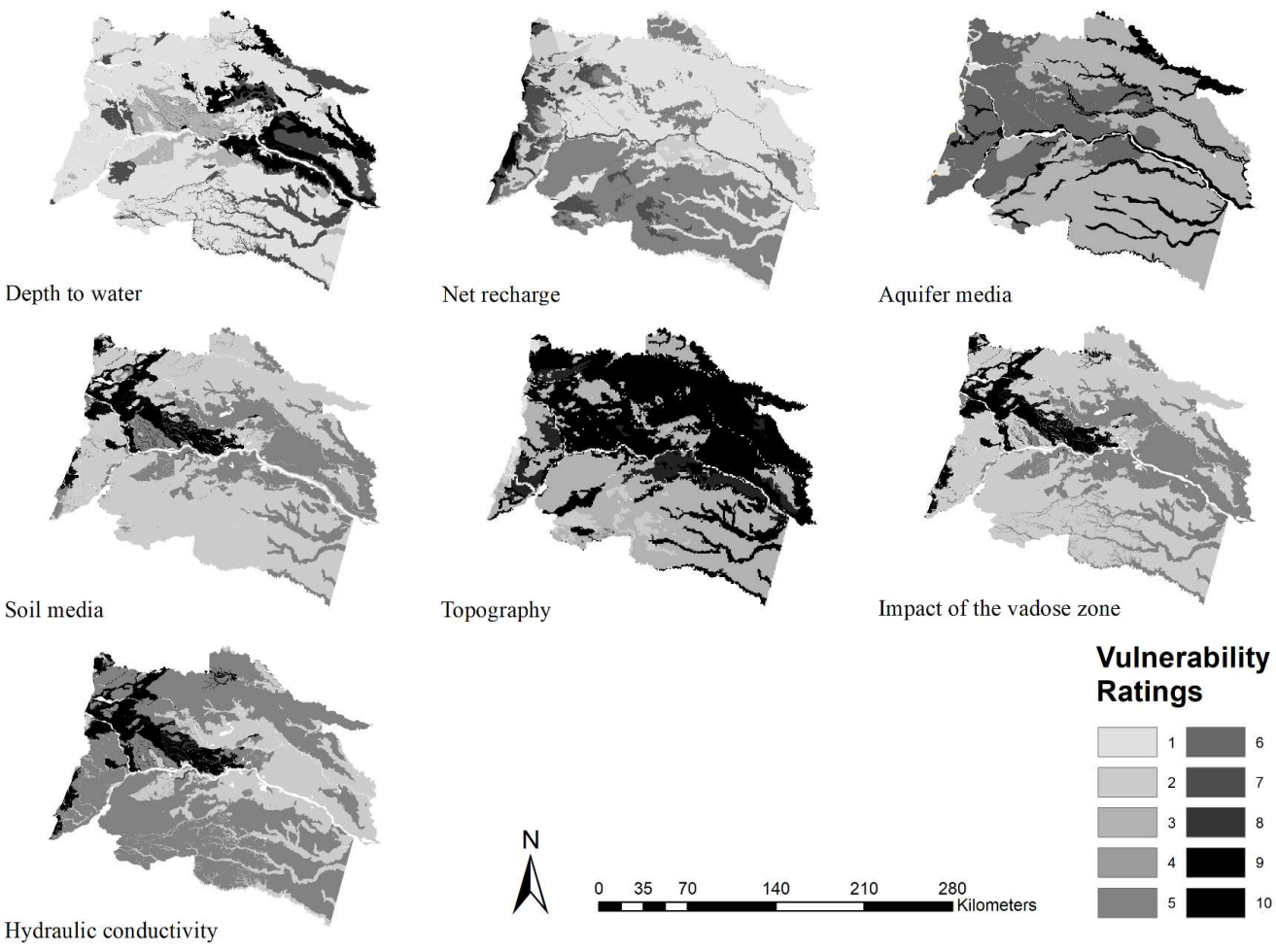


Fig. 2. Impact of the seven parameters used in the modified DRASTIC model contrasted in terms of their individual rating scores.

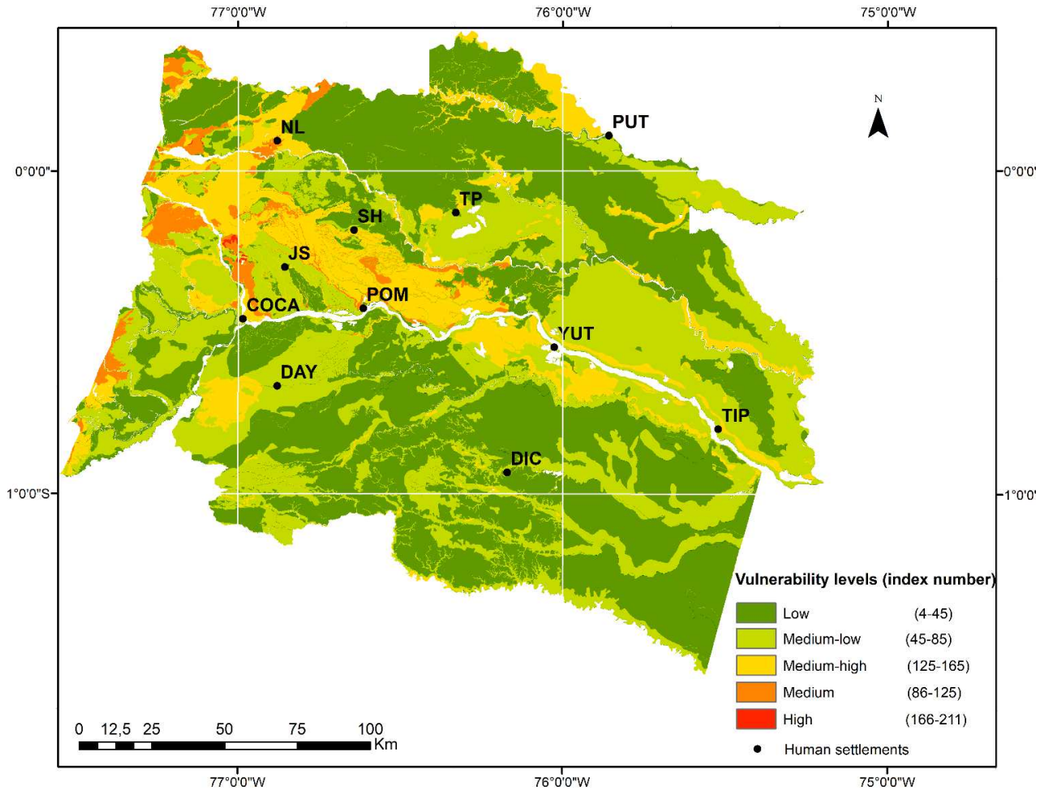


Fig. 3. Vulnerability levels and corresponding classes for the NEA using equal intervals. Areas corresponding to high, medium-high, medium, medium-low, and low vulnerabilities represent 0.06% (22 km^2), 2.7% (909 km^2), 13.05% (4435 km^2), 35.9% (12 120 km^2), and 48.28% (16 389 km^2), respectively. Nueva Loja (NL), Shushufindi (SH), Coca, Pompeya (POM) are in the highly vulnerable zones, followed by Tarapoa (TP) and Tiputini (TIP). The abbreviations for human settlements follow those in Fig. 1.

3.2 Risk mapping

The cost-distance analysis provides the potential distance of the contamination plume, which describes the transport of pollutants from the source (oil pit). This procedure can map different situations. In this study, to simulate the risk posed by a contamination plume, four risk maps of potential MDRs based on the TPH mass contents of pits were retained for illustration purposes (Figs. 4A–4D). The potential

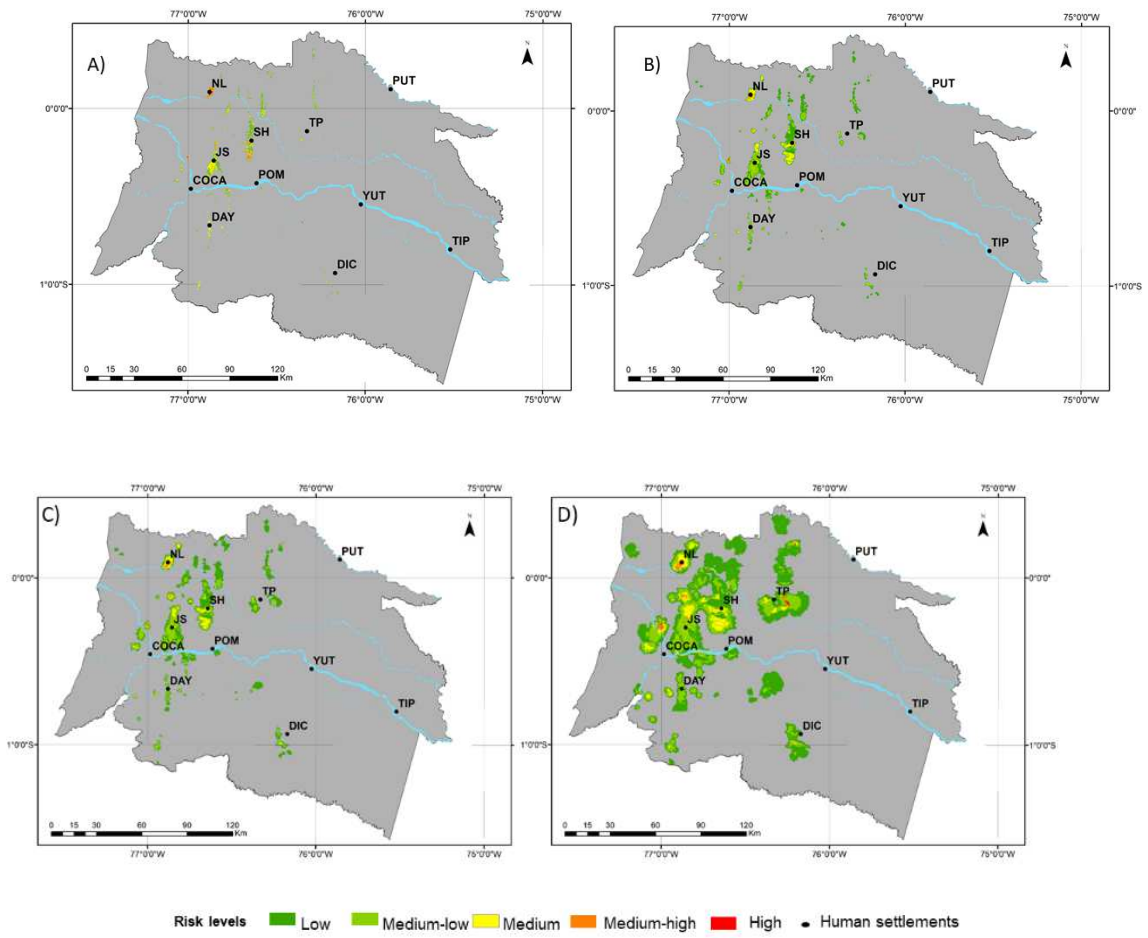
contaminated surface increased as the length of the MDR increased (Fig.4). In contrast, Table 3 shows that the change in the surface at each risk level was not necessarily related to the MDR.

