# Analysis of flow patterns and flow mechanisms in soils

Christina Bogner

# Analysis of flow patterns and flow mechanisms in soils

Dissertation

Co-directed by the University of Bayreuth and the University of Avignon

submitted to the

Faculty of Biology, Chemistry and Geosciences

of the University of Bayreuth

to attain the degree of

Dr. rer. nat

and

Dr. de l'Université d'Avignon

Presented by

CHRISTINA BOGNER

born December 13, 1977

in Nawoi (Uzbekistan)

Bayreuth, April 2009

This doctoral thesis was prepared at the Department of Soil Physics, University of Bayreuth, and at the Hydrogeological Laboratory, University of Avignon, between September 2004 and April 2009. It was supervised by Prof. Dr. Bernd Huwe and Prof. Dr. Yves Travi.

This is a full reprint of the dissertation submitted to attain the academic degree of Doctor of Natural Sciences (Dr. rer. nat.) and approved by the Faculty of Biology, Chemistry and Geosciences of the University of Bayreuth.

Date of submission: April 21, 2009
Date of defence (disputation): July 6, 2009

Doctoral Committee:

| | |
|---|---|
| Prof. Dr. Egbert Matzner | Chairman |
| Prof. Dr. Bernd Huwe | 1st reviewer |
| Dr. habil. Isabelle Cousin | 2nd reviewer |
| Prof. Dr. Yves Travi | |
| Prof. Dr. Stefan Peiffer | |

*To Alf and Michael*

## Summary

Matrix flow and preferential flow can occur concurrently in the same soil. Both flow regimes produce typical flow patterns that can be visualised in dye tracer experiments. To extract quantitative information from dye tracer studies a vast variability of approaches exists. One of them is to describe dye patterns by the so called dye coverage function, i.e. the percentage of stained area per soil depth. Based on extreme value statistics the dye coverage function can be reinterpreted as a probability function to find the tracer in a certain depth. Therefore, the two-parametric probability distribution $1 - H$, $H$ being the generalised Pareto distribution, can be fitted to the dye coverage function. The form parameter of this distribution serves as a risk index for vertical solute propagation.

We did tracer experiments with Brilliant Blue FCF at three different study sites: in a Norway spruce forest in southeast Germany, in a tropical mountain rainforest in southern Ecuador and on an agricultural field in southern France. We tested the ability of the risk index to summarise main information obtained in dye tracer studies and characterise flow patterns in different soils under varying boundary conditions.

Our results suggest that the risk index is to some degree invariant to changing experimental conditions (such as irrigation rate). The initial soil moisture, however, seems to have a large influence on the risk index. It is difficult to adjust the parameters of the generalised Pareto distribution when the dye coverage function fluctuates or does not decrease monotonically. This might be due to tortuosity of paths, varying flow mechanism or changing soil physical properties (stratification). Thus, in stratified soil, we restricted the analysis to the lowest part of the profile. Since the theory of the risk index is based on extreme values of vertical solute propagation it is the lowest part of the profile that is the most interesting.

We propose to combine the two parameters of the generalized Pareto distribution and to use the complete distribution $1 - H$ to estimate the risk of vertical solute propagation in soils. Despite a certain resistance to changes of experimental conditions, the risk index is not an intrinsic soil parameter. Since the flow regime in the same soil can be dominated either by preferential flow or by uniform matrix flow, the risk of vertical solute propagation will change. It is a

physical reality and not a default in the risk index theory. The adjusted parameters of the generalised Pareto distribution will capture the dominant flow regime as reflected by tracer flow patterns. Bearing in mind the boundary conditions of the tracer experiment like irrigation rate, the tracer employed, soil initial moisture or type of vegetation (permanent or seasonal, deep rooted or shallow rooted) it is possible to compare different study sites or to consider the same site at different boundary conditions and to access the risk of vertical solute propagation.

Pattern analysis based on the risk index for vertical solute propagation revealed the occurrence of preferential flow at the German study site. To gain insight in flow mechanisms and possible impacts of preferential flow on soil chemistry we analysed soil texture, fine root density, soil bulk density, exchangeable cations, pH and total C and N contents in preferential flow paths and soil matrix. Results from linear mixed-effects models suggested that at this study site roots constituted main preferential flow paths and induced macropore flow, especially in the topsoil. In the subsoil root density decreased and inhomogeneous infiltration from preferential flow paths into the soil matrix caused non-uniform flow. There were no textural differences between the flow domains, but smaller bulk densities in preferential flow paths. This is probably due to a higher soil organic matter content in preferential flow paths. We found smaller pH values, more Ca, more Mg, more C and more N in preferential flow paths. Compared to the adjacent soil matrix, more Al and more Fe (but small absolute amounts) were found in the subsoil where macropore flow along root channels decreases and heterogeneous matrix flow dominates. These distinct chemical properties can be explained by root activity and translocation of solutes and DOC (dissolved organic carbon) via preferential flow paths. During transport along preferential flow paths contact time between DOC and soil is reduced so that DOC is transported to greater depth where it potentially forms organo-mineral associations. If this holds true, preferential flow is a mechanism that promotes C sequestration in subsoil and does not only influence its immediate environment around paths, but also underlying subsoil horizons.

A major outcome of this thesis is the large number of images of flow patterns from different soils. Further studies could employ recent dimensionality reduction techniques to investigate whether there is a low dimensional structure underlying these images.

# Zusammenfassung

Matrixfluss und präferentieller Fluss können in ein und demselben Boden gleichzeitig auftreten. Beide Fließregime erzeugen charakteristische Fließmuster, die in Versuchen mit Farbtracern sichtbar gemacht werden können. Es existiert eine Reihe von Methoden, um Tracerversuche quantitativ auszuwerten. Eine davon ist die Beschreibung der Fließmuster durch die so genannte Deckungsgradfunktion, den Anteil der gefärbten Fläche pro Tiefe. Die Methoden der Extremwertstatistik erlauben eine Neuinterpretation der Deckungsgradfunktion als eine Wahrscheinlichkeitsfunktion, den Tracer in einer bestimmten Tiefe anzutreffen. Demzufolge kann die zweiparametrige Wahrscheinlichkeitsfunktion $1 - H$ ($H$: verallgemeinerte Paretoverteilung) an die Deckungsgradfunktion angepasst werden. Der Formparameter dieser Verteilung dient als Risikoindex für vertikale Ausbreitung von gelösten Substanzen.

Tracerversuche mit Brilliant Blue FCF wurden an drei unterschiedlichen Standorten durchgeführt: in einem Fichtenwald in Südostdeutschland, einem Bergregenwald in Südostecuador und an einem landwirtschaftlichen Standort in Südfrankreich. Es wurde überprüft, ob die wichtigsten Ergebnisse aus Tracerversuchen auf unterschiedlichen Böden und bei verschiedenen Randbedingungen mithilfe des Risikoindex beschrieben werden können.

Die Ergebnisse zeigen eine gewisse Unabhängigkeit des Risikoindex von experimentellen Randbedingungen (wie z. B. Beregnungsintensität). Dagegen scheint die Bodenfeuchte eine zentrale Rolle zu spielen. Schwierigkeiten bei der Anpassung der Parameter der verallgemeinerten Paretoverteilung ergeben sich, wenn die Deckungsfunktion fluktuiert oder nicht monoton fallend ist. Dies kann möglicherweise auf die Tortuosität von Fließpfaden, variierenden Fließmechanismen oder sich verändernden bodenphysikalischen Eigenschaften (Stratifikation) zurückgeführt werden. Daher wurde die Musteranalyse in stratifizierten Böden auf den Unterboden begrenzt. Da die dem Risikoindex zugrunde liegende Theorie auf den Extremwerten der vertikalen Ausbreitung von gelösten Stoffen basiert, gilt das Hauptinteresse dem untersten Teil des Bodenprofils.

Wir schlagen vor, die beiden Parameter der verallgemeinerten Wahrscheinlichkeitsverteilung zu nutzen, um das Risiko der vertikalen Ausbreitung von gelösten Stoffen in Böden abzuschätzen. Obwohl der Risikoindex eine gewisse Toleranz gegenüber sich ändernden Randbedingungen zeigt, ist er kein intrinsischer Bodenparameter. Da das Fließgeschehen in ein und demselben Boden sowohl vom Matrix- als auch vom präferentiellen Fluss dominiert werden kann, ändert sich das Risiko der vertikalen Ausbreitung von gelösten Stoffen. Dies ist physikalische Realität und kein Fehler in der Theorie des Risikoindex. Die angepassten Parameter der verallgemeinerten

Paretoverteilung erfassen das durch den Tracer sichtbar gemachte dominante Fließregime. Unter der Berücksichtigung der Randbedingungen des Tracerexperiments wie Beregnungsintensität, des verwendeten Tracers, Bodenfeuchte oder Art der Vegetation (einjährig, mehrjährig oder perennierend, tiefwurzelnd oder flachwurzelnd) ist es möglich, unterschiedliche Standorte zu vergleichen oder denselben Standort unter verschiedenen Randbedingungen zu betrachten und das Risiko der vertikalen Ausbreitung von gelösten Stoffen abzuschätzen.

Extremwertstatistikgestützte Musteranalyse zeigte das Auftreten von präferentiellem Fluss auf dem Standort in Südostdeutschland. Um die Fließmechanismen und mögliche Auswirkungen des präferentiellen Flusses auf die Bodenchemie aufzudecken, wurden Textur, Feinwurzeldichte, Trockenraumdichte, austauschbare Kationen, pH, Gehalt an totalem C und N in präferentiellen Fließwegen und Bodenmatrix analysiert. Ergebnisse aus gemischten Modellen zeigen, dass auf diesem Standort präferentielle Fließwege durch Wurzeln gebildet werden, und zwar hauptsächlich im Oberboden. Im Unterboden nimmt die Durchwurzelung ab, und heterogene Infiltration aus den präferentiellen Fließpfaden in die Bodenmatrix führt zu ungleichmäßigem Matrixfluss. Es wurden keine signifikanten Unterschiede in der Textur gefunden. Allerdings ist die Trockenraumdichte in den präferentiellen Fließwegen geringer als in der Bodenmatrix, wahrscheinlich bedingt durch den erhöhten Gehalt an organischer Materie. Weiterhin wurden in den präferentiellen Fließwegen niedrigere pH-Werte, höherer Gehalt an Ca, Mg, C und N gemessen. Im Vergleich zur umgebenden Bodenmatrix wurde im weniger durchwurzelten und von heterogenem Matrixfluss dominierten Unterboden höherer Gehalt an Al und Fe (allerdings kleine absolute Mengen) festgestellt. Diese klar unterschiedlichen chemischen Eigenschaften lassen sich durch Wurzelaktivitäten und den Transport von gelösten Substanzen (darunter auch DOC: gelöster organischer Kohlenstoff) durch präferentielle Fließwege erklären. Während des Transports ist die Kontaktzeit zwischen dem DOC und dem Boden verkürzt, so dass der Kohlenstoff in tiefere Bodenhorizonte transportiert wird, in denen er eventuell organo-mineralische Komplexe bilden kann. Dies würde bedeuten, dass präferentieller Fluss unter Umständen die Kohlenstoff-Sequestration im Unterboden begünstigen könnte, und nicht nur seine unmittelbare Umgebung, sondern auch die tiefer liegenden Bodenhorizonte beeinflusst.

Ein wichtiges Ergebnis dieser Untersuchungen ist die große Anzahl an Bildern der Fließmuster in verschiedenen Böden. In nachfolgenden Arbeiten könnte mit den neuesten Methoden der Reduktion der Dimension untersucht werden, ob diesen Bildern eine niedrigdimensionale Struktur zugrunde liegt.

# Résumé

Des écoulements matriciels et des flux préférentiels peuvent se produire concurremment dans le même sol. Ces deux régimes d'écoulements se manifestent par des empreintes de flux caractéristiques qu'on peut visualiser par des essais de traçage. Afin d'extraire l'information quantitative des essais de traçage un grand nombre de méthodes existe. On peut, entre autre, décrire les empreintes de traceur par ce qu'on appelle la fonction de couverture, c'est à dire le pourcentage de région teintée par un traceur coloré en fonction de la profondeur du sol. En utilisant la statistique des valeurs extrêmes cette fonction peut être réinterprétée comme une fonction exprimant la probabilité de trouver le traceur à une profondeur donnée. Ainsi, la fonction de probabilité à deux paramètres $1 - H$, $H$ étant la distribution de Pareto généralisée, peut être ajustée. Le paramètre de forme de cette fonction est utilisé comme indice de risque de propagation verticale des solutés.

Nous avons effectué des essais de traçage au Bleu Brillant FCF sur trois sites différents : dans une forêt d'épicéa dans le sud-est de l'Allemagne, dans une forêt tropicale humide montagnarde en Équateur et sur un champ agricole au sud de la France. Nous avons examiné la capacité de l'indice de risque à rassembler l'information principale des essais de traçage et à caractériser les empreintes de flux dans des sols différents, sous conditions aux limites diverses.

Nos résultats indiquent que l'indice de risque est, dans une certaine mesure insensible aux changements des conditions aux limites (comme l'intensité d'irrigation). Par contre, l'humidité initiale du sol semble influencer cet indice de façon importante. L'ajustement des paramètres de la fonction Pareto généralisée s'avère difficile si la fonction de couverture fluctue ou ne décroît pas de manière monotone. Ceci peut être dû à la tortuosité des chemins d'écoulement, à la variation des mécanismes de flux ou aux changements de propriétés physiques du sol (stratification). Ainsi, dans des sols stratifiés nous avons restreint l'analyse à la partie inférieure du profil de sol. En effet, étant donné que la théorie de l'indice de risque est basée sur les valeurs extrêmes de propagation verticale de solutés c'est la partie inférieure qui est la plus intéressante.

Nous proposons de combiner les deux paramètres de la fonction Pareto généralisée et d'utiliser la distribution $1 - H$ complète afin d'estimer le risque de propagation verticale des solutés dans le sol. Bien que l'indice de risque montre une certaine invariance vis-à-vis du changement des conditions aux limites il n'est pas un paramètre intrinsèque de sol. Comme le régime d'écoulement dans un même sol peut être dominé soit par le flux matriciel soit par le flux préférentiel le risque de propagation verticale des solutés change. Ceci est une réalité physique et non un défaut dans la théorie de l'indice de risque. Les paramètres de la fonction de Pareto généralisée capturent le régime d'écoulement dominant représenté par

les empreintes du traceur. En prenant en compte les conditions aux limites de l'essai de traçage comme l'intensité d'irrigation, le traceur utilisé, l'humidité initiale du sol ou la nature de la végétation (pérenne ou saisonnière, type d'enracinement) il est ainsi possible de comparer des sites différents ou des résultats obtenus sur le même site sous conditions aux limites différentes et d'estimer le risque de propagation verticale de solutés.

L'analyse d'image d'empreintes de flux basée sur le risque de propagation verticale de solutés a révélé l'existence d'écoulements préférentiels sur le site allemand. Afin de comprendre les mécanismes de flux ainsi que les impacts éventuels des flux préférentiels sur la chimie du sol nous avons analysé la texture du sol, la densité racinaire, la densité du sol, la concentration des cations échangeables, le pH, et les teneurs en C et N total dans les chemins préférentiels et la matrice du sol. Les résultats de la modélisation indiquent que sur ce site les racines constituent les chemins préférentiels et créent les écoulements le long des macropores, surtout dans la partie supérieure du sol. Dans la partie inférieure la densité racinaire diminue et l'infiltration hétérogène à partir des chemins préférentiels dans la matrice provoque un écoulement non-uniforme. Nous n'avons constaté aucune différence significative de texture, mais des différences de densité du sol dans les chemins préférentiels par rapport à celle de la matrice. Ceci est probablement dû à la quantité de matière organique plus élevée dans les chemins préférentiels. Nous avons également trouvé des pH plus acides, plus de Ca, plus de Mg, et plus de C et de N dans les chemins préférentiels. Comparé à la matrice, des quantités plus importantes d'Al et de Fe (mais de petites quantités absolues) ont été trouvés dans la partie inférieure du sol où l'écoulement préférentiel le long des macropores créés par les racines diminue et le flux matriciel hétérogène domine. Ces propriétés chimiques distinctes peuvent s'expliquer par l'activité racinaire et la translocation de solutés et du carbone organique dissous (COD) le long des chemins préférentiels. Le temps de contact entre le COD et le sol étant réduit il est transporté plus bas dans le profil où il peut potentiellement créer des complexes organo-minéraux. Ainsi, l'écoulement préférentiel est un mécanisme qui peut promouvoir la séquestration de C en sous-sol et n'influence pas uniquement son environnent immédiat, mais aussi les horizons sous-jacents.

Un des acquis majeurs de cette thèse est le nombre important d'images d'empreintes de flux issues des sols différents. Dans les études qui suivront les méthodes récentes de réduction de dimensionnalité peuvent être employées afin de trouver d'éventuelles structures de basse dimensionnalité dans ces images.

# Acknowledgements

Writing acknowledgements always feels good because the hardest part of the work is done. Many people have supported me during my PhD and I would like to thank everyone who helped me with an advice, during field work, in the lab or just by listening to my problems.

I am sincerely grateful to my first advisor Prof. Dr. Bernd Huwe for many hours of fruitful discussions, for his good ideas, for finding excellent titles for papers, for his sense of humour and for the freedom to develop my own concepts. I am grateful to Prof. Dr. Martin Schlather who helped me to lose my fear of statistics. I even begin to like R! A special thanks to Andreas Kolb for his powerful assistance during filed work and his ability to find missing spare parts, and to Iris Schmiedinger for her support in the lab, in the field and a sympathetic ear. Thanks to Prof. Dr. Michael Hauhs for the opportunity to work at his department, and to Folkert Bauer for the lively exchange of insight. A special thanks to all members of the technical and the scientific staff of the DFG research group 562 "Dynamics of soil processes under extreme meteorological boundary conditions" for field site coordination.

