

Study of groundwater salinity in the HAHA syncline by the Kohonen self-organized classification (Essaouira, Morocco)

Zakaria Ouzerbane (✉ ouzerbanegeophy@gmail.com)

Moulay Ismail University: Universite Moulay Ismail <https://orcid.org/0000-0001-6339-0630>

Soumia Loulida

Cadi Ayyad University Faculty of Science

Abdellah El Hmaidi

Universite Moulay Ismail

Ali Essahlaoui

Universite Moulay Ismail Faculte des Sciences

Mostafia Boughalem

Université Ain Timouchent Algéria

Habiba Ousmana

Universite Moulay Ismail Faculte des Sciences

Mohamed Berrada


Universite Moulay Ismail Faculte des Sciences

Research Article

Keywords: Kohonen classification, ACP, Plio-Quaternary, Aquifer, Essaouira, Contamination.

Posted Date: March 29th, 2021

DOI: <https://doi.org/10.21203/rs.3.rs-257053/v1>

License:  This work is licensed under a Creative Commons Attribution 4.0 International License. [Read Full License](#)

Abstract

The coastal aquifer of the Essaouira syncline (Morocco) was studied to identify the main processes at the origin of the salinization of groundwater. In fact, a multicriteria analysis approach on hydrochemical data and physicochemical parameters of the Plio-Quaternary aquifer was used to understand their spatio-temporal variation and their origins. Currently, integrated water resources management has become paramount to both local, regional, national and international. This management is favored by extreme hydrological events (droughts or floods) which can have direct influences on human, economic and political aspects. Appropriate management of a resource requires its evaluation. The statistical study by Kohonen's self-organized classification (SOM) of hydrochemical data for the years 1995 and 2009 is used to process 47 samples distributed over the entire study area; it showed an evolution of the values of the parameters. Physico-chemical as a function of time and space with an increase in the values of the parameters, from the center of the study area towards the southwest.

Introduction

Groundwater remains the main water resources for drinking water consumption, agricultural exploitation, and for industrial and tourist use by inhabitants of towns and rural areas in coastal areas with arid climates and semi-arid. Several factors are involved in the degradation of electrical conductivity and consequently the quality of groundwater, these factors can be of natural origin (leaching, evaporation, dissolution, marine aerosols) or of anthropogenic origin (marine intrusion, over-fertilization, water drainage, wastewater spreading, automobile emissions), and sometimes the combination of anthropogenic and natural origins (Andreasen and Fleck 1997, Oulaaross 2009).

Generally in nature, the groundwater in a relation with the surface water infiltrated and percolated in the unsaturated zone to reach the aquifer they contain (saturated zone), the effective infiltration rate to influence the piezometric evolution and consequently the evolution of the water quality of the reservoir. Indeed, a large effective infiltration causes an increase in the piezometric level, an increase in its flow and leads to the dilution of its mineralization (Urish and Frohlich 1990; Sherif et al. 2006 ; Attwa and Zamzam 2020). On the contrary, a weak or absent effective infiltration combined with intensive pumping causes a lowering of the water table and a mineral overconcentration of the groundwater (Chafouq et al. 2018). These waters are often threatened by contamination by pollutants of different origins which can be biological, chemical or physical.

The Essaouira basin has seen, like other regions of Morocco, a significant decrease in water supplies in quantity and quality. This situation has resulted in the reduction of agricultural productivity and the degradation of several ecosystems. However, this basin has an aquifer system formed by a set of aquifers of unequal size. These aquifers can offer a natural regulating capacity which makes them valuable in ensuring a safe regular supply. The reserve also makes it possible to meet seasonal needs through temporary overexploitation to the extent that recovery is possible (Chamchati et al., 2013).

The diffuse pollution generated by marine waters is amplified by point sources of pollution mainly represented by human and agro-industrial activities. As a result, the salt contents recorded in certain places greatly exceed the standard of drinking water or even irrigation (Bahir et al. 2001). This situation risks ultimately jeopardizing the sustainability of agricultural activity as a whole, harms public health and compromises the self-purification power of groundwater. These waters circulate slowly through the subsoil, so much so that pollution from human activities can persist over long periods that can range from a few years to several decades or even longer for certain specific aquifers. In other words, in coastal areas, groundwater resources require special attention to minimize saltwater intrusions, either locally, or on a regional scale. The extent of salt water intrusion depends on the geometry, structure and properties of the aquifer, the stresses applied to the aquifer (pumping, etc.) and the flow of fresh water and salty exchanging with the surrounding environment (Karroum et al., 2017). This situation is all the more critical as the groundwater in this study area (Essaouira coastal area) constitutes the only water resources available to local populations (Ouzerbane et al. 2014 ; Bahir et al. 2019).

This study will take place in the context of regional planning. It will make it possible to prepare, confirm or supplement the orientations and choices necessary for a harmonious development of the region taking into account, as far as possible, all the needs and the corresponding constraints.

Geographical And Geological Location

The Essaouira basin is a coastal area located in the center-west of Morocco. It belongs to the Moroccan Atlantic margin and covers approximately an area of 20,000 km² of the eastern Mesozoic passive margin of the central Atlantic Ocean. The basin is limited to the north by the Tensifriver, to the south by the western High Atlas, to the east by the plains of Imi-N-Tanoute and Chichaoua and to the west by the Atlantic Ocean. It is a plateau, slightly raised to the South and to the North, gently sloping down towards the Atlantic Ocean then suddenly lowering to give rise to the coastal zone with dune relief.

The study area known as the synclinal Essaouira is part of the coastal area of the basin. It is bounded to the north by Qsob river, to the south by Tidzriver, to the east by the reliefs of Haha North and by the diapir of Tidzi and to the west by the Atlantic Ocean. It is not very rugged and is characterized by a relief of low hills with altitudes between 0 and 800 meters, modeled by a low density hydrographic network which flows into the Atlantic Ocean (Fig. 1).

Geologically, the formations outcropping in the synclinal zone of Essaouira are generally Quaternary and Plio-Quaternary in age (Fig. 2). Going from the South to the North of the study area, the stratigraphic series extends from the Triassic to the Quaternary (Ouzerbane 2015 ; Ouzerbane et al. 2019 ; Cohen et al. 2020).

The Triassic formations outcrop in the southeastern part of the study area at the level of the Tidzi region in a NE-SW direction forming the Tidzidiapir over a large area while in the extreme south of the area. study these formations outcrop over a smaller area along the Aghbalou river. These formations are mainly known by saliferous red clays, doleritic basalts and reddish sandstone pelites (Suter 1958).

The Jurassic outcrops are formed mainly by limestones and gypsiferous marl-limestones dating from the Upper Jurassic of the Jbelhadid located north of the syncline of Essaouira (Medina 1994 ; Piqué 1994 ; Roch, 1931 ; Taj-Eddine, 1991).

The Cretaceous formations outcrop over the entire Essaouira basin with thicknesses varying from one place to another. In the study area, the Cretaceous outcrops in the east by gray marls and lumachelliclimestones of Senonian age, dolomitic limestones with flint of Turonian age, limestones, marls and gypsiferous marls of the Cenomanian along the Tidziriver. Towards the west, the outcrops of the Albian correspond to green marls surmounted by dolomitic limestones of Maestrichtian age. The Lower Cretaceous formations also outcrop to the NW of the study area (Choubertand Faure-Muret1962 ;Duffaud et al. 1966 ; Rey et al. 1950).