The different risk levels determine the surface extent and change in the function of the downslope and flow direction of the local territory, that is, the surface at higher risk levels (shown in red on the maps) decreased as part of the MDR, but low to moderate MDRs constantly increased the impacted surface. Compared with the other risk categories, the 'intermediate' MDRs (Fig. 4B and 4C) were associated with larger surface in regions with low–moderate risk levels relative to the total risk surface.

The 'closest distance' MDR (Fig. 4A) indicates that the risk hotspots with the highest level of contamination were found in Nueva Loja, while medium- to high-risk zones were dispersed near Shushufindi and Joya de los Sachas. However, in the intermediate MDRs, hotspots were also observed to the south of Shushufindi, in Tarapoa, to the north of Joya de los Sachas, and to the east of El Coca.

For the 'longest' MDR, the results indicate a low to moderate risk of contamination over 13% of the NEA surface. The surface extents of the risk zones for our maps are presented in Table 3.

302



303

304 **Fig. 4.** Risk maps for pollutant discharge from pits into groundwater as a function of the maximum distance transport: a) closest distance to source, MDR of 500
305 m and b) 2000 m; c) intermediate distance (MDR of 5000 m); and d) longest path of contamination, MDR of 10 000 m.

Table 3.

Surface (km^2) corresponding to each risk level, depending on the maximum distance of the plume. The figures illustrating each risk MDR are indicated.

	Distance (m)			
Category	**Closest**	**Intermediate**		**Longest**
Risk level	**500**	**2000** (Δ)	**5000** (Δ)	**10 000** (Δ)
Low	53	421 (368)	866 (445)	2871 (2005)
Moderate low	99	252 (153)	424 (172)	1315 (891)
Moderate	85	80 (5)	117 (37)	322 (205)
Moderate high	29	13 (-16)	20 (7)	57 (37)
High	5	0.5 (-4.5)	0.2 (-4.3)	1.3 (1.1)
Total surface at risk (km^2)	271	766.5	1427.2	4566.7
Figures	4A	4B	4C	4D

4. Discussion

4.1. Risk components: vulnerability and hazard

The overall vulnerability was medium-low, which corresponds well with the highly impervious and argillaceous characteristics of Amazonian soils. These characteristics can reduce soil permeability owing to the aquifer formation differences between sandstone (Curaray) and quartz pebble bearing (Chambira) matrices. In this study, vulnerability was mostly sensitive to the combined effects of soil parameters. The infiltration rate explained 71% of the potential information loss, probably because hydraulic conductivity was not integrated (effective weight/empirical weight ratio of 0.71). The sensitivity analysis indicated that the groundwater vulnerability assessment was robust, suggesting that the available quantitative data was integrated into the model in a satisfactory manner. The intricacies of the Drastic model and sensitivity analysis are detailed in Appendix A.

322 The impact of the hazard assessment on the overall risk estimates is determined
323 by the estimated TPH content per pit, which was 3859.6 t (median = 14.57 t; minimum
324 = 7.18 kg; maximum = 5712 t) and 39 052.3 t (median = 14.57 t; minimum = 7.18 kg;
325 maximum = 5712 t) for the 2016 and 1997 datasets, respectively. In official data for
326 1997, the oil-related discharge amount was 48 183 t, which is ~20% higher than our
327 TPH content for 1997 (MAE-PRAS, 2016). This difference may partly relate to the
328 fact that the official value refers to crude oil, while this study considers only
329 hydrocarbons. Only a fraction of these quantities can propagate because of the low
330 solubility of heavier crude oil. Maurice et al., (2019) found that only the toluene
331 concentration was high in deep groundwater wells. Additionally, the pollutant fraction
332 that could eventually reach the water table, as indicated by the MDRs, would be more
333 accurately defined using transport-related data of oil-associated contaminants (Table.
334 A2). Our methodology was not sensitive to inaccuracies in the oil content estimation
335 because the cost-distance analysis, which emphasises the relative risk, primarily
336 depends on the volume and location of a pit (agreeing with the TPH mass content
337 evaluation conducted in this study) based on combining semi-quantitative data to map
338 several risk scenarios (La Manna et al., 2013). The number of pits with information
339 on the hydrocarbon content (6%) and volume characteristics (i.e. surface (92%),
340 volume (15%), and depth (8%)) suggests that the accuracy of the hazard assessment
341 could improve with permanent data collection schemes (Durango-Cordero et al.,
342 2018; MAE-PRAS, 2016).

4.2. Land use planning implications

MDR mapping is useful for decision-makers to anticipate the environmental impact of spatial planning (Lahr & Kooistra, 2010). The cost-distance analysis results showed that while the surface at risk increased, the proportion of high-risk surface decreased; thus, a nonlabile contaminant will have a higher concentration but affect a smaller surface than a labile contaminant. The relative delta obtained from the MDRs (i.e. percentage of surface change from one MDR to the next) indicated that surface with low to with moderately-low levels of risk increased, while surface moderately-high levels of risk remained similar (Table 3). This finding might relate to the fact that micro-slopes act as barriers throughout the Ecuadorian Amazonian basin, and/or to the intrinsic properties of the algorithm used.

Although TPH samples were evaluated using standardised methods (Cabrera, 2008), their application has been internationally criticised because of the lack of scientific rigor (Buccina et al., 2013). The extrapolated TPH concentration of the MONOIL-project protocol (2016) was approximately 7 times lower than the observed with ADF data, i.e., 30.4 for the Auca-08 pit versus 209.7 for ADF data. Although this finding does not indicate that the ADF data are reliable or valid for the TPH concentration, it suggests an expected time decay in the TPH concentration. It is reasonable to have a lower TPH content after 19 years. Spatial planners and local decision-makers can use locally collected data for risk assessment when certain criteria are met; for example, based on the social trust of local regulatory institutions, recognition of local communities, and *post hoc* environmental impact assessments (Bronfman et al., 2009; Gómez, 2013; Otieno-odawa & Kaseje, 2014; Schweitzer,

2008). The relatively short MDR of 500 m to 2000 m provides better spatial insights for agriculture and water management, which could impact local public health if improperly planned (Lahr & Kooistra, 2010). This relates to the high hydrophobicity of hydrocarbon compounds, biological and microbial degradation (Maddela et al., 2016; Singh & Chandra, 2014), and organic matter content (MAGAP-SIGTIERRAS, 2015), and because the capacity of the MDR for hydrocarbons is reduced by soil adsorption in Amazonian environments. In addition to the inconclusive results regarding the correlation between the incidence of disease (e.g. cancer) and oil pollution, controversies also remain regarding spatial exposure to oil activities (Hurtig & San Sebastián, 2002; Moolgavkar et al., 2014). It is advisable to communicate potential pollution in the medium-high–high risk surface of the 500 m MDR. This issue should be given special attention with respect to agriculture, specifically cacao plantations, whose grain quality might be affected by cadmium uptake (Barraza et al., 2017).

4.3. Limitations and next steps

The spatial coincidence of the soil-related parameters indicated that the total vulnerability was influenced by their cumulative mean variation (cMVI = 3.5%), rather than their individual weighted factors, which were low (evaluated in the sensitivity analysis). These are similar subsurface texture zones and are spatially co-dependent on soil type. There was a lack of data to rank the parameter for the depth to the water table in some zones with a ‘considerable’ depth to the water table (>10 m), which probably relates to the use of shallow monitoring systems (MAGAP-SIGTIERRAS, 2015). More accurate data would improve this data-driven approach

389 across Ecuadorian landscapes (Ribeiro et al., 2017), and knowledge of groundwater
390 vulnerability under human activities. Only 6.5% (87 out of 1337) of the studied oil
391 pits had relevant data (Table 2), and only four aquifers were rigorously identified (Fig.
392 1B), representing 2% of the NEA. Aquifers in the Napo basin showed variable average
393 flow rates of 0.89–13 mL s^{-1}, with a water depth of up to 55 m (SENAGUA-ESPOL,
394 2017). The identification of other aquifers should continue to enhance vulnerability
395 analyses in future research. To better explore the importance of all existing parameters
396 and variables in the vulnerability assessment and to make this study more informative,
397 we used a modified DRASTIC model instead of more simplified methods such as the
398 SI (Ribeiro et al., 2017), GOD, and AVI methods (Merchant, 1994; Secunda et al.,
399 1998; Shrestha et al., 2017). Social vulnerability should also be considered in future
400 studies to obtain a more integrated perspective (Frigerio et al., 2016).