This thesis would not have been possible in this form without the support of my second supervisor Prof. Dr. Yves Travi. He has the magic ability to resolve the most difficult administrative problems and believes that international cooperation is a chance worth taking. I want to express my gratitude to Dr. Stéphane Ruy for these fruitful years of cooperation, his help with administrative issues, assistance during experimental work and the warm-hearted welcome in his office. It was a real pleasure to work with you! I also would like to thank Dr. Vincent Marc for his help in organising my experimental work and his assistance with stable isotope issues. I am grateful to the technical staff of the INRA Avignon for their assistance during field experiments, and to Jean-Claude Gaudu for having wonderfully documented them.

I thank all my colleagues and members of the Department of Soil Physics, the INRA Avignon and the Hydrogeological Laboratory of Avignon for the nice and friendly work climate. Moreover, I would like to thank Mr Christoph Simon, IDS GmbH and MVTec Software GmbH for the opportunity to work with the software HALCON. And last but not least, I want to thank my parents for their support.

I dedicate my thesis to my husband Alf and our little son Michael. Thank you for your love, support and understanding. Without your help I would not have managed these four and a half years of hard work!

# Table of contents

# List of figures

# List of tables

# Chapter 1

# General introduction

## *1.1 Water flow in soils*

Two distinct categories of water flow in soils can be identified: uniform flow (1911) and non-uniform i.e. preferential flow (Lawes *et al.*, 1882). The former describes a relatively slow movement of water through soil micropores. The latter one covers all phenomena where water flows through localised pathways bypassing a portion of the soil matrix (Hendrickx & Flury, 2001). Although discovered in the late 19th century, preferential flow has been considered for a long time as exceptional. Nowadays, it has been recognized that uniform and preferential flow can occur concurrently in the same soil. Whether matrix flow or preferential flow dominates depends on:

(i) Intensity and variability of rainfall. Gish *et al*. (2004), for instance, showed that an irrigation rate of 4.4 mm $h^{-1}$ could generate preferential transport of bromide at steady state conditions in the field. Tymchak & Torres (2007) analysed the effect of variable rainfall intensity on timing and magnitude of soil water content response. Variable rainfall rates produced faster wetting responses: a possible indicator for preferential flow.

(ii) Water repellency. In a bromide tracer experiment Hedrickx *et al*. (1993) demonstrated that solutes travelled faster to the groundwater in a water-repellent soil than in a wettable soil. While for initial water content below a critical value preferential flow paths developed, for moister conditions uniform flow could occur (Ritsema & Dekker, 2000).

(iii) Soil structure and initial soil moisture. Structured soils are more susceptible to develop preferential flow than unstructured ones. In

contrast, the effect of initial soil moisture is less evident (Flury *et al.*, 1994; Vogel *et al.*, 2006).

(iv) Biological factors such as distribution of roots (Mitchell *et al.*, 1995) and earthworm burrows (Farenhorst *et al.*, 2000; Shuster *et al.*, 2002; Weiler & Naef, 2003) that increase flow rates by creating macropores.

Preferential flow can further be classified in macropore flow, unstable flow and funnel flow. Macropore flow describes water movement along root channels, earthworm burrows, soil fissures, cracks or large inter-aggregate voids in cultivated soils (Beven & Germann, 1982). A detailed review on macropore flow can be found in Jarvis (2007). Unstable flow is common in coarse-textured soils and can be induced by variations in texture, water repellency, air entrapment or continuous non-ponding infiltration. Lateral redirection of water caused by textural boundaries is referred to as funnel flow (Hendrickx & Flury, 2001).

In summary, non-uniform flow results in rapid water movement and solute transport to greater depths than predicted by the classical concept of flow through unsaturated soils (Richards' equation). The impact on groundwater quality might be considerable if preferential flow – especially through continuous macropores – becomes a dominant mechanism of transport of agrochemicals to greater depths (Jarvis, 2007).

## 1.2 Visualization and analysis of flow patterns

Uniform and preferential flow lead to different flow patterns that can be visualised by dye tracer experiments (e.g. Pickering *et al.*, 1988; Flury *et al.*, 1994; Forrer *et al.*, 2000). Brilliant Blue FCF – a food dye – is frequently used in vadose zone hydrology for such tracing studies, because it is easily seen against most soil colours and has acceptable toxicological characteristics for environmental use (Flury & Flühler, 1994). Among numerous studies using staining techniques to analyse water flow in soils two major groups can be distinguished, one using images of stained profiles for pattern analysis and another deriving concentration maps of the dye tracer. Our research is part of the first group.

Flow regime (uniform or preferential) and flow patterns are interrelated. Weiler & Flühler (2004) for instance used the width of stained objects in order to identify flow regimes in macroporous soils. To distinguish between different

transport mechanisms Kulli *et al*. (2003) performed a hierarchical clustering of dye patterns. Other investigations were based on the proportion of staining, the dye coverage (Flury *et al.*, 1994). Öhrström *et al*. (2002), for instance, analysed spatial variability of preferential pathways at catchment scale and found that variability of dye patterns increased with scale. To investigate the influence of water repellency on flow patterns Lipsius & Mooney (2006) compared dye coverage functions by analysis of variance. Numerous other studies analysing the percentage of stained area as a function of soil depth exist (e.g. Weiler & Naef, 2003; Sander & Gerke, 2007). This variability of techniques shows clearly the difficulty to summarize the relevant information from dye tracer studies in a compact way. Thus appreciation of susceptibility of soils to induce preferential flow and of a potential vulnerability of groundwater is often based on expert knowledge. This lack of theoretical foundation makes comparisons between soils difficult and subjective.

To overcome this problem Schlather & Huwe (2005) introduced a new theoretical base for the dye coverage function. They used the peak-over-threshold approach that usually serves to investigate extreme events in time-series analysis and reinterpreted the dye coverage as a probability function to find the tracer below a certain depth. By fitting the *Generalised Pareto Distribution* – a two parametrical extreme value distribution – to the dye coverage they defined the form parameter as a risk index for the vulnerability of groundwater by pollutants. Thus, they summarised data from a dye staining experiment in one real-valued quantity. Furthermore, the authors hypothesised that the risk index was an intrinsic soil parameter showing some persistence of its value against small variations of boundary, i.e. experimental and initial conditions.

## *1.3 Objectives of the thesis*

The first goal of the theses was to consider in detail the behaviour of the risk index for different experimental and initial conditions. We wanted to know whether extreme value statistics could be used as a tool to analyse and classify flow patterns under varying boundary conditions. This first study was conducted in a Norway spruce forest in southeast Germany (Chapter 2.1).

The second objective was the application of the risk index to stained patterns obtained from different soils in different ecosystems. We worked in a tropical mountain rainforest in southern Ecuador – an environment characterised by a varying degree of anthropogenic disturbance – and on an agricultural filed with shallow ground water in southern France (Chapter 2.2 and Chapter 2.3).

And last but not least, after analysis of flow patterns we investigated the underlying flow processes and their impacts on soil chemistry at the German study site in details (Chapter 3).

These aims required the refinement of field, laboratory and data analysis methods. For field wok, the sprinkler proposed by Ghodrati *et al*. (1990) had to be modified in order to improve its portability and flexibility of control. A laboratory method to rapidly determine the concentration of Brilliant Blue in soil samples was lacking (Appendix A). And finally, we adapted image analysis techniques originally developed for industrial machine vision to process pictures of stained patterns in soils.

## *1.4 Experimental sites*

### 1.4.1 Waldstein in the Fichtelgebirge (Germany)

The experimental site is located in Waldstein in the Fichtelgebirge (775 m a.s.l), a mountain ridge in southeast Germany, densely forested by the Norway spruce (*Picea abies*). The mean annual temperature is 5.3 °C and annual precipitation 1156.5 mm. Main soil groups are Haplic Poszols, Dystric Cambisols and Hostosols (IUSS Working Group WRB, 2007) developed on granite or gneiss bedrock. The content of rock fragments varies strongly within the site and with soil depths and ranges between few percent in the topsoil and up to 75 vol% in the subsoil. Extremely acidic pH values (in $CaCl_2$) between 2.6 (organic horizons) and 4.3 (subsoil) were measured. High deposition of $SO_4$ during several decades resulted in a strong acidification of the soil. A detailed description of the site can be found in Gerstberger *et al*. (2004); Foken (2003) gives an overview on climatic and meteorological conditions.

### 1.4.2 ECSF research station (Ecuador)

The investigated area is situated in the Cordillera Real, in the eastern part of South Ecuadorian Andes and belongs to the "Reserva Biologíca San Francisco" that extends between 1800 m and 3160 m a.s.l. It comprises undisturbed tropical mountain rain forest and its anthropogenic replacement ecosystem where the forest has been cleared by slash and burn. The centre of the research area is the "Estacion Cientiíica San Francisco" (ECSF). The annual temperature averages 19-13 °C between 1100-2200 m and 13-6 °C up to 3800 m. Since the spatial rainfall structure is highly heterogeneous, the annual precipitation in the studied area varies between 383 mm and >6000 mm. The soils are highly heterogeneous due to the diversity of parent material (metasiltsontes, sandstones, quartzites and schists) and altitudinal gradients. Additionally, the area displays a high landslide activity; therefore, the stone content in the topsoil varies between 80% on landslides and less than 10% in undisturbed forest soils. More details on the studied area are given in Beck *et al*. (2008) and references herein.

### 1.4.3 INRA experimental site in Avignon (France)

The research area is located on an alluvial plane at the confluence of Rhone and Durance rivers at 30 m a.s.l (Monjuvent, 1991). The alluvial sediments constitute an important aquifer with a shallow water table that can rise up to 3 m below the surface. The mean annual maximum and minimum temperature equals 19.3 °C and 9.1 °C respectively and the mean annual precipitation is 722.4 mm (Météo France). The soil is a Calcisol (IUSS Working Group WRB, 2007) with a pH (in $H_2O$) of 8.5. It is highly structured due to important content of clay (>35%) and develops deep cracks during desiccation. Detailed overview on soil physical and chemical properties can be found in Di Pietro *et al.*, (2004).

## *1.5 Sy nopsis*

### 1.5.1 Behaviour of the risk index under different experimental and initial conditions (Chapter 2.1)

To consider in detail the behaviour of the risk index for different experimental and initial conditions we did five tracer experiments in a Norway spruce forest in southeast Germany. We applied a tracer solution consisting of Brilliant Blue and the reference tracer potassium iodide at 32 mm $hour^{-1}$ or 64 mm $hour^{-1}$ on plots of

about 2 $m^2$. The reference tracer was used to verify that Brilliant Blue was not substantially retarded and reflected well the flow patterns of water. Indeed, rainfall simulations with Brilliant Blue and iodide on soil columns extracted at the same study site showed clearly that when matrix flow dominated Brilliant Blue was heavily retarded with regard to iodide. In this case the transport behaviour of Brilliant Blue differs from flow behaviour of infiltrating water (Appendix B). The day after the irrigation, six vertical 1 m × 1 m soil profiles were excavated at intervals of 20 cm in the central part of the plot and photographed with a CCD camera. One to three profiles were treated with iron(III) nitrate and starch to visualize iodide (Lu & Wu, 2003). To account for changing light conditions in a forest environment we took photographs in the lossless RAW format in order to adjust white balance *a posteriori* in appropriate graphical software. Images of soil profiles were corrected for radial and perspective distortion using the software PTGui (New House Internet Services B.V., 2005), classified in stained and non stained parts and the dye coverage function $p(d)$ (number of stained pixels per depth $d$) was calculated in Matlab 7.1 (The MathWorks, 2005b). The distribution $1 - H$ ($H$ being the generalised Pareto distribution) was fitted to the dye coverage function in order to estimate the risk index as defined by Schlather & Huwe (2005).

In this study we slightly modified the interpretation of the risk index. We think that the form parameter of the generalized Pareto distribution should be interpreted as a risk index for vulnerability of groundwater to pollutants only in regions with fairly homogeneous geological material between the soil surface and the water table, as in sedimentary basins with shallow water tables. The groundwater at our site is 8 to 10 m below the surface, and so we prefer to qualify $\xi_r$ as a risk index for vertical solute propagation.

Our results support the hypothesis formulated by Schlather and Huwe (2005) that the risk index is to some degree invariant to changing experimental conditions (such as irrigation rate). The initial soil moisture, however, seems to have a large influence on the risk index. We propose to combine the two parameters of the generalized Pareto distribution to estimate the risk of vertical solute propagation in soils. A complete probability distribution $1 - H$ allows us to compare plots with various initial and experimental conditions or different tracers.

As mentioned by Schlather & Huwe (2005), in stratified soils when the flow process changes as a result of varying physical properties the dye coverage function cannot be represented by one single distribution $1 - H$. Therefore, we used only the lowest part of the profile to calculate the risk index. This accords with the limit law of the extreme value theory stating that the behaviour of the process at great depths is independent of the behaviour near the origin (Schlather & Huwe, 2005). Additionally, as our analysis is based on extreme values of vertical solute propagation it is the lowest part of the profile that is the most interesting. To correctly access the risk of vertical solute propagation the depth where $1 - H$ was adjusted should be explicitly taken into account. Provided that the dye coverage function decreases monotonically, the estimated risk for vertical solute propagation can serve to classify soils.

### 1.5.2 Application of the risk index to stained patterns obtained from different soils in different ecosystems (Chapter 2.2 and Chapter 2.3)

To apply the risk index to soils in a different ecosystem we did five tracer experiments in a mountain rainforest of south Ecuador (Chapter 2.2). Most tracer studies documented in literature were carried out on agricultural soils. To the authors' knowledge, there are only few studies in young stony forest soils, landslides or in tropical soils in general. Therefore, the aim of this study was to investigate water flow in disturbed and undisturbed tropical soils, with special emphasis on stony forest soils. We did tracer experiments on a new landslide and on an old one, at two sites in a primary forest and on pastures. The field work was adjusted to difficult conditions in a mountainous tropical region and only one tracer (Brilliant Blue) was used. Image processing was similar to the first study except that the camera did not allow taking images in the RAW format.

Infiltration depth and pattern complexity were higher on the old landslide than on the new one probably due to soil regeneration. Indeed, on the new landslide soil structure was destroyed by mass movement, producing a more or less heterogeneous mixture of soil material and stones. Pedogenetic processes and plant activities recreated soil structure on the old landslide, thus increasing the occurrence of preferential flow especially along bio-macropores such as root

channels or earthworm burrows. This was well reflected by the risk of vertical solute propagation that was higher on the old landslide.

In the primary forest, stones amplified the development of preferential flow by funnelling water along their walls and reducing the cross-section for water flow thus inducing a higher risk of solute propagation. Stained patterns observed on pastures suggested that infiltrability was reduced by compaction of the soil surface due to changes in land use. This could explain the smaller amount of dye penetrated into the soil. The calculated risk index indicates a low propagation risk at this site.

From this study we conclude that the complexity of stained patterns and infiltration depths are different at disturbed and undisturbed study sites. Our data indicate that soil structure, stone content, plant roots system and possibly land use are controlling factors for water flow in soils of this study area. The index proposed by Schlather & Huwe (2005) serves as a useful characteristic of flow regime and as a risk index for vertical solute propagation in a variety of soils.

Thereafter, we characterised the flow regime on an agricultural study site in southern France (Chapter 2.3). In cultivated soils tillage is an important factor that influences processes on soil surface, is important for generation of surface runoff and affects soil erodibility and depression storage. Rough soil surfaces tend to store more water thus reducing runoff and increasing infiltration. While many studies on soil surface processes or flow patterns in soils exist, works relating both are rare in literature. We thus combined stereophotogrammetry to measure soil surface roughness and dye staining techniques on a tilled and compacted plot and a non-tilled one with a well developed system of open cracks to investigate processes at soil surface during simulated rainfall, analyse the resulting flow patterns and to relate both to tillage treatment. A special emphasise was given to connectivity of processes above and below the plough pan.

For this study, image analysis techniques originally used in machine vision were adapted for tracer studies. We calibrated the camera to correct radial and geometric distortion in images of stained soil profiles more accurately and rapidly. Additionally, the classification method to segment images in stained and non stained parts was adjusted. Since image analysis could be considerably speeded up soil profiles were excavated at intervals of 10 cm.

Our results showed that on the non-tilled plot soil roughness increased after tracer application probably due to surface consolidation. However, on the tilled plot it did not change because the applied simulated rainfall had not enough energy to induce significant erosion. Important surface storage was observed on the tilled plot, especially on the compacted part. Despite tillage preferential flow along cracks occurred on both plots and macropores buried below tillage pan functioned as preferential flow paths.

We could demonstrate that the studied soil was susceptible to deep vertical solute propagation at dry no till conditions when cracks are open, connected down to the subsoil and infiltration rates through macropores are important. Keeping in mind the shallow water table at this site, the risk for groundwater contamination is high. Tillage destroys macroporosity in the topsoil and disconnects it from the subsoil, thus reducing the amount of solutes infiltrating into greater depths. However, it does not prevent macropore flow so that buried macropores could still function as preferred transport paths and allow solute and pesticide leaching down to the groundwater.

### 1.5.3 Flow processes and their impacts on soil chemistry (Chapter 3).

Pattern analysis based on the risk index for vertical solute propagation revealed the occurrence of preferential flow at the German study site. The second part of this thesis is devoted to statistical analysis of mechanisms of preferential flow (Chapter 3.1) and its impacts on soil chemistry (Chapter 3.2). We did three tracer experiments and qualitatively identified the dominant flow regime based on stained patterns. We sampled soil material and analysed soil texture, fine root density and soil bulk density from preferential flow paths and the soil matrix and tested whether these parameters differed significantly between regions of preferential flow and soil matrix and could give insight into mechanisms of water flow.