Material And Method

The water samples were taken from the plio-quaternary water table in the coastal zone of Essaouira between Qsob river and Tidzi river at the level of 35 wells (Table 1) during 1995 (Mennani 2001) and at the level of 12 wells (Table 2) in 2009 (Chamchati 2014). The physical parameters relate to the piezometric level (PL), the temperature (T) and the electrical conductivity (CE).The chemical parameters concern on the one hand the major cations (calcium, magnesium, sodium and potassium) and on the other hand the major anions (chlorine, nitrates, sulphates and HCO₃⁻). The data has been processed by advanced statistical analysis techniques. The Kohonen Self-Organizing Topological Map (SOM) classification method was used to understand and visualize the spatial and temporal distribution of the samples.

Principal component analysis (ACP) and hierarchical classification (CHA of SOM) by topological maps were used to validate the classification by SOM.

Kohonen Self-Organizing Map (SOM)

This method is based on the neural networks of Kohonen (2001). These are common tools increasingly used for multivariate data processing and provide practical visualization results. A SOM map consists of units, called neurons, connected on a regular grid, usually 2D hexagonal grid. It allows the partition of a global training data set (input data space) into a reduced number of subsets, having certain statistical characteristics in common. Each subset is represented by a weight vector which has as many components as the input data vectors. A weight vector is a vector in the virtual input data space and corresponds to a neuron of the network card (output space card). A special feature of SOM algorithms over other grouping methods is the preservation of the topology of the input data space. This conservation of topology allows samples with similar characteristics to be placed together on the map. Correlations and relationships between samples and variables can be easily visualized using the SOM component visualization plans (Vesanto 1999). The map of accuracy in topology preservation is evaluated by the mean quantization error (QE) and topological error (TE) (Kiviluoto1996 ;Kohonen2001).

Component planes show the values of each component in each single neuron and how each input vector varies in the plane of the output space. They allow the detection of correlated variables when viewing several component plans at the same time. Patterns of the same color between variables mean that a variable is increasing or decreasing. In this case, these are positive correlations. Conversely, negative correlations between variables will have the same patterns, but of opposite color distribution (Fig. 3). In this study, the SOM card is used to process all 47 samples.

Table 1
Physico-chemical data of samples taken from 35 wells during 1995 (Mennani 2001).

Ref	Samples	Stations	X	Y	Z	H (m)	PL (m)	T (°C)	CE 25° (µs/cm)	HCO ³⁻ (mg/l)	Cl ⁻ (mg/l)	SO ₄ ²⁻ (mg/l)	Ca ²⁺ (mg/l)	Mg ²⁺ (mg/l)	
MENNANI 2001 (Partner Dec 1995)	140/51	1	87300	103600	60	50	10	19	2700	363.6	738.4	177.8	179.6	56.6	
	20/51	2	88150	96050	92	66	26	22	3250	209.8	1029.5	221.2	159.2	76.3	
	135/51	3	92450	97900	110	68	42	22	2500	219.6	788	164.8	155.2	73.7	
	116/51	4	100650	96000	200	178.5	21.5	22	2610	285.5	671	143.5	194.4	69.1	
	33/51	5	96800	96750	164	96	68	21	2340	285.5	646.1	84.3	172	43.2	
	327/51	6	88800	88800	130	106	24	21.5	1870	162.3	422.5	78.6	92	27.8	
	138/51	7	87850	92830	109	84	25	21	3960	192.8	1487.5	166.4	298.4	104.6	
	M33	8	91150	102300	78	60.8	17.2	20.5	1890	346.5	436.7	187.1	107.2	73	
	103/51	9	94820	102175	99	73	26	21	2100	314.8	532.5	213.3	112.2	85.2	
	272/51	10	97170	100760	105.5	77.5	28	20.5	2280	295.2	600	243.3	108	98.4	
	133/51	11	87800	98800	70	44	26	22	4710	307.4	1462.6	402.2	295.2	150	
	3/51	12	81400	93400	18	13	5	19	2550	412.4	646.1	163.4	117.2	52.6	
	M20	13	85590	93700	85	5	80	23	3130	383.1	972.7	163.9	164	61.9	
	28/51	14	97200	91800	225	176	49	23	3360	339.2	1136	137.5	227.2	109.4	
	125/51	15	93250	94100	155	95	60	24	3330	226.9	1228.3	106.4	279.6	63.1	
	104/51	16	97400	99750	115	78.6	36.4	22	2100	268.4	546.7	168.3	152	53.3	
	261/51	17	96420	99250	124	80	44	23	2000	266	436.7	178.4	146.8	48.7	
	149/51	18	85100	105800	40	2	38	22	3580	353.8	1075.7	216.3	104.8	101.3	
	M5	19	86550	102700	60	31	29	22	1590	205	429.6	100.1	112.8	55.2	
	M10	20	81250	92800	12	3.7	8.3	19	2700	390.4	777.5	131.8	184.4	19.4	
	18/51	21	87800	99200	78	60	18	21	5040	302.6	1693.4	418.3	320.8	76.8	
	132/51	22	89750	98500	99	60.5	38.5	15.5	2580	226.9	830.7	163.4	185.6	75.1	
	137/51	23	86750	94450	87	68.5	18.5	21	3490	175.7	1320.6	129.3	248	91.2	
	21/51	24	89400	91400	89.6	61.9	27.7	20	4460	207.4	1728.9	162	376	61.4	
	124/51	25	93650	94450	148	86	62	23	1860	241.6	586.8	45.5	140.4	67.9	
	126/51	26	94800	96950	137	79	58	24.5	3240	222	1189.3	132	306	64.1	
	127/51	27	94400	97800	125	73	52	19	3570	180.6	1359.7	131.5	325.6	106.3	
	M31	28	92900	97300	125	72	53	23	3220	102.5	1221.2	147.9	291.2	121.7	
	130/51	29	91350	101800	85	64.62	20.38	22	2170	324.5	500.6	205.9	112	77.3	
	49/51	30	90950	102700	66	51	15	20.5	2340	369.7	532.5	244.6	102	104.4	
	93/51	31	92370	101900	98	68.42	29.58	21.5	2150	339.2	489.9	202.9	111.2	128.6	
	260/51	32	96550	100800	106	70.2	35.8	20.5	2090	287.9	539.6	238.9	109.6	71.3	
	27/51	33	95500	91300	208	173.7	34.3	21	2000	358.7	433.1	229.4	108	106.1	
	236/51	34	97500	99650	120	105.8	14.2	21	2380	342.8	582.2	241.4	120	118.1	
	11/51	35	80450	96450	8	5.3	2.7	18	4830	263.5	1654.3	290.7	172.4	101	

Table 2
Physico-chemical data of samples taken from 12 wells during 2009 (Chamchati 2001).

Ref	Samples	Stations	X	Y	Z	H (m)	NP (m)	T (°C)	CE 25° (µs/cm)	HCO ³⁻ (mg/l)	Cl ⁻ (mg/l)	SO ₄ ²⁻ (mg/l)	Ca ²⁺ (mg/l)	Mg ²⁺ (mg/l)
CHAMCHATI, 2014 (Partner oct 2009)	B1	36	85500	105620	22	0	22	20	2440	146.4	475.7	128.6	40	38.64
	138/51	37	87850	92830	109	5	104	21	3520	109.8	766.8	124.8	38	45.96
	M33	38	91150	102300	78	29	49	20	2040	207.4	347.9	104.64	40	30.24
	B5	39	90290	102260	102	48	54	19	2249	176.9	390.5	102.72	42	39.84
	B7	40	95150	104490	97	27.8	69.2	20	770	67.1	74.55	27.36	30	24.12
	B9	41	93410	102680	114	47	67	19	1763	219.6	241.4	121.44	44	35.04
	M61	42	91200	100750	90	34	56	23	1720	195.2	255.6	119.52	46	41.04
	103/51	43	94820	102170	99	22.5	76.5	21	1671	176.9	217.97	111.84	38	31.44
	272/51	44	97170	100760	105.5	0	105.5	20	2180	189.1	355	129.6	42	38.64
	133/51	45	87800	98800	70	38	32	22	2550	146.4	442.33	96	40	42.36
	3/51	46	81400	93400	18	4	14	19	2130	195.2	344.35	44.64	34	30.24
	B39	47	84980	111080	23	5	18	23	3060	207.4	372.75	168.48	66	59.28

Ascending hierarchical classification of the SOM card (SOM-ACH)

It is an automatic classification method used in data analysis, from a set of individuals, its goal is to distribute these individuals into a certain number of classes (Fig. 4). These are hierarchical classification methods of techniques for sharing the training dataset in certain classes according to their proximity. The first iteration of an ascending hierarchical classification (AHC) algorithm is to combine the two closest individuals. The elements (individuals or group of individuals) pairs are combined according to their proximity to arrive at a single class. Finally, a hierarchy of groups is obtained. Each level of the hierarchy represents a partition of particular data on disjoint groups. For the present study, the hierarchical classification SOM (SOM-HC) allows us to group the similar neurons of the SOM map according to the distances between the vectors of corresponding weight. Each particular set of data will be assigned to the corresponding group of its neuron.

