

Statistical Analysis of Groundwater and Slate Samples with Special Influences on Rare Earth Elements

Dissertation
Zur Erlangung des akademischen Grades doctor rerum naturalium
(Dr. rer. nat)

Vorgelegt dem Rat der Chemisch- Geowissenschaftlichen Fakultät der
Friedrich- Schiller- Universität Jena



seit 1558

Von M.Sc. Anahita Pourjabbar
Geboren am 31.01.1981 in Teheran, Iran

Gutachter:

Prof. Dr. Georg Büchel

Institute of Geosciences, Friedrich Schiller University, Jena, Germany

Prof. Dr. Jürgen W. Einax

Institute of Inorganic and Analytical Chemistry, Friedrich Schiller University, Jena, Germany

Tag der öffentlichen Verteidigung: 02.04.2012

Zusammenfassung

Umweltkontaminationen und umweltrelevante Einflüsse auf die Gesundheit der Menschen sind wichtige Studiengebiete in den modernen analytischen Wissenschaften. Bergbauaktivitäten sind als Quelle für Kontaminationen in Bezug auf die umgebende Umwelt z.B. Wasser, Böden und Pflanzen bekannt. Die Art und die Auswirkungen der Kontaminationen hängen dabei sowohl von der Geologie und Geografie als auch von der Bergbautechnik wie Abbauverfahren und von den Verarbeitungsprozessen ab. Während und nach den Bergbauaktivitäten müssen daher Umweltstudien durchgeführt und Sanierungspläne entwickelt werden.

Diese hier vorgelegte Arbeit ist Teil eines großen Projektes der Universität Jena, das sich mit Sanierungsstrategien in der ehemaligen Uranbergbauregion von Ronneburg (Ostthüringen) beschäftigt. Die spezifischen Bedingungen in den Bergbauhinterlassenschaften wie niedriger pH-Wert und hohe Konzentration von Schwermetallen stellen weiterhin das Potenzial für eine kontaminierte Umwelt auch in Zukunft dar.

Multivariate Statistik incl. Clusteranalysen (R- und Q-Modus), Korrelations- und Faktorenanalysen, multivariate Statistik, Zeitreihenanalysen und Fuzzy-Analysen wurden benutzt, um die Kontaminationen selbst und die Muster der Kontaminationen im Untersuchungsgebiet zu untersuchen. Ebenfalls war das Ziel der statistischen Untersuchungen, die Beziehungen zwischen den Kontaminationen im Grundwasser, im Bodenwasser (Ammoniumnitrat- und Wasserextrakte) und den gelaugten paläozoischen Schieferbruchstücken. Mit Ausnahme der Schieferbruchstücke wurden die Wasser- und Bodenproben vom Testfeld „Gessenwiese“ genommen. Das Testfeld ist im nördlichen Bereich der Aufstandsfläche der Gessenhalde eingerichtet worden. Es besteht aus einer ca. 2000 m² große Fläche, auf der von Seiten der Angewandten Geologie und der Mikrobiologie der Universität Jena Experimente zur Remediation durchgeführt werden. Die Untersuchungen zielten vor allem auf die Fragestellung nach den möglichen Quellen der Kontaminationen und damit der Optimierung der Sanierungsstrategien. In Anbetracht der umfangreichen Kenntnisse über das generelle Verhalten der Schwermetalle wurden Vergleichsstudien an vergleichbaren Daten an zwei Standorten in Schweden durchgeführt. Um das Gesamtziel des Projektes zu erreichen, wurden die Seltenen Erden Elemente (SEE) und ihr spezifisches geochemisches Verhalten mit in die Untersuchungen aufgenommen. Die SEE sind im Untersuchungsgebiet reichhaltig vertreten (bis zu 8148 µg/l).

Die Grundlage für die statistischen Untersuchungen stellen die analytischen Ergebnisse von 174 Grundwasserproben, 120 Laugungsproben, 90 Bodenwasser und 10 Schiefer dar. Mit Hilfe der statistischen Untersuchungen konnten folgende Aussagen zur räumlichen Verteilung der Kontaminationen getroffen werden. Bezüglich der SEE-Konzentration in den Grundwasserproben des Testfeldes ist festzuhalten, dass die Konzentrationen generell von Süden nach Norden zunehmen. Es ergeben sich zwei Gruppen: SEE und Al, Cu, F⁻, Fe, Li, PO₄³⁻, Sc, Th, Ti, U und Y sind einer Hauptgruppe zuzuordnen und Ca, Cd, Co, Cl⁻, DOC, HCO₃⁻, Mg, Mn, Ni, Pb, pH, Zn, und SO₄²⁻ einer zweiten. Der pH-Wert und die lokale Geologie (unterschiedliche Verteilung der unterschiedlichen glazigenen Sedimenttypen) sind die Hauptfaktoren, die die SEE-Konzentration und die Verteilung im Untersuchungsgebiet bestimmen. Die Proben der zentralen und südlichen Bereiche gehören zu einer Domäne. Der südliche Bereich ist dominiert durch glazigenen Sand und der zentrale Bereich durch mehr bindige Materialien wie Ton und Silt. Es konnten keine saisonalen Effekte erkannt werden, die die Konzentration beeinflussen. Die Fluktuation des Grundwasserspiegels, der auf die Niederschläge

mit einer gewissen Zeitdifferenz von ein paar Tagen reagiert (z.B. Juni-Oktober 2007 in dieser Studie), beeinflusst die SEE-Konzentration zwar sehr wenig doch signifikant, die andere Elementverteilung jedoch nicht.

Mit Hilfe der statistischen Analysen wurde der Frage nachgegangen, woher die Kontamination der ursprünglich unkontaminierten Auftragsböden kommt. Der Vergleich der SEE-Muster zwischen dem Grundwasser und den Wasser- und Ammoniumextrakten in dem aufgetragenen ursprünglich unkontaminierten Böden (mehrere Dezimeter) zeigt semiquantitativ betrachtet ähnliche Muster. Alle Proben zeigen eine Anreicherung der mittleren SEE und eine Abreicherung der leichten in Relation zu den mittleren. Alle Proben mit Ausnahme der Wasser-Extrakte zeigen eine positive Ce-Anomalie, manche Proben eine leicht positive Gd-Anomalie. Bei den Wasserextrakten zeigt sich eine Gruppe, der SEE, Cu, U und Y zugehören. Im Unterschied zum Grundwasser sind die Elemente Al, Fe und Th nicht in der gleichen Gruppe mit den SEE. Diese Elemente bilden zusammen mit Pb und PO_4^{3-} eine andere Hauptgruppe. In den Ammoniumextrakten sind die SEE mit Ausnahme von La ähnlich zu Al, Cu, Fe, Pb, U und Y. Eine zweite Hauptgruppe schließt La, Mg, Mn, S und Zn ein. Das zugehörige Dendrogramm ist ähnlich dem des Grundwassers. Die Nicht-Übereinstimmung zwischen Grundwasser und den Wasserextrakten kann durch die unterschiedlichen Systeme erklärt werden. Tatsächlich ist der Austausch der löslichen Elemente zwischen Wasser und Boden hoch. So kann auch Grundwasser den Boden beeinflussen, z.B. durch „return flow“ bzw. durch kapillaren Aufstieg und zusätzlich spielen Mikroben und Pflanzen in Bezug auf die Schwermetalle eine Rolle.

Mit Hilfe von Fuzzy-Analytik wurde der Frage nachgegangen, ob die Schieferfragmente, die man heute noch in dem Auftragssubstrat vereinzelt finden kann, die Ursache für die Kontaminationen sind: Die Schieferfragmente können hiernach nicht die Ursache der Schwermetallkontamination in den Boden- und Grundwasserproben sein. Zwar existiert insbesondere eine sehr gute Übereinstimmung zwischen den Schieferproben und der Bodenproben MF3 in 100 cm Tiefe, bei den anderen Bodenproben MF1 und MF2 und MF 3 in 60 und 40 cm Tiefe ist die Ähnlichkeit zwar vorhanden, doch müssen hier auch andere Quellen zur Erklärung herangezogen werden. Die zugrundeliegenden Prozesse sind komplizierter als gedacht. Die Kontamination ist zwar ursprünglich auf den Laugungseffekt der paläozoischen Armerze auf der Gessenhalde zurückzuführen, doch gibt es mehrere Phasen der Infiltration und Neubildung von Präzipitaten, bei denen sich die Muster deutlich ändern konnten und nicht alle diese Änderungen lassen sich mit Hilfe der Statistik zurückverfolgen.

Abstract

Contamination of the environment and the environmental impact on human health is one of the most important fields of studies in modern analytical science. Mining activities are known as a source of contamination on the surrounding environment such as water, soil, and plants. The type and the degree of contamination are related to the geology and geography, as well as to the mining technique such as excavation, and to the mineral processing methods. Hence, during and after the period of mining activities, environmental studies, and remediation plans are necessary. This study is a part of a larger project that focuses on the remediation strategy for Ronneburg, a former uranium mining site in eastern Germany. The specific conditions of this abandoned mining site, such as low pH range and high concentration of heavy metals, gives the study area the potential conditions for a contaminated environment, which explains why the spread of contamination is still on-going. Multivariate statistical analyses including cluster analysis (R- and Q-mod), cross-correlation, factor analysis, multivariate outlier detection, time series analysis, and fuzzy clustering, have been used to investigate the contamination distribution/pattern of the study area, and also to investigate the relation between samples of various media such as groundwater, soil water, soil (leached with water and ammonium nitrate), and slate samples. Excepting the slates, the samples were collected from a Gessenwiesse (Test site). This test site (about 2000m²) were established at the base of the former leaching heap (Gessenhalde) in order to be used as a test area for geological and microbiological studies as well as remediation strategy. The results were used to identify the probable contamination sources and also to improve and optimize the remediation strategy. In order to expand the knowledge of the general behavior of heavy metals, similar investigations have been performed on a data set from a similarly contaminated site located in Sweden. To achieve the goal, the focus of the study was on the behavior of rare earth elements (REE). The reason is the high concentration of these elements (total REE: 8148.4 µg/l in groundwater) and also their specific geochemical behavior.

The statistical methods were used for 174 groundwater samples, 33 soil water samples, 90 soil samples, and 10 slate samples. By means of classical and fuzzy statistical analysis, the following results were obtained:

The REE concentration in the groundwater samples of the test site is increasing from south to north. A general south-north gradient exists, although the REE concentration is not completely homogeneous. With regards to the results of clustering, the analytical parameters are in two main clusters: REE and Al, Cu, F⁻, Fe, Li, PO₄³⁻, Sc, Th, Ti, U and Y form one main cluster; and Ca, Cd, Co, Cl⁻, DOC, HCO₃⁻, Mg, Mn, Ni, Pb, pH, SO₄²⁻ and Zn form another main cluster. The pH and local sedimentology are the important factors that affect the REE concentration and distribution in the studied area. The samples collected from the central and southern part of the test site belong to different data domain, which emphasizes the role of the local sedimentology. The southern part of the test site is dominated by sand, the middle and the northern parts by silt and silty sand. No specific seasonal effect was found by using statistics. Groundwater fluctuation shows a meaningful correlation to precipitation data during the period of increased precipitation (June-October 2007). The time lag of groundwater response to precipitation is an average of three to four days. This time lag affects the REE concentration after the time lag period very slightly, but it has no effect on general contamination distribution.

The comparison of the REE patterns (groundwater, soil leached with water, and soil leached with ammonium nitrate) shows similar qualitative patterns. All samples exhibit a MREE-enrichment and LREE-depletion with respect to the HREE. Furthermore, all samples, excepting some water-leached samples, show a positive Ce anomaly. Moreover, some samples also exhibit a slight positive Gd anomaly.

The clustering of data from soil leached with water consists of REE, Cu, U and Y. Unlike the groundwater samples, the elements Al, Fe and Th are not in the same cluster with REE. These elements together with Pb and PO_4^{3-} are forming another main cluster. With regard to the ammonium-nitrate leached samples, REE with exception of La are clustered with Al, Cu, Fe, Pb, U and Y. The second main cluster includes La, Mg, Mn, S and Zn. The dendrogram resulting from this data set is similar to that of the groundwater samples.

The dissimilarities observed between the soil leachates and the groundwater can be explained by the difference of their systems. The site is influenced by fluctuating groundwater levels that lead to sporadic return flow during high precipitation events. While the fluctuating groundwater level can make possible a high degree of exchange of (soluble) elements between water and soil, it may also be possible that the action of microorganisms and plants play a role in the fate of metals within the upper soil level.

With regards to these processes and the results using fuzzy analysis, the composition of the slates is not the sole factor leading to heavy metal contamination in the soil water and soil samples. There is a strong statistical similarity between slates samples and two soil water samples (MF1 and MF2), while a great dissimilarity between the sampling point MF3 and other samples. This dissimilarity is likely due to the local soil composition near the sampling points MF1, 2 and 3. This hypothesis is also supported by dissimilarity observed between samples from different depths of the same sampling point, namely MF3. The soil water samples collected from MF3 at a depth of 100 cm below the surface level was not within the same clusters as the samples from the depths of 60 and 30 cm. Furthermore, the MF3/100cm soil water samples formed a cluster together with the heavy metals rather than slate samples, meaning that the slates do not have a statistical relationship to the contamination at this depth and location. In summary, these statistical analyses lead to the conclusion that the composition of the slates is the source of the heavy metals but the role of other processes is also important. The contamination in this area repeatedly undergoes various processes such as chemical precipitation and re-dissolution, due to various infiltration mechanisms that are on-going. To investigate all these mechanisms requires more detailed studies combined with statistical analysis.

چکیده

معدنکاری و فعالیت های جانبی آن از آلاینده های زیست محیطی شناخته شده به شمار می آیند. این صنعت همواره تهدیدی برای محیط زیست اطراف خود می باشد و هرگونه سهل انگاری در پالایش آلودگی های ناشی از آن، خسارت های جبران ناپذیری را برای محیط زیست اطراف به همراه خواهد داشت. میزان و نوع این آلودگی ها وابسته به عواملی گوناگون چون کانی/ کانه استخراج شده، روش استخراج و کانه آرای و خصوصیات زمین شناختی منطقه می باشد. از این رو، توجه به پالایش محیط زیست در طول حیات یک معدن و پس از آن امری است مهم.

مطالعه حاضر، بخشی از پروژه ای بزرگ با هدف پالایش آلودگی های زیست محیطی به جا مانده از معدن متروکه رونبورگ، واقع در شرق جمهوری فدرال آلمان می باشد. نهشته اروانیوم رونبورگ با تولیدی معادل 113000 تن ماده معدنی استخراج شده سومین معدن اورانیوم بزرگ جهان در زمان فعالیت (1949 تا 1990 میلادی) خود به شمار می آمده است. با گذشت دو دهه از خاتمه معدنکاری و همچنین انجام اقدامات اساسی جهت پالایش و جلوگیری از نشر بیشتر آلاینده ها، محدوده مورد نظر همچنان در زمره نواحی آلوده به شمار می آید؛ که از ویژگی های آن می توان به میزان اسیدیته و غلظت بالای فلزات سنگین در انواع نمونه های برداشت شده اشاره کرد. در این پژوهش، نمونه های برداشت شده از خاک، آب زیرزمینی، آب خاکی، و اسلیت مورد پردازش های آماری چند متغیره شامل، خوشه بندی سلسه مراتبی کلاسیک و فازی، آنالیز فاکتوری، همبستگی تقاطعی، مطالعه روند زمانی، و جدایش چند متغیره نمونه های خارج از رده قرار گرفته اند.

هدف از مطالعه تعیین شباهت ها و تفاوت های آماری نمونه ها با تمرکز بر رفتار فلزات سنگین و به خصوص عناصر نادر خاکی، می باشد؛ و از نتایج به دست آمده برای بررسی سرچشمه آلودگی، تفسیر الگوی پراکندگی آلاینده ها (فلزات سنگین)؛ و همچنین بهینه سازی طراحی عملیات پالایش استفاده گردیده است. در ابتدا، نمونه های مذکور از محدوده ای موسوم به «منطقه آزمایش» برداشت شده اند. منطقه آزمایش به وسعت تقریبی 2000 متر مربع، در سال 2004 میلادی بر بقایای به جا مانده از تپه تخلیه زهاب اسیدی معدن رونبورگ، و به منظور انجام مطالعات زمین شناسی و میکروبیولوژی بنا گردیده است.

در ادامه و در راستای گسترش بانک داده های زیست محیطی با تمرکز بر الگو و پراکندگی آلاینده های زیست محیطی در قاره اروپا، مطالعه ای مشابه بر روی نمونه های آب سطحی از سه محدوده به جا مانده از معادن متروک واقع در کشور سوئد انجام گردید.

نتایج به دست آمده از بررسی 90 نمونه خاکی، 174 نمونه آب زیرزمینی، 33 نمونه آب خاکی و 10 نمونه اسلیت به شرح زیر می باشد:

غلظت عناصر نادر خاکی در منطقه آزمایش دارای روند تدریجی افزایشی ملایم از جنوب به شمال می باشد. این روند کلی در حالی مشاهده می شود که غلظت این عناصر و سایر فلزات سنگین ناهمگن است و در مواردی مقادیر بالایی از این عناصر در شمال و در شرق محدوده اندازه گیری شده اند. بر اساس نتایج حاصل از خوشه بندی سلسله مراتبی کلاسیک نمونه های آب زیرزمینی، عناصر در دو خوشه اصلی قرار میگیرند: عناصر نادر خاکی، به همراه آلومینیوم، آهن، اورانیوم، ایتریوم، توریم، تیتان، سزیم، لیتیم، مس و همچنین آنیون های فسفات و فلوئورید در یک خوشه اصلی؛ و عناصر روی، سرب، کادمیوم، کبالت، کلسیم، منگنز، منیزیم، نیکل و آنیون های کلرید و کلسیت به هم راه عامل اسیدیته (pH) در دیگر خوشه اصلی قرار دارند. با توجه به خوشه بندی به دست آمده، می توان عامل اسیدی و زمین شناسی منطقه را از فاکتور های اصلی کنترل کننده پراکندگی آلاینده ها به شمار آورد. این تئوری با کمک نتایج حاصل از جدایش چند متغیره نمونه های خارج از رده تایید می شود. بر اساس نتایج این روش، نمونه های برداشت شده از شمال و جنوب به دو دسته جداگانه تعلق دارند. نکته قابل اهمیت در جدایش این دسته ها، همپوشانی آنها با رسوب پوششی منطقه آزمایش است: رسوب شن در بخش جنوبی تا میانی، و رسوب سیلت در بخش شمالی. بر خلاف تصور، و با وجود تاثیر نوسان سطح آب زیرزمینی بر میزان غلظت عناصر نادر خاکی، الگوی پراکندگی آلاینده ها همبستگی معنی داری با میزان های بارش فصلی ندارد، و می توان پراکندگی آلاینده ها پایدار دانست.

مقایسه نتایج پردازش های آماری بر روی انواع نمونه های مطالعه شده، نشان دهنده شباهت الگوی عناصر نادر خاکی در نمونه های آب زیرزمینی و نمونه های خاکی (پس از انحلال در آب و نیترات آمونیوم) می باشد. وجود حوضه آب زیرزمینی بسیار کم عمق و جریان برگشتی، از عواملی هستند که این مقایسه را معنی دار می کنند. این الگو، دارای مقدار افزاینده عناصر نادر خاکی میانه، و کاهنده عناصر نادر خاکی سبک در برابر عناصر نادر خاکی سنگین است. بر اساس پردازش های آماری، نمونه

های خاک (پس از انحلال در نیترات آمونیوم و آب)، به ترتیب نتایج مشابه و متفاوتی با نتایج حاصل از آب زیرزمینی نشان داده اند. برای مثال، در نمونه های خاک محلول در نیترات آمونیوم، نتایج خوشه بندی کاملاً مشابه با نمونه های آب زیرزمینی است. این در حالی است که، در خوشه بندی نمونه های خاکی محلول در آب، عناصر آلومینیوم، آهن و توریم در یک خوشه اصلی به همراه عناصر نادر خاکی جای نگرفته اند؛ بلکه به همراه سرب و فسفات یک خوشه اصلی جداگانه تشکیل می دهند.

مطالعات مشابه بر روی نمونه ای اسلیت (به عنوان سرچشمه احتمالی آلودگی) به همراه روش خوشه بندی فازی تقاطعی، بیانگر تشابه های آماری میان این نمونه ها و نمونه های آب خاکی است. نکته قابل توجه این است که این تشابهات در مورد تمام نمونه های آب خاکی صادق نمی باشد. از سه محل نمونه گیری شده، تنها دو محل رفتار آماری کاملاً مشابه با نمونه های اسلیت (پس از انحلال در آب) نشان می دهند؛ و محل سوم رفتاری کاملاً جداگانه دارد. وجود این رفتار جداگانه را می توان به ترکیب لایه های رسوبی موجود در محل نمونه گیری نسبت داد. این وتئوری زمانی تقویت پیدا میکند که نمونه مذکور (نمونه با رفتار متفاوت) مورد پردازش های دقیق تر قرار میگیرند و وجود خوشه بندی های متفاوت بر اساس عمق نمونه برداری آشکار می گردد. بدین ترتیب که در این محل نمونه گیری، نمونه های گرفته شده از عمق 100 سانتیمتری زیرسطح زمین، در یک خوشه و سایر نمونه های برداشت شده از عمق 30 و 60 سانتیمتر در خوشه ای دیگر قرار دارند. بر خلاف تصور، نمونه های برداشت شده از عمق 100 سانتیمتری این مکان نمونه برداری دارای همبستگی آماری معنی دار با عناصر نادر خاکی هستند؛ و نه نمونه های اسلیت. این نتیجه دور از انتظار، به همراه نتایج به دست آمده از سایر پردازش ها، این اندیشه را تقویت میکند که سرچشمه آلودگی های پایدار در منطقه نه وجود سنگ بستر آلوده بلکه وجود فعل و انفعالات شیمیایی پیوسته در منطقه است که آزاد سازی فلزات سنگین در محدوده را به طور مستمر بر عهده دارند.

Table of Contents

Introduction

1	Introduction	1
2	Study area	1
3	Geology and hydrology setting	4
4	Aim of study	5
5	Material and method	7
5.1	water movement within porous	7
5.1.1	Groundwater	7
5.1.2	Soil water	8
5.2	Data sets	9
5.2.1	Groundwater sample	9
5.2.2	Soil water sample	11
5.2.3	Slate samples	11
5.2.4	Soil samples	11
5.2.5	Groundwater level fluctuation and precipitation	12
5.3	Methods	12
5.3.1	Multivariate outlier detection	13
5.3.2	Cluster analysis	14
5.3.3	Factor analysis	14
5.3.4	Cross-correlation	14
5.3.5	Partial last square	15
5.3.6	Fuzzy clustering	15
5.3.7	Time series analysis	17
6	Results and discussion	18
6.1	Elements distribution/ pattern in groundwater	18
6.2	Elements distribution in soil	21
6.3	REE distribution in slate samples	23
6.4	Heavy metals in lake Norrtorpsjön and lake Pölen	23
6.5	The behavior of heavy metals before and after remediation	24
7	Conclusion	25
	References	27

Manuscript 1

Statistical evidence of groundwater contamination and significant factors with special emphasis of REE in a former uranium mine, eastern Thuringia, Germany

	Abstract	33
1	Introduction	34
2	Geology and hydrogeological setting	35
3	Material and methods	35
3.1	Sampling and data set	35
3.2	Statistics	37
3.2.1	Data preparation	37
3.2.2	Multivariate outlier detection	38
3.2.3	Cluster analysis	39
3.2.4	Cross-correlation	40
4	Results and discussion	41
4.2	Elements distribution	41
4.3	REE and related analytical parameters	41
4.4	Data domain	45
4.5	Groundwater level fluctuation	47
5	Conclusion	48
	References	51

Manuscript 2

Top soil heavy metals and rare earth elements contamination at a remediated former uranium mining site, eastern Thuringia, Germany

	Abstract	56
--	----------	----

1	Introduction	57
2	Material and methods	58
2.1	Sampling and sample preparation	58
2.1.1	Groundwater sample	58
2.1.2	Soil samples	59
2.1.3	Analysis	60
2.2	Statistical methods	62
2.2.1	Hierarchical cluster analysis	62
2.2.2	Factor analysis	62
3	Results	63
3.1	General chemical analysis	63
3.2	REE pattern	64
3.3	Cluster analysis	65
3.4	Factor analysis	68
4	Discussion	69
5	Conclusion	71
	References	73

Manuscript 3

Fuzzy, hierarchical cross-clustering of slate samples contaminated with heavy metals

	Abstract	76
1	Introduction	77
2	Sampling site description	78
3	Material and methods	79
3.1	Fuzzy clustering	79
3.2	Fuzzy divisive hierarchical clustering	80
3.3	Fuzzy hierarchical cross-clustering algorithm	81
3.4	Samples and analysis	81
4	Results and discussion	83
4.1	Fuzzy hierarchical clustering	84
4.2	Fuzzy hierarchical cross-clustering	88
5	Conclusion	91
	References	94

Manuscript 4

Distribution of rare earth elements and other metals in a stratified acidic pit lake in black shales 45 years after mine closure

	Abstract	98
1	Introduction	99
2	Material and methods	99
2.1	The field site	100
2.2	Sampling and analysis	101
3	Results and discussion	101
3.1	General hydrochemistry	101
3.2	Trace metals	104
3.3	Rare earth elements	105
3.4	Statistical interpretation	105
3.5	Environmental consequence	106
4	Conclusion	106
	References	107

Manuscript 5

Dry Covers on Historical Sulphidic Mine Waste – Long Term Statistic Performance Assessment of Surface Water Quality

	Abstract	109
1	Introduction	110
2	Material and methods	110
2.1	The Bersbio site	110
2.2	Remediation	112

2.3	Surface water monitoring	112
2.3.1	Chemical analysis	112
2.3.2	Statistical analysis	112
2.3.2.1	Hierarchical cluster analyses	112
2.3.2.2	Time series analyses	113
3	Results	115
3.1	Lake Gruvsjön	115
3.2	Kuntebo	117
3.3	Lake Risten	118
4	Discussion and conclusion	120
	References	122

List of Tables

Introduction

Table 1	The maximum, minimum, and standard deviation of the analytical results of	2
---------	---	---

Manuscript1

Table 1	Numbers of censored data in analytical result and used replacement method	38
Table 2	Analytical parameters and descriptive analysis of the analytical results	42

Manuscript2

Table 1	Chemical analyses of groundwater samples and leached soil samples	61
Table 2	Factor loading matrix after Varimax rotation of groundwater and leached soil	69

Manuscript3

Table1	The minimum, maximum, and standard deviation of the analytical results	83
Table 2	Fuzzy divisive hierarchical clustering of samples	85
Table 3	Fuzzy divisive hierarchical clustering of variables	87
Table 4	Membership degrees of samples and variables to the clusters	89

Manuscript 4

Table1	General hydrochemical parameters	102
Table2	Principal cation concentrations	103
Table 3	Selected trace metal concentrations	104
Table 4	Selected trace rare earth elements concentrations	105

Manuscript 5

Table 1	Selected rare earth element concentration	114
---------	---	-----

List of of Figures

Introduction

Fig. 1	Gessenhalde and the test site	3
Fig. 2	Location of the groundwater, soil water, soil and slate samples	10
Fig. 3	Distribution of Al, Mn , REE and U	19
Fig. 4	The groundwater fluctuation of sampling points 6 and 13	21
Fig. 5	Three phases of infiltration:	26

Manuscript1

Fig.1	Map of the test site Gessenwiese	36
Fig. 2	Result of the R-mode cluster analysis	42
Fig. 3	Scatter plot showing REE concentration	44
Fig. 4	Results of the Q-mode cluster analysis	46
Fig. 5	Scatterplot showing REE versus pH	46
Fig. 6	Groundwater level and daily precipitation	48
Fig. 7	Cross-correlation of groundwater level	49

Manuscript2

Fig. 1	Location map of the studied area	59
Fig. 2	PAAS-normalized REE patterns of selected groundwater	64
Fig. 3	PAAS--normalized REE patterns of selected samples; leached by water	65
Fig. 4	PAAS-normalized REE patterns of samples from an amended plot	65
Fig. 5	Dendrogram of groundwater samples	66
Fig. 6	Dendrogram of water leached samples;	67
Fig. 7	Dendrogram of ammonium nitrate leached samples	67
Fig. 8	The distribution of REE in the soil sampling zone regarding	68
Fig. 9	Schematic representation of underground processes	72

Manuscript3

Fig. 1	Sampling locations	82
Fig. 2	Q-mode cluster analysis	84
Fig. 3	Scatter plots of the membership degrees of sub-clusters of A1 and A2 by	88
Fig. 4	Scatter plots of the membership degrees of sub- clusters of A2	88
Fig. 5	Scatter plots of the membership degrees of sub- clusters of A1 and A2	89
Fig. 6	Three phases of infiltration	92

Manuscript 4

Fig. 1	Map of the area	100
Fig. 2	Examples of depth profiles, lake Norrtorpssjön	101
Fig. 3	Examples of concentration profiles of selected elements	103

Manuscript 5

Fig. 1	Map of the study area	111
Fig. 2	Q-mode cluster analysis of Lake Gruvsjön	115
Fig. 3	R-mode cluster analysis of Lake Gruvsjön	115
Fig. 4	Time series analyses with measured and predicted concentration	116
Fig. 5	Q-mode cluster analysis of Kuntebo creek	117
Fig. 6	R-mode cluster analysis of Kuntebo creek	117
Fig. 7	Time series analyses of Kuntebo data	118
Fig. 8	Q-mode cluster analysis of Lake Risten	119
Fig. 9	R-mode cluster analysis of Lake Risten	119
Fig. 10	Time series analyses of Lake Risten data	120

Introduction

1 Introduction

Contamination of the environment and the environmental impact on human health is one of the most important fields of studies in modern analytical science. It is known that many industries can seriously contaminate the environment and endanger human health. Mining activities are known as a source of contamination on the surrounding environment such as water, soil, and plants. The type and the degree of contamination is related to the geology and geography as well as the mining technique such as excavation, and the mineral processing methods (Bhattacharya et al., 2006; Chopin and Alloway, 2007; Gomes and Favas, 2006; Lee, 2003; Morento et al., 2007; Rodríguez et al., 2009). Acid mine drainage (AMD) results from a mining process and it affects the surrounding environments. It is polluted water that typically contains high levels of metals, including heavy metals, e.g. rare earth elements (REE) (Hadley and Snow, 1974). It is produced by the oxidation of sulfide minerals, chiefly pyrite or iron disulfide (FeS_2). This is a natural chemical reaction which can proceed when minerals are exposed to air and water and is found around the world. Although remediation strategies vary from one location to another depending on the environmental, economic, and technical situation, the monitoring of water, plant, and soil contamination levels is always helpful to evaluate and improve the remediation process (Bozau et al., 2007; Elias and Gulson, 2003; Franklin and Fernandes, 2011; Otte and Jacob, 2008). Such characterizations studies are composed by data tasks that are influenced by many factors and thus, multivariate analyses are helpful tools, since they place the factors into more or less homogeneous groups so that the relation between the groups is revealed.

The current study is a part of a project that focuses on an abandoned-mine area located in eastern Thuringia, Germany. Although the mining activities were stopped in 1990s, the surrounding area is still contaminated by heavy metals. The specific conditions such as low pH and high concentration of certain heavy metals make this area a potential condition for contaminated environment which is still on-going (Table 1). These concentrations are remarkably high compared to the permissible threshold for drinking water (Bundesgesetzblatt, 2011). The contamination in various media such as groundwater, soil water, surface water, soil, and plants was measured and reported in previous studies (Carlsson and Büchel, 2005; Grawunder, 2010; Haferburg, 2007; Horn, 2003; Lonschinski, 2009; Lorenz, 2009; Mirgorodsky et al., 2010; Mirgorodsky et al., 2012; Ollivier et al., 2010). Identify the source of contamination as well as the contamination distribution in the different media, and also the relation between them are important tasks. In this study, the data from the various media (groundwater, soil water, soil, and slate samples) were analyzed statistically in order to find the relationship between the media, and relying upon these results we are able to identify the probable contamination sources and hence to optimize the remediation strategy.

2 Study area

Gessenhalde: The former uranium mine district in eastern Thuringia, with more than 113,000 tons of mined uranium, was the third-largest uranium producer in the world (Jakubick et al., 2002; Lange, 1995).

Mining activities started in 1949 and ended in 1990 after the re-unification of Germany. The remnants of the mine include: a) a large open pit mine, Lichtenberg (an open pit with more than

Table 1 The maximum, minimum, and standard deviation of the analytical results of the samples that were used in this study; 174groundwater samples; 53 soil water samples, 90 soil samples, and 10 slates

	Min.	Max.	Stdev.	Min.	Max.	Stdev.	Min.	Max	Stdev.	Min.	Max	Stdev.	Min.	Max	Stdev.
	Groundwater			NH ₄ NO ₃ -leached soil			H ₂ O-leached soil			Soil Water			H ₂ O-leached slate		
Al	0.21	308.12	50.97	26.5	254.3	40.7	0.1	97.2	1.6	7.87	294.30	73.10	11.32	661.58	213.44
Ba	13.6	70.3	10.7	8.92	66.14	10.08	-	-	-	5.00	46.67	10.60	0.34	1.98	0.48
Ca	77.33	672.24	59.69	-	-	-	43.3	659.2	141.8	195.67	573.33	71.85	13.26	2553.86	804.48
Cd	3.92	601.21	81.56	0.17	0.55	0.07	-	-	-	16.67	403.00	78.36	0.20	0.98	0.28
Ce	3.9	4127.2	566.4	1.51	6.12	0.80	0.003	0.1	0.05	81.63	1183.67	226.86	0.12	0.90	0.32
Co	85.3	20120.2	3331.8	-	-	-	0.01	1.83	0.02	227.67	9605.00	1793.06	0.18	3.48	1.04
Cu	0.4	4562.2	631.51	0.37	7.35	1.38	0.01	0.82	0.09	150.00	9864.50	3031.79	0.25	21.42	6.97
Dy	1.01	566.04	96.04	0.24	1.23	0.15	0.003	0.021	0.003	17.53	452.67	99.47	0.10	0.60	0.18
Er	57.69	310.92	54.21	0.14	0.73	0.08	0.02	0.011	0.001	10.11	269.00	60.42	0.11	0.90	0.25
Eu	0.11	92.54	16.44	0.02	0.44	0.04	0.00	0.004	0.00	3.39	80.47	18.05	0.10	0.70	0.21
Fe	0.01	180.47	23.84	0.03	1.71	0.28	0.21	9.11	1.50	0.00	500.10	90.30	12.76	3488.90	1108.71
Gd	0.91	548.25	97.49	0.37	1.46	0.18	0.000	0.011	0.002	18.90	465.00	104.28	0.10	0.70	0.20
Ho	0.26	115.88	19.89	0.03	0.32	0.04	0.000	0.004	0.001	3.61	95.73	21.19	0.10	0.60	0.18
La	1.39	423.34	57.69	0.14	10.56	0.15	0.00	0.04	0.01	9.20	153.07	32.15	0.16	0.90	0.28
Li	0.81	1.64	0.25	-	-	-	0.01	0.41	0.06	-	-	-	0.16	2.21	0.74
Lu	0.01	34.21	6.44	0.07	0.29	0.02	0.00	0.001	0.00	1.27	33.23	7.94	0.10	0.20	0.05
Mg	57.61	3660.33	746.45	207.25	1051.06	161.15	26.23	677.61	138.62	37.33	1369.50	249.34	9.69	667.88	198.66
Mn	51.2	705.1	153.5	46.2	278.7	57.9	0.22	174.54	31.48	4.77	559.00	97.45	0.24	2.78	0.81
Na	1.50	51.41	13.23	-	-	-	2.5	35.5	9.2	1.80	31.53	7.86	4.75	122.49	36.18
Nd	1.83	1073.24	169.11	0.8	3.4	0.3	0.0	0.041	0.01	35.20	828.00	181.54	0.11	1.12	0.40
Ni	957.95	56260.47	1109.13	5.1	24.8	4.1	-	-	-	1286.33	36690.00	6638.88	1.42	18.66	5.15
Pb	0.01	33.21	6.15	0.77	0.6	0.13	0.001	0.011	0.001	0.24	18.60	2.66	0.13	3.67	1.07
Pr	0.42	220.55	32.21	0.12	0.61	0.07	0.00	0.009	0.002	6.20	143.03	30.28	0.10	0.60	0.15
Sc	0.43	47.46	6.79	-	-	-	0.002	0.9	0.1	1.03	77.30	14.94	0.12	1.00	0.26
Sm	0.52	316.41	55.99	0.21	1.03	0.12	0.001	0.012	0.002	11.73	278.33	62.87	0.10	0.90	0.23
Sr	0.03	2.44	0.48	1.01	5.25	0.55	0.2	0.87	0.67	133.00	990.00	233.26	0.19	41.45	12.56
Tb	0.03	91.24	15.62	0.02	0.36	0.03	0.0	0.021	0.002	2.86	71.73	15.76	0.10	0.50	0.16
Th	0.04	21.45	2.38	-	-	-	0.0	0.1	0.015	0.02	29.00	5.39	0.18	0.65	0.22
Ti	0.22	13.17	3.13	-	-	-	0.003	0.2	0.03	0.16	13.70	3.76	-	-	-
Tm	0.12	38.80	6.94	0.01	0.23	0.02	0.0	0.001	0.00	1.33	34.77	8.09	0.10	0.20	0.04
U	0.22	3411.1	532.03	0.04	0.23	0.03	0.001	0.040	0.006	34.45	6676.50	1312.37	0.14	0.56	0.16
Y	8.2	4223.2	673.8	1.55	7.69	0.93	0.01	0.13	0.02	113.00	3326.67	670.54	0.16	1.55	0.53
Yb	0.46	226.41	41.57	0.02	0.53	0.06	0.001	0.007	0.001	8.40	216.33	50.41	0.10	0.60	0.19
Zn	0.41	15.28	3.22	1.26	8.37	1.44	0.31	6.52	1.55	-	-	-	-	-	-
Cl ⁻	2.36	575.32	156.42	-	-	-	-	-	-	-	-	-	-	-	-
F ⁻	0.66	83.11	12.14	-	-	-	-	-	-	-	-	-	-	-	-
SO ₄ ²⁻	422.14	16758	1020.25	96.23	2044.82	400.28	30.1	1454.6	309.5	-	-	-	-	-	-
pH	3.27	5.46	0.47	-	-	-	-	-	-	-	-	-	-	-	-
PO ₄ ³⁻	0.03	0.84	0.12	-	-	-	-	-	-	-	-	-	-	-	-
HCO ₃ ⁻	0.76	53.21	7.63	-	-	-	-	-	-	-	-	-	-	-	-
DOC	0.89	20.64	2.57	-	-	-	-	-	-	-	-	-	-	-	-

230 m depth, 1.6 km length and 0.9 km width, which is filled now); b) an underground mining system, with maximum depth 900 m, and 3000 km of underground galleries; c) several waste rock piles in which acid mine drainage (AMD) occurred (Wismut GmbH, 1994a, b). Between 1971 and 1978, the waste rock piles were leached with AMD and later with sulfuric acid (10 g/l) in order to extract the uranium. Gessenhalde was the only leaching heap built up by Ordovician and Silurian shales with a low-grade ore mineralization (uranium content <300 g/ton) (Rüger and Dietel, 1998). The leach pad was sealed with 0.6 m of local loam and compacted in order to prevent infiltration. This seal was covered by a one-meter-thick layer of coarse waste rock containing low grade of uranium mineralization from Lichtenberg as a drainage layer during the leaching process. Of the total area of 28.7 ha, approximately 17 ha was directly on top of the local loam and 12 ha was on top of waste rock. Between 1971 and 1978, AMD (pH 2.7–2.8) was used to leach the two layers of mainly siliceous shale (Silurian radiolarite, 4 Mm³). In 1978, a third layer of mainly leather shale (Ordovician) was applied (2.8 Mm³) on top of the two older layers and, a change to dilute sulfuric acid (10 g/l) instead of AMD to enhance the leaching process was adopted. Thus, all three layers consisted of black shale. Drainage of the heap was designed to transport the leaching solution to collection ponds. It is probable that these drainage gullies were not completely sealed so that some contamination (uranium, REE, and other heavy metals) infiltrated the underlying soil to a great depth, and leachate seeped through the lining of the Gessenhalde and accumulated in the Quaternary sediments underneath (Wismut GmbH, 1994a, b). The leaching process was stopped in 1989, and in 1990s, the leaching heap (Gessenhalde) was placed into the open pit, Lichtenberg and a layer of allochthonic top soil was added as a last remediation step. A few years later, in 2003 and 2004, residual heavy metal contamination was measured in water (groundwater and surface water) and in the upper soil layer as reported by Carlsson and Büchel (2005). This shows that the remediation was not complete. Still, in many locations at the site, it is apparent that plants are affected by high metal concentration indicating that contamination is present in the root zone, the upper 30 cm of soil. In 2004, the test site “Gessenwiese” was created in the northern part of the base area of the former leaching heap, Gessenhalde, with the aim of improving remediation strategies for heavy-metal contaminated areas with special emphasis on REE contamination (Carlsson and Büchel, 2005).

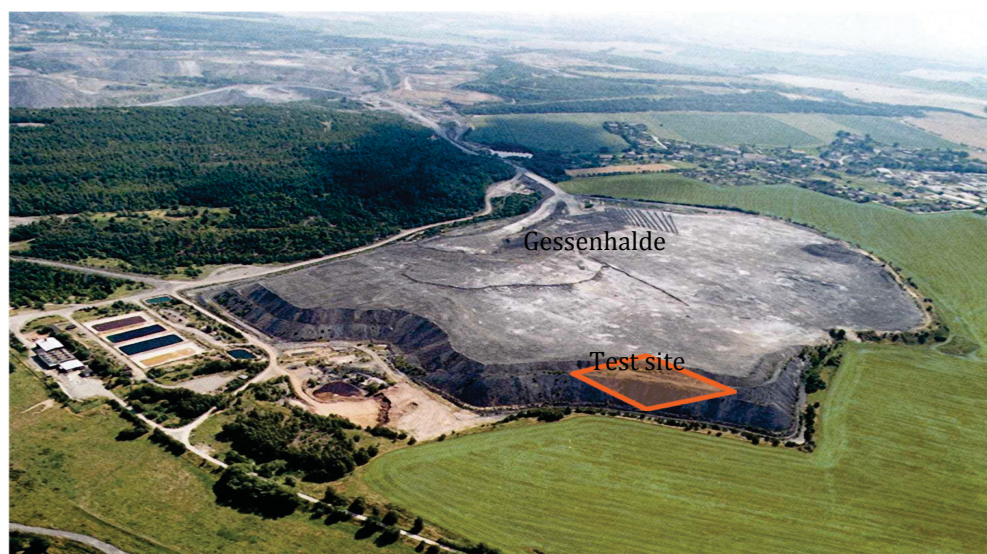


Fig. 1 Gessenhalde and the test site which is located at the base area of the Gessenhalde (former leaching heap) in order to improve the remediation strategy

To extend the collaboration of the project within other similar occasions, a short period (three months) of this project was done in two data sets from Sweden: Kvarntorp, stratified acidic pit lake, and Bersbo, historical sulphidic mine waste. The data of each area was studied separately and with different methods. The main aim of choosing these areas to study is the similar environmental and geological situation of these areas to the test site (Gessenwiese).

Kvarntorp: In Sweden, organic rich black shales are commonly known as alum shales because of their content of alum ($\text{KAl}(\text{SO}_4)_2$), which was used during the 18th and 19th centuries for the production of dyes and rag paper (Eklund et al., 1995). Hence, in many regions where black shales are available, the remains of past operations are visible as piles of burnt shale. Production of hydrocarbons from the organic rich black shale (up to 18% organic carbon) in the Kvarntorp area in south central Sweden started during the Second World War. It is located 20 km south of Örebro, in south central Sweden, 200 km from Stockholm. The former mining area covers 8 km². The black shale horizon that has a thickness of 5-15 m is of late Cambrium age. Usually it is found beneath an Ordovician limestone but at this site the black shale has been lifted to the surface due to faulting. Hence, it is easily accessible and well suited for open pit mining. The depth of the deepest pit lake is roughly 30 m. That is the reason this is thermally stratified during summers and winters but interrupted by turnovers in autumn and spring. This is an opportunity to study the impact of fundamental hydrochemical conditions on the metal distribution and compare the two lakes, one of which is continuously neutralised by the intruding alkaline leachates from the industrial waste deposit. The entire data set was evaluated with Hierarchical cluster analysis in R-mode and Q-mode to examine the relationships between variables and samples, respectively.

Bersbo: Historical sulphidic mine waste is the single largest source of environmental pollution of metals in Sweden. Bersbo is located some 250 km SSW of Stockholm, Sweden, in the municipality of Åtvidaberg. Mining for copper is documented from the 14th century but probably began earlier. Rational mining was introduced in 1765, and during the peak production in 1850-1870 Bersbo was the largest copper producer in Sweden. Copper was refined in the nearby refinery in Åtvidaberg. The ore that could be extracted with the technology of the time ceased in the late 1800s, and the mining was stopped in 1902. The eleven shafts were kept dry until the early 1940s. For a more complete description see Allard et al. (1987). The database contains some 50,000 quality observations at present. A statistical evaluation of surface quality changes after remediation in 1988 of a historical sulphidic mine site in Bersbo, Sweden, has been made in the present project, as a follow-up of an extensive monitoring program.

3 Geology and hydrology setting

According to Wismut GmbH (1994a, b) and Geletneky et al. (2002), the uranium deposit of Ronneburg, is a strata- controlled structure bound deposit. The Paleozoic host rock consists of argillaceous and siliceous black shale with intercalations of dolomitic and phosphorite nodules beds. The main back- shale horizon lies below Ordovician carbonaceous sandy shale and overlies Silurian carbonate rocks. The rocks contain up to 7wt% sulfides, 5-9 wt% organic carbon, and 40-60 ppm uranium. Hence, it was decided to relocate to the lowermost part of the Lichtenberg open pit and mixed with granulated anhydrous lime (Jakubick et al., 1997; Paul et al., 2003). In June 1995 the last leached ore was removed (Wismut GmbH, 1994a, b).

Below the new- contoured base area (at the southern part of the Test site), Quaternary glacial sediments with at least 10 m thickness were found, including four layers: a) a graded bedding of silty and gravelly sand beneath a sand layer; b) silt; c) clayey silt/ varved clay; and finally d) an allochthonic soil material that was added up to the other layer during remediation. The southern part of the test site is dominated by unit (a), the middle test site by unit (b), and to a certain extent by unit (c). The north of the test site is sandy also, but this sand has a higher proportion of silt. The allochthonic soil material (top soil) is a very heterogeneous, unsorted, and unlayered material that covers the area (Grawunder et al., 2009). The sand in the north and the silty material in the middle test site are connected within a facies change, while the sand in the south appears more as an overlying unit.

Groundwater is flowing through sandy glacial sediments at the northern and the southern parts, and through thin layers of a perched aquifer, especially within the silty at the central part of the test site. The groundwater level on the test site fluctuates from about 2.5 to 3.5 m below the surface (261- 262 m above the sea level) in the south and middle, and from 0.2 to 1.0 m (261 m above the sea level) in the northern test site (Lonschinski, 2009). The water infiltrating the aquifer follows the topographic gradient through the test site from the south to the north. The more specific flow direction is from the south–southwest to the north–northeast. It is not possible to specify one or two flow paths, because the groundwater is flowing through very thin layers of gravel and sand. Shallow aquifers (about 0.9 m depth) were identified in the middle of the test site.

4 Aim of study

As it was mentioned in Section 1 and 2, various studies have been performed in the study area in order to characterize and understand the contamination conditions; and hence, to improve the remediation strategy. In order to achieve to this goal, some questions must be answered, such as: What is the source of contamination? What is the contamination distribution in the various data sets (e.g., groundwater, soil water, and soil)? What is the relation (similarities and/or dissimilarities) between these data sets?

To investigate all above questions, the focus of the work was on REE. There are two reasons to focus on REE in this study: a) high concentration of REE in the area (Table 1); b) specific geochemical behavior of REE. Rare earth elements are a set of seventeen elements in the periodic table including lanthanides, as well as scandium and yttrium. The attraction of using REE to investigate the geochemical problems is that they form the coherent group of trace metals whose properties change systematically across the series La through Lu (Brookins, 1989). The pattern of REE indicates some geochemical processes in the area (Merten et al., 2005). They are often grouped into light rare earth elements (LREE; La to Pm), middle rare earth elements (MREE; Sm to Dy) and heavy rare earth elements (HREE; Ho to Lu). The current study has been done in three stages as follows:

a) Statistical evidence of groundwater contamination and significant factor with special influence of REE

In the literatures, the role of certain factors that control heavy metal contamination in different environments has been studied (Köher et al., 2005; Ling et al., 2008; Silva et al., 2009). Several detailed studies of the test site have been performed regarding the extent of heavy metal contamination (Carlsson and Büchel, 2005; Grawunder, 2010; Lonschinski, 2009; Lorenz, 2009;

Wismut GmbH, 1994 a). However, none of these studies considered the factors from statistical perspective nor did they examine the general contamination distribution or pattern of the site.

The aim of this stage was to investigate the significant factors that control contamination distribution at the test site. Various factors such as chemical analyses, pH, precipitation, geology, and groundwater fluctuation were taken into account to define the significant factors that influence the contamination distribution. This investigation is important because the results can be used to plan the future sampling (location and frequency) and also predict the contamination distribution in future. Using these results, the remediation strategy can be optimized and improve the project economically.