The data were sampled hierarchically on three plots, in several profiles and different horizons. Since some horizons were bypassed by the flow we had to deal with missing values. Hierarchical sampling might induce dependencies in data so that classical statistical techniques like the analysis of variance are not applicable. Most studies that investigate differences between preferential flow paths and soil

matrix ignore the hierarchical nature of data sampled from different plots and use the paired t-test of mean values or its non-parametric equivalent and test different depths separately. We propose to employ mixed-effects models and to consider all plots and all depths in one single analysis. Mixed-effects models can account for fixed-effects representing parameters of the entire population or certain repeatable levels of experimental factors (like horizon) and for random-effects associated with individual experimental units drew at random from a population (like plots or profiles). Furthermore, they are robust against missing values (Pinheiro & Bates, 2000).

Our data showed that at this study site roots constituted main preferential flow paths and induced macropore flow, especially in the topsoil. In the subsoil, root density decreased and inhomogeneous infiltration from preferential flow paths into the soil matrix caused unstable flow. Due to the large sand content (i.e. high permeability) the dye spread from preferential flow paths into the soil matrix creating large stained objects. We found no significant differences in soil texture between preferential flow paths and soil matrix. In contrast, fine root density was higher in preferential flow paths indicating the importance of roots channels as macropores. Soil bulk density was lower in preferential flow paths probably because of higher organic matter content. Root turnover is an important source of soil carbon and decomposition of dead roots is a major input to soil organic matter (Tate *et al.*, 1993; Guo *et al.*, 2005). Soil bulk density is known to decrease with increasing content of organic matter (e.g. Balland *et al.*, 2008).

Root macropores promote preferential transport of solutes from the organic horizons to the subsoil. Furthermore, roots are known to strongly influence their immediate environment, the rhizosphere, by exudation of organic compounds. As a consequence distinct chemical compartments might develop with gradients in the transition zone between the soil matrix and preferential flow paths. For that reason, we analysed exchangeable cations, pH, and total C and N contents in the same soil samples to elucidate eventual impacts of preferential flow on soil chemistry by means of mixed-effects modelling.

Brilliant Blue ($C_{37}H_{34}N_2Na_2O_9S_3$, molar mass 792.9 g $mol^{-1}$) is an organic molecule consisting of 56% C and 4% N. Sorption of the dye on soil particles affects the C and N contents of soil and should be corrected. Usually, Brilliant

Blue concentrations are determined by extracting the dye with a water acetone solution or a 0.5 M $K_2SO_4$ (e.g. Bundt *et al.*, 2001). This is a laborious procedure with changing accuracy due to varying mass recovery (Forrer *et al.*, 2000). We developed a method to measure the content of Brilliant Blue by visible diffuse reflectance spectroscopy (VIS-DRS) directly on soil samples without extraction (Appendix A). We corrected the content of total C and N for presence of Brilliant Blue prior to mixed-effects modelling.

We found smaller pH values more Ca, more Mg, more C and more N in preferential flow paths. Compared to the adjacent soil matrix, more Al and more Fe (but small absolute amounts) were found in the subsoil where macropore flow along root channels decreases and heterogeneous matrix flow dominates. These distinct chemical properties can be explained by root activity and translocation of solutes via preferential flow paths. Higher Ca and Mg concentration in preferential flow paths are probably due to transport from the soil surface after liming. Smaller pH values could be explained by transport of acid soil solution from organic horizons along preferential flow paths. Higher Al and Fe concentration in the subsoil probably results from release and translocation of these solutes during podzolisation. Rhizodeposition of organic compounds, decomposition of dead roots and transport of DOC from organic to mineral horizons are major sources of organic C input to the soil (e.g. Kuzyakov & Komansky (2000)). Higher root densities in preferential flow paths lead to a higher C input through roots, but also facilitate preferential transport of DOC. Indeed, there is strong experimental indication of transport of DOC via preferential flow paths at our study site (Schulze *et al*. 2009). DOC is strongly adsorbed in soils by Al and Fe oxides/hydroxides and clay minerals (Kalbitz *et al.*, 2000). During transport along preferential flow paths contact time between DOC and soil is reduced so that DOC is transported to greater depth where it potentially form organo-mineral associations. If this holds true, preferential flow is a mechanism that promotes C sequestration in subsoil. We conclude that preferential flow does not only influence its immediate environment around paths, but also underlying subsoil horizons.

### 1.5.4 Concluding remarks

This thesis began with the idea to summarize the main information from dye tracer experiments in one real-valued index, the risk index for vulnerability of groundwater to pollutants developed by Schlather & Huwe (2005). Bearing in mind the boundary conditions of the tracer experiment like irrigation rate, the tracer employed, soil initial moisture or type of vegetation (permanent or seasonal, deep rooted or shallow rooted) it is possible to compare different study sites or the same site at different boundary conditions and to access the risk of vertical solute propagation using extreme value statistics.

By means of pattern analysis we identified main flow mechanisms and investigated flow processes and their impacts on soil chemistry. We have successfully applied a particular type of statistical analysis, the mixed-effects model to reveal distinct physico-chemical properties in preferential flow paths and soil matrix. Based on this type of analysis conclusions about differences between preferential flow paths and soil matrix are not affected by between-plot or between-sample variability.

A major outcome of this thesis is the large number of images of flow patterns from different soils. Images, even binary ones, are high dimensional objects. Further studies could employ recent dimensionality reduction techniques to investigate whether there is a low dimensional structure underlying these images (Weinberg & Saul, 2006).

## 1.6 List of manuscripts and specification of own contributions

This thesis includes six manuscripts. Two of them are published, three are submitted to the European Journal of Soil Science and one is in preparation for the Soil & Tillage Research. The list below details the contributions of all co-authors.

**Manuscript 1**

| | | |
|---|---|---|
| Authors | C. Bogner, B. Wolf, M. Schlather & B. Huwe | |
| Title | Analysing flow patterns from dye tracer experiments in a forest soil using extreme value statistics | |
| Status | published | |
| Journal | European Journal of Soil Science | |
| Contributions | | |
| Bogner | 40% | idea, methods, data collection, data analysis, manuscript writing, figures, discussion, editing, corresponding author |
| Wolf | 35% | idea, data collection, data analysis, discussion |
| Schlather | 15% | idea, methods, manuscript writing, discussion, editing |
| Huwe | 10% | idea, discussion, editing |

**Manuscript 2**

| | | |
|---|---|---|
| Authors | C. Bogner, S. Engelhardt, J. Zeilinger & B. Huwe | |
| Title | Visualization and analysis of flow patterns and water flow simulations in disturbed and undisturbed tropical soils | |
| Status | published | |
| Book | Gradients in a Tropical Mountain Ecosystem of Ecuador (eds. Beck, E., Bendix, J., Kottke, I., Makeschin, F. & Mosandl, R.) | |
| Contributions | | |
| Bogner | 60% | idea, data collection, data analysis, manuscript writing, figures, discussion, editing, corresponding author |
| Engelhardt | 10% | idea, data collection |
| Zeilinger | 20% | idea, data collection, data analysis, discussion, editing |
| Huwe | 10% | idea, discussion, editing |

**Manuscript 3**

| | | |
|---|---|---|
| Authors | C. Bogner, M. Mirzaei, S. Ruy & B. Huwe | |
| Title | Relating flow patterns and processes on soil surface in an agricultural soil | |
| Status | in preparation | |
| Journal | Soil & Tillage Research | |
| Contributions | | |
| Bogner | 65% | idea, methods, data collection, data analysis, manuscript writing, figures, discussion, editing, corresponding author |
| Mirzaei | 10% | idea, methods, data collection |
| Ruy | 15% | idea, methods, data collection, discussion, editing |
| Huwe | 10% | idea, discussion, editing |

**Manuscript 4**

| | | |
|---|---|---|
| Authors | C. Bogner, D. Gaul, A. Kolb, I. Schmiedinger & B. Huwe | |
| Title | Investigating flow mechanisms in a forest soil by mixed-effects modelling | |
| Status | Submitted 7 April 2009 | |
| Journal | European Journal of Soil Science | |
| Contributions | | |
| Bogner | 70% | idea, methods, data collection, data analysis, manuscript writing, figures, discussion, editing, corresponding author |
| Gaul | 10% | data collection, data analysis, discussion |
| Kolb | 5% | data collection |
| Schmiedinger | 5% | data collection |
| Huwe | 10% | idea, discussion, editing |

**Manuscript 5**

| | | |
|---|---|---|
| Authors | C. Bogner, W. Borken & B. Huwe | |
| Title | Impact of preferential flow on soil chemistry in a forest soil | |
| Status | Submitted 7 April 2009 | |
| Journal | European Journal of Soil Science | |
| Contributions | | |
| Bogner | 80% | idea, methods, data collection, data analysis, manuscript writing, figures, discussion, editing, corresponding author |
| Borken | 10% | idea, discussion, editing |
| Huwe | 10% | idea, discussion, editing |



**Manuscript 6**

| | | |
|---|---|---|
| Authors | C. Bogner, I. Schmiedinger & B. Huwe | |
| Title | Rapid analysis of Brilliant Blue by diffuse reflectance spectroscopy | |
| Status | Submitted 7 April 2009 | |
| Journal | European Journal of Soil Science | |
| Contributions | | |
| Bogner | 80% | idea, methods, data collection, data analysis, writing, figures, discussion, editing, corresponding author |
| Schmiedinger | 10% | method, data collection |
| Huwe | 10% | idea, discussion, editing |

## *1.7 References*

Balland, V., Pollacco, J.A.P. & Arp, P.A. 2008. Modeling soil hydraulic properties for a wide range of soil conditions. *Ecological Modelling*, **219,** 300-316.

Beck, E., Makeschin, F., Haubrich, F., Richter, M., Bendix, J. & Valerezo, C. 2008. The Ecosystem (Reserva Biológica San Francisco). In: *Gradients in a Tropical Mountain Ecosystem of Ecuador* (eds. Beck, E., Bendix, J., Kottke, I., Makeschin, F. & Mosandl, R.), pp. 1-13. Springer Verlag, Berlin Heidelberg.

Beven, K. & Germann, P. 1982. Macropores and water flow in soils. *Water Resources Research*, **18,** 1311-1325.

Bundt, M., Jaggi, M., Blaser, P., Siegwolf, R. & Hagedorn, F. 2001. Carbon and nitrogen dynamics in preferential flow paths and matrix of a forest soil. *Soil Science Society of America Journal*, **65,** 1529-1538.

Di Pietro, L., Doussan, C., Ruy, S., Gaudu, J.C., Perrin, P., Pallut, R. & Renard, D. 2004. *Infiltration, recharge, écoulements préférentiels et transport de polluants vers les nappes.* Rapport N° 1. Centre de recherche, INRA Avignon, Avignon.

Farenhorst, A., Topp, E., Bowman, B.T. & Tomlin, A.D. 2000. Earthworm burrowing and feeding activity and the potential for atrazine transport by preferential flow. *Soil Biology & Biochemistry*, **32,** 479-488.

Flury, M. & Flühler, H. 1994. Brilliant Blue FCF as a dye tracer for solute transport studies - a toxicological overview. *Journal of Environmental Quality*, **23,** 1108-1112.

Flury, M., Flühler, H., Jury, W.A. & Leuenberger, J. 1994. Susceptibility of soils to preferential flow of water: a field study. *Water Resources Research*, **30,** 1945-1954.

Foken, T. 2003. *Lufthygienisch-Bioklimatische Kennzeichnung des oberen Egertales*. Bayreuther Forum Ökologie. Vol. 100, 69+XLVIII.

Forrer, I.E., Papritz, A., Kasteel, R., Flühler, H. & Luca, D. 2000. Quantifying dye tracers in soil profiles by image processing. *European Journal of Soil Science*, **51,** 313-322.

Gerstberger, P., Foken, T. & Kalbitz, K. 2004. The Lehstenbach and Steinkreuz catchments in NE Bavaria, Germany. In: *Biogeochemistry of Forested Catchments in a Changing Environment* (ed. Matzner, E.), pp. 15-44. Springer Verlag, Berlin Heidelberg.

Ghodrati, M., Ernst, F.F. & Jury, W.A. 1990. Automated spray system for application of solutes to small field plots. *Soil Science Society of America Journal*, **54,** 287-290.

Gish, T.J., Kung, K.-J.S., Perry, D.C., Posner, J., Bubenzer, G., Helling, C.S., Kladivko, E.J. & Steenhuis, T.S. 2004. Impact of preferential flow at varying irrigation rates by quantifying mass fluxes. *Journal of Environmental Quality*, **33,** 1033-1040.

Green, W.H. & Ampt, G.A. 1911. Studies on soil physics, 1. The flow of air and water through soils. *The Journal of Agricultural Science*, **4,** 1-24.

Guo, L.B., Halliday, M.J., Siakimotu, S.J.M. & Gifford, R.M. 2005. Fine root production and litter input: Its effects on soil carbon. *Plant and Soil*, **272,** 1-10.

Hendrickx, J.M.H. & Flury, M. 2001. Uniform and Preferential Flow Mechanisms in the Vadose Zone. In: *Conceptual Models of Flow and Transport in the Fractured Vadose Zone* (ed. Council, N.R.), pp. 149-187. National Academy Press, Washington, DC.

Hendrickx, J.M.H., Dekker, L.W. & Boersma, O.H. 1993. Unstable wetting fronts in water-repellent field soils. *Journal of Environmental Quality*, **22,** 109-118.

IUSS Working Group WRB. 2007. World reference base for soil resources 2006, first update 2007. World Soil Resources Reports No. 103.

Jarvis, N.J. 2007. A review of non-equilibrium water flow and solute transport in soil macropores: principles, controlling factors and consequences for water quality. *European Journal of Soil Science*, **58,** 523-546.

Kalbitz, K., Solinger, S., Park, J.H., Michalzik, B. & Matzner, E. 2000. Controls on the dynamics of dissolved organic matter in soils: A review. *Soil Science*, **165,** 277-304.

Kulli, B., Stamm, C., Papritz, A. & Fluhler, H. 2003. Discrimination of flow regions on the basis of stained infiltration patterns in soil profiles. *Vadose Zone Journal*, **2,** 338-348.

Kuzyakov, Y. & Domanski, G. 2000. Carbon input by plants into the soil. Review. *Journal of Plant Nutrition and Soil Science - Zeitschrift für Pflanzenernährung und Bodenkunde*, **163,** 421-431.

Lawes, J.B., Gilbert, J.H. & Warrington, R. 1882. *On the amount and composition of the rain and drainage water collected at Rothamsted*. Williams, Clowes & Sons, London.

Lipsius, K. & Mooney, S.J. 2006. Using image analysis of tracer staining to examine the infiltration patterns in a water repellent contaminated sandy soil. *Geoderma*, **136,** 865-875.

Lu, J. & Wu, L. 2003. Visualizing bromide and iodide water tracer in soil profiles by spray methods. *Journal of Environmental Quality*, **32,** 363-367.

Météo France. <http://france.meteofrance.com/france/climat_france>

Mitchell, A.R., Ellsworth, T.R. & Meek, B.D. 1995. Effect of root systems on preferential flow in swelling soil. *Communications in Soil Science and Plant Analysis*, **26,** 2655-2666.

Monjuvent, G. 1991.*Carte géologique de la France à 1:50 000. 940. Avignon*. Bureau de recherches géoologiques et minières, Paris.

New House Internet Services B.V. 2005. PTGui version 5.5. http://www.ptgui.com/. Accessed on January 2007.

Niemann, T. 2005. PTLens version 6.4. http://epaperpress.com/ptlens/. Accessed on January 2007.

Öhrstöm, P., Persson, M., Albergel, J., Zante, P., Nasri, S., Berndtsson, R. & Olsson, J. 2002. Field-scale variation of preferential flow as indicated from dye coverage. *Journal of Hydrology*, **257,** 164-173.

Pickering, N.B., Richard, T.L., Parlange, J.Y., Waltman, W.J. & Kromm, L. 1988. *Dye tracer analysis of preferential flow to groundwater*. Department of Agricultural and Biological Engineering, Staff Report 88-2. Cornell University, Ithaca, N. Y.

Pinheiro, J.C. & Bates, D.M. 2000. *Mixed-Effects Models in S and S-Plus*. Springer, New York.

Ritsema, C.J. & Dekker, L.W. 2000. Preferential flow in water repellent sandy soils: principles and modeling implications. *Journal of Hydrology*, **231,** 308-319.

Sander, T. & Gerke, H.H. 2007. Preferential flow patterns in paddy fields using a dye tracer. *Vadose Zone Journal*, **6,** 105-115.

Schlather, M. & Huwe, B. 2005. A risk index for characterising flow pattern in soils using dye tracer distributions. *Journal of Contaminant Hydrology*, **79,** 25-44.

Schulze, K., Borken, W. & Matzner, E. 2009. Dynamics of dissolved organic $^{14}C$ signature in throughfall and soil solution of a Norway spruce forest. *submitted to Biogeochemistry*.

Shuster, W.D., Subler, S. & McCoy, E.L. 2002. The influence of earthworm community structure on the distribution and movement of solutes in a chisel-tilled soil. *Applied Soil Ecology*, **21,** 159-167.

Tate, K.R., Ross, D.J., Obrien, B.J. & Kelliher, F.M. 1993. Carbon storage and turnover, and respiratory activity, in the litter and soil of an old-growth southern beech (Nothofagus) forest. *Soil Biology & Biochemistry*, **25,** 1601-1612.

The MathWorks, Inc. 2005b. Matlab version. 7.1. http://www.mathworks.com

Tymchak, M.P. & Torres, R. 2007. Effects of variable rainfall intensity on the unsaturated zone response of a forested sandy hillslope. *Water Resources Research*, **43,** 13.