401 The information provided in this study can be used to identify vulnerable
402 surface that could be investigated and used for validation. Selecting sites including
403 old oil pits within different risk zones for research at finer spatial scales would be a
404 step forward. As data become available for assessing aquifer vulnerability factors (e.g.
405 porosity, hydraulic conductivity, hydraulic gradient, and groundwater velocity)
406 (Kazakis & Voudouris, 2015) and factors related to the physicochemical properties of
407 contaminants (e.g. cross-sectional area, partition fractions, biodegradation,
408 sedimentation, and chemical complexion), a multi-phase flow approach using in situ
409 measurements would benefit from the use of specialised software (e.g. Mod-flow and
410 Hydrus-2D) (Ngo et al., 2014) and stochastic techniques (e.g. the Markov chain
411 model) (Weigand et al., 2001). It would be informative to apply such models directly

412 downstream of the monitored surface to evaluate the probability of pollutant leaching
413 into specific water tables due to pit leakage.

414 *5. Conclusion*

415 This study constructively evaluated oil pit-associated risks of groundwater
416 contamination for integration in land use planning within the NEA. An aquifer
417 vulnerability map for the NEA was constructed with respect to any contaminant (i.e.
418 hazard), and then the potential pollution from oil pits was estimated. In addition, the
419 spatial risk assessment proposed different MDR values for potential pollutant
420 propagation. Based on a combination of pit locations within different risk zones,
421 several potential sampling sites could be selected using these risk maps depending on
422 the solubility and lability of hydrocarbons and other oil-associated pollutants
423 measured in the TPH analysis. The risk map may be combined with information on
424 other human activities and social variables (Frigerio et al., 2016), for instance,
425 overlaying it with a land use map (Al-Adamat et al., 2003; Rahman, 2008; Secunda et
426 al., 1998).

427 This study was based on a large spatial dataset, physicochemical properties of
428 soil, and hydrocarbon characteristics to improve the contamination risk assessment of
429 aquifers. Future research could integrate refined weights, and data gathering may
430 improve risk mapping. Local decision-makers and spatial planners may avoid
431 potential adverse impacts within high-risk zones at a MDR of 500 m or 2000 m by
432 encouraging further investigations in these surfaces, including validation of
433 contaminant plumes. Land use plans should be revised in terms of potentially

relocating activities from medium-high to high-risk surfaces designated for agriculture, increasing potable water services, and intensifying risk communication to improve public health.

Acknowledgements

The authors wish to thank the officials from PRAS and ADF, and the communities in the NEA who provided advice during field visits to improve the quality of this study. The authors acknowledge GIS assistance from Edna Melo-Ramirez and Barbara Romeo. The English language was checked and revised by the professional services of American Journal Experts (AJE). Many thanks to the EPOC team, Hélène Budzinski, Karyn Le Menach, and Marie Hélène Devier in Bordeaux (France) for analysing the PAH concentrations of pit soils.

Funding

This study was funded and conducted within the ANR-MONOIL French Research Programme (ANR-13-SENV-0003-01). PhD of Dr. Durango-Cordero was funded by the Ecuadorian National Secretary of Higher Education, Science, and Technology (SENESCYT).

References

Al-Adamat, R. A. N., Foster, I. D. L., & Baban, S. M. J. (2003). Groundwater vulnerability and risk mapping for the Basaltic aquifer of the Azraq basin of Jordan using GIS, Remote sensing and DRASTIC. *Applied Geography*, *23*(4), 303–324. https://doi.org/10.1016/j.apgeog.2003.08.007

Aller, L., Lehr, J. H., & Petty, R. (1987). *DRASTIC: A standardized Sytem to Evaluate Ground Water Pollution Using Hydrogeologic Settings* (pp. 38–57).

Amazon Defence Front. (2008). *Breve Historia de las Operaciones de Texaco en el Oriente Ecuatoriano 1964 -1990*.

Andreo, B., Goldscheider, N., Vadillo, I., Vías, J. M., Neukum, C., Sinreich, M., Jiménez, P., Brechenmacher, J., Carrasco, F., Hötzl, H., Perles, M. J., & Zwahlen, F. (2006). Karst groundwater protection: First application of a Pan-European Approach to vulnerability, hazard and risk mapping in the Sierra de Líbar (Southern Spain). *Science of the Total Environment*, *357*(1–3), 54–73. https://doi.org/10.1016/j.scitotenv.2005.05.019

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*(1–3), 127–140. https://doi.org/10.1016/j.scitotenv.2004.11.005

Barraza, F., Maurice, L., Uzu, G., Becerra, S., López, F., Ochoa-Herrera, V., Ruales, J., & Schreck, E. (2018). Distribution, contents and health risk assessment of metal(loid)s in small-scale farms in the Ecuadorian Amazon: An insight into impacts of oil activities. *Science of the Total Environment*, *622–623*, 106–120. https://doi.org/10.1016/j.scitotenv.2017.11.246

Barraza, F., Schreck, E., Lévêque, T., Uzu, G., López, F., Ruales, J., Prunier, J., Marquet, A., & Maurice, L. (2017). Cadmium bioaccumulation and gastric bioaccessibility in cacao: A field study in areas impacted by oil activities in Ecuador. *Environmental Pollution*, *229*, 950–963. https://doi.org/10.1016/j.envpol.2017.07.080

Barraza, F., Uzu, G., Jaffrezo, J. L., Schreck, E., Budzinski, H., Le Menach, K., Dévier, M. H., Guyard, H., Calas, A., Perez, M. I., Villacreces, L. A., & Maurice, L. (2020). Contrasts in chemical composition and oxidative potential in PM10 near flares in oil extraction and refining areas in Ecuador. *Atmospheric Environment*, *223*, 117302. https://doi.org/10.1016/j.atmosenv.2020.117302

Bronfman, N. C., López, E., & Dorantes, G. (2009). An empirical study for the direct and indirect links between trust in regulatory institutions and acceptability of hazards. *Safety Science*, *47*(5), 686–692. https://doi.org/10.1016/j.ssci.2008.09.006

Buccina, S., Chene, D., & Gramlich, J. (2013). Accounting for the environmental impacts of Texaco's operations in Ecuador: Chevron's contingent environmental liability disclosures. *Accounting Forum*, *37*(2), 110–123. https://doi.org/10.1016/j.accfor.2013.04.003

Burbano, N., Becerra, S., & Pasquel, E. (2015). *Introducción a la Hidrogeología del Ecuador* (Vol. 91).

http://www.serviciometeorologico.gob.ec/Publicaciones/Hidrologia/HIDROGEOLOGIA_2 EDICION_2014.pdf

Busico, G., Kazakis, N., Cuoco, E., Colombani, N., Tedesco, D., Voudouris, K., & Mastrocicco, M. (2020). A novel hybrid method of specific vulnerability to anthropogenic pollution using multivariate statistical and regression analyses. *Water Research*, *171*, 115386. https://doi.org/10.1016/j.watres.2019.115386

Cabrera, R. (2008). *Extrapolation of sampling results from all sites within the Texaco-Chevron oil concession blocs*.

Coronel Vargas, G., Au, W. W., & Izzotti, A. (2020). Public health issues from crude-oil production in the Ecuadorian Amazon territories. In *Science of the Total Environment* (Vol. 719, p. 134647). Elsevier B.V. https://doi.org/10.1016/j.scitotenv.2019.134647

Durango-Cordero, J., Saqalli, M., & Parra, R. & Elger, A. (2019). Spatial inventory of selected atmospheric emissions from oil industry in Ecuadorian Amazon: insights from comparisons among satellite and institutional datasets. *Journal of Safety Science*.