To achieve this goal, analysis of an extensive set of data was necessary. This data set includes various types of samples and information that were collected over a period of a several years. Multivariate analysis is a suitable tool for such statistical analysis because variables and features can be taken into account and the relation between them can be studied.

b) Top soil heavy metals and rare earth elements contamination

As mentioned above, investigation the source of contamination is an important task, since it affects the contamination pattern/distribution, and the results are helpful to plan remediation and predict the future environmental condition of the site. The hydrogeological conditions of the study area, such as the shallow groundwater level and high groundwater fluctuation, support the hypothesis that the contaminated- top soil could be a source of groundwater contamination. To investigate this hypothesis, two sets of data, soil samples and groundwater samples, were studied to discover the similarities between the contamination patterns. The study also helped to select the agent that behaved more similarly to the groundwater with respect to leaching. Investigating the similarities of contamination patterns in the groundwater samples and the soil leachate reveals the influence that the contaminated soil has on the groundwater.

c) Fuzzy hierarchical cross-clustering of soil and slate samples contaminated with heavy metals and rare earth elements

Several hypotheses for the sources that influence the contamination are still under investigation. One hypothesis is that the Silurian- Ordovician slates are the source of contamination. Previous studies have been performed using soil water samples, soil samples, and slates from the test site (Lonschinski, 2009; Pasalic, 2011; Wagner, 2010). However, these studies did not focus on the relation between different types of samples in order to verify the hypothesis. Presently at the site, there is no means by which to access the original, parent material and thus the only option is sample the material of the contoured area. Furthermore, both the low number of slate samples compared with the number of soil water samples and the existence of uncertainty result a gap of information in the data sets. These uncertainties arise for various reasons: for example, complexation or water- rock interaction that cannot be quantified. Under such conditions, a study relying on fuzzy logic is a more practical method because it is flexible and can consider more possible relationships between the parameters (Demicco and Klir, 2003). The results are helpful to better understand the source of contamination in the area and to test the hypothesis.

d) Dry Covers on historical sulphidic mine waste; long term statistic performance assessment of surface water quality

The changes of surface water composition during the period 1985-96 (before and after the remediation), were statistically evaluated with data from two adjacent locations (Lake Gruvsjön and the

Kuntebo creek), and a downstream lake (Lake Risten). Cluster and time series analyses were used to study the surface water composition. The behavior of the elements is different before and after the remediation period, based on the statistical analysis. These differences can be seen in the Q-mode cluster analysis as well as in the time series analysis. The Q-mode clusters were generally good representative of the samples that were taken before or after the remediation period.

e) Distribution of rare earth elements and other metals in a stratified acidic pit lake in black shale 45 years after mine closure

The pattern and distribution of metal contamination in acidic pit lake were studied in black shales. Water samples were collected (profiles and transects) in some of the pit lakes and analyzed with respect to general hydrochemical parameters and metals. To summarise, the water chemistry in Lake Norrtorpsjön is heavily influenced by the input of elements from three major sources (black shale, alkaline solid waste, municipal waste) in relation to Lake Pölen that serves as a reference for the conditions governed by the surrounding shale/limestone. The relationship between the chemical analyses depends on the chemical environment and the properties of the individual element.

5 Material and methods

The samples of this study were collected from porous media. Following, the movement of water in porous media where the samples of this study were collected from will be explained.

5.1 Water movement within porous media

Porous media includes natural soils; unconsolidated sediments, and sedimentary rocks. The main interest in a porous media is its ability to hold and transmit water.

5.1.1 Groundwater

At a regional scale and undeveloped conditions, water typically enters the subsurface at high regional topographic level and leaves it at low topographic level. It is important to note that the high and low topographic levels are equivalent to high and low potential energy. The flow path between the recharge and discharge zone is controlled by the host medium permeability distribution. If water enters and leaves the subsurface at adjacent topographic level, the flow system is known as local system. In such condition, the travel time varies from months to tens of years. If there are intervening of topographic level between recharge and discharge zones, the system is called intermediate system. In this case, the travel time varies from tens of years to hundreds of years. A flow system that travels from a high regional topography to a low regional topography is called a regional flow system. The travel time for such system varies from order of thousands to hundred thousand years.

Hydraulic conductivity and hydraulic head are important factors that control the groundwater movement. Darcy's Law is the relationship that explains fluid flow in porous media. Hydraulic conductivity (K) is the ability of geological material to transport groundwater and is measured in volume per unit time per cross-sectional area. Hydraulic head (H) is a specific measurement of water pressure or total energy per unit weight. It is usually measured as a water surface elevation, expressed in units of length, but represents the energy at the entrance (or bottom) of a piezometer. Groundwater movement is always in the downward direction of the hydraulic head gradient. If there is no hydraulic head gradient, there is no flow.

Based on Darcy's law, the total discharge, Q , is equal to the product of the permeability of the medium, k , the cross-sectional area to flow, A , the pressure gradient, ∇P , and λ mobility. The parameter k , has the dimension of length squared. The negative sign is needed because a fluid flows from high pressure to low pressure. Equation 1 is the Darcy law.

$$q = - \lambda k \nabla P \quad \text{Eq. (1)}$$

where, q is the flux, ∇P is the pressure gradient vector, λ is mobility which is a function of density and viscosity, and k is permeability of the medium.

Hydraulic conductivity is a property of porous media that describes the ease with which water can move through the pores or fractures. Hydraulic conductivity depends on the permeability of the material and on the fluid properties. The permeability is a function of soil properties such as pore and particle size (grain size) distributions, soil texture, and the saturation degree. This factor describes the variations through space within a geological formation. Hydraulic conductivity has units of velocity. One important factor that influences this parameter is cation composition. A good example is clay soils that can exhibit either a flocculated or dispersed structure. Fewer high valence cations, like Ca^{2+} , comparing with Na^+ , is needed in a layer to balance the negative charge of the clays. Hence, the resulting ion layer will be thinner, and tend to flocculate and create a larger pore spaces and hence greater permeability. On the other hand, low valence cations, e.g., Na^+ , result in an expanded cation layers and dispersed clay. Such soils, show poor structure and with low permeability and poor discharge. Hydraulic head is the mechanical energy per unit weight and quantifies in units of length. Equation 2 is the hydraulic head (Freeze and Cherry, 1979; Todd and Maysy, 2005).

$$h = v^2/2g + P/\gamma + z \quad \text{Eq. (2)}$$

where $v^2/2g$ is velocity head which is the kinetic energy per unit weight, P/γ is pressure head which is the measure of the ability of the fluid to do work, z is the elevation head which is the potential energy per unit weight.

5.1.2 Soil water

The water-unsaturated zone, or the vadose zone, extends from the ground surface down to the water table. The shallow part of this zone is called soil water zone. The porosity and permeability of soil is a function of soil texture and structure. Soil structure is a function of physical shape and size of the grains. It is also influenced by soil chemistry, since soil minerals have electrical charge on their surface and this electrical charge influences the stability of soil structure. The process of water entering from the ground surface into the vadose zone is called infiltration. Not all of the water that infiltrates the unsaturated zone reaches the water table. A significant amount is returned to the atmosphere through evaporation and transpiration from plants. In the vadose zone, only a part of the void space is filled with water. Coarse-textured, sandy soil may hold about 10- 20 percent of the water saturation after a long period of drainage. However, fine- textured, silt and clay, may hold about 90 percent saturation. The remaining part is filled with soil air. Hence, the hydraulic conductivity is low comparing with the saturated zone, because the pores that are filled by air do not transmit the water. The permeability and hence hydraulic conductivity of the soil is a non-linear function of the water content. Also, since the large-water filled pores empty out first, the unsaturated hydraulic conductivity

decreases rapidly as the volumetric water content (the volume of the water as a ratio to the total volume of the soil pores) decreases. This situation is in contrast with the saturated zone that all of the pore space is filled with water. In the vadose zone, water flow is dependent upon water content, negative suction or matric potential (soil pressure or capillary potential) and gravity potential. It is also important to note that Darcy's law is not valid for a non-Newtonian fluid. Regardless Darcy or non-Darcy flow, hydrogen and covalent bonding between water and clay particle and ion adsorption affect the water flow in soil. It is important to note that soil structure and heterogeneities (e.g. preferential flow paths) are factors that leads to non-Darcy flow in vadose zone. In the root zone, the pores that are formed by roots and animal burrows are large. These macro-pores form a preferential flow path for water movement horizontally and vertically (Beven and Germann, 1982). This situation makes the infiltration faster than what might be expected regarding the hydraulic conductivity of soil. Another type of potential flow path is called fingering. It occurs when an infiltrating fluid splits into downward-reaching finger due to the instability caused by viscosity variation. The other type of potential flow path is funneling (Kung, 1990) that occurs below the root zone regarding to the stratified soil or sediment profiles. Sloping, coarse sand layers embedded in fine-sand layers can slow down the downward infiltration of water. The sloping layer will direct the flow to the end of the layer.

At small volumetric water content, water movement in soil is slow, even under large gradient, because the water must move across the thin layer of adsorbed water (a thin film of water on the surface of the soil grains) as it passes from one pendular ring (accumulating water at the contact points between grains that represent the smallest pore-space openings in the soil) to another one. As the water content increases, the pendular ring grows and the thickness of pellicular water (a thin skin of adsorbed water covering the grains) film increases. With a film thickness of about 500 to 1000 Å, the water is free to move under imposed energy gradients. As the water content continues to increase, the air becomes isolated in individual pockets in the larger pores, and flow of the air phase will not be possible longer. Hence, under normal condition of unsaturated zone, full water saturation is not achieved due to the entrapped air (Tindall and Kunkel, 1999).

5.2 Data sets

5.2.1 Groundwater samples

Contamination was measured in groundwater in the study area (Grawunder et al., 2009). Furthermore, the groundwater level is shallow in this area. In the duration of high precipitation, at the middle part of the test site, groundwater fluctuates highly and return flow occurs. This fact makes groundwater a suitable tool to understand the relation between various media. The present study deals with 174 groundwater samples that were collected from 33 (GTF2-34) sampling points (Fig.2). Samples were collected seasonally from 2004 to 2008, a total of 30 samples were collected during the month of April in the years 2005 and 2007; 14 samples in total were collected during May of 2005 and 2006; 84 samples were collected in September of 2006 to 2008, and; 46 samples were collected in December of 2004 and 2006. The Eh, pH, electrical conductivity (EC), and temperature were measured on-site using portable instruments (WTW, pH320; WTW, LF320; WTW, external thermocouple). In the field, all samples except those for HCO_3^- analysis were filtered using 0.45 µm cellulose acetate filters (Sartorius), with glass fiber pre-filters (Sartorius). Samples for element analysis were acidified with HNO_3 (65%, subboiled) to pH < 2. All samples were kept cool (6°C) until analysis.

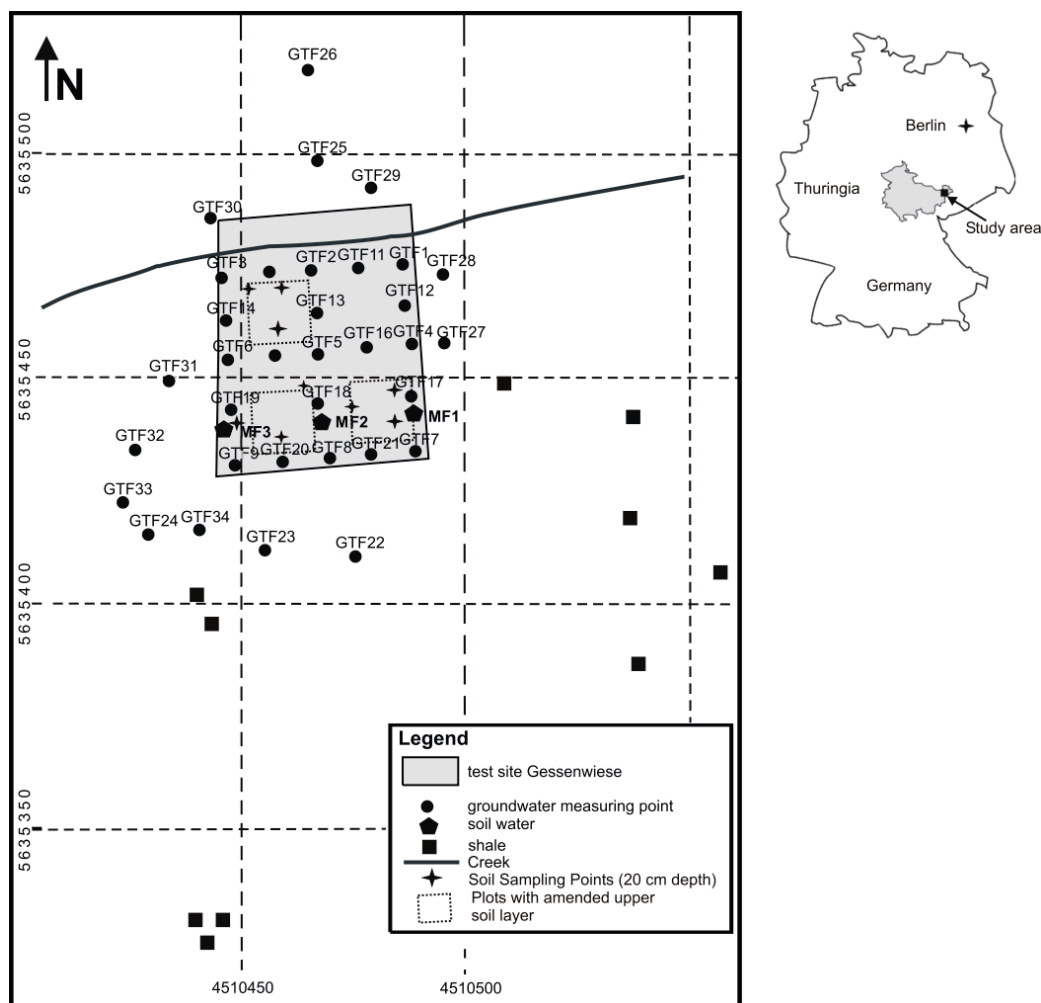


Fig. 2 Location of the groundwater, soil water, soil and slate samples from the test site and the surrounding area

The analysis of Al, Ca, Fe, Mg, Mn, Na, Sr, and Zn were performed with ICP-OES (Spectroflame, Spretro); Ba, Cd, Co, Cu, Li, Ni, Pb, REE (La-Lu), Sc, Th, Ti, U, Y, and the analysis of Zn with ICP-MS (until 2007: PQ3-S, Thermo Elemental, subsequently, X-Series II, Thermo Fisher Scientific). The anions Cl^- , F^- , Cl^- and SO_4^{2-} were analyzed using ion chromatography (DX120, Dionex), while PO_4^{3-} was measured with photometry (Hach- DR/4000U). For the determination of HCO_3^- , an automatic titration system was used (Titrimo 716 DM, Metrohm). Dissolved organic carbon (DOC) was measured on 0.45 μm -filtered samples (DIMA-TOC 100, Dimatec until 2006; and then multi N/C 2100, Analytik Jena). In order to check the accuracy of the analytical results, the instrument drift was monitored and corrected using Be, Ru and Re as internal standards. Each sample was measured three times. First, an outlier test is performed on each of the three runs (Grubbs test, 90% significance, criterion 1.15). Then, for the remaining runs, the mean and standard deviation are calculated. Furthermore, analytical quality is checked by the use of standard reference materials: the standard reference materials SCREE and PPREE are used for the analysis of the REEs (Verplanck et al., 2001). For other heavy metals, the standard reference materials SPS-SW2 (LGC standards) and NIST1643e (LGC standards) were used. Only values between 90-110 % for these reference materials are accepted. Detection limits are calculated according to the "3 sigma criterion".

$$DL = Dil. Fact. * \sqrt{2s_0/S} \quad \text{Eq. (3)}$$

where, $S = I/c$, Dil. Fact. is the dilution factor, s_0 = standard deviation, S = sensitivity, I = intensity and c = concentration.

5.2.2 Soil water Samples

The present study concerns 53 soil water samples that were collected from 3 sampling points (MF1, MF2, MF3) (Fig. 2), from three depths: 30 cm, 60 cm and 100 cm below the surface level (237, 207, 257 m below sea level, respectively). The sampling device employs the principle of tension-controlled under pressure; each sampling event took 14 to 15 days. During sampling, the water was collected in glass bottles from a depth of about 30 cm below the surface. Samples were collected between 2005 and 2007. The samples were prepared and analyzed in a similar manner as the groundwater samples described in Section 5.2.1. Table 1 presents the results of the analysis.

5.2.3 Slate Samples

Section 3 discusses the possibility that the slates could be a source of contamination. This possibility was studied more in detail using a total of 10 slate samples, collected from the former Gessenhalde and surrounding area. The samples were kept in 40°C for five days. The weathered and non-weathered portions were separated, and the non-weathered portion was crushed using a hammer. A plastic sieve was used to fraction the samples into two different sizes 0.63-2 mm (middle size) and 2 mm-2 cm (coarse size).

Furthermore, the slate samples were leached with water (Ultrapure water was obtained with a Purelab Plus system from USF Elga Seral (Ransbach-Baumbach, Germany)). The amount of three grams of the milled material were mixed with 30 ml of leachate (water) and shaken in overhead shaker for 24 hours (ELU safety lock, Edmund Bühler). The samples were centrifuged for 15 min, and acidified with nitric acid. The elements Al, Ca, Fe, Mg, Mn, Na, Sr, and Zn were analyzed with ICP-OES (Spectroflame, Spectro); Ba, Cd, Co, Cu, Li, Ni, Pb, REE (La-Lu), Sc, Th, Ti, U, Y, and Zn were analyzed using ICP-MS (until 2007 PQ3-S, Thermo Elemental, 2007: X-Series II, Thermo Fisher Scientific). For accuracy, they were treated like the groundwater and the soil water samples as described in Section 5.2.1. Table 1 presents the result of the analysis.

5.2.4 Soil samples

The top soil is contaminated in the area and is in contact with groundwater as a result of shallow groundwater level and return fellow during the period of high precipitation. In order to investigate the relation between the soil and the other media, 80 soil samples (each weighing about 20-100 g) were collected from the test site, nine samples were collected from the amended soils and one sample was taken from the surroundings (Fig.2). All samples were obtained from a depth of about 20 cm. Furthermore, different zones within the test site were amended down a depth that varied from 20 to 90 centimeters, by adding of allochthonic non-contaminated soil material. These amended soils are categorized based on their different organic carbon content. The leaching solutions consisted of water (Ultrapure water was obtained with a Purelab Plus system from USF Elga Seral (Ransbach-Baumbach, Germany)) and ammonium nitrate (1M, p.a., Roth).

Water was used as a leaching agent to remove the water soluble fraction, which represents the fraction that can be washed out by rain. Ammonium nitrate is used to leach the plant-available fraction of soluble metals, and it is useful in order to estimate the effect on flora caused by heavy metals. The soil samples were dried in porcelain plates either at 40°C in the drying oven or at room temperature. The specimens were sieved to below 2 mm, and kept in a dry place inside a plastic container. The amount of 3-4 g of the material was placed into 50 mL polyethylene test tubes and the leaching solution was added to reach a solid-liquid ratio of 1:10. The suspension was tumbled for 24 h at about 20 rpm (Overhead shaker- ELU safety lock, Edmund). For each experiment, blanks containing only leaching solution were prepared. The samples were then centrifuged 15 min at 2500 rpm, afterwards, 15 mL of each sample were filtered through a 0.45 µm-cellulose acetate filter. Then, aliquots of the filtered samples were acidified with HNO₃ (65%, suprapur, Merck) to a pH below 2 and kept at 6°C until analysis. Using the remaining supernatant, the pH and the electrical conductivity were measured using respectively pH meter pH197 (WTW), and LF320 (WTW). All samples were analyzed in triplicate, and the mean of the obtained values were used for statistical analysis.

5.2.5 Groundwater level fluctuation and precipitation

The groundwater levels were monitored every 15 minutes from 2004-2009 by data loggers in 11 monitoring wells GTF2, 3, 4, 6, 7, 8, 12, 13, 15, 18 and 25 (AquiLite ATM10 by Aquitronic (GTF2, 18); beaver, Aquitronic (for other wells)). The monitoring wells were constructed simply; the borehole was drilled, the PVC casing (2-inch diameter) was inserted, and the surface was sealed to avoid the rain/surface water entering the wells. The depths of the wells varied to a maximum of 7 m. Geological profiles are available for the boreholes GTF10 to 34.

The daily precipitation data used in this study were provided by the Wismut GmbH from a nearby measuring station (Lonschinski, 2009).

5.3 Methods

Prior to using multivariate methods, censored data were replaced by estimated values and the univariate outliers were defined to make a complete set of data. Multiple censored data were reported in analytical results for Cr, Cu, F⁻, Fe, Pb, PO₄³⁻, Th, Ti, Sc and U. The reason for reporting multiple censored data is that the instrument detection limit is not constant, but rather it is adjusted according to the sample's dilution factor. The estimating method was applied based on the numbers of the censored data for each parameter. The replacing values for substitution methods are 3/4 of each detection limit. One-half, seven-tenths of the detection limit are other among common substitution values used. However, any single value between zero and the detection limit is as good as another (Lee and Helsén, 2005). Parameters with more than 10% censored data (Sanford et al., 1993) were replaced by the values that were calculated by ROS method in R software (Lee and Helsén, 2005). ROS is a method that is applicable for water data sets with multiple censored data. This technique is based on a "robust" semi-parametric method developed by Helsén and Cohen (1988). It computes a linear regression for data or logarithms of data versus their normal scores in a normal probability plot. The regression parameters use detected observations. Due to the definition of normal scores, fitting this line is fitting a normal distribution. The regression equations are fitted to the detected observations on the probability plot, and values of individual censored observations are predicted from

the regression models based on their normal scores. This method is available in R software and can be used for any data set with a maximum of 80 percent values censored (Lee and Helsen, 2005).

Also Empirical Cumulative Distribution Function (ECDF) and Cumulative Probability (CP) (Filzmoser, 2005) have been plotted for the analytical parameters (variables) and univariate outlier values were defined for each. ECDF plot shows the variables along the x- axis and the probability of the empirical cumulative distribution function along y-axis. One of the main advantages of this function is that all single data points are visible, and any unusual high or low value can be identified. For calculating the outliers boundary, classic (mean \pm 2 standard deviation) and robust (median \pm Mahalanobis distance) formula can be applied to detect the univariate outliers. Furthermore, combination of CP plot with the ECDF can define the best result (Reimann et al., 2009). By plotting CP, the direct visual estimation of the median or any other value is possible. These functions are available in the package StatDa by Filzmoser (2010), in R software environment. The defined values were replaced by nearest smaller neighbor value in each variable to reduce the influence of outliers. Subsequent to the univariate methods described above, multivariate methods (multivariate outlier detection, cluster analyses and cross- correlation) were used.

5.3.1 Multivariate outlier detection

The main aim of using multivariate outlier detection is defining the observations (samples) which have different structures (Filzmoser, 2005). In this case the outliers are not only the data that are very high or low in relation with the other data, but their shape and domain are different. Generally these observations are resulting from secondary geological processes; and do not have higher or lower concentration in one variable. Hence, multivariate outliers are not necessarily univariate outliers and vice versa (Reimann et al., 2009). The function “symbol.plot” from “mvoutlier” package by Moritz and Filzmoser (2009), in R environment was used in this study to define multivariate outliers. This function is based on robust Mahalanobis distance (MD) (Eq. 3), minimum covariance determinant (MCD) estimator and adjusted quantile (for more detail about these techniques and formula, the reader is referred to Filzmoser (2005)).

$$MD(x) = \sqrt{(x - \bar{X})^T S^{-1} (x - \bar{X})} \quad \text{Eq. (4)}$$

where, MD(x) is Mahalanobis distance, \bar{x} is the mean, x is multivariate vector, **S** is covariance matrix, and T is transmitted matrix.

The outcome is a numerical matrix that reports the MD of the observations and a two-dimensional diagram which shows the ellipsoids corresponding to the 25%, 50%, 75% and adjusted quantile of the chi- squared distribution. Observations with different MD levels are shown by different symbols. The symbols that are plotted outside of the adjusted ellipsoid are potentially multivariate outlier. To run the function, different variables must be selected as the entry data set. Since the aim of this study is focused on REE distribution and its relation with other analytical parameters, three variables that were considered were: a) REE (La to Lu); b) other metals (Al, Cu, Fe, Zn); and c) pH. This selection is based on the result of the previous studies and also the concentration of the metals (Liang et al., 2008; Rönnback et al., 2008).

5.3.2 Cluster analysis

As mentioned above, cluster analyses (R- and Q-mode) were used. Squared Euclidean Distance, SED, (Eq. 5) was used to measure the distance between the parameters. After selecting the measurement, linkage method must be selected, Ward linkage was used for this data (Einax et al., 1997; Derde and Massart, 1982). Z-standardization (Eq. 6) was performed on the data to avoid the influence of data scale (Davis, 2002).

$$SED = \sqrt{(x_1 - y_1)^2 + (x_2 - y_2)^2} \quad (5)$$

Where, x, y are different observations in the data matrix.

$$Z = \frac{(x - \bar{x})}{s} \quad (6)$$

where, x is a raw score to be standardized, \bar{x} is the mean of the population, s is the standard deviation of the population.

The principal aim of this technique is to partition multivariate observations into a number of meaningful, multivariate homogeneous groups. The analytical parameters placed in one main or sub cluster can be a clue to investigate the significant factors. Hence, to find the similarities between REE and other analytical parameters, R-mode cluster analysis was used. The Ward distance of 25 is the clustering criterion. In order to investigate more about the geological and seasonal factors, Q-mode cluster analysis was used as well as multivariate outlier detection method. By applying this method, the similarity between the samples was defined based on the location, and also compared in different sampling seasons. As the analytical results are based on the concentration, each group represents the samples that have similar concentration of the analytical result.

5.3.3 Factor analysis

Factor analysis is used to find and interpret the hidden complex and possibly relationships between the observations (samples) and features (variables) in the data sets. Correlating samples are converted to the factors that are themselves non-correlated. The result of factor analysis is reported by a matrix of factor loading. Features with low loading have only slight influences on the factors; features with high negative or positive loadings determine the factors. Using factor analysis the common factor structures that explain the main part of the variance of data must be found (Einax et al., 1997).

In order to evaluate the results of factor analysis, KMO index is used. The KMO (by Kaiser-Meyer-Olkin) is an index for comparing the magnitude of the partial correlation coefficients. Small values of KMO measure indicate that a technique such as factor analysis may not be a good idea. Kaiser (1974) has indicated that KMOs below 0.5 are not accepted.

5.3.4 Cross-correlation

The hydrograph of the boreholes based on the groundwater level fluctuations from 2004 to 2008 were plotted using Matlab (Version 7.6) to study the general hydrographs. Generally, groundwater levels fluctuate according to the characteristics of precipitation events such as amount, duration and intensity. Some other factors that can influence the groundwater level are topography and sedimentology and stratigraphy. Groundwater level fluctuation in response to precipitation events can

be used to estimate the geological factors that control heavy metal concentration in groundwater, such as dilution. To reach this aim, the cross-correlation of groundwater levels monitoring data and daily precipitation was calculated. Cross-correlation is most appropriately used to compare two series that may have a temporal dependency between them (Eq. 7).

$$r_m = \text{Cov}_{1,2} / s_1 s_2 \quad (7)$$

Where, $\text{cov}_{1,2}$ is the covariance between the overlapped portions of sequence 1 and 2, and, s_1 and s_2 are the corresponding standard deviations.

This analysis will provide a correlation between two series or two waveforms. The aim of correlation analysis is to compare one or more functions and to calculate their relation with respect to a change of lag in time or distance (Einax et al., 1997). It means the observation of one series is correlated with another series at various lags and leads. Cross-correlation helps to identify variables which are leading indicators of other variables or how much one variable is predicted to change in relation to the other variables (Davis, 2002). Significant correlation at 95% confidence level is a criterion to test the hypothesis. As the two data series must be in a same size, the daily average of groundwater level fluctuation was used. By using this technique, time lag and correlation coefficient are recognized. Time lag is an important factor for such topics (Hölting and Coldewez, 2005). The time lag which occurs because rainfall needs time to reach to the groundwater can be determined by cross-correlation technique. The cross-correlation function in Matlab R 2008a was used for this purpose.

5.3.5 Partial least squares (PLS)

Sometimes is called "Projection to Latent Structures" because of its general strategy. The X variables (the predictors) are reduced to principal components, as are the Y variables (the dependents). The components of X are used to predict the scores on the Y components, and the predicted Y component scores are used to predict the actual values of the Y variables. In constructing the principal components of X, the PLS algorithm iteratively maximizes the strength of the relation of successive pairs of X and Y component scores by maximizing the covariance of each X-score with the Y variables. PLS is a predictive technique which can handle many independent variables, even when there are more predictors than cases and even when predictors display multicollinearity. Overall, PLS is favored as a predictive technique and not as an interpretive technique, except for exploratory analysis as a prelude to an interpretive technique such as multiple linear regression or structural equation modeling (Wold, 1981 and 1985).

5.3.6 Fuzzy clustering

Most fuzzy clustering algorithms are objective-function based (Bezdek, 1984 and 1987). In objective function-based clustering, each cluster is represented by a cluster centroid. The cluster centroid is computed by the clustering algorithm and may or may not appear in the dataset. The partitioning of the data points into different clusters is depending on the membership degree. It is computed based on the distance of the data points to the cluster centroids. The closer a data point lies to the centroid of a cluster, the higher is its degree of membership to this cluster. Hence, the aim

when dividing a data set into “c” number of clusters is to minimize the distances between the data points to the cluster centroids while maximizing the degrees of membership.

The focus of this study is fuzzy divisive hierarchical clustering approach, which uses only cluster centroids and a Euclidean distance function in comparison to the cross-clustering algorithm which produces not only a fuzzy partition of the soil water and slate samples, but also a fuzzy partition of the considered metals.

5.3.6.1 Fuzzy divisive hierarchical clustering

In general, fuzzy clustering algorithm explained by Bezdek (1980 & 1981) and later on by Bezdek et al (1987) and Sabin (1987). It can be formulated as follows:

Let $X = \{x^1, \dots, x^n\} \subset \mathbf{R}^p$ be a finite set of feature vectors, where n is the number of objects (measurements), p is the number of the original variables, $x^j = [x_1^j, x_2^j, \dots, x_p^j]^T$ and $L = (L^1, L^2, \dots, L^c)$ is prototypes (supports) of which characterizes one of the c clusters composing the cluster substructure of the data set. A partition of X into c fuzzy clusters will be performed by minimizing the objective function.

$$J(P, L) = \sum_{i=1}^c \sum_{j=1}^n (A_i(x^j))^2 d^2(x^j, L^i) \quad (8)$$

where $P = \{A_1, \dots, A_c\}$ is the fuzzy partition, $A_i(x^j) \in [0, 1]$ represents the membership degree of feature point x^j to cluster.

$A_i, d(x^j, L^i)$ is the distance from a feature point x^j to the prototype of cluster A_i , defined by the Euclidean distance norm:

$$d(x^j, L^i) = \|x^j - L^i\| = \left[\sum_{k=1}^p (x_k^j - L_k^i)^2 \right]^{1/2} \quad (9)$$

The optimal fuzzy set will be determined by using an iterative method where J is successively minimized with respect to A and L . Supposing that, L is given, the minimum of the function $J(\cdot, L)$ is obtained for:

$$A_i(x^j) = \frac{C(x^j)}{\sum_{k=1}^c \frac{d^2(x^j, L^i)}{d^2(x^j, L^k)}}, i = 1, \dots, c \quad (10)$$

Where C is a fuzzy set from X and

$$C(x^j) = \sum_{i=1}^c A_i(x^j). \quad (11)$$

It is easy to observe that $C(x^j) \leq 1, j = 1, 2, \dots, n$. For a given P , the minimum of the function $J(P, \cdot)$ is obtained for:

$$L^i = \frac{\sum_{j=1}^n [A_i(x^j)]^2 x^j}{\sum_{j=1}^n [A_i(x^j)]^2}, i = 1, \dots, c \quad (12)$$

The above formula allows one to compute each of the p components of L^i (the center of the cluster i). Elements with a high degree of membership in cluster i (i.e., close to cluster i 's center) will

contribute significantly to this weighted average, while elements with a low degree of membership (far from the center) will contribute almost nothing (Dumitrescu et al., 1994; Pop et al., 1995, 1996; Sarbu et al., 1993; Sarbu et al., 2007).

5.3.6.2 Fuzzy hierarchical cross-clustering algorithm

Building the classification binary tree is as follows: the nodes of the tree are labeled with a pair (C, D) , where C is a fuzzy set from a fuzzy partition of objects and D is a fuzzy set from a fuzzy partition of characteristics. The root node corresponds to the pair (X, Y) . In the first step the two sub-nodes (A_1, B_1) and respectively (A_2, B_2) will be computed by using the cross-classification algorithm. It is important to note that these two nodes will be effectively built only if the fuzzy partitions $\{A_1, A_2\}$ and $\{B_1, B_2\}$ describe real clusters. For each of the terminal nodes of the tree it is tried to determine partitions having the form $\{A_1, A_2\}$ and $\{B_1, B_2\}$. In this way the binary classification tree is extended with two new nodes, (A_1, B_1) and (A_2, B_2) . The ending of this processes is when no more structure of real cluster (either for the set of objects or for the set of characteristics) can be determined. The final fuzzy partitions will contain the fuzzy sets corresponding to the terminal nodes of the binary classification tree. This algorithm, which is called Fuzzy Hierarchical Cross-Clustering (FHCC), is a useful algorithm when it is desired to identify the relationships between different classes of samples and different classes of variables.

5.3.7 Time series analyses

A time series is a set of observations obtained by measuring a single variable regularly over a period of time. The form of the data for a typical time series is a single sequence or list of observations representing measurements taken at regular intervals. These sets of data have a natural temporal ordering. This makes time series analysis distinct from other common data analysis problems, in which there is no natural ordering of the observations.

Time series analysis comprises methods for analyzing time series data in order to extract meaningful statistics and other characteristics of the data. It is also distinct from spatial data analysis, where the observations typically relate to geographical locations. One of the most important reasons for doing time series analysis is to try to forecast future values of the series. A model of the series that explained the past values may also predict whether and how much the next few values will increase or decrease. The ability to make such predictions successfully is important. Methods for time series analyses may be divided into two classes: a) frequency-domain methods that include spectral analysis and recently wavelet analysis, and b) time-domain methods that include auto-correlation and cross-correlation analysis. Time series models will generally reflect the fact that observations close together in time will be more closely related than observations further apart. Models for time series data can be classified into three main methods such as the autoregressive (AR) models, the integrated (I) models, and the moving average (MA) models. These three classes depend linearly on previous data points. Combinations of these approaches produce autoregressive moving average (ARMA) and autoregressive integrated moving average (ARIMA) models. The autoregressive fractionally integrated moving average (ARFIMA) model generalizes the former three. Data are generally divided into two series: an estimation period, and a validation period. A model can be developed on the basis of the observations in the estimation period and the model can be tested and evaluated in the

validation period. The results are shown by plots and tables that may include stationary R-square, R-square (R²), root mean square error (RMSE), mean absolute error (MAE), mean absolute percentage error (MAPE), maximum absolute error (MaxAE), maximum absolute percentage error (MaxAPE) and normalized Bayesian information criterion (BIC). The plots would represent residual autocorrelations and partial autocorrelations. Results for individual models may be expressed in terms of forecast values, fit values, observed values, upper and lower confidence limits, residual autocorrelations and partial autocorrelations.

6 Results and discussion

6.1 Elements distribution/ pattern in groundwater

At the test site, element concentrations are lower in the southern part (GTF7, 8, 20, 21, 22) and increase to the north which is about the groundwater flow direction. For example, from south to north, the mean concentration of Mn is increasing from GTF22 (26 ± 13 mg/l) to GTF8 (52 ± 10 mg/l), GTF18 (114 ± 21 mg/l), GTF5 (253 ± 142 mg/l) to GTF13 (509 ± 98 mg/l). GTF25 (156 ± 64 mg/l) and GTF26 (43 ± 43 mg/l) are lower concentrated than the central part of the test site. This trend is quite similar for Co, Cl⁻, Mg, SO₄²⁻, and Zn. The distribution of Al, Cu, Fe, REE, Th, and U in the study area is more heterogeneous with only a slight gradient from south to north. Generally the highest concentrations of the heavy metals were measured in sampling points that are mostly located in the central part of the test site; GTF25 rather in the north. The general distribution in seasonal data sets is similar (Fig. 3).

The groundwater sample at the test site are in acidic range (pH: 3.2- 5.4) and oxic (Eh: 340- 715 mV). Redox potential decreases when travelling from the vadose zone to the saturated zone depending on different factors such as soil structure, porosity, permeability, composition of soil, distribution of organic matter, depth, frequency of infiltration, depth of water table and temperature. In a recharge area with silty or clayey soils, shallow groundwater does not contain detectable dissolved oxygen. In areas with little or no soil overlying permeable fractured rock, dissolved oxygen at detectable levels commonly persists far in to the flow system (Freeze and Cherry, 1979). The occurrence of dissolved oxygen in shallow groundwater in sandy structure is a result of low content of organic matter and relatively rapid infiltration.

One possible process of a solute transport in groundwater is called advective transport or convection. In this process, dissolved solids are transported along with the flowing groundwater. The amount of the transported solute is a function of its concentration in the groundwater and the quantity of groundwater flow. Due to the heterogeneity of geology, advective transport in different strata can result in solute fronts spreading at different rates in each stratum. For example, if a water sample is taken from a borehole that contains groundwater from different strata, the sample is a representation of all strata. Generally the groundwater movement can be faster or slower than the average linear velocity because of two different reasons: a) through the pores, the fluid moves faster in the center of the pores than along the edges; b) in large pores the fluid moves faster than small pores. Since, contaminated water moves in different velocity, mixing occurs along the flow path.

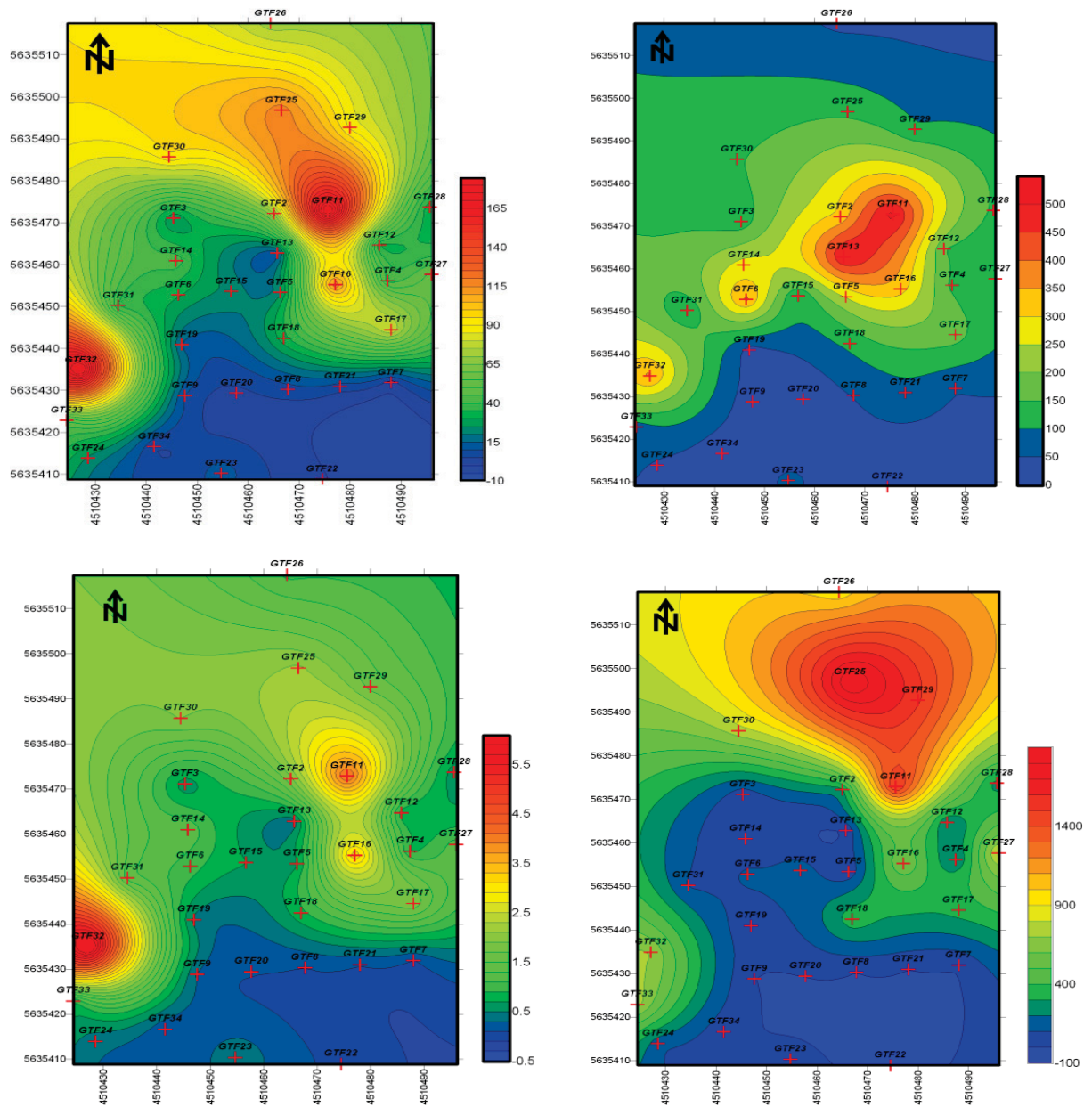


Fig. 3 Distribution of Al, Mn (upper left and right respectively), REE and U (lower left and right respectively) in the groundwater samples of the test site. The concentrations are lower in the southern part and increase to the north part of the test site. The distributions of the elements are stable and have inverse relation to pH range

Regarding the results of clustering, the analytical parameters are in two main clusters: REE and Al, Cu, F^- , Fe, Li, PO_4^{3-} , Sc, Th, Ti, U and Y in one main cluster, and Ca, Cd, Co, Cl^- , DOC, HCO_3^- , Mg, Mn, Ni, Pb, pH, Zn, and SO_4^{2-} in another main cluster. The pH of water determines the solubility. Metals are more soluble in water at a low pH range. The metals precipitate as metal hydroxides, since high pH corresponds to high hydroxide concentrations. However, every dissolved metal has a distinct pH at which the optimum precipitation will occur, as well as the presence of other metals' species. For example, in a previous study of the test site by Grawunder (2010) it was shown that in the presence of Fe (0.01 M) and absence of Al, the metals Cd, Co, and Ni, are soluble at pH below 7. Similarly, Cu and REE are soluble at pH below 5, and a slight amount of them precipitate at the pH range of 5 to 7; and a slight amount of U precipitate at pH range from 3.5 to 4. The behavior of the elements in this study with regard to the pH range will be explained in detail in manuscript 1. Using Q-mode clustering,

the samples were mainly divided into two main clusters. GTF 3, 7, 8, 9, 10, 13, 15, 18, 19, 20, 21, 23, 24 and 26 are the members of the first main cluster and GTF 2, 4, 5, 6, 11, 12, 16, 17 and 25 are located in the second main cluster. The similarity between the outcomes of different data sets is an indication of no seasonal effect in the study area. Moreover, using multivariate outlier detection methods, two groups of samples were defined. GTF7, 8, 9, 19, 20, and 21 are defined as the first sampling domain and GTF2, 4, 5, 6, 12 and 13 are defined as second domain. These two domains represent the contamination patterns that are influenced by different types of sediments that cover the sampling location. The first domain is located at the south of the test site, which is dominated by sand. The second domain is located mainly in the centre of the test site within the unit of silt and clayey silt/ varved clay with the facial interlocking with the northern unit. The samples were not defined based on the sampling season. The sedimentology and stratigraphy have an important role on mass transportation and hence, the elements distribution/ pattern. Porous media that have colloidal- size particle can exhibit ion exchange. Ion exchange process is almost limited to colloidal particles because they have a large electrical charge relative to their surface size. Ion substitution creates a charge imbalance on the surface of the particle. Hence, the ions of the opposite charge accumulate on the surface in order to compensate the charge imbalance. The general hydrographs of groundwater level fluctuation were compared with regards to the season and also precipitation range from 2006 to 2008. It is interesting to note that the hydrographs can be divided into two main groups, although there are some differences in details. Similar to the results that got by statistics, one group consists of the sampling points of the south of the test site (GTF6, 7, 8, 9, 20, 21, 22, 23, 24, 32, 33 and 34) and the second group is including the samples of central to the north of the studied area (GTF 2, 3, 4, 5, 11, 12, 13, 14, 15, 16, 17, 18, 19, 27 and 30). As it is shown in Fig. 4, the hydrograph of GTF3, has a smooth profile with mildly varying fluctuation in groundwater level; while the hydrograph of GTF6 has a saw-tooth profile with repeated triangular transitions and widely varying fluctuations.

Regardless the minor fluctuation, the hydrographs of the first group show a general increasing in the period of December 2007 to August 2008, and a sharp decreasing in October 2008. However, the second group doesn't have such a significant increasing in this period. These small differences between the hydrographs of each group are more apparent during the period of the highest groundwater fluctuations (June to October 2007). The maximum variation of precipitation (0- 38 mm/day) was reported for this duration and can be noted on the precipitation graph. It is interpreting that the groundwater level fluctuation is a response to precipitation event but this response is more noticeable when a higher precipitation occurs. As known from the field work, the thin aquifers are running through the different geological units. It implies the role of geological units on groundwater fluctuation. The flow needs more time to react with the sediments to elution REE in the silty central part on the test site. Comparing the local geological profiles surrounding the wells, shows that the local glacial sediments composition, and hence its hydraulic parameter have an important role of REE concentration. The increase of REE concentration from the south to the north is more noticeable when groundwater level remains within a fine grain layer, rather than remaining within a coarse grain layer.

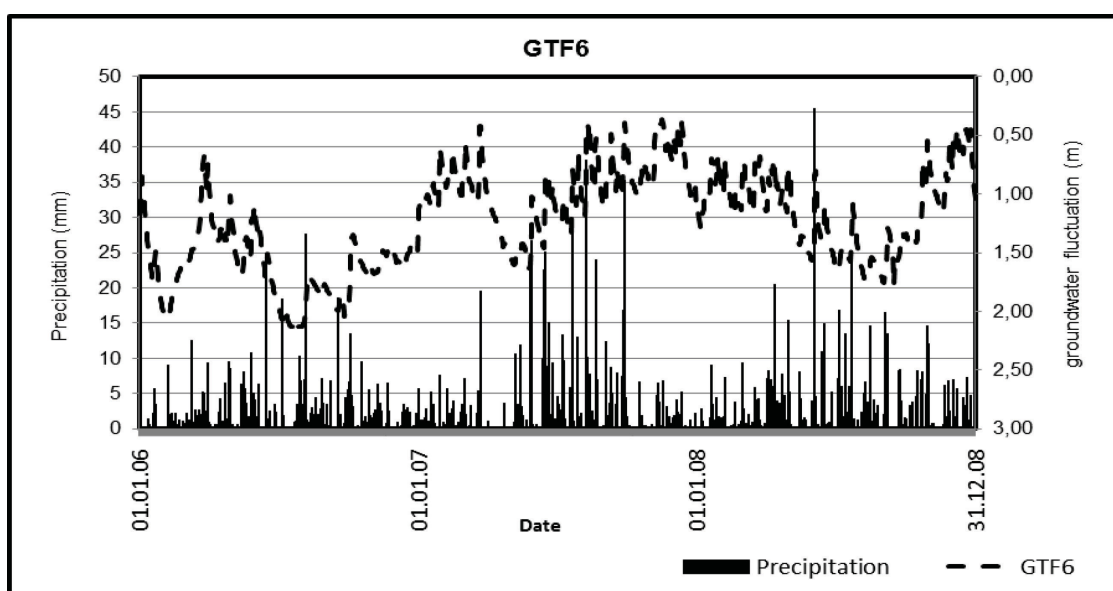
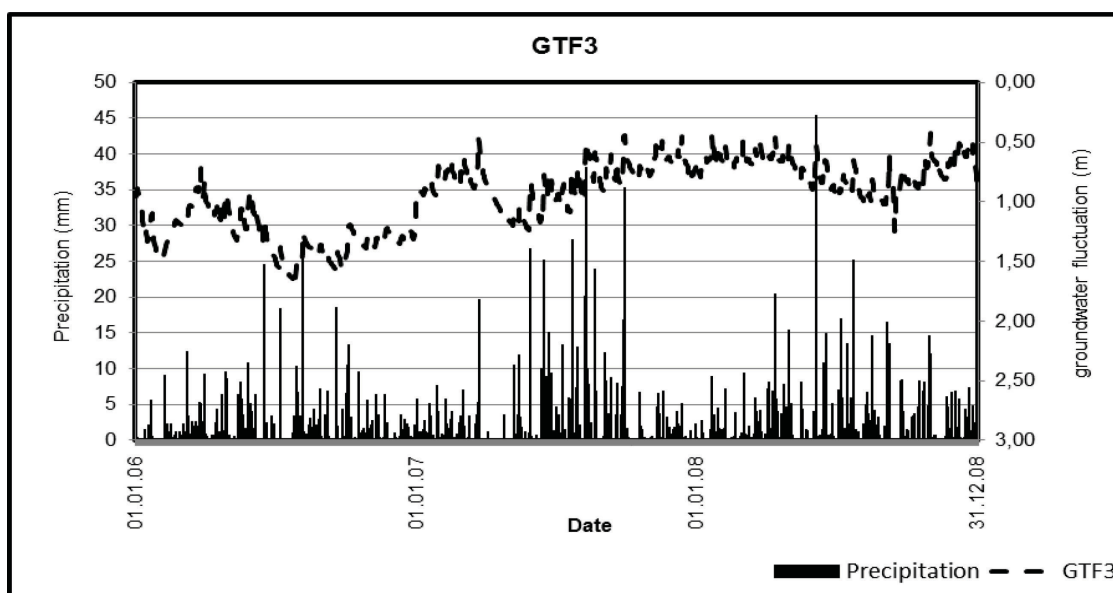


Fig. 4 The groundwater level fluctuation of sampling points 6 and 3 from 2006 to 2008. The hydrograph of GTF3, has a smooth profile with mildly varying fluctuation in groundwater level; while the hydrograph of GTF6 has a saw-tooth profile with repeated triangular transitions and widely varying fluctuations

6.2 Elements distribution in soil

The distribution of the contamination as well as pH range, and REE pattern in top soil (water- leached soil samples and ammonium nitrate-leached soil samples) were studied and compared to the contamination distribution in groundwater samples in order to investigate the influence of groundwater on the top soil contamination. According to the results, the samples leached with ammonium nitrate exhibit pH between 3.8 and 4.3, whereas groundwater pH ranges between 3.2 and 5.4. The water- leached samples are all above 4.5 to 5.3. Water had a pH value of 5.5 after equilibration with the air, and ammonium nitrate had a pH of 4.7. The soil samples obtained at the locations that were amended as a part of a prior project exhibited higher pH values when leached with ammonium nitrate (4.8 to 6.2) and water (7.6 to 7.8).