Vogel, H.-J., Cousin, I., Ippisch, O. & Bastian, P. 2006. The dominant role of structure for solute transport in soil: experimental evidence and modelling of structure and transport in a field experiment. *Hydrology and Earth System Sciences*, **10,** 495-506.

Weiler, M. & Flühler, H. 2004. Inferring flow types from dye patterns in macroporous soils. *Geoderma*, **120,** 137-153.

Weiler, M. & Naef, F. 2003. An experimental tracer study of the role of macropores in infiltration in grassland soils. Hydrological Processes, 17, 477-493.

Weinberger, K.Q. & Saul, L.K. 2006. Unsupervised learning of image manifolds by semidefinite programming. International Journal of Computer Vision, 70, 77-90.

# Chapter 2

# Analysis of flow patterns

## *2.1 Analysing flow patterns from dye tracer experiments in a forest soil using extreme value statistics*

C. BOGNER[a], B. WOLF[a], M. SCHLATHER[b] & B. HUWE[a]

[a]*Soil Physics Group, University of Bayreuth, 95440 Bayreuth and* [b]*Institute for Mathematical Stochastics, Georgia Augusta University Göttingen, Maschmühlenweg 8-10, 37073 Göttingen, Germany*

Correspondence: C. Bogner. E-mail: christina.bogner@uni-bayreuth.de

Short title:
*Analysing dye patterns by extreme value statistics*

Received 31 January 2007; revised version accepted 13 September 2007

Published in the *European Journal of Soil Science*, February 2008, **59**, 103–113

British Society of Soil Science, Blackwell Publishing

doi: 10.1111/j.1365-2389.2007.00974.x

The definitive version is available at: www.blackwell-synergy.com

## Summary

Preferential flow of water in soil is now recognized as a common phenomenon. It results in complex flow patterns that can be visualized by dye tracers and increases the risk of pollutants' reaching greater depths. We analyzed the behaviour of a risk index for vertical solute propagation based on extreme value theory. This risk index can be calculated from binary images of dye stained soil profiles and is defined as the form parameter of the generalized Pareto distribution. We did five tracer experiments with Brilliant Blue and iodide under changing experimental (variable initial soil moisture) and initial conditions (different irrigation rates). Our results indicate some persistence of the risk index against small changes of experimental conditions such as the irrigation rate. On the other hand, it seems to be affected by initial soil moisture. Comparisons of Brilliant Blue and iodide patterns show that the form parameter alone is not sufficient to estimate the risk of vertical solute propagation. Therefore we propose to combine the risk index with the scale parameter of the generalized Pareto distribution.

### 2.1.1 Introduction

Although preferential flow of water in soil was discovered in the late 19th century (Schumacher, 1864; Lawes *et al.*, 1882), it was considered for a long time as exceptional. Today, it is regarded as a common phenomenon that depends on the spatial heterogeneity and intensity of rainfall (Gish *et al.*, 2004), water repellency (Hendrickx *et al.*, 1993; Ritsema & Dekker, 2000; Wang *et al.*, 2000), soil structure (Flury *et al.*, 1994; Kulli *et al.*, 2003; Vogel *et al.*, 2006) and biological factors such as the distributions of roots (Mitchell *et al.*, 1995) and earthworm burrows (Farenhorst *et al.*, 2000; Shuster *et al.*, 2002; Weiler & Naef, 2003). Preferential flow results in complex flow patterns that can be visualized by dye tracers. Brilliant Blue is frequently used in vadose zone hydrology for such tracing studies although its adsorption behaviour is non-linear and depends on soil properties (Ketelsen & Meyer-Windel, 1999; German-Heins & Flury, 2000; Kasteel *et al.*, 2002). However, it is readily seen against most soil colours and has acceptable toxicological characteristics for environmental use (Flury & Flühler, 1994; Mon *et al.*, 2006).

Usually, the main information obtained from dye stained profiles are binary images – photographs of soil profiles that are classified in stained (black) and unstained (white) parts. They are used for qualitative description of flow regimes and for the visualization of preferential flow (Öhrstöm *et al.*, 2002; Kulli *et al.*, 2003; Weiler & Naef, 2003). Recent studies, however, took a quantitative approach to tracer studies in soils by establishing dye concentration maps (Aeby *et al.*, 1997; Forrer *et al.*, 1999; Forrer *et al.*, 2000). This method needs calibration because the same dye concentration has different hues depending on soil colour. Forrer *et al.* (2000) reported 203 calibration samples in a 'fairly uniform Eutric Cambisol' on an agricultural field. Morris & Mooney (2004) used 100 samples to assess concentrations in a small intact soil block (200 mm × 200 mm × 200 mm).

Such a calibration becomes complicated for soils with progressively changing colours because the number of calibration samples increases rapidly. This is the case at our study site. Indeed, we can distinguish four or five different main hues in our soil and varying degrees of their combinations. Each of these hues needs its own calibration between the possible concentration range of Brilliant Blue and the resulting RGB (Red, Green and Blue) values on images. Therefore, we needed some other approach to obtain quantitative information on flow processes from stained profiles, one that does not require any information on dye concentrations in the soil. Schlather & Huwe (2005) propose a risk index for groundwater vulnerability to pollutants based on extreme value theory. It can be calculated from binary images of dye stained soil profiles and does not require any additional information on soil properties. The goal of our study is to consider in detail the behaviour of the risk index for different experimental and initial conditions.

## 2.1.2 Materials and methods

### *Dye tracer experiments*

We did five tracer experiments in a Norway spruce forest in southeast Germany. The soil is a Cambisol or a Cambic Podzol with loam or sandy loam above loamy sand. The stone content is medium to high and the pH is 4. We used Brilliant Blue FCF and iodide as tracers. The latter served as reference because Brilliant Blue may be retarded with regard to infiltrating water as a result of adsorption on soil particles. Bowman (1984) reported that the sorption behaviour of iodide is similar to that of bromide, which is considered as the most suitable tracer for water

movement in soil (Flury & Wai, 2003). In order to have the same spatial resolution of flow patterns for both tracers, we visualised iodide by a spray method proposed by Lu & Wu (2003). Following this we applied a solution of iron(III) nitrate and starch directly on the excavated soil profile. Iron(III) oxidized iodide to iodine, which formed a dark-blue complex with starch. This method worked well; however, the time reported by Lu & Wu (2003) of about 1 to 2 hours for the colour reaction was not sufficient for a good contrast to Brilliant Blue dye, and we let it develop during the night. Lu & Wu (2003) also proposed a visualization method for bromide, but the Prussian blue complex formed has a blue colour that would be too difficult to distinguish from Brilliant Blue.

We applied 64 mm of tracer solution on plots of about 2 $m^2$ using a sprinkler similar to that proposed by Ghodrati *et al.* (1990). The irrigation rate was either 32 mm $hour^{-1}$ (referred to as 'low') or 64 mm $hour^{-1}$ ('high'), and the concentration of both tracers in the solution was 5 g $l^{-1}$. The maximum 10-minutes intensity recorded at the study site between 1999 and 2006 was 22 mm and the maximum 1-hour intensity was 54 mm. So the applied irrigation rate was fairly high but not unrealistic. Before the experiment, plots 1 and 2 were covered for approximately 2 weeks and are referred to as 'moist', plots 4 and 5 were covered for approximately 5 weeks and are called 'dry'. The initial matric potential before the plots were covered was −157 hPa at 0.2 m, −53 hPa at 0.3 m and −14 hPa at 1.0 m depth. Plot 3 was not covered and represented the actual field moisture conditions of the study site. Here, the matric potential before tracer application was −52 hPa at 0.2 m, −46 hPa at 0.3 m and −25 hPa at 1.0 m depth. Plot 2 was additionally irrigated with 64 mm of water just before tracer application. Prior to irrigation, we removed the spruce cones as they covered a large portion of the soil surface, but left the litter untouched. Table 2.1.1 summarizes the experimental boundary conditions.

The day after the irrigation, six vertical 1 m × 1 m soil profiles were excavated at intervals of 20 cm in the central part of the plot. We lit them by halogen projectors to supplement the natural daylight in the forest and photographed them with a CCD camera in RAW format. In this lossless format the image is not processed by the camera software and must be transformed in JPEG or TIFF by appropriate graphics software. Thus finer control is gained over white balance, sharpness or colour space. A rectangular frame and a grey scale were placed around the profiles for later correction of distortion and white balance

adjustment. Soil samples were taken for texture analysis in the laboratory, and Figure 2.1.1 summarizes the results. Nine profiles (one, two or three per plot) were treated with the indicator solution of iron(III) nitrate and starch to visualize iodide. They were photographed the same way as Brilliant Blue patterns.

Table 2.1.1 Experimental conditions for dye tracer experiments.

| Plot | Initial moisture | Irrigation rate /mm hour$^{-1}$ |
|---|---|---|
| 1 | 'moist'[a] | 64 |
| 2 | 'moist' | 64[b] |
| 3 | 'natural'[c] | 64 |
| 4 | 'dry'[d] | 64 |
| 5 | 'dry' | 32 |

[a]covered for approximately 2 weeks
[b]pre-irrigated with 64 mm of water just before tracer application
[c]not covered
[d]covered for approximately 5 weeks

In some sections of plot 3, large blocks of stone prevented us from digging deep enough. So we were obliged to diminish the spacing between profiles to 10 cm to prepare them in sections without blocks. Nevertheless, only four profiles had the desired depth of about 1 m and were suitable for further analysis.

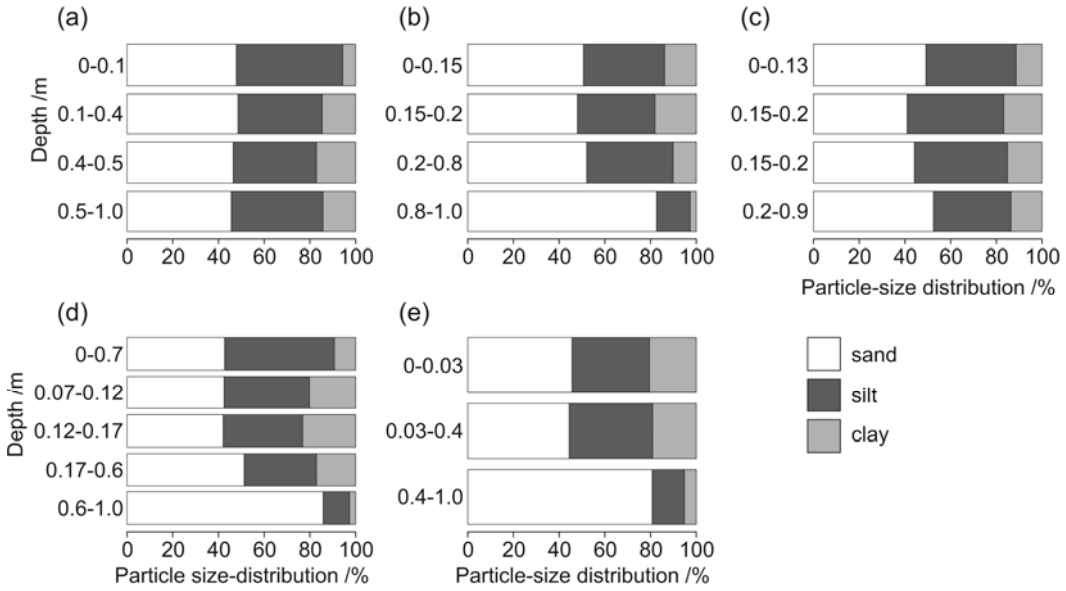



Figure 2.1.1 Particle-size distributions of the soil fine fraction on plots 1 (a) to 5 (e). Sand fraction is defined as 2000 - 63 µm, silt 63 - 2 µm and clay < 2 µm. The different depth sections correspond to soil horizons.

*Image processing*

The profiles were lit by halogen projectors, with the result that the colour temperature of the images differed from that of daylight. Therefore white balance was adjusted in Photoshop CS2 RAW-Converter (Adobe, 2005) via the grey scale. Then the photographs were corrected for perspective and radial distortion with the software PTGui (New House Internet Services B.V., 2005). Radial distortion is due to imperfections of the lens and was modelled by a fourth degree polynomial:

$$r_{\text{src}} = a\, r_{\text{dest}}^4 + b\, r_{\text{dest}}^3 + c\, r_{\text{dest}}^2 + (1-a-b-c)\, r_{\text{dest}}\,, \tag{2.1.1}$$

where $r_{\text{src}}$ is the radius between a pixel and the center of the original image (source, measured in pixels), and $r_{\text{dest}}$ is the radius in the corrected image (destination, measured in pixels). The radii $r_{\text{src}}$ and $r_{\text{dest}}$ are scaled such that the value 1 corresponds to:

$$\frac{1}{2}\ \max(\text{width}, \text{height})\,, \tag{2.1.2}$$

of the image. Parameters *a*, *b* and *c* are so-called lens parameters and can be adjusted in PTGui. Furthermore, the software PTLens (Niemann, 2005) offers a database of these values for many different types of cameras. Besides radial distortion, perspective distortion occurred because of rotation of the camera with respect to the photographed profile. We corrected it by setting control points all along the sides of the rectangular frame and adjusting them to horizontal and vertical lines. To decrease computing time, the image size was reduced such that 1 cm corresponded approximately to six pixels. This reduction did not affect further calculations as preliminary tests with different image resolutions had shown. In some images parts of the plot surfaces or shadows of the frame were visible. These regions would disturb further processing and were cut off. So the upper boundary of the output image corresponded to the first line in the photograph where the plot surface was no longer visible. Using Matlab 7.1 (The MathWorks, 2005b) and the Image Processing Toolbox (The MathWorks, 2005a), we extracted the blue patterns from Brilliant Blue stained images by a colour-based segmentation by *k*-means clustering (MacQueen, 1967) in the CIE 1976 *L*a*b** colour space. Segmentation of iodine–starch patterns by *k*-means clustering algorithm was not good enough because the colour of the iodine–starch complex

and that of the upper soil horizons were similar. Therefore we tried a classification based on hyper cuboids, an approach implemented in HALCON (MVTec Software GmbH, 2005). Finally, after segmentation, we generated binary images with stained parts in black and non-stained in white and calculated the dye coverage function $p(d)$ (the number of stained pixels per depth $d$).

### *Extreme value model*

Schlather & Huwe (2005) proposed a method for quantitative analysis of images from dye tracer experiments based on extreme value theory (see Coles, 2001, for an introduction). They applied the generalized Pareto distribution, a limit distribution of the extreme value theory with two parameters, to an idealized model of dye drops that run through soil along paths. The main idea is that the maximum depth $z$ of a dye stained path after $n$ drops converges to the so called *generalized extreme value distribution* (GEV), if $n \to \infty$ and the following assumptions as stated in Schlather & Huwe (2005) are satisfied:

'(i) any drop stains the path continuously up to the travel distance;

(ii) $z$ is in the maximum domain of attraction of the generalised extreme value distributions (Resnick, 1987);

(iii) the travel distances of the drops are independent and identically distributed.'

Statement (ii) means that the maxima of $z$ become GEV distributed. Excesses below greater depths converge to the *generalized Pareto distribution H*:

$$H(d, \xi_r, s) = 1 - \left\{1 + \frac{\xi_r (d - D)}{s}\right\}^{-\frac{1}{\xi_r}}, \qquad (2.1.3)$$

where $D$ is the threshold depth beyond which the data are assumed to follow closely the Pareto distribution, $d$ is the profile depth ($d$ and $D$ are measured in pixels on a photograph, $d > D$), $\xi_r$ is the form parameter ($\xi_r \in \mathbb{R}$) and $s$ is the scale parameter ($s > 0$), such that $(1 + \xi_r (d - D) / s) > 0$. Schather & Huwe (2005) argued that the dye coverage function $p(d)$ is an estimate of the probability that a path is stained at least down to this depth, modulo a constant factor $m$. The distribution $1 - H$ is fitted to the normalized dye coverage function $p(d)/m$ and describes the conditional probability that a path is still stained to a depth $d$, given that it is stained to the depth $D$ (for $d > D$). The form parameter $\xi_r$ is defined as a

risk index for vulnerability of groundwater to pollutants. Although the theoretical model describes drops travelling along distinct paths, Schather & Huwe (2005) stated that it could be applied both to preferential and matrix flow. In the case of matrix flow, paths are replaced by micropaths and drops by infinitesimal volumes of dye (the terms 'micro' and 'infinitesimal' are used in the sense of Marshall *et al.*, 1996). So the model always describes the predominant flow regime.

In this study we slightly modified the interpretation of the risk index. We think that the form parameter of the generalized Pareto distribution should be interpreted as a risk index for vulnerability of groundwater to pollutants only in regions with fairly homogeneous geological material between the soil surface and the water table, as in sedimentary basins with shallow water tables. The groundwater at our site is 8 to 10 m below the surface, and so we prefer to qualify $\xi_r$ as a risk index for vertical solute propagation.

The parameter $\xi_r$ determines the form of the generalized Pareto distribution. If it is negative, the distribution has an upper end point, i.e. dye infiltration stops before attaining a certain depth and the dye coverage function reaches zero. In this case, there is a low risk of solute's propagating in greater depths. If $\xi_r$ is positive then the distribution has no finite upper point, it decreases slowly and does not reach zero. Therefore the risk of solute propagation is high. Values of $\xi_r$ around 0 describe a transition zone. The scale parameter $s$ 'stretches' the distribution and can easily be interpreted for negative form parameters. Given a fixed negative $\xi_r$, $s$ depends monotonically on the maximum infiltration depth, i.e. the deeper the maximum infiltration depth the larger the value of $s$. So for the same value of $\xi_r$, the risk of solute propagation increases with larger values of $s$ (Figure 2.1.2). For positive form factors, $s$ is difficult to measure in the field. But for a fixed positive $\xi_r$, the portion of the stained area in a certain depth is greater for larger values of $s$. Indeed, as $s$ 'stretches' the distribution, larger values of $1 - H$, i.e. larger portions of stained pixels, can be found deeper in the soil.