Durango-Cordero, Saqalli, M., Laplanche, C., Locquet, M., & Elger, A. (2018). Spatial analysis of accidental oil spills using heterogeneous data: A case study from the North-Eastern Ecuadorian Amazon. *Sustainability (Switzerland)*, *10*(12). https://doi.org/10.3390/su10124719

EPA. (2014). *Compilation of Publicly Available Sources of Voluntary Management Practices for Oil and Gas Exploration & Production (E&P) Wastes As They Address Pits, Tanks, and Land Application*. https://www.epa.gov/hw/proper-management-oil-and-gas-exploration-and-production-waste

ESRI. (2013). *Desktop Help 10.5 - How cost distance tools work*. http://desktop.arcgis.com/en/arcmap/10.3/tools/spatial-analyst-toolbox/how-the-cost-distance-tools-work.htm

Fried, J. J., Muntzer, P., & Zilliox, L. (1979). Groundwater Pollution by Transfer of Oil Hydrocarbons. *Groundwater*, *17*(6), 586–596. https://onlinelibrary.wiley.com/doi/epdf/10.1111/j.1745-6584.1979.tb03359.x

Frigerio, I., Ventura, S., Strigaro, D., Mattavelli, M., De Amicis, M., Mugnano, S., & Boffi, M. (2016). A GIS-based approach to identify the spatial variability of social vulnerability to seismic hazard in Italy. *Applied Geography*, *74*, 12–22. https://doi.org/10.1016/j.apgeog.2016.06.014

Gleyze, J. F. (2002). *Le Risque* (Insititut Geographique National (ed.)). Insititut Geographique National. http://recherche.ign.fr/labos/cogit/pdf/RAPPORTS/Gleyze_rapport_risque.pdf

Gogu, R. ., & Dassargues, A. (2000). Current and future trends in groundwater vulnerability assessment using overlay and index methods. *Environmental Geology*, *39*(April), 549–559. https://doi.org/https://doi.org/10.1007/s002540050466

Gómez, M. A. (2013). The Global Chase : Seeking the Recognition and Enforcement of the Lago Agrio Judgment Outside of Ecuador. *Stan. J. Complex Litig*, *1*(429), 429. http://ecollections.law.fiu.edu/faculty_publications%0Ahttp://ecollections.law.fiu.edu/faculty_publications/42

Hamza, M. H., Added, A., Rodríguez, R., Abdeljaoued, S., & Ben Mammou, A. (2007). A GIS-based DRASTIC vulnerability and net recharge reassessment in an aquifer of a semi-arid region (Metline-Ras Jebel-Raf Raf aquifer, Northern Tunisia). *Journal of Environmental Management*, *84*(1), 12–19. https://doi.org/10.1016/j.jenvman.2006.04.004

Hurtig, A. K., & San Sebastián, M. (2002). Geographical differences in cancer incidence in the Amazon basin of Ecuador in relation to residence near oil fields. *International Journal of Epidemiology*, *31*(5), 1021–1027. https://doi.org/10.1093/ije/31.5.1021

Jaderne Houssou, N. L., Durango-Cordero, J., Bouadjio-Boulic, A., Morin, L., Maestripieri, N., Ferrant, S., Belem, M., Peláez, J. I., Saenz, M., Lerigoleur, E., Elger, A., Gaudou, B., Maurice, L., & Saqalli, M. (2019). Synchronizing histories of exposure and demography: The construction of an agent- based model of the ecuadorian amazon colonization and exposure to oil pollution hazards. *Jasss*, *22*(2). https://doi.org/10.18564/jasss.3957

Kaasschieter, E. F. (1999). Solving the Buckley – Leverett equation with gravity in a heterogeneous porous medium. *Computational Geosciences*, *3Shresh*, 23–48. https://doi.org/10.1023/A:1011574824970

Kaffa, N. S., Sukojo, B. M., & Handayani, H. H. (2021). Tiga Dihaji Dam Access Route Plan Alternative using Geographic Information System (Cost Distance Method). *IOP Conference Series: Earth and Environmental Science*, *731*(1), 012007. https://doi.org/10.1088/1755-1315/731/1/012007

Kazakis, N., & Voudouris, K. S. (2015). Groundwater vulnerability and pollution risk assessment of porous aquifers to nitrate: Modifying the DRASTIC method using quantitative parameters. *Journal of Hydrology*, *525*, 13–25. https://doi.org/10.1016/j.jhydrol.2015.03.035

Khan, M. A. I., Biswas, B., Smith, E., Naidu, R., & Megharaj, M. (2018). Toxicity assessment of fresh and weathered petroleum hydrocarbons in contaminated soil- a review. In *Chemosphere* (Vol. 212, pp. 755–767). Elsevier Ltd. https://doi.org/10.1016/j.chemosphere.2018.08.094

La Manna, L., Greslebin, A. G., & Matteucci, S. D. (2013). Applying cost-distance analysis for forest disease risk mapping: Phytophthora austrocedrae as an example. *European Journal of Forest Research*, *132*(5–6), 877–885. https://doi.org/10.1007/s10342-013-0720-3

Lahr, J., & Kooistra, L. (2010). Environmental risk mapping of pollutants: State of the art and communication aspects. *Science of the Total Environment*, *408*(18), 3899–3907. https://doi.org/10.1016/j.scitotenv.2009.10.045

Laraque, A., Ronchail, J., Cochonneau, G., Pombosa, R., & Guyot, J. L. (2007). Heterogeneous distribution of rainfall and discharge regimes in the Ecuadorian Amazon basin. *Journal of Hydrometeorology*, *8*(6), 1364–1381. https://doi.org/10.1175/2007JHM784.1

Lodwick, W. A., Monson, W., & Svoboda, L. (1990). Attribute error and sensitivity analysis of map operations in geographical informations systems: Suitability analysis. *International Journal of Geographical Information Systems*, *4*(4), 413–428. https://doi.org/10.1080/02693799008941556

Maddela, N., Burgos, R., Kadiyala, V., Carrion, A., & Bangeppagari, M. (2016).

Removal of petroleum hydrocarbons from crude oil in solid and slurry phase by mixed soil microorganisms isolated from Ecuadorian oil fields. *International Biodeterioration & Biodegradation*, *108*, 85–90. https://doi.org/10.1016/J.IBIOD.2015.12.015

MAE-PRAS. (2016). *Pasivos ambientales y reparación ambiental integral: Experiencias de gestion en el Ecuador* (2nd ed.). Ministerio del Ambiente. http://www.ambiente.gob.ec/programa-de-reparacion-ambiental-y-social-pras/#

MAGAP-SIGTIERRAS. (2015). *Manual of metodologies for the geopedological study and mapping at scale of 1:25 000.* http://metadatos.sigtierras.gob.ec/pdf/Metodologia_Geopedologia_16122015.pdf

Maurice, L., López, F., Becerra, S., Jamhoury, H., Le, K., Dévier, M., Budzinski, H., Prunier, J., Juteau-martineau, G., Ochoa-herrera, V., Quiroga, D., & Schreck, E. (2019). Drinking water quality in areas impacted by oil activities in Ecuador: Associated health risks and social perception of human exposure. *Science of the Total Environment*, *690*, 1203–1217. https://doi.org/10.1016/j.scitotenv.2019.07.089

Merchán-Rivera, P., Chiogna, & Gabriele. (2017). *Assessment of contamination by petroleum hydrocarbons from oil exploration and production activities in Aguarico , Ecuador* .

Merchant, J. W. (1994). GIS-Base Groundwater Pollution Hazard Assessment: Critical Review of the DRASTIC Model. *Photogrammetric Engineering & Remote Sensing*, *60*(9), 1117–1127. https://www.asprs.org/wp-content/uploads/pers/1994journal/sep/1994_sep_1117-1127.pdf

Metzger, P., Ercole, R. D., Metzger, P., & Ercole, R. D. (2008). Enjeux territoriaux et vulnerabilité : une approche operationnelle. *Colloque Interdisciplinaire " Vulnerabilites Sociétales, Risques et Environnement : Comprendre et Les Evaluer", May 2008, Toulouse, France.*

Miguez-Macho, G., & Fan, Y. (2012). The role of groundwater in the Amazon water cycle: 1. Influence on seasonal streamflow, flooding and wetlands. *Journal of Geophysical Research Atmospheres*, *117*(15), 1–30. https://doi.org/10.1029/2012JD017539

Mimi, Z. A., & Assi, A. (2009). Intrinsic vulnerability, hazard and risk mapping for karst aquifers: A case study. *Journal of Hydrology*, *364*(3), 298–310. https://doi.org/10.1016/j.jhydrol.2008.11.008

Mimi, Z. A., Mahmoud, N., & Madi, M. A. (2012). Modified DRASTIC assessment for intrinsic vulnerability mapping of karst aquifers: A case study. *Environmental Earth Sciences*. https://doi.org/10.1007/s12665-011-1252-0

Moolgavkar, S. H., Chang, E. T., Watson, H., & Lau, E. C. (2014). Cancer mortality and quantitative oil production in the Amazon region of Ecuador, 1990-2010. *Cancer Causes and Control*, *25*(1), 59–72. https://doi.org/10.1007/s10552-013-0308-8

Napolitano, P., & Fabbri, A. G. (1996). Single-parameter sensitivity analysis for aquifer vulnerability assessment using DRASTIC and SINTACS. *IAHS Publications-Series of Proceedings and Reports-Intern Assoc Hydrological Sciences*, *235*(235), 559–566.