The Σ REE concentration in groundwater samples varies approximately from 30 to 8000 $\mu\text{g/L}$. The concentration of Σ REE in samples leached with water varies from 0.007 to 123 $\mu\text{g/L}$ leachate corresponding to 0.00007 to 1.23 $\mu\text{g/g}$ soil. Samples leached with ammonium nitrate contain a higher concentration of REE; 4- 7 $\mu\text{g/g}$ range (400-700 $\mu\text{g/L}$ leachate) with an average of 6.5 $\mu\text{g/g}$ soil (650 $\mu\text{g/L}$ leachate). This higher concentration could be due to cation exchange, and also the role of clay minerals. Clay minerals are negatively-charged colloids in soil that play a role in ion exchange. Clay can adsorb cations of specific type and amount. The total amount will balance the charge deficiency of the particle. Although ion exchange does not affect the structure of the clay particle, physical and physiochemical change in soil properties may occur. With regards to the comparison of the REE patterns of the data, generally, the patterns are similar qualitatively. All samples exhibit a MREE- enrichment and LREE- depletion with respect to the HREE. Furthermore, all samples except of some water-leached samples have a positive Ce anomaly. The REE pattern from the groundwater samples, exhibit a clear MREE enrichment and LREE depletion relative to the HREE. All groundwater samples exhibit a slight positive Ce anomaly ($\text{Ce}/\text{Ce}^* = 2.4 \pm 0.6$). Moreover, some samples exhibit also a slight positive Gd anomaly ($\text{Gd} / \text{Gd}^* = 1.3 \pm 0.06$). REE patterns of water-leached samples also display a MREE enrichment, with a positive Gd anomaly ($\text{Gd}/\text{Gd}^*=1.3 \pm 0.6$). Furthermore, many samples display a positive Ce anomaly ($\text{Ce}/\text{Ce}^*=0.8 \pm 0.6$). It is interesting to note that the Ce anomaly is higher for samples obtained from deeper horizons (lower than 20 cm) as compared to near-surface ($\text{Ce}/\text{Ce}^*>2$ vs $\text{Ce}/\text{Ce}^* <1.5$). LREE are also depleted relative to HREE (more explanation in manuscript 2).

With regards to the results of clustering, the data of water-leached samples reveals three main clusters. REE show similarities to Cu, U and Y; unlike the groundwater samples, the elements Al, Fe and Th are not in the same cluster with REE. These elements together with Pb and PO_4^{3-} are forming another main cluster. In the ammonium nitrate-leached samples, the analytical parameters formed two main clusters. REE with exception of La are similar to Al, Cu, Fe, Pb, U and Y. The second main cluster includes La, Mg, Mn, S and Zn. The dendrogram resulting from this data set is similar to that of the groundwater samples. With regard to the results of Q-mode clustering, data from samples leached with water formed three main clusters. The first cluster reflects the samples obtained from the zones with normal soil. The second cluster consists of data from samples collected from the zone within the test site that having a higher level of contamination. The data from samples leached with ammonium nitrate formed two main clusters as the case for the water-leached samples. The clusters correspond to the zones of the normal soil and to the higher level of REE. The mentioned similarities and dissimilarities could be explained due to the pH range of the samples, since pH influences the process of precipitation and co-precipitation. In this study heavy metals are typically precipitated as hydroxides, which is the most common heavy metal precipitation. Other precipitation such as carbonate and sulphate cannot occur in this area due to the pH range and present ions. For example, metal sulfides are very insoluble. Therefore, metals can be precipitated in presence of sulfide ions (S^{2-}). Metal sulfides have much lower solubility than the corresponding metal hydroxides; hence, resulting in lower metal concentrations in the contaminant water. Sulfide precipitation is always occurring under alkaline conditions to promote sulfide ion formation. Precipitation of a heavy metal may cause the co-precipitation of another metal. Co-precipitation is a process in which a solute is removed from a solution containing other ions which are in the process of precipitating. These

co-precipitated ions are incorporated into the solid by adsorption on the surface of the growing particles or substitution within the crystal lattice. Adsorption is one of the principle mechanisms of co-precipitation.

6.3 REE distribution in slate samples

In order to find the most comparable samples with slates, various samples including groundwater and soil water samples were studied. For this study, classic multivariate study (cluster classification and PLS modeling) were used. Three slates leached (water-leached, AMD-leached and sulfuric acid-leached) as well as groundwater, and soil water samples were studied by mentioned statistical methods. The outcomes were not significant and no meaningful relation between the studied samples was detected. The only data set that has an acceptable significance was soil water samples in comparison to water-leached slates. It should be taken into account that the processes that are present in soil water samples are much more complicated than soil, and groundwater. The reason is that transport through the unsaturated zone is more complicated than in saturated zone, because there are different phases of interest such as soil, water, air and contaminants. Similar to groundwater, the main transport processes are diffusion and advective transport. However, the contaminant in soil can exist in four phases: dissolved within soil moisture, sorbed onto soil particle, a separate contaminant phase, and a vapor.

However, fuzzy analysis could find some relationships between soil water samples and slates samples leached by water. The slate samples are more or less close to the majority of samples collected in sampling point MF1 and MF2. According to the membership degrees, it is observed that there is a similarity between the slate samples, and a large difference between the samples collected in sampling point MF3 at the depth of 100 cm and the rest of samples. The elements Cd, Co, Cu, Li, Ni, REE and Y are in one cluster. The other cluster includes Cs, Fe, K, Pb, Sc, Th, and Ti. The last cluster includes Ba, Ca, Mg, Mn, Na and Sr.

6.4 Heavy metals in lake Norrtorpsjön and lake Pölen

The data that were collected from two lakes, Lake Norrtorpsjön and Lake Pölen were analysed statistically with the similar classical, statistical methods that the data from Gessenhalde, were studied. Comparing the results will help to define the factors that control the contamination distribution. In Lake Norrtorpsjön there is a decline from pH 7.4 at 1 m to 6.4 at 8 m and further down to pH 5.9 at 27 m. This is in contrast to the constant pH of 3.2 found in the entire Lake Pölen. The highest concentrations of Na (60-80 mg/L) and K (25-40 mg/L) were measured at depth of 1 m. The elements Mg, Fe and Mn have almost identical depth profiles, with low concentrations in the epilimnion and an abrupt increase in the thermocline. Such concentration profiles would indicate formation of settling particles above the thermocline, followed by dissolution/release below it. Lake Pölen is close to equilibrium with respect to Fe-hydrite (S.I. 0.08). In Lake Norrtorpsjön Fe-hydrite was oversaturated above the thermocline (S.I. 4.9), but under saturation limiting concentrations below it. The redox potentials is in the range of 0.40 mV to -150 m. In such oxic region, large particles of Fe-(hydro)oxides are formed. It is also evident that the bottom waters of the lake contain a large amount of dissolved Fe. The concentrations of Al in the acidic Lake Pölen (750-800 µg/L) do not represent saturation with hydroxides or hydroxysulphate solid phases. The conditions are the

opposite in Lake Norrtorpsjön where several modifications of hydroxides and hydroxysulphates are oversaturated. This could be the reason for the rather low concentrations of Al (90-135 µg/L) as well as its retention in the filters throughout the water column. Manganese had a similar depth dependence as Fe in Lake Norrtorpsjön, with moderate concentrations above the thermocline (200 µg/L) and high (8000 µg/L) in the bottom water. Precipitation of Al-hydroxides and hydroxyl sulphates as well as of Fe-hydroxides can result as higher pH in Lake Norrtorpsjön. In Lake Pölen most of the REE are found at rather high concentrations, in relation to non-shale environments. The conditions are quite different in Lake Norrtorpsjön where only Dy, Er and Sc are present at concentrations above the detection limit (more detail in manuscript 4). According to the statistical results, in Lake Norrtorpsjön, the samples were classified in two main clusters. Cluster A represents samples from a depth of 9 m and deeper; cluster B were samples between 5 m and 8 m, and cluster C represents samples above 5 m. The physical interpretation is of course that the three compartments, epilimnion, thermocline and hypolimnion, were identified with this strategy for data evaluation. Based on R-mode cluster analysis the analytical parameters are divided into two main clusters using the analytical data. Two main sub-clusters were identified within first main cluster. The one sub-cluster includes Ca, Fe, Ga, Mg, Mo, Na, Ni, Sc and U while the second sub-cluster includes Al, K, Pb, V and Zn. Also second main cluster is divided into two main sub-clusters. The first sub-cluster includes Ce, Eu, Gd, La, Lu, Nd, Pr and Y while the second sub-cluster includes Cd, Cu, Ga, pH and V. In brief, the analysis indicates that the relationship between the parameters considered here possibly have different origins and different behaviour in different compartments of the lake. A general interpretation would be that the relationship between them depends on the chemical environment and the properties of the individual element. One critical parameter is depth and another is pH. The pattern also indicates that in order to better understand the system it should be sampled under different hydrochemical conditions.

6.5 The behavior of heavy metals before and after remediation, data from lake Gruvsjön, Kuntebo creek, and Lake Risten

The samples that were taken from the surrounding area of an old mine, before and after remediation period were analyzed by statistical methods in order to evaluate the influence of the remediation. The differences between the behaviour of heavy metals can be seen in the Q-mode cluster analyses as well as in the time series analyses. The Q-mode clusters were generally good representative of the samples that were taken before or after the remediation period. The samples from Lake Gruvsjön were divided into three main clusters using Q-mode cluster analysis. Cluster one includes the samples that were taken in 1990s (from 1990 to 95); the second cluster as well as the third cluster include the samples that were taken in the period of 1988-95. Based on R-mode clustering, the analytical parameters are divided into two main clusters: cluster one includes Al, Cd, Cu, Mn, Pb and Zn, and cluster two includes Fe and pH. The measured concentrations of Al, Cu and Zn have similar behaviour and, decreased noticeably after remediation period. There is not a noticeable difference between the measured concentration of Mn and Pb before and after remediation period. The measured concentration of Fe increased noticeably after the remediation period, and a peak of concentration was measured in 1995.

The samples from Kuntebo were divided into two main clusters by Q-mode cluster analyses. Cluster one include the samples that were taken before 1990 and cluster two includes the samples were taken after this year. The analytical parameters were divided into two main clusters by R-mode cluster analysis. The first cluster includes Al, Cd, Cu Fe, Mn, Pb, and Zn. The second cluster has only one member: pH. The concentration of Al, Cd, Cu and Pb, decreased noticeably after the remediation period, and the forecasted values are higher than the measured ones after remediation. However, the concentration of Zn does not change noticeably after the remediation period, and the forecasted values are less than the measured ones.

The samples from Lake Risten were divided into two clusters using Q-mode cluster analyses. Cluster one includes the samples that were taken before 1988, and cluster two contains samples that were taken after this time. By means of R-mode clustering, the analytical parameters were divided into four main clusters. Cluster one includes Cu, Pb and Zn. Cluster two includes Al, Cd and Fe. Cluster three and four have only one member, Mn and pH respectively. A significant decrease occurred after the remediation period (mainly in 1990) for each of the elements Cu, Pb and Zn. Elements Al, Cd and Fe behave in a similar way as Al and Cd from Kuntebo and Lake Gruvsjön. The only exception is Mn, which shows a stable concentration before and after the remediation period (more detail in manuscript 5).

7 Conclusion:

It can be concluded that different sources are contaminating the test site. The main sources of contaminations spread from the various phases of infiltration that occurred from 1970s to 1990s, the geochemical processes (e.g., precipitation, co-precipitation), and the composition of the glacial sediments. Some other minor factors also influence the concentration of the heavy metals, for example, groundwater level fluctuation that is a result of seasonal precipitation, return flow of groundwater, and the pH range. As is shown in Fig. 5, the AMD that was used for leaching uranium has infiltrated the clay material that was used under the heap to contain the leachate, as well as to the underlying glacial sediments; hence, contaminants have reached the groundwater. This first phase of infiltration caused a secondary mineralization and thus released contamination to the groundwater (Fig. 5.1). After remediation and the removal of the leaching heap, the pit was left open and thus rainwater led to a second phase of infiltration forming above the groundwater level (Fig. 5.2). More heavy metals were released from the ponded water and mineralized. Due to the remediation, the contaminated glacial sediment was excavated down to 10m and the area was filled with non-contaminated top soil and groundwater level was changed, and the third phase of infiltration was formed (Fig. 5.3).

This remediation work was expected to cease the release of heavy metals to the surrounding environment. However, the concentration of heavy metals that were measured several years after the remediation was completed contradicts this expectation. In order to understand the reason for the ongoing contamination, various factors were considered. It is important to note that the physicochemical parameters of the area varied during the different phases of infiltration. Hence, in each phase various chemical processes are possible that can lead to the release of various species of metals. Another important factor that influences the contamination is the sedimentology and stratigraphy. Higher REE concentrations were measured in middle part of the test site where silt and clay silts are located. The

critical role of the sedimentology and stratigraphy is the main reason for the stability of the contamination distribution of the current groundwater table, although the concentrations of the heavy metals also vary due to the groundwater level fluctuation. Higher concentrations of heavy metals were measured during the period of decreased rainfall. With regards to this fact, it is probable that the change of groundwater level due to remediation (during 1990s) has also influenced the degree of heavy metal concentration.

Although multivariate statistical analysis were helpful to define the relation between the data sets from various media, defining one source as the main source of contamination was not easily possible. By means of fuzzy classification, the relation between the leached Ordovician-Silurian slates samples

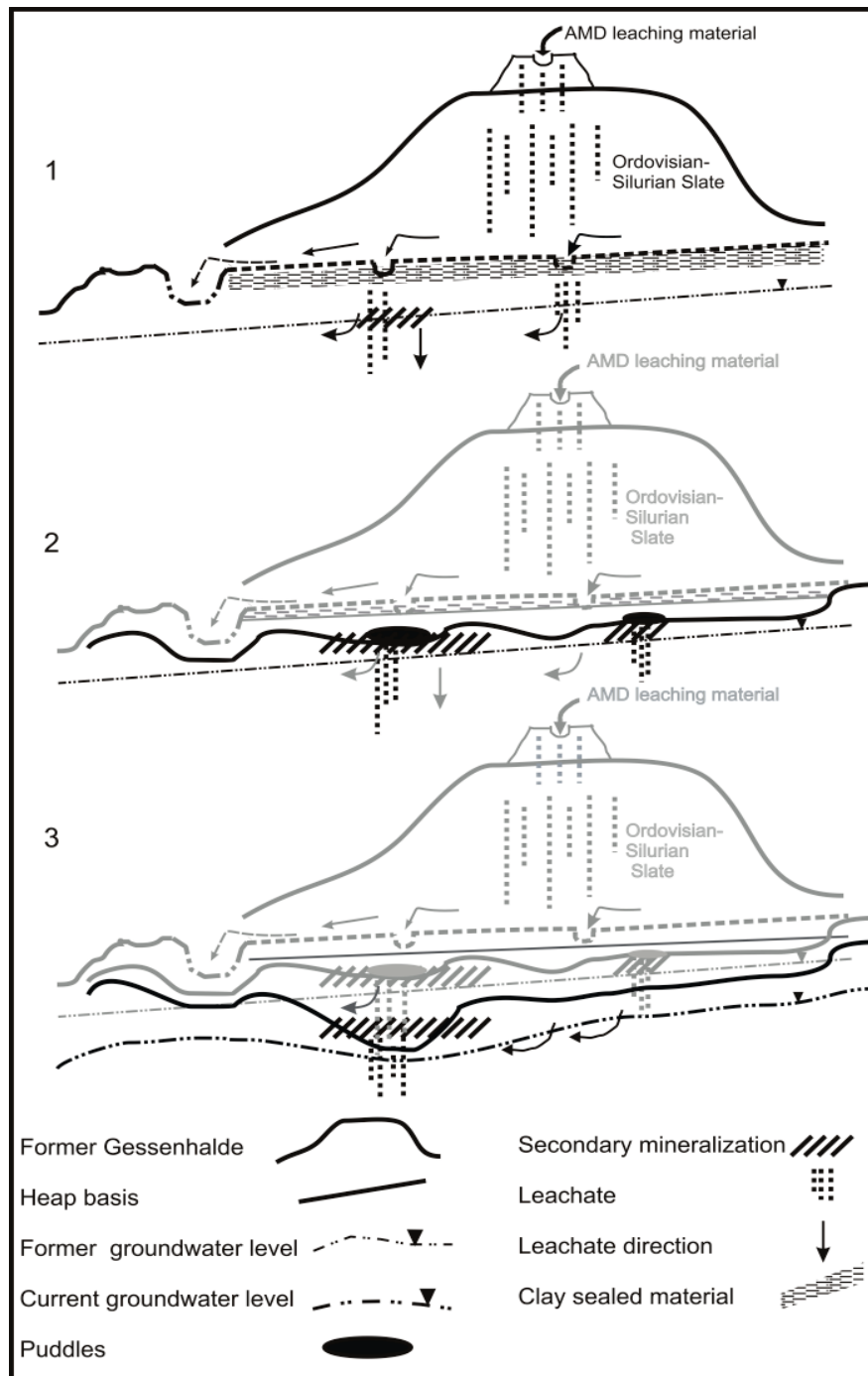


Fig. 5 Three phases of infiltration: 1) AMD infiltration into the heap and reached to the groundwater, 2) secondary mineralization and releasing the heavy metals and infiltrating into the groundwater, 3) current situation of the area after remediation and excavating the glacial sediments

and soil water samples was identified. However, to define the slates as the main source of contamination measured in soil water samples is not proved yet. It is more possible that the condition of the slates (weathered or non-weathered) and also the physicochemical situation of the media surrounding the slates have more influence of the contamination concentration. To summarize, evaluating the phases of contamination is very helpful in order to interpret the current contamination condition, in some occasions multivariate statistics is a helpful tool. With regards to the stability of the contamination distribution/pattern, the intervals of sampling can be reduced to two times per year. It is suggested to design a systematic sampling net/profile that covers all sampling domains (as described in Section 6.1); with three to four samples from each domain, and that has overlap with other geological/geochemical information (e.g., borehole profiling, variety of sample types, monitoring groundwater level fluctuation, etc.). In this case, more (geo)statistical methods can be performed and, hence, more details can be investigated with less physical work. It will help the project to decrease the expenses and optimizing the project economically.

References:

- Beven K, Germann P (1982). Macropores and water flow in soils. *Water Res. Res.* 18: 1311–1325.
- Bezdek J (1980). A convergence theorem for the fuzzy ISODATA clustering algorithms. *IEEE transactions on pattern analysis and machine intelligence.* 2 (1): 1–8.
- Bezdek J (1981). *Pattern recognition with fuzzy objective function algorithms*, Plenum Press.
- Bhattacharya A, Routh J, Jacks G, Bhattacharya P, Mörtz M (2006). Environmental assessment of abandoned mine tailings in Adak, Västerbotten district (northern Sweden). *App. Geo.* 21: 1760–1780.
- Bezdek J, Hathaway R, Sabin S, Tucker W (1987). Convergence theory for fuzzy c means: counter examples and repairs. *IEEE T. Syst. Man Cyb.* 17 (5): 873–877.
- Brookins DG (1989). Aqueous geochemistry of rare earth elements. *Reviews in Miner. Geochem.* 21(1): 201-225.
- Bozau E, Bechstedt T, Friese K, Frömmichen R, Herzsprung P, Koschorreck M, Meier J, Völkner C, Wendt-Potthoff K, Wieprecht M, Geller W (2007). Biotechnological remediation of an acidic pit lake: modeling the basic processes in a mesocosm experiment. *J. Geochem. Explor.* 92 (2-3): 212-221.
- Bundesgesetzblatt Jahrgang (2011). Teil I, Nr. 21. Bonn (In German).
- Carlsson E, Büchel G (2005). Screening of residual contamination at former uranium heap leaching site, Thuringia, Germany. *Chem. Erde- Geochem.* 65: 75-95.
- Chopin EIB, Alloway BJ (2007). Distribution and mobility of trace elements in soils and vegetation around the mining and smelting areas of Tharsis, Riotinto and Huelva, Iberian Pyrite Belt, SW Spain. *Water Air and Soil Poll.* 182: 245–261.
- Davis JC (2002). *Statistics and data analysis in geology*. John Wiley & Sons. New York, United states of America.
- Demicco R, Klir G (2003). *Fuzzy logic in geology*. Elsevier.
- Derde MP, Massart DL (1982). Extraction of information from large data sets by pattern recognition. *Fres. Z. Anal. Chem.* 313: 484-495.
- Dumitrescu D, Sârbu C, Pop HF (1994). *Anal Lett* 27:1031–1054
- Einax JW, Zwaniger HW, Geiß S (1997). *Chemometrics in environmental analysis*. VCH.
- Elias RW, Gulson B (2003). Overview of lead remediation effectiveness. *Sci. Total. Environ.* 303 (1-2):1-13.
- Filzmoser P, Garrett RG, Reiman C (2005). Multivariate outlier detection in exploration geochemistry. *Comput. Geo. Sci.* 31(5): 579-587.
- Filzmoser P, Steiger B (2010). Package "StatDa", <http://www.statistik.tuwien.ac.at/StatDA/R-scripts/index.html>. March 2008.
- Franklin MR, Fernandes HM (2011). Identifying and overcoming the constraints that prevent the full implementation of decommissioning and remediation programs in uranium mining sites. *J. Environ. Radioactiv.* 2011:1-7.
- Freeze RA, Cherry JA (1979). *Groundwater*. Prentice Hall, Inc.
- Gomes MEP, Favas PJC (2006). Mineralogical controls on mine drainage of the abandoned Ervedosa tin mine in north-eastern Portugal. *Appl. Geochem.* 21: 1322–1334.
- Geletneky J W, Büchel G, Paul M (2002). Impact of acid rock drainage in a discrete catchment area of former uranium mining site of Ronneburg (Germany). *Tailings Mine Waste.* 02: 67-74.

- Grawunder A (2010). Hydrogeochemistry of rare earth elements in an acid mine drainage influenced area. Doctoral dissertation, Friedrich Schiller university of Jena.
- Hadley R, Snow D (1974). Water Resources and Problems Related to Mining. American Water Resource Association, MN.
- Hair JF, Anderson RE, Tatham RL, Black WC (1992). Multivariate data analysis (3rd edition). Macmillan, New York, Unites States of America.
- Helsel DR, Cohn TA (1988). Estimation of descriptive statistics for multiply censored water quality data. Water Res. Res. 24: 1997-2004.
- Hafeburg G (2007). Studied on heavy metals resistance of bacterial isolates from a former uranium mining area. Doctoral dissertation, Friedrich Schiller university of Jena.
- Horn K (2003). Das Migrationsverhalten von Radon im Sanierungs gebiet der osttüringischen Uranbergbauregion. Doctoral dissertation, Friedrich Schiller university of Jena (In German).
- Hölting B, Coldewez WG (2005). Hydrogeologie. Elsevier. Munich, Germany.
- Jakubick A, Jenk U, Kahnt R (2002). Modeling of mine flooding and consequences in the mine hydrogeological environment: Flooding of the Koenigstein mine, Germany. Environ. Geol. Water. S. 42(2): 222-234.
- Kaiser H (1965). Zum Problem der Nachweigrenzer. Workshop: Moderne Methoden der anorganischen Analyse, 05- 07.10. 1965. Düsseldorf, Germany (In German).
- Kaiser HF (1974). An index of factorial simplicity. Psychometrika. 39: 31-36.
- Kung, K.J (1990). Preferential flow in a sandy vadose zone, 2: Mechanisms and implications. Geoderma. 22: 59–71.
- Köher SJ, Harouiya N, Chairat C, Oelkers EH (2005). Experimental studies of REE fractionation during water–mineral interactions: REE release rates during apatite dissolution from pH 2.8 to 9.2. Chem. Geol. 222 (3-4): 168-182.
- Lange G (1995). Die Uranlagerstätte Ronneburg, Zeitschr. Zeitschrift für Geologische Wissenschaften 23: 517-526 (In German).
- Lee CH (2003). Assessment of contamination load on water, soil and sediment affected by the Kongjujeil mine drainage, republic of Korea. Environ.Geo. 44: 501–515.
- Lee L, Helsel D (2005). Statistical analysis of water-quality data containing multiple detection limits: S-language software for regression on order statistics. Comput. Geo. Sci. 31(10): 1241-1248.
- Liang-qi L, Ci-an S, Xiang-li X, Yan-hong L (2008). REE behavior and effect factors in AMD-type acidic groundwater at sulfide tailings pond, BS nickel mine, W.A. T Non ferr. Metal. Soc. 18(4): 955-961.
- Lonschinski M (2009). Schwermetalle im System Boden-Wasser-Pflanze mit Hinblick auf Fraktionierungsprozesse der Seltenen Erden Elemente auf der Aufstandsfläche der Gessenhalde im ehemaligen Uranbergbaurevier Ronneburg (Ostturingen). Doctoral dissertation, Friedrich Schiller university of Jena (In German).
- Lorenz C (2009). Untersuchungen zum geomikrobiellen Schwermetalltransfer und simultanen Monitoring seltener Erden Elemente (in German). Doctoral dissertation, Friedrich Schiller university of Jena.
- Moritz G, Filzmoser P (2009). Package “mvoutlier for Multivariate Outlier Detection”, <http://www.statistik.tuwien.ac.at/StatDA/R-scripts/index.html>.

- Moreno M, Oldroyd A, McDonald L, Gibbons W (2007). Preferential fractionation of trace metals-metalloids into PM10 re-suspended from contaminated gold mine tailings at Rodalquilar, Spain. *Water Air Soil Poll.* 179: 93–105.
- Otte ML, Jacob DL (2008). Mine area remediation. *Encyclopedia of Ecology*. 2397-2409.
- Pop HF, Dumitrescu D, Sârbu C (1995). A Study of roman pottery (terra sigillata) using hierarchical fuzzy clustering. *Anal. Chim. Acta.* 310:269–279
- Pop HF, Sârbu C (1966). A new fuzzy regression algorithm. *Anal. Chem.* 68:771–780
- Pasalic S (2011). Untersuchung von Bodensubstraten im Bereich von Saugkerzenanlagen auf dem Testfeld Gessenwiese bei Ronneburg. Biological project module. Friedrich Schiller university of Jena (In German).
- Reimann C, Filzmoser P, Garrett RG, Dutter R (2009). *Statistical data analysis explained: applied environmental statistics with R*. John Wiley & Sons.
- Rodríguez L, Ruiz E, Alonso-Azcárate J, Rincón J (2009). Heavy metal distribution and chemical speciation in tailings and soils around a Pb–Zn mine in Spain. *J. Environ. Manage.* 90:1106–1116.
- Rüger F, Dietel W (1998). Vier Jahrzehnte Uranbergbau um Ronneburg. *Lapis*, 7(8): 14-18 (In German).
- Rönnback P, Åström M, Gustafsson JP (2008). Comparison of the behavior of rare earth elements in surface waters, overburden groundwater and bedrock groundwater in two granitoidic settings, eastern Sweden. *Appl. Geochem.* 23(7): 1862-1880.
- Sârbu C, Dumitrescu D, Pop HF (1993). Selection and optimal combination of solvent systems in TLC. *Rev. Chim. (Bucharest)*. 44:450–459.
- Sârbu C, Zehl K, Einax JW (2007). Fuzzy divisive hierarchical clustering of soil data using Gustafson-Kessel algorithm. *Chemometr .Intell. Lab Syst.* 86:121–129
- Sabin M (1987). Convergence and consistency of fuzzy c-means/ISODATA algorithms. *IEEEET Pattern Anal.* 9(5): 661-668.
- Sanford RF, Pierson CT, Crovelli RA (1993). An objective replacement method for censored geochemical data. *Math. Geol.* 25 (1):59-80.
- Schippers A, Hallmann S, Wentzien S, Sand W (1995). Microbial diversity in uranium mine waste heaps. *Appl. Environ. Microb.* 61(8): 2930- 2935.
- Silva EF, Bobos I, Matos JX, Patinha C, Rei AP, Fonseca EC (2009). Mineralogy and geochemistry of trace metals and REE in volcanic massive sulfide host rocks, stream sediments, stream waters and acid mine drainage from the Lousal mine area (Iberian Pyrite Belt, Portugal). *Appl. Geochem.* 24(3): 383-401.
- Tindall JA, Kunkel JR (1999). *Unsaturated zone hydrology for scientists and engineers*. Prentice-hall, New Jersey, United States of America.
- Todd DK, Mays LW (2005). *Groundwater hydrology* (3rd edition). John Wiley & Sons. New York, United States of America.
- Verplanck PL, Antweiler RC, Nordstrom DK, Taylor H E (2001). Standard reference water samples for rare earth element determinations. *Appl. Geochem.* 16: 231-244.
- Wagner S (2010). *Geoschemische Untersuchungen an Paläozoischen Gesteinen aus dem Gebiet der ehemaligen Gessenhalde bei Ronneburg ehemaliges ostthüringisches Uranerzbergbaurevier*. Diploma dissertation, Friedrich Schiller university of Jena (In German).

- Wismut GmbH (1994a). Entwurf Sanierungskonzept Standort Ronneburg.-Stand März. Chemnitz (in German).
- Wismut GmbH (1994b). Sanierungskonzept Standort Ronneburg.-Stand Dezember. Chemnitz (in German).
- Wold H (1981). The fix-point approach to interdependent systems. North- Holland. New York, United States of America.
- Wold H (1985). Partial least squares, Encyclopedia of statistical sciences. John Wiley & Sons. New York, United States of America. 82: 581-59.

Manuscript 1

Statistical evidence of groundwater contamination and significant factor with special influence of REE in former uranium mining site, eastern Thuringia, Germany

Statistical evidence of groundwater contamination and significant factors with special emphasis of REE in a former uranium mine, eastern Thuringia, Germany

Anahita Pourjabbar¹ - Anja Grawunder¹ - Dirk Merten¹ - Jürgen .W. Einax² - Georg Büchel¹

¹ Institute of Geosciences, Friedrich Schiller University, Jena, Germany

² Institute of Inorganic and Analytical Chemistry, Friedrich Schiller University, Jena, Germany

Abstract

Acid Mine Drainage (AMD) process in the former uranium mining area in Eastern Thuringia, Germany has contaminated the groundwater by uranium, rare earth elements (REE) and other heavy metals. The focus of this work is defining the significant factors such as physicochemical, geological and also seasonal effects that control the REE distribution, and its relation to the other elements in groundwater. 174 groundwater samples were collected from the basement area of a former leaching heap. Multivariate outlier detection, hierarchical- cluster and cross-correlation analyses were used to define the relation between REE and other analytical parameters. Groundwater level fluctuation and daily precipitation were studied to investigate the role of (hydro) geology and seasonal factors. The pH and, sedimentology and stratigraphy of the area are the most important factors that affect the REE distribution in this area. Also some elements such as Al, Cu, Fe, U and Y show similarities with REE in the cluster analysis. The pH has an inverse correlation with REE. The pH value of 4.5 can be considered as a boundary for REE partitioning in this study. No specific seasonal effect was found by statistics. However, groundwater fluctuation shows a meaningful correlation to precipitation data for a time period of elevated precipitation. The time lag shows a quick response to precipitation. It can be concluded that the geological composition and also precipitation are significant factors that affect REE concentration in the test site. Although precipitation is a seasonal dependent factor but this factor does not affect the REE distribution and data domain.

Key words: Cluster analysis; Multivariate outlier; Cross-Correlation; Groundwater Contamination; Heavy Metals; Rare Earth Elements; Precipitation.

1 Introduction

The former uranium mining district in Eastern Thuringia, Germany with more than 113,000 tons of mined uranium, was the third-largest uranium producer in the world (Jakubick et al., 2002; Lange, 1995). Mining activities started in 1949 and ended in 1990 after re-unification of Germany. The remnants of the mining include: a) a large open pit mine (the Lichtenberg pit with more than 230 m depth, 1.6 km length and 0.9 km width which is filled now), b) an underground mining system going down to 900 m with 3000 km of underground galleries, and c) several waste rock piles in which acid mine drainage (AMD) occurred (Wismut GmbH, 1994).

A uranium leaching heap that was a part of the mining activities at Lichtenberg is called Gessenhalde. Waste rocks with a low grade of uranium mineralization (uranium content <300 g/ton; Rüger and Dietel, 1998) were leached with AMD and sulfuric acid (10 g/l) (Wismut GmbH, 1994). During the leaching process, leachate seeped through the lining of the Gessenhalde and accumulated in the Quaternary sediments underneath. The leaching process was stopped in 1989, and in 1990s, the leaching heap (Gessenhalde) was placed into the open pit, Lichtenberg, along with an underlying layer of contaminated Quaternary sediments of variable thickness (up to about 10 m).

Due to the leaching process, surface water, seepage water and groundwater were highly mineralized with a pH value of about 3 and the water has been contaminated by uranium, rare earth elements (REE) and other heavy metals. In 2004 the test site "Gessenwiese" was created in the northern part of the base area of the former leaching heap, Gessenhalde, with the aim of improving remediation strategies for heavy metal contaminated areas (Carlsson and Büchel, 2005).

REE have been extensively studied in earth science fields, since they can record subtle geochemical processes in natural systems (Leybourne et al., 2000). The attraction of using REE to investigate the geochemical problems is that they form the coherent group of trace metals whose properties change systematically across the series La through Lu (Brookins, 1989). They are often grouped into light rare earth elements (LREE; La to Pm), middle rare earth elements (MREE; Sm to Dy) and heavy rare earth elements (HREE; Ho to Lu). Recent studies have focused on REE concentration and the factors that control REE distribution in surface waters, e.g., river, lake and mine surface water (Bozau et al., 2004; Elbaz and Pupy, 1999; Gammons et al., 2003; Protano and Riccobono, 2002; Zhao et al., 2007). Furthermore, traces of REE were used in previous studies as a tool to understand the processes in soil- water- plant systems (Kimoto et al., 2006; Merten et al., 2005).

Studies on REE geochemistry of waters, including seawater, river water, groundwater, lake water, and rain water indicate significant differences in composition of dissolved REE in different water columns (Goldstein and Jacobsen, 1988; Sholkovitz, 1993; Sholkovitz et al., 1994; Iderfield et al., 1990). In general, REE concentrations in AMD-influenced areas are higher than the areas not influenced by AMD (Bozau et al., 2004; Dia et al., 2000; Merten et al., 2005; Miekeley et al., 1992; Smedly, 1991; Tang and Johannesson, 2006). The significant factors are mainly the composition of the host rocks in the source region and intensity of chemical weathering, the distribution of different particulates (e.g. colloids) in the water column, pH and hydrochemical composition of the water, especially metal (Al, Mn and Fe) contents and complexation (Ding et al., 2007; Elderfield et al., 1981; Lewis et al., 1998; Liang et al., 2005; Rönnback et al., 2008; Steinmann and Stille, 2008; Wood, 1990).

This paper studies the REE concentration and distribution and the factors that play a role such as the physicochemical, geological, and seasonal parameters, in groundwater of the test site and its relation to the other metals. Generally the difficulty in conducting environmental studies relates to the fact that they involve complex and interacting factors, which cannot be easily isolated and studied individually. In such situations multivariate analyses are helpful tools, since they place the objects into more or less homogeneous groups so that the relation between the groups is revealed. During the last decades multivariate statistical methods have been applied to study the analyses of large data quantities of chemical and environmental systems. Such techniques can be used in REE studies also, as there are multi factors that affect REE concentration and distribution. Although there are many studies of REE distribution, few of them applied extensive statistical analyses to their data set.

2 Geological and hydrogeological setting

The geological setting was described in previous studies (Carlsson and Büchel, 2005; Grawunder et al., 2009). Below the test site, Quaternary glacial sediments with at least 10 m thickness were found, including four units: a) a graded bedding of silty and gravelly sand at the base and sand at the top, b) silt, c) clayey silt/ varved clay, and d) an allochthonic soil material added up to the area during remediation. The southern part of the test site is dominated by unit (a); the middle test site by unit (b), and to a certain extent, by unit (c). The north of the test site is sandy also, but this sand has a higher proportion of silt (Fig. 1). The allochthonic soil material is a very heterogeneous, unsorted, and unlayered material that covers the area (Grawunder et al., 2009). The sand in the north and the silty material in the middle test site are connected, while the sand in the south appears more as an overlying unit. The distributive province was expected to be limno-fluvial.

Groundwater is flowing through sandy glacial sediments at the northern and the southern parts, and through thin layers of a perched aquifer, especially within the silty at the central part of the test site. The groundwater level on the test site fluctuates from about 3.5 to 2.5 m below the surface (262-261 m above the sea level) in the south and middle; and from 1.0 to 0.2 m (261 m above the sea level) in the northern test site (Lonschinski, 2009). The water infiltrating the aquifer follows the topographic gradient through the test site from south to north. The more specifically flow direction is from the south–southwest to the north–northeast. It is not possible to specify one or two flow paths, because the groundwater is flowing through very thin layers of gravel and sand. Shallow aquifers (about 0.9 m depth) were identified in the middle of the test site.

3 Materials and methods

3.1 Sampling and data set

The present study deals with 174 groundwater samples which were collected from 33 (GTF2-34) sampling points (Fig.1). Samples were collected seasonally from 2004 to 2008, a total of 30 samples were taken in April 2005 and 2007; 14 samples were taken in May 2005 and 2006; 84 samples were taken in September 2006 to 2008; and 46 samples were taken in December 2004 to 2006. Eh, pH, electrical conductivity (EC), and temperature were measured on site using portable instruments (WTW, pH320; WTW, LF320; WTW, external thermocouple. In the field, all samples except those for HCO_3^- analysis were filtered using 0.45 μm cellulose acetate filters (Sartorius), with glass fiber

prefilters (Sartorius). Samples for element analysis were acidified with HNO_3 (65%, subboiled) to $\text{pH} < 2$. All samples were kept cool (6°C) until analysis.

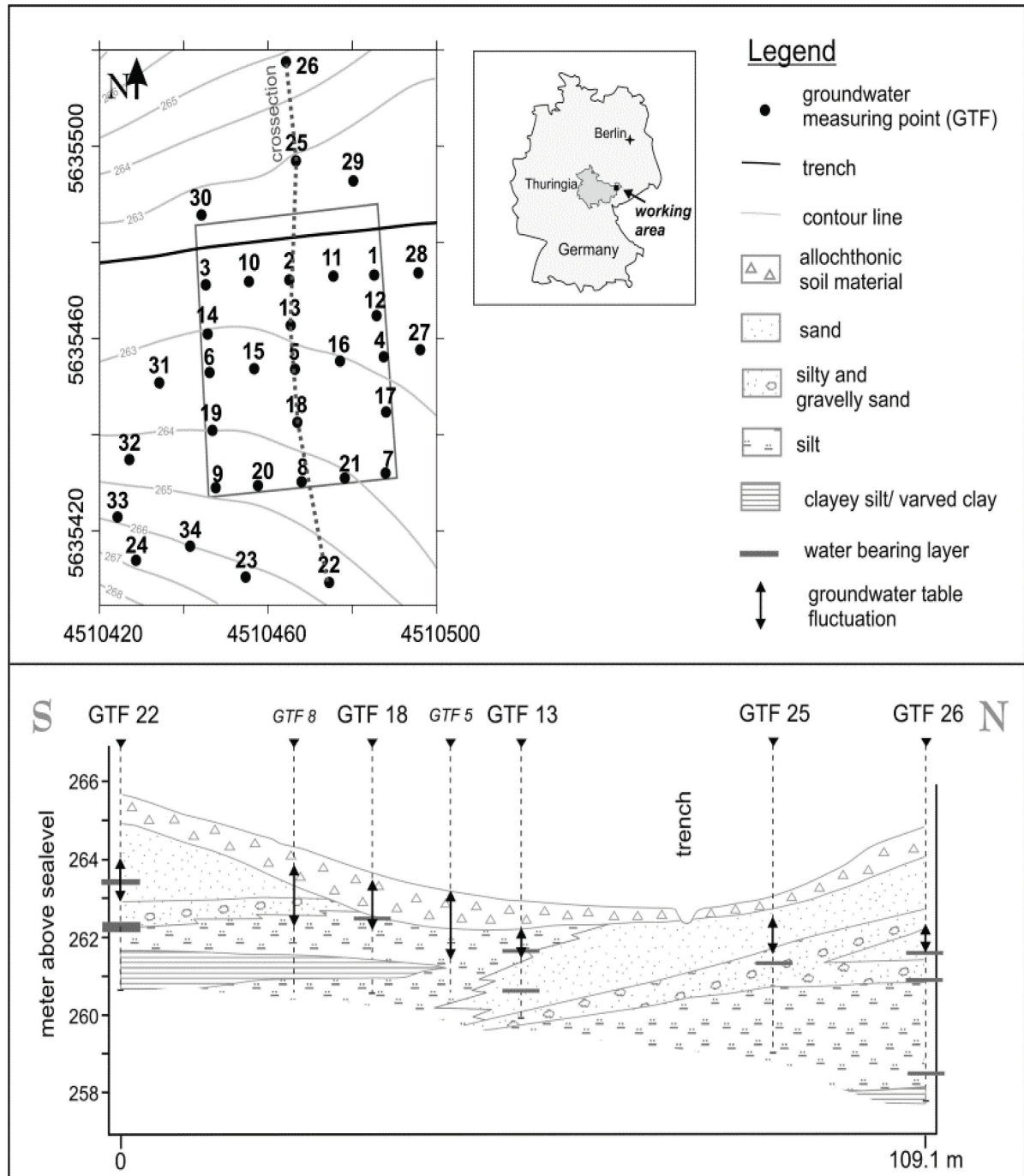


Fig. 1 Upper left picture: Map of the test site, Gessenwiese, with location of the groundwater measuring points (the coordinate system is Gauß-Krüger system). Lower picture: S-N- cross-section through the test site. It becomes apparent that the northern and the southern part of the test site are dominated by sandy and the middle test site rather by silty material. The units sand and silty/gravelly sand are within a graded bedding. Groundwater is running through thin layers. The geological information of GTF5 and 8 are not available

The chemical analysis of Al, Ca, Fe, Mg, Mn, Na, Sr, and Zn were performed with ICP-OES (Spectroflame, Spectro); Ba, Cd, Co, Cu, Li, Ni, Pb, REE (La-Lu), Sc, Th, Ti, U, Y, and Zn with ICP-MS (until 2007: PQ3-S, Thermo Elemental, subsequently: X-Series II, Thermo Fisher Scientific).

The anions Cl^- , F^- , Cl^- and SO_4^{2-} were analyzed using ion chromatography (DX120, Dionex), while PO_4^{3-} was measured with photometry (Hach- DR/4000U). For the determination of HCO_3^- an automatic titration system was used (Titrimo 716 DM, Metrohm). Dissolved organic carbon (DOC) was measured on 0.45 μm filtered samples (DIMA-TOC 100, Dimatec until 2006; and then multi N/C 2100, Analytik Jena). In order to check the accuracy of the analytical results, the instrument drift was monitored and corrected by using Be, Ru and Re as internal standards. Each sample was measured three times. First, an outlier test is performed on each of the three runs (Grubbs test, 90% significance, criterion 1.15). Then, for the remaining runs, the mean and standard deviation are calculated. Furthermore, analytical quality is checked by the use of standard reference materials: the standard reference materials SCREE and PPREE are used for the analysis for REEs (Verplanck et al., 2001). For other heavy metals, the standard reference materials SPS-SW2 (LGC standards) and NIST1643e (LGC standards) were used. Only values between 90-110 % for these reference materials are accepted. Detection limits (DL) are calculated according to the "3 sigma criterion" as follow (Kaiser, 1965):

$$\text{DL} = \text{Dil. Fact.} * 3 * \sqrt{2s_0/s} \quad (1)$$

$$S = I / c$$

Where, Dil. Fact. is dilution factor, s_0 is standard deviation, S is sensitivity, I is intensity, and c is concentration.

The groundwater level fluctuations were monitored every 15 minutes from 2004 to 2009 by data loggers in 11 boreholes GTF2, 3, 4, 6, 7, 8, 12, 13, 15, 18 and 25 (AquiLite ATM10 by Aquitronic (GTF2, 18); beaver, Aquitronic (for other wells)). The boreholes were constructed simply; the hole was drilled, the PVC casing (2 inch diameter) was inserted, and the surface was sealed to avoid the rain/surface water entering the boreholes. The depths of the boreholes varied to a maximum of 7 m. Moreover, geological profiles are available for the boreholes GTF10 to 34.

The daily precipitation data used in this study were supplied by the Wismut GmbH from a nearby measuring point.

To identify the role of seasonal effects, samples were separated into three data sets as spring (April, May), autumn (September) and winter (December). The reason to separate the data into seasonal data sets is to investigate the seasonal dependence between groundwater level fluctuation and precipitation in relation to REE distribution and concentration. All statistical analyses have been used in all data sets. Furthermore, the (hydro) geological information was studied as an aid to interpret the outcomes of statistical studies. This data includes groundwater level fluctuation, precipitation data and geological profiles. Since there are different glacial layers and each of them has its specific grain size distribution and concentration range of the elements, it is helpful to study these details in order to investigate their effects.

3.2 Statistics

3.2.1 Data preparation

Prior to using multivariate methods, censored data were replaced by estimated values and the univariate outliers were defined to make a complete set of data. Censored data are the data that are recorded as below specified analytical reporting limits (detection limit) due to measurement capacities

or practical concerns. If multiple censoring thresholds are present, then the data are called “multiply censored”. In this study, multiple censored data were reported in the analytical results for Cr, Cu, F⁻, Fe, Pb, PO₄³⁻, Th, Ti, Sc and U. The reason for reporting multiple censored data is that the instrument detection limit is not constant, but rather it is adjusted according to the sample’s dilution factor. The estimating method was applied based on the numbers of the censored data for each parameter (Table 1). The replacing values for substitution methods are 3/4 of each detection limit. One-half, seven-tenths of the detection limit are other common substitution values used. However, any single value between zero and the detection limit is as good as another (Lee and Helsén, 2005). Parameters with more than 10% censored data (Sanford et al., 1993) were replaced by the values that were calculated by ROS method in R software (Lee and Helsén, 2005). ROS is a method that is applicable for water data sets with multiple censored data. This technique is based on a “robust” semi-parametric method developed by Helsel and Cohen (1988). It computes a linear regression for data or logarithms of data versus their normal scores in a normal probability plot. The regression parameters use detected observations. Due to the definition of normal scores, fitting this line is fitting a normal distribution.

Table 1 Numbers of censored data in analytical result and used replacement method of groundwater samples; N= 174

Element/anion	Numbers of Censored Data	Detection limit	Replacement Method
Cu	3	<0.5	Simple Substitution
Pb	4	<0.1, <0.05	Simple Substitution
PO ₄ ³⁻	1	<0.05	Simple Substitution
Ti	5	<0.1, 0.2, 0.4, 2, 5	Simple Substitution
U	5	<0.3	Simple Substitution
Cr	50	<0.2, 0.3, 1, 5, 6	ROS
Fe	67	<0.006, 0.02, 0.03, 0.04, 0.1, 0.2	ROS
F ⁻	29	<1, 1.5, 2, 3, 4, 8	ROS
Th	32	<0.1, 0.2,	ROS

The regression equations are fitted to the detected observations on the probability plot, and values of individual censored observations are predicted from the regression models based on their normal scores. This method is available in R software and can be used for any data set with a maximum of 80 percent values censored (Lee and Helsén, 2005).

Also Empirical Cumulative Distribution Function (ECDF) and Cumulative Probability (CP) (Filzmoser, 2005) have been plotted for the analytical parameters (variables) and hence, univariate outlier values were defined for each. ECDF plot shows the variables along the x- axis and the probability of the empirical cumulative distribution function along y-axis. One of the main advantages of this function is that all single data points are visible, and any unusual high or low value can be identified. For calculating the outliers boundary, classic (mean±2 standard deviation) and robust (median± Mahalanobis distance) formula can be applied to detect the univariate outliers. Furthermore, combination of CP plot with the ECDF can define the best result (Reimann et al., 2009). By plotting CP, the direct visual estimation of the median or any other value is possible. These functions are available in the package StatDa by Filzmoser (2010) in R software environment. The defined values were replaced by nearest smaller neighbor value in each variable to reduce the influence of outliers.

Subsequent to the univariate methods described above, multivariate methods (multivariate outlier detection, cluster analyses and cross- correlation) were used.

3.2.2 Multivariate outlier detection

The main aim of using multivariate outlier detection is defining the observations (samples) which have different structures (Filzmoser, 2005) e.g. various contamination pattern that are caused by different factors. In this case the outliers are not only the data that are very high or low in relation with the other data, but their shape and domain are different. Generally these observations are resulting from secondary geological processes, mineralization or contamination; and do not have higher or lower concentration in one or more variable (analytical parameters). Hence, multivariate outliers are not necessarily univariate outliers and vice versa (Reimann et al., 2009). In this study, multivariate outliers were detected in order to investigate the samples that are influenced by different multivariate factors for example surrounded by different glacial sediments or maybe are influenced by an unknown factor that should be studied in detail later. These data might exhibit different contamination pattern. Investigating various contamination patterns is helpful to know more about the factors that influence the contamination and its pattern.

To reach to this aim, the function “symbol.plot” from “mvoutlier” package by Moritz and Filzmoser (2009) in R environment was used in this study to define multivariate outliers. This function is based on robust Mahalanobis distance (MD) (Eq. 1), minimum covariance determinant (MCD) estimator and adjusted quantile (for more detail about these techniques and formula the reader is referred to Filzmoser (2005).

$$MD = \sqrt{(x - \bar{x})^T S^{-1} (x - \bar{x})} \quad (1)$$

where, MD(x) is Mahalanobis Distance, \bar{x} is the mean, x is multivariate vector, S is covariance matrix, and T is transmitt matrix.

The outcome is a numerical matrix that reports the MD of the observations and a two dimensional diagram which shows the ellipsoids corresponding to the 25%, 50%, 75% and adjusted quantile of the chi-squared distribution. Observations with different MD levels are shown by different symbols. The symbols that are plotted outside of the adjusted ellipsoid are potentially multivariate. To run the function, different variables must be selected as the entry data set. Selecting the variable is flexible and is depending on the goal of study. Since the aim of this study is focused on REE distribution and its relation with other analytical parameters, three variables that were considered were: a) REE (La to Lu), b) other metals (Al, Cu, Fe, Zn), and c) pH. This selection is mainly based on the result of the previous studies and also the concentration of the metals (Liang et al., 2008; Rönnback et al., 2008; Steinmann et al., 2008; Ding et al., 2007; Quinn et al., 2004; Wood, 1990; Lewis et al., 1998; Elderfield et al., 1981).

3.2.3 Cluster analysis

As mentioned above, hierarchical cluster analysis was used in order to find and making visible structures within observations and variables. In order to find structures in a data set or to reveal similarities of observations, a similarity measurement is needed (Einax et al., 1997). In this study Squared Euclidean distance (Eq. 2) was used to measure the distance between the parameters. After selecting the measurement, linkage method must be selected. Ward linkage was used for this data (Einax et al., 1997; Derde and Massart, 1982). One important aspect of preprocessing is to ensure the comparability of the variables by avoiding different measurement units for the variables. Hence, Z-standardization (Eq. 3) was performed on the data to avoid the influence of data scale (Davis, 2002).

$$SED = \sqrt{(x_1 - y_2)^2 + (y_1 - y_2)^2} \quad (2)$$

where, x, y are different observations in the data matrix.

$$Z = (x - \bar{x}) / s \quad (3)$$

Where, x is a raw score to be standardized, \bar{x} is the mean of the population, and s is the standard deviation of the population.

The principal aim of this technique is to partition multivariate observations into a number of meaningful, multivariate homogeneous groups. The analytical parameters placed in one main or sub cluster can be a clue to investigate the significant factors. Hence, to find the similarities between REE and other analytical parameters, R-mode cluster analysis was used. The Ward distance of 25 is the clustering criterion.