Schlather & Huwe (2005) affirmed that the scale parameter $s$ depends strongly on experimental conditions such as the amount of sprinkled tracer solution or the time between irrigation and excavation of profiles. The behaviour of the form parameter $\xi_r$ under changing initial or experimental conditions is not clear, even though it seems to show some persistence against small variations.

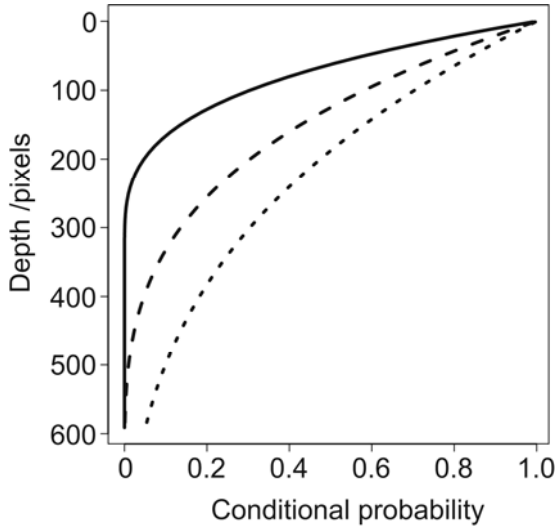



Figure 2.1.2 Effect of increasing values of scale parameter $s$ on the probability distribution $1 - H$. Form parameter $\xi_r$ is fixed to −0.3, scale parameter $s$ equals 100 (solid line), 200 (dashed line) and 300 (dotted line), $D$ equals 0.

For a reliable estimation of the risk index of a soil, Schlather & Huwe (2005) proposed taking at least 15 pictures. We used 28 pictures from five different experiments. Our goal was not to characterize the site but rather to understand the behaviour of the risk index under various initial and boundary conditions.

### *Parameter estimation*

As stated above, the generalized Pareto distribution describes excesses below greater depths, so we may not consider processes near the soil surface. We first estimated the parameter $\xi_r$ and then $s$. The estimation of $\xi_r$ is complex and recalls the ideas of Schlather & Huwe (2005) that are summarized in the following (see also the extension package SoPhy ver 1.0.25 (Schlather, 2005) of R (R Development Core Team, 2007)). Assume that we know the threshold $D$ beyond which the data follow (approximately) the Pareto distribution. Then the parameter $\xi_r$ and the parameter $s$ can simultaneously be estimated by a non-linear parameter optimization, e.g. non-linear least squares or maximum likelihood. Since we do not know $D$, $\xi_r$ has to be estimated for a range of values of $D$. We might expect to get the same value of $\xi_r$ for any $D$, except for some (small) error. This, however is not true, and ideally $\xi_r(D)$ behaves as in the sketch in Figure 2.1.3.

The horizontal line designates the true value of $\xi_r$ and the circles indicate the estimated values of $\xi_r$ depending on the threshold depth $D$. Three areas of $D$ can be distinguished, marked by the two vertical lines. The middle part gives the correct estimation of $\xi_r$. To the left, $D$ is not large enough, so that we are outside the assumed asymptotics, i.e. the assumption that the data can be approximated by a Pareto distribution below such a threshold $D$ is wrong. To the right, the number

of data available below the (large) threshold of *D* is small, so that larger variations in the estimation are visible.

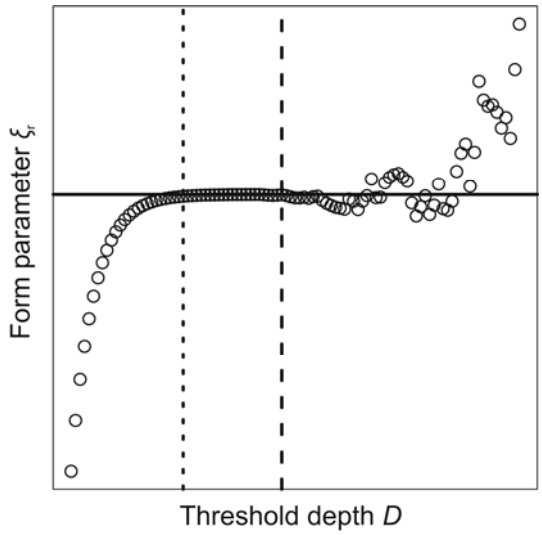



Figure 2.1.3 Schematic evaluation of the form parameter $\xi_r$ with changing threshold depth D. The horizontal line designates the true value of $\xi_r$ and the two vertical lines mark three areas of D: in the middle $\xi_r$ is correctly estimated; to the left, D is not large enough and the data cannot be approximated by a Pareto distribution; to the right, the number of data is small, so that larger variations in the estimation are visible.

Schlather & Huwe (2005) aimed (i) to find the middle part, (ii) to estimate $\xi_r$ from the middle part and (iii) to do it automatically. To achieve (iii) they suggested to take as middle part the values of *D*, where the maximum number of stained pixels of $p(d)$ ($d > D$) lies between 50% and 80%. For robustness, the median of the corresponding values of $\xi_r(D)$ is taken to get a final estimate for $\xi_r$. In contrast to $\xi_r$, the scale parameter will depend on *D* even under idealized conditions. Hence, we cannot get a final estimate for *s* in a similar way. Instead, we chose as the value of *D* the depth where $p(D)$ equalled 80%, and we estimated *s* in a next step whilst keeping $\xi_r$ and *D* fixed. The maximum likelihood estimator is frequently used for fitting parameters, since it is possible to calculate confidence intervals because of its approximate normality (Coles, 2001). However, the maximum likelihood estimator did not behave well for our data, and so we preferred the least-squares estimator until a better estimator that provides confidence intervals is found.

For stratified soil as at our study site, to find the middle part to estimate $\xi_r$ was more complex. The dye coverage functions were multimodal with a more or less pronounced second maximum in the lower soil (see Figure 2.1.4d, the first maximum is at the soil surface, the second one in about 130 pixels depth). So, following the proposition in Schlather & Huwe (2005), we applied the Pareto distribution only to the lowest part of the soil profile. Undocumented comparisons between different fitting procedures showed best results when we used the part of

the dye coverage function where $p(D)$ lies between 0% and 80% of the number of stained pixels at its second maximum. We took the median of these values to calculate the final $\xi_r$. We estimated the scale parameter $s$ in R (R Development Core Team, 2007) by unweighted non-linear least-squares regression using the form parameter $\xi_r$ determined in SoPhy (Schlather, 2005) and taking the depth where $p(D)$ equals 80% of the number of stained pixels at the second maximum, as $D$. In forthcoming versions of SoPhy, the final estimation of the scale parameter $s$ will be implemented. As a measure of goodness of fit, we calculated the coefficient of determination $R^2$ defined as

$$R^2 = 1 - \frac{\sum (p(d) - \hat{p}(d))^2}{\sum (p(d) - \bar{p})^2}, \qquad (2.1.4)$$

where $p(d)$ is the number of stained pixels in the depth $d$ ($d > D$), $\hat{p}(d)$ is the estimated number of stained pixels in the depth $d$ and $\bar{p}$ is the mean number of stained pixels in the part of the profile used for fitting the $1 - H$ distribution. The coefficient of determination can be negative if the enumerator is larger than the denominator i.e. if the adjusted function fits the data worse than a strait line through the mean of the data (Kvalseth, 1985).

### 2.1.3 Results

*Qualitative analysis of flow patterns*

In the following section we adopt the nomenclature proposed by Weiler & Flühler (2004) to describe flow processes based on the appearance of flow patterns. Figure 2.1.4 shows examples of binary images and their corresponding dye coverage functions. The soil profiles we excavated had a litter layer up to 10 cm thick and its first few centimetres were homogenously stained on all plots. The infiltration front broke into preferential paths in the lower part of the litter layer. Thus infiltration into the loamy upper soil (see Figure 2.1.1 for soil texture data) was inhomogeneous, and water flow bypassed large portions of the soil matrix in the upper 10 to 20 cm of the profiles. Accordingly, the dye coverage function decreased rapidly. In the upper soil, we found blue stained roots, indicating that there had been macropore flow in root channels.

The maximum of the dye coverage function was represented by large stained spots found between 20 and 40 cm. Texture analysis did not indicate any abrupt

changes, but the root system was less dense. So one possible explanation is that macropore flow in root channels decreased and the flow regime changed to predominantly heterogeneous matrix flow. Further studies should investigate if the root system is really responsible for this transformation of flow regime.

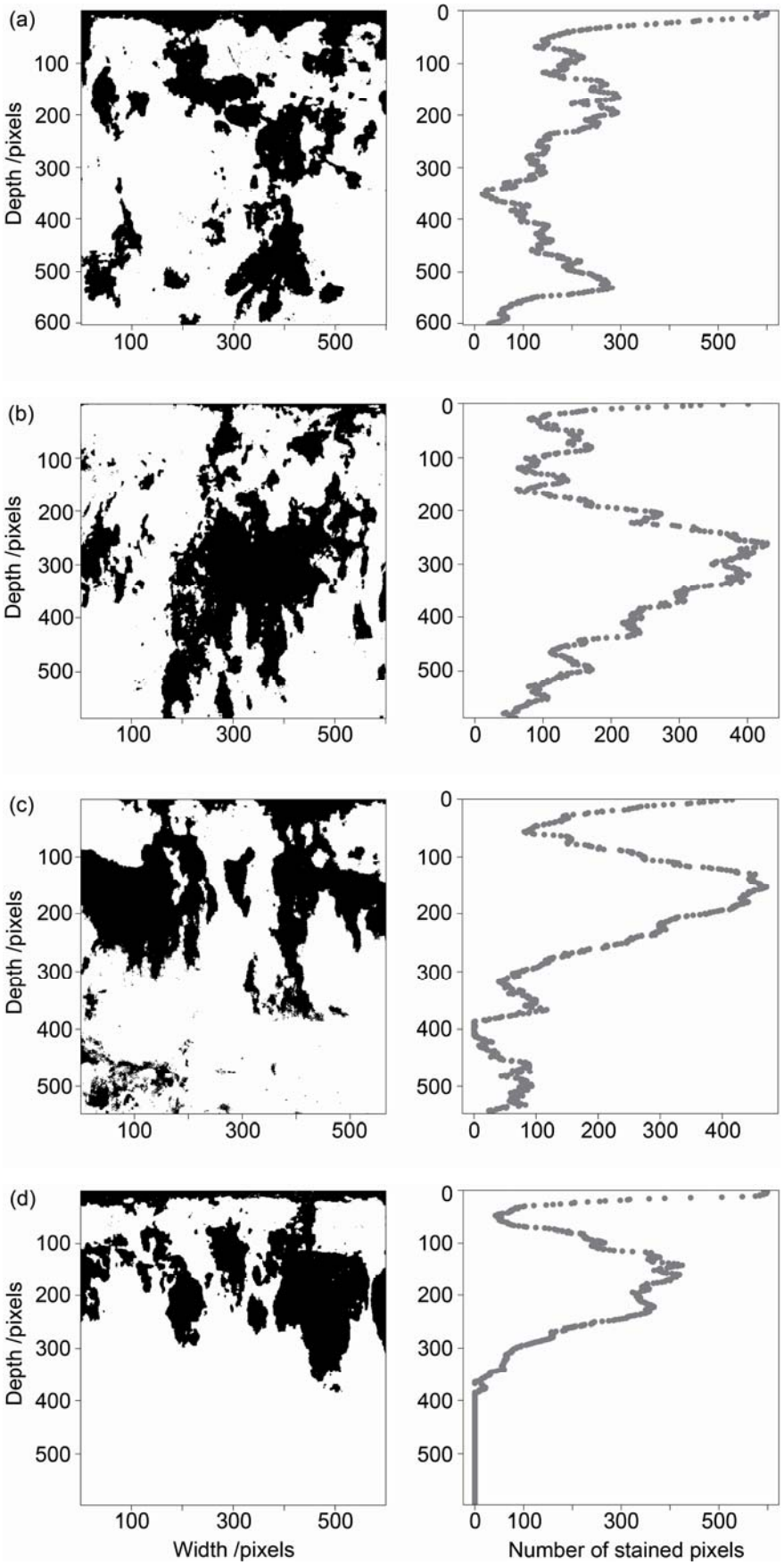
In the lower soil, heterogeneous matrix flow and fingering dominated, but water flowed along macropores containing both dead and living roots when these were encountered. This was especially the case on plots 1 and 2. The exchange of water and solute between macropores and soil matrix was greater on 'moist' plots (1, 2 and 3) than on 'dry' plots (4 and 5). The effect of pre-irrigation on plot 2 supported this observation, since the stained spots on this plot were larger than on plot 1.

Plots 1, 2 and 3 were stained down to the bottom of the profile, i.e. 1 m, whereas on plots 4 and 5 dye infiltration stopped at between 70 and 80 cm. On plot 3 less dye infiltrated in greater depths than on plots 2 and 3 as indicated by a smaller portion of blue stained surface. The surfaces of the stones in plot 3 served as preferential flow routes and were stained.

Figure 2.1.5 shows an example of Brilliant Blue and iodine−starch patterns on plot 4. In the upper 10 cm of the soil, there was no significant difference between the two tracers, neither in the location of the tracers inside the profile nor in the covered surface (see Figure 2.1.5, 0 to 40 pixels depth). But lower in the soil the iodine−starch spots were larger and the infiltration depth of iodide was greater than that of Brilliant Blue.

Two critical aspects remain when we compare the infiltration depths of Brilliant Blue and iodide. First, the redistribution time was different for Brilliant Blue and iodine−starch profiles, as the last were allowed to react over night. Lu & Wu (2003) stated in their work that 1 to 2 hours of reaction are sufficient for the development of the iodine−starch complex. At our site the first colour reaction was visible after approximately 2 hours. So there was indeed fixation of iodide after only a few hours of reaction. Once fixed, iodide becomes much less mobile as the molecules of the iodine−starch complex are large. But the contrast to Brilliant Blue, especially in areas stained by both tracers, was too low, and therefore, the iodine−starch complex was allowed to develop over night. Thus, even if there was a difference in redistribution times, it was less than 12 hours.

Second, the minimum concentration still visible on a profile might be different for the two tracers, and so the actual infiltration depth could be greater.



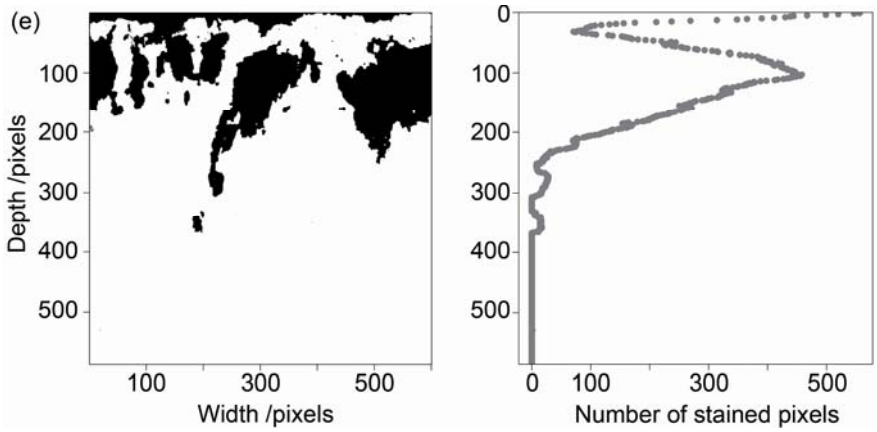



Figure 2.1.4 Example images from dye tracer experiments on plots 1 (a) to 5 (e). Six pixels correspond to 1 cm. Left column: binary images from Brilliant Blue stained profiles with blue parts in black and non-stained regions in white. Right column: the corresponding dye coverage functions.

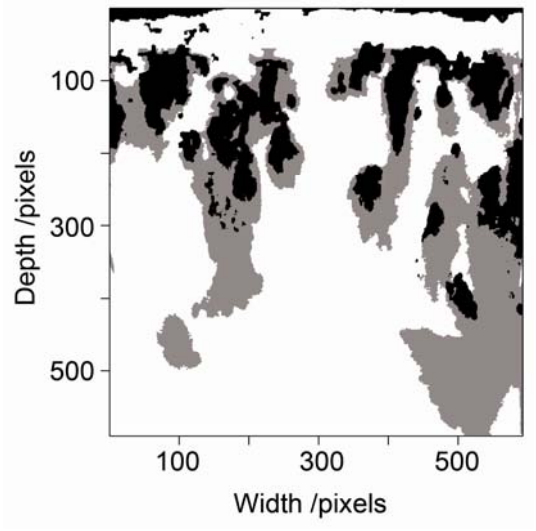



Figure 2.1.5 Example of Brilliant Blue (black) versus iodine−starch (grey) patterns on plot four. Six pixels correspond to 1 cm.

## *Risk indices*

In order to balance small fluctuations in the dye coverage function, we superposed all Brilliant Blue stained profiles of the same plot. Figure 2.1.6 shows these superposed profiles and the fitted distribution $1 - H$. Table 2.1.2 presents the risk indices $\xi_r$ and the scale parameters $s$. To demonstrate the variations of the risk indices between profiles of one plot, minimum and maximum of estimates of $\xi_r$ on single profiles are shown in the columns 'Minimum of $\xi_r$' and 'Maximum of $\xi_r$'. We used only single profiles where $1 - H$ was successfully fitted (visual check) and where $R^2$ exceeded 0.5 to calculate the minima and maxima.

Except on plot 3, the calculated risk indices are negative. It means that the dye infiltration will stop before reaching a certain depth. Thus there is a low risk of solute moving to greater depth below the analysed part of the profile. The risk

index on plot 3 equals 0. Here, the dye coverage function decreases exponentially and does not reach 0, but the amount of dye carried to greater depth below the analysed profile might be negligible (Schlather & Huwe, 2005).