Correlation matrix

The correlation matrix is the matrix of correlation coefficients, calculated on several variables taken in pairs. The resulting matrix is symmetric and the elements on the diagonal are equal to 1, the covariance of a variable with itself being equal to its variance. It is used to easily identify the links between the variables. When the covariance is zero, it is certain that the variables are independent and when it is equal to 1 the correlation is established. Intermediate values are more difficult to interpret.

However, a positive value indicates a positive correlation between two variables while a negative value indicates opposite correlations. This coefficient varies between - 1 and + 1; the intensity of the linear relation will therefore be all the stronger as the value of the coefficient is close to + 1 or to - 1, and all the weaker as it is close to 0.

Principal Component Analysis (ACP)

The objective of this methodology is to determine the factors that influence the variability of the parameters of the Essaouira basin. The ACP method used for this study is based on the interpretation of the correlation matrix as well as the various factors obtained as a result of data processing. The choice of the main axes takes into account the reduction in the number of factors. This number is such that the cumulative sum of the contributions is significant (75% which represents three quarters of the total inertia). Indeed, two variables are correlated when their correlation coefficient is greater than or equal to 0.7.

In addition, at the factorial design level, variables are only representative when they are close to the end of these factors. When two variables are correlated, the variation of one results in the variation of the other. Analyses of variance were used to judge the significance of the relationships highlighted by the factor analysis.

Results And Discussion

Kohonen self-organizing map and correlation matrix

The distribution maps, taken from the Kohonen map, make it possible to visualize the distribution of the samples according to the hydro-chemical parameters (Table 3, 4, 5).

Dark cells in red represent high values, while blue cells represent low values. The values are represented with a logarithmic scale. A topological map of 48 cells (8 columns x 6 rows) was selected for this analysis with quantification errors $q_e = 0.338$ and topography errors $t_e = 0.000$ (Fig. 5). Each card is representative of the corresponding variable. Consequently, it makes it possible to study the correlations between the variables.

Table 3
 Classification data and visualization of the spatial and temporal distribution of samples (Classes 1, 2).

Classes	Neurons	Stations	PL (m)	T (°C)	CE 25° (µs/cm)	HCO ₃ ⁻ (mg/l)	Cl ⁻ (mg/l)	SO ₄ ²⁻ (mg/l)	Ca ²⁺ (mg/l)	Mg ²⁺ (mg/l)	Na ⁺ (mg/l)	K ⁺ (mg/l)	rMg/rCa	i.e.b
1	1	36	22,00	20,00	2440,00	146,40	475,70	128,60	40,00	38,64	266,80	5,46	0,97	0,43
1	1	45	32,00	22,00	2550,00	146,40	442,33	96,00	40,00	42,36	236,44	3,90	1,06	0,46
1	1	47	18,00	23,00	3060,00	207,40	372,75	168,48	66,00	59,28	219,42	19,50	0,90	0,36
1	9	39	54,00	19,00	2249,00	176,90	390,50	102,72	42,00	39,84	207,69	10,53	0,95	0,44
1	10	44	105,50	20,00	2180,00	189,10	355,00	129,60	42,00	38,64	212,98	4,29	0,92	0,39
1	10	46	14,00	19,00	2130,00	195,20	344,35	44,64	34,00	30,24	254,38	5,85	0,89	0,24
1	17	38	49,00	20,00	2040,00	207,40	347,90	104,64	40,00	30,24	221,95	13,65	0,76	0,32
1	18	6	24,00	21,50	1870,00	162,30	422,50	78,60	92,00	27,80	253,00	5,10	0,30	0,39
1	25	19	29,00	22,00	1590,00	205,00	429,60	100,10	112,80	55,20	200,10	12,10	0,49	0,51
1	25	40	69,20	20,00	770,00	67,10	74,55	27,36	30,00	24,12	36,80	3,12	0,80	0,46
1	25	41	67,00	19,00	1763,00	219,60	241,40	121,44	44,00	35,04	138,69	4,68	0,80	0,41
1	25	42	56,00	23,00	1720,00	195,20	255,60	119,52	46,00	41,04	127,42	5,46	0,89	0,48
1	25	43	76,50	21,00	1671,00	176,90	217,97	111,84	38,00	31,44	135,70	4,68	0,83	0,36
1	42	8	17,20	20,50	1890,00	346,50	436,70	187,10	107,20	73,00	167,90	1,20	0,68	0,61
1	42	17	44,00	23,00	2000,00	266,00	436,70	178,40	146,80	48,70	241,50	2,30	0,33	0,44
1	42	25	62,00	23,00	1860,00	241,60	586,80	45,50	140,40	67,90	218,50	2,00	0,48	0,62
2	4	30	15,00	20,50	2340,00	369,70	532,50	244,60	102,00	104,40	287,50	1,60	1,02	0,46
2	4	34	14,20	21,00	2380,00	342,80	582,20	241,40	120,00	118,10	264,50	2,00	0,98	0,54
2	5	1	10,00	19,00	2700,00	363,60	738,40	177,80	179,60	56,60	333,50	3,10	0,32	0,54
2	6	18	38,00	22,00	3580,00	353,80	1075,70	216,30	104,80	101,30	655,50	4,70	0,97	0,39
2	6	37	104,00	21,00	3520,00	109,80	766,80	124,80	38,00	45,96	438,15	10,53	1,21	0,41
2	11	29	20,38	22,00	2170,00	324,50	500,60	205,90	112,00	77,30	299,00	1,20	0,69	0,40
2	11	31	29,58	21,50	2150,00	339,20	489,90	202,90	111,20	128,60	230,00	1,60	1,16	0,53
2	13	12	5,00	19,00	2550,00	412,40	646,10	163,40	117,20	52,60	391,00	4,70	0,45	0,39

Table 4
Classification data and visualization of the spatial and temporal distribution of samples (Classes 2, 3).