To investigate more about the geological and seasonal factors, Q-mode cluster analysis was used as well as multivariate outlier detection method. By applying this method, the similarity between the samples was defined based on the location; and also compared in different sampling seasons. As the analytical results are based on the concentration, each group represents the samples that have similar concentration of the analytical result.

3.2.4 Cross- correlation

The hydrograph of the boreholes based on the groundwater level fluctuations from 2004 to 2008 were plotted using Matlab (Version 7.6) to study the general hydrographs. Generally, groundwater levels fluctuate according to the characteristics of precipitation events such as amount, duration and intensity. Some other factors that can influence the groundwater level are topography and local geology. Groundwater level fluctuation in response to precipitation events can be used to estimate the factors that control heavy metal concentration in groundwater, such as dilution. To reach this aim, cross-correlation of groundwater levels monitoring data and daily precipitation was calculated. Cross-correlation is most appropriately used to compare two series that may have a temporal dependency between them (Eq. 4).

$$r_m = \text{Cov}_{1,2} / s_1 s_2 \quad (4)$$

where, $\text{cov}_{1,2}$ is the covariance between the overlapped portions of sequence 1 and 2, and s_1 and s_2 are the corresponding standard deviations.

This analysis will provide a correlation between two series or two waveforms. The aim of correlation analysis is to compare one or more functions and to calculate their relation with respect to a change of lag in time or distance (Einax et al., 1997). It means the observation of one series is correlated with another series at various lags and leads. Cross-correlation helps to identify variables which are leading indicators of other variables or how much one variable is predicted to change in relation to the other variables (Davis, 2002). Significant correlation at 95% confidence level is a criterion to test the hypothesis. As the two data series must be in a same size, the daily average of groundwater fluctuation was used. By using this technique, time lag and correlation coefficient are recognized. Time lag is an important factor for such topics (Hölting and Coldewez, 2005). The time lag occurs because rainfall needs time to reach to the groundwater and can be determined by cross- correlation technique. The cross-correlation function in Matlab R 2008a was used for this purpose.

4 Results and discussion

4.1 General hydrogeochemistry

The groundwater sample at the test site are in acidic range (pH: 3.2-5.4) and oxic (Eh: 340-715 mV) (Table 2). The general water type is Mg-(Ca)-SO₄. Mg occurring as main cation in most measuring points is in the range of 58 and 3660 mg/l, followed by Ca with values from 77 to 672 mg/l. Highest metal concentrations were found for Mn (166 ± 153 mg/l) > Al (42.1 ± 50.9 mg/l) > Ni (15.7 ± 1.1 mg/l) > Zn (4.0 ± 3.2 mg/l) > Co (3.9 ± 3.3 mg/l) (Tab.2). Most of the Fe concentration is low, often 67% below detection limit. The maximum Fe concentration was 180 mg/l in GTF25. Uranium which was leached during mining operations is in the range of 0.2 and 3411 µg/l. The highest concentration of REE was measured in GTF16 (8148.4 µg/l, December 2006), and the lowest concentration in GTF7 (11.4 µg/l, September 2007). Sulfate that occurs as main anion due to pyrite oxidation in the host rock and use of sulfuric acid during leaching operations varies between 422 and 16,758 mg/l, followed by Cl⁻ (2.3-575 mg/l). HCO₃⁻ was compared to sulfate only present in low concentrations especially in measuring points in the southern area, ranging generally between 0.7 and 53 mg/l. The dissolved organic carbon (DOC) ranges between 0.8 and 20.6 mg/l.

The standard deviation, kurtosis and skewness were also calculated to investigate the distribution of the measured parameter. As it is shown in the Table 2, measured parameters have non-normal distribution.

4.2 Elements distribution

At the test site, element concentrations are lower in the southern part (GTF7, 8, 20, 21, 22) and increase to the north what is about the groundwater flow direction (south-west to north-east). For example, along the north-south profile shown in Fig. 1, the mean concentration of Mn is increasing from GTF22 (26 ± 13 mg/l) to GTF8 (52 ± 10 mg/l), GTF18 (114 ± 21 mg/l), GTF5 (253 ± 142 mg/l) to GTF13 (509 ± 98 mg/l). GTF25 (156 ± 64 mg/l) and GTF26 (43 ± 43 mg/l) are lower concentrated than the central part of the test site. This trend is quite similar e.g. for Co, Mg, or Zn, and also for Cl⁻ or SO₄²⁻. The distribution of Al, Cu, Fe, REE, Th, and U in studied area is more heterogeneous with only a slight gradient from south to north. For example, for total REE along the profile (Fig.1), the mean concentration increases from GTF22 (44 ± 8 µg/l) to GTF8 (99 ± 19 µg/l) to GTF18 (849 ± 48 µg/l). Then, it varies more strongly in GTF5 (646 ± 652 µg/l) and GTF13 (343 ± 293 µg/l) and is higher again in GTF25 (2369 ± 1101 µg/l) and GTF26 (1055 ± 673 µg/l). Generally the highest concentrations of these elements were measured in GTF11, 16, 17, 25 and 32. These measuring points are mostly located in the central part of the test site; GTF25 rather in the north. The general distribution in seasonal data sets is similar.

4.3 REE and related analytical parameters

Generally, the analytical parameters are in two main clusters. The dendrograms of all data sets are very similar and thus only the dendrogram of the spring data set is shown as an example (Fig. 2). It is important to note that the distribution and speciation of heavy metals in complex system is influenced by interactive and competing factors; and describing the current situation accurately, needs more detail geochemical investigation which is not the aim of this study.

Table 2 Analytical parameters and descriptive analysis of groundwater samples of the test site; N=174

Parameter	Unit	Mean	Standard Deviation	Skewness	Kurtosis	Min.	Max.
Al	mg/l	42.1	50.9	2.3	5.9	0.2	308.4
Ba	µg/l	28.2	10.7	0.9	0.5	13.2	70.3
Ca	mg/l	4,367	60	-0.8	8.7	77.2	672.4
Cd	µg/l	97.8	80.4	2.0	7.1	3.9	601
Ce	µg/l	411	560	3.1	13.3	3.9	4,127.2
Co	µg/l	3,943	3,341	1.6	3.3	85	20,120
Cu	µg/l	453	64	2.9	11.9	0.4	4,562
Dy	µg/l	85	96	2.3	6.5	0.9	565.8
Er	µg/l	48	54	2.2	5.9	0.57	310.9
Eu	µg/l	14.1	16.4	2.3	6.2	0.1	92.5
Fe	mg/l	6.1	23.8	5.7	33.8	0.05	180.0
Gd	µg/l	83	97	2.3	6.1	0.9	547.7
Ho	µg/l	17.7	19.9	2.3	6.1	0.2	115.8
La	µg/l	53	57	2.9	13.9	1	423
Li	µg/l	0.4	0.3	1.9	5.3	0.8	1.6
Lu	µg/l	5.4	6.4	2.2	5.4	0.1	34.6
Mg	mg/l	1131	746	0.9	0.6	57.6	3,660.4
Mn	mg/l	166	153	1.4	1.3	5.19	704.6
Na	mg/l	21.2	13.2	0.4	-1.2	1.5	50.9
Nd	µg/l	148	167	2.5	8.1	1.8	1,073.3
Ni	µg/l	15,728	1,108	1.2	1.6	957	56,260
Pb	µg/l	4.1	6.2	2.6	7.2	0.1	33.2
Pr	µg/l	28.2	32.2	2.6	9.7	0.04	220.41
Sc	µg/l	7.8	6.8	2.4	9.5	0.4	47.2
Sm	µg/l	48	56	2.34	6.42	0.05	315.83
Sr	mg/l	0.63	0.49	0.14	-0.11	0.04	2.44
Tb	µg/l	13.84	15.61	2.33	6.31	0.04	91.22
Th	µg/l	1.24	2.43	5.88	43.72	0.03	21.41
Ti	µg/l	4.9	3.1	0.2	-0.5	0.2	12.6
Tm	µg/l	6.0	6.9	2.2	5.5	0.1	38.8
U	µg/l	290	532	2.9	10.2	0.2	3,411.1
Y	µg/l	611.1	673	2.3	6.9	8.2	4223
Yb	µg/l	35.3	41	2.2	5.4	0.4	226.4
Zn	mg/l	4.03	3.23	1.6	2.2	0.4	15.28
Cl ⁻	mg/l	178	157	0.7	-0.8	2.3	575.1
F ⁻	mg/l	12.5	12.14	2.3	7.12	0.6	83.3
SO ₄ ²⁻	mg/l	5,675	2,070	1.1	1.1	421.7	16,757.6
pH	-	4.2	0.5	0.5	-0.3	3.2	5.4
PO ₄ ³⁻	mg/l	0.2	0.1	2.6	7.2	0.1	0.8
HCO ₃ ⁻	mg/l	9.0	1.3	3.03	9.1	0.7	53.2
DOC	mg/l	3.7	2.6	2.7	11.4	0.8	20.6

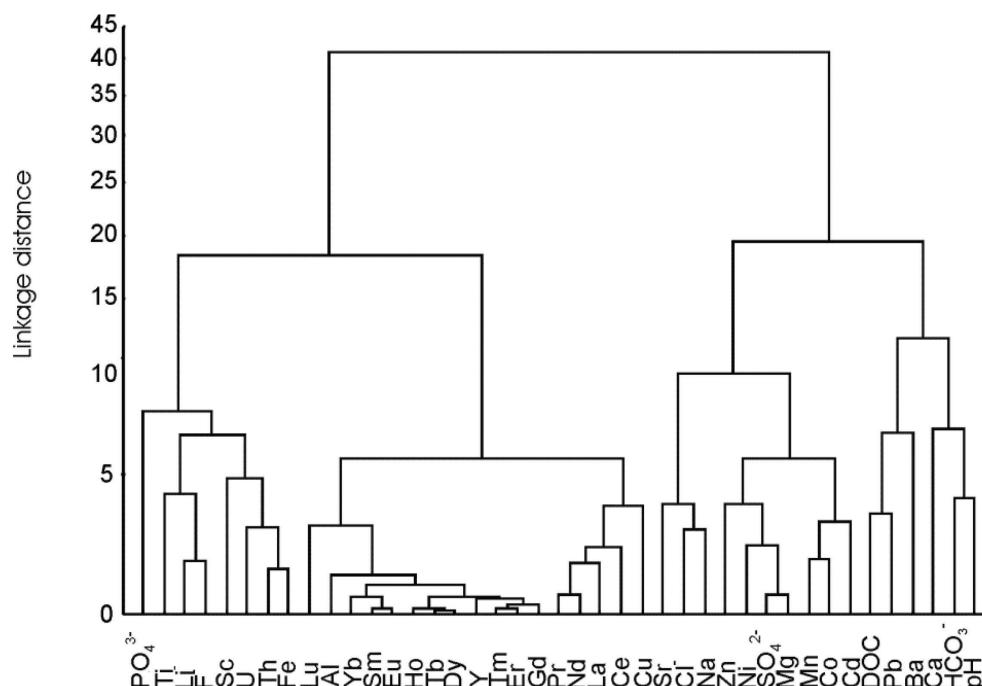


Fig. 2 Result of the R-mode cluster analysis showing the similarities between the analytical parameters of the groundwater samples in the spring data set. The analytical parameters are divided into two main clusters: REE and similar parameters; pH and similar parameters

REE and Al, Cu, F^- , Fe, Li, PO_4^{3-} , Sc, Th, Ti, U and Y are in one main cluster. The relation between the heavy metals Al and Fe and REE is probably more indirectly coupled with pH. Both Al and Fe (hydr)-oxide are good bounds for metals including REE. While Al precipitates, REE can be scavenged and hence, co-precipitated. In this process, HREE preferentially bind to Al-hydroxides compare to LREE (Quinn et al., 2004). The REE behavior towards Fe-hydroxides was described previously by Aström (2001). Verplanck et al. (2004) stated that Fe-hydroxides do not partition REE below a pH of 5.1. Liang et al. (2008) adjusted this boundary pH downwards to 4. In the scatter plot for the groundwater samples in this study (Fig.3), it becomes apparent that the boundary pH for REE mobility at the test site is about 4.5. Y shows similarity with REE in the studied area and this can be traced back to their similar chemical characteristics (Brookins, 1989).

In acidic environment F^- and PO_4^{3-} can complex HREE as well as Al and Fe (Aström, 2001). Lei et al. (2008) described the affinity of HREE to Al (hydroxides), whereas LREE have greater affinity to Fe (hydroxides). However, in this study both Al and Fe show correlation with HREE rather than LREE. Due to their similar ion radius, Li and Fe have a potential for ion exchange.

The frequent association of Cu with organic colloids in groundwater was shown by several studies (Sanudo-Wilhelmy et al., 2002; Jensen et al., 1999; Freedman et al., 1996; Pauwels, et al., 2002). At the pH values below 5 and the current Eh range, uranium can be formed as UO_2^{2+} . Above pH value 5, $UO_2 [OH]$ are formed (GSJ, 2005). UO_2 forms highly soluble hydroxide and carbonate complexes of varying stoichiometry as a function of pH and the partial pressure of CO_2 .

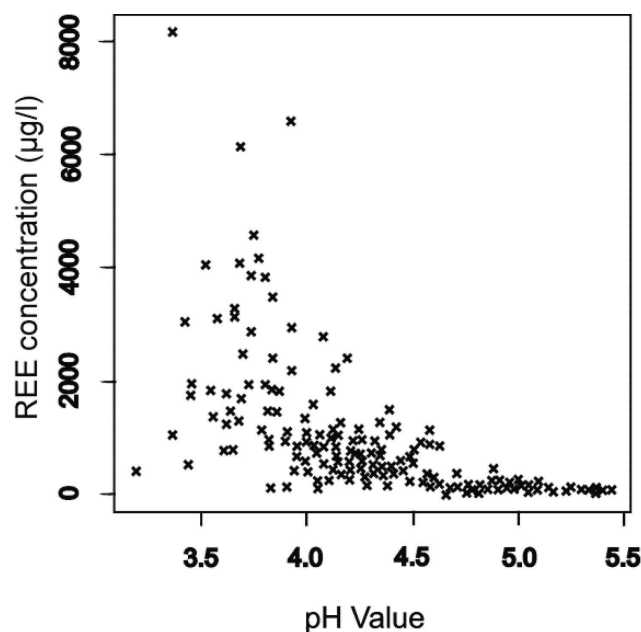


Fig. 3 Scatter plot showing REE concentration ($\mu\text{g/l}$) versus pH value of groundwater samples. REE has inverse relation with pH value. Below the value of 4.5, the REE precipitate. N=174

At low pH, adsorption of UO_2^{2+} is generally negligible and increases with increasing pH usually in the pH range of 4 to 6. Uranium-phosphate complex at the pH range of 4 to 9 was shown by Bruno and Casas, (1991). At pH 2 and higher, U(IV) is present in the precipitated form, i.e. it is immobilized and is not migrating as a dissolved species in the environment (Arnold et al., 2010), whereas U(VI) is much more soluble and may migrate as aqueous species in the environment. Sulfate is also able to form strong complexes with uranium and thereby strongly affects its speciation and migration (Hennig et al., 2007).

The second main cluster consists of Ca, Cd, Co, Cl^- , DOC, HCO_3^- , Mg, Mn, Ni, Pb, pH, Zn, and SO_4^{2-} .

Ion SO_4^{2-} is an important ligand for metal transportation in solutions. Manganese, Ni, Co are influenced by SO_4^{2-} (Grawunder, 2010). Since there is a wide range of Eh in the studied area, Mn can occur either as Mn^{2+} or MnO_2 . Activities of dissolved metals decrease as pH increases and metals are removed by precipitation. At high Eh, MnO_2 is formed and, hence the Mn concentration decreases in groundwater. At the contrary, at low Eh, Mn^{2+} ions are formed. As Carlsson and Büchel (2005) described the studied area, Mn^{2+} can form complex with SO_4^{2-} . It is also known that pH and HCO_3^- are closely related through the chemical equilibrium of CO_2 in water, and Co can co-precipitate with Mn hydroxides (Murry and Dillard, 1979). At the test site, Mn-oxides were already mentioned by Burkhardt et al., (2009) and the affinity of Co towards Mn-oxide was found by sequential extraction by Grawunder et al., (2009). Furthermore, Burkhardt et al., (2009) stated that a black Mn-oxide rich layer formed directly in the range of groundwater level fluctuation. Hence, precipitation could cause a decrease in the elements concentration in groundwater. In the current situation of the test site, pH is low enough to induce Mn-oxide dissolution and increases the Mn and Co concentrations in groundwater. This process is also valid for other heavy metals as Ni or Zn.

In the current situation Ni and Cd are present as Ni^{2+} and Cd^{2+} ; at such a low pH, their concentration got increased in groundwater. At pH values below 5.4, the possibility of Ni adsorption is

low; above this pH value and in the presence of dissolved organic carbon, Ni can be adsorbed on the surface of goethite (the presence of goethite in the study area was shown by Carlson and Büchel, (2005)). The criterion pH value for such situation for Cd is 6.5 (Weirich et. al, 2000). It is also important to note that in presence of sulfate, Cd adsorption by goethite is high. Sulfate promotes the adsorption of Cd depending on sulfate concentration and ionic strength, respectively. Ion Cd^+ adsorption is an inverse function of pH (Hoins et al., 1993).

Zinc usually occurs as Zn^{2+} and may precipitate as $\text{Zn}(\text{OH})_2$, ZnCO_3 or ZnS . Zinc is one of the most mobile heavy metals in surface waters and groundwater because it is present as soluble compounds at neutral and acidic pH values. Zinc readily precipitates under reducing conditions and in highly polluted systems when it is present at very high concentrations, and may co-precipitate with hydrous oxides of Fe or Mn (Smith et al., 1995). Sorption to sediments or suspended solids, including Fe- hydrous and Mn- oxides, clay minerals, and organic matter, is the primary fate of Zn in aquatic environments. Sorption of Zn increases as pH increases and salinity decreases.

Lead is known to form stable aqueous complexes with Cl^- , OH^- and SO_4^{2-} . Lead sulfate complexes (PbSO_4 and $(\text{Pb}(\text{SO}_4)_2)^{2-}$), and Pb-chloride complexes (PbCl^+ , PbCl_2), are typically considered in aqueous speciation modeling efforts (Hem and Durum, 1973; Hem, 1976; Marani et al., 1995; Pierrard et al., 2002; Rozan et al., 2003). Dissolved organic carbon may also form stable complexes with Pb and play an important role in governing Pb mobility in groundwater systems. At low pH and oxidizing conditions (current situation), Pb-sulfate is stable. Lead is usually not a metal of concern at mining-related sites where AMD is produced. This is because the weathering of metal-sulfides, in addition, to generating acidity also produces high concentrations of sulfate, which results in the precipitation of anglesite (PbSO_4) (Zanker et al., 2002). Lead adsorbs onto clay minerals, and poorly ordered Fe- and Al-containing hydroxypolymer coatings on natural aquifer sediments (Sposito, 1984; Coston et al., 1995; O'Reilly and Hochella, 2003). In the pH and Eh range of the study area, Mg presents as Mg^{2+} . Carlson and Büchel (2005) have shown that dolomite is the source of Mg^{2+} at this location. Hence, the expected reaction can be: $\text{CaMg}(\text{CO}_3)_2 + \text{H}_2\text{O} \leftrightarrow \text{Ca}^{2+} + \text{Mg}^{2+} + 2\text{HCO}_3^- + 2\text{OH}^-$.

4.4 Data domain

Using Q-mode clustering, the samples were mainly divided into two main clusters. GTF 3, 7, 8, 9, 10, 13, 15, 18, 19, 20, 21, 23, 24 and 26 are the members of the first main cluster and GTF 2, 4, 5, 6, 11, 12, 16, 17 and 25 are located in the second main cluster (Fig. 4). The members of the first cluster have lower concentration of REE and the members of the second cluster have higher REE concentration and mainly are located in the central to the northern of the test site. The similarity between the outcomes of different data sets is an indication of no seasonal effect in the studies area.

Moreover, using multivariate outlier detection methods, two groups of samples were defined. GTF7, 8, 9, 19, 20, and 21 are defined as the first sampling domain and GTF2, 4, 5, 6, 12 and 13 are defined as second domain. This identification is based on the plotted diagrams (Fig. 5 as an example shows the REE~ pH) as described in section 3.2.2. These two domains represent the contamination patterns that are influenced by different types of sediments that cover the sampling location. The first domain is located at the south of the test site, which is dominated by sand. The second domain is located mainly in the centre of the test site within the unit of silt and clayey silt/ varved clay with the facial interlocking with the northern unit. The samples were not defined based on the sampling

season. For example, all the samples collected from sampling point GTF7, were identified in one domain. Hence, no specific seasonal effect is evident.

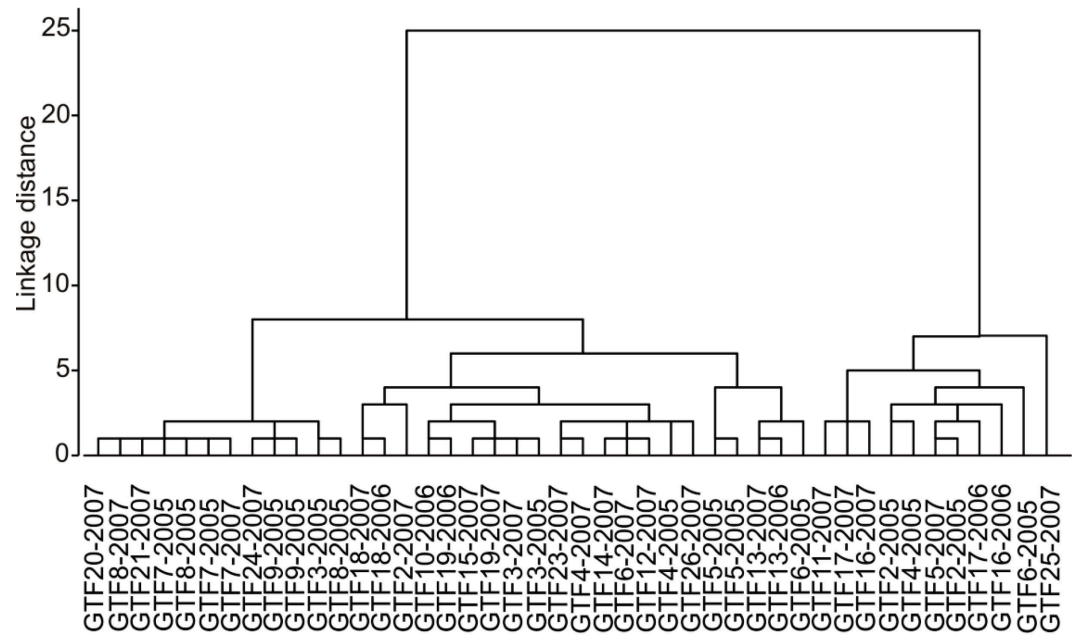


Fig. 4 Results of the Q- mode cluster analysis show similar groundwater samples in the spring data set. The samples are divided into two main clusters, and clustered by sampling location rather than sampling time

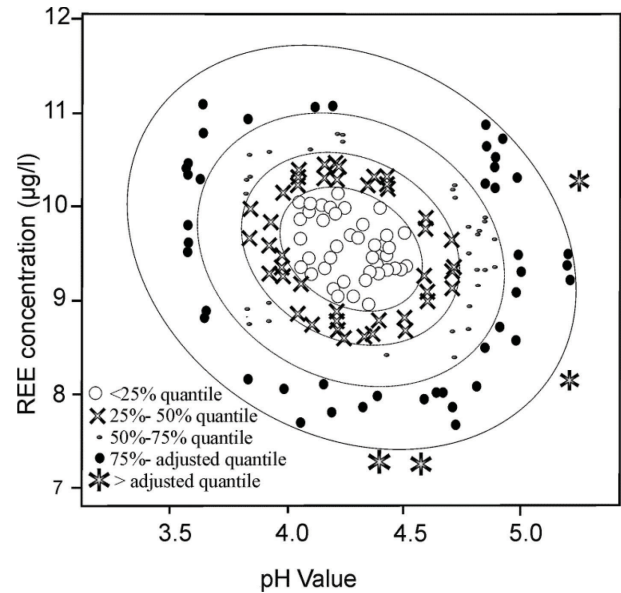


Fig. 5 Scatterplot showing REE versus pH value of groundwater samples, used for multivariate outlier detecting with ellipses corresponding to the 25%, 50%, 75% and adjusted quintile. The samples out of the quantile ellipsoid are known as multivariate outliers

The lower concentration in the south can be a result of a stronger bonding of metals in this area. Iron and Mn-hydroxides are good bounds for metals due to sorption or (co)precipitation (Liang et al., 2008; Ulrich et al., 2006; Coppin et al., 2002; Ohta and Kawabe, 2001; Spark et al., 2001). The occurrence of different hydr-oxides and clay minerals was described in the working area previously (Burkhardt et al., 2009; Grawunder et al., 2009; Carlsson and Büchel, 2005). After Coppin et al.,

(2002) especially HREE sorbs preferentially on smectite and kaolinite clay minerals. Bau (1999) reported preferential Ce enrichment at low pH which is also known to occur for Mn-hydr-oxides (Ohta and Kawabe, 2001). Sorption to or co-precipitation on Al- or/and Fe-hydr-oxides is a pH-dependent process as well. A decrease in pH would lead to dissolving of such minerals but also desorption of metals from hydr-oxides and clay minerals, resulting on a higher element concentration in groundwater. The relationship between the REE concentration and electrical conductivity (EC) were also studied. The EC of water is commonly used to reveal dilution effects of precipitation on subsurface water. Comparing the EC range and the REE concentration shows that increases in the EC correlate with increases in the REE concentration.

4.5 Groundwater level fluctuation

The REE and other metals distribution, and also the outcome of statistical studies show an increasing concentration of the REE from south to north. Furthermore, a broader statistical study could not show any probable seasonal effect. Since, investigating the seasonal effect is important, a more focus investigation using groundwater level fluctuation and its response to precipitation was studied.

As it mentioned in section 3.1., the groundwater level fluctuations were monitored every 15 minutes and the hydrographs of the measured wells are available. The general hydrographs were compared regarding the season and also precipitation range from 2006 to 2008. It is interesting to note that the hydrographs can be divided into two main groups, although there are some differences in details. Similar to the results that got by statistics, one group consists of the sampling points of the south of the test site (GTF7, 8, 9, 20, 21, 22, 23, 24, 32, 33 and 34) and the second group is including the samples of central to the north of the studied area (GTF2, 3, 4, 5, 6, 11, 12, 13, 14, 15, 16, 17, 18, 19, 27 and 30). Regardless the minor fluctuation, the hydrographs of the first group show a general increasing in the period of December 2007 to August 2008, and a sharp decreasing in October 2008. However, the second group doesn't have such a significant increasing in this period. Figure 6 shows the hydrograph of GTF25 and the precipitation graph of the same time period.

As it was mentioned above, there are two general hydrograph types. However, some minor differences are noticeable. These small differences between the hydrographs of each group are more apparent during the period of the highest groundwater level fluctuations (June-October 2007). The maximum variation of precipitation (0-38 mm/day) was reported for this duration and can be noted on the precipitation graph. To find the response of groundwater level fluctuation to the precipitation, cross-correlation between them was calculated. Time lag is another important factor that is recognizable by cross-correlation and is helpful to interpret the results. Moreover, to be able to interpret the resulting graphs, the significant correlation coefficient must be known which is depending on the data set size. Based on Dörfel 1990, a correlation coefficient bigger than 0.2, is significant for the data sets of this study. A 20-days period was considered for defining the time lag, since the ability of cores-correlation to define a large time lag is limited by a small data set (June-Oct. 2007). Although the general hydrographs were different for various data sets, the general graphs of time lags are similar. However, the time lags differ from one sampling point (borehole) to the other. The calculated coefficient also varies slightly between the sapling points; but, the average of 0.4 is defined for all of them. This coefficient shows a meaningful correlation for these data sets (Dörffel, 1990). Studying the time lags more in detail, shows that a longer time lag is resulted for the wells that contain higher REE

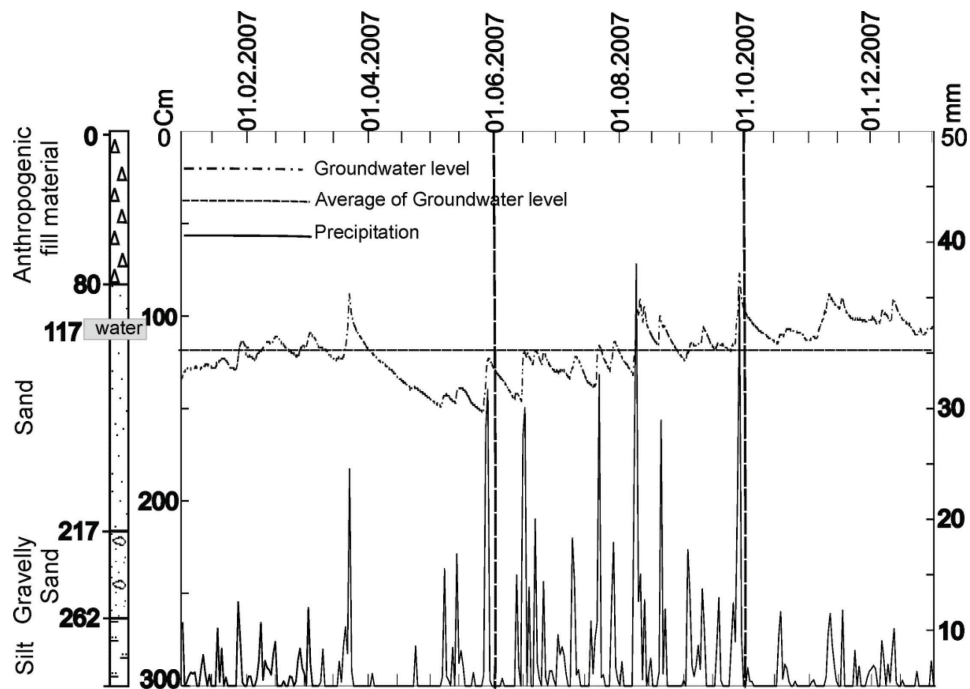


Fig. 6 Groundwater level and daily precipitation (2007) with the geological profile of GTF25; the highest precipitation is reported in a short period of the year (from June to October). The highest groundwater fluctuation occurred in this period too. The water bearing layer found during drilling operations is marked in the left. Signatures of sediments are the same as in Fig.1.

Concentration and are located in the silty part of the test site. For example, the time lag is about three days for GTF7, 8 and 3; and four days for GTF6, 15, 18 and 25. The borehole GTF13 has a time lag about one day. Figure 7 shows the cross-correlation of GTF25 as an example. It is interpreted that the groundwater level fluctuation is a response to precipitation but this response is more noticeable when a higher precipitation occurs. As known from the field work, the thin aquifers are running through the different geological units (Fig. 1). It implies the role of geological units on groundwater level fluctuation.

Furthermore, flow needs more time to react with the sediments to elute REE in the silty central part on the test site. As the precipitation data was reported daily, any shorter time lag cannot be recognizable by the cross-correlation. Comparing the local geological profiles surrounding the boreholes, shows that the local composition of the glacial sediments and hence its hydraulic parameter have an important role of REE concentration. The increasing REE concentration from the south to the north is more noticeable when groundwater level remains within a fine-grain layer, rather than remaining within a coarse-grain layer.

Furthermore, comparing the REE concentration and groundwater level fluctuation shows a slight inverse relation between them. It means that increasing the groundwater level fluctuation, decreases the REE concentration by dilution, although this in/decreasing is not big to affect on the REE distribution.

5 Conclusion

Based on the present study, generally the REE concentration in the study area is increasing from south to north. Although a general south-north gradient exists, but the REE concentration is not completely homogenate.

The pH and, sedimentology and stratigraphy of the study area are the important factors that affect

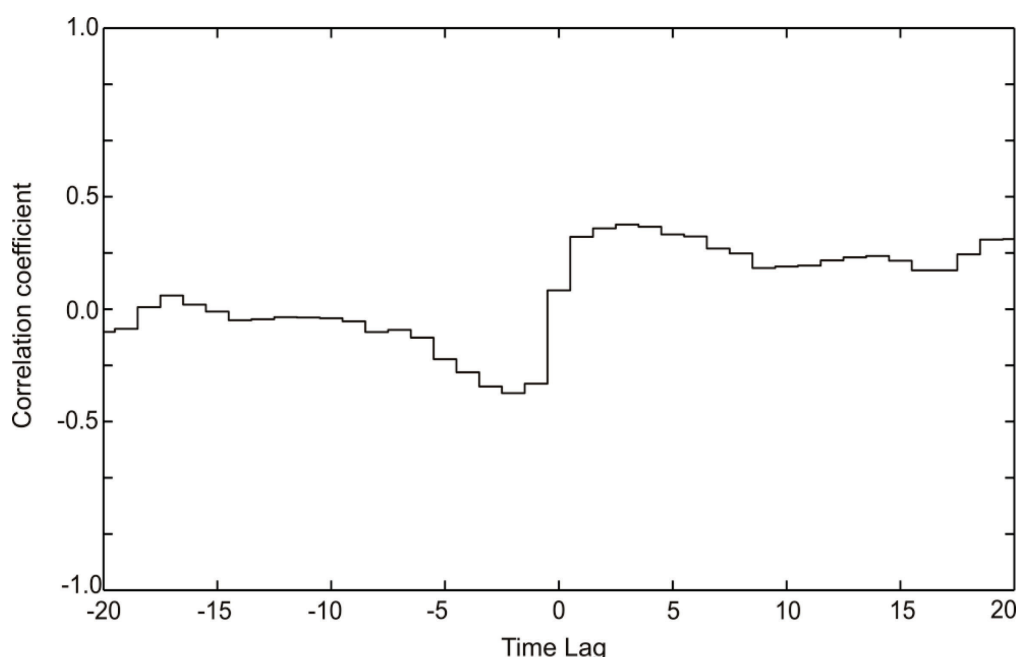


Fig. 7 Cross-correlation of groundwater level (daily average) and precipitation of a short period (June-October 2007) of GTF25; x axis is correlation coefficient and Y axis is time lag (day). The maximum of correlation coefficient in this period of time is 0.4. The time lag is the positive peak of the graph which shows a short time lag about 4 days

REE concentration and distribution in the studied area. The pH has an inverse correlation with REE concentration. The value of 4.5 can be considered as the boundary pH value in this study. The outcomes of statistical studies show that the samples in the central part and southern part belong to different data domain, which emphasis the role of local geology. The southern test site is dominated by sand, the middle and the northern test site by silt and silty sand. No specific seasonal effect was found by using statistics. Groundwater fluctuation shows a meaningful correlation to precipitation data in the period of increased precipitation. The time lag shows an average of three to four days response to precipitation. This variation affects on REE concentration after the time lag period very slightly, but it has no effect on general contamination distribution. By considering these results and also the geological profiles of the boreholes, we can come to a conclusion that the geological composition and also precipitation are important factors that affect REE distribution and concentration in the test site. However, the general contamination pattern is stable and is not affected by seasonal depended factors. Regarding this conditions, sampling in different seasons and different precipitation period is not critical for studying REE (contamination) distribution in the test site. For future sampling it is suggested to reduce sampling intervals (it can be reduced to two times per year), but rather collect samples in a fixed sampling interval. Since the general trend of contamination is known, the sampling locations can be reduced too. It is suggested to design a systematic sampling net/profile that covers all sampling domains (as described before) with three to four samples from each domain and also has overlap with other geological/ geochemical information (e.g. borehole profiling, taking different type of samples, monitoring groundwater fluctuation, etc.). In this condition, more (geo)statistical methods can be done and hence, more details can be investigated with less works. It will help the project to decrease the expenses and optimizing the project economically.

Acknowledgement

The authors would like to acknowledge Wismut GmbH, providing the precipitation data, also I. Kamp, G. Weinzierl and G. Rudolph. This project is a part of the graduate research training group, GRK 1257/1 (Alteration and element mobility at the microbe-mineral interfaces), and Jena School for Microbial Communication (JSMC) and financially supported by the German Research Society (DFG).

References

- Åström M (2001). The effect of acid soil leaching on trace element abundance in a medium-sized stream, w. Finland. *Appl. Geochem.* 16: 387–396.
- Bau M (1999). Scavenging of dissolved yttrium and rare earths by precipitating iron oxy-hydroxide: experimental evidence for Ce oxidation, Y-Ho fractionation, and lanthanide tetrad effect. *Geochim. Cosmochim. Acta.* 63(1): 67-77.
- Bozau E, Leblanc M, Seidel JL, Stark H J (2004). Light rare earth elements enrichment in an acidic mine lake (Lusatia, Germany). *Appl. Geochem.* 19 (3): 261-271.
- Brookins DG (1989). Aqueous geochemistry of rare earth elements. *Reviews in Miner and Geochem.* 21(1): 201-225.
- Bruno J, Casas I (1991). The kinetics of dissolution of UO_2 under reducing conditions and the influence of an oxidized surface layer: application of a continuous flow-through reactor. *Geochem. Cosmochim. AC.* 55: 647–658.
- Burkhardt E M, Meißner S, Merten D, Büchel G, Küsel K (2009). Heavy metal retention and microbial activities in geochemical barriers formed in glacial sediments subjacent to a former uranium mining leaching heap. *Chem. Erde. Geochem.* 69 (Supplement 2): 21-34.
- Carlsson E, Büchel G (2005). Screening of residual contamination at former uranium heap leaching site, Thuringia, Germany. *Chem. Erde. Geochem.* 65: 75-95.
- Coppin F, Berger G, Bauer A, Castet S, Loubet M (2002). Sorption of lanthanides on smectite and kaolinite. *Chem. Geol.* 182 (1): 57-68.
- Coston JA, Fuller CC, Davis JA (1995). Pb^{2+} and Zn^{2+} adsorption by a natural Al- and Fe-bearing surface coating on an aquifer sand. *Geochem. Cosmochim. AC.* 59(17): 3535-3547.
- Davis JC (2002). *Statistics and data analysis in geology.* John Wiley & Sons, New York, United States of America.
- Derde MP, Massart DL (1982). Extraction of information from large data sets by pattern recognition. *Fres. Z. Anal. Chem.* 313: 484-495.
- Dia A, Gruau G, Olivie-Lauquet G, Riou C, Molenat J, Curmi P (2000). The distribution of rare earth elements in groundwaters: assessing the role of source-rock composition, redox changes and colloidal particles. *Geochem. Cosmochim. AC.* 64(24): 4131-4151.
- Ding S, Liang T, Yan JC, Zhang ZL, Huang ZC, Xie YN (2007). Fractionations of rare earth elements in plants and their conceptive model. *Sci. China Ser. C.* 50: 47-55.
- Dörffel, Klaus, (1990). *Chemometrische strategien in der Analytik.* VCH, Leipzig, Germany (In German).
- Einax JW, Zwanziger HW, Geiß S (1997). *Chemometrics in Environmental Analysis.* VCH.
- Elbaz-Poulichet F, Dupuy C (1999). Behavior of rare earth elements at the freshwater-seawater interface of two acid mine rivers: the Tinto and Odiel (Andalucia, Spain). *Appl. Geochem.* 14 (8): 1063-1072.
- Elderfield H, Hawkesworth CJ, Greave MJ, Calvert S (1981). Rare earth element geochemistry of oceanic ferromanganese nodules and associated sediments. *Geochim. Cosmochim. AC.* 45: 513-528.
- Elderfield H, Upstill-Goddard R, Sholkovitz, ER (1990). The rare earth elements in rivers, estuaries, and coastal seas and their significance to the composition of ocean waters. *Geochim. Cosmochim. AC.* 54 (4): 971-991.
- Filzmoser P, Garrett RG, Reiman C (2005). Multivariate outlier detection in exploration geochemistry. *Comput. Geo. Sci.* 31(5): 579-587.

- Filzmoser P, Steiger B (2010). Package "StatDa", <http://www.statistik.tuwien.ac.at/StatDA/R-scripts/index.html>. Appeared on March 2008.
- Freedman YE, Ronen D, Long GL (1996). Determination of Cu and Cd content of groundwater colloids by sampling graphite furnace atomic absorption spectrometry. *Environ. Sci. Technol.* 30: 2270-2277.
- Gammons CH, Wood SA, Jonas JP, Madison JP (2003). Geochemistry of the rare-earth elements and uranium in the acidic Berkeley Pit lake, Butte, Montana. *Chem. Geo.* 198 (3-4): 269-288.
- Goldstein SJ, Jacobsen SB (1988). Rare earth elements in river waters. *Earth. Planet. Sci. Lett.* 89: 35-47.
- Grawunder A, Lonschinski, M, Merten D, Büchel G (2009). Distribution and bonding of residual contamination in glacial sediments at the former uranium mining leaching heap of Gessen/Thuringia, Germany. *Chem. Erde. Geochem.* 69: 5-19.
- Grawunder A (2010). Hydrogeochemistry of rare earth elements in an acid mine drainage influenced area. Doctoral dissertation, Friedrich Schiller university of Jena.
- Helsel DR, Cohn TA (1988). Estimation of descriptive statistics for multiply censored water quality data. *Water Res. Res.* 24: 1997-2004.
- Hem JD, Durum WH (1973). Solubility and occurrence of lead in surface water. *J. Amer. Wat. Works Assn.* 65:562-568.
- Hem JD (1976). Inorganic chemistry of Leas in water, USGS, Washington DC: US department of interior.
- Hoins U, Charlet L, Sticher H (1993). Ligand effect on the adsorption of heavy metals: the sulfate-cadmium-goethite case. *Water, Air and Soil Poll.* 68: 241–255.
- Hölting B, Coldewez WG (2005). Hydrogeologie. Elsevier. Munich, Germany.
- Jakubick A, Jenk U, Kahnt R (2002). Modeling of mine flooding and consequences in the mine hydrogeological environment: flooding of the Koenigstein mine, Germany. *Environ. Geol. Water. S.* 42 (2): 222-234.
- Jensen DL, Ledin A., Christensen TH (1999). Speciation of heavy metals in landfill-leachate polluted groundwater. *Water Res.* 33: 2642-2650.
- Kaiser H (1965). Zum Problem der Nachweigrenzer. Workshop: Moderne Methoden der anorganischen Analyse. Düsseldorf, Germany (In German).
- Kimoto A, Nearing MA, Zhang XC, Powell DM (2006). Applicability of rare earth element oxides as a sediment tracer for coarse-textured soils. *Catena.* 65: 214-221.
- Lange G (1995). Die Uranlagerstätte Ronneburg, *Zeitschr. Zeit. Geo. Wisse.* 23: 517-526 (In German).
- Lee L, Helsel D (2005). Statistical analysis of water-quality data containing multiple detection limits: S-language software for regression on order statistics. *Comput. Geo. Sci.* 31(10): 1241-1248.
- Lewis AJ, Komninou A, Yardley BWD, Palmer MR (1998). Rare earth element speciation in geothermal fluids from Yellowstone national park, Wyoming, USA. *Geochim. Cosmochim. Acta.* 62: 657- 663.
- Leybourne MI, Goodfellow WD, Boyle DR, Hall GM (2000). Rapid development of negative Ce anomalies in surface waters and contrasting REE patterns in groundwater associated with Zn-Pb massive sulphide deposits. *Appl. Geochem.* 15 (6): 695-723.
- Liang-qi L, Ci-an S, Xiang-li X, Yan-hong L (2008). REE behavior and effect factors in AMD-type acidic groundwater at sulfide tailings pond, BS nickel mine, W.A. T. *Nonferr. Metal Soc.* 18 (4): 955-961.
- Liang T, Zhang S, Wang L, Kung HT, Wang Y, Hu A, Ding S (2005). Environmental biogeochemical behaviors of rare earth elements in soil–plant systems. *Environ. Geochem. Hlth.* 27: 301-311.

- Lonschinski M (2009). Schwermetalle im System Boden-Wasser-Pflanze mit Hinblick auf Fraktionierungsprozesse der Seltenen Erden Elemente auf der Aufstandsfläche der Gessenhalde im ehemaligen Uranbergbaurevier Ronneburg (Ostturingen). Doctoral dissertation. Friedrich Schiller university of Jena (In German).
- Marani D, Macchi G, Pagano M (1995). Lead precipitation in the presence of sulphate and carbonate: testing of thermodynamic predictions. *Water Res.* 9:1085-1092.
- Merten D, Geletneky J, Bergmann H, Haferburg G, Kothe E, Büchel G (2005). Rare earth element patterns: a tool for understanding processes in remediation of acid mine drainage. *Chem. Erde. Geochem.* 65: 97-114.
- Miekeley N, Dejesus HC, Dasilveira CLP, Linsalata P, Morse R (1992). Rare-earth elements in groundwaters from the Osamu Utsumi Mine and Morro-Do-Ferro analog study sites, Pocos-De-Caldas, Brazil. *J. Geochem. Explor.* 45 (1-3): 365-387.
- Moritz G, Filzmoser P (2009). Package "mvoutlier for Multivariate Outlier Detection", <http://www.statistik.tuwien.ac.at/StatDA/R-scripts/index.html>.
- Murray JW, Dillard JG (1979). The oxidation of cobalt (II) adsorbed on manganese dioxide. *Geochim. Cosmochim. AC.* 43: 781-787.
- Ohta A, Kawabe I (2001). REE (III) adsorption onto Mn-dioxide (δ -MnO₂) and Fe-oxyhydroxide: Ce(III) oxidation by δ -MnO₂. *Geochim. Cosmochim. AC.* 65 (5): 695-703.
- O'Reilly SE, Hochella M F (2003). Lead sorption efficiencies of natural and synthetic Mn and Fe-oxides. *Geochem. Cosmochim. AC.* 67:4471–4487.
- Pauwels H, Tercier- Weber M, Arenas M (2002). Chemical characteristics of groundwater around two massive sulphide deposits in an area of previous mining contamination, Iberian Pyrite belt, Spain. *J. Geochem. Explor.* 75: 17-41.
- Pierrard JC, Rimbault J, Aplincourt M (2002). Experimental study and modeling of lead solubility as a function of pH in mixtures of ground waters and cement waters. *Water Res.* 36: 879–890.
- Protano G, Riccobono F (2002). High contents of rare earth elements (REEs) in stream waters of a Cu-Pb-Zn mining area. *Environ. Pollut.* 117(3): 499-514.
- Quinn KA, Byrne RH, Schijf J (2004). Comparative scavenging of yttrium and the rare earth elements in seawater: Competitive influences of solution and surface chemistry. *Aquat. Geochem.* 10: 59-80.
- Reimann C, Filzmoser P, Garrett RG, Dutter R (2009). *Statistical data analysis explained: applied environmental statistics with R.* John Wiley & Sons.
- Rozan TF, Luther III GW, Ridge D, Robinson R (2003). Determination of Pb complexation in oxic and sulfidic waters using pseudovoltammetry. *Environ. Sci. Technol.* 37: 3845-3852.
- Rönnback P, Åström M, Gustafsson JP (2008). Comparison of the behavior of rare earth elements in surface waters, overburden groundwater and bedrock groundwater in two granitoidic settings, Eastern Sweden. *Appl. Geochem.* 23(7): 1862-1880.
- Rüger F, Dietel W (1998). Vier Jahrzehnte Uranbergbau um Ronneburg. *Lapis.* 7(8): 14-18 (In German).
- Sanford RF, Pierson CT, Crovelli RA (1993). An objective replacement method for censored geochemical data. *Math. Geol.* 25 (1): 59-80.
- Sañudo-Wilhelmy S, Olsen KA, Scelfo JM, Foster T D, Flegal AR (2002). Trace metal distributions off the Antarctic peninsula in the Weddell sea. *Mar. Chem.* (77): 157-170.

- Sholkovitz ER (1993). The geochemistry of the rare earth elements in the Amazon river estuary. *Geochim. Cosmochim. Acta*. 57: 2181-2190.
- Sholkovitz ER, Landing WM, Lewis BL (1994). Ocean particle chemistry: the fractionation of rare earth elements between suspended particles and seawater. *Geochim. Cosmochim. Acta*. 58: 1567-1579.
- Smedley PL (1991). The geochemistry of rare earth elements in groundwater from the Carnmenellis area, southwest England. *Geochim. Cosmochim. Acta*. 55 (10): 2767-2779.
- Smith LA, Means JL, Chen A, Alleman B, Chapman CC, Tixier JS, Brauning SE, Gavaskar AR, Royer MD (1995). Remedial options for metals-contaminated sites. Lewis. Boca Raton. United States of America.
- Spark KM, Johnson BB, Wells J (1995). Characterizing heavy-metal adsorption on oxides and oxyhydroxides. *Eur. J. Soil. Sci.* 46 (4): 621-631.
- Sposito G (1984). *Surface Chemistry of Soils*. Oxford University press. Oxford, England.
- Steinmann M, Stille P (2008). Controls on transport and fractionation of the rare earth elements in stream water of a mixed basaltic-granitic catchment basin (Massif Central, France). *Chem. Geol.* 254: 1-18.
- Tang Jaj, Johannesson KH (2006). Controls on the geochemistry of rare earth elements along a groundwater flow path in the Carrizo Sand aquifer, Texas, USA. *Chem. Geol.* 225 (1-2): 156-171.
- Ulrich KU, Rossberg A, Foerstendorf H, Zänker H, Scheinost AC (2006). Molecular characterization of uranium (VI) sorption complexes on iron (III)-rich acid mine water colloids. *Geochim. Cosmochim. Acta*. 70 (22): 5469-5487.
- Verplanck PL, Antweiler RC, Nordstrom DK, Taylor H E (2001). Standard reference water samples for rare earth element determinations. *Appl. Geochem.* 16: 231-244.
- Verplanck PL, Nordstrom DK, Taylor HE, Kimball BA (2004). Rare earth element partitioning between hydrous ferric oxides and acid mine water during iron oxidation. *Appl. Geochem.* 19: 1339-1354.
- Weirich DB, Behra P, Sigg L (2000). Adsorption of Copper, Nickel, and Cadmium on Goethite in the presence of organic ligands. *Aquat. Geochem.* 9: 65-85.
- Wismut GmbH (1994a). Entwurf Sanierungskonzept Standort Ronneburg.-Stand März. Chemnitz, Germany (In German).
- Wood SA (1990). The aqueous geochemistry of the rare-earth elements and yttrium: review of available low-temperature data for inorganic complexes and the inorganic REE speciation of natural waters. *Chem. Geol.* 82: 159-186.
- Zänker H, Moll H, Richter W, Brendler V, Hennig C, Reich T, Kluge A, Huttig G (2002). The colloid chemistry of acid rock drainage solution from an abandoned Zn-Pb-Ag mine. *Applied Geochemistry*, 17: 633-648.
- Zhao F, Cong Z, Sun H, Ren D (2007). The geochemistry of rare earth elements (REE) in acid mine drainage from the Sitai coal mine, Shanxi province, North China. *Int. J. Coal. Geol.* 70 (1-3): 184-192.