Table 2.1.2 Calculated risk indices for superposed profiles.

| Plot | $\xi_r$ | s | $R^2$ | Minimum of $\xi_r$ [a] | Maximum of $\xi_r$ [a] |
|---|---|---|---|---|---|
| 1 | −0.9 | 377 | −0.49[b] | 0.4 | 1.3 |
| 2 | −1.1 | 334 | 0.62 | −1.0 | 0 |
| 3 | 0 | 94 | 0.99 | 0 | 0.1 |
| 4 | −0.3 | 118 | 0.97 | −0.9 | 0.0 |
| 5 | −0.2 | 61 | 0.98 | −0.1 | 0.5 |

[a]Minimum and maximum of $\xi_r$ for single (not superposed) profiles show the variation of the risk index within the plot. Results are presented for profiles where $1 - H$ was successfully fitted (visual check) and where $R^2$ exceeded 0.5.
[b]A negative $R^2$ indicates that the adjusted curve fits the data worse than a straight line through the mean value of the data.

On plots 1 and 2, several Brilliant Blue profiles had a multimodal dye coverage function with two similar maxima in the lower soil (as in Figure 2.1.4a, depths 200 and 530 pixels respectively), so even superposing them did not result in a monotonically decreasing function. Especially on plot 1, the third maximum appeared in the last third of the profile and led to a poor fit. One possible reason for this is the automated procedure to determine the starting point of the fit. Another reason is that the generalized Pareto distribution does not reflect flow processes in soils completely as the model theory is based on idealized assumptions. The distribution $1 - H$ is a monotonically decreasing function, and the quality of the fit depends strongly on the monotonicity of the dye coverage function. Therefore, the model is not suitable for dye coverage functions with pronounced multimodal behaviour as on plot 1.

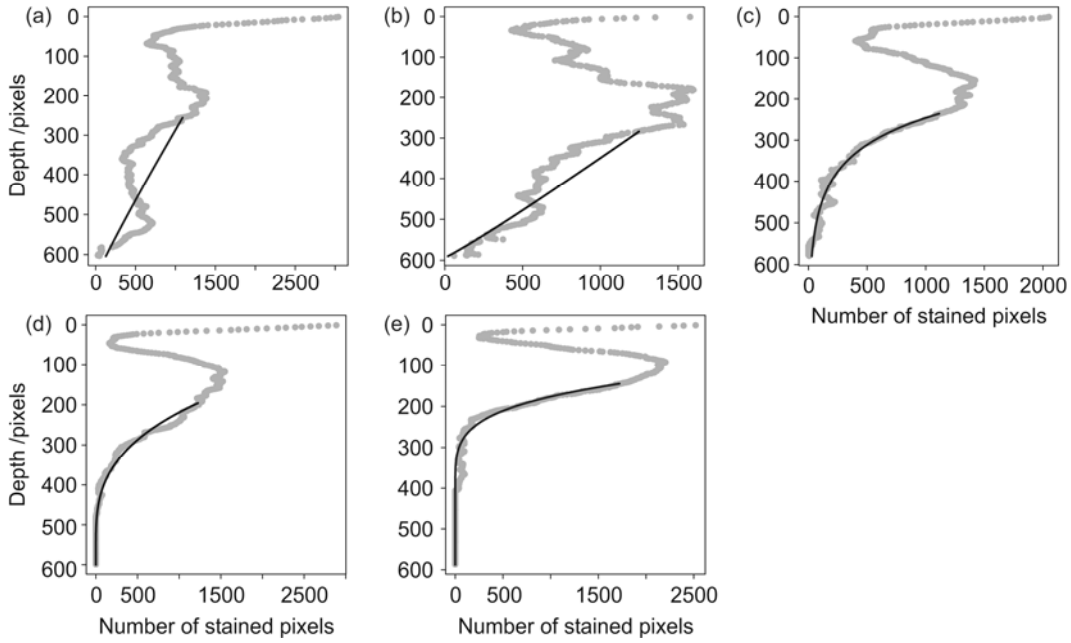



Figure 2.1.6 Dye coverage function of superposed profiles (dots) and fitted distribution $1 - H$ (line) on plot 1 (a) to plot 5 (e). Six pixels correspond to 1 cm.

We did not superpose iodine−starch stained profiles because there were too few. Risk indices based on these patterns were compared to those of Brilliant Blue of the same profile. The maximum difference in length between Brilliant Blue and iodine−starch profiles on the same plot was 26 pixels or 4 cm. Table 2.1.3 shows the results for profiles where $1 - H$ was successfully fitted (visual check) and where $R^2$ exceeded 0.5 for Brilliant Blue as well as for iodine-starch patterns.

Table 2.1.3 Risk indices for profiles with Brilliant Blue and iodide-starch patterns.

| | Brilliant Blue | | | Iodide-starch | | |
|---|---|---|---|---|---|---|
| Plot[a] | $\xi_r$ | s | $R^2$ | $\xi_r$ | s | $R^2$ |
| 2 | -0.9 | 46 | 0.85 | -1.0 | 245 | 0.71 |
| 3 | 0.1 | 74 | 0.89 | 3.5 | 5 | 0.54 |
| 4 | -0.5 | 159 | 0.87 | -1.4 | 659 | 0.80 |
| 5 | 0 | 58 | 0.96 | -0.9 | 261 | 0.96 |

[a]Adjustments on plot 1 did not give satisfying results.

Except on plot 3, where $R^2$ was small, risk indices $\xi_r$ for iodine−starch are less than those for Brilliant Blue patterns, indicating a lower risk for propagation of iodide. This is in contradiction with the greater infiltration depth of this solute and is discussed below.

### 2.1.4 Discussion

*Risk index for stratified soils*

As mentioned by Schlather & Huwe (2005), especially in stratified soils with pronounced differences in physical properties between horizons, the application of the generalized Pareto distribution to several strata is problematic. At our study site, the stratification seems to be due to changes in root distribution between 20 and 40 cm depth. Macropore flow that starts in the upper soil ends as matrix flow in lower horizons with a less dense root system. When the flow process changes as a result of varying physical properties the dye coverage function cannot be represented by one single distribution $1 - H$. The evident solution is to use only the lowest part of the profile to fit the distribution as we did it in our study. This accords with the limit law of the extreme value theory stating that the behaviour of the process at great depths is independent of the behaviour near the origin (Schlather & Huwe, 2005).

Furthermore, Schlather & Huwe (2005) stated in their paper that preferential flow is frequently linked to a positive risk index and matrix flow to a negative one. This is not longer true for stratified soils because only the lowest part of the profile is considered. At our study site, despite the occurrence of macropore flow in the upper soil, the calculated risk indices are negative, as the distribution $1 - H$ is fitted only to the lowest strata. And there the dominant flow regime is inhomogeneous matrix flow. So for correct assessment of risk of vertical solute propagation, the analysed profile depth should be taken into account.

*Combination of form and scale parameters*

Smaller risk indices for propagation of iodide are in contradiction with the greater infiltration depth of this solute. It is not surprising that the form parameter changes, as Brilliant Blue and iodide have different sorption characteristics. Especially in the lower soil where heterogeneous matrix flow and fingering dominated, Brilliant Blue was retarded with regard to iodide and their respective dye coverage functions differed in shape. We can resolve the contradiction by using both parameters, $\xi_r$ and $s$, to estimate the risk of vertical solute propagation. As mentioned before, the scale parameter $s$ 'stretches' the generalized Pareto distribution. So for the same risk index $\xi_r$ the probability to find stained pixels at a certain depth increases with larger values of $s$. The combination of $\xi_r$ and $s$

determines a complete probability distribution. Figure 2.1.7a shows the real difference between the risk of solute propagation based on Brilliant Blue (solid line) and iodine−starch patterns (dashed line) (see Figure 2.1.5 for patterns). The length of both profiles used for the adjustment of the generalised Pareto distribution differs by only 7 pixels (about 1 cm). The dashed line is situated right of the solid line indicating a higher risk for iodide propagation.

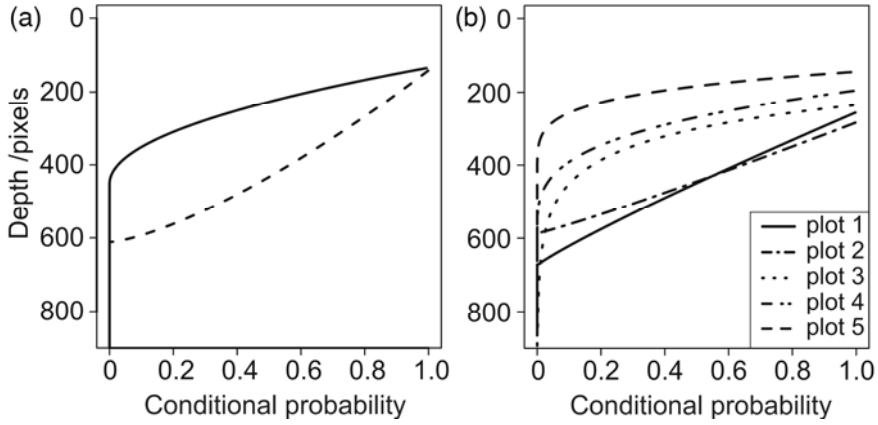



Figure 2.1.7 (a): Probability distribution for Brilliant Blue (solid line) and iodine−starch patterns (dashed line) on plot 4. (b): Probability distributions of the superposed Brilliant Blue stained profiles. The depth of the profiles is about 600 pixels, so in case of plots 1, 2 and 3 the dye reached the bottom of the profile.

The same procedure should be applied to assess the risk of Brilliant Blue propagation. Plots 4 and 5 are good examples. Here, the estimated risk indices are similar, but the scale parameters vary as predicted by Schlather & Huwe (2005) because of changing experimental and initial conditions. So the estimated actual risk of solute propagation is different on these plots. In Figure 2.1.7b we show the probability distributions $1 - H$ for the five superposed Brilliant Blue profiles (see Table 2.1.2 for parameters). For correct interpretation the fitted distributions are plotted for the part of the profile they were calculated for, i.e. beyond the threshold depth $D$. Despite the negative risk indices, it is clear that preferential flow is responsible for deep infiltration of the tracer. Based on our data, the estimated risk for solute propagation tends to increase from ‘dry’ to ‘moist’ plots.

### *Dependence of the risk index on boundary conditions*

In our experiment, changing the irrigation intensity from 64 mm hour$^{-1}$ to 32 mm hour$^{-1}$ seems to not affect the risk index significantly. Indeed, $\xi_r$ on plot 4 was −0.3 and on plot 5 −0.2 (Table 2.1.2). But according to the theory, the scale

parameter $s$ changes (halves) as experimental conditions are modified. The combination of the two parameters indicates a higher risk of solute propagation on plot 4, i.e. for the higher irrigation rate.

It is more difficult to see the effect of pre-irrigation on plot 2 as the fit is unsatisfactory. Moreover, it cannot be compared to plot 1 that has similar initial moisture conditions because the distribution $1 - H$ could not be fitted properly on this plot either. But as indicated by the dye coverage function, the stained surface was larger on the pre-irrigated plot 2 than on plot 1. This would support the hypothesis that the initial soil moisture is an important factor. Compared to 'dry' plots 4 and 5, the risk index on plots 1 and 2 is lower, but $s$ is much larger. The risk on plot 3 is the highest in accord with the highest initial moisture content. The tail of the distribution decreases exponentially ($\xi_r = 0$), and the dye penetrates deeper than on other plots. But as stated in Schlather & Huwe (2005), the transported mass might be negligible. Indeed, the dye-covered surface in the lower part of the profile is more important on plots 1 and 2 than on plot 3. So after combining $\xi_r$ and $s$, the complete distribution $1 - H$ supports a higher risk of vertical solute propagation for moist initial conditions.

Actually, initial soil moisture seems to be a crucial factor. Dye coverage functions on 'moist' plots fluctuated more than on 'dry' ones, so the quality of the fit was poorer, especially for single profiles. More rapid flow velocities and interactions between preferential flow paths and a moister matrix are one possible explanation. Hence, flow patterns are more complex, and the resulting dye coverage functions do not decrease monotonically.

Finally, the risk index seems to depend on the tracer, but we need more data (iodine−starch stained profiles) to verify this. Although experimental and initial conditions for Brilliant Blue and iodide are the same, their risk indices and scale parameters tend to differ in the same profile. This phenomenon is probably due to different physical properties of the two tracers (especially their sorption behaviour) and is not a characteristic of the risk index.

### 2.1.5 Conclusions

We varied experimental and initial conditions for our tracer experiments and used two different tracers, Brilliant Blue and iodide, to study the behaviour of the risk index $\xi_r$. Our results support the hypothesis formulated by Schlather and Huwe

(2005) that the risk index is to some degree invariant to changing experimental conditions (such as irrigation rate) and that the scale parameter *s* strongly depends on them. The initial soil moisture, however, seems to have a large influence on the risk index.

We propose to combine the two parameters of the generalized Pareto distribution to estimate the risk of vertical solute propagation in soils. The scale factor *s* reflects the maximum infiltration depth (for negative risk indices) or the amount of stained area at a certain depth (for positive risk indices). This information is important to assess correctly the risk and should be taken into account. Furthermore, a complete probability distribution $1 - H$ allows us to compare plots with different initial and experimental conditions or various tracers. For stratified soils, only the lowest part of the profile was used for the adjustment of the $1 - H$ distribution. Thus the depth should be explicitly included when one interprets the risk of vertical solute propagation.

To deal with strongly fluctuating or not decreasing dye coverage functions, the theory should be improved to account for tortuosity of flow paths. Provided that the dye coverage function decreases monotonically, the estimated risk for solute propagation can serve to classify soils. In stratified soils, different flow regimes can occur in different regions of the profile. To asses the risk for vertical solute propagation the generalized Pareto distribution should be fitted to the lowest part of the soil profile. But it could possibly be applied to single horizons as well to characterize the various flow regimes within the profile. Further studies should help to identify homogenous zones of flow patterns corresponding to different flow regimes or reflecting different physical soil properties. Weiler & Flühler (2004), for instance, proposed an interesting classification approach using stereology. Applied to such homogenous zones, fitting results of the distribution $1 - H$ could be improved.

### 2.1.6 Acknowledgements

We thank Andreas Kolb and Iris Schmiedinger for their assistance during field and laboratory work, Tobias Zuber for providing data of soil matric potential and Professor Foken for the precipitation data. We also thank the reviewers for their scrutiny of our script and suggestions for improving it; and MVTec for providing

a research-license for the software HALCON. This project was financially supported by the Deutsche Forschungsgemeinschaft (DFG FOR 562).

### 2.1.7 References

Adobe 2005. Photoshop version CS2. http://www.adobe.com. Accessed on January 2007.

Aeby, P., Forrer, J., Steinmeier, C. & Flühler, H. 1997. Image analysis for determination of dye tracer concentrations in sand columns. Soil Science Society of America Journal, 61, 33-35.

Bowman, R.S. 1984. Evaluation of some new tracers for soil-water studies. Soil Science Society of America Journal, 48, 987-993.

Coles, S. 2001. An Introduction to Statistical Modeling of Extreme Values. Springer-Verlag, London.

Farenhorst, A., Topp, E., Bowman, B.T. & Tomlin, A.D. 2000. Earthworm burrowing and feeding activity and the potential for atrazine transport by preferential flow. Soil Biology and Biochemistry, 32, 479-488.

Flury, M. & Flühler, H. 1994. Brilliant Blue FCF as a dye tracer for solute transport studies − a toxicological overview. Journal of Environmental Quality, 23, 1108-1112.

Flury, M. & Wai, N.N. 2003. Dyes as tracers for vadose zone hydrology. Reviews of Geophysics, 41, 2-1-2-37.

Flury, M., Flühler, H., Jury, W.A. & Leuenberger, J. 1994. Susceptibility of soils to preferential flow of water: a field study. Water Resources Research, 30, 1945-1954.

Forrer, I.E., Kasteel, R., Flury, M. & Flühler, H. 1999. Longitudinal and lateral dispersion in an unsaturated field soil. Water Resources Research, 35, 3049-3060.

Forrer, I.E., Papritz, A., Kasteel, R., Flühler, H. & Luca, D. 2000. Quantifying dye tracers in soil profiles by image processing. European Journal of Soil Science, 51, 313-322.

German-Heins, J. & Flury, M. 2000. Sorption of Brilliant Blue FCF in soils as affected by pH and ionic strength. Geoderma, 97, 87-101.

Ghodrati, M., Ernst, F.F. & Jury, W.A. 1990. Automated spray system for application of solutes to small field plots. Soil Science Society of America Journal, 54, 287-290.

Gish, T.J., Kung, K.-J.S., Perry, D.C., Posner, J., Bubenzer, G., Helling, C.S., Kladivko, E.J. & Steenhuis, T.S. 2004. Impact of preferential flow at varying irrigation rates by quantifying mass fluxes. Journal of Environmental Quality, 33, 1033-1040.

Hendrickx, J.M.H., Dekker, L.W. & Boersma, O.H. 1993. Unstable wetting fronts in water-repellent field soils. Journal of Environmental Quality, 22, 109-118.

Kasteel, R., Vogel, H.-J. & Roth, K. 2002. Effect of non-linear adsorption on the transport behaviour of Brilliant Blue in a field soil. European Journal of Soil Science, 53, 231-240.

Ketelsen, H. & Meyer-Windel, S. 1999. Adsorption of Brilliant Blue FCF by soils. Geoderma, 90, 131-145.

Kulli, B., Gysi, M. & Flühler, H. 2003. Visualizing soil compaction based on flow pattern analysis. Soil and Tillage Research, 70, 29-40.