http://hydrologie.org/redbooks/a235/iahs_235_0559.pdf
National Board of Hydrocarbons. (2016). *Exploration: production subprocess oil wells sampling by oil operator*. http://www.historico.secretariahidrocarburos.gob.ec/banco-de-informacion-petrolera/#
Ngo, V. V, Michel, J., Gujisaite, V., Latifi, A., & Simonnot, M.-O. (2014). Parameters describing nonequilibrium transport of polycyclic aromatic hydrocarbons through contaminated soil columns: estimability analysis, correlation, and optimization. *Journal of Contaminant Hydrology*, *158*, 93–109. https://doi.org/10.1016/j.jconhyd.2014.01.005
Otieno-odawa, C. F., & Kaseje, D. O. (2014). Validity and reliability of data collected by community health workers in rural and peri- urban contexts in Kenya. *BMC Health Services Research*, *14*(Suppl 1), S5. https://doi.org/10.1186/1472-6963-14-S1-S5
Paca, V. H. da M., Espinoza-Dávalos, G. E., Hessels, T. M., Moreira, D. M., Comair, G. F., & Bastiaanssen, W. G. M. (2019). The spatial variability of actual evapotranspiration across the Amazon River Basin based on remote sensing products validated with flux towers. *Ecological Processes 2019 8:1*, *8*(1), 1–20. https://doi.org/10.1186/S13717-019-0158-8
Rahman, A. (2008). A GIS based DRASTIC model for assessing groundwater vulnerability in shallow aquifer in Aligarh, India. *Applied Geography*, *28*(1), 32–53. https://doi.org/10.1016/j.apgeog.2007.07.008
Ramirez, M. I., Arevalo, A. P., Sotomayor, S., & Bailon-Moscoso, N. (2017). Contamination by oil crude extraction – Refinement and their effects on human health. In *Environmental Pollution* (Vol. 231, pp. 415–425). Elsevier Ltd. https://doi.org/10.1016/j.envpol.2017.08.017
Ribeiro, L., Pindo, J. C., & Dominguez-Granda, L. (2017). Assessment of groundwater vulnerability in the Daule aquifer, Ecuador, using the susceptibility index method. *Science of the Total Environment*, *574*, 1674–1683. https://doi.org/10.1016/j.scitotenv.2016.09.004
Rosário, F. F. do, Custodio, E., & Silva, G. C. da. (2016). Hydrogeology of the Western Amazon Aquifer System (WAAS). *Journal of South American Earth Sciences*, *72*, 375–386. https://doi.org/10.1016/j.jsames.2016.10.004
San Sebastián, M., Armstrong, B., Córdoba, J. A., & Stephens, C. (2001). Exposures and cancer incidence near oil fields in the Amazon basin of Ecuador. *Occupational and Environmental Medicine*, *58*(8), 517–522. https://doi.org/10.1136/oem.58.8.517
San Sebastián, M., & Hurtig, A. . (2005). Oil development and health in the Amazon basin of Ecuador: the popular epidemiology process. *Social Science & Medicine (1982)*, *60*(4), 799–807. https://doi.org/10.1016/j.socscimed.2004.06.016
Sarah, a., Julius, U., & Francis, C. (2015). Determination of Some Petrophysical Properties of Reservoir Rocks in the Niger Delta. *Journal of Scientific Research and Reports*, *5*(5), 388–401. https://doi.org/10.9734/JSRR/2015/15096
Saxton, K. E., & Rawls, W. J. (2006). Soil Water Characteristic Estimates by

Texture and Organic Matter for Hydrologic Solutions. *Soil Science Society of America Journal*, *70*(5), 1569. https://doi.org/10.2136/sssaj2005.0117

Schweitzer, L. (2008). Accident frequencies in environmental justice assessment and land use studies. *Journal of Hazardous Materials*, *156*(1–3), 44–50. https://doi.org/10.1016/j.jhazmat.2007.11.125

Seabra, V. S., Da Silva, G. C., & Cruz, C. B. M. (2009). The use of geoprocessing to assess vulnerability on the east coast aquifers of Rio de Janeiro State, Brazil. *Environmental Geology*, *57*(3), 665–674. https://doi.org/10.1007/s00254-008-1345-6

Secunda, S., Collin, M. ., & Melloul, A. . (1998). Groundwater vulnerability assessment using a composite model combining DRASTIC with extensive agricultural land use in Israel's Sharon region. *Journal of Environmental Management*, *54*(1), 39–57. https://doi.org/10.1006/jema.1998.0221

SENAGUA-ESPOL. (2017). *Hidrología y plan hidráulico regional de la demarcación hidrográfica de Napo*.

Shrestha, S., Kafle, R., & Pandey, V. P. (2017). Evaluation of index-overlay methods for groundwater vulnerability and risk assessment in Kathmandu Valley, Nepal. *Science of the Total Environment*, *575*, 779–790. https://doi.org/10.1016/j.scitotenv.2016.09.141

Singh, K., & Chandra, S. (2014). Treatment of Petroleum Hydrocarbon Polluted Environment Through Bioremediation: A Review. *Pakistan Journal of Biological Sciences*, *17*(1), 1–8. https://doi.org/10.3923/pjbs.2014.1.8

Stigter, T. Y., Ribeiro, L., & Dill, A. M. M. C. (2006). Evaluation of an intrinsic and a specific vulnerability assessment method in comparison with groundwater salinisation and nitrate contamination levels in two agricultural regions in the south of Portugal. *Hydrogeology Journal*, *14*(1–2), 79–99. https://doi.org/10.1007/s10040-004-0396-3

Udoh, J., & Ekanem, E. (2011). GIS based risk assessment of oil spill in the coastal areas of Akwa Ibom State, Nigeria. *African Journal of Environmental Science and Technology*, *5*(3), 205–211. https://doi.org/10.4314/ajest.v5i3.71928

US Army Corps. (1998). *Water Resources Assessment of Ecuador US Army Corps of Engineers*. moz-extension://2469b4a6-52f9-4d46-b046-13f11a5cca40/enhanced-reader.html?openApp&pdf=https%3A%2F%2Fwww.sam.usace.army.mil%2FPortals%2F46%2Fdocs%2Fmilitary%2Fengineering%2Fdocs%2FWRA%2FEcuador%2FEcuador%2520WRA%2520English.pdf

Verutes, G. M., Johnson, A. F., Caillat, M., Ponnampalam, L. S., Peter, C., Vu, L., Junchompoo, C., Lewison, R. L., & Hines, E. M. (2020). Using GIS and stakeholder involvement to innovate marine mammal bycatch risk assessment in data-limited fisheries. *PLoS ONE*, *15*(8 August). https://doi.org/10.1371/journal.pone.0237835

Vrba, J., & Zaporozec, A. (1994). *Guidebook on mapping groundwater vulnerability* (International Contributions to Hydrogeology (ed.); 16th ed.). H.Heise. http://agris.fao.org/agris-search/search.do?recordID=US201300287789

Wasserstrom, R. (2013). Deforestation, Agrarian Reform and Oil Development in Ecuador, 1964-1994. *Natural Resources*, *04*(01), 31–44.

https://doi.org/10.4236/nr.2013.41004
Weigand, H., Totsche, K. U., Huwe, B., & Kögel-Knabner, I. (2001). PAH mobility in contaminated industrial soils: a Markov chain approach to the spatial variability of soil properties and PAH levels. *Geoderma*, *102*(3–4), 371–389. https://doi.org/10.1016/S0016-7061(01)00043-X
Zhang, P., Qin, C., Hong, X., Kang, G., Qin, M., Yang, D., Pang, B., Li, Y., He, J., & Dick, R. P. (2018). Risk assessment and source analysis of soil heavy metal pollution from lower reaches of Yellow River irrigation in China. *Science of the Total Environment*, *633*, 1136–1147. https://doi.org/10.1016/j.scitotenv.2018.03.228

735 **Appendix A. Supplementary data**

736

737 **Individual parameters for the modified DRASTIC model**

738

739 **Table A1.** Parameter definition, specific treatment and reference source (Aller et al., 1987)

Parameter	Definition	Specific treatment for the NEA	Reference
Depth to the water table	A measure of depth from the soil surface to the water table. A deeper water table level entails a lower probability for contamination: the deeper the water table, the lesser the rating.	No data, sand banks, and human settlements (1984 data) were excluded. A 'not evident' water-table depth class was reported in data from 2015, which potentially reaches a depth of tens of metres (i.e. deeper than the pit bottom). This was considered the less vulnerable class of five classes (n = 5).	(Hamza et al., 2007) Fig. A.1
Final recharge	Represents the amount of water per unit of land surface that can reach the water table. The main source is precipitation, which serves pollutant transport. The greater the recharge, the greater the contamination potential, and hence the rating value increases.	The recharge was built from a four-weighted parameter sub-model, which integrated the average annual rainfall, soil type, elevation, and topography. Elevation and slope data were extracted from the DEM. Although the sub-model did not estimate a quantity for the recharge value, it could index vulnerability classes (n = 5).	(Seabra et al., 2009) Fig. A.2
Aquifer media	Refers to the saturated zone of soil, where the higher the permeability (larger grain size) and the lower the attenuation potential, the greater the contamination potential.	Permeability (n = 7) and formation types (n = 11) were reclassified into rating scores (n = 6)	Figs. A.3 and A.4
Soil media	Indicates the uppermost portion of the vadose zone with significant biological activity. Additionally, the less clay shrinks and swells, and the	The thicker the grain, the higher the permeability, thus the higher the index rating (n = 3).	(Hamza et al., 2007) Fig. A.5

	smaller the grain size, the less the contamination potential.		

741 **Table A1.** (continued).