Manuscript 2

**Top soil heavy metal and rare earth elements contamination at a remediation
former uranium mining site, eastern Thuringia, Germany**

Top soil heavy metals and rare earth elements contamination at a remediated former uranium mining site, eastern Thuringia, Germany

Anahita Pourjabbar¹ - Tsilla Boisselet¹ - Dirk Merten¹ - Jürgen .W. Einax² - Georg Büchel¹

¹ Institute of Geosciences, Friedrich Schiller University, Jena, Germany

² Institute of Inorganic and Analytical Chemistry, Friedrich Schiller University, Jena, Germany

Abstract

This paper deals with heavy metal contamination in groundwater and soil affected by former uranium mining activities in Eastern Thuringia, Germany. Acid mine drainage (AMD) contaminated the groundwater-soil system with high concentrations of uranium, Rare Earth Elements (REE) and other metals. Over time, other processes such as groundwater level fluctuations and capillary rise continued the metal exchange between soil and groundwater. The solubility and mobility of heavy metals in soil and groundwater plays a key role for prediction of their bioavailability and toxicity. Data of 190 groundwater samples were collected from the area to provide the characteristics of groundwater. Furthermore, 99 soil samples have been eluted with water and ammonium nitrate. 40 chemical and physico-chemical parameters including pH, and metals including Σ REE were analyzed in groundwater and leached-soil. The data were subjected to cluster and factor analyses. By using these two methods, the statistical relations between the analytical parameters that control the heavy metals distribution were determined. The REE patterns obtained through normalization to Post Archean Australian Shale (PAAS) were used as a tool to study the water-soil interaction.

With regards to the comparison of the REE patterns of the data sets, generally, the patterns are similar qualitatively. All samples exhibit middle rare earth elements (MREE)-enrichment and light rare earth elements (LREE)-depletion with respect to the heavy rare earth elements (HREE). Furthermore, all samples except of some water-leached samples have a positive Ce anomaly.

With regards to the statistical analyses, there are high similarities between REE and Al and Cu. Furthermore, the coherence between the experimental data and the natural samples including the correlation between the elements is shown. The groundwater samples are statistically more similar to soil samples leached with ammonium nitrate than to the soil samples leached with water. It indicates that ammonium nitrate is a better represent to simulate the contaminated groundwater rather than water.

Key words: Heavy Metals, REE, Statistical Analysis, Water-Soil System, Leachate, Groundwater-Soil-Water contamination

1 Introduction

The former uranium mine districted in Eastern Thuringia, Germany was the third-largest uranium producer in the world (Jakubick et al., 2002; Kahlert, 1992; Lange, 1995; Wismut GmbH, 1994). The mining activities caused changes in the hydrogeology of the area. The excavation activities introduced the rock to oxidizing conditions. Hence, under the influence of oxygen, rain water and bacterial reactions, sulphuric acid was produced. This sulphuric acid as well as pyrite oxidation led to high sulphate concentrations in the drainage water of the heaps. This acidic solution, which is enriched with heavy metals, is called acid mine drainage (AMD). Uranium and various other heavy metals became mobile under such conditions (Wismut GmbH, 1994; Geletneky, 2002).

Among several heaps in the area, Gessenhalde was the only leaching heap built up by Ordovician and Silurian shales with a low grade of ore mineralization ($< 300 \text{ U g/ton}$) (Rüger and Dietel, 1998). Between 1971 and 1978, the materials were leached with AMD and later with sulfuric acid (10 g/l) in order to extract the Uranium. The leach pad was sealed with 0.6 m of loam and was compacted in order to prevent infiltration. This seal was covered by a one-meter-thick layer of coarse waste rock containing low grade of uranium mineralization from Lichtenberg, as a drainage layer during the leaching process. Drainage hills were designed to transport the leaching solution to collection ponds. It is probable that these were not completely sealed, and some contamination infiltrated the under-lying soil. In 1989, leaching was stopped (Wismut GmbH, 1994). In the 1990s, the heap was removed and used to fill the nearby open pit, Lichtenberg. The area was left uncovered, which led to form puddles prior 10 m upper layer of the underlying Quaternary sediments were excavated, and a layer of allochthonic top soil was added in order to re-couture the area as a last remediation step. A few years later, in 2003 and 2004, the evidence of residual heavy metal contamination was measured in water and in that upper soil layer as reported by Carlsson and Büchel (2005). It showed that the remediation was not completed. Still, in many locations at the site it is visible that plants are affected by high metal concentration indicating that contamination is present in the root zone, the upper 30 cm of soil.

In 2004 the test site “Gessenwiese” was created in the northern part of the base area of the former leaching heap, Gessenhalde, with the aim of monitoring the groundwater and soil parameters and improving bio-remediation strategies (with the help of growing selected plants, soil amendments and adding microorganism to improve the plants growth) for low level heavy metals contamination (Büchel, et al., 2005; Neagoe et al., 2005). Different zones within the test site were amended to a depth that varied from 20 to 90 centimeters. Three plots were amended by addition of top soil, compost soil, and a non-amended plot was used as a control.

In this study, we want to investigate the source of the contamination, and the way contamination spread to the surface soil after short time. Since the added soil was allochthonic, the water could be the most probable source, although usually soil water and groundwater belong to different systems. Therefore, the heavy metal content and contamination patterns of surface soil samples and groundwater samples are compared.

The northern and the southern part of the test site are dominated by sand, while the middle part is dominated by fine-grained silty sediments. Between the depth of 0.70–1.60 m clayey silt/varved clay appears from glacial Quaternary sediments (Grawunder et al., 2009). The groundwater flow direction

is south–south-west to north–north-east. It is not possible to specify distinctive flow paths because the groundwater is flowing through very thin layers of gravel and sand.

The top soil is characterized by the presence of different heavy metals and REE (REE, La-Lu). The values up to $0.05\mu\text{g}$ water soluble $\Sigma\text{REE/g}$ soil are considered as normal level content for this site and will be described as 'normal' soil in this study. This corresponds mostly to ammonium nitrate leachate values of $6\mu\text{g/g}$ soil. During drilling for groundwater samples, Fe and Mn precipitation was observed above the groundwater level. Lonschinski (2009) describes the saturation of the soil at a depth of 30 cm to be at a high level during the wet season after the winter. Furthermore, over a period of three years, groundwater level fluctuation was monitored (Lonschinski, 2009). The groundwater level fluctuations occurred by recharge, cause changes in metals concentrations (due to dilution) and make it difficult to assess the performance of a remediation strategy. Rare earth element signatures, however, are more reliable indicators of changes in soil-water composition that are due to a remediation approach.

The aim of this study is to investigate the similarities of contamination patterns in groundwater samples and soil leachate to show the influence of groundwater on contaminated soil, and the estimation of leachate parameters of the soil in order to predict groundwater contamination.

Soil leaching with selected leaching agent is generally used as DEV S4 (DIN 38414-4(1984) or DIN EN 12457-4 (2003)) method to evaluate the contaminant release from a given soil. This was extended by Zeien and Brümmer (1989) to the sequential extraction which aimed to evaluate also further stronger binding forms and possible releases of metal contaminants.

2 Materials and methods

To achieve the goal of the study, all samples, including soil-leached samples and groundwater samples, were analyzed for metal concentrations and the data were processed with multivariate statistics such as cluster analyses and factor analysis. Furthermore, REE patterns were compared, as it is an established method for understanding geochemical processes. The pattern obtained through normalization of the REE with PAAS (McLennan, 1989) is a tool used to study water-rock-interactions, to determine the source of contamination (Merten et al., 2005). REE are usually separated into light REE (LREE; La-Nd or Pm), middle REE (MREE; Sm-Dy) and heavy REE (HREE; Ho-Lu).

2.1 Sampling and sample preparation

2.1.1 Groundwater samples

For groundwater sampling at the test site, 33 boreholes (labeled as GTF (1 to 34)) were installed in a grid pattern, each separated by a distance of about 10 m. In total, 175 groundwater samples were collected from December 2004 to September 2008. Eh, pH, electrical conductivity (EC), and temperature were measured on site using portable instruments (WTW, pH320; WTW, LF320; WTW, external thermocouple). Samples for anions (except for HCO_3^-) and cations were filtered using glass fiber pre-filters (Sartorius) and cellulose acetate filters ($0.45\mu\text{m}$, Sartorius). Water samples for cations analysis were acidified with HNO_3 (65%, subboiled) to $\text{pH} < 2$. All samples were kept at 6°C until analysis. The HCO_3^- analysis was performed on the same day of sampling by titration (Titrimo 716 DM, Metrohm), whereas Cl^- , F^- , SO_4^{2-} and PO_4^{3-} analysis was performed by ion chromatography (DX-120, Dionex) (Lonschinski, 2009).

2.1.2 Soil samples

Figure 1 shows the sampling location of groundwater and soil samples. The 80 soil samples (each weighing about 20-100 g) were collected from the test site, nine samples were collected from the amended soils and one sample was taken from the surroundings. All samples were obtained from a depth of about 20 cm (more explanation in Section 3.3, Fig. 8). A soil sampling zone, about 5 m x 4 m, was divided to 80 sampling points using a grid system: eight West-East grid lines that were labeled from A to H, and 10 North-South grid lines that were labeled from 1 to 10. The distance between the grid lines (and hence, sampling locations) is 50 cm.

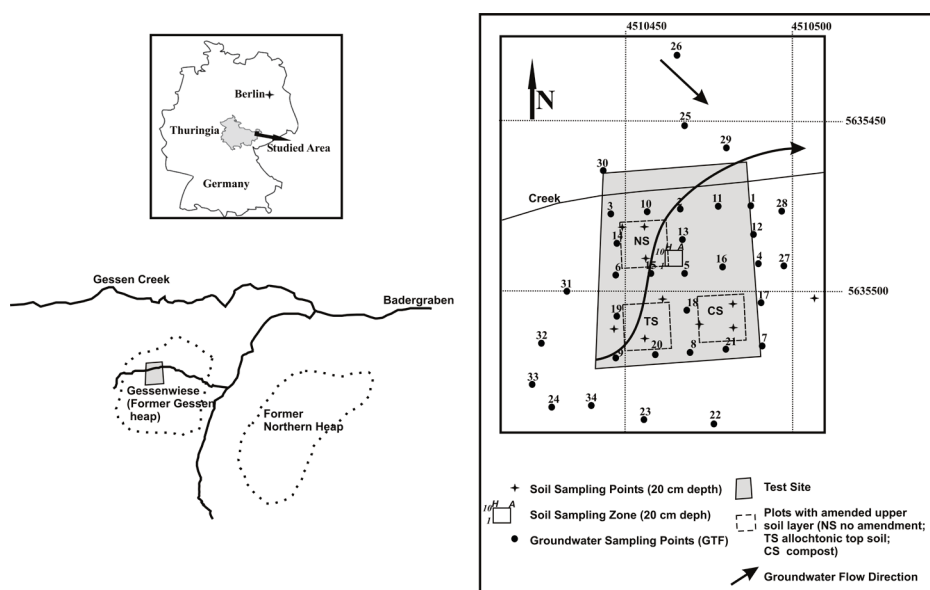


Fig. 1 Location map of the studied area. Sampling points for groundwater and soil samples at the former heap, Gessenhalde, in the former uranium mining area in Ronneburg, Germany

Furthermore, three zones within the test site were amended. These three amended-soils are categorized based on their different organic carbon content, and were marked as TS (addition of Top Soil), and CS (addition of Compost Soil), which had the highest organic content; NS (addition of Non-amended Soil) was the control soil, corresponding to the remaining surface soil samples. The thickness of the mentioned zones varies from 20 to 90 cm.

The leaching solutions consisted of water (Ultrapure water was obtained with a Purelab Plus system from USF Elga Seral (Ransbach-Baumbach, Germany)) and ammonium nitrate (1M, p.a., Roth).

Water was used as a leaching agent to remove the water soluble fraction, which represents the fraction that can be washed out by rain. Ammonium nitrate is used to leach the plant-available fraction of soluble metals, and it is useful in order to estimate the effect on flora caused by heavy metals. In general, other salt solutions such as 1M CaCl_2 could also be used to leach the soluble fraction. The reason for using salt solutions is that precipitation water becomes enriched with dissolved salts when it is in contact with soil. Hence, a salt solution can represent surface water. However, because our study is focusing on the effect of precipitation and the dissolution of salt, water was used as a leachate. Stronger leaching agents can be used, for example during the sequential extraction, whose properties are characterized by lower pH and higher complexation properties (such as ammonium

acetate, EDTA), but they are used to estimate the binding of metals (such as Mn oxides or humic acid fraction), which is not of relevance in this study.

The soil samples were dried in porcelain plates either at 40°C in the drying oven or at room temperature. The specimens were sieved to below 2 mm, and kept in a dry place inside a plastic container. The amount of 3-4 g of the material was placed into 50 mL polyethylene test tubes and the leaching solution was added to reach a solid-liquid ratio of 1:10. The suspension was tumbled for 24 h at about 20 rpm (Overhead shaker-ELU safety lock, Edmund). For each experiment, blanks containing only leaching solution were prepared. The samples were then centrifuged 15 min. at 2500 rpm; afterwards, 15 mL of each sample were filtered through a 0.45 µm-cellulose acetate filter. Then, aliquots of the filtered samples were acidified with HNO₃ (65%, suprapur, Merck) to a pH below 2 and kept at 6°C until analysis. Using the remaining supernatant, the pH and the electrical conductivity were measured using respectively pH meter pH197 (WTW), and LF320 (WTW). All samples were analyzed in triplicate, and the mean of the obtained values were used for statistical analysis.

2.1.3 Analysis

Aluminum, Ca, Fe, Mg, Mn, Na and Sr were analyzed with coupled plasma-optical emission spectrometry (ICP-OES (Spectroflame, Spectro)); Al, Ba, Cd, Co, Cu, Fe, Li, Ni, Pb, REE, Sc, Sr, Th, Ti, U, Y, and Zn were analyzed with Inductively Coupled Plasma–Mass Spectrometry ICP-MS (until 2007 PQ3-S, Thermo Elemental, then: X-Series II, Thermo Fisher Scientific). The anions Cl⁻, F⁻, Cl⁻ and SO₄²⁻ were analyzed using ion chromatography (DX120, Dionex), and PO₄³⁻ with photometry (Hach- DR/4000U). For determination of HCO₃⁻ an automatic titration system was used (Titrino 716 DM, Metrohm). Dissolved organic carbon (DOC) was measured on 0.45 µm filtered samples (DIMA-TOC 100, Dimatec until 2006; and then multi N/C 2100, Analytik Jena). To check the accuracy and precision of the analytical results the instrument drift is monitored and corrected by using Be, Ru and Re as internal standards. Each sample is measured three times. First, an outlier test is performed on each of the three runs (Grubbs test, 90% significance, criterion 1.15). Then, for the remaining runs, the mean and standard deviation are calculated. Furthermore, analytical quality is checked by the use of standard reference materials: the standard reference materials SCREE and PPREE are used for REEs (Verplanck et al., 2001). For other heavy metals, the standard reference materials SPS-SW2 (LGC standards) and NIST1643e (LGC standards) are used. Only values between 90-110 % for these reference materials are accepted. Detection limits (DL), Eq.1, are calculated according to the "3 sigma criterion" (Kaiser, 1965) as follow:

$$DL = Dil.Fact.* 3 * \sqrt{2s_0/s} \quad (1)$$

$$S = I / c$$

where, Dil. Fact. is dilution factor, s₀ is standard deviation, S is sensitivity, I is intensity, and c is concentration.

Table 1 presents the results of the analysis and analytical information. Gadolinium and Ce anomalies were calculated respectively according to Eq. 2 and Eq. 3 (Rabiet et al., 2009). These equations are used to measure the divergence or anomaly of an element to its expected value relative to the other REE.

where, N is the measured amount of the element normalized on PAAS, and * is expected value.

Table 1 Chemical analyses of groundwater samples and leachate resulting from ammonium nitrate- and water-leached soil samples; Symbol '-' denotes non measured analytical parameters

	Groundwater			NH ₄ NO ₃ leachate			Water leachate		
	Min.	Max.	Stdev.	Min.	Max.	Stdev.	Min.	Max.	Stdev.
Al	0.21	308.12	50.97	26.5	254.3	40.7	0.1	97.2	1.6
Ba	13.6	70.3	10.7	8.92	66.14	10.08	-	-	-
Ca	77.33	672.24	59.69	-	-	-	43.3	659.2	141.8
Cd	3.92	601.21	81.56	0.17	0.55	0.07	-	-	-
Ce	3.9	4127.2	566.4	1.51	6.12	0.80	0.003	0.1	0.05
Co	85.3	20120.2	3331.8	-	-	-	0.01	1.83	0.02
Cu	0.4	4562.2	631.51	0.37	7.35	1.38	0.01	0.82	0.09
Dy	1.01	566.04	96.04	0.24	1.23	0.15	0.003	0.021	0.003
Er	57.69	310.92	54.21	0.14	0.73	0.08	0.02	0.011	0.001
Eu	0.11	92.54	16.44	0.02	0.44	0.04	0.00	0.004	0.00
Fe	0.01	180.47	23.84	0.03	1.71	0.28	0.21	9.11	1.50
Gd	0.91	548.25	97.49	0.37	1.46	0.18	0.000	0.011	0.002
Ho	0.26	115.88	19.89	0.03	0.32	0.04	0.000	0.004	0.001
La	1.39	423.34	57.69	0.14	10.56	0.15	0.00	0.04	0.01
Li	0.81	1.64	0.25	-	-	-	0.01	0.41	0.06
Lu	0.01	34.21	6.44	0.07	0.29	0.02	0.00	0.001	0.00
Mg	57.61	3660.33	746.45	207.25	1051.06	161.15	26.23	677.61	138.62
Mn	51.2	705.1	153.5	46.2	278.7	57.9	0.22	174.54	31.48
Na	1.50	51.41	13.23	-	-	-	2.5	35.5	9.2
Nd	1.83	1073.24	169.11	0.8	3.4	0.3	0.0	0.041	0.01
Ni	957.95	56260.47	1109.13	5.1	24.8	4.1	-	-	-
Pb	0.01	33.21	6.15	0.77	0.6	0.13	0.001	0.011	0.001
Pr	0.42	220.55	32.21	0.12	0.61	0.07	0.00	0.009	0.002
Sc	0.43	47.46	6.79	-	-	-	0.002	0.9	0.1
Sm	0.52	316.41	55.99	0.21	1.03	0.12	0.001	0.012	0.002
Sr	0.03	2.44	0.48	1.01	5.25	0.55	0.2	0.87	0.67
Tb	0.03	91.24	15.62	0.02	0.36	0.03	0.0	0.021	0.002
Th	0.04	21.45	2.38	-	-	-	0.0	0.1	0.015
Ti	0.22	13.17	3.13	-	-	-	0.003	0.2	0.03
Tm	0.12	38.80	6.94	0.01	0.23	0.02	0.0	0.001	0.00
U	0.22	3411.1	532.03	0.04	0.23	0.03	0.001	0.040	0.006
Y	8.2	4223.2	673.8	1.55	7.69	0.93	0.01	0.13	0.02
Yb	0.46	226.41	41.57	0.02	0.53	0.06	0.001	0.007	0.001
Zn	0.41	15.28	3.22	1.26	8.37	1.44	0.31	6.52	1.55
Cl ⁻	2.36	575.32	156.42	-	-	-	-	-	-
F ⁻	0.66	83.11	12.14	-	-	-	-	-	-
SO ₄ ²⁻	422.14	16758	1020.25	96.23	2044.82	400.28	30.1	1454.6	309.5
pH	3.27	5.46	0.47	-	-	-	-	-	-
PO ₄ ³⁻	0.03	0.84	0.12	-	-	-	-	-	-
HCO ₃ ⁻	0.76	53.21	7.63	-	-	-	-	-	-
DOC	0.89	20.64	2.57	-	-	-	-	-	-

2.2 Statistical methods

Statistical methods can be used to interpret complex relationships between geochemical parameters when a large amount of data is available. In this study, statistical methods including hierarchical cluster and factor analyses were used for groundwater and soil samples separately. Z-standardization was performed on the data to avoid the influence of data scale (Davis, 2002).

2.2.1 Hierarchical cluster analysis

To find the statistical similarity between REE and other analytical parameters, hierarchical cluster analysis was used. Comparing the results of the different data sets (groundwater, water-leachate and ammonium nitrate-leachate samples) is helpful to learn more details statistical behavior of the data sets. The obtained results were used to interpret the source of contamination.

Hierarchical cluster analysis is a common multivariate technique used in geochemical statistical studies. Cluster analysis can be applied as an “exploratory data analysis tool” to better understand the multivariate behavior of a data set. The principal aim of this technique is to partition multivariate observations into a number of meaningful, multivariate homogeneous clusters. By applying this technique, a large amount of data reduces into a few clusters. Hence, interpreting the clusters is easier, as each cluster contains some observations (samples or variables) that are similar to each other. The outcomes of a cluster analysis are shown by a special plot called dendrogram. The hierarchical clustering method uses the dissimilarities /similarities or distances between objects when forming the clusters. Therefore, two important measurements for this output are distance and linkage measures. It should be noted that diverse techniques can yield different clusters, even when using exactly the same data.

The most straight-forward way of computing distances between objects in a multi-dimensional space is to compute the Squared Euclidean distance (SED) (Davis, 2002) according to Eq. 4.

$$SED = \sum_{i=1}^n (x_i - y_i)^2 \quad (4)$$

where, x, y are different observations in the data matrix.

Furthermore, the linkage method should be defined. Ward is a linkage method that is commonly used for geochemical data, and thus was used in this study.

With geochemical data, cluster analysis can be used in different ways: R-mode clustering that can be used to cluster the variables (e.g., to detect geochemical relations between the variables) and Q-mode can be used to cluster the samples (e.g., to assign samples to different domains) to arrive at more homogenous data subsets for further data analysis. The distance value of 15 is considered as a criterion to separate the main clusters.

2.2.2 Factor analysis

Factor analysis, a well-known statistical technique, is a powerful tool to study the interrelationship among the various components. It compresses the total information content of the multivariate data in terms of a few factors. The information gained about the interdependencies can be used later to reduce the set of variables in a dataset. R-mode factor analysis was used to describe the relationship among the analytical parameters. Principal component and Varimax factor rotation were applied in this study. Principal component analysis seeks a linear combination of variables such that the maximum variance is extracted from the variables. It then removes this variance and seeks a second

linear combination which explains the maximum proportion of the remaining variance, and so on. This is called the principal axis method and results in orthogonal (uncorrelated) factors. A Varimax rotation is a change of coordinates used in factor analysis that maximizes the sum of the variances of the squared loadings. This scheme is used for orthogonal rotation which reduces the factors as follows: for each factor, high loadings (correlations) will result for a few variables; the rest will be near zero. Varimax rotation is often used in surveys to see how groupings of questions (items) measure the same concept.

The numbers of factors for each data set were defined by a Scree-plot of each data set. A Scree-plot helps the analyst to visualize the relative importance of the factors. A sharp drop in the plot signals means that subsequent factors are ignorable. By using this technique in this study, the number of variables reduces to a maximum of 4. The criterion for assigning an analytical parameter to a factor is 0.8 factor value. All values below 0.8 should be ignored. For more detailed description of the mathematical principles and factor analysis, see Krumbein et al. (1965), Johnson et al. (1988), and Bakaç (2000).

3 Results

3.1 General chemical analyses

The first parameter to be discussed is the pH range of samples. The samples leached with ammonium nitrate exhibit pH between 3.8 and 4.3, whereas groundwater pH ranges between 3.2 and 5.4. The water leached samples are all above 4.5 to 5.3. Water had a pH value of 5.5 after equilibration with the air, and ammonium nitrate had a pH of 4.7. The soil samples obtained at locations that were amended as part of a prior project (TS and CS) exhibited higher pH values when leached with ammonium nitrate (4.8 to 6.2) and water (7.6 to 7.8). The EC of groundwater samples varies from 760 to 15,800 $\mu\text{S}/\text{cm}$. Samples leached with water exhibit low electrical conductivity between 59.2 to 831 $\mu\text{S}/\text{cm}$. The samples leached with ammonium nitrate display an electrical conductivity between 89,900 and 100,100 S/cm . Ammonium nitrate had an EC of 99,775 $\mu\text{S}/\text{cm}$, whereas water showed an EC of almost 0 (1.2 $\mu\text{S}/\text{cm}$).

The results of chemical analyses of groundwater, the soil leachate samples (ammonium nitrate and water) and statistical data are presented in Table 1. The ΣREE concentration in groundwater samples approximately varies from 30 to 8000 $\mu\text{g}/\text{L}$. The concentration of ΣREE in samples leached with water varies from 0.007 to 123 $\mu\text{g}/\text{L}$ leachate corresponding to 0.00007 to 1.23 $\mu\text{g}/\text{g}$ soil. However 28.5 % of the measurements are below the detection limit (0.00005 $\mu\text{g}/\text{g}$ soil or 0.005 $\mu\text{g}/\text{L}$). The concentration of single REE varies between 0.00001 (detection limit) to 0.115 $\mu\text{g}/\text{g}$ soil or 0.001 to 115 $\mu\text{g}/\text{L}$ leachate with an average value of 0.001 $\mu\text{g}/\text{g}$ soil or 0.01 $\mu\text{g}/\text{L}$ leachate. Samples leached with ammonium nitrate contain a higher concentration of REE; 4-7 $\mu\text{g}/\text{g}$ range (400-700 $\mu\text{g}/\text{L}$ leachate) with an average of 6.5 $\mu\text{g}/\text{g}$ soil (650 $\mu\text{g}/\text{L}$ leachate).

The groundwater at the test site is in acidic range and oxic (Eh: 340- 715 mV). The general water type is $\text{Mg}-(\text{Ca})-\text{SO}_4$. Magnesium occurring as main cation in most measuring points is in the range of 58 and 3660 mg/l , followed by Ca with values from 77 to 672 mg/l . Highest metal concentrations were found for Mn ($166 \pm 153 \text{ mg}/\text{l}$) > Al ($42.1 \pm 50.9 \text{ mg}/\text{l}$) > Ni ($15.7 \pm 1.1 \text{ mg}/\text{l}$) > Zn ($4.0 \pm 3.2 \text{ mg}/\text{l}$) > Co ($3.9 \pm 3.3 \text{ mg}/\text{l}$) (Table 1). The Fe concentrations generally are low, 67% of the data are below

detection limit. However the maximum of Fe concentration was 180 mg/l in GTF25. Uranium which was leached during mining operations is in the range of 0.2 and 3411 µg/l. The highest concentration of REE was measured in GTF16 (8148.4 µg/l, December 2006), and the lowest concentration in GTF7 (11.4 µg/l, September 2007). Sulfate that occurs as main anion due to pyrite oxidation in the host rock and use of sulfuric acid during leaching operations varies between 422 and 16,758 mg/l, followed by Cl⁻ (2.3- 575 mg/l). Bicarbonate was compared to SO₄²⁻ present only in low concentrations especially in measuring points in the southern area, ranging generally between 0.7 and 53 mg/l. The dissolved organic carbon (DOC) ranges between 0.8 and 20.6 mg/l.

3.2 REE patterns

With regards to the comparison of the REE patterns of the data, generally, the patterns are similar qualitatively. All samples exhibit a MREE enrichment and LREE depletion with respect to the HREE. Furthermore, all samples except of some water leached samples have a positive Ce anomaly. The REE pattern from the analyzed groundwater samples (Fig. 2); exhibits a clear MREE enrichment and LREE depletion relative to the HREE. All groundwater samples exhibit a slight positive Ce anomaly ($Ce/Ce^* = 2.4 \pm 0.6$). Moreover, some samples exhibit also a slight positive Gd anomaly ($Gd / Gd^* = 1.3 \pm 0.06$).

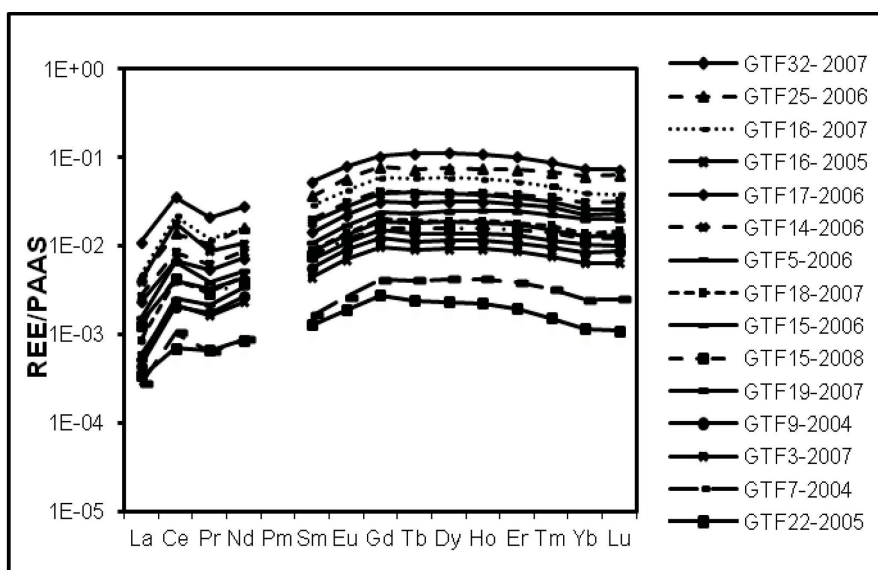


Fig. 2 PAAS-normalized REE patterns of selected groundwater samples. $N_{total}=175$. The concentration range varies by almost two orders of magnitude whereas the patterns are rather similar for all samples; the slight differences are the results of varying sampling locations with respect to the groundwater flow direction

REE patterns of water-leached samples also display a MREE enrichment (Fig. 4), with a positive Gd anomaly ($Gd/Gd^*=1.3 \pm 0.6$). Furthermore, many samples display a positive Ce anomaly ($Ce/Ce^*=0.8 \pm 0.6$). It is interesting to note that the Ce anomaly is higher for samples obtained from deeper horizons (lower than 20 cm) as compared to near-surface ($Ce/Ce^*>2$ vs $Ce/Ce^* < 1.5$). Light-REE are also depleted relative to HREE.

Water-leached samples can clearly be discriminated between normal soil and high level contaminated samples that are called hot spots (Fig. 3).

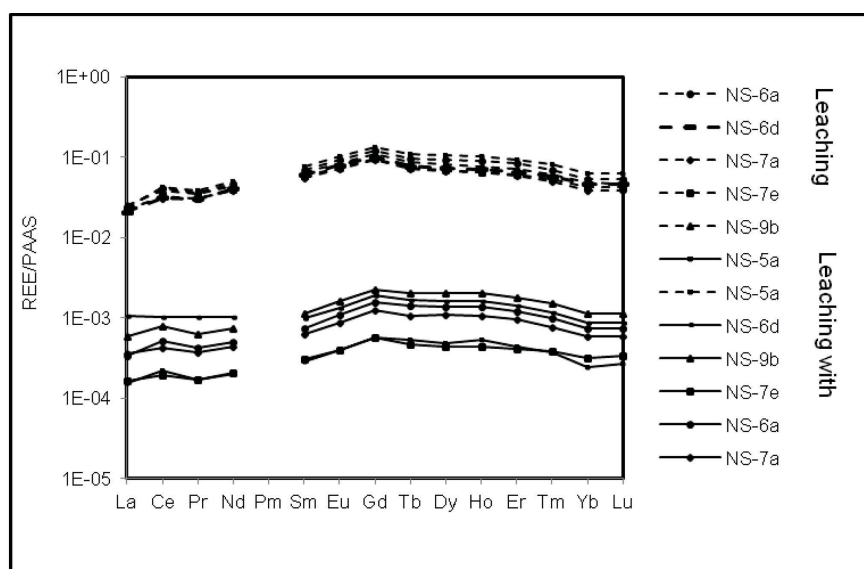


Fig. 3 PAAS-normalized REE patterns of selected samples; leached by water and ammonium nitrate. $N_{total}=90$ for each leachate. The concentration for water-leached samples ranges over one order of magnitude, due to different soil type and soluble compounds. The REE pattern is generally the same with slight variations that reflect the different soil amendments. Ammonium nitrate-leached soil samples are all very similar; the eluted metal concentrations are always in the same range. Soils have similar slightly bound metal amounts

However, more interesting is that different soil treatments display different pattern. Indeed, samples from an amended plot (CS or TS) show an enrichment of the elements Pr to Gd compared to the heavier ones. Furthermore, no Ce anomaly (positive or negative) is recognizable (Fig. 4). The samples that were leached with ammonium nitrate reveal a higher concentration of REE (Fig. 3). A positive Ce anomaly is present in all samples ($Ce/Ce^*=1.3 \pm 0.2$), as well as a positive Gd anomaly ($Gd/Gd^*=1.40 \pm 0.04$).

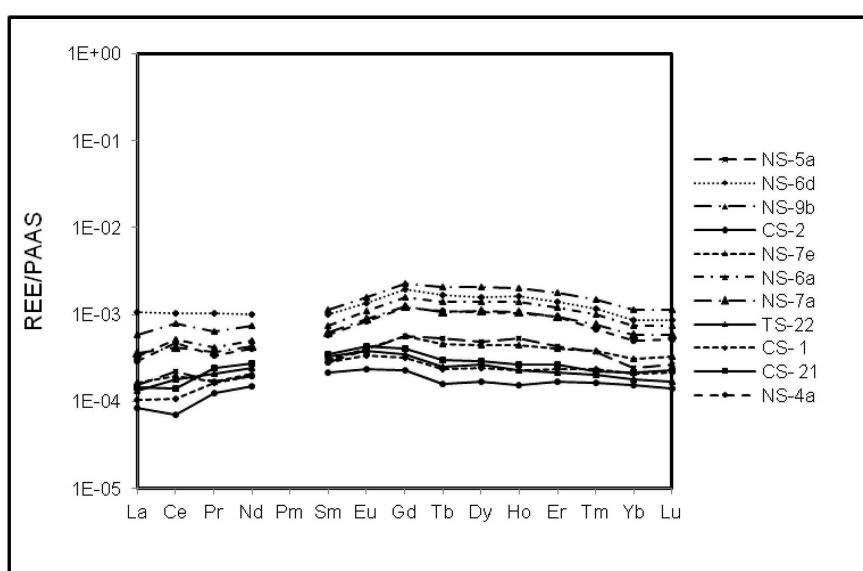


Fig. 4 PAAS-normalized REE patterns of samples from an amended plot (CS or TS); these samples show an enrichment of the elements Pr to Gd compared to the heavier ones and no Ce anomaly (positive or negative) is recognizable

3.3 Cluster analysis

In this study, the linkage distance of 15 was considered as a criterion to group the members. In groundwater samples (Fig. 5), generally the analytical parameters are divided into three main clusters. Aluminum and Y have great statistical similarity with REE and fall into one sub-cluster. Moreover, Cu, Fe, Th and U are placed in the same main cluster with REE. Furthermore, there is a great statistical similarity among Mg, Mn, HCO_3^- , SO_4^{2-} and Zn, and these elements compose a second main cluster with pH.

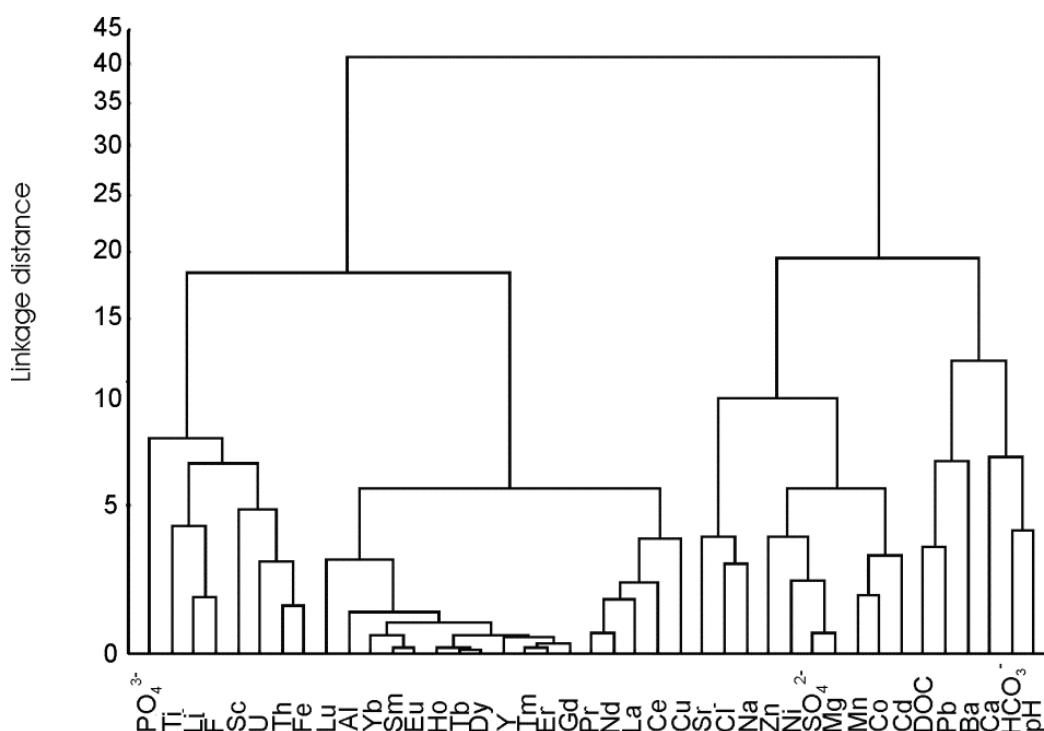


Fig. 5 Dendrogram of groundwater samples; Ward linkage, and squared Euclidean distance. The linkage distance value of 25 is considered as a criterion to separate the main groups. Hence, the analytical parameters are divided into two main clusters

Cluster analysis of the data resulting from water-leached samples reveals three main clusters (Fig. 6). Here, REE show statistical similarities to Cu, U and Y, and unlike the groundwater samples, the elements Al, Fe and Th are not in the same cluster with REE. These elements together with Pb and PO_4^{3-} are forming another main cluster. The difference between the clusters indicates that a different physicochemical process is active in the groundwater versus the water leachate.

In the ammonium nitrate-leached soil samples, the analytical parameters formed two main clusters (Fig. 7). REE with exception of La are statistically similar to Al, Cu, Fe, Pb, U and Y. The second main cluster includes La, Mg, Mn, S and Zn. The dendrogram resulting from this data set is similar to that of the groundwater samples. Based on this observation, it can be assumed that some similar mechanisms such as cation exchange with NH_4^+ and dissolution are active in groundwater formation and ammonium nitrate leaching. It is important to note that similar results of the dendrograms are corresponding to how the analytical results are similar. These similarities are defined by their placing in main or sub-clusters but not the general shape of the dendrograms. In this study the reason that the dendrograms of groundwater samples and ammonium nitrate-leached samples are called similar is due to the statistical similarities between REE and other metals. For examples, in the dendrogram of water-leached soil samples, Al, Fe and Th are not statistically similar with REE.

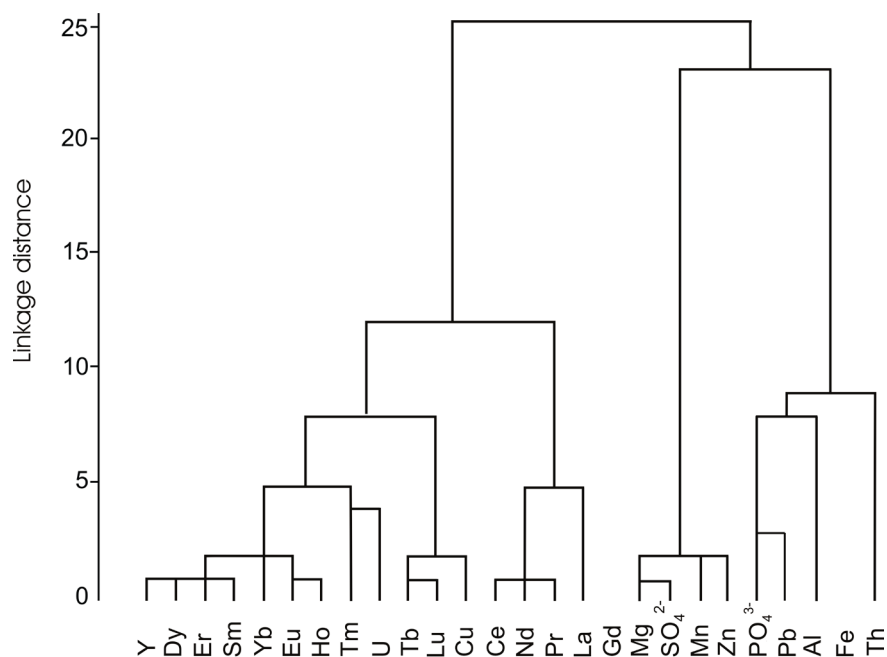


Fig. 6 Dendrogram of water-leached soil samples; Ward linkage, and squared Euclidean distance. The linkage distance value of 15 is considered as a criterion to separate the main clusters. The analytical parameters are divided into three main clusters

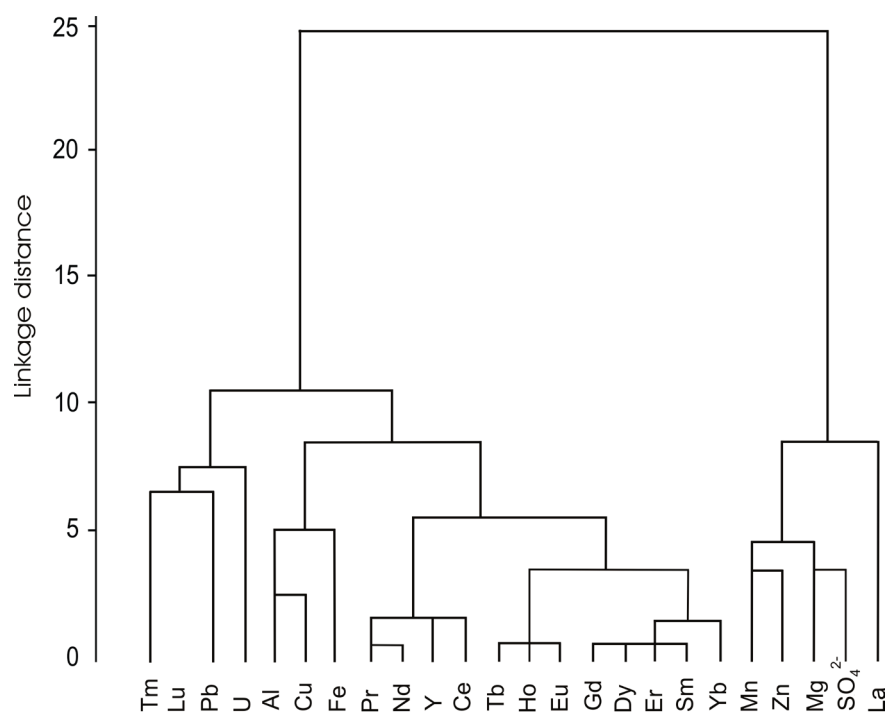


Fig. 7 Dendrogram of ammonium nitrate-leached soil samples; Ward linkage, and squared Euclidean distance. The linkage distance value of 15 is considered as a criterion to separate the main clusters. The analytical parameters are divided into two main clusters

However, these two elements have statistical similarities with REE in groundwater samples and ammonium nitrate-leached soil samples.

Q-mode cluster analysis was used for two data sets, water-leached and ammonium nitrate-leached soil samples. Data from the soil samples leached with water formed three main clusters. The first cluster reflects the samples obtained from the zones with normal soil. The second cluster consists of

data from samples collected from the zone within the test site that having a higher level of contamination. This can be observed in Fig. 8, a1, a2, where cluster one corresponds to low REE concentrations and cluster two with higher REE concentrations.

The data from samples leached with ammonium nitrate formed two main clusters as the case for the water leached samples. The clusters correspond to the zones of the normal soil and to the higher level of REE. This can be observed in Fig 8, b1, b2.

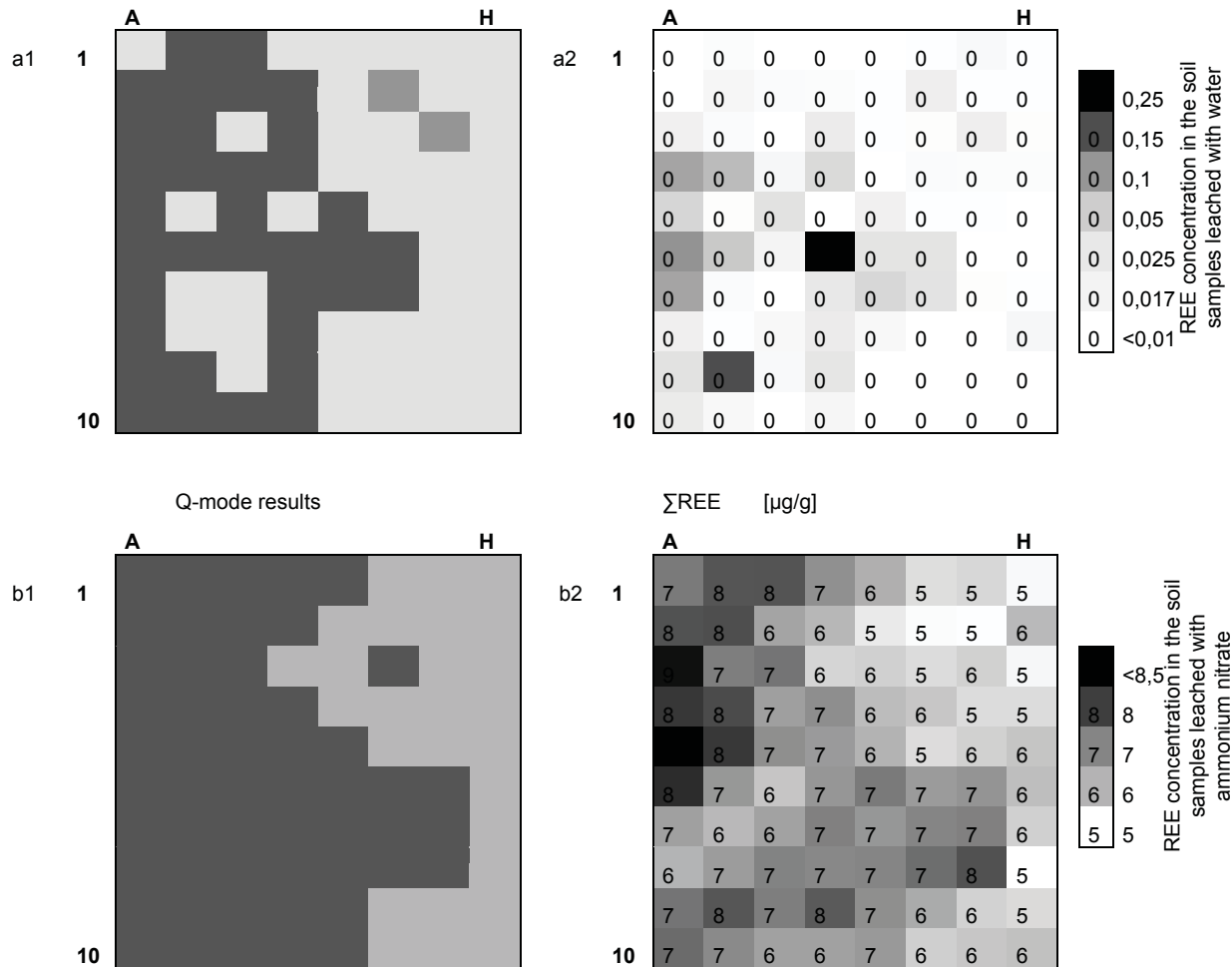


Fig. 8 (a1) the distribution of REE in the soil sampling zone with regards to water-leached data (upper left); (a2) shows the Q-mode clustering result of soil samples leached by water (upper right). Mainly the data were divided into two clusters 1 and 2. Cluster 1 is composed of samples with low concentration of REE and cluster 2 is composed of the samples with concentration of REE; (b 1) the distribution of REE in the soil sampling zone with regards to ammonium nitrate-leached data (lower left); (b 2) shows the Q-mode clustering result of soil samples leached by ammonium nitrate. Mainly the data were divided into two clusters 1 and 2 (lower right)

3.4 Factor Analysis

The factor analysis for groundwater data shows three factors that explain 90.2% of the variance (Table 2a). The first factor explains 76.87% of the variance and is highly loaded with Al, Cu and REE. The second factor explains 7.43% the variance and is loaded with Mn and Mg. The third factor explains 5.2% of the variance and is loaded with Fe and Th.

For the water leachate samples (Table 2b) four factors can explain 72.90% of the variance. The first factor explains 48.25% of the variance and is highly loaded with the Cu, REE (except Ce, Gd, La,

Nd, Pr), U and Y. Factor 2 explains 13.01% of the variance and is loaded with Gd, La, Mg, Mn, S and Zn. Factor 3 explains 8.50% of the variance and is loaded with Fe and Pb. The last factor explains 3.22% of the variance and is loaded with Th.

For ammonium nitrate leachate samples (Table 2c) three factors can explain 80.80% of the variance. The first factor explains 58.38% of variance and loaded by Cu, REE (except La) and Y. Factor 2 explain 12.24% of variance and it is loaded with Mg, Mn and Zn. The third factor explains 10.18% of the variance and is loaded with U.