Kvalseth, T.O. 1985. Cautionary note about R2. American Statistician, 39, 279-285.

Lawes, J.B., Gilbert, J.H. & Warrington, R. 1882. On the amount and composition of the rain and drainage water collected at Rothamsted. Williams, Clowes & Sons, London.

Lu, J. & Wu, L. 2003. Visualizing bromide and iodide water tracer in soil profiles by spray methods. Journal of Environmental Quality, 32, 363-367.

MacQueen, J.B. 1967. Some methods for classification and analysis of multivariate observations. In: Proceedings of the 5th Berkeley Symposium on Mathematical Statistics and Probability, (eds. L.M. Le Cam & J. Neyman), Volume I, pp. 281-297. University of California Press, Berkeley, CA.

Marshall, T.J., Holmes, J.W. & Rose, C.W. 1996. Soil Physics. Cambridge University Press, Cambridge.

Mitchell, A.R., Ellsworth, T.R. & Meek, B.D. 1995. Effect of root systems on preferential flow in swelling soil. Communications in Soil Science and Plant Analysis, 26, 2655-2666.

Mon, J., Flury, M. & Harsh, J.B. 2006. Sorption of four triarylmethane dyes in a sandy soil determined by batch and column experiments. Geoderma, 133, 217-224.

Morris, C. & Mooney, S.J. 2004. A high-resolution system for the quantification of preferential flow in undisturbed soil using observations of tracers. Geoderma, 118, 133-143.

MVTec Software GmbH 2005. Halcon version 7.1. http://www.mvtec.com/halcon/. Accessed on January 2007.

New House Internet Services B.V. 2005. PTGui version 5.5. http://www.ptgui.com/. Accessed on January 2007.

Niemann, T. 2005. PTLens version 6.4. http://epaperpress.com/ptlens/. Accessed on January 2007.

Öhrstöm, P., Persson, M., Albergel, J., Zante, P., Nasri, S., Berndtsson, R. & Olsson, J. 2002. Field-scale variation of preferential flow as indicated from dye coverage. Journal of Hydrology, 257, 164-173.

R Development Core Team 2007. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna. http://www.R-project.org. Accessed on January 2007.

Resnick, S.I. 1987. Extreme Values, Regular Variation, and Point Processes. Springer, New York.

Ritsema, C.J. & Dekker, L.W. 2000. Preferential flow in water repellent sandy soils: principles and modeling implications. Journal of Hydrology, 231, 308-319.

Schlather, M. & Huwe, B. 2005. A risk index for characterising flow pattern in soils using dye tracer distributions. Journal of Contaminant Hydrology, 79, 25-44.

Schlather, M. 2005. SoPhy: some soil physics tools for R. http://www.r-project.org/, contributed extension package

Schumacher, W. 1864. Die Physik des Bodens in ihren theoretischen und praktischen Beziehungen zur Landwirtschaft. Wiegandt und Hempel, Berlin.

Shuster, W.D., Subler, S. & McCoy, E.L. 2002. The influence of earthworm community structure on the distribution and movement of solutes in a chisel-tilled soil. Applied Soil Ecology, 21, 159-167.

The MathWorks, Inc. 2005a. Image Processing Toolbox version. 5.1. http://www.mathworks.com/products/image/. Accessed on January 2007.

The MathWorks, Inc. 2005b. Matlab version. 7.1. http://www.mathworks.com

Vogel, H.-J., Cousin, I., Ippisch, O. & Bastian, P. 2006. The dominant role of structure for solute transport in soil: experimental evidence and modelling of structure and transport in a field experiment. Hydrology and Earth System Sciences, 10, 495-506.

Wang, Z., Wu, Q.J., Wu, L., Ritsema, C.J., Dekker, L.W. & Feyen, J. 2000. Effects of soil water repellency on infiltration rate and flow instability. Journal of Hydrology, 231, 265-276.

Weiler, M. & Flühler, H. 2004. Inferring flow types from dye patterns in macroporous soils. Geoderma, 120, 137-153.

Weiler, M. & Naef, F. 2003. An experimental tracer study of the role of macropores in infiltration in grassland soils. Hydrological Processes, 17, 477-493.

## 2.2 Visualization and analysis of flow patterns and water flow simulations in disturbed and undisturbed tropical soils

C. BOGNER[a], S. ENGELHARDT[a], J. ZEILINGER[a,b] AND B. HUWE[a]

[a]*Soil Physics Group, University of Bayreuth, 95440 Bayreuth and* [b]*Institute of Plant Physiology, University of Bayreuth, 95440 Bayreuth, Germany*

Correspondence: C. Bogner. E-mail: christina.bogner@uni-bayreuth.de

Published in *Gradients in a Tropical Mountain Ecosystem of Ecuador* (eds. Beck, E., Bendix, J., Kottke, I., Makeschin, F. & Mosandl, R.), pp. 387-396. Springer Verlag, Berlin Heidelberg. 2008.

doi: 10.1007/978-3-540-73526-7

The original publication is available at www.springerlink.com

## Abstract

We studied water flow in disturbed and undisturbed tropical soils with Brilliant Blue dye tracer experiments. Combining the advantages of computer based image analysis with extreme value statistics we estimated a risk index for vertical solute propagation in these soils. The complexity of stained patterns and infiltration depths were different on disturbed and undisturbed study sites. And our data indicate that soil structure, stone content, plant roots system and possibly land use are controlling factors for water flow in soils. Additionally, we conducted a simulation study using Hydrus 2D to analyse the influence of soil texture heterogeneity and stone content on water flow. We show that stones are one of the reasons for heterogeneities in soil moisture distribution, as they create preferential flow paths and increase water flow velocities.



## 2.2.1 Introduction

The complexity of flow patterns and occurrence of preferential flow in soils depend upon spatial heterogeneity of the upper boundary condition, heterogeneous distribution of soil hydraulic parameters and soil structure (Flury *et al.*, 1994; Kulli *et al.*, 2003). In a tropical rainforest, the canopy transforms the spatially almost uniform rainfall to heterogeneous patterns of throughfall on soil surface. These patterns continue in the soil amplified by its own heterogeneity. More details on canopy interactions and water relations are given by Wilcke *et al.* (2008).

One of the most powerful methods to study water flow and solute transport in soils is to perform dye tracer experiments (Ghodrati & Jury, 1990; Flury *et al.*, 1994; Flury & Wai, 2003). For this purpose the dye Brilliant Blue is frequently used for its low toxicity and good visibility against the background colour of most soils (Flury & Flühler, 1995; German-Heins & Flury, 2000). Usually, pictures of stained patterns serve for a qualitative illustration of preferential flow. Recent works, however, propose a more quantitative approach based on modern image processing techniques (Schwartz *et al.*, 1999; Forrer *et al.*, 2000; Weiler & Flühler, 2004).

Most tracer experiments documented in literature were carried out on agricultural soil. To the authors' knowledge, there are only a few studies in stony forest soils (Schulin *et al.*, 1987; Buchter *et al.*, 1997) or in tropical soils

(Reichenberger *et al.*, 2002; Renck & Lehmann, 2004). The aim of the present research was to study water flow in disturbed and undisturbed tropical soils, with special emphasis on stony forest soils. We focused on undisturbed primary forest and two types of typical disturbances in the study area: natural landslides and pastures as a form of human land use. Thus we investigated a new and an old landslide, two sites in the primary forest and one on pastures. Combining the advantages of computer-based image analysis with extreme value statistics we estimated a risk index for vertical solute propagation in soil, as proposed by Schlather & Huwe (2005). Furthermore, we chose one of the primary forest sites for more detailed studies. We performed two-dimensional small-scale simulations using the model Hydrus-2D (Simunek *et al.*, 1999) and analyzed the influence of soil texture heterogeneity and stone content on water flow.



### 2.2.2 Methods

*Definition of the Risk Index*

Schlather & Huwe (2005) propose a model based on extreme value statistics. It describes the dye coverage function $p(d)$, the number of stained pixels ($p$) in depth $d$, as being an estimate of the probability to find stained pixels in at least this depth, up to a multiplicative constant $m$. The authors fit the tail distribution $1-H$ to the function $p(d)/m$, $H$ being the generalized Pareto distribution:

$$H(d, \xi_r, s) = 1 - \left(1 + \frac{\xi_r d}{s}\right)^{-1/\xi_r}, \qquad (2.2.1)$$

where $s$ is the scale parameter and $s > 0$, $\xi_r$ is the form parameter and $\xi_r \in \mathbb{R}$ and $d$ is the depth in the profile (measured in the image in pixels) such that $(1 + \xi_r d/s) > 0$. They propose $\xi_r$, called the form parameter of the generalized Pareto distribution, as a risk index for the vulnerability of groundwater to pollutants. Given the experimental and pedological conditions, three classes of values for $\xi_r$ can be distinguished:

1. If $\xi_r < 0$, $H$ has an upper end-point and the dye tracer does not exceed a certain depth.

2. If $\xi_r > 0$, $H$ has an infinite upper end point and the water table is surely reached.

3. If $\xi_r = 0$, $H$ decreases exponentially and the water table is reached, but the transported mass might be negligible.

We suppose that the parameters of the Pareto distribution vary with depth if the flow regime changes (i.e. from preferential to matrix flow). Therefore, $\xi_r$ should only be interpreted as a risk index for groundwater vulnerability for regions where the flow regime stays the same between the soil surface and the water table. This might be the case in sedimentary basins with rather homogeneous geological material and a shallow water table. In our mountainous and heterogeneous study area, the situation is different. Thus we propose to take the form parameter of the Pareto distribution as a risk index for vertical solute propagation in soil and a characteristic of the flow regime for given experimental conditions.

### *Dye Tracer Experiments and Image Processing*

We investigated five different sites: a new landslide (west of Q2), a cambisol on an old landslide (behind the ECSF building), a dystric skeletic cambisol (site we named "primary forest I" in Q2), a dystric leptosol ("primary forest II" on T2 at 2000 m) both in the primary forest, and a regosol on the pastures (100 m above road level, opposite the ECSF). See Chapter 1 for more details on the research area and Chapter 9 for further information on soils.

At each study site, a plot of approximately 2 $m^2$ was chosen. We eliminated litter and grass from the soil surface and applied 40 mm of a 10 g/l concentrated Brilliant Blue solution, using a spray system similar to the one described by Ghodrati *et al*. (1990). The irrigation intensity was 55 mm/h. The day after irrigation several vertical soil profiles were excavated and photographed within a metallic frame of 1 $m^2$. After geometrical correction the image size was reduced in a way that 1 cm corresponded approximately to six pixels and some parts like plot surface or a shadow of the frame were cut away. Using Matlab ver. 7.1 software (The MathWorks, 2005b) and Image Processing Toolbox ver. 5.1 (The MathWorks, 2005a), we extracted the blue patterns by a color-based segmentation and generated binary images with stained parts in black and non-stained in white. From these images, the dye coverage function $p(d)$ was calculated. Finally, the risk index for vertical solute propagation (for a definition, see Section 30.2.1) was

estimated by the software package SoPhy ver. 1.0.25 (Schlather, 2005) written in R (R Development Core Team, 2005).

### *Water Flow Simulations*

Next to the dye tracer experiment site primary forest I, a 4.0×1.5 m soil profile (Figure 2.2.1) was prepared. Soil texture was estimated based on a 10 cm grid. Additionally, bulk density, water retention curves and hydraulic conductivity were measured in the laboratory.

The matric potential-water content function and the hydraulic conductivity function were calculated with Rosetta (Schaap *et al.*, 2001), a computer program for estimating soil hydraulic parameters. Two-dimensional simulations were performed with Hydrus-2D (Simunek *et al.*, 1999), a model for water flow, heat and solute transport in two-dimensional variably saturated media that uses the finite element method to solve the Richards' equation.

Figure 2.2.1 Photograph of a 4.0×1.5 m soil profile. The *red squares* show two digitized sections used for water flow simulations. Photograph by J. Zeilinger

Simulation Run 1

In Hydrus-2D, we excluded the upper part of the profile in Fig. 30.1 (grass and litter) and generated a 4.0×1.0 m soil profile with texture distribution determined in the field. For this profile, stones were not taken into account and we focussed on soil texture heterogeneity and its influence on flow patterns. A constant infiltration rate of 2 cm/day (d) was applied on the upper boundary of the profile; the lower boundary condition was free drainage.

Simulation Run 2

Two 40×40 cm large sections of the profile were chosen (Figure 2.2.1, red squares) and the stones were digitized and simulated as internal zero flux

boundaries to investigate their influence on the flow regime. We applied the same constant infiltration rate of 2 cm/d on the upper boundary of each section.

### 2.2.3 Results and Discussion

*Dye Tracer Experiments*

We chose one characteristic profile from each dye tracer experiment to calculate the risk index $\xi_r$. Figure 2.2.2 shows the rectified original photographs and Figure 2.2.3 the binary images. Except for the new landslide, we had difficulties to fit one single distribution $1-H$ to the whole dye coverage function $p(d)$, i.e. from the top to the bottom of the soil profile, as different flow regimes occurred in different parts of the profile. So we limited the fit to the lowest part of the profile. The estimated risk indices and the depths from which on we fitted the distribution $1-H$ are given in Table1. In order to plot the estimated distribution $1-H$, we had to optimize the parameter $s$ of the Pareto distribution. This is not yet implemented in the software package SoPhy (Schlather, 2005) and was done directly in R (R Development Core Team, 2005). Figure 2.2.3 shows the function $p(d)$ and the fitted distribution $1-H$ for each profile.

On the *new landslide*, the stained patterns were rather simple. Only a small amount of dye infiltrated below 50 cm and just a few preferential pathways were visible. The estimated risk index is negative indicating a low risk of vertical solute propagation. The pattern complexity increased on the *old landslide*. Here, the dye infiltration was limited to few points and the first 20 cm of the topsoil were bypassed showing highly preferential flow. We found blue-stained plant roots indicating that water flowed along root channels. Deeper in the soil, larger blue stains occurred and the profile was stained down to a depth of 1 m. Accordingly, the risk index is positive which means that there is a high risk of vertical solute propagation.

This pronounced difference between the stained patterns on the new landslide and on the old one is probably due to soil regeneration. Indeed, on the new landslide soil structure was destroyed by mass movement, producing a more or less heterogeneous mixture of soil material and stones. Pedogenetic processes and plant activities recreated soil structure on the old landslide, thus increasing the occurrence of preferential flow especially along bio-macropores such as root

channels or earthworm burrows. More details on soil properties of landslides and plant succession can be found in Bussmann et al. (Bussmann *et al.*, 2008).


Figure 2.2.2 Vertical soil profiles (1×1 m) of the Brilliant Blue dye tracer experiments. **a** New landslide, **b** old landslide, **c** primary forest I, **d** primary forest II, **e** pastures. Photographs by C. Bogner

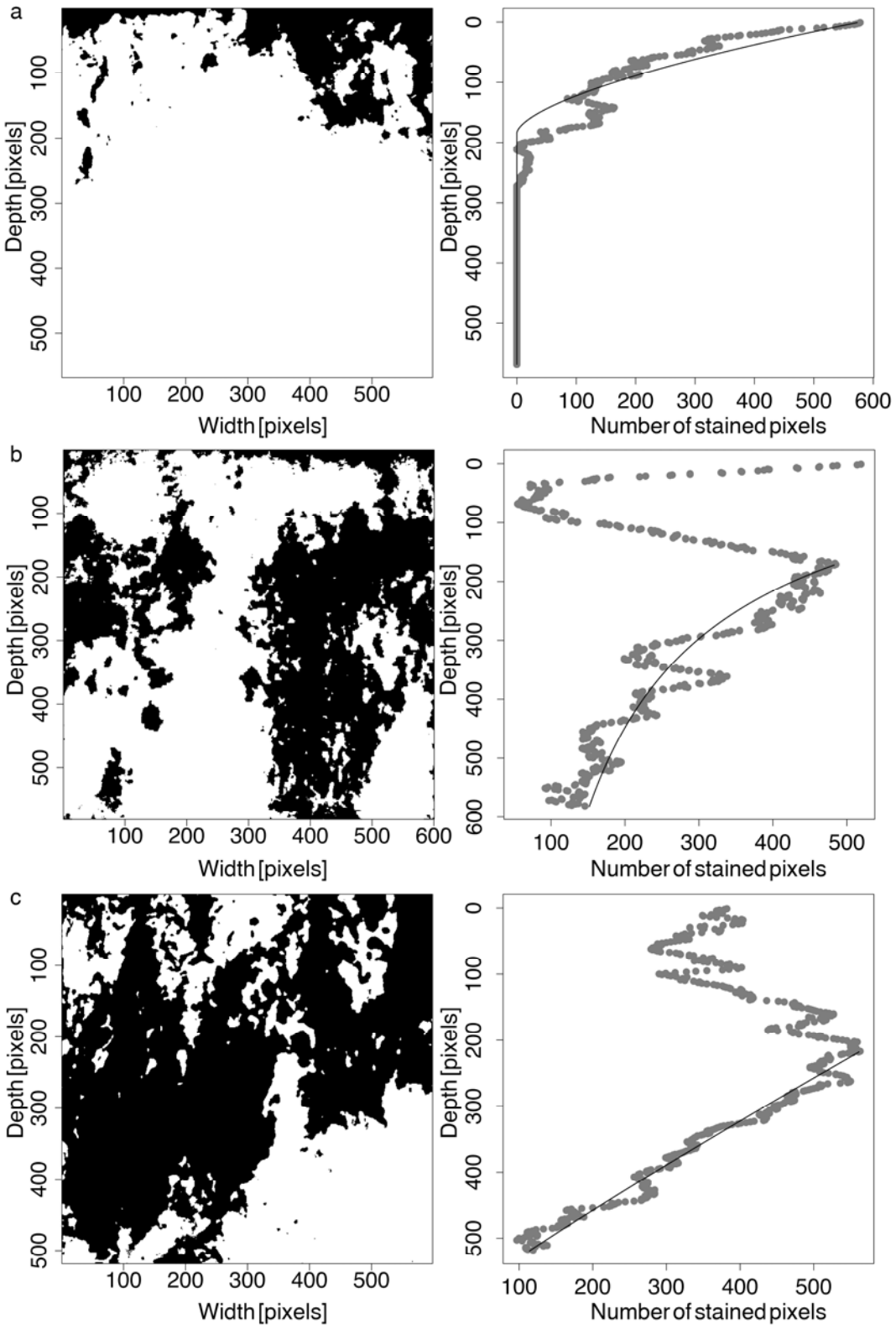



Figure 2.2.3 *Left column* Binary images (segmented images where blue-stained parts are colored black and non-stained parts are white) of the stained patterns from dye tracer experiments on the sites. **a** New landslide, **b** old landslide, **c** primary forest I, **d** primary forest II, **e** pastures. *Right column* The function $p(d)$, i.e. number of stained pixels with depth (*dotted lines*), and the fitted Pareto distribution (*solid lines*)