Parameter	Definition	Specific treatment for the NEA	Reference
Topography	The ratings were assigned following the premise that the less the percentage slope, the higher the local contamination potential due to the length of time that water and pollutants remained in a particular surface.	No modifications from the original DRASTIC index classes (n = 5).	Fig. A.6
Impact to the vadose zone	Depicted by unsaturated material, which controls and attenuates the passage of pollutants to the saturated zone. The thicker the particles, the higher the rating.	Data from 1984 were used to define the average values for null-data surfaces in GIS layers for 2015. The category of soil texture at depth (measured within the upper 50 cm of the soil profile) was used to define the impact of the vadose zone, and the superficial soil texture was used to rate null-data surfaces (n = 3).	(Hamza et al., 2007). Fig. A.5
Infiltration rate	Hydraulic conductivity, which is a lateral flow velocity parameter of the water table, was not available for the study surfaces; thus, it was replaced with the infiltration rate, which is defined by the vertical flow of rainfall through the soil surface column.	This modification is plausible because the depth to the water table is >10 m in most surfaces of the NEA and the maximum depth of pits is 6 m. It was assumed that water flows in a vertical direction at shallow depths, with discontinuities in zones of subsurface and surface waters (n = 3).	(Miguez-Macho & Fan, 2012). Fig. A.5

742

Figures of the single parameters used in the Drastic model

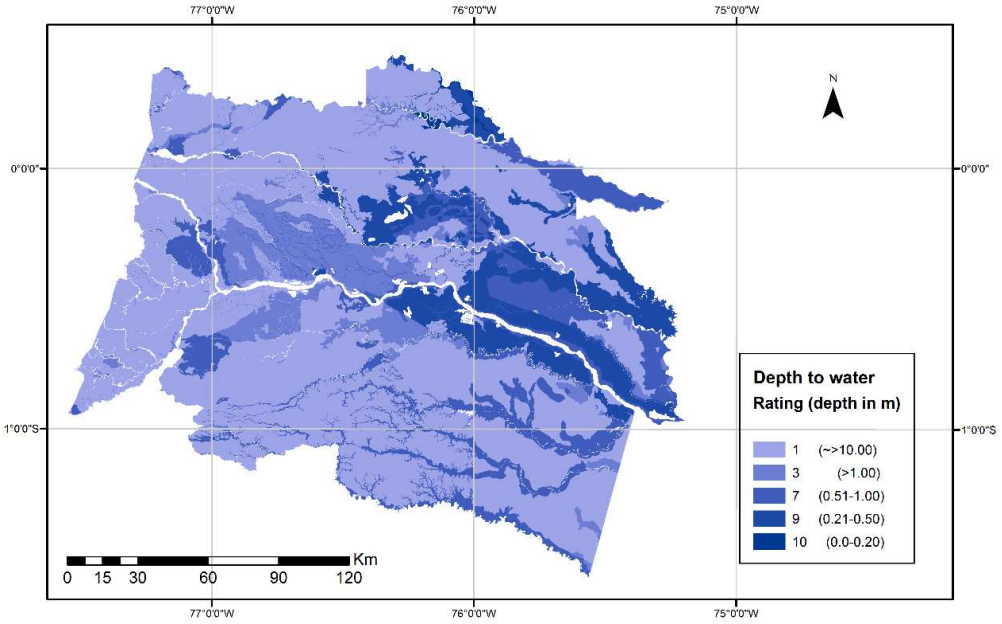


Fig. A.1. Depth to the water table. Showing measured values and assigned ratings.

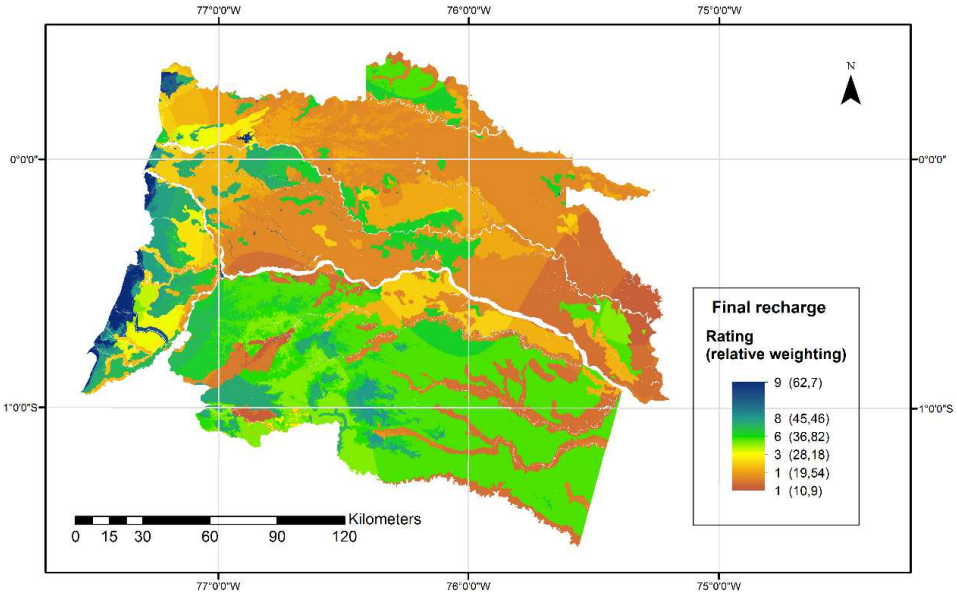


Fig. A.2. Final recharge approximated by the sub-model with the corresponding relative weighting and assigned ratings.

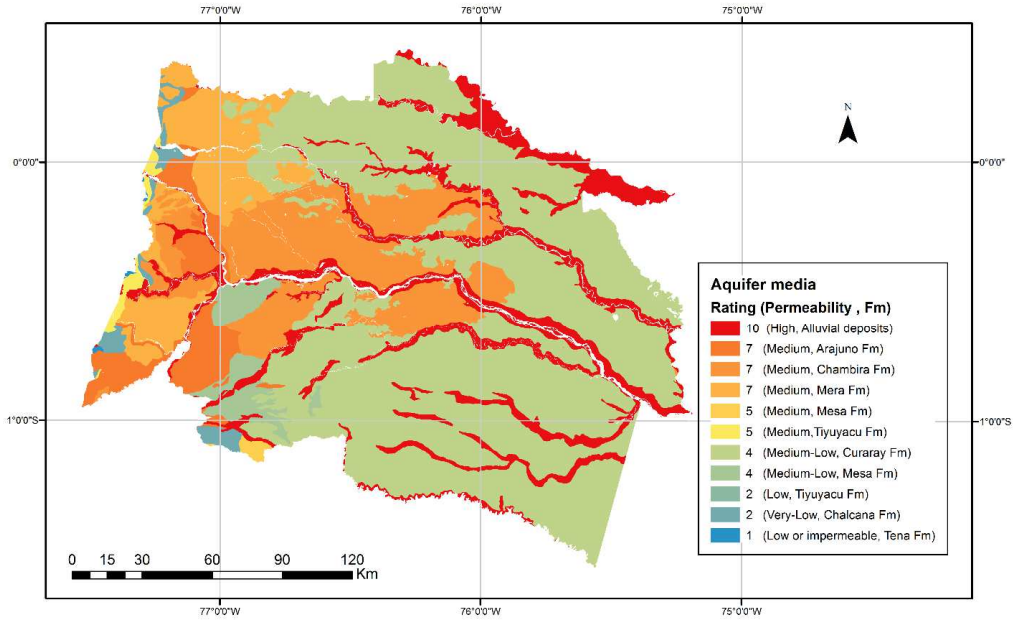


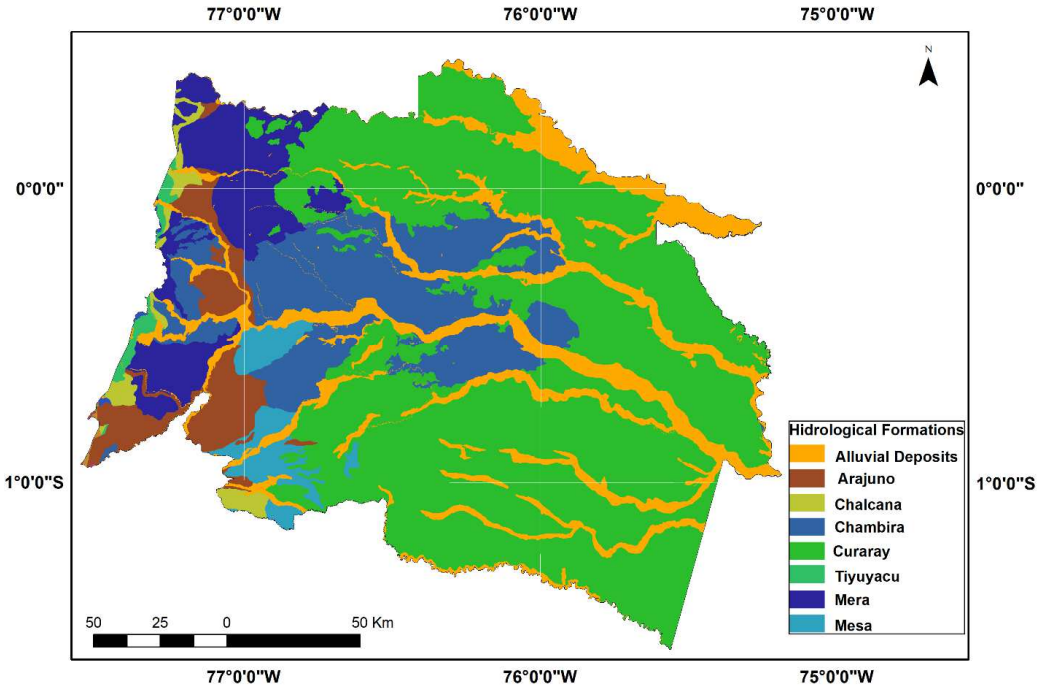
Fig. A.3. Hydrogeological settings depicting categories of permeability and formation types with the corresponding ratings.

Aquifer formations (Fm) in Figure A.4:

- Curaray Fm: Miocene tidalites (sandy to silty), argillaceous/tuffaceous lutite, and fine plastered sandstone.
- Chambira Fm: Pliocene quartz pebble-bearing conglomerates are included in a quartz-rich argillaceous matrix. The basal part contains trough cross-bedded and matrix-supported conglomerates. The main flow directions range from SW–NE to E–W.
- Arajuno Fm: Ranging from Miocene fine–coarse-grained sandstones to silty tidalites with two drainage directions: WNW–ESE to W–E and N to S.