Table 2 (a) Factor loading matrix after Varimax rotation of groundwater samples; N= 175; (b) Factor loading matrix after Varimax rotation of water-leached samples; N= 90; (c) Factor loading matrix after Varimax rotation of ammonium nitrate leached-samples; N= 90

(a)				(b)					(c)			
	1	2	3		1	2	3	4		1	2	3
Dy	0.94	0.26	0.18	Tb	0.97	-0.13	-0.02	-0.1	Ce	0.96	0.18	0.03
Tb	0.94	0.27	0.18	Cu	0.94	0.08	0.00	0.15	Gd	0.96	0.14	0.24
Nd	0.94	0.28	0.11	Lu	0.94	-0.24	0.03	0.03	Dy	0.95	0.14	0.26
Pr	0.94	0.30	0.06	Eu	0.93	0.26	0.12	-0.07	Nd	0.95	0.23	0.14
Ho	0.94	0.26	0.20	Er	0.92	0.35	0.07	-0.02	Sm	0.94	0.14	0.28
Er	0.94	0.25	0.22	Yb	0.92	0.3	0.08	0.03	Er	0.94	0.12	0.31
Y	0.94	0.27	0.19	Dy	0.91	0.4	0.12	-0.02	Y	0.93	0.28	0.09
Gd	0.94	0.27	0.21	Ho	0.89	0.3	0.10	0.08	Cu	0.92	-0.09	0.06
Eu	0.94	0.26	0.21	Sm	0.88	0.4	0.23	0.03	Yb	0.92	0.06	0.38
Sm	0.94	0.27	0.21	Y	0.87	0.5	0.10	-0.00	Pr	0.91	0.31	0.14
Tm	0.93	0.24	0.25	Tm	0.83	0.1	0.08	-0.82	Eu	0.90	0.09	0.40
Yb	0.93	0.23	0.30	U	0.80	.22	0.34	0.02	Ho	0.89	0.10	0.41
Lu	0.92	0.22	0.30	Nd	0.79	0.52	0.25	0.00	Tb	0.88	0.10	0.43
Al	0.92	0.26	23	Ce	0.77	0.58	0.19	-0.01	Al	0.77	0.10	0.42
Ce	0.91	0.32	-0.03	Pr	0.71	0.59	0.29	-0.05	Fe	0.87	-0.24	-0.08
La	0.90	0.329	0.01	S	0.16	0.95	-0.18	0.03	Tm	0.74	0.01	0.59
Cu	0.85	0.06	0.17	Mg	0.12	0.93	-0.17	0.01	Lu	0.72	-0.00	0.61
Zn	0.79	0.48	0.12	Mn	0.18	0.92	-0.14	0.03	Pb	0.52	-0.08	0.30
U	0.70	0.05	0.62	Zn	0.05	0.92	-0.15	0.09	Zn	0.20	0.93	0.04
pH	-0.67	0.13	-0.11	La	.036	0.82	0.09	-0.05	Mn	0.24	0.84	0.03
Mn	0.37	0.84	-0.06	Gd	0.40	0.82	0.24	0.00	Mg	0.07	0.80	0.38
Mg	0.46	0.82	0.18	Fe	-0.04	-0.21	0.94	-0.05	La	-0.22	0.67	-0.42
S	0.52	0.79	0.19	Pb	0.16	-0.16	0.83	0.14	U	0.16	0.09	0.82
Fe	0.27	-0.05	0.86	Al	0.39	0.08	0.77	-0.09				
Th	0.33	0.18	0.81	Th	0.02	0.11	0.02	0.97				
Pb	0.13	0.24	-0.26									

4 Discussion

The comparison of REE patterns in different data sets shows that generally the patterns are qualitatively similar. In all samples, an MREE-enrichment and LREE-depletion relative to HREE is noticeable. Furthermore, all samples show a positive Ce anomaly. However, the leachate samples obtained using ammonium nitrate reveal a higher concentration of REE.

One noticeable difference between REE patterns is the slight positive Gd anomaly. In addition, a difference between the data sets is the variation of the LREE depletion relative to HREE. In

groundwater samples this variation is less noticeable as compared to the leachate samples. The variation between the patterns for ammonium nitrate-leached samples is not as high as for groundwater.

Q-mode cluster analysis resulted in clusters that reflect both contamination level and soil type for the two data sets that were obtained using leaching. For example, one cluster corresponds to the high level of contamination. Furthermore, cluster analysis showed that there are greater similarities between groundwater samples compared to ammonium nitrate-leached samples than compared to water-leached soil samples. Also, the pH ranges of these data sets are more similar to each other than to the one of water-leached samples. We note similar results using factor analysis, with the exception of La. Lanthanum is not in the same factor with other members of REE for the ammonium nitrate-leached samples. This exception of La from the REE cluster and the factors is not expected and should be investigated further. However, it is worth noting that this unexpected result for La was previously observed for the same test site (Lonschinski, 2009).

The LREE-depletion relative to HREE is a typical feature of an AMD-influenced area as described by Lei et al. (2008). Middle REE-enrichment is one effect of pyrite oxidation (Grawunder, 2010). Furthermore, preferential sorption onto clay minerals of HREE compared to LREE can also contribute to that fractionation. The higher concentration of REE in samples that were leached with ammonium nitrate could be due to cation exchange with NH_4^+ .

As the cluster analysis showed, REE and some other metals are clustered along with Al and Fe. This can be interpreted as an effect of Al- and Fe-hydroxides; the existence of hydroxides and clay minerals at the test site has been previously documented (Burkhardt et al., 2009; Carlsson and Büchel, 2005; Grawunder et al., 2009). Hence, many features of the studied hydrogeological system are connected to the fine grained soil material and its high content in hydroxides, specially of Fe and Al.

The pH is a known factor regarding metal mobility, higher metal mobility occur at lower pH (Aström, 2001; Semhi et al., 2009; Shan Xiao-quan, 2002). For the pH conditions of the test site, Al- and Fe-hydroxides are possible in soil since precipitation occurs at pH values above 4. The water-leached samples are all above 4.5, as well as most of the ammonium nitrate leached samples, whereas groundwater shows some lower pH, around 3. This could explain some of the differences observed between the soil and the groundwater, specially with regard to the REE patterns.

Quinn et al. (2005) found a bonding mechanism between HREE and Al-hydroxide as compared with MREE and LREE. Rare earth elements behavior towards Fe-hydroxides has been described previously, which states that REE can be enriched through sorption and co-precipitation with Fe-oxy-hydroxides, even at pH as low as pH 4 (Aström, 2001). Verplanck et al., (2004) stated that Fe-hydroxides do not fractionate REE below a pH of 5.1. However, Lei et al., (2008) adjusted this critical pH downwards to 4. Above this pH, REE tend to fractionate. A previous study at the test site showed that the critical pH for REE mobility at the test site is about 4.5 (manuscript 1).

The observed anomalies of the REE patterns can also be explained by the physico-chemical characteristics of this system.

Positive Gd anomaly is generally known to be a characteristic of anthropogenic-influenced material (Möller, 2002), although at this site we have a geogenic origin. Indeed, geogenic Gd anomaly are the result of a greater stability of inner-sphere Gd-complexes, Gd being more stable than its neighbors

(Coppin et al., 2002). Amino-NH₂-group containing complexes are one example in which Gd is more stable than other REE (Hennebrüder, 2003). In systems with increasing salinity, the relative amount of Gd increases in the soluble phase compared to the other REE since those decrease by co-precipitation. Small geogenic anomalies can be caused where Fe-hydroxide precipitation takes place (Hötz et al., 2008). A positive Ce anomaly is a common REE characteristic, as Ce is often found bound to oxy-hydroxides (Steinmann and Stille, 2008). Here it is probable that the positive Ce anomaly is the result of leaching a metal-enriched soil layer (the observed dark cemented Fe-Mn layer) that was likely formed by oxidation processes or pH changes, and complex Ce and other metals by co-precipitation. Indeed, the dark color of some layers in the soil is also an indication of high Fe contents (over 65% for a hardpan layer) and are likely due to secondary precipitation of Fe, which is in turn due to groundwater fluctuation (Carlsson and Büchel, 2005).

The occurrence of this anomaly above the groundwater level can be explained by capillary rise, from the groundwater, as it has been described by Grawunder (2010). Indeed, the capillary rise from the groundwater to higher soil levels is possible due to the presence of fine grained material (Lonschinski et al., 2010). Additionally, the observed red-colored solution at some locations of the test field suggest the occurrence of a diffusive return flow, i.e., upward transport of groundwater to the surface, and spreading via overland flow on the surface.

A schematic of the different processes involved in the spreading of contamination to the upper soil level is shown on Fig 9.

A variation in the degree of the LREE depletion relative to HREE for leachate obtained using ammonium nitrate as compared to that of groundwater. This is primarily the result of dilution effects due to rain fall events. A secondary effect is the variability in soil composition over a large geographic area; hence water is flowing through soil that may consist of different soil types. Additionally, retention time influences processes, redox conditions may change due to water saturation. Lonschinski (2009) describes the saturation of the soil at a depth of 30 cm to be at a high level during the wet season after the winter.

5 Conclusion

We have used multivariate statistical analysis to explain the probable mechanism of contaminant mass transport for the period beginning with the closing of the mine until today. Leachate from the former heaps infiltrated into deeper soil, transported dissolved metals, and has formed a zone at the former groundwater level containing high metals concentrations, which in turn has led to secondary precipitation of metals in that zone. This precipitation has also been found at the current groundwater level, which has acted as a source of contamination for the allochthonic soil layer that replaced the former heap. The presence of minerals that are rich in Al- and Fe-hydroxides has played a role in the formation of complexes with some of the REE, and can be a cause of the observed Ce anomalies.

Furthermore, during periods of increased rainfall, the elevated groundwater level causes return flow which is highly mineralized and hence, contaminated. In addition, because the subsurface is composed of very fine-grained material, capillary rise also transports mineralized water to the ground

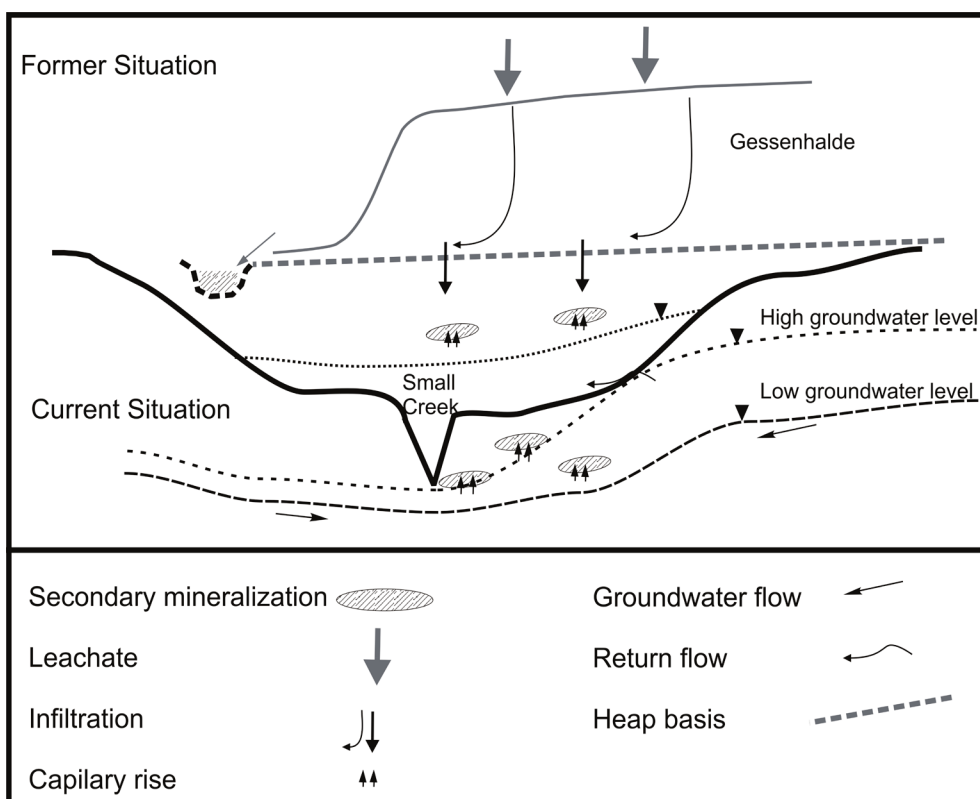


Fig. 9 Schematic representation of underground processes that cause contaminant exchange between groundwater and surface soil

surface. The dissimilarities observed between the soil leachates and the groundwater can be explained by the difference of their systems. The site is influenced by fluctuating groundwater levels that lead to sporadic return flow during high precipitation events. While the fluctuating groundwater level can make possible a high degree of exchange of (soluble) elements between water and soil, it may also be possible that the action of microorganisms and plants play a role in the fate of metals within the upper soil level.

With regards to the outcomes, based on the results presented in this paper, for sites with similar characteristics as the Gessenhalde ammonium nitrate solution can be a good one for leaching experiments.

Acknowledgments:

This project is part of the Jena School of Microbial Communication (JSMC) and graduate research training group GRK 1257/1 (Alteration and element mobility at the microbe-mineral interfaces) financially supported by the German Research Society (DFG).

References

- Åström M (2001). The effect of acid soil leaching on trace element abundance in a medium-sized stream, w. Finland. *App.Geo.*16: 387–396.
- Aubert D, Stille P, Probst A (2001). REE fractionation during granite weathering and removal by waters and suspended loads: Sr and Nd isotopic evidence. *Geochim. Cosmochim. AC.* 65: 387- 406.
- Bakaç M (2000). Factor analysis applied to a geochemical study of suspended sediments from the Gediz river, western Turkey. *Environ. Geochem. Hlth.* 22: 93-111.
- Burkhardt EM, Meißner S, Merten D, Büchel G, Küsel K (2009). Heavy metal retention and microbial activities in geochemical barriers formed in glacial sediments subjacent to a former uranium mining leaching heap. *Chem. Erde.* 69: 21-34.
- Büchel G, Bergman H, Eberna G, Kothe E (2005). Geomicrobiology in remediation of mine waste. *Chem. Erde.* 65 (S1): 1-5.
- Carlsson E, Büchel G (2005). Screening of residual contamination at a former uranium heap leaching site, Thuringia, Germany. *Chem. Erde.* 65 (S1): 75–95.
- Coppin F, Berger G, Bauer A, Castet S, Loubet M (2002). Sorption of lanthanides on smectite and kaolinite. *Chem. Geol.*182 (1): 57-68.
- Davis JC (2002). *Statistics and data analysis in geology.* John Wiley & Sons. New York, United States of America.
- Geletneky JW (2002). Hydrogeologische/Hydrologische Untersuchung einer Prä-Flutungssituation am Beispiel des Gessentals im ehemaligen ostthüringischen Uranbergbauggebiet. Doctoral dissertation, Friedrich Schiller university of Jena (In German).
- Grawunder A, Lonschinski M, Merten D, Büchel G (2009). Distribution and bonding of residual contamination in glacial sediments at the former uranium mining leaching heap of Gessen/Thuringia, Germany. *Chem. Erde.* 69: 5-19.
- Grawunder A (2010). Hydrogeochemistry of rare earth elements in an acid mine drainage influenced area. Doctoral dissertation, Friedrich Schiller university of Jena.
- Hennebrüder K, Wennrich R, Mattusch J, Stärk H, Engewald W (2003). Determination of gadolinium in river water by SPE pre concentration and ICP-MS. *Talanta.* 63: 309-316.
- Jakubick A, Jenk U, Kahnt R (2002). Modeling of mine flooding and consequences in the mine hydrogeological environment: flooding of the Königstein mine, Germany. *Environ. Geol. Water S.* 42(2): 222-234.
- Johnson AR, Wichern WD (1988). *Applied multivariate statistical analysis* (2nd
- Kaiser H (1965). Zum Problem der Nachweigrenzer. Workshop: Moderne Methoden der anorganischen Analyse. Düsseldorf, Germany (In German).
- Kahlert J (1992). Wismut und die Folgen des Uranbergbaus: eine Tagung der Friedrich-Ebert-Stiftung. Proc. Conf. Friedrich-Ebert foundation. Bonn, Germany (In German).
- Krumbein WC, Graybill FA (1965). *An introduction to statistical models in geology.* Mc Graw-Hill. New York, United States of America.
- Lange G (1995). Die Uranlagerstätte Ronneburg, *Zeitschr. Z. Geo. Wiss.* 23: 517-526.

- Lei L, Song C, Xie X, Li Y (2008). REE behavior and effect factors in AMD-type acidic groundwater at sulfide tailings pond, BS nickel mine, W.A. T. Nonfer. Metal Soc. 955-961.
- Lonschinski M (2009). Schwermetalle im System Boden-Wasser-Pflanze mit Hinblick auf Fraktionierungsprozesse der Seltenen Erden Elemente auf der Aufstandsfläche der Gessenhalde im ehemaligen Uranbergbaurevier Ronneburg (Ostturingen). Doctoral dissertation. Friedrich Schiller university of Jena, (In German).
- Lonschinski M, Knöller K, Merten D, Büchel G (2010). Flow dynamics of groundwater and soil water in the former heap Gessenhalde at the uranium mining area of Ronneburg: a stable isotope approach. Hydrol. Process. 6(25): 186-172.
- Mc Lennan SM (1989). Rare earth elements in sedimentary rocks: influence of provenance and sedimentary processes. Mineral. Geochem. 21:169-200.
- Merten D, Geletneky J, Bergmann H, Haferburg G, Kothe E, Büchel G. (2005). Rare earth element patterns: a tool for understanding processes in remediation of acid mine drainage. Chem. Erde. 65: 97-114.
- Möller P, Paces T, Dulski P, Morteami G (2002). Antropogenetic Gd in surface water drainage system, and the water supply of the city of Prague, Czeck Republic. Environ. Sci. Technol. 36: 2387–2394.
- Rabiet M, Brissaud F, Seidel JL, Pistre S, Elbaz-Poulichet F (2009). Positive gadolinium anomalies in waste water treatment plant effluents and aquatic environment in the Hérault watershed (South France). Chemosphere. 8: 1057-1064.
- Neagoe A, Ebona G, Carlsson E (2005). The effect of soil amendments on the plants in an area affected by acid mine drainage. Chem. Erde. 65(S1): 115-129.
- Rüger F, Dietel W (1998). Vier Jahrzehnte Uranbergbau um Ronneburg. Lapis. 7(8):14–18, (In German)
- Semhi K, Chaudhuri S, Clauer N (2009). Fractionation of rare earth elements in plants during experimental growth in varied clay substrates. Appl. Geochem. 24: 447-453.
- Shan Xiao-quan LJ, Wen B (2002). Effect of organic acids on adsorption and desorption of rare earth elements. Chemosphere. 47:701-710.
- Steinmann M, Stille P (2008). Controls on transport and fractionation of the rare earth elements in stream water of a mixed basaltic-granitic catchment basin (Massif Central, France). Chem. Geol. 254: 1-18.
- Verplanck PL, Nordstrom DK, Manning AH, Caine JS, Plumlee GS, Hunt AG, Bove DJ (2004). Linking geochemical and hydrologic models of ground water in two Rocky Mountain catchments: Straight creek, NM and Handcart. Geol. Soc. of America. 36(7): 539. Ed.: Prentice Hall, Englewood Cliffs, New Jersey, United States of America.
- Wismut GmbH (1994a). Entwurf Sanierungskonzept Standort Ronneburg.-Stand März. Chemnitz, Germany (In German).
- Zeien H, Brümmer GW (1989). Chemical extraction for identification of binding-forms of heavy metals Mitt. Dt. Bodenkundl. Ges. 59: 505-510

Manuscript 3

Fuzzy, hierarchical cross-clustering of slate samples contaminated with heavy metals

Fuzzy, hierarchical cross-clustering of slate samples contaminated with heavy metals

Anahita Pourjabbar¹ - Costel Sarbu² - Jürgen .W. Einax³ - Georg Büchel¹

¹ Institute of Geosciences, Friedrich Schiller University, Jena, Germany

²University of Babes-Bolyai Cluj- Napoca, Romania

³ Institute of Inorganic and Analytical Chemistry, Friedrich Schiller University, Jena, Germany

Abstract

In this paper, the characteristics of soil water and slate samples were critically analyzed using statistical clustering techniques. The main aim of this study was to investigate the source of contamination in an area close to a uranium-abandoned mine in Germany. The mining activities were abandoned in the 1990s, and the surrounding area was remediated. However, the contamination is still detectable in water, soil and plants. Hence, investigating the source of the current contamination is important task. In order to achieve the goal, two data series composed of the results from chemical analysis of both soil water samples and slate samples were analyzed using the the mentioned statistical methods. Two fuzzy clustering algorithms were used such as Fuzzy Divisive Hierarchical Clustering (FDHC) of samples, and Fuzzy Hierarchical Cross-Clustering (FHCC). The purpose for using these methods was to investigate the (dis)similarities between the two data sets. By using the cross-clustering algorithm, it is possible to identify which metals are responsible for the similarities or differences observed between different groups of samples.

By means of fuzzy clustering, the relation between the leached Ordovician-Silurian slates samples (10 samples) and soil water samples (53 samples) was identified: The slate samples are very similar and appear to be more close to the soil water samples collected in sampling point MF1 and MF2. The fuzzy cross-clustering approach allows the qualitative and quantitative identification of the characteristics (metal concentration) responsible for the observed similarities and differences between all the samples.

Key words: Fuzzy Clustering; Fuzzy Cross-Clustering; Soil Water Samples; Slate Sample; Metals; REE

1 Introduction

Contamination of the environment and the environmental impact on human health is one of the most important fields of study in modern analytical science. It is also known that many industries can seriously contaminate the environment and endanger human health (Hosono et al., 2010; Nocolas et al., 2000). Mining activities are known as a source of contamination on their surrounding environment, including water, soil, and plants. The type and the degree of contamination is related to the geology and geography as well as the mining technique such as excavation, and the mineral processing methods (Bhattacharya et al., 2006; Chopin and Alloway, 2007; Gomes and Favas, 2006; Lee, 2003; Morento et al., 2007; Rodríguez et al., 2009). Acid mine drainage (AMD) results from a mining process and it affects the surrounding environments. It is a polluted water that typically contains high levels of metals, including heavy metals, e.g. rare earth elements (REE) (Hadley and Snow, 1974). It is produced by the oxidation of sulfide minerals, chiefly pyrite or iron disulfide (FeS_2). This is a natural chemical reaction which can proceed when minerals are exposed to air and water and is found around the world. Although remediation strategies vary from one location to another depending on the environmental, economic, and technical situation, the monitoring of water, plant, and soil contamination levels is always helpful to evaluate and improve the remediation process (Bozau et al., 2007; Elias and Gulson, 2003; Franklin and Fernandes, 2011; Otte and Jacob, 2008).

In such studies where several variables (e.g., physicochemical parameters, composition of the host rocks, precipitation, hydrogeological processes, etc.) are involved, multivariate analyses are helpful tools. The reason is that they place the variables into more or less homogeneous groups so that the relation between the groups is revealed. However, the outcomes of classic multivariate statistical analysis are not always well-suited. Several parameters can disturb the outcome of the statistical analysis; for example, insufficient numbers of samples, non-accurate chemical analysis, or sampling methods. Furthermore, some affective parameters are either not reported numerically or have not similar measurement intervals; water-rock interaction or precipitation (usually have different measurement intervals) are examples of such a case and for this reason, the study area was ideal for the application of more advanced statistical techniques. Fuzzy methods are helpful in such cases because it is flexible and can consider more possible relationships between the parameters (Demicco and Klir, 2003).

Fuzzy logic, which is based on the ideas of fuzzy set theory by L.A. Zadeh 1965, provides a method to formalize reasoning when dealing with vague terms. Traditional computing requires finite precision that is not always possible in real world scenarios. Fuzzy logic is based on membership functions or degrees of truthfulness and falsehoods; which is not only either 0 or 1 but all the numbers that fall in between. A membership degree of 0 is a degree of an item which is not in the set, and 1 is a degree of an item which is in the set. A membership degree between 0 and 1 is of an item that is thought to be in the set. For example: "A" is a fuzzy set described by the items, u , where the set $A = \{u/a(u) \mid u \in U\}$; A is a set of items combined with their degrees of being in U , and $a(u)$ is a membership function between 0 and 1 which derives the degree that each u is or is not in U . It is important to note that membership functions are not necessarily based on statistic distributions.

Fuzzy logic can be successfully applied to geological studies because the study of geology itself is

based on inferences and decisions that are often incomplete and/or uncertain. Moreover, geological processes are complex, non-linear and possibly non-deterministic. Their non-deterministic nature may arise not from randomness, but from organized complexity. Another good example is a frontier region, where the available geological data are regional and details or hard evidences are inadequate. Fuzzy logic offers an alternative to statistical modeling in geology that is more computationally efficient and more intuitive for geologists than complicated numerical models consisting of a few sets of differential equations.

The main idea behind the current study is the investigation of the contamination sources of an area close to a leaching heap (Gessenhalde) of an abandoned mine in east Thuringia, Germany. This leaching heap was composed of Silurian slates. Acid mine drainage was used to leach uranium. Later on, up to 10 m of excavated material of the leaching heap were placed within the open pit mine (Lichtenberg) in order to remediate the area. However, the surrounding area is still contaminated by heavy metals. Special conditions at this area, namely low pH and high concentration of some heavy metals, lead to a high potential for a contaminated environment. The contamination was measured and reported for various media such as groundwater, soil water, surface water, soil, and plants (Carlsson and Büchel, 2005; Grawunder, 2010; Haferburg, 2007; Horn, 2003; Lonschinski, 2009; Lorenz, 2009; Mirgorodsky et al., 2010; Mirgorodsky et al., 2012; Olivier et al., 2010).

The hypothesis is that the Ordovician and Silurian slate leachates influence the environmental contamination (heavy metals including rare earth elements (REE)) in this area. In order to determine the influence that these slates may have on the contamination, chemical analyses stemming from soil water and slate samples (leached with water) were studied using statistical methods. Previous studies have been performed on soil water samples, soil samples, and slates in the study area focused on the chemical analyses (Lonschinski, 2009; Pasalic, 2011; Wagner, 2010). However, there was no application of statistical methods to the data that focused on the relation between different types of samples in order to study the hypothesis. Fuzzy Clustering (FC) and Fuzzy Cross Clustering (FCC) methods were used to study these samples. The reasons to select fuzzy techniques rather than classical statistical techniques in this study are: (a) the low number of slate samples compared with the number of soil water samples; (b) the existence of uncertainty in the data sets. These uncertainties have various reasons, for example complexation or water-rock interaction that are not reported numerically in the data sets. Hence, the data set is not a complete set of information. Furthermore, there is currently no access to the parent material and the only possibility is to sample from the material of the current contoured area. Since not all the parameters affecting these processes are present in the data set, classic statistical analysis is not well-suited.

2 Sampling site description

The former uranium mining site district in Eastern Thuringia and Saxony, Germany with more than 113,000 tons of mined uranium, was the third-largest uranium producer in the world (Jakubick et al., 2002; Lange, 1995). Mining activities started in 1949 and ended in 1990 after the re-unification of Germany. The remnants of the mining included a large open pit mine (which is filled now), Lichtenberg, with more than 200 m depth, 1.6 km length and 0.6 km width, an underground mining system going down to 900 m with 3000 km of underground galleries and several waste rock piles in

which acid mine drainage (AMD) occurred (Wismut GmbH, 1994). Among several heaps in the area, the Gessenhalde was the only leaching heap built up by Ordovician and Silurian slates. Gessenhalde was constructed as a production site for leaching low-grade ore. Waste rocks with a low grade of uranium mineralization, uranium content <300 g/ton, (Rüger and Dietel, 1998) were leached with AMD and with sulfuric acid (10 g/l) (Wismut GmbH, 1994). During the leaching process, leachate seeped through the lining of the Gessenhalde and accumulated in the Quaternary sediments underneath. The leach pad had been sealed with 0.6 m of loam and was compacted in order to prevent infiltration. This seal was covered by a one-meter-thick layer of coarse waste rock containing low grade of uranium mineralization from Lichtenberg, as a drainage layer during the leaching process. Drainage gullies were designed to transport the leaching solution to collection ponds. It is probable that these gullies were not completely sealed, and that some contamination infiltrated the under-lying soil to a great depth. In 1989, leaching was stopped (Wismut GmbH, 1994). In the 1990s, the heap was removed and used to fill the nearby open pit, Lichtenberg. The area was left uncovered, which led to the formation of puddles before 10 m of the underlying Quaternary sediments were excavated and a layer of uncontaminated material, including allochthonic top soil, was used to re-contour the bottom of the pit. A few years later, in 2003 and 2004, the evidence of residual heavy metal contamination was measured within that upper soil layer as reported by Carlsson & Büchel (2005).

The contamination is still detectable in the area, in groundwater, soil water, surface water, soil, and plants (Carlsson and Büchel, 2005; Grawunder, 2010; Hafeburg, 2007; Horn, 2003; Lonschinski, 2009; Lorenz, 2009; Mirgorodsky et al., 2010; Mirgorodsky et al., 2012; Ollivier et al., 2010). Hence, investigation of the source of contamination is important, since it can help to design a suitable remediation method for this area. Various material and complicated processes could influence the contamination, such as: AMD-leaching material, composition of Silurian-Ordovician slates, composition of underlying Quaternary sediments, or the recent leaching process. Since all these facts can influence the contamination, it is difficult to prove that one factor is the main source of contamination and reject another one. However, the aim of the study is to investigate if the compositions of the slates have an influence on the contamination.

3 Material and methods

3.1 Fuzzy clustering

Most fuzzy clustering algorithms are objective-function based (Bezdek 1984, 1987). In objective function-based clustering, each cluster is represented by a cluster centroid. The cluster centroid is computed by the clustering algorithm and may or may not appear in the dataset. The partitioning of the data points into different clusters is depending on the membership degree. It is computed based on the distance of the data points to the cluster centers. The closer a data point lies to the center of a cluster, the higher is its degree of membership to this cluster. Hence, the aim when dividing a data set into “c” number of clusters is to minimize the distances between the data points to the cluster centers while maximizing the degrees of membership.

The focus of this study is on fuzzy divisive hierarchical clustering and fuzzy hierarchical cross-clustering approach. Fuzzy divisive hierarchical clustering uses only cluster centers and a Euclidean

distance function in comparison to the cross-clustering algorithm which produces not only a fuzzy partition of the soil water and slate samples, but also a fuzzy partition of the considered metals.

3.2 Fuzzy divisive hierarchical clustering

In general, fuzzy clustering algorithm explained by Bezdek 1980 & 1981 and later on by Bezdek et al. (1987) and Sabin (1987). It can be formulated as follows: Let $X = \{x^1, \dots, x^n\} \subset \mathbf{R}^p$ be a finite set of feature vectors, where n is the number of objects (measurements), p is the number of the original variables, $x^j = [x_1^j, x_2^j, \dots, x_p^j]^T$ and $L = (L^1, L^2, \dots, L^c)$ be prototypes (supports) of which characterizes one of the c clusters composing the cluster substructure of the data set.

A partition of X into c fuzzy clusters will be performed by minimizing the objective function

$$J(P, L) = \sum_{i=1}^c \sum_{j=1}^n (A_i(x^j))^2 d^2(x^j, L^i) \quad (5)$$

Where $P = \{A_1, \dots, A_c\}$ is the fuzzy partition, $A_i(x^j) \in [0, 1]$ represents the membership degree of feature point x^j to cluster. $A_i, d(x^j, L^i)$ is the distance from a feature point x^j to the prototype of cluster A_i , defined by the Euclidean distance norm

$$d(x^j, L^i) = \|x^j - L^i\| = \left(\sum_{k=1}^p (x_k^j - L_k^i)^2 \right)^{1/2} \quad (6)$$

The optimal fuzzy set will be determined by using an iterative method where J is successively minimized with respect to A and L .

Supposing that, L is given, the minimum of the function $J(\bullet, L)$ is obtained for:

$$A_i(x^j) = \frac{C(x^j)}{\sum_{k=1}^c \frac{d^2(x^j, L^i)}{d^2(x^j, L^k)}}, i = 1, \dots, c \quad (7)$$

Where C is a fuzzy set from X and

$$C(x^j) = \sum_{i=1}^c A_i(x^j). \quad (8)$$

It is easy to observe that $C(x^j) \leq 1, j = 1, 2, \dots, n$. For a given P , the minimum of the function $J(P, \bullet)$ is obtained for:

$$L^i = \frac{\sum_{j=1}^n [A_i(x^j)]^2 x^j}{\sum_{j=1}^n [A_i(x^j)]^2}, i = 1, \dots, c \quad (9)$$

The above formula allows one to compute each of the p components of L^i (the center of the cluster i). Elements with a high degree of membership in cluster i (i.e., close to cluster i 's center) will contribute significantly to this weighted average, while elements with a low degree of membership (far from the center) will contribute almost nothing (Dumitrescu et al., 1994; Pop et al., 1995, 1996; Sarbu et al., 1993; Sarbu et al., 2007).

3.3 Fuzzy hierarchical cross-clustering algorithm

Building the classification binary tree is as follows: the nodes of the tree are labeled with a pair (C, D) , where C is a fuzzy set from a fuzzy partition of objects (samples in this study) and D is a fuzzy set from a fuzzy partition of characteristics (variables in this study). The root node corresponds to the pair (X, Y) . In the first step the two sub-nodes (A_1, B_1) and respectively (A_2, B_2) will be computed by using the cross-classification algorithm. It is important to note that these two nodes will be effectively built only if the fuzzy partitions $\{A_1, A_2\}$ and $\{B_1, B_2\}$ describe real clusters. For each of the terminal nodes of the tree it is tried to determine partitions having the form $\{A_1, A_2\}$ and $\{B_1, B_2\}$. In this way the binary classification tree is extended with two new nodes, (A_1, B_1) and (A_2, B_2) .

The ending of this processes is when no more structure of real cluster (either for the set of objects or for the set of characteristics) can be determined. The final fuzzy partitions will contain the fuzzy sets corresponding to the terminal nodes of the binary classification tree. This algorithm, which is called fuzzy hierarchical cross-clustering (FHCC), is a useful algorithm when it is desired to identify the relationships between different classes of samples and different classes of variables.

3.4 Samples and analysis

As described in section 1 and 2, the aim of this work is to investigate if the slates are the source of contamination in the study area by studying the relation between different types of data with slate samples. In order to find the most comparable samples with slates, data sets including groundwater and soil water samples were studied. Firstly, the classic multivariate study (cluster classification and PLS modeling) were used. Three slates leached (water leached, AMD leached and sulfuric acid leached) as well as groundwater and soil water samples were studied by mentioned statistical methods. The outcomes were not significant and no meaningful relation between the studied samples was detected. The only data set that has an acceptable significance was soil water samples in comparisson to water leached slates.

The present study concerns 53 soil water samples that were collected from 3 sampling points (MF1, MF2, MF3) from three depths: 30 cm, 60 cm and 100 cm below the surface level (237, 207, 257 m below sea level, respectively) (Fig.1). The sampling device employs the principle of tension-controlled under pressure; each sampling event took 14 to 15 days. During sampling, the water was collected in glass bottles from a depth of about 30 cm below the surface. Samples were collected between 2005 and 2007.

The Eh, pH, electrical conductivity (EC), and temperature were measured on-site using portable instruments (WTW, pH320; WTW, LF320; WTW, external thermocouple). In the field, all samples except those for HCO_3^- analysis were filtered using 0.45 μm cellulose acetate filters (Sartorius), with glass fiber pre-filters (Sartorius). Samples for element analysis were acidified with HNO_3 (65%, subboiled) to $\text{pH} < 2$. All samples were kept cool (6°C) until analysis. The analysis of Al, Ca, Fe, Mg, Mn and Na were performed with ICP-OES (Spectroflame, Spectro); Ba, Cd, Co, Cu, Li, Pb, REE (La-Lu), Sc, Th, U, Y, and the analysis of Zn with ICP-MS (until 2007: PQ3-S, Thermo Elemental, subsequently: X-Series II, Thermo Fisher Scientific).

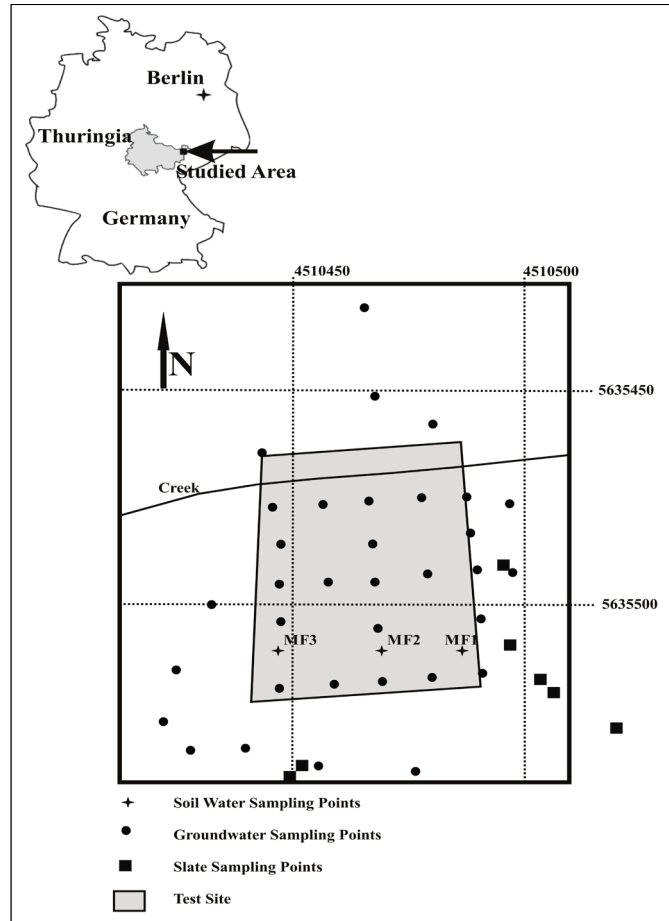


Fig. 1 Sampling locations, 174 groundwater from the test site, 10 Slate samples from the test site and surrounding area, and 53 soil water samples from 3 sampling location at 3 different depths (30, 60 and 100 cm)

In order to check the accuracy of the analytical results, the instrument drift was monitored and corrected by using Be, Ru and Re as internal standards. Each sample was measured three times. First, an outlier test is performed on each of the three runs (Grubbs test, 90% significance, criterion 1.15). Then, for the remaining runs, the mean and standard deviation are calculated. Furthermore, analytical quality is checked by the use of standard reference materials: the standard reference materials SCREE and PPREE are used for the analysis of the REEs (Verplanck et al., 2001). For other heavy metals, the standard reference materials SPS-SW2 (LGC standards) and NIST1643e (LGC standards) were used. Only values between 90-110 % for these reference materials are accepted. Detection limits are calculated according to the "3 sigma criterion" (Kaiser, 1965).

$$DL \equiv \text{Dil.Fact.} \cdot 3 \cdot \sqrt{\frac{2s_0}{S}} \quad (10)$$

Where $S = I/c$; Dil. Fact. is the dilution factor, s_0 = standard deviation, S = sensitivity, I = intensity, c = concentration.

The slates were studied more in detail using a total of 10 slate samples, collected from the former Gessenhalde and surrounding area. The samples were kept in 40°C for five days. The weathered and non-weathered portions were separated, and the non-weathered portion was crushed using a hammer. A plastic sieve was used to fraction the samples into two different sizes from 0.63 to 2 mm (middle size) and from 2 mm to 2 cm (coarse size). Furthermore, the slate samples were leached with water (Ultrapure water was obtained with a Purelab Plus system from USF Elga Seral (Ransbach-

Baumbach, Germany)). The amount of three grams of the milled material were mixed with 30 ml of leachate (water) and shaken in overhead shaker for 24 hours (ELU safety lock, Edmund Bühler). The samples were centrifuged for 15 min., and acidified with nitric acid. The elements were analyzed as it was described above for soil water samples. For accuracy, both samples type (soil water and slate samples) were treated like the groundwater as described in Introduction section. Table 1 presents the result of the analysis.

Table 1 The minimum, maximum, and standard deviation of the analytical results for the studied samples: 53 soil water samples and 10 slate samples leached with water

	Soil Water			water-leached slate		
	Min.	Max	Stdev.	Min.	Max	Stdev.
Al	7.87	294.30	73.10	11.32	661.58	213.44
Ba	5.00	46.67	10.60	0.34	1.98	0.48
Ca	195.67	573.33	71.85	13.26	2553.86	804.48
Cd	16.67	403.00	78.36	0.20	0.98	0.28
Ce	81.63	1183.67	226.86	0.12	0.90	0.32
Co	227.67	9605.00	1793.06	0.18	3.48	1.04
Cu	150.00	9864.50	3031.79	0.25	21.42	6.97
Dy	17.53	452.67	99.47	0.10	0.60	0.18
Er	10.11	269.00	60.42	0.11	0.90	0.25
Eu	3.39	80.47	18.05	0.10	0.70	0.21
Fe	0.00	500.10	90.30	12.76	3488.90	1108.71
Gd	18.90	465.00	104.28	0.10	0.70	0.20
Ho	3.61	95.73	21.19	0.10	0.60	0.18
La	9.20	153.07	32.15	0.16	0.90	0.28
Li	-	-	-	0.16	2.21	0.74
Lu	1.27	33.23	7.94	0.10	0.20	0.05
Mg	37.33	1369.50	249.34	9.69	667.88	198.66
Mn	4.77	559.00	97.45	0.24	2.78	0.81
Na	1.80	31.53	7.86	4.75	122.49	36.18
Nd	35.20	828.00	181.54	0.11	1.12	0.40
Ni	1286.33	36690.00	6638.88	1.42	18.66	5.15
Pb	0.24	18.60	2.66	0.13	3.67	1.07
Pr	6.20	143.03	30.28	0.10	0.60	0.15
Sc	1.03	77.30	14.94	0.12	1.00	0.26
Sm	11.73	278.33	62.87	0.10	0.90	0.23
Sr	133.00	990.00	233.26	0.19	41.45	12.56
Tb	2.86	71.73	15.76	0.10	0.50	0.16
Th	0.02	29.00	5.39	0.18	0.65	0.22
Ti	0.16	13.70	3.76	-	-	-
Tm	1.33	34.77	8.09	0.10	0.20	0.04
U	34.45	6676.50	1312.37	0.14	0.56	0.16
Y	113.00	3326.67	670.54	0.16	1.55	0.53
Yb	8.40	216.33	50.41	0.10	0.60	0.19

4 Results and discussion

Based on the outcomes of classical, statistical methods, the only possible relationship between the studied samples is the one between slate samples (water leached) and soil water samples. However, these results were not statistically significant (Figure 2 a, b). For this reason, fuzzy methods were used for the data set in order to better investigate possible relations in more detail and hopefully produce clusters with higher statistical significance.

Sub-cluster A2 includes A21, and A22 which is a representative of slate samples and sampling point MF1 and MF2, with a few samples from MF3. It is worth noting in Table 2 that the slate samples clustered together with the majority of samples collected at sampling points MF1 and MF2. It is also observed that there is a similarity among the slate samples, and a large dissimilarity among the samples collected at sampling point MF3. These samples (MF3) are classified into two clusters in accordance with their sampling depth. The samples from the depth of 100 cm are locating in one cluster and the remaining samples from MF3 are locating the other clusters. The scatter plots that are presented in Figures 3a-b support the previous statements very well.

The final clustering obtained by applying fuzzy, hierarchical clustering of the 36 variables, with auto-scaling of data, are shown in Table 3. The variables were divided into two main clusters A1 and A2. The cluster A2 includes five sub-clusters (A211, A212, A221, A222). The cluster, A1 is included only one variable, Ni. The sub-clusters A211 and A212 include Al, Ca, Co, Fe, K, Mg, Sr and U. The clusters A22, and the cluster A222 contain Ba, Cd, Cr, Cs, Li, Mn, Na, Pb, Sc, Th, Ti, Y and REE. The statements above are very well-supported by the patterns depicted in Figure 4a-b.

Table 2 Fuzzy divisive hierarchical clustering of samples (maximum membership degree with bold)

Sample Name	Code	Clusters						
		A ₁₁₁₁	A ₁₁₁₂	A ₁₁₂₁	A ₁₁₂₂	A ₁₂	A ₂₁	A ₂₂
MF1/30- 06.09.2005	1	0.01	0.01	0.02	0.01	0.14	0.56	0.24
MF1/60- 06.09.2005	2	0.00	0.00	0.01	0.00	0.05	0.72	0.21
MF1/100- 06.09.2005	3	0.00	0.01	0.01	0.00	0.07	0.72	0.18
MF2/30- 06.09.2005	4	0.00	0.00	0.00	0.00	0.03	0.78	0.18
MF2/60- 06.09.2005	5	0.00	0.00	0.01	0.00	0.05	0.77	0.17
MF2/100- 06.09.2005	6	0.00	0.00	0.01	0.00	0.04	0.70	0.24
MF3/30- 06.09.2005	7	0.04	0.05	0.05	0.05	0.35	0.24	0.22
MF3/60- 06.09.2005	8	0.06	0.11	0.06	0.07	0.54	0.08	0.07
MF3/100- 06.09.2005	9	0.56	0.08	0.03	0.03	0.13	0.09	0.07
MF1/60- 09.18.2005	10	0.00	0.01	0.01	0.01	0.06	0.74	0.17
MF1/100- 09.18.2005	11	0.00	0.01	0.01	0.01	0.10	0.68	0.18
MF2/30- 09.18.2005	12	0.00	0.00	0.00	0.00	0.03	0.84	0.12
MF2/60- 09.18.2005	13	0.00	0.00	0.00	0.00	0.04	0.82	0.13
MF2/100- 09.18.2005	14	0.00	0.00	0.00	0.00	0.03	0.82	0.14
MF3/30- 09.18.2005	15	0.01	0.01	0.01	0.01	0.14	0.45	0.38
MF3/60- 09.18.2005	16	0.02	0.04	0.03	0.03	0.76	0.07	0.06
MF3/100- 09.18.2005	17	0.00	0.65	0.06	0.07	0.13	0.05	0.04
MF1/30- 01.09.2006	18	0.00	0.00	0.01	0.00	0.06	0.73	0.18
MF1/60- 09.18.2005	19	0.00	0.00	0.00	0.00	0.02	0.78	0.19
MF1/100- 09.18.2005	20	0.00	0.00	0.00	0.00	0.04	0.77	0.19

MF2/30- 09.18.2005	21	0.00	0.00	0.00	0.00	0.02	0.63	0.34
MF2/60- 09.18.2005	22	0.00	0.00	0.00	0.00	0.02	0.78	0.17
MF2/100- 09.18.2005	23	0.00	0.00	0.00	0.00	0.02	0.79	0.18
MF3/30- 09.18.2005	24	0.01	0.01	0.01	0.01	0.14	0.41	0.41
MF3/60- 09.18.2005	25	0.01	0.02	0.02	0.02	0.59	0.20	0.14
MF3/100- 09.18.2005	26	0.02	0.03	0.53	0.02	0.23	0.10	0.07
MF1/30- 05.28.2006	27	0.00	0.00	0.01	0.00	0.05	0.58	0.35
MF1/60- 05.28.2006	28	0.00	0.00	0.01	0.00	0.05	0.39	0.55
MF1/100- 05.28.2006	29	0.00	0.01	0.01	0.00	0.06	0.67	0.25
MF2/30- 05.28.2006	30	0.00	0.00	0.00	0.00	0.03	0.58	0.38
MF2/60- 05.28.2006	31	0.00	0.00	0.00	0.00	0.02	0.72	0.26
MF2/100- 05.28.2006	32	0.00	0.00	0.00	0.00	0.02	0.73	0.25
MF3/30- 05.28.2006	33	0.01	0.02	0.02	0.02	0.27	0.34	0.32
MF3/60- 05.28.2006	34	0.00	0.00	0.01	0.00	0.09	0.51	0.37
MF3/100- 05.28.2006	35	0.03	0.04	0.32	0.099	0.24	0.15	0.11
MF1/30- 11.13.2006	36	0.00	0.00	0.01	0.00	0.06	0.73	0.19
MF1/60- 11.13.2006	37	0.00	0.00	0.01	0.00	0.04	0.76	0.18
MF1/100- 11.13.2006	38	0.00	0.00	0.00	0.00	0.04	0.8	0.15
MF2/30- 11.13.2006	39	0.00	0.00	0.01	0.00	0.05	0.74	0.18
MF2/60- 11.13.2006	40	0.00	0.00	0.01	0.01	0.06	0.74	0.18
MF2/100- 11.13.2006	41	0.00	0.00	0.00	0.00	0.03	0.79	0.17
MF3/30- 11.13.2006	42	0.02	0.03	0.03	0.03	0.39	0.26	0.24
MF3/60- 11.13.2006	43	0.01	0.03	0.02	0.03	0.72	0.11	0.08
MF3/100- 11.13.2006	44	0.06	0.07	0.00	0.57	0.16	0.08	0.06
MF1/30- 09.23.2007	45	0.00	0.01	0.01	0.01	0.09	0.43	0.44
MF1/60- 09.23.2007	46	0.00	0.00	0.01	0.00	0.05	0.32	0.61
MF1/100- 09.23.2007	47	0.00	0.01	0.01	0.01	0.09	0.58	0.28
MF2/30- 09.23.2007	48	0.00	0.00	0.00	0.00	0.02	0.61	0.36
MF2/60- 09.23.2007	49	0.00	0.00	0.01	0.00	0.04	0.71	0.23
MF2/100- 09.23.2007	50	0.00	0.00	0.00	0.00	0.03	0.62	0.34
MF3/30- 09.23.2007	51	0.00	0.01	0.0	0.01	0.12	0.36	0.50
MF3/60- 09.23.2007	52	0.00	0.01	0.01	0.01	0.09	0.46	0.42
MF3/100- 09.23.2007	53	0.44	0.10	0.04	0.07	0.16	0.10	0.08
S1	54	0.01	0.01	0.01	0.01	0.08	0.20	0.70
S2	55	0.01	0.01	0.01	0.01	0.09	0.22	0.64

S3	56	0.02	0.02	0.03	0.02	0.18	0.30	0.42
S4	57	0.01	0.02	0.02	0.02	0.15	0.34	0.43
S5	58	0.01	0.01	0.01	0.01	0.09	0.25	0.62
S6	59	0.01	0.01	0.01	0.01	0.07	0.18	0.72
S7	60	0.01	0.02	0.03	0.02	0.16	0.32	0.43
S8	61	0.01	0.02	0.02	0.02	0.13	0.27	0.53
S9	62	0.01	0.01	0.01	0.01	0.10	0.24	0.61
S10	63	0.01	0.02	0.02	0.02	0.15	0.26	0.52

Table 3 Fuzzy divisive hierarchical clustering of variables (maximum membership degree in bold)

Analytical parameters	code	Clusters				
		A_1	A_{211}	A_{212}	A_{221}	A_{222}
Al	1	0.01	0.10	0.50	0.19	0.20
Ba	2	0.00	0.01	0.00	0.00	0.98
Ca	3	0.06	0.23	0.34	0.20	0.17
Cd	4	0.00	0.00	0.00	0.00	0.99
Co	5	0.07	0.42	0.12	0.24	0.14
Cr	6	0.00	0.01	0.00	0.01	0.98
Cs	7	0.00	0.01	0.01	0.01	0.97
Cu	8	0.19	0.29	0.17	0.20	0.15
Fe	9	0.04	0.17	0.38	0.20	0.20
K	10	0.05	0.12	0.46	0.19	0.19
Li	11	0.00	0.19	0.05	0.57	0.19
Mg	12	0.03	0.10	0.55	0.17	0.14
Mn	13	0.00	0.00	0.00	0.01	0.98
Na	14	0.02	0.19	0.19	0.28	0.33
Ni	15	0.10	0.00	0.00	0.00	0.00
Pb	16	0.00	0.01	0.01	0.01	0.97
Sc	17	0.00	0.01	0.01	0.00	0.98
Sr	18	0.01	0.36	0.10	0.35	0.17
Th	19	0.00	0.01	0.01	0.01	0.97
Ti	20	0.00	0.01	0.01	0.01	0.97
U	21	0.02	0.38	0.11	0.32	0.18

Y	22	0.00	0.30	0.05	0.55	0.10
REE	37	0.02	0.57	0.02	0.31	0.08

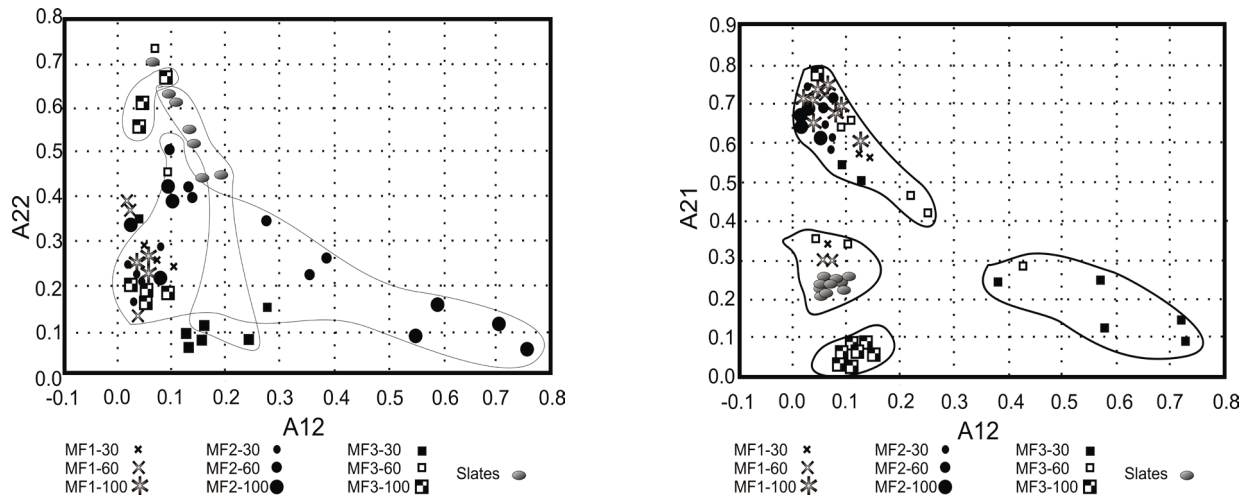


Fig. 3 Scatter plots of the membership degrees of sub-clusters of A1 and A2 by fuzzy hierarchical clustering; (a) is the similarities between sub-clusters A12 and A22; (b) is the similarities between sub-clusters A12 and A21. It is observed that there is a large similarity between the slate samples, and a large difference between the samples collected in sampling point MF3 at the depth of 100 cm and the rest of samples.