Classes	Neurons	Stations	PL (m)	T (°C)	CE 25° (µs/cm)	HCO ₃ ⁻ (mg/l)	Cl ⁻ (mg/l)	SO ₄ ²⁻ (mg/l)	Ca ²⁺ (mg/l)	Mg ²⁺ (mg/l)	Na ⁺ (mg/l)	K ⁺ (mg/l)	rMg/rCa	i.e.b
2	14	2	26,00	22,00	3250,00	209,80	1029,50	221,20	159,20	76,30	471,50	7,00	0,48	0,54
2	14	13	80,00	23,00	3130,00	383,10	972,70	163,90	164,00	61,90	414,00	9,80	0,38	0,56
2	19	9	26,00	21,00	2100,00	314,80	532,50	213,30	112,20	85,20	207,00	2,70	0,76	0,61
2	19	32	35,80	20,50	2090,00	287,90	539,60	238,90	109,60	71,30	218,50	2,30	0,65	0,59
2	19	33	34,30	21,00	2000,00	358,70	433,10	229,40	108,00	106,10	165,60	2,30	0,98	0,61
2	20	10	28,00	20,50	2280,00	295,20	600,00	243,30	108,00	98,40	253,00	2,70	0,91	0,57
2	29	3	42,00	22,00	2500,00	219,60	788,00	164,80	155,20	73,70	287,50	2,30	0,47	0,63
2	37	4	21,50	22,00	2610,00	285,50	671,00	143,50	194,40	69,10	299,00	1,90	0,36	0,55
2	37	22	38,50	15,50	2580,00	226,90	830,70	163,40	185,60	75,10	322,00	2,30	0,40	0,61
2	40	14	49,00	23,00	3360,00	339,20	1136,00	137,50	227,20	109,40	333,50	3,50	0,48	0,70
2	40	15	60,00	24,00	3330,00	226,90	1228,30	106,40	279,60	63,10	368,00	3,10	0,23	0,70
2	40	26	58,00	24,50	3240,00	222,00	1189,30	132,00	306,00	64,10	356,50	3,10	0,21	0,70
2	40	28	53,00	23,00	3220,00	102,50	1221,20	147,90	291,20	121,70	368,00	2,70	0,42	0,70
2	43	16	36,40	22,00	2100,00	268,40	546,70	168,30	152,00	53,30	287,50	3,50	0,35	0,47
2	44	5	68,00	21,00	2340,00	285,50	646,10	84,30	172,00	43,20	299,00	2,30	0,25	0,53
2	46	20	8,30	19,00	2700,00	390,40	777,50	131,80	184,40	19,40	402,50	4,70	0,11	0,48
3	8	11	26,00	22,00	4710,00	307,40	1462,60	402,20	295,20	150,00	678,50	3,50	0,51	0,53
3	8	21	18,00	21,00	5040,00	302,60	1693,40	418,30	320,80	76,80	713,00	3,10	0,24	0,58
3	8	35	2,70	18,00	4830,00	263,50	1654,30	290,70	172,40	101,00	736,00	35,10	0,59	0,53
3	16	24	27,70	20,00	4460,00	207,40	1728,90	162,00	376,00	61,40	586,50	2,30	0,16	0,66
3	24	7	25,00	21,00	3960,00	192,80	1487,50	166,40	298,40	104,60	517,50	3,90	0,35	0,65
3	31	23	18,50	21,00	3490,00	175,70	1320,60	129,30	248,00	91,20	471,50	3,90	0,37	0,64
3	32	27	52,00	19,00	3570,00	180,60	1359,70	131,50	325,60	106,30	391,00	2,70	0,33	0,71

The correlation matrix expresses the different correlations between the analyzed variables (Table 5). The set of correlations is relatively weak. Indeed, more than 77% of the correlation coefficients are less than 0.53. It follows that the analyzed variables are not strongly correlated with each other. This finding significantly reduces the redundancy of information and at the same time justifies the relevance in the choice of these variables to conduct the study.

The strong correlations observed link variables of the same class (Table 5), namely:

- The electrical conductivity (CE 25 ° C) shows positive correlations with significant R coefficients respectively with Na²⁺ (R = 0.94), Cl⁻ (R = 0.93), Ca²⁺ (R = 0.72), SO₄²⁻ (R = 0.56) and Mg²⁺ (R = 0.48).
- The bicarbonates HCO₃⁻ (mg / l) are slightly positively correlated with sulphates (R = 0.51).
- Chlorides are positively correlated with Na⁺ (R = 0.89), Ca²⁺ (R = 0.88), ieb (R = 0.61), Mg²⁺ (R = 0.53), SO₄²⁻ (R = 0, 46) and negatively correlated with the rMg / rCa ratio.
- Sulphates are relatively positively correlated with Mg²⁺ (mg / l) (R = 0.64) and Na⁺ (mg / l) (R = 0.58).
- Calcium is positively correlated with the base exchange index i.e.b. (R = 0.76), sodium Na⁺ (R = 0.65), Mg²⁺ (mg / l) (R = 0.49) and negatively correlated with the rMg / rCa ratio (R = - 0.75).
- Magnesium Mg²⁺ is positively correlated with the base exchange index i.e. b. (R = 0.55) and to a lesser degree with Na⁺ (mg / l) (R = 0.44).
- The rMg / rCa ratio is slightly negatively correlated with the core exchange index i.e.b. (R = - 0.49).

Table 5
Correlation coefficient between the physicochemical parameters.

	PL (m)	T (°C)	CE 25° (µs/cm)	HCO ₃ ⁻ (mg/l)	Cl ⁻ (mg/l)	SO ₄ ²⁻ (mg/l)	Ca ²⁺ (mg/l)	Mg ²⁺ (mg/l)	Na ⁺ (mg/l)	K ⁺ (mg/l)	rMg/rCa	i.e,b
PL (m)	1,00	0,21	-0,20	-0,42	-0,19	-0,42	-0,18	-0,29	-0,28	-0,09	0,15	-0,02
T (°C)		1,00	0,03	-0,02	0,06	-0,02	0,17	0,16	-0,04	-0,17	-0,09	0,22
CE 25° (µs/cm)			1,00	0,09	0,93	0,56	0,72	0,48	0,94	0,28	-0,35	0,38
HCO ₃ ⁻ (mg/l)				1,00	0,08	0,51	0,12	0,35	0,19	-0,15	-0,11	0,09
Cl ⁻ (mg/l)					1,00	0,46	0,88	0,53	0,89	0,12	-0,55	0,61
SO ₄ ²⁻ (mg/l)						1,00	0,34	0,64	0,58	0,06	-0,01	0,20
Ca ²⁺ (mg/l)							1,00	0,49	0,65	-0,19	-0,75	0,76
Mg ²⁺ (mg/l)								1,00	0,44	-0,10	0,04	0,55
Na ⁺ (mg/l)									1,00	0,26	-0,39	0,25
K ⁺ (mg/l)										1,00	0,13	-0,27
rMg/rCa											1,00	-0,49
i,e,b												1,00

Hierarchical classification ACH - SOM

The SOM hierarchical classification method applied consists in grouping the neurons as best as possible so as to give a more global view of the SOM map. To obtain a good partition, one can manually choose to cut at a level where the branches of the tree (Dendrogram) are long, indicating that the data contained in the classes are very different.

When you have a lot of data, you just need to visualize the nodes closest to the root. Another possible cutting criterion is to choose according to the distance between the classes. The line of the figure shows a cutoff at 1.2 (Fig. 6). However, it is difficult to automatically determine what the correct cutting value is. Another criterion is to cut according to the number of classes obtained by the SOM card.

The dendrogram obtained by ACH-SOM shows the presence of 3 classes (Fig. 6). The first class (cluster I) includes neurons 1, 2, 10, 17, 18, 25 and 29 which correspond to stations 3, 6, 36, 38, 40, 41, 42,43, 44, 45, 46 and 47. These stations represent the entire 2009 campaign.

The second class (cluster II) is characterized by neurons 7, 8, 9, 15, 16, 23 and 24 which correspond to stations 7, 11, 21, 24, 35 and 39. These stations belong to the 1995 campaign and are located in the SW part of the study area. The third class (cluster III) presents the greatest number of neurons with 3, 4, 5, 6, 11, 12, 13, 14, 19, 20, 21, 22, 26, 27, 28 and 30 which correspond to the stations 1, 2, 9, 10, 12, 13, 18, 29, 30, 31, 32, 33, 34 and 37. These stations are part of the 1995 campaign and are located mainly in the NE part of the study area around the Qsobriver.