- Mesa Fm: Late Pliocene clastic deposits with medium to coarse grain sizes.
- Mera Fm: Poorly constrained thick Quaternary conglomerates, tuffaceous sand, and argillaceous rocks (thickness of 45 m) covered by volcanoclastic deposits.
- Tiyuyacu Fm: Ranging from Eocene conglomerate sandy lutites (fine–coarse-grained) to mudstones. Typical channel-filling organisation with upward fining, 10-m-thick sequences, and variable thicknesses (150–548 m).
- Alluvial deposit Fm (Quaternary).
- Chalcana Fm: Late Oligocene to Early Miocene reddish shales intercalated with rare fine-grained and thin sandstone beds displaying trough cross-bedded stratifications and horizontal laminations with variable thicknesses (255-455 m).

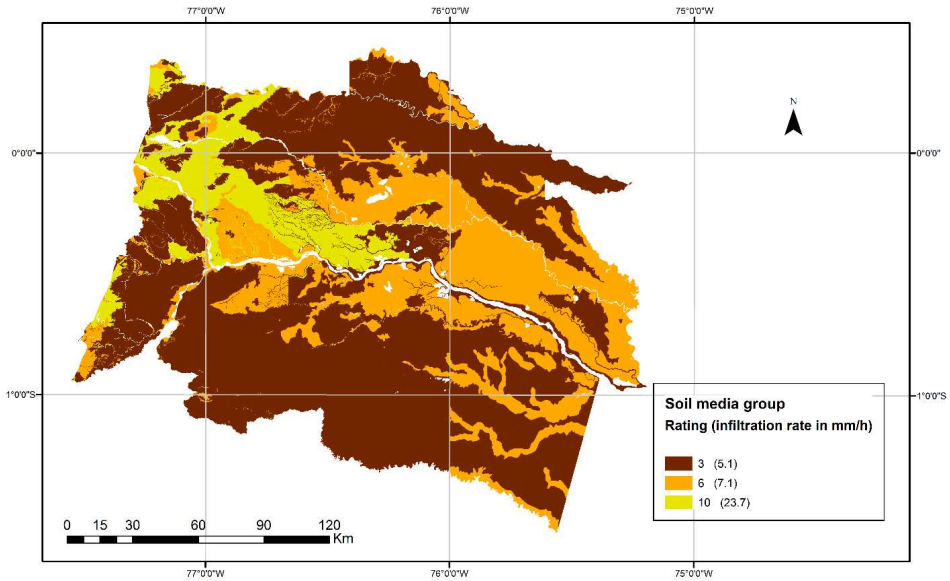
Rating scores: Alluvial deposits were given high ratings and are located mostly in riverbanks and sediments. The medium scores of the seven ratings correspond to the Arajuno, Chambira, and Mena formations in the western and central parts of the NEA. Minimum ratings scores were assigned to the Tena, Tiyuyacu, and Chalcana formations.



781

782 **Fig. A.4.** Formations (Fm) corresponding to the proposed Western Amazon Aquifer System (Rosário

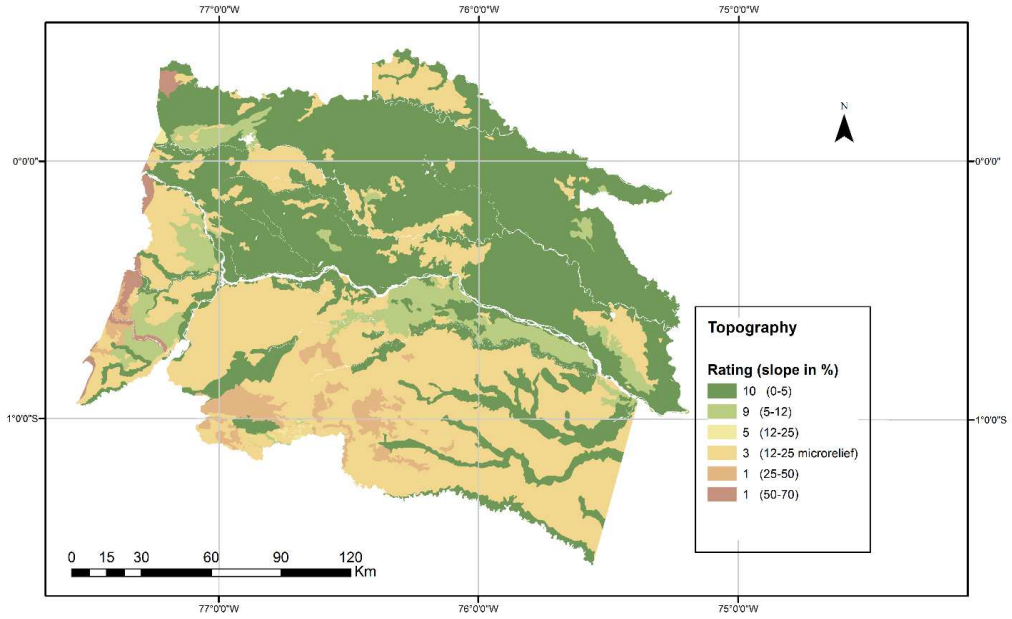
783 et al., 2016).



784

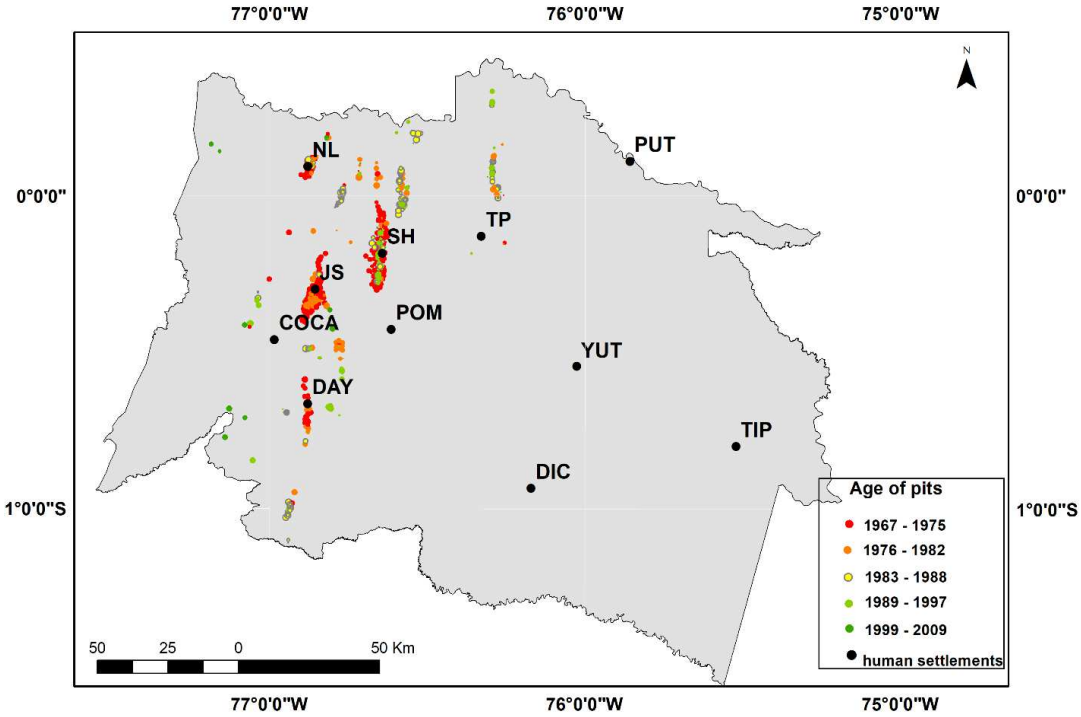
785 **Fig. A.5**. Soil media layers, corresponding to the vadose zone, with the soil type and infiltration rate in

786 parentheses.



787

788 **Fig. A.6**. Topography with the respective percentage slope and rating scores.



789

790 **Fig. A.7.** Locations of pits with the year when operation initiated (PRAS, 2015).