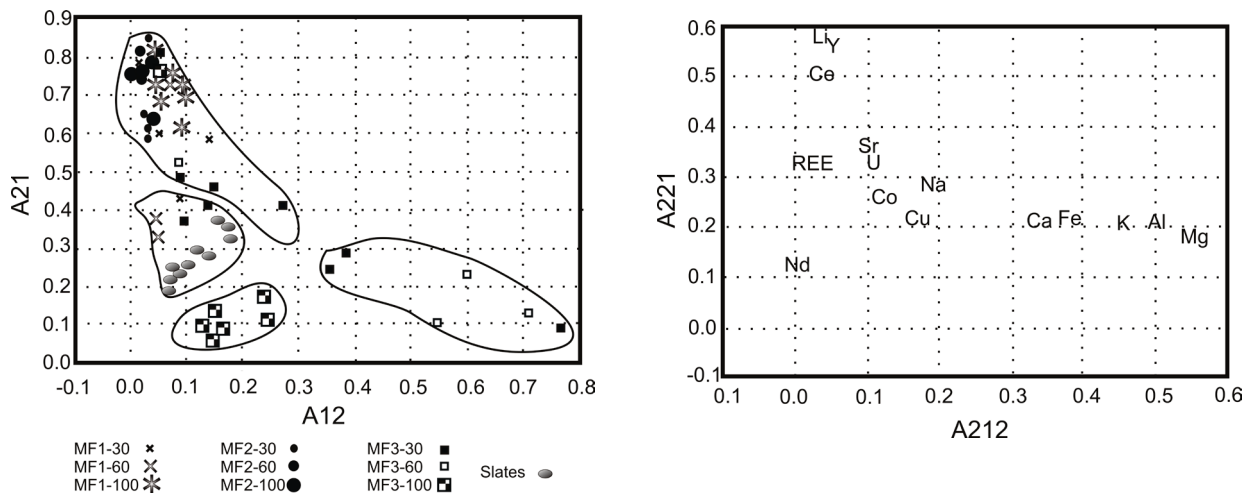


Fig.4 Scatter plots of the membership degrees of sub-clusters A2 based on the partitions obtained by fuzzy hierarchical clustering of the analytical parameters of both data sets. (a) is the similarities between sub-cluster A21 and A12; (b) is the similarities between sub-clusters A221 and A212

4.2 Fuzzy hierarchical cross-clustering

This method was applied to the variables (element concentrations including concentration of REE) of the 63 samples (slates samples and soil water samples), along with the application of auto-scaled data. The results obtained are presented in Table 4. Similar to the fuzzy hierarchical method, the variables and samples were divided into two main clusters A1 and A2. These two main clusters are included four sub-clusters (A11, A12, A21 and A22). The variables that are associated to sub-cluster

A11 are Al and U. This cluster includes the majority of the samples collected at sampling point MF3, depth 100 cm. The elements in sub-cluster A12 are Cd, Co, Cu, Li, Ni, Y, and REE. This cluster includes samples collected MF3/30, MF3/60 (sampled at different times). The third sub-cluster A21 includes Cr, Cs, Fe, K, Pb, Sc, Th, Ti, and the majority of samples collected at sampling point MF1 and MF2 (MF1/30, MF1/60, MF1/100, MF2/30, MF2/60, MF2/100 that were sampled at different times). The last sub-cluster, A22, includes Ba, Ca, Mg, Mn, Na, and Sr. This cluster contains all the slate samples (MF3/30, MF3/60, MF1/100, MF1/30, MF1/60, S1 to S10). Figure 5a-b supports these observations.

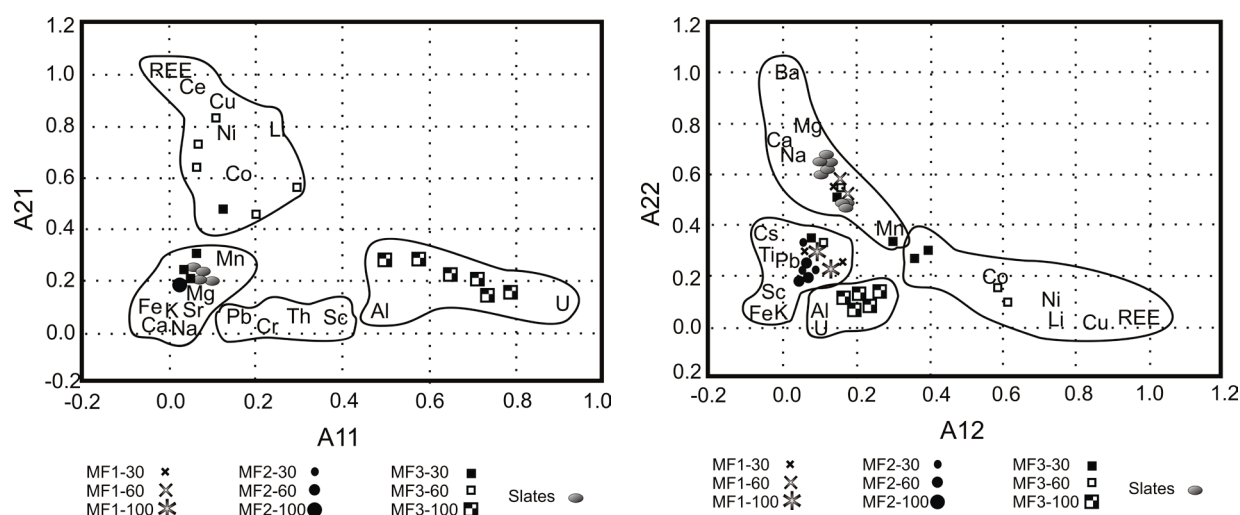


Fig. 5 Scatter plots of the membership degrees of sub-clusters of A1 and A2 based on the partitions obtained by fuzzy hierarchical cross-clustering. (a) is the similarities between sub-clusters A11 and A21; (b) is the similarities between sub-clusters A12 and A22. The partitioning of the samples in different clusters is very similar with those obtained by the fuzzy hierarchical clustering algorithm

Table 4 Membership degrees of samples and variables to the clusters by fuzzy cross-clustering method. The bold numbers show the highest membership degree for each sub-cluster

Sample Name / Analytical parameters	Code	Clusters			
		A ₁₁	A ₁₂	A ₂₁	A ₂₂
MF1/30- 06.09.2005	1	0.05	0.14	0.56	0.24
MF1/60- 06.09.2005	2	0.02	0.05	0.72	0.21
MF1/100- 06.09.2005	3	0.02	0.07	0.72	0.18
MF2/30- 06.09.2005	4	0.01	0.03	0.78	0.18
MF2/60- 06.09.2005	5	0.02	0.05	0.76	0.17
MF2/100- 06.09.2005	6	0.02	0.04	0.70	0.24
MF3/30- 06.09.2005	7	0.19	0.35	0.24	0.22
MF3/60- 06.09.2005	8	0.30	0.54	0.08	0.07
MF3/100- 06.09.2005	9	0.71	0.13	0.90	0.07
MF1/60- 09.18.2005	10	0.02	0.06	0.74	0.17
MF1/100- 09.18.2005	11	0.03	0.10	0.68	0.18
MF2/30- 09.18.2005	12	0.01	0.03	0.84	0.12
MF2/60- 09.18.2005	13	0.01	0.04	0.82	0.13

Sample Name / Analytical parameters	Code	Clusters			
		A ₁₁	A ₁₂	A ₂₁	A ₂₂
MF2/100- 09.18.2005	14	0.01	0.03	0.82	0.14
MF3/30- 09.18.2005	15	0.04	0.14	0.45	0.38
MF3/60- 09.18.2005	16	0.11	0.76	0.07	0.06
MF3/100- 09.18.2005	17	0.78	0.13	0.05	0.04
MF1/30- 01.09.2006	18	0.02	0.06	0.73	0.18
MF1/60- 09.18.2005	19	0.01	0.02	0.78	0.19
MF1/100- 09.18.2005	20	0.01	0.04	0.77	0.19
MF2/30- 09.18.2005	21	0.01	0.02	0.63	0.34
MF2/60- 09.18.2005	22	0.01	0.02	0.78	0.19
MF2/100- 09.18.2005	23	0.01	0.02	0.79	0.18
MF3/30- 09.18.2005	24	0.03	0.14	0.41	0.41
MF3/60- 09.18.2005	25	0.06	0.59	0.20	0.13
MF3/100- 09.18.2005	26	0.60	0.23	0.10	0.07
MF1/30- 05.28.2006	27	0.02	0.05	0.58	0.35
MF1/60- 05.28.2006	28	0.02	0.05	0.39	0.55
MF1/100- 05.28.2006	29	0.02	0.06	0.68	0.25
MF2/30- 05.28.2006	30	0.01	0.03	0.58	0.38
MF2/60- 05.28.2006	31	0.01	0.02	0.72	0.26
MF2/100- 05.28.2006	32	0.013	0.02	0.73	0.25
MF3/30- 05.28.2006	33	0.062	0.27	0.34	0.32
MF3/60- 05.28.2006	34	0.021	0.10	0.51	0.37
MF3/100- 05.28.2006	35	0.50	0.24	0.15	0.11
MF1/30- 11.13.2006	36	0.02	0.06	0.73	0.19
MF1/60- 11.13.2006	37	0.02	0.0	0.76	0.18
MF1/100- 11.13.2006	38	0.01	0.04	0.80	0.15
MF2/30- 11.13.2006	39	0.02	0.05	0.75	0.18
MF2/60- 11.13.2006	40	0.02	0.06	0.74	0.18
MF2/100- 11.13.2006	41	0.01	0.03	0.80	0.16
MF3/30- 11.13.2006	42	0.11	0.39	0.26	0.24
MF3/60- 11.13.2006	43	0.09	0.72	0.11	0.09
MF3/100- 11.13.2006	44	0.70	0.16	0.08	0.06
MF1/30- 09.23.2007	45	0.04	0.09	0.44	0.44
MF1/60- 09.23.2007	46	0.02	0.05	0.32	0.61
MF1/100- 09.23.2007	47	0.03	0.09	0.60	0.28
MF2/30- 09.23.2007	48	0.01	0.02	0.61	0.36
MF2/60- 09.23.2007	49	0.01	0.04	0.71	0.26
MF2/100- 09.23.2007	50	0.01	0.03	0.62	0.34
MF3/30- 09.23.2007	51	0.03	0.11	0.36	0.50
MF3/60- 09.23.2007	52	0.02	0.09	0.46	0.42

Sample Name / Analytical parameters	Code	Clusters			
		A ₁₁	A ₁₂	A ₂₁	A ₂₂
MF3/100- 09.23.2007	53	0.65	0.16	0.10	0.08
S1	54	0.03	0.08	0.20	0.69
S2	55	0.04	0.09	0.22	0.64
S3	56	0.09	0.18	0.30	0.42
S4	57	0.08	0.15	0.34	0.43
S5	58	0.04	0.09	0.24	0.62
S6	59	0.03	0.07	0.18	0.72
S7	60	0.08	0.16	0.32	0.43
S8	61	0.06	0.13	0.27	0.53
S9	62	0.04	0.10	0.24	0.61
S10	63	0.07	0.15	0.26	0.52
Al	Al	0.49	0.05	0.37	0.08
Ba	Ba	0.01	0.00	0.02	0.98
Ca	Ca	0.01	0.01	0.28	0.70
Cd	Cd	0.04	0.95	0.00	0.00
Co	Co	0.16	0.58	0.10	0.16
Cr	Cr	0.19	0.02	0.74	0.04
Cs	Cs	0.00	0.007	0.58	0.42
Cu	Cu	0.14	0.85	0.01	0.00
Fe	Fe	0.00	0.00	0.96	0.04
K	K	0.03	0.02	0.77	0.18
Li	Li	0.25	0.74	0.00	0.00
Mg	Mg	0.06	0.08	0.12	0.72
Mn	Mn	0.14	0.29	0.20	0.36
Na	Na	0.056	0.032	0.27	0.64
Ni	Ni	0.14	0.73	0.05	0.08
Pb	Pb	0.14	0.06	0.44	0.36
Sc	Sc	0.38	0.03	0.52	0.07
Sr	Sr	0.09	0.06	0.21	0.64
Th	Th	0.29	0.03	0.62	0.05
Ti	Ti	0.01	0.00	0.70	0.28
U	U	0.91	0.07	0.02	0.01
Y	Y	0.02	0.98	0.00	0.00
REE	REE	0.00	0.99	0.00	0.00

5 Conclusions

As is shown in Fig. 6, the AMD that was used for leaching uranium has infiltrated the clay material that was used under the heap to contain the leachate, as well as to the underlying glacial sediments;

hence, contaminants have reached the groundwater. This first phase of infiltration caused a secondary mineralization and thus released contamination to the groundwater (Fig. 6.1). After remediation and the removal of the leaching heap, the pit was left open and thus rainwater led to a second phase of infiltration forming above the groundwater level (Fig. 6.2). More heavy metals were released from the ponded water and mineralized. Due to the remediation, the contaminated glacial sediment was excavated down to 10m and the area was filled with non-contaminated top soil and groundwater level was changed, the third phase of infiltration was formed (Fig. 6.3).

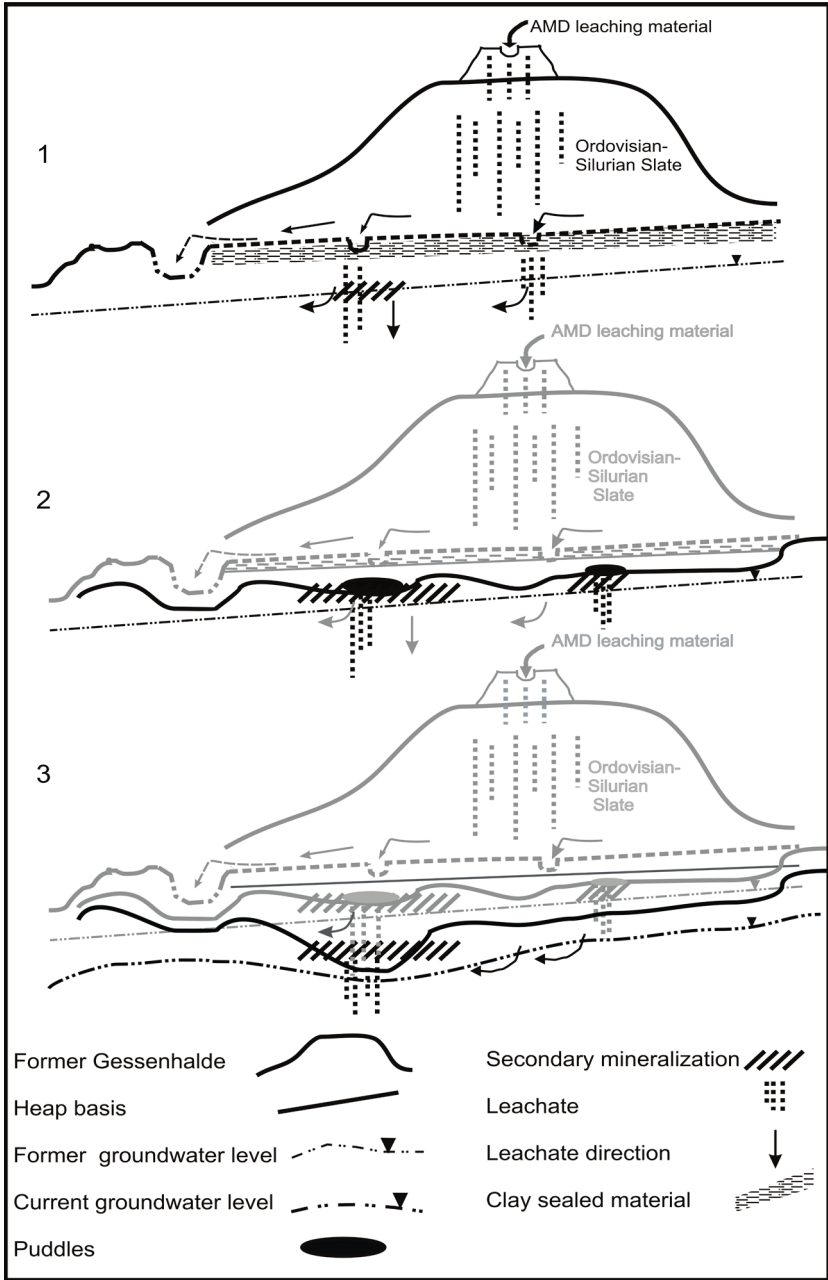


Fig. 6 Three phases of infiltration: (1) first phase of infiltration that caused a secondary mineralization and thus released contamination to the groundwater; (2) second phase of infiltration forming above the groundwater level; (3) third phase of infiltration caused by the ponded water

By means of fuzzy clustering, the relation between the leached Ordovician-Silurian slates samples and soil water samples was identified: fuzzy clustering of the 53 soil water and 10 slate samples based

on concentration of some heavy metals allow an objective interpretation of their similarities and differences, respectively. It is very interesting to observe the similarity of slate samples and their relation with soil water samples considering the membership degrees. The slate samples are very similar and appear to be more close to the soil water samples collected in sampling point MF1 and MF2. The fuzzy cross-clustering approach allows the qualitative and quantitative identification of the variables (metal concentration) responsible for the observed similarities and differences between all the samples. Considering this aspect it is extremely useful to point out the strong association of sampling point MF3 with Cd, Co, Cu, Ni, Li, REE and U.

In closing, the obtained results clearly underline the efficiency of the fuzzy clustering algorithms for the comparison of different soil water and slate samples and allowing associating the most characteristic chemical elements to each sample cluster. However, defining the slates as the main source of contamination measured in soil water samples is not proved yet. It is more possible that the condition of the slates (weathered or non-weathered) and also the physicochemical situation of the media surrounding the slates have more influence of the contamination concentration.

Acknowledgments

The authors would like to acknowledge S. Wagner, M. Lonschinski for sampling, also I. Kamp, G. Weinzierl and G. Rudolph. This project is a part of the graduate research training group, GRK 1257/1 (Alteration and element mobility at the microbe-mineral interfaces), and Jena School for Microbial Communication (JSMC) and financially supported by the German Research Society (DFG).

References

- Bezdek J (1980). A convergence theorem for the fuzzy ISODATA clustering algorithms. *IEEE T. Pattern Anal.* 2 (1): 1–8.
- Bezdek J (1981). *Pattern recognition with fuzzy objective function algorithms*, Plenum Press. New York, United States of America.
- Bezdek J, Hathaway R, Sabin S, Tucker W (1987). Convergence theory for fuzzy c-means: counter examples and repairs. *IEEE T. Syst. Man Cyb.* 17 (5): 873–877.
- Bhattacharya A, Routh J, Jacks G, Bhattacharya P, Mörtz M (2006). Environmental assessment of abandoned mine tailings in Adak, Västerbotten district (northern Sweden). *App. Geo.* 21: 1760–1780.
- Bozau E, Bechstedt T, Frieze K, Frömmichen R, Herzsprung P, Koschorreck M, Meier J, Völkner C, Wendt-Potthoff K, Wieprecht M, Geller W (2007). Biotechnological remediation of an acidic pit lake: modeling the basic processes in a mesocosm experiment. *J. Geochem. Explor.* 92 (2-3):212-221.
- Carlsson E, Büchel G (2005). Screening of residual contamination at former uranium heap leaching site, Thuringia, Germany. *Chem. Erde. Geochem.* 65: 75-95.
- Chopin EIB, Alloway BJ (2007). Distribution and mobility of trace elements in soils and vegetation around the mining and smelting areas of Tharsis, Riotinto and Huelva, Iberian Pyrite Belt, SW Spain. *Water Air Soil Poll.* 182: 245–261.
- Demicco R, Klir G (2003). *Fuzzy logic in Geology*. Elsevier.
- Dumitrescu D, Sârbu C, Pop HF (1994). A fuzzy divisive hierarchical clustering algorithm for the optimal choice of sets of solvent systems. *Anal. Lett.* 27:1031–1054
- Elias RW, Gulson B (2003). Overview of lead remediation effectiveness. *Sci. Total. Environ.* 303 (1-2):1-13.
- Franklin MR, Fernandes HM (2011). Identifying and overcoming the constraints that prevent the full implementation of decommissioning and remediation programs in uranium mining sites. *J. Environ. Radioactiv.* In press.
- Gomes MEP, Favas PJC (2006). Mineralogical controls on mine drainage of the abandoned Ervedosa tin mine in north-eastern Portugal. *Appl. Geochem.* 21: 1322–1334.
- Grawunder A (2010). *Hydrogeochemistry of rare earth elements in an acid mine drainage influenced area*. Doctoral Thesis, Friedrich Schiller University of Jena.
- Hadley R, Snow D (1974). *Water resources and problems related to mining*. American water resource association.
- Hafeburg G (2007). *Studied on heavy metals resistance of bacterial isolates from a former uranium mining area*. Doctoral dissertation, Friedrich Schiller university of Jena. P:189.
- Horn K (2003). *Das Migrationsverhalten von Radon im Sanierungsgebiet der ostturingischen Uranbergbauregion*. Doctoral dissertation, Friedrich Schiller university of Jena. P:120 (In German).
- Hoson T, Su C C, Siringan F, Amano A, Shin-ichi Onodera S (2010). Effects of environmental regulations on heavy metal pollution decline in core sediments from Manila Bay. *Mar. Pollut. Bull.* 60: 780–785.
- Jakubick A, Jenk U, Kahnt R (2002). Modeling of mine flooding and consequences in the mine hydrogeological environment: Flooding of the Koenigstein mine, Germany. *Environ. Geol. Water. S.* 42(2): 222-234.

- Kaiser H (1965). Zum Problem der Nachweigrenzer. Workshop: Moderne Methoden der anorganischen Analyse, 05- 07.10. 1965. Düsseldorf, Germany (In German).
- Lange G (1995). Die Uranlagerstätte Ronneburg. Z. Geol. Wiss. 23: 517-526 (In German).
- Lee CH (2003). Assessment of contamination load on water, soil and sediment affected by the Kongjujeil mine drainage, Republic of Korea. Environ. Geo. 44: 501–515.
- Lonschinski M (2009). Schwermetalle im System Boden-Wasser-Pflanze mit Hinblick auf Fraktionierungsprozesse der Seltenen Erden Elemente auf der Aufstandsfläche der Gessenhalde im ehemaligen Uranbergbaurevier Ronneburg (Ostthüringen). Doctoral Thesis, Friedrich Schiller University of Jena (In German).
- Lorenz C (2009). Untersuchungen zum geomikrobiellen Schwermetalltransfer und simultanen Monitoring seltener Erden Elemente. Doctoral Thesis, Friedrich Schiller University of Jena, (In German).
- Mirgorodsky D, Ollivier D, Merten D, Büchel G, Willscher S, Wittig J, Jablonski L, Werner P (2010). Phytoremediation experiments on a slightly contaminated test field of a former uranium mining site. Proceedings of IMWA 2010: "Mine Water & Innovative Thinking", Sydney/Canada. 587-590.
- Moreno M, Oldroyd A, McDonald L, Gibbons W (2007). Preferential fractionation of trace metals-metalloids into PM10 re-suspended from contaminated gold mine tailings at Rodalquilar, Spain. Water Air Soil Poll. 179: 93–105.
- Nicholas MJ, Clift R, Azarpagic A, Walker FC, Porter DE (2000). Determination of best available techniques for integrated pollution prevention and control: A life cycle. Process. Saf. Environ. 78 (3): 193-203.
- Otte ML, Jacob DL (2008). Mine area remediation. Encyclopedia of Ecology. 2:397-2409.
- Pasalic S (2011). Untersuchung von Bodensubstraten im Bereich von Saugkerzenanlagen auf dem Testfeld Gessenwiese bei Ronneburg. Biological project module, Friedrich Schiller University of Jena, (In German).
- Pop HF, Dumitrescu D, Sârbu C (1995). A study of Roman pottery terra sigillata using hierarchical fuzzy clustering Anal. Chim. Acta. 310:269–279
- Pop HF, Sârbu C (1966). A new fuzzy regression algorithm. Anal. Chem. 68:771–780
- Rodríguez L, Ruiz E, Alonso-Azcárate J, Rincón J (2009). Heavy metal distribution and chemical speciation in tailings and soils around a Pb–Zn mine in Spain. J. Environ. Manage. 90:1106–1116.
- Rüger F, Dietel W (1998). Vier Jahrzehnte Uranbergbau um Ronneburg. Lapis. 7(8): 14-18, (In German).
- Sabin M (1987). Convergence and consistency of Fuzzy c-means/ISODATA algorithms. IEEEET Pattern Anal. 9(5).
- Sârbu C, Dumitrescu D, Pop HF (1993). Selection and optimal combination of solvent systems in TLC. Rev. Chim. 44:450–459.
- Sârbu C, Zehl K, Einax JW (2007). Fuzzy divisive hierarchical clustering of soil data using Gustafson-Kessel algorithm. Chemometr .Intell. Lab Syst. 86:121–129
- Verplanck PL, Antweiler RC, Nordstrom DK, Taylor H E (2001). Standard reference water samples for rare earth element determinations. Appl. Geochem. 16: 231-244
- Wagner S (2010). Geoschemische Untersuchungen an Paläozoischen Gesteinen aus dem Gebiet der ehemaligen Gessenhalde bei Ronneburg ehemaliges ostthüringisches Uranerzbergbaurevier. Diploma thesis. Friedrich Schiller University of Jena, (In German).

Wismut GmbH (1994). Proposal of remediation of Ronneburg site. Chemnitz, Germany (In German).

Manuscript 4

**Distribution of rare earth elements and other metals in a stratified acid pit lake
in black shales 45 years after mine closure**

Distribution of rare earth elements and other metals in a stratified acidic pit lake in black shales 45 years after mine closure

Stefan Karlsson¹ - Viktor Sjöberg¹ - Anahita Pourjabbar² - Anna Grandin¹ - Bert Allard¹

¹ Man-Technology-Environment Research Centre, Örebro University, Sweden

² Institute of Geoscience, Friedrich Schiller University, Jena, Germany

Abstract

Open pit mining of black shale for oil extraction was undertaken between 1942 and 1966 at Kvarntorp, Sweden; and the pits are now water filled. The shale is rich in sulphides (5-8%), so the leachates from it usually exhibit a pH range of 3 to 3.5. The shale contains a range of metals at concentrations in the range of 50 to 500 g/ton, including rare earth elements and uranium. Water samples were collected (profiles and transects) from some of the pit lakes and analyzed with respect to general hydrochemical parameters and metals. The obtained results indicate an incongruent release of metals from the shale. Heterogeneous distributions of trace metals, including the rare earth elements, were observed with concentrations that varied with depth and with the distance from water exposed shale horizons. This heterogeneity reflects pH, redox conditions, concentration of dissolved organic matter, the lake stratification, presence of adsorbing particulate matter and the formation of precipitates that accumulate on the bottom creating secondary metal rich sediments. The incongruent release of metals from the weathering shale is also demonstrated from the accumulation and distribution in secondary minerals that are precipitated directly on shale surfaces above the water level.

Key Words: Black Shales, Pit Lake, Rare Earth Elements, Hydrogeochemistry

1 Introduction

In Sweden the organic rich black shales are commonly known as alum-shales because of their content of alum ($K Al(SO_4)_2$), which was sought for during the 18th and 19th centuries for the production of dyes and rag paper (Eklund et al., 1995). Hence, in many regions where black shales are available, the remains of past operations are visible as piles of burnt shale. Production of hydrocarbons from the organic rich black shale (up to 18% organic carbon) in the Kvarntorp area in south central Sweden started during the Second World War. The shale was excavated in open pits, crushed, sieved and pyrolysed, and the volatile organic hydrocarbon fraction was recovered. The heated shale residues were put on a single pile together with a fine-grained shale fraction that was discarded since it was not suitable for pyrolysis. An alternative method for recovery was in-situ heating by inserting electrodes into the shale. The production continued until the mid 1960s, when it ceased because of the production costs (Dyni, 2006). In the later stage of the operation the shale was also extracted for uranium recovery, but only some 62 tonnes were produced by sulphuric acid leaching. The original shale as well as the different wastes has a high pollution potential because of the rather high concentrations of metals such as U, V, Ni and Mo in combination with pyrite and other acid generating metal sulphides. These elements, together with the REEs, make the materials economically attractive for future extraction from both the waste fractions as well as the pristine shale. Hence, it is essential to understand the chemical mechanisms behind metal release and redistribution in this complex setting.

The two pit lakes in this study receive water that has passed through untreated waste shale as well as the original strata in the surrounding soil. A fraction of the shallow groundwater that reaches one of the lakes has passed through a backfill of alkaline waste material from the production of construction materials (cement) from limestone. There are also some contributions of soil water from the adjacent municipal waste landfill. The depth of the deepest pit lake is roughly 30 m. One pit lake is thermally stratified during summers and winters but interrupted by turnovers in autumn and spring. These conditions provide an opportunity to study the impact of fundamental hydrochemical conditions on the metal distribution and compare the two lakes.

2 Materials and methods

2.1 The field site

Kvarntorp is located 20 km south of Örebro, in south central Sweden, 200 km from Stockholm. The former mining area covers about 8 km². The late- Cambrian, black-shale horizon has a thickness of 5 to 15 m. This horizon, usually, is exposed beneath an Ordovician limestone. However, at this site the black-shale horizon, exposed at the surface level due to faulting. Hence, it is easily accessible and well suited for open pit mining. Major mineral components of the shale are quartz (18-34%), illite (24-41%), K-feldspars (3-8%), chlorite (1-5%), calcite (1-4%), as well as pyrite (5-17%) and organic carbon (5-18%). The area is one of the most polluted sites in Sweden from emissions of sulphuric acid, metals and hydrocarbons, including PAHs, PCBs and dioxins, from the shale processing period (SWEKO VIAK, 2005).

The two pit lakes are shown in Fig. 1. Lake Pölen (S1) is a small acidic, shallow lake with a maximum depth of 8 m; lake Norrtorpsjön (S2) is a larger, deeper lake with a maximum depth of 30 m which receives alkaline leachates from the waste deposits north and west of the lake.

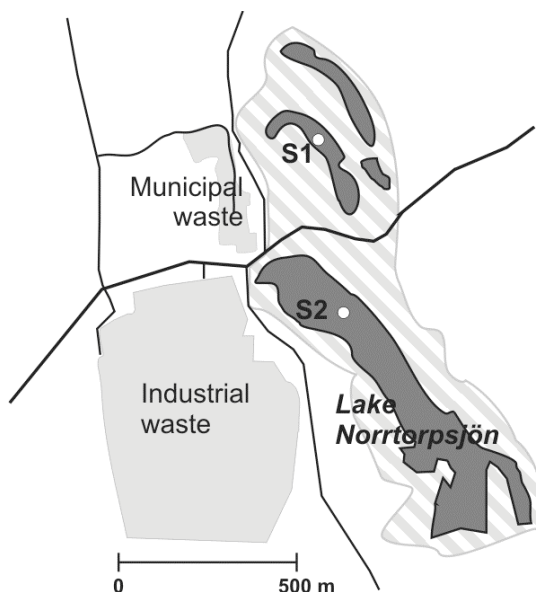


Fig. 1 Map of the area with the pit lakes. Lake Pölen (S1) is a small acidic, shallow lake with a maximum depth of 8 m; lake Norrtorpsjön (S2) is a larger, deeper lake with a maximum depth of 30 m which receives alkaline leachates from the waste deposits north and west of the lake

2.2 Sampling and analysis

The two pit lakes were sampled in mid-June and July 2011. On both occasions a conventional water sampler was used and the samples were poured into one-liter polypropylene bottles and transported to the laboratory for further processing. In lake Norrtorpsjön samples were taken at the depth of one meter (epilimnion), eight meter (thermocline), 12 meter (hypolimnion) and 27 meter (bottom water). In the lake Pölen, the samples were collected at the depth of one and three meter, because of its shallow depth. Samples for speciation (metals bound to humic and fulvic acids) were collected in a one-liter triplicate. DEAE anion-exchange resin was added on the field site. Sub-samples of the dissolved phase were collected after eight hours; and acidified with sub-boiled nitric acid.

Phase separations were made in a clean room, to minimize the risk for contamination. They were made by filtration through 47 mm polycarbonate membranes with defined pore sizes of 1.0, 0.4 and 0.2 μm , respectively. Plastic syringes were used, and the filtration was made below the filter clogging point to avoid the risk of secondary redistributions between the filter cake and the passing sample. The filtrates were then acidified and stored until analysis. Samples for “total” element analysis were only acidified with nitric acid. Electrical conductivity and pH were measured in-situ as well as in the laboratory, with conventional electrodes. Moreover, temperature profiles were measured. Principal anions (Br^- , Cl^- , F^- , NO_2^- , NO_3^- , PO_4^{3-}) were quantified by ion-chromatography using a Dionex AS12A column and a running buffer with 10.5 mM Na_2CO_3 /0.5 mM NaHCO_3 . Principal and selected metals were analyzed by ICP-MS (Agilent 7500 cx) operated with a Micromist nebulizer. The REEs were determined with the same instrument but with an ultrasonic nebulizer unit equipped with a dryer (CETAC U6000AT+) to minimize interferences from oxides and hydrides. All sample treatment and

metal analyses were performed in a clean room (class 10/100). Dissolved organic carbon (DOC) and inorganic carbon were quantified with a TOC-analyzer (Shimadzu TOC-V CPH).

The selected speciation procedure is based on exchange of anionic organic matter on a DEAE-cellulose adsorbent. The resin was cleaned and conditioned to match the sample pH and ionic strength. In the field an excess of the adsorbent was added as slurry to a one-litter sample bottle immediately after sampling. Triplicate samples were prepared for each depth. The bottles were shaken intermittently for one hour and transported to the laboratory where the exchanger was allowed to settle. A sub-volume (50 ml) was decanted and acidified to one-percent nitric acid and stored until analysis.

3 Results and discussion

3.1 General hydrochemistry

The profiles of temperature and electrical conductivity show that lake Norrtorpsjön was stratified at the sampling occasion, with the thermocline extending from eight to 11 m. Thus, the lake had in principle the three classical compartments. For the principal anions the common pattern was low concentrations in the epilimnion and a pronounced increase towards the thermocline (Fig. 2). In the hypolimnion the concentrations were rather constant, except for Cl^- that showed a decrease. For most anions, the concentrations are quite high (Table 1).

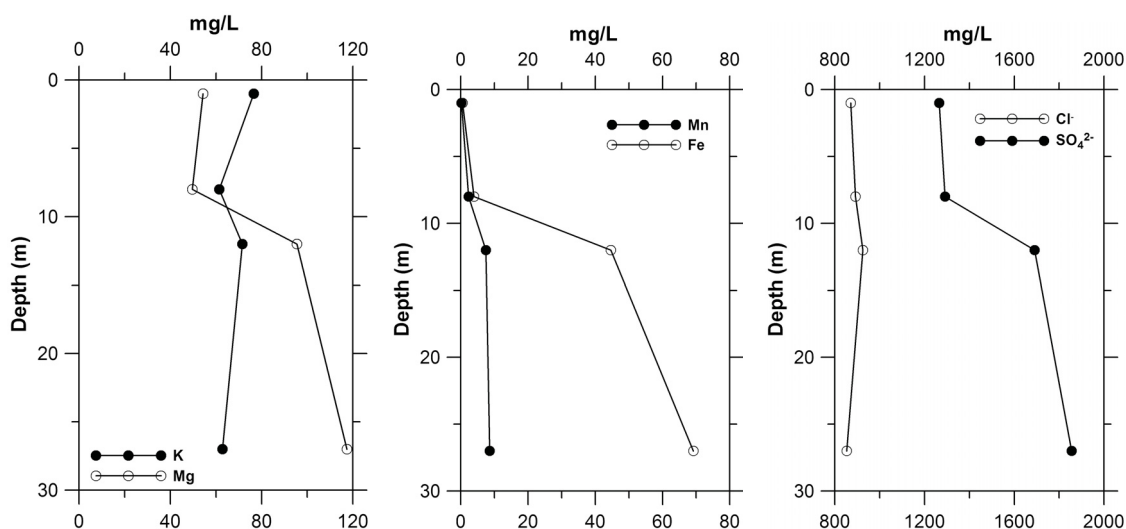


Fig. 2 Examples of depth profiles of lake Norrtorpsjön

It is also evident that there were different anion sources. High SO_4^{2-} concentrations are expected as a result of S^{2-} weathering of the shale, but this source cannot account for the high concentrations of Cl^- and NO_3^- as well as the presence of Br^- . The proximity to the municipal waste treatment site is a clear indication of the origin of these ions (Fig. 1). The elevated concentrations of fluoride are most likely of both shale and municipal waste origin. Lake Pölen was not deep enough to develop any thermocline, and the concentrations of anions were similar at the depth of one meter below the surface and 0.5 m above the bottom, respectively. This lake is not influenced by seepage water from the municipal waste. Its concentrations of notably F^- , Br^- and NO_3^- are low, while the contamination levels in Lake Norrtorpsjön are due to the diffuse intrusion of leachates from the municipal waste as well as from the alkaline industrial waste. Analyses of shallow groundwater and leachates from the

waste areas confirm this explanation. The alkaline waste in the tributary to Lake Norrtorpsjön has had a large impact on the pH regimes, as compared to Lake Pölen. In lake Norrtorpsjön there is a decline of pH range from 7.4 at the depth of one meter to 6.4 at the depth of eight 8 meter and further down to pH value of 5.9 at the depth of 27 m. This decrease of pH value is an indication of acid producing processes below the thermocline. This is in contrast to the constant pH value of 3.2 measured in the entire Lake Pölen. This impact is also reflected by the concentrations of inorganic carbon at ranges of 20-34 mg/l and 1.5 mg/l in Lake Norrtorpsjön and Lake Pölen, respectively. Dissolved organic carbon was below 0.01 mg/L in Lake Norrtorpsjön and slightly above 1.0 mg/l in Lake Pölen.

Table 1 General hydrochemical parameters (BD: below detection)

Site	Depth m	pH	El.Cond MS/cm	IC mg/L	TOC mg/L	F ⁻ mg/L	Cl ⁻ mg/L	Br ⁻ mg/L	SO ₄ ²⁻ mg/L	NO ₃ ⁻ mg/L
L. Norrtorpsjön	1	7.42	1.978	19.7	<0.01	0.88	871	1.52	1265	3.18
	8	6.76	2.135	24.3	<0.01	1.03	892	1.47	1291	6.21
	12	6.42	2.512	33.8	<0.01	1.38	925	2.76	1690	8.68
	27	5.97	2.947	27.4	<0.01	1.39	853	2.55	1855	BD
L.Pölen	1	3.18	1.796	1.52	<0.01	0.63	792	BD	1406	BD
	3	3.18	1.834	1.05	<0.01	0.64	819	BD	1412	BD

The impact of the different wastes in the tributary is reflected by the concentration levels as well as depth profiles of the principal cations (Tables 1 and 2; Fig. 2). The concentrations of Na (60-80 mg/L) and K (25-40 mg/L) are highly elevated, and definitely higher than expected from only the shale (Karlsson, 2011) but rather common in this calcite dominated environment. The highest concentrations of the two mentioned elements were measured at the depth of one meter. Calcium concentrations are also high (450-550 mg/L). Both lakes are close to saturation with respect to gypsum (S.I. -0.02). None of these elements were retained by the filters or by the anion-exchangers. The reason could be due to fact that these elements are mainly present as dissolved cations. Magnesium, Fe and Mn exhibited almost identical depth profiles, with low concentrations in the epilimnion and an abrupt increase in the thermocline (Fig. 2). Such concentration profiles would indicate formation of settling particles above the thermocline, followed by dissolution/release below it. lake Pölen is close to equilibrium with respect to ferrihydrite (S.I. 0.08). In lake Norrtorpsjön ferrihydrite was oversaturated above the thermocline (S.I. 4.9) but under saturation limiting concentrations below it (S.I. -4.4). Evidently the redox potential was not low enough to induce precipitation of Fe-SO₄²⁻, as indicated by an oxygen saturation of 2-5% in the bottom water. Here, there is a lower concentration of these elements in the filtrates 0.40 and 0.20 µm, respectively. The appearance of the particulate/colloidal phase corresponds to the depths where the oxygen saturation goes down from 85-90% to less than 20%. This corresponds to measured redox potentials in the range of 0.40 mV to -150 mV, respectively. Consequently, reductive dissolution would be feasible below the thermocline. The absence of filterable fractions above the thermocline indicates two possible mechanisms. In this oxic region, large particles of Fe-(hydro)oxides are formed; and quickly settled below the thermocline under the impact of gravity. It is also possible that photo reduction maintains Fe in its divalent state in the photic zone and formation of particles takes place only during

the night (not included in this sampling). Irrespective to the mechanism, it is evident that the bottom waters of the lake contain a large amount of dissolved Fe that would be divalent.

Table 2 Principal cation concentrations

Site	Depth m	Al µg/L	Ca mg/L	Fe mg/L	K mg/L	Mg mg/L	Mn mg/L	Na mg/L
L. Norrtorpsjön	1	135	520	0.55	37.8	54.3	0.18	76.5
	8	92	496	3.91	32.0	49.6	2.35	61.4
	12	113	503	44.6	28.3	95.4	7.47	71.5
	27	92	454	69.2	31.2	117.4	8.58	62.8
L. Pölen	1	850	334	5.0	6.9	40.4	792	16.0
	3	778	321	4.7	6.3	38.7	819	15.5

The concentrations of Al in the acidic lake Pölen (750-800 µg/L) do not represent saturation with hydroxides or hydroxysulphate solid phases. However, in lake Norrtorpsjön several modifications of hydroxides and hydroxysulphates are oversaturated. This could be a reason for the rather low concentrations of Al (90-135 µg/L) as well as its retention in the filters throughout the water column (Fig. 3). Manganese had a similar depth dependence as Fe in Lake Norrtorpsjön, with moderate concentrations above the thermocline (200 µg/L) and high concentration (8000 µg/L) in the bottom water. Manganese was dissolved throughout the water column which is a confirmation of the idea that no solubility limiting stoichiometric phases were identified by geochemical modelling

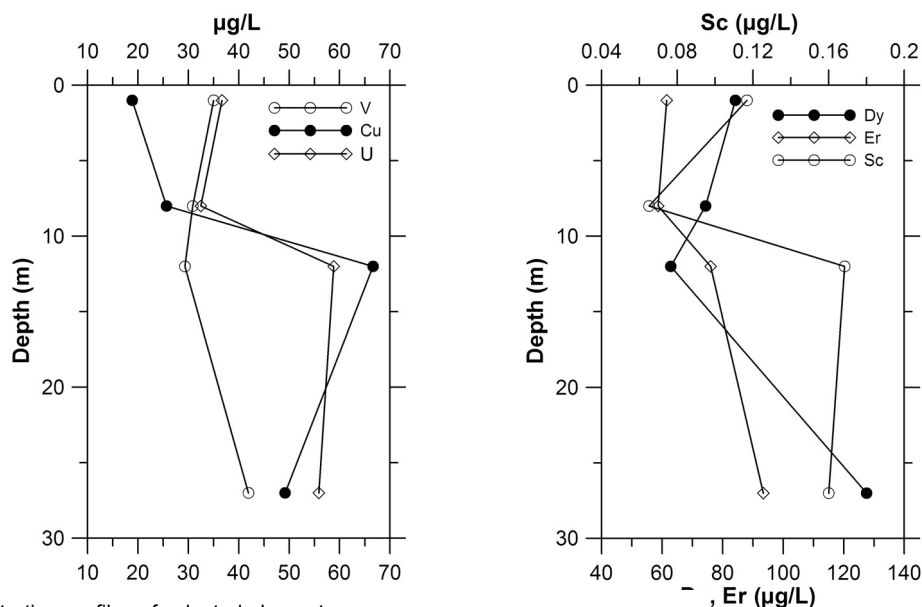


Fig. 3 Concentration profiles of selected elements

To summarise, the water chemistry in lake Norrtorpsjön is heavily influenced by the input of elements from three major sources (black shale, alkaline solid waste, municipal waste) in compared to lake Pölen that serves as a reference for the conditions governed by the surrounding shale/limestone. As a result of the alkaline waste the higher pH in Lake Norrtorpsjön induces precipitation of Al-hydroxides and hydroxyl-sulphates as well as of Fe-(hydro)oxides. In combination with oxygen depletion below the thermocline the Fe-phases are dissolved due to reduction of trivalent Fe. They

are maintained at these high concentrations since oxidative decomposition of organic matter is not high enough to induce reducing conditions where Fe-sulphides form.

3.2 Trace metals

In lake Norrtorpsjön the depth profiles for Ni (15-70 µg/L), Co (15-70 µg/L), Mo (10-25 µg/L) and U (40-60 µg/L) correlated with those of Fe, i.e. low concentrations above the thermocline and higher in the bottom water (Table 3). All of these metals exhibited filterable fractions (0.2 µm) corresponding to up to 30% of the total concentration which showed a tendency to increase below the thermocline. No stoichiometric solid phases were identified according to saturation conditions which would contribute to their particulate/colloidal species, except for Co. Above the thermocline, the Co concentrations were in equilibrium with malachite and oversaturated with respect to tenorite and cupric ferrite. Only U exhibited lower concentrations after the sample had passed through the anion-exchanger. Above the thermocline 5-15% was retained and below it increased to 30-40%. These results are in quite poor agreement with the 6.5% of negatively charged U complexes that was identified from equilibrium modelling. In the acidic lake Pölen, the concentrations were slightly higher for Ni and Co (70 µg/L), lower for Mo (0.6 µg/L) and not different for U (22 µg/L). A more prominent difference was the lower abundance in the filtrates for all of them, not exceeding 5%, most likely an effect by the lower pH. Only U was retained by the anion exchanger and then up to some 30%.

Table 3 Selected trace metal concentrations (BD below detection)

Site	Depth m	Cd µg/L	Cu µg/L	Mo µg/L	Ni µg/L	Pb µg/L	U µg/L	V µg/L	Zn µg/L
L. Norrtorpsjön	1	0.097	18.8	9.4	18.8	0.40	36.6	35.0	199
	8	0.042	25.6	13.6	25.6	0.10	32.4	30.8	176
	12	0.054	66.6	23.5	66.6	18.8	58.8	29.3	193
	27	0.046	49.2	12.6	49.2	1.20	55.9	41.9	185
L.Pölen	1	0.406	69.0	0.6	69.0	1.79	22.8	38.5	216
	3	0.363	60.0	BD	59.9	1.40	22.4	18.8	204

Zinc had quite constant concentrations (180-200 µg/L) throughout the profile in Lake Norrtorpsjön. About 10% was retained by the 0.20 µm filter and no anionic species was identified. Since its concentrations were below saturation, its distribution would be controlled by sorption. This is apparently a rather incomplete redistribution mechanism in this lake system in spite of the circumneutral pH. In lake Pölen, the concentrations were in the upper range (220 µg/L). Zinc was not retained either by the filters or by the anion exchanger. A similar behaviour would be expected for Cd, because of its chemical similarity with Zn. Although the Cd concentrations are much lower (0.04-0.1 µg/L) it has the highest concentrations at one meter depth and rather stable concentrations below. In all samples about 50% is retained by the 0.2 µm filter, which indicates the importance of particle transport for this element. Sorption process would be working since no saturation with stoichiometric phases was measured. In fact, the only similarity with Zn is that no anionic species was measured. In lake Pölen, the slightly higher Cd concentrations (0.5 µg/L) were present in species that passed through the filters in a cationic form. Lead in Lake Norrtorpsjön exhibited the lowest concentration at

the surface (0.4 µg/L) and increased slightly below the thermocline (1.5 µg/L) while in lake Pölen it was homogeneously distributed (2 µg/L). There was, however, a great qualitative difference since lead in lake Norrtorpsjön was quantitatively retained by the filters (0.4 µm, 0.2 µm) while filtration had no impact on samples from Lake Pölen. In spite of the different pH values the concentrations of V are rather similar in both lakes (20-40 µg/L) and the element was retained by the 0.2 µm filter with up to 25%.

3.3 Rare earth elements

Surrounding the black shale is a source for the rare earth elements (REE), including Sc and Y. Lowering of pH value would mobilize relatively high concentrations of REE, which should be reflected by their concentrations in adjacent water bodies. This is true for lake Pölen where most of the REE exhibit rather high concentrations, in relation to non-shale environments (Table 4).