Classification by SOM card

Kohonen's self-organizing SOM card also enabled 3 classes to be identified (Fig. 7). Class 1 is formed by neurons 1, 9, 10, 17, 18, 25 and 42 corresponding to stations 6, 8, 17, 19, 25, 36, 38, 39, 40, 41, 42, 43, 44, 45, 46 and 47. These stations mainly include the samples of the 2009 campaign. Class 2 is made up of neurons 4, 5, 6, 11, 13, 14, 19, 20, 29, 37, 40, 43, 44 and 46 corresponding to stations 1, 2, 3, 4, 5, 9, 10, 12, 13, 14, 15, 16, 18, 20, 22, 26, 28, 29, 30, 31, 32, 33, 34 and 37. These stations mainly belong to the 1995 campaign and are located in the NE part around Qsobriver. Class 3 contains neurons 8, 16, 24, 31 and 32 corresponding to stations 7, 11, 21, 23, 24, 27, 35. These stations, also belonging to the year 1995 and are located in the SW part of the study zone. This classification confirms the results obtained by the ascending hierarchical classification of the ACH-SOM self-organizing map.

Principal component analysis (ACP)

Study of the variables in the correlation circle

The correlation circle corresponds to a projection of the variables on a two-dimensional plane made up of two factors. Only variables located far from the center of the circle are taken into account. If they are close to each other, then they are significantly positively correlated (R close to 1). If they are orthogonal with respect to each other, then they are significantly uncorrelated (R close to 0). If they are symmetrically opposed with respect to the center, then they are relatively negatively correlated (R close to -1).

The analysis in the space of the variables was carried out in the factorial designs PC1- PC2 (Fig. 8) and PC1- PC3 (Fig. 9) which give a satisfactory account of the data structure and explain 71.5% some information.

The first principal component (PC1): A first grouping located in the positive values includes the elements Ca^{2+} , Cl^- , CE 25°C, Na^+ , Mg^{2+} and to a lesser degree i.e.b.

A second grouping located in the negative values includes PL (m) and rMg / rCa. The second principal component (PC2) correlates slightly with T (°C) in the positive part and with K^+ and HCO_3^- in the negative part.

The factors PC1 and PC3: A group of elements (Na^+ , Cl^- , SO_4^{2-} , Ca^{2+} , i.e.b.) are closer and therefore show a positive correlation.

Projection of individuals

The axes express a significant percentage of the information with respectively 42% by PC1 and 16.1% by PC2; which gives a total of 58.1% of the information (Fig. 10, 11)

The analysis of the descriptive statistical characteristics, the physicochemical variables used, reveals that the groundwater in the study area has temperatures ranging between 15.5 and 24.5 ° C, with an average of 21.02°C. The electrical conductivity oscillates between 770 and 5040 $\mu\text{s} / \text{cm}$ with an average of 2682.19 $\mu\text{s} / \text{cm}$ (Table 6).

Class 1 shows the highest Piezometric Level (PL), while class 3 shows the highest concentrations of Electrical Conductivity (CE), Cl^- and Na^+ compared to the other two classes (Table 6).

Table 6
Basic statistical quantities (Min, avg, max) for the physico-chemical parameters of different classes.

		PL	T	CE 25°	HCO ₃ ⁻	Cl ⁻	SO ₄ ²⁻	Ca ²⁺	Mg ²⁺	Na ⁺	K ⁺	rMg/rCa	i.e.b
		(m)	(°C)	($\mu\text{s}/\text{cm}$)	(mg/l)	(mg/l)	(mg/l)	(mg/l)	(mg/l)	(mg/l)	(mg/l)	(mg/l)	(mg/l)
Total base	Min	2.7	15.5	770	67.1	74.55	27.36	30	19.4	36.8	1.2	0.11	0.24
	Avg	38.52	21.02	2682.19	251.30	744.93	164.08	146.62	69.17	323.10	5.19	0.61	0.52
	Max	105.5	24.5	5040	412.4	1728.9	418.3	376	150	736	35.1	1.21	0.71
Classe I	Min	14	19	770	67.1	74.55	27.36	30	24.12	36.8	1.2	0.30	0.24
	Avg	46.21	21	1986.44	196.81	364.40	109.03	66.32	42.72	196.20	6.49	0.75	0.43
	Max	105.5	23	3060	346.5	586.8	187.1	146.8	73	266.8	19.5	1.06	0.62
Classe II	Min	5	15.5	2000	102.5	433.1	84.3	28	19.4	165.6	1.2	0.10	0.39
	Avg	37.54	21.25	2675.83	293.01	769.77	177.79	158.06	78.17	331.34	3.57	0.59	0.55
	Max	104	24.5	3580	412.4	1228.3	244.6	306	128.6	655.5	10.53	1.21	0.70
Classe III	Min	2.7	18	3490	175.7	1320.6	129.3	172.4	61.4	391	2.3	0.16	0.53
	Avg	24.27	20.29	4294.29	232.86	1529.57	242.91	290.91	98.76	584.86	7.79	0.36	0.61
	Max	52	22	5040	307.4	1728.9	418.3	376	150	736	35.1	0.59	0.71

Conclusion

The statistical study by Kohonen's self-organizing classification (SOM) of hydro-chemical data from the years 1995 in Mennani(2001) and 2009 in Chamchati (2014), showed an evolution of the values of the physicochemical parameters as a function of time (between 1995 and 2009) and space with an increase in the values of the parameters going from the center of the study area (Qsob river) to the southwest (Atlantic Oceans).

This increase is generally due to the influence of the Qsobriver (dilution of the water table), the diapiric influence (flush and / or sub-flush diapirs) and the influence of the Atlantic Ocean.

To get an idea of the classification of the samples and their sampling position, the coordinates of the stations were projected on the geological map of the study area. This projection showed that the class 1 samples are grouped around the Qsobriver. Feeding the water table by infiltration of fresh water from this river will dilute the concentrations of water in this class, which explains the low values of their parameters (Fig. 12).

The 2nd class is generally located in the eastern part where the Tidzidiapir outcrops. The erosion of the clay-salt outcrops of Triassic age (Diapir de Tidzi) will influence the synclinal basin and may contaminate the Plio Quaternary water table; this explains the increase in the values of the physico-chemical parameters compared to the 1st class.

The 3rd class is characterized by high values in most of the parameters probably related to its proximity to the Atlantic coast (infiltration of marine waters).

Declarations

Competing Interests

The authors declare that they have no competing interests.

Funding

This work is a part of a PhD thesis funded by Faculty of Science-Meknes, Moulay Ismail University Morocco.

Authors Contributions

All authors have made a significant contribution to this research. Zakaria Ouzerbane, Abdellah El Hmaid and Ali Essahlaoui conceived the research idea and designed the framework of the research. Zakaria Ouzerbane and Soumia Loulida participate in the fieldwork. Zakaria Ouzerbane, Soumia Loulida and Boughalem Mostafia wrote the paper. Mohamed Berrada and Habiba Oussmana contributed Data processing and model creation. Ouzerbane Zakaria revised the paper.

Acknowledgements

The logistic support during exploration and sampling was funded by the Faculty of Science and Technology, Sultan Moulay Slimane, University - Béni Mellal (Morocco). We are grateful to an anonymous reviewer for the fruitful remarks that helped to improve the manuscript.