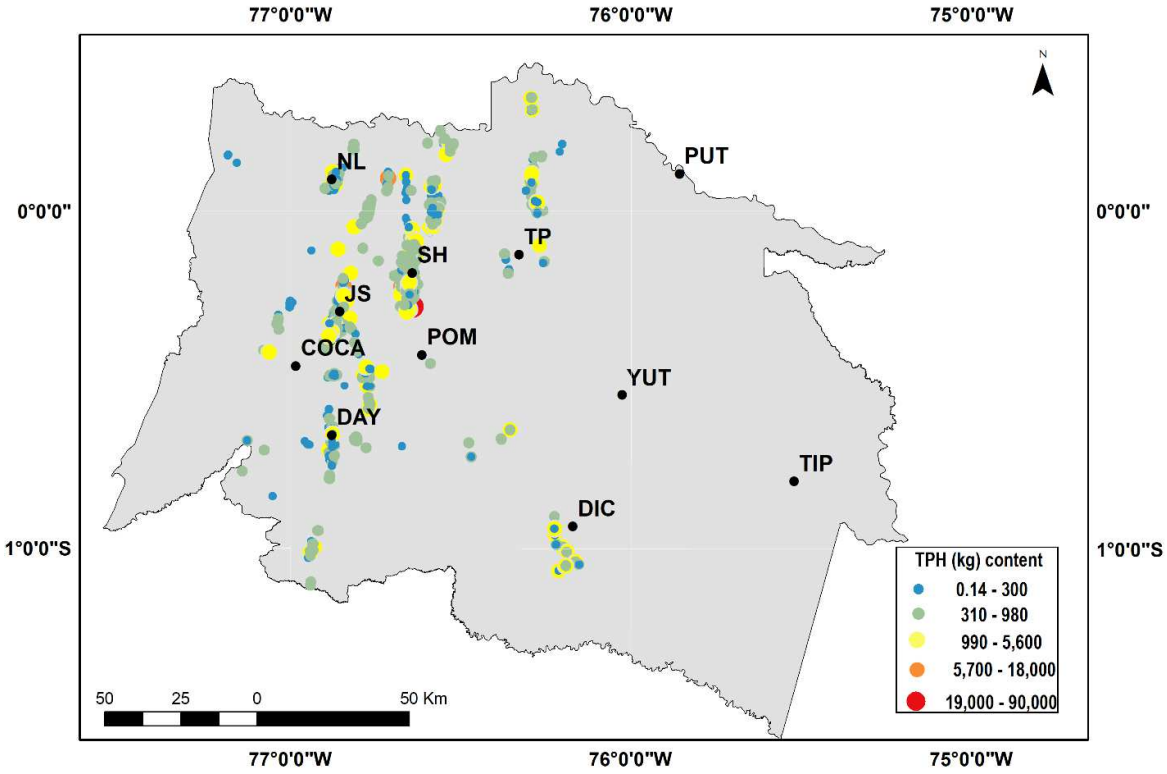


Fig. A.8. TPH contents calculated in unlined oil pits combining original and extrapolated data. Displayed using Natural Break Jenks. Abbreviations of human settlements: DAY: Dayuma; NL: Nueva Loja; SH: Shushufindi; TP: Tarapoa; POM: Pompeya; PUT: Putumayo; COCA: Puerto Francisco de Orellana; JS: Joya de Los Sachas; YUT: Yuturi; DI: Dícaro; TIP: Tiputini.

Standard procedure and equations for performing sensitivity analysis

The *map-removal analysis* was used to measure the sensitivity of the overall vulnerability index for removing one layer. The *single-parameter analysis* was used to calculate the 'real' weight of a parameter, expressed in percentage variation using only one parameter (repeated in different zones (n = 300) of the studied area), and compared it with the 'empirical' weight given by the original DRASTIC model. The ratings of each parameter that could influence the model according to site-specific conditions were evaluated to provide new site-specific rating scores (Kazakis & Voudouris, 2015). The general procedure and equations for the sensitivity analysis are

805 explained in Eq. (A1) and Eq. (A2), respectively. The map removal and single-
806 parameter analyses were performed upon selecting 'unique condition subareas' (n =
807 ~300) of at least 10-pixels in size and recalculating the DRASTIC index n-times upon
808 obtaining the variable rating scores, while the weights remain unchanged (Napolitano
809 & Fabbri, 1996).

810 $$S = \left(\left| \frac{V}{N} - \frac{V'}{n} \right| \div V \right) \times 100 \qquad \text{(A1)}$$

811 where S is the sensitivity value expressed in terms of the percentage variation from
812 the overall index; V and V' are the originally computed and tested vulnerability
813 indices, respectively; and N and n are the number of data layers used to compute V
814 and V', respectively. The originally computed vulnerability index represents the model
815 using all parameters, whereas the tested one is the index resulting from the removal
816 of one parameter. Values closer to zero indicate a lower contribution to the overall
817 vulnerability index; conversely, values greater than zero indicate a reduced
818 contribution.

819 $$W = [(\mathrm{P}r \times \mathrm{P}w) \div V)] \times 100 \quad \text{(A2)}$$

820 where W is the effective weight of a given parameter, Pr is the rating of the parameter,
821 Pw is the weight of the parameter, and V is the vulnerability index, as computed in
822 Equation (1). For instance, the D parameter is multiplied by the corresponding weight
823 of 5 and divided by the overall DRASTIC index.

824 **Results and discussion on the sensitivity analysis**

825 Table A2 shows the variations in the vulnerability index using the *single-map*
826 *removal* approach. The variation was highest upon removal of the parameter for the
827 depth to the water table (vulnerability index mean variation = 1.7%), and was slightly
828 higher than that of the final recharge parameter, which is important and had a high
829 weighting. The lowest average variation occurred when removing the parameters for
830 soil and aquifer media.

831 **Table A2**. Influence of the various parameters on DRASTIC index variations according to the single-
832 map removal sensitivity analysis (SD: standard deviation).

Parameter removed	Variation in vulnerability index (%)			
	Mean	Minimum	Maximum	SD
Depth to the water table (D)	1.7	0.2	3.8	0.9
Final recharge (R)	1.6	0	3.6	0.6
Aquifer media (A)	0.6	0	2.0	0.5
Soil media (S)	0.8	0	1.6	0.4
Topography (T)	1.0	0	2.3	0.6
Impact on saturated zone (I)	1.6	0	3.8	1.1
Infiltration rate (C)	1.1	0.1	2.4	0.5

833 The mean effective weight indicates the extent of departure from the weights
834 given by the empirical DRASTIC model according to the specific vulnerability
835 conditions of the subareas (n for $V' = 300$). Accordingly, the impacts on the vadose
836 zone, infiltration rate, and aquifer media had higher mean effective weights. Typical
837 higher weighted parameters (e.g. depth to the water table and final recharge) were
838 assigned lower effective weights (Table A3). Topography has an empirically low
839 contribution to overall vulnerability ($w = 1$), yet it is also an input for final recharge
840 indexing. This property may result in an increase in the effective weight of the overall
841 index.

842 **Table A3.** Single parameter sensitivity statistics.

Parameter	Empirical weight	Empirical Weight (%)	Effective weight (%)				
			Mean	Minimum	Maximum	SD	Redefined weights
D	5	21.7	14.8	2.9	37.2	11.4	3.4
R	4	17.4	9.6	2.5	36.0	9.2	2.2
A	3	13.0	16.2	7.9	26.5	4.4	3.7
S	2	8.7	9.4	4.9	15.2	2.7	2.2
T	1	4.4	8.2	0.6	14.3	3.9	1.9
I	5	21.7	23.6	12.3	37.0	6.7	5.4
C	3	13.0	18.1	6.0	28.9	6.2	4.2

843

844 Al-Adamat et al. (2003) successfully applied the DRASTIC model without
845 integrating all parameters, notably hydraulic conductivity. Sensitivity analysis
846 indicates the relative importance of each parameter on the final results by removing
847 them or adjusting weights to the specific information of a study site (Kazakis &
848 Voudouris, 2015; Napolitano & Fabbri, 1996; Rahman, 2008). The results of the
849 single-map removal analysis suggest that weighting factors were more relevant in
850 defining vulnerability than the rating scores. For instance, the variation of the
851 vulnerability index was highest when removing the parameter for the depth to the
852 water table (1.7%), which could be attributed to the higher empirical weights ($R = 5$)
853 and not to the high rating score, contrary to other studies (Babiker et al., 2005). The
854 effective weight variation analysis showed that the infiltration rate and aquifer media
855 parameters were weighted higher than their empirical values of 4.2 and 3.7,
856 respectively. Aquifer media has two components: (1) soil thickness (SI Appendix, Fig.
857 A.4), which was available; and (2) transmissivity, which was not available (Hamza et
858 al., 2007; Kazakis & Voudouris, 2015).

Protocol for estimating the composition of an oil related pollutants and the results for soil of the AUCA-08 oil pit

Parent Petroleum Aromatic Hydrocarbons (PAHs) and their alkylated forms were determined by gas chromatography–mass spectrometry (GC–MS, HP 5890 series II) without supplementary treatment. The quantification limits were calculated as 10 times the reagent blank standard deviation (SD), and quality control analysis was performed using supra-pure water enriched with 10 µL of a reference solution containing a mixture of 16 native molecules of PAHs. In addition, 15 µL of a solution containing 16 deuterated PAHs was added to each sample to correct for losses due to manipulation. The recoveries obtained in the enriched water were 101 ± 7%.

Sediments were freeze-dried and subsequently treated by microwave extraction and purification in a column using dichloromethane and isooctane. The PAHs concentrations in the final extracted solution (~300 µL) were determined using inductively coupled plasma (ICP)–MS (HP 5890 series II).

Reagent blanks and certified materials: Deuterated solutions were analysed with the samples. A reference solution containing 18 deuterated PAHs was diluted to obtain 2 mg L^{-1} of each molecule, and 30 µL of this solution was added to 3 g of the sample before extraction. The average recovery of the certified materials was 91 ± 10%. TPH was added as the reference value.

878 **Table A4.** PAH concentrations (single chemical compounds).

Polycyclic Aromatic Hydrocarbons	Concentration (mg kg^{-1})	Contribution (%)
Naphthalene	0.022	0.09
Acenaphthylene	0.000025	0.00
Acenaphthene	0.84	3.21
Fluorene	1.78	6.79
Phenanthrene	6.5	24.76
Anthracene	0.33	1.28
Fluoranthene	-	-
Pyrene	-	-
Benzo[a]anthracene	0.84	3.18
Chrysene + Triphenylene	10.48	39.95
Benzo[b,j,k]fluoranthene	2.62	9.97
Benzo[a]pyrene	0.45	1.70
Indeno[1,2,3-cd]pyrene	0.33	1.25
Dibenzo[a,h]anthracene + Dibenzo[a,c]anthracene	0.30	1.16
Benzo[g,h,i]perylene	1.75	6.68
Σ15 PAHs	**26.24**	**100.00**
Low molecular weight (more soluble)	9.48	36.12
High molecular weight (less soluble)	16.76	63.88
TPH	**1050**	**-**

879