Table 4 Selected rare earth element concentration (BD below detection)

Site	Depth m	Ce µg/L	Dy µg/L	Er µg/L	Eu µg/L	Gd µg/L	La µg/L	Nd µg/L	Sc µg/L	Sm µg/L	Tb µg/L	Y µg/L	Yb µg/L
L. Norrtorpsjön	1	BD	0.08	0.06	DB	DB	DB	0.02	0.11	DB	DB	0.01	DB
	8	BD	0.07	0.06	DB	DB	DB	DB	0.07	DB	DB	DB	DB
	12	BD	0.06	0.08	DB	DB	DB	0.01	0.17	DB	DB	0.09	DB
	27	BD	0.13	0.09	DB	DB	DB	DB	0.17	DB	DB	0.06	DB
L.Pölen	1	12.0	15.4	9.91	0.31	1.87	3.77	6.84	0.05	1.41	0.24	0.40	0.72
	3	12.9	16.9	10.8	0.31	1.98	0.30	7.38	b.d.	1.52	0.23	0.41	0.81

Moreover, they are not significantly retained by the filters or anion exchanger. The conditions are quite different in lake Norrtorpsjön where only Dy, Er and Sc are present at concentrations above the detection limit with the equipment used in this study. Possibly, it is related to the fact that these elements are bound to the particulate fraction with up to 70% of the total content. The distribution is evidently related to the redox conditions; since the particulate fractions were found below the thermocline. In addition, the concentrations in the untreated samples are roughly twice as high below the thermocline as above it. Another feature in lake Norrtorpsjön is that, these elements are retained by the anion exchangers. Up to 20% of the content in the untreated sample was measured in the bottom water while it was just a few percent at the surface. If these observations are consistent with environmental conditions or artefacts from the sample pre-treatment remains to be elucidated. The difference in pH value between the systems, seems to generate particulate carrier phases that redistribute the elements within the system.

3.4 Statistical interpretation

The entire data set was evaluated with hierarchical cluster analysis in R- and Q-mode to examine the relationships between variables and samples, respectively. In lake Norrtorpsjön the samples were grouped in three main clusters. The first main cluster represents samples from a depth of nine meter and deeper; the second main cluster included that samples that were collected from the depth of five to eight meter, and the third main cluster represents the samples that were collected from the surface

up to a depth of five meter. The physical interpretation is due to the fact of three compartments, epilimnion, thermocline and hypolimnion. With regards to the R-mode cluster analysis, the analytical parameters are divided into two main clusters. The first main cluster includes two sub-clusters. The first sub-cluster includes Ca, Fe, Ga, Mg, Mo, Na, Ni, Sc and U while the second sub-cluster includes Al, K, Pb, V and Zn. The second main cluster is divided into two main sub-clusters as well. The first sub-cluster includes Ce, Eu, Gd, La, Lu, Nd, Pr and Y while the other one includes Cd, Cu, Ga, pH and V. In brief, with regards to the obtained results, the relationship between the analytical parameters could have different origins and different behaviour in different compartments of the lake. Furthermore, the relationship between the analytical parameters could be due to the chemical environment and the properties of the individual element. One critical parameter is the depth that the samples were collected; the other parameter is pH value. The obtained results indicate that in order to better understand the system; the samples should be collected under different hydrochemical conditions.

3.5 Environmental consequences

Present data are not sufficient for a complete spatial or temporal evaluation. However, some remarks can be made concerning the impact of the pit lake on downstream systems. During stratification only the epilimnion is active in the lateral transport process. The retention time would be a fraction of the theoretical residence time. Individual rain storms would have a large impact on the downstream transport of the elements in the aqueous phase. The rather high pH value in lake Norrtorpsjön results in a general accumulation, or at least a temporal retention. The retention could be counteracted by the redox conditions. In a “natural” system with this depth the primary production in combination with the inflow of carbon would lead to an accumulation of organic matter at the sediment surface. During stratification, oxygen would be consumed. With regards to the high sulphate concentrations in the system, sulphide formation could occur.

4 Conclusions

The neutralised pit lake, lake Norrtorpsjön retains a number of metals from the surrounding sources during summer stratification because of its rather high pH value, which facilitates adsorption processes. Zinc and vanadium are exceptions due to weak adsorption properties and speciation, respectively. Both allochthonous and autochthonous sources of carbon are too low to allow for a low redox potential which is low enough to induce sulphide formation. Hence, metals that are accumulating below the thermocline are available for further transport downstream once the autumn circulation begins.

Acknowledgement

The authors express their gratitude to Örebro University for financial support.

References

- Dyni JR (2006). Geology and resources of some world oil shale deposits. U.S. G. S. Scientific investigations report. 2005-5294.
- Eklund M, Bergbäck B, Lohm U (1995). Reconstruction of historical cadmium and lead emissions from Swedish alum works. *Sci. Total Environ.* 170: 21-30.
- Karlsson L (2011). Natural weathering of shale products from Kvarntorp. Bachelor Dissertation. School of Science and Technology, Örebro University.
- McKnight DM, Bencala KE (1988). Diel variations in iron chemistry in an acidic mountain stream in the Colorado Rocky Mountains. *Arctic. Alpine Res.* 20: 492–500.
- SWCO VIAK (2005). Serpentindammsystemet Kvarntorp: SAMmanfattande rapport avseende restaureringsåtgärder som har genomförts i serpentindammsystemet, Kvarntorp, (In Swedish).

Manuscript 5

Dry covers on historical sulphidic mine waste; long term statistic performance assessment of surface water quality

Dry covers on historical sulphidic mine waste – long term statistic performance assessment of surface water quality

Anahita Pourjabbar¹ - Stefan Karlsson² - Bert Allard²

¹ Institute of Geoscience, Friedrich Schiller University, Burgweg 11, 07749 Jena, Germany, anahita.pourjabbar@uni-jena.de

² Man-Technology-Environment Research Centre, Örebro University, SE-701 82 Örebro, Sweden, stefan.karlsson@oru.se

Abstract

In 1988 two sub-sections of a historic mine site with sulphidic ore residues (Fe, Cu, Zn, Cd and Pb) at Bersbo, Sweden, were covered with compacted illitic clay and cement stabilised fly ash, respectively. Changes in surface water composition during the period from 1985 to 1996, before and after the remediation, have been statistically evaluated with data from two adjacent locations (Lake Gruvsjön and the Kuntebo creek), and a downstream lake (Lake Risten). The analysis is focused on R- and Q-mode clusters and time trends using concentrations of Al, Cd, Cu, Fe, Mn, Pb, and Zn, as well as pH (as a predictor in time trend). Data from Kuntebo and Lake Risten represent two main Q-mode clusters based on the sampling dates (before and after remediation period). However, the Lake Gruvsjön data from 1989 form a third cluster. Two R-mode clusters are observed in the Kuntebo data, one is included Al, Cd, Cu and Pb (1), and the other one is included Fe, Mn, Zn and pH (2). Corresponding clusters of Lake Gruvsjön data include: Al, Cd, Cu, Pb, Zn, and Fe (1) and pH (2), respectively. Lake Risten data represent four clusters: Cu, Pb and Zn (1), Al, Cd and Fe (2), Fe (3), and pH (4). The predicted trends are compared with measured concentrations from the period after the remediation in 1988. A general conclusion is that the remediation strategy was effective and successful, based on the statistical analysis of data. The only exception is Fe that did not change significantly after remediation.

Key Words: ARD, sulphidic mine waste, covering, statistical analysis, Fe, Cu, Zn, Cd, Pb

1 Introduction

Historical sulphidic mine waste is a single largest source of environmental pollution of metals in Sweden. It has been estimated that some 8,000 sites are in need for remediation according to national standards. Many of them are found at remote sites with low populations, while in other cases entire cities are built upon them. Several projects aim to determine the most favourable remediation strategy, but until now only a few sites allow for long term performance assessment. Overviews on the conditions after remediation are given by Karlsson and Bäckström (2003, 2005). A statistical evaluation of surface quality changes after remediation in 1988 of a historical sulphidic mine site in Bersbo, Sweden, has been made in the present project, as a follow-up of an extensive monitoring program.

2 Materials and Methods

2.1 The Bersbo site

Bersbo is located some 250 km SSW of Stockholm, Sweden, in the municipality of Ätvidaberg. Mining for copper is documented from the 14th century but probably began earlier. Rational mining was introduced in 1765, and during the peak production from 1850 to 1870, Bersbo was the largest copper producer in Sweden. Copper was refined in the nearby refinery in Ätvidaberg. The ore that could be extracted with the technology of the time ceased in the late 1800s, and the mining was stopped in 1902. For a more complete description see Allard et al. (1987).

The copper ore was mainly magnetic chalcopyrite that was confined to narrow veins together with pyrite, and sphalerite in a leptite/granite host rock. Karlqvist and Qvarfort (1979) reported that the element concentrations were in the ranges of 0.5-3% Cu, 1-3% Zn, 0.5-1% Pb, around 20% Fe and 25% S. A detailed description of the mineralogy is given by Tegengren (1924). The grain size of the waste ranged from silt to rock, and the waste was put in the immediate vicinity of the shafts. These piles had a hydraulic conductivity of 10-3 m/s, and consequently, the material was in full contact with atmospheric oxygen and rainwater.

Because of a ridge that runs in north-south direction through the site, approximately half of the drainage from the waste is directed towards the east (the Kuntebo creek, K) and west (Lake Gruvsjön), respectively (Fig. 1). There were two major waste piles with approximate volume of 700,000 m³, that cover 0.2-0.3 km² (Allard et al., 1991).

The two catchments have different hydrological properties. The westerly one reaches the small Lake Gruvsjön after just some 100 m. The lake has an average depth of 8 m and a theoretical residence time of approximately 10 months. The easterly catchment is drained by the small Kuntebo creek (some 1 m wide) with highly variable discharge and retention times. The annual variation at location Kuntebo varies from 0 l/sec during the dry summer to about 150 l/sec during intense snow melt. An ordinary late summer shower of 15 mm usually gave an increase from 1-2 l/sec to 30-40 l/sec, depending on original level of the water table. The corresponding flows for the outlet of Lake Gruvsjön are 1 l/sec and 1500 l/sec.

Peak flow occurs during snow melt, usually in April. Total precipitation reaches 650 mm of which some 200 mm forms surface water runoff. A comprehensive presentation of the hydrological conditions before the remediation is given by Sandén et al. (1987, 1997).

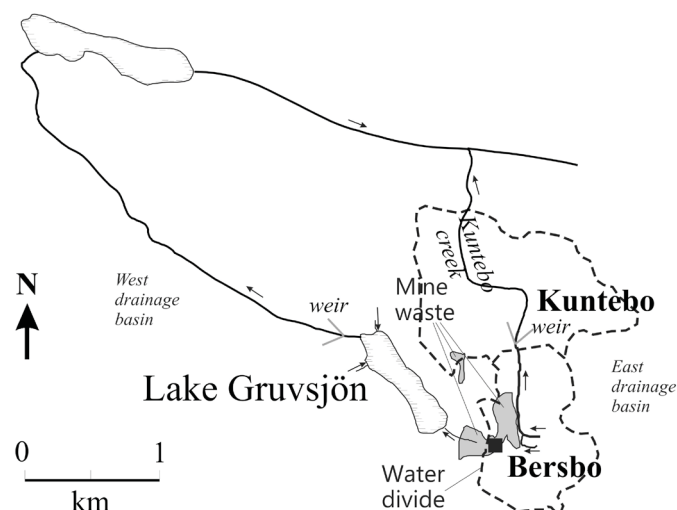


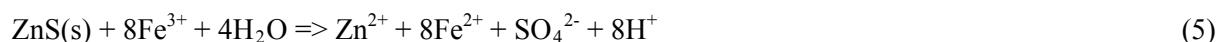
Fig. 1 Map of the study area

Rather coarse till dominates the soil, but in lower parts lenses of illitic clay are found. This is also reflected by the groundwater type which is of $\text{Ca}^{2+}\text{-SO}_4^{2-}\text{-CO}_3^{2-}$ with pH value about 7.5 and an alkalinity of 2.5-3 meq/L (Allard et al., 1987). The higher vegetation consists of mixed stands with coniferous and deciduous tree species with a layer of mainly vaccinium species as well as grasses and mosses on the soil.

In this area, the oxidative weathering of pyrite, as well as other sulphide minerals, has been summarised by Stumm and Morgan (1996):



Reactions 1 to 3 are representative of the weathering pathway before the remediation since there was no limitation of either air or oxygen. The reactions are usually catalysed by bacteria which results in a rate some 10,000 higher than for the abiotic processes.



Reaction 4-5 are driven by Fe(III) as the electron acceptor, and presence of air is not required.

2.2 Remediation

In the years 1987 to 89 the site was remediated by filling all available shafts with mine waste and sealing them as well as putting the remaining material in two piles that were contoured for optimum stability (Lundgren and Lindahl 1991). All exposed soil was limed. The waste on the westerly side

was covered with a sealing layer of 0.25 m cement stabilised fly ash and the easterly one with 0.5 m compacted illitic clay from the area. For both materials the target permeability was 10⁻⁹ m/s or lower. Identical 2 m protective layers of till was put on the impervious layers. Vegetation was allowed to establish, and after a couple of years it consisted of grasses, shrubs and plants of Scots pine. A complete technical description of the technical procedure is given in Lundgren (1990) and Lundgren and Lindahl (1991).

2.3 Surface water monitoring

The strategy for monitoring of surface water quality relied upon weekly samples from some 15 locations along the watercourses. Water discharge was measured continuously during the period of 1983 to 1988 at the weirs, and the water level at the outlet of Lake Risten was noted on sampling occasions. The weirs were also equipped with automatic samplers that collected up to four daily samples. With exception of automatic sampling all samples were collected manually in 1 L plastic (polyethylene or polypropylene) bottles that were acid washed and thoroughly rinsed with 18 M Ω - water. Samples for metal analysis were acidified with nitric acid (Suprapur). After 1994, water quality measurement became less frequent because of economical constraints. The number of sampling sites and sampling frequencies were reduced successively and after 1996 only a few observations are available. The database contains some 50,000 quality observations at present.

2.3.1 Chemical analysis

Electrical conductivity and pH were measured in the laboratory with conventional electrodes. Metals included in the study were Al, As, Ca, Cd, Cr, Cu, Fe, K, Mg, Mn, Na, Ni, Pb and Zn. Before 1991 the metals were quantified by atomic absorption spectrophotometry (PerkinElmer 5000 Zeeman), either flame or furnace. This technique was replaced by ICP-MS (PerkinElmer Elan Sciex 5000) in 1991 after validation of the analytical performance. Metal analysis was performed on acidified original samples as well as on 0.40 μ m filtrates (not discussed here since the data that were used for statistical analysis are non- filtered data).

2.3.2 Statistical analyses

Statistical methods such as clustering analyses (Q- and R-mode) as well as time series analysis were used to classify the samples with respect to concentrations of Al, Cd, Cu, Fe, Mn, Pb, Zn, as well as pH.

2.3.2.1 Hierarchical Cluster analysis

Hierarchical Cluster analysis is a common multivariate technique in geochemical statistical studies. Cluster analysis can be applied as an “exploratory data analysis tool” to better understand the multivariate behavior of a data set. The principal aim of this technique is to partition multivariate observations into a number of meaningful, multivariate homogeneous groups. By applying this technique, a large amount of data reduces into a few clusters. Interpreting the clusters is easier, as each cluster contains some observations (samples or variables) that are similar to each other. The outcomes of this cluster analysis are shown by a special plot called dendrogram. The hierarchical clustering method uses the dissimilarities /similarities or distances between objects when forming the

clusters. Therefore, two important measurements for this output are distance and linkage measures. It should be noted that diverse techniques can yield different groupings, even when using exactly the same data.

2.3.2.2 Time series analyses

A time series is a set of observations obtained by measuring a single variable regularly over a period of time. The form of the data for a typical time series is a single sequence or list of observations representing measurements taken at regular intervals. These sets of data have a natural temporal ordering. It makes time series analysis distinct from other common data analysis problems, in which there is no natural ordering of the observations.

Time series analysis comprises methods for analyzing time series data in order to extract meaningful statistics and other characteristics of the data. It is also distinct from spatial data analysis, where the observations typically relate to geographical locations. One of the most important reasons for doing time series analysis is to try to forecast future values of the series. A model of the series that explained the past values may also predict whether and how much the next few values will increase or decrease. The ability to make such predictions successfully is important. Methods for time series analyses may be divided into two classes: (a) frequency-domain methods that include spectral analysis and recently wavelet analysis, and (b) time-domain methods that include auto-correlation and cross-correlation analysis.

Time series models will generally reflect the fact that observations close together in time will be more closely related than observations further apart. Models for time series data can be classified into three main methods such as the autoregressive (AR) models, the integrated (I) models, and the moving average (MA) models. These three classes depend linearly on previous data points.

Combinations of these approaches produce autoregressive moving average (ARMA) and autoregressive integrated moving average (ARIMA) models. The autoregressive fractionally integrated moving average (ARFIMA) model generalizes the former three.

Data are generally divided into two series: An estimation period, and a validation period. A model can be developed on the basis of the observations in the estimation period and the model can be tested and evaluated in the validation period. The results are shown by plots and tables that may include stationary R-square, R-square (R^2), root mean square error (RMSE), mean absolute error (MAE), mean absolute percentage error (MAPE), maximum absolute error (MaxAE), maximum absolute percentage error (MaxAPE) and normalized Bayesian information criterion (BIC). The plots would represent residual autocorrelations and partial autocorrelations. Results for individual models may be expressed in terms of forecast values, fit values, observed values, upper and lower confidence limits, residual autocorrelations and partial autocorrelations.

The data in this study originate from the samples that were taken from three sampling locations: The two lakes (Lake Gruvsjön and Lake Risten) and the creek (Kuntebo). The focus of the study is on the samples that were taken during the period of 1985 to 96. Missing data were replaced and a seasonal effect test was done prior to the statistical analysis. The data were analyzed by SPSS 17. Since the remediation processes have been performed from 1987 to 1989, the data were divided into two groups, representing the estimation and the validation period, respectively. The samples that were taken before the period of 1985 to 1987 were considered as belonging to the estimation period;

and the samples that were taken in the next years were considered as belonging to the validation period. Analytical parameters such as Al, Cd, Cu, Fe, Mn, Pb, Zn and pH were determined separately in each lake and in the creek. Prior to processing the data by times series analyses, a model type was selected. A model statistics table shows the goodness of the chosen model assessed from different parameters as mentioned above. Table 1 shows such results; SPSS 17 examines various model types and suggests the optimal one

Table 1 Model types and model fit statistics

Gruvsjön	Model Type	Model Fit statistics		Ljung-Box Q(18)	
		R-squared	Statistics	DF	Sig.
Al	ARIMA(0,0,0)	0,73	41,23	18,00	0,00
Cd	ARIMA(0,0,0)	0,72	17,58	18,00	0,48
Cu	ARIMA(0,0,0)	0,64	13,26	18,00	0,78
Fe	ARIMA(0,1,1)	0,45	11,24	17,00	0,84
Mn	ARIMA(0,0,0)	0,81	22,09	18,00	0,23
Pb	ARIMA(0,0,0)	0,87	24,77	18,00	0,13
Zn	ARIMA(0,0,0)	0,67	17,61	18,00	0,48

Kuntebo	Model Type	Model Fit statistics		Ljung-Box Q(18)	
		R-squared	Statistics	DF	Sig.
Al	Winters' Multiplicative	0,51	14,23	15,00	0,51
Cd	Simple Seasonal	0,61	27,88	16,00	0,03
Cu	Winters' Multiplicative	0,79	23,71	15,00	0,07
Fe	Simple Seasonal	0,19	11,70	16,00	0,76
Mn	Simple Seasonal	0,49	17,38	16,00	0,36
Pb	Winters' Multiplicative	0,68	21,57	15,00	0,12
Zn	Winters' Multiplicative	0,75	7,48	15,00	0,94

Risten	Model Type	Model Fit statistics		Ljung-Box Q(18)	
		R-squared	Statistics	DF	Sig.
Al	Winters' Multiplicative	0,72	31,11	16,00	0,01
Cd	Simple Seasonal	0,63	14,89	16,00	0,53
Cu	Winters' Multiplicative	0,77	17,65	16,00	0,34
Fe	Simple Seasonal	0,83	27,86	16,00	0,03
Mn	Simple Seasonal	0,61	19,03	15,00	0,21
Pb	Winters' Multiplicative	0,88	21,55	16,00	0,16
Zn	Winters' Multiplicative	0,84	22,00	15,00	0,11

3 Results

All mentioned statistical analyses were used for the three sampling sites, Lake Gruvsjön, Kuntebo and Lake Risten. None of the data sets show significant seasonal effects with regards to statistical analysis.

3.1 Lake Gruvsjön

The samples from this lake were divided into three main clusters by Q-mode cluster analysis (Fig. 2). Cluster one includes the samples that were taken in 1990s (from 1990 to 1995); the second cluster

as well as the third cluster include the samples that were taken from 1988 to 1995. Based on R-mode clustering, the analytical parameters are divided into two main clusters: cluster one includes Al, Cd, Cu, Mn, Pb and Zn, and cluster two includes Fe and pH (Fig. 3).

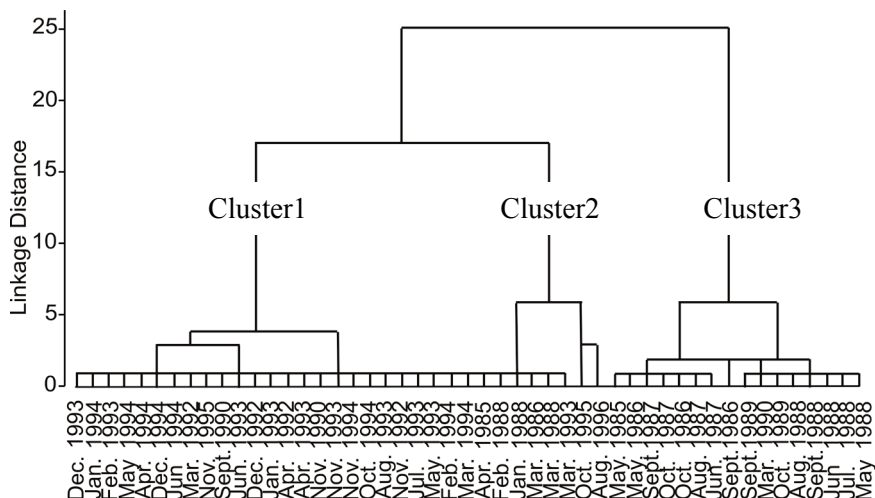


Fig. 2 Q-mode cluster analysis of Lake Gruvsjön. The samples from this lake were divided into three main clusters. Cluster one includes the samples that were taken in 1990s (from 1990 to 1995); the second cluster as well as the third cluster include the samples that were taken from 1985 to 1996

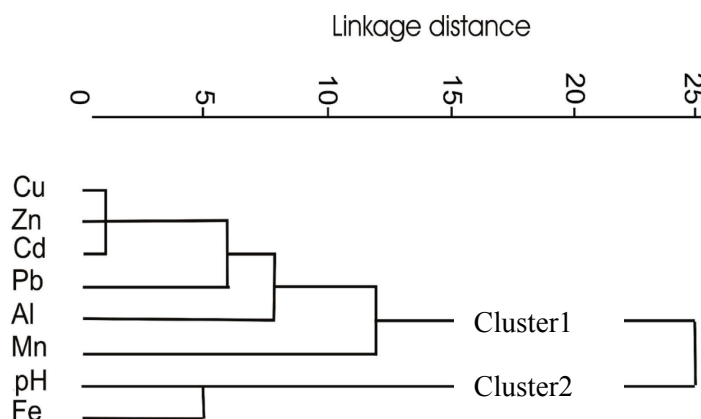


Fig. 3 R-mode cluster analysis of Lake Gruvsjön. The analytical parameters are divided into two main clusters: cluster one includes Al, Cd, Cu, Mn, Pb and Zn; and cluster two includes Fe and pH

Based on time series analyses, Al, Cu and Zn have similar behavior. The measured concentrations had both increasing and decreasing changes before the remediation period. This increasing and decreasing of the concentrations of the elements are also visible during the remediation period (1987- 1989). However, after this period the concentrations decreased noticeably. The forecasted concentration for the period of remediation is higher than the measured concentration. The behaviour of Cd is similar to that of the mentioned elements except two peaks of concentrations that were measured after remediation period from 1990 to 1991. There are similarities between the behavior of Mn and Pb. There is not a noticeable difference between the measured concentration before and after remediation period. However, two peaks of concentration were measured in 1990 and 1991. The forecasted value for Pb is very close to the measured one. However, the forecasted value of Mn is less than the measured one. The behaviour of Fe is unique and has no similarities to the other elements. The measured concentration of Fe increased noticeably after the remediation period, and a peak of concentration was measured in 1995. Figure 4 shows the described trends.

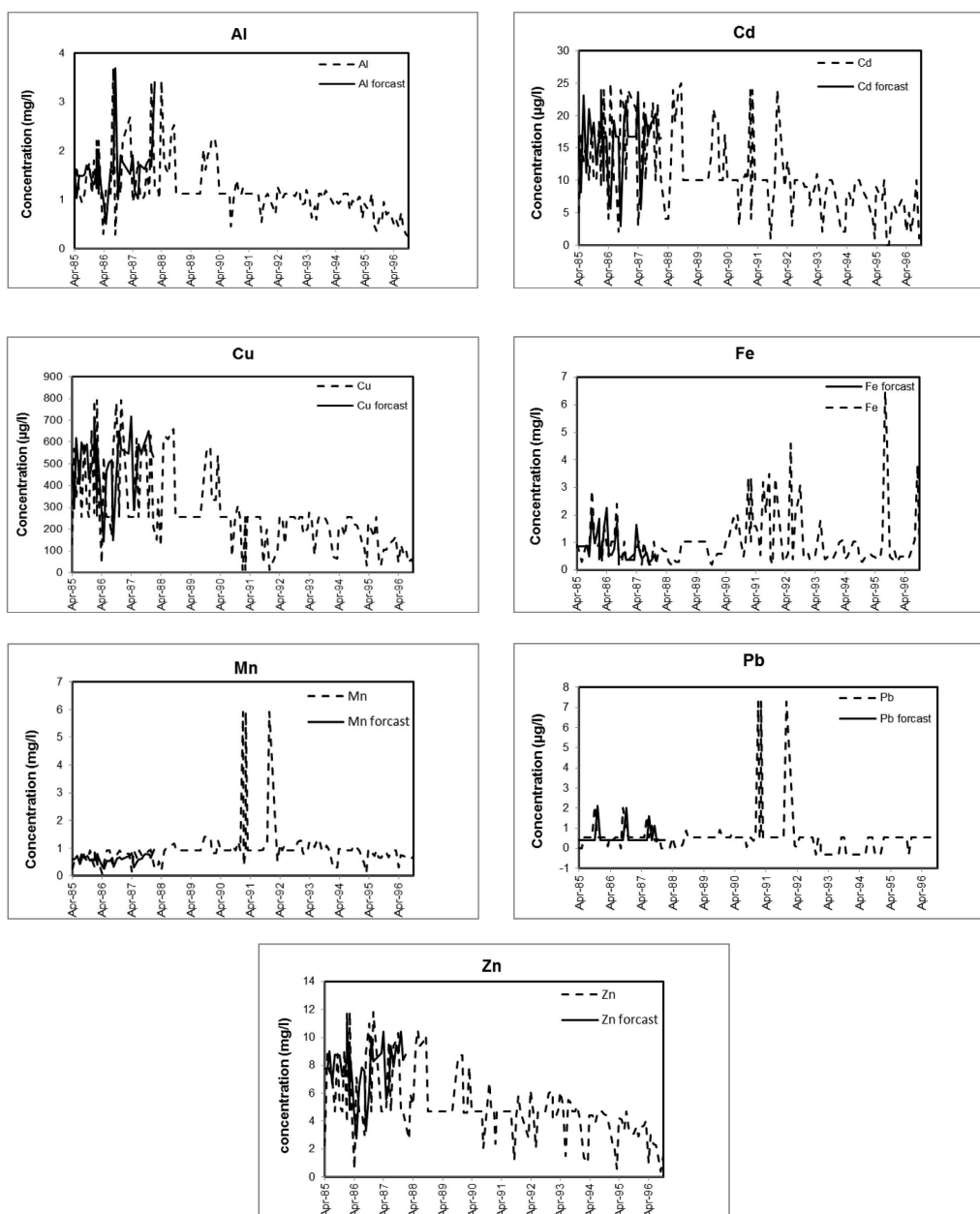


Fig. 4 Time series analyses with measured and predicted concentration from 1985 to 1996. The solid line shows the predicted concentration for the period of 1985 to 1988

3.2 Kuntebo

The samples from this creek were divided into two main clusters by Q-mode cluster analyses (Fig. 5). Cluster one includes the samples that were taken before 1990 and cluster two includes the samples were taken after this year. The analytical parameters were divided into two main groups by R-mode cluster analysis (Fig. 6). Cluster one has two sub-clusters. Sub-cluster one includes Al, Cd, Cu and Pb; and second sub-cluster includes Fe, Mn, and Zn. Cluster two has only one member, pH.

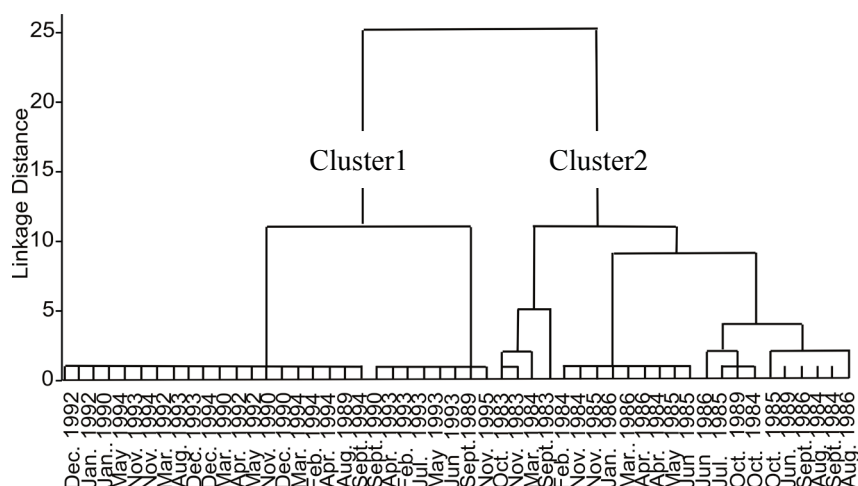


Fig. 5 Q-mode cluster analysis of Kuntebo creek. The samples were divided into two main clusters. Cluster one include the samples that were taken before 1990 and cluster two includes the samples were taken after this year

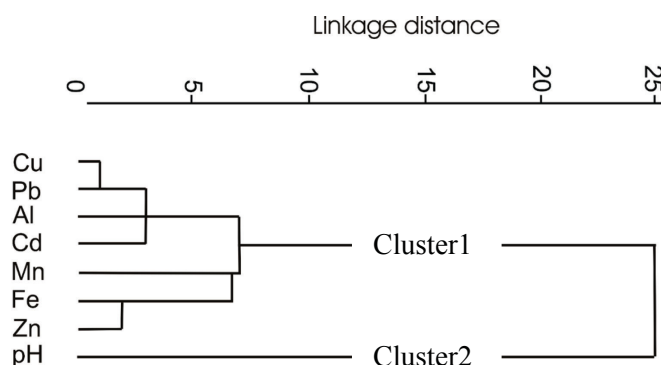


Fig. 6 R- mode cluster analysis of Kuntebo creek. The analytical parameters are divided into two main clusters. Cluster one has two sub- clusters. Sub-cluster one includes Al, Cd, Cu and Pb, and sub-cluster two includes Fe, Mn, and Zn. Cluster two has only one member, pH.

Similar to the behaviour of the elements from Lake Gruvsjön, increasing and decreasing the concentrations are visible for all elements from Kuntebo before the remediation period. A similar behaviour is observed for Al, Cd, Cu and Pb. The concentration of each of these elements, decreased noticeably after the remediation period, and the forecasted values are higher than the measured ones after remediation. However, the concentration of Zn does not change noticeably after the remediation period, and the forecasted values are less than the measured ones. The behaviour of Fe and Mn are similar. Regardless the wavy behaviour, the concentration of each element decreased after the remediation period. However, two peaks of high concentrations were measured in 1993 and 1995. Figure 7 shows the described trends.

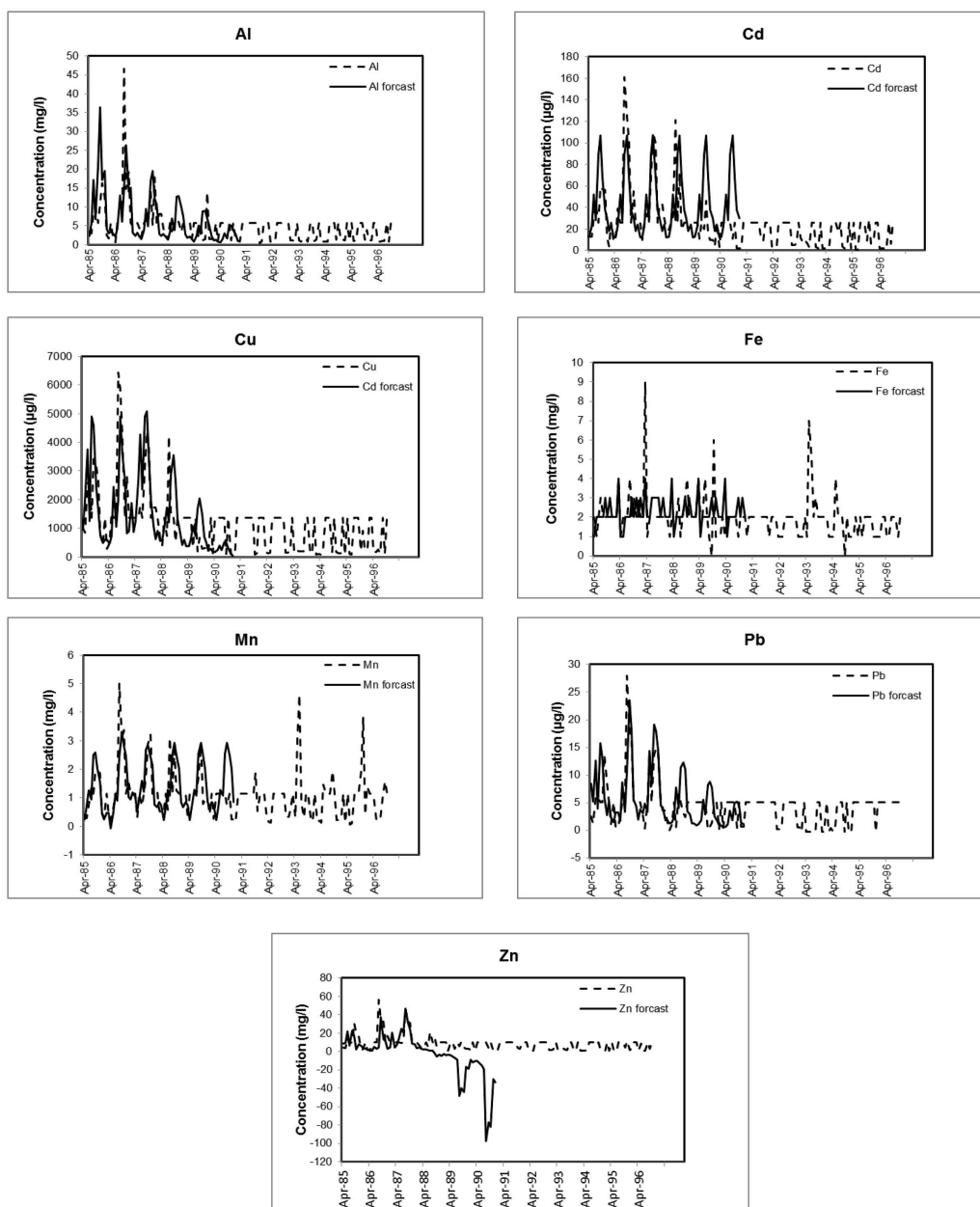


Fig. 7 Time series analyses of Kuntebo data, measured and predicted concentration from 1985 to 1996. The solid line shows the predicted concentration for the period of 1985 to 1991.

3.3 Lake Risten

The samples from Lake Risten were divided into two clusters by Q-mode cluster analyses. Cluster one includes the samples that were taken before 1988, and cluster two contains samples that were taken after this time (Fig. 8). The R-mode cluster analysis was used to classify the analytical parameters. By means on this method, the analytical parameters were divided into four main groups.

The analytical parameters were divided into two main groups. Cluster one includes all elements; and cluster two has only pH (Figure 9).

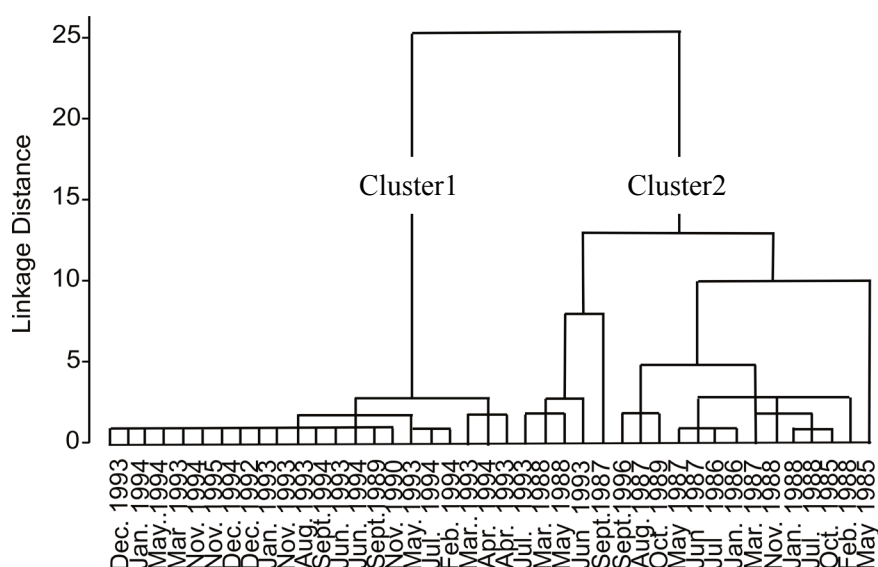


Fig. 8 Q-mode cluster analysis of Lake Risten. The samples from this lake were divided into two main clusters. Cluster one includes the samples that were taken before 1988, and cluster two contains samples that were taken after this time

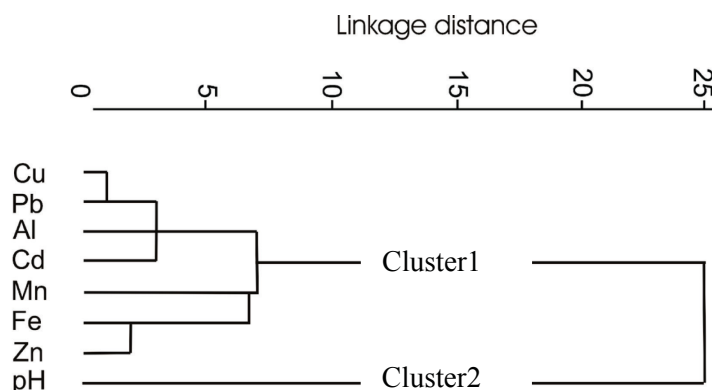


Fig. 9 R-mode cluster analysis of Lake Risten. The analytical parameters were divided into two main groups. Cluster one includes all elements; and cluster two has only pH

Based on the time series analyses, Cu, Pb and Zn exhibit a similar behaviour. A significant decrease occurred after the remediation period (mainly in 1990) for each of these elements, and the forecasted values are significantly higher than measured values. Three elements Al, Cd and Fe behave in a similar way as Al and Cd from Kuntebo and Lake Gruvsjön. The only exception is Mn, which shows a stable concentration before and after the remediation period. Figure 10 shows the described trends.

However, this exception will not influence the main outcome (the concentrations of the elements decreased after remediation period). The similarities of analytical parameters were investigated by R-mode clustering. The similar behaviour of the elements has a physicochemical explanation either by strong interactions with particulate/organic matter (Cu, Pb, Zn, Al, Cd, Fe) and their tendency to be the main element in stationary and mobile carrier phases (Al, Fe). In the Bersbo system Mn remains in an under-saturated state and has a very limited abundance in the particulate phase.

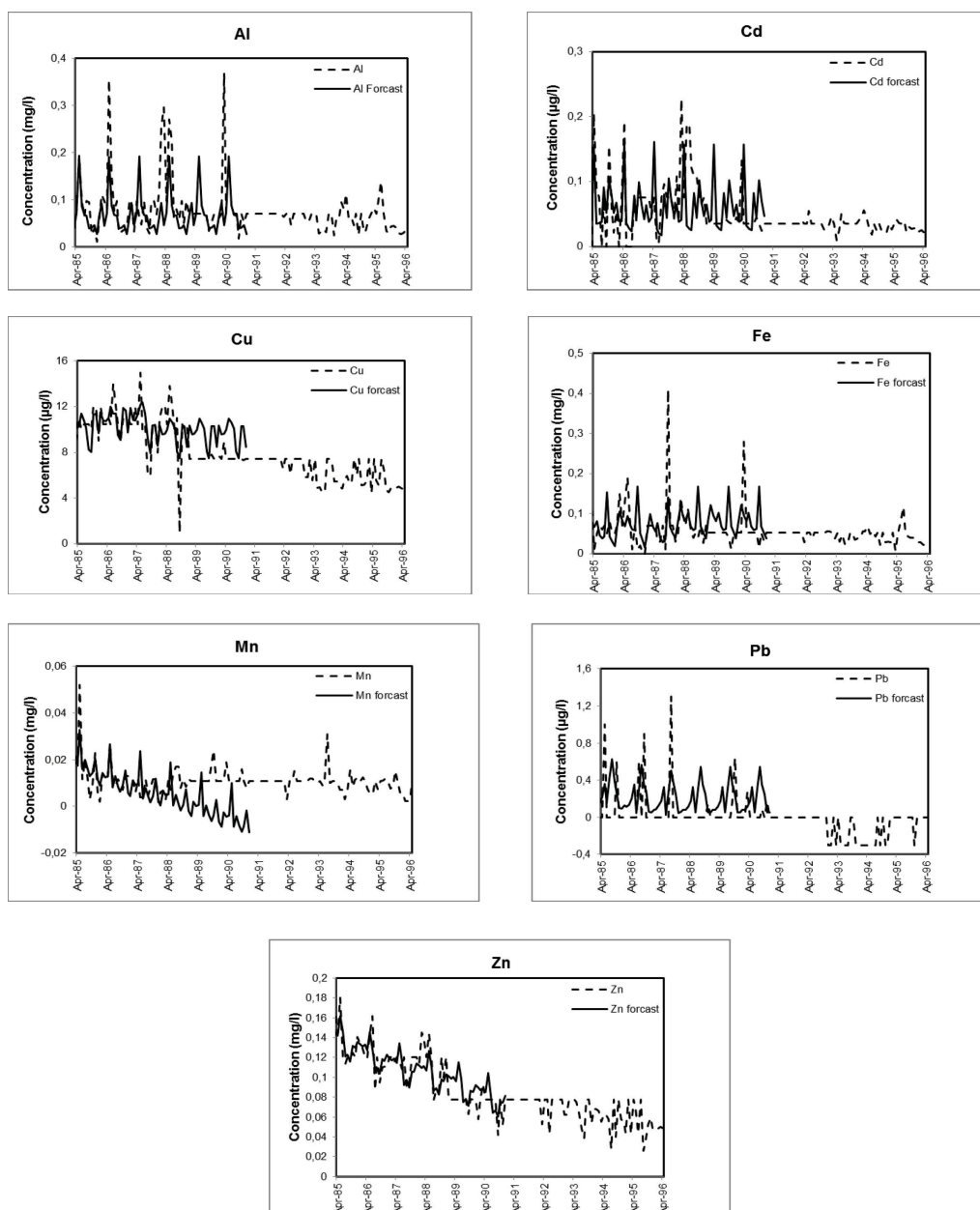


Fig. 10 Time series analyses of lake Risten data, measured and predicted concentration from the period 1985-1996. The solid line shows the predicted concentration for the period of 1985 to 1988

4 Discussion and conclusion

The behaviour of the elements is different before and after the remediation period, based in the statistical analysis. These differences can be seen in the Q-mode cluster analyses as well as in the time series analyses. The analysed data are concentrations of single elements, and the results represent concentration changes, not mass fluxes. The Q-mode clusters were generally good representative of the samples that were taken before or after the remediation period. An exception is the Q-mode dendrogram of Lake Gruvsjön, where the samples from the remediation period are clustered into two separate groups.

The rapid increase in pH in combination with annual water flow changes from 0 to 150 l/s at the Kuntebo site are the main reasons for the increase- decrease level of the metals concentration. In addition, previous studies have identified the quantitative impact of metal redistribution between stationary and mobile carrier phases, respectively, in relation to the transport in the dissolved phase.

The single most important factor that control the redistribution processes is pH, as identified by the statistical analysis and previous mechanistic interpretations (Karlsson and Bäckström, 2003). This also highlights the relevance of the pH dependent control for formation of carrier phases through precipitation (hydrous oxides of Al and Fe) as well as adsorption to stationary phases. In addition, the adsorption of trace elements to the solids is a highly pH dependent process. The slightly different outcome for Lake Gruvsjön is attributed to the continued weathering of sulphides in the covered waste where precipitated Fe(III) serves as the oxidizing agent in the absence of air. When groundwater that has passed through the deposit enters into Lake Gruvsjön there is a massive precipitation of Fe-hydroxides, which is a highly acid generating process. This negative effect is illustrated by the time series analysis.

With regards to the statistical outcomes, the remediation strategy significantly lowered the concentrations of the toxic metals, particularly at the Kuntebo sampling site. It is also clear that weathering of the waste could not be stopped by covering only, with respect to the increase of Fe concentrations at the inlet of Lake Gruvsjön. Consequently, it had the largest impact on the water balance. The statistical analysis also suggested the possibility of other mechanisms of mobility control and/or metal sources after the remediation. These circumstances must be fully investigated in order to complete a performance assessment. Hence, the overall success of the remediation should be evaluated in a long term perspective in order to include physicochemical aspects of the system as well as political ambitions.

Acknowledgements

Data from the the extensive Bersbo monitoring program, partly operated by the Swedish Environmental Protection Agency, was at our disposal, which is gratefully acknowledged.

References

- Allard B , Bergström S , Brandt M , Karlsson S , Lohm U, Sandén P (1987). Environmental impacts of an old mine tailings deposit- hydrochemical and hydrological background. Nord. Hydrol. 18: 270-290.
- Sandén P, Karlsson S, Düker A, Ledin A, Lundman L, Pettersson C (1997). Variations in general hydrochemistry, trace metal concentration and transport during a rain storm event in a small catchment. J. Geochem. Explor. 58: 145-155.
- Allard B, Karlsson S, Lohm U (1991). Full scale restoration of a mine waste deposit, a case study of waste management strategy and decision making. Proc. Int. Conf. Environ. Conse. of hazar. Waste dispos. Stockholm, Sweden. 2: 57-66.
- Karlsson S, Bäckström M (2003). Surface water quality in Bersbo, Sweden;fifteen years after amelioration of sulphidic waste. Proc. Int. Conf. Min. Environ. Sudbury, Canada.
- Karlsson S, Bäckström M (2005). Sediment quality in Bersbo, Sweden, Sixteen years after amelioration of sulphidic waste. Proc. Int. Conf. Biogeochem. of Trace Elemen. Adelaide, Australia. 614-615
- Karlqvist L, Qvarfort U (1979). The Bersbo project, effects of mining activities on the Bersbo area. Åtvidaberg. (in Swedish)
- Lundgren T (1990). The first full scale project in Sweden to abate acid mine drainage from old activities. Proc. Int. Conf. GAC-MAC. Vancouver, Canada.
- Lundgren T, Lindahl L Å (1991). The efficiency of covering the sulphidic waste rock in Bersbo, Sweden. Proc. Int. Conf. Abatement of Acid Min. Drain. Montreal, Canada.
- Tegengren, F.R. (1924) Sveriges ädlare metaller och bergverk. SGU Ser. Ca, Nr 17. Norstedts, Stockholm.
- Sandén, P., Karlsson, S. and Lohm, U. (1987) Environmental impacts of an old mine tailings deposit – metal concentrations and water pathways. Nord. Hydrol. 16 (301-312).
- Stumm W, Morgan JJ (1996). Aquatic Chemistry. John Wiley and Sons. New York, United States of America.

Ich erkläre,

dass mir die geltende Promotionsordnung der Fakultät bekannt ist;

dass ich die Dissertation selbst angefertigt und alle von mir benutzten Hilfsmittel, persönlichen Mitteilungen und Quellen in meiner Arbeit angegeben habe;

dass mich folgende Personen bei der Auswahl und Auswertung des Materials sowie bei der Herstellung des Manuskripts unterstützt habe

dass die Hilfe eines Promotionsberaters nicht in Anspruch genommen wurde und dass Dritte weder unmittelbar geldwerte Leistungen von mir für Arbeiten erhalten haben, die im Zusammenhang mit dem Inhalt der vorgelegten Dissertation stehen;

dass ich die Dissertation noch nicht als Prüfungsarbeit für eine staatliche oder andere wissenschaftliche Prüfung eingereicht habe;

dass ich nicht die gleiche, eine in wesentlichen Teilen ähnliche oder eine andere Abhandlung bei einer anderen Hochschule als Dissertation eingereicht habe

Jena, 07.02.2012,

Acknowledgment:

I would like to express my gratitude for the support and help of all my colleagues.

Above all, I would like to express my sincere appreciation to my supervisors Prof. Dr. Georg Büchel and Prof. Dr. Jürgen W. Einax. This dissertation would have not been possible without their help, support and patience.

I would also like to thank Prof. Dr. Bert Allard, and Dr. Stefan Karlsson for their immense support, kindness, and hospitality during my three-month research visit at the University of Örebro, Sweden.

I also would like to acknowledge Prof. Dr. Costel Sarcbu, the University of Babes-Bolyai Cluj-Napoca, Romania, for his advice and guidance in applying his fuzzy algorithm. This method proved to be extremely useful for the statistical work and thus benefited the project tremendously.

I also would like to thank my colleague, Dr. Dirk Merten for his friendly advice and help during these years.

I am also grateful for the help of my colleagues at the chemistry laboratory and those who work at the test site, winter and summer, to collect the samples.

This project was a part of the graduate research training group, GRK 1257/1 (Alteration and element mobility at the microbe-mineral interfaces), and Jena School for Microbial Communication (JSMC) and financially supported by the German Research Society (DFG). I appreciate the support that they provided for this project and I would also like to acknowledge Wismut GmbH for providing the information of the studied area.

Finally, I would like to express my gratitude to my parents, my sister and brother, who have supported me unequivocally my entire life and to my husband for his personal support and great patience at all times.