Ethical Approval (Not applicable)

Consent to Participate (Not applicable)

Consent to Publish (Not applicable)

References

- Andreasen DC and Fleck WB (1997) Use of different potential sources of chloride and shallow, unconfined aquifer affected by brackish-water intrusion. *Hydrogeology journal*, V. 5, n°2: 17-26.
- Attwa M, Zamzam S (2019) An integrated approach of GIS and geoelectrical techniques for wastewater leakage investigations: Active constraint balancing and genetic algorithms application, *Journal of Applied Geophysics*, 175, 103992. <https://doi.org/10.1016/j.jappgeo.2020.103992>
- Bahir et al. 2019 Bahir, M., Ouazar, D., Ouhamdouch, S., 2019. Dam effect on groundwater characteristics from area under semi-arid climate: case of the Zerrar dam within Essaouira basin (Morocco). *Carbonates and Evaporites*, 34(3), 709-720. <https://doi.org/10.1007/s13146-019-00497-0>
- Bahir M, Jalal M, Mennani A (2001) Pollution nitraté des eaux souterraines du bassin synclinal d'Essaouira. *Journal of Environmental Hydrology*, 18, 9, 1-9.
- Chamchati H and Bahir M (2013). Potential hydrogeological, environment and vulnerability to pollution of the Plio-Quaternary aquifers of the coastal basin of Essaouira (Morocco). *Journal of Environmental Earth Sciences*, 3(10), 170-185.
- Chamchati H (2014) Evolution, gestion et protection des ressources en eau en zones semi-arides ; Exemple du bassin d'Essaouira. Thèse de l'Université Cadi Ayyad, Faculté des Sciences Semlalia. Marrakech, 143p.
- Choubert G and Faure-Muret A (1956) Lexique stratigraphique du Maroc. Notes Mém. Serv. Géol. Maroc, 134, 165p.
- Cohen K.M, Finney S.C, Gibbard P.L, Fan J.-X (2020) The ICS International Chronostratigraphic Chart. *Episodes*, 36, 199-204. <http://www.stratigraphy.org/ICSchart/ChronostratChart2020-03.pdf>
- Duffaud F, Brun & Planchot B (1966) Bassin du sud-ouest marocain. In :Reyre (ed), Bassins sédimentaire littoral africain. Ass. Serv. Géol. Africain. 1ère partie, Paris, p. 5-12.
- Karroum M, El Guettafi M, El Mandour A, Wilske C, Himi M, Casas A, (2017) Geochemical processes controlling groundwater quality under semi-arid environment: A case study in central Morocco. *Science of the Total Environment*, 609, 1140-1151. <https://dx.doi.org/10.1016/j.scitotenv.2017.07.199>
- Kohonen T (2001) Self-organizing maps. 3rd ed. Berlin, Springer.
- Kiviluoto K (1996) Topology preservation in self-organizing maps, IEEE International conference Neural Networks, pp. 294-299.
- Medina F (1994) Evolution structurale de Haut Atlas Occidental et des régions voisines du Trias à l'actuel, dans le cadre de l'ouverture de l'Atlas Atlantique Centrale et de la collision Afrique-Europe. Thèse d'état. Université Mohammed V, Rabat, 272p.
- Mennani A (2001) Apport de l'Hydrochimie et de l'Isotopie à la connaissance du fonctionnement des aquifères de la zone côtière d'Essaouira (Maroc occidental). Thèse Doct. Es. Sci, Uni. Cadi Ayyad, Marrakech, 172p.
- Oulaaross Z (2009) Etude climatologique, hydrogéologique et géophysique du Sahel Côtier des Doukkala (Maroc). Apport de l'analyse statistique et de l'inversion des données géoélectriques à l'étude du biseau salé de la lagune de Sidi Moussa. Thèse en cotutelle de l'Université Chouïab Doukkali et Université

Bordeaux 1, El jadida. 276p.

Ouzerbane Z, Boughalem M, El Hmaidi A, Aïfa T, Essahlaoui A, El Ouali A, Najine A, Redouani F, Wafik A (2019) Application of GIS to study the physiographic factors and the water resources in the watersheds of Essaouira, Morocco. *Journal of Water Sciences & Environment Technologies*, 4(1), 480-490, ISSN: 2508-9250,

Ouzerbane Z (2015) Contribution de la géophysique, des SIG et de l'analyse multicritère à la reconnaissance des ressources en eau et leur vulnérabilité à la pollution. Cas du sous-bassin côtier d'Essaouira (Maroc). Thèse de Doctorat, Université Moulay Ismail, Meknès, Morocco, 304p.

Ouzerbane Z, Najine A, Aïfa T, El Hmaidi A, Essahlaoui A (2014) Apport de l'hydrogéophysique et le SIG à l'étude des ressources en eaux dans la partie nord côtière du bassin d'Essaouira-Maroc. Colloque international des Utilisateurs de SIG, TAZA GIS815 Days, Morocco. *Infrastructure des données spatiales et Application des SIG*, 300-306. ISBN : 978-3-8381-4837-3.

Piqué A (1994) Géologie du Maroc. Les domaines régionaux et leur évolution structurale. Presse Universitaire Maghrébien Edition, Marrakech, 284p.

Rey J, Canerot J, Peybernes B, Taj-Eddine K, Rahhali I, Roch E (1950) Histoire stratigraphique du Maroc. Note et mémoire. Ser. Géol. du Maroc, Rabat, N°80, 440p.

Roch E (1931) Carte géologique provisoire des Abda, des Jbillets occidentales, de la zone synclinale de Mogador et de l'Atlas occidental.

Sherif M, El Mahmoudi A, Garamoon H, Shetty A (2006) Geoelectrical and hydrogeochemical studies for delineating seawater intrusion in the outlet of Wad, Ham, UAE. *Environmental Geology*, 49, 536–551

Suter G (1958) Géologie du plateau d'Akermoud (Maroc occidental). Note et mémoire du service géologie. Maroc, Tome 16, N° 143, p. 17-34.

Taj-Eddine K (1991) Le Jurassique terminal et le Crétacé basal dans l'Atlas Atlantique (Maroc), Biostratigraphie, Sédimentologie, Stratigraphie séquentielle et Géodynamique. Thèse d'Etats-Science. Université Cady Ayyad. Marrakech, 285p.

Urish, D W, Frohlich, R.K (1990) Surface electrical resistivity in coastal groundwater exploration. *Geoexploration*, 26, 267–289.

Vesanto J (1999) SOM-based data visualization methods. *Intelligent Data Analysis*, 3, 111-126.

Figures

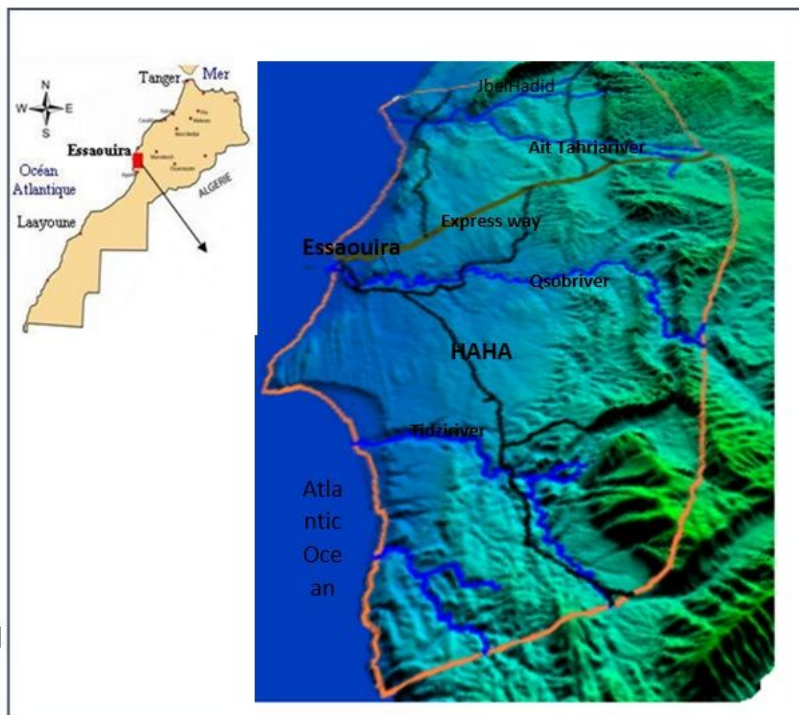


Figure 1

Situation map of HAHA zone of Essaouira in 3D mode by Global Mapper.

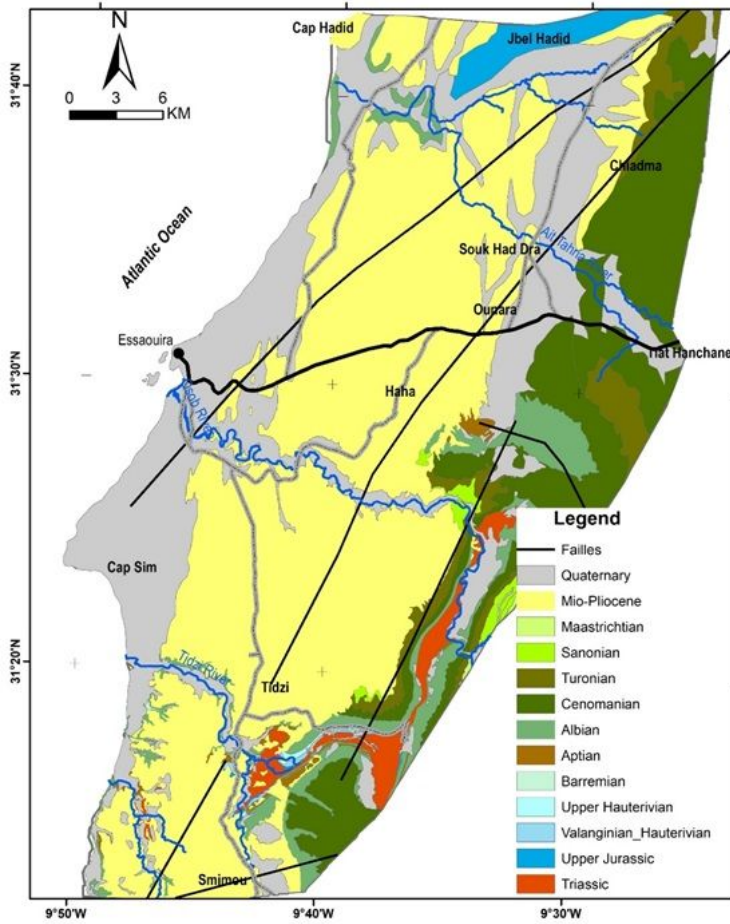


Figure 2

Carte géologique de la zone côtière d'Essaouira.

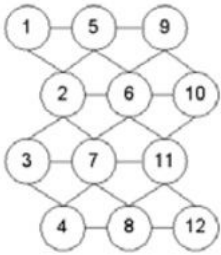


Figure 3

Neighborhood relations on the SOM.

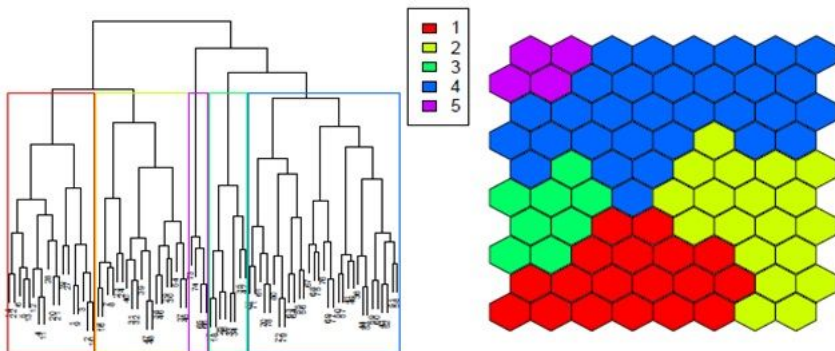


Figure 4

Example of hierarchical classification and separation of groups in SOM.

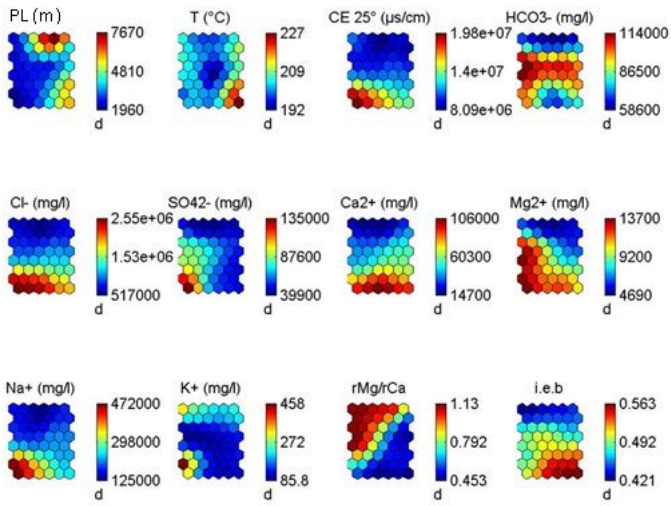


Figure 5

Gradient of values of the hydro-chemical parameters studied on the Kohonen map.

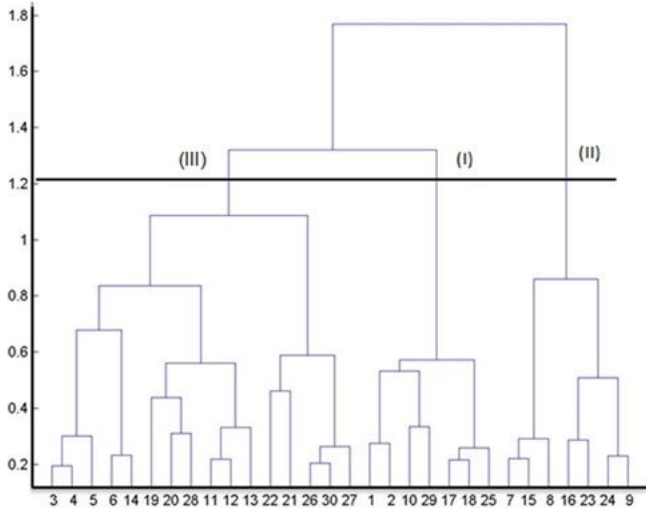


Figure 6

Dendrogram obtained with the hierarchical classification ACH-SOM.

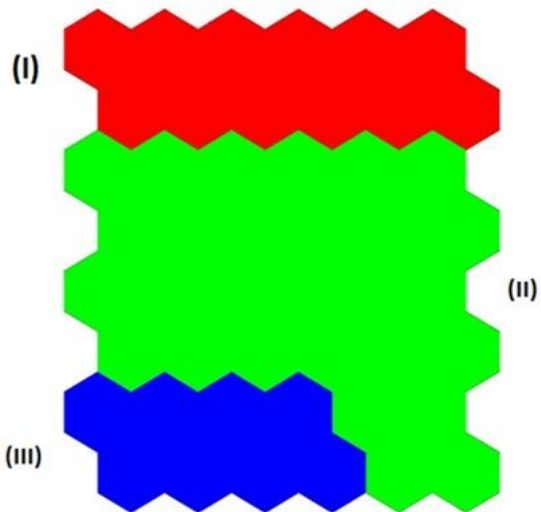


Figure 7

Class distribution on Kohonen's SOM map.

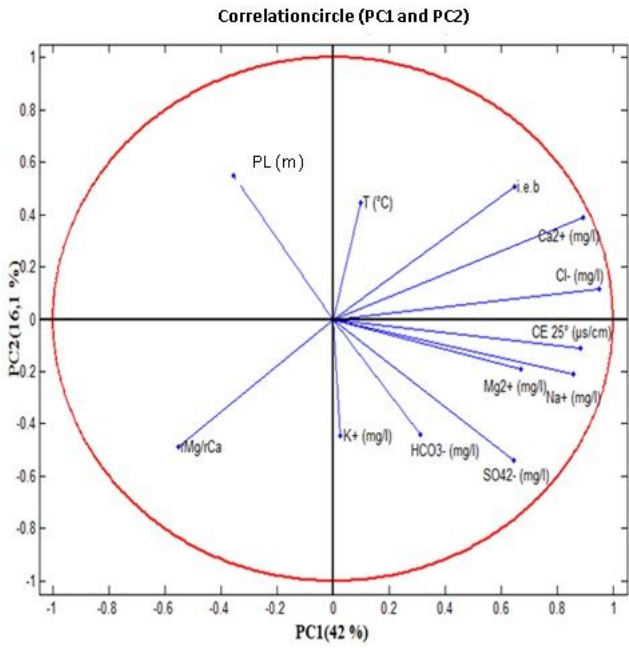


Figure 8

Graphical representation of the data in the factorial plane PC1 × PC2.

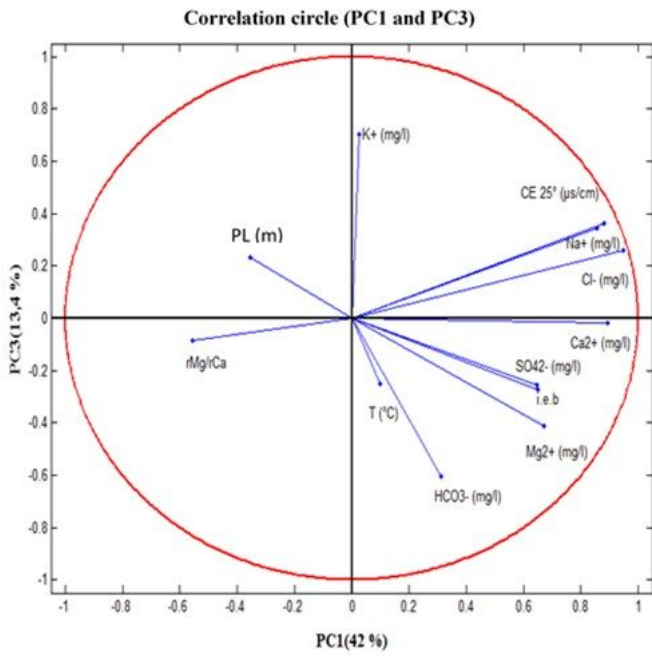


Figure 9

Graphical representation of the data in the factorial plane PC1 × PC3.

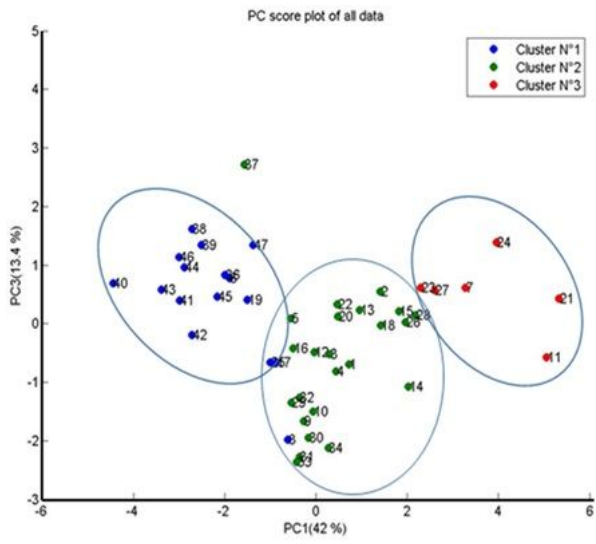


Figure 10

Projection of individuals on the factorial plane PC1 x PC3.

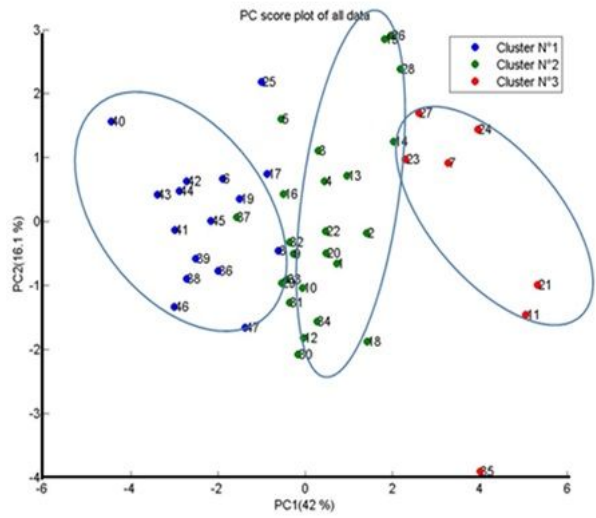


Figure 11

Projection of individuals on the factorial plane PC1 x PC2.

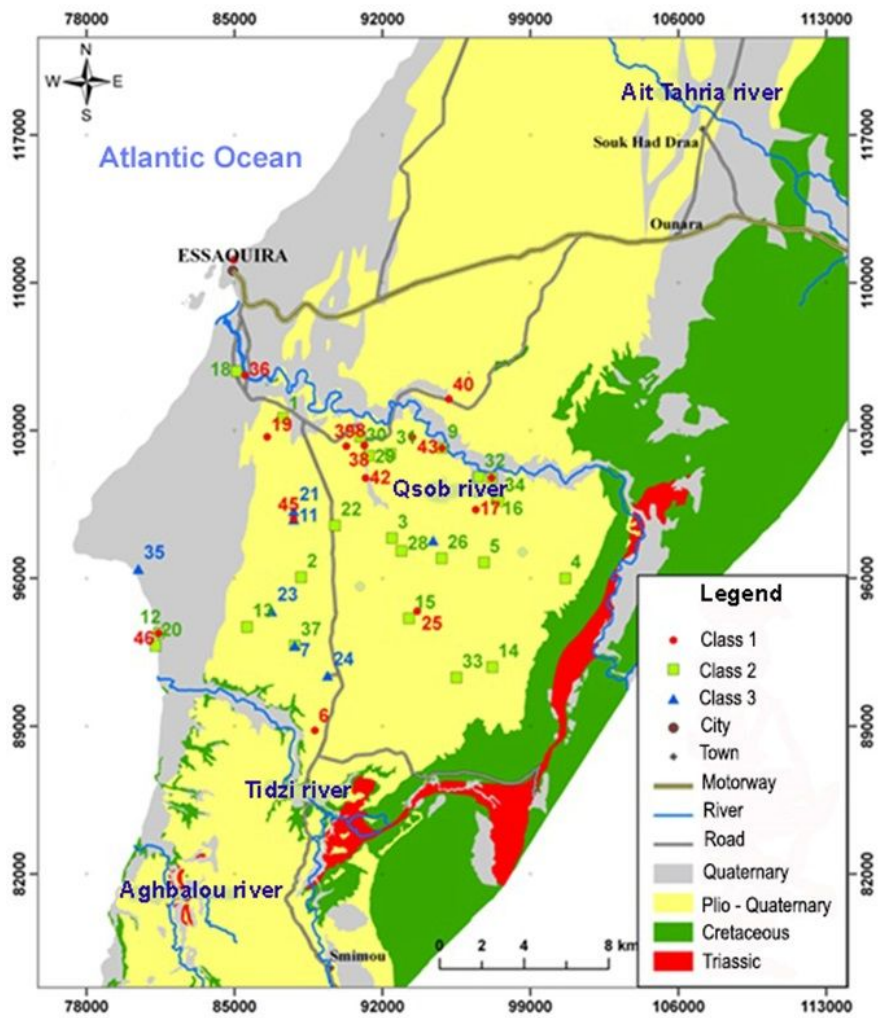


Figure 12

Projection of the 3 classes on the geological map.