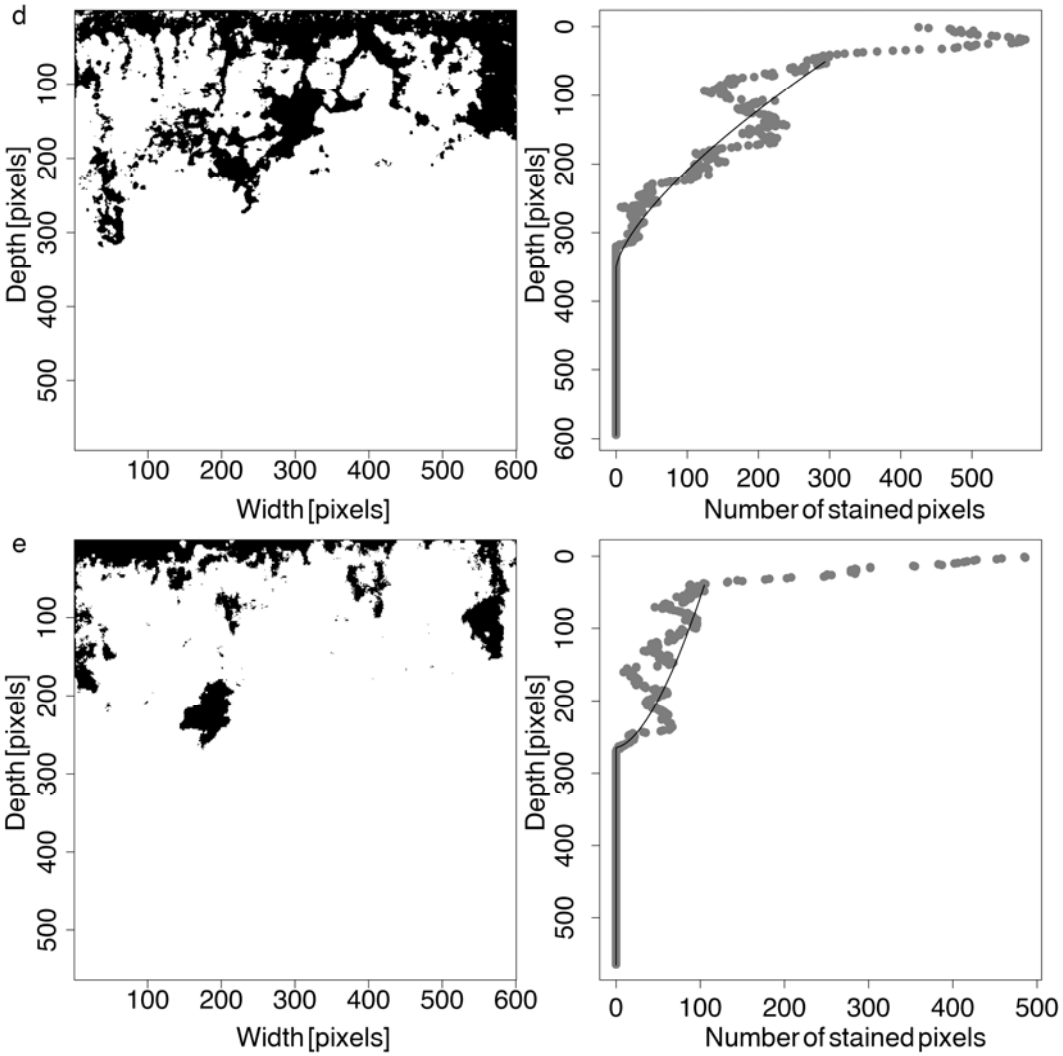



Figure 2.2.3 (continued)

Table 2.2.1 Estimated risk indices for the dye tracer experiments

| | New landslide | Old landslide | Primary forest I | Primary forest II | Pasture |
|---|---|---|---|---|---|
| Risk index ($\xi_r$) | −0.62 | 0.79 | −0.91 | −0.69 | −1.66 |
| Depth (cm)[a] | 4 | 32 | 50 | 9 | 7 |

[a]Approximate depths from which the distribution $1-H$ was fitted to the dye coverage function $p(d)$.

At the *primary forest I* site, we found similarly complicated patterns with localized infiltration as on the old landslide. Moreover, the soil at the primary forest I site contained a lot of stones which amplified the development of preferential flow, as will be discussed in Section 0. The interpretation of the risk

index for this site is difficult. It seems contradictory that the risk index is negative although the tracer reached the bottom of the profile (Figure 2.2.3 c). But one should keep in mind that the risk index only describes whether there is an end point for tracer infiltration or not and tells nothing about the infiltration depth. The function $p(d)$ decreases rapidly in the lower part of the profile and goes towards zero. So, there was an end point for tracer infiltration, but it exceeded the visible depth of the profile. Therefore, the risk of vertical solute propagation is greater than for the new landslide or the primary forest II site despite a more negative risk index. But it is still lower than for the old landslide, as there the dye coverage function did not tend towards zero at the bottom of the profile. For profiles as complicated as the primary forest I site, the second parameter of the Pareto distribution $s$ plays an important role. Schlather & Huwe (2005) mention that, for a given risk index, the parameter $s$ depends monotonically on the maximum depth of the dye tracer front. So for complex patterns we suggest to take $s$ into account, in order to correctly estimate the risk index for vertical solute propagation. This is not yet possible within the software package SoPhy (Schlather, 2005) and should be implemented in upcoming versions.

At the *primary forest II* site and on *pastures*, we found a compact top soil of about 30 cm. The dye stained surface on the first mentioned was greater partially because of a higher stone content in the top soil. Indeed, stones constituted preferential flow surfaces and were colored. But we can not exclude that infiltrability on pastures was reduced by compaction of the soil surface due to changes in land use. This could explain the smaller amount of dye penetrated into the soil. Several studies reported a decrease in saturated conductivity or an increase in bulk density after the primary rainforest was slashed and burnt and used as pastures (Elsenbeer *et al.*, 1999; Martinez & Zinck, 2004; Ziegler *et al.*, 2004). This form of anthropogenic disturbance is typical for our study area and further work is required to understand its influence on soil hydrologic properties. The calculated risk index indicates a low propagation risk at the primary forest II site and on pastures.

### *Simulation Study*

The small-scale heterogeneities of soil texture led to a non-homogeneous distribution of soil moisture. We observed the development of relatively dry areas

in the lower left corner of the profile and areas of stagnation in the right part. This produces different environments for chemical reactions.

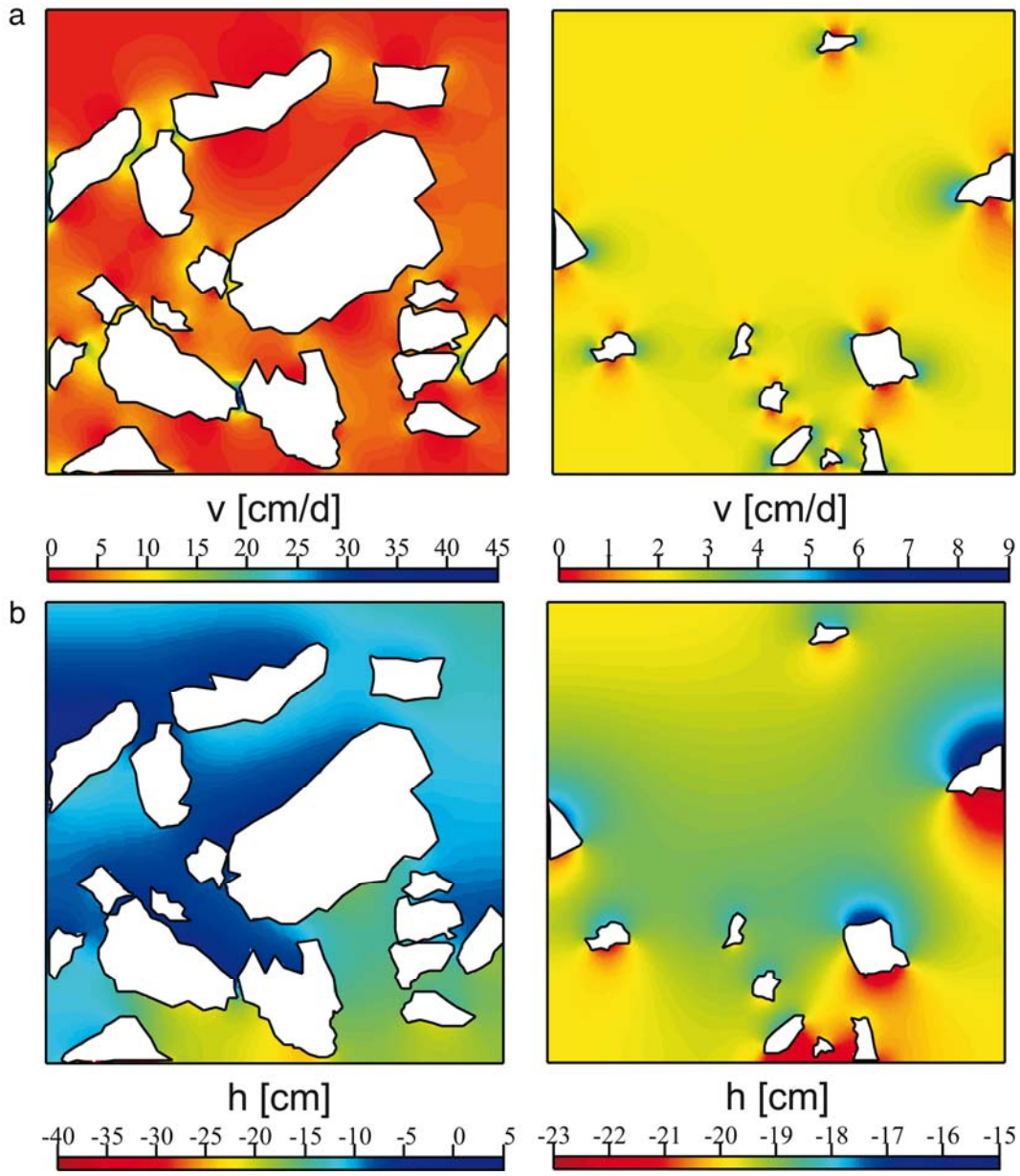



Figure 2.2.4 Results of water simulation run 2: distribution of flow velocities (**a**) and matric potential (**b**) in the two digitized sections (Fig. 30.1) at steady state. Boundary conditions: upper = constant infiltration rate of 2 cm/d; lower = free drainage; stones = internal boundaries of zero flux

Stones have a high influence on the flow regime. This is an important aspect as in our mountainous study area soils have a high stone content. Figure 2.2.4 shows flow velocities and pressure heads of simulation run 2 at steady state. We observed complex patterns in their distributions, with pronounced differences between the two sections. A high stone content leads to high velocities near the stones and especially in the gaps between them, creating preferential flow paths.

Stones cause a higher grade of differentiation in the soil with relatively wet areas on their tops and dryer areas below them. As a consequence, they modify water flow and transport of solutes like nutrients and pollutants and have therefore an influence on chemical processes in soils.

### 2.2.4 Conclusions

We studied water flow in disturbed and undisturbed tropical soils with dye tracer experiments. Soil structure, stone content, plant root systems and possibly land use are controlling factors for water flow in soils. Accordingly, the complexity of stained patterns and infiltration depths of the dye varied between the different study sites: rather simple patterns with a few preferential flow paths on the *new landslide* and on *pastures*, complex patterns at the *primary forest II* site and deep infiltration at the *primary forest I* site and on the *old landslide*.

The index proposed by Schlather & Huwe (2005) can serve as a useful characteristic of flow regime and as a risk index for vertical solute propagation in a variety of soils. The estimated indices for the *new landslide*, the *primary forest II* site and the *pastures* were negative. Despite the high irrigation intensity and the important amount of solution applied, the tracer did not exceed a certain depth and there is a low risk of vertical solute propagation to deeper soil regions or contamination of groundwater on these sites. In contrast, the index on the *old landslide* was positive, indicating a high propagation risk. Future studies should investigate the robustness of the risk index concerning the experimental conditions and analyze the role of litter for the development of preferential flow paths.

The detailed simulation study conducted on *primary forest I* site showed that stones were one of the reasons for heterogeneities in soil moisture distribution, creating preferential flow paths and increasing flow velocities.

Preferential water flow results in heterogeneous soil moisture distribution and has therefore several ecological implications. In general, preferential flow means a heterogeneous water supply and the coexistence of zones with high and low oxidation potential. Thus preferential flow influences multiple factors such as root growth, C- and N-mineralization, denitrification or humus accumulation and leads to rapid leaching of nutrients. In a non-homogeneously moist soil, the soil air phase is discontinuous and therefore the oxygen supply of plant roots and

microorganisms could be interrupted. Chemical and physico-chemical reactions, for example cation exchange, kinetic sorption processes or nutrient exchange between mobile and immobile water, precipitations and oxidations are also concerned by the occurrence of preferential flow.

### 2.2.5 Acknowledgements

We thank Martin Schlather for his help with the software package SoPhy and R, and Benjamin Wolf for having written the scripts in R and Matlab.

### 2.2.6 References

Buchter, B., Hinz, C. & Leuenberger, J. 1997. Tracer transport in a stony hillslope soil under forest. Journal of Hydrology, 192, 314-320.

Bussmann, R.W., Wilcke, W. & Richter, M. 2008. Landslides as Important Disturbance Regimes — Causes and Regeneration. In: Gradients in a Tropical Mountain Ecosystem of Ecuador, pp. 319-330.

Elsenbeer, H., Newton, B.E., Dunne, T. & Moraes, J.M.d. 1999. Soil hydraulic conductivities of latosols under pasture, forest and teak in Rondonia, Brazil. Hydrological Processes, 13, 1417-1422.

Flury, M. & Flühler, H. 1995. Tracer Characteristics of Brilliant Blue Fcf. Soil Science Society of America Journal, 59, 22-27.

Flury, M., Flühler, H., Jury, W.A. & Leuenberger, J. 1994. Susceptibility of soils to preferential flow of water: a field study. Water Resources Research, 30, 1945-1954.

Flury, M. & Wai, N.N. 2003. Dyes as tracers for vadose zone hydrology. Reviews of Geophysics, 41, 2-1-2-37.

Forrer, I.E., Papritz, A., Kasteel, R., Flühler, H. & Luca, D. 2000. Quantifying dye tracers in soil profiles by image processing. European Journal of Soil Science, 51, 313-322.

German-Heins, J. & Flury, M. 2000. Sorption of Brilliant Blue FCF in soils as affected by pH and ionic strength. Geoderma, 97, 87-101.

Ghodrati, M., Ernst, F.F. & Jury, W.A. 1990. Automated spray system for application of solutes to small field plots. Soil Science Society of America Journal, 54, 287-290.

Ghodrati, M. & Jury, W.A. 1990. A field-study using dyes to characterize preferential flow of water. Soil Science Society of America Journal, 54, 1558-1563.

Kulli, B., Gysi, M. & Flühler, H. 2003. Visualizing soil compaction based on flow pattern analysis. Soil & Tillage Research, 70, 29-40.

Martinez, L.J. & Zinck, J.A. 2004. Temporal variation of soil compaction and deterioration of soil quality in pasture areas of Colombian Amazonia. Soil & Tillage Research, 75, 3-17.

Reichenberger, S., Amelung, W., Laabs, V., Pinto, A., Totsche, K.U. & Zech, W. 2002. Pesticide displacement along preferential flow pathways in a Brazilian Oxisol. Geoderma, 110, 63-86.

Renck, A. & Lehmann, J. 2004. Rapid water flow and transport of inorganic and organic nitrogen in a highly aggregated tropical soil. Soil Science, 169, 330-341.

Schaap, M.G., Leij, F.J. & van Genuchten, M.T. 2001. ROSETTA: a computer program for estimating soil hydraulic parameters with hierarchical pedotransfer functions. Journal of Hydrology, 251, 163-176.

Schlather, M. 2005. SoPhy: some soil physics tools for R. http://www.r-project.org/, contributed extension package

Schlather, M. & Huwe, B. 2005. A risk index for characterising flow pattern in soils using dye tracer distributions. Journal of Contaminant Hydrology, 79, 25-44.

Schulin, R., Vangenuchten, M.T., Flühler, H. & Ferlin, P. 1987. An Experimental-Study of Solute Transport in a Stony Field Soil. Water Resources Research, 23, 1785-1794.

Schwartz, R.C., McInnes, K.J., Juo, A.S.R. & Cervantes, C.E. 1999. The vertical distribution of a dye tracer in a layered soil. Soil Science, 164, 561-573.

1999. The HYDRUS-2D software package for simulating two-dimensional movement of water, heat, and multiple solutes in variably saturated media.

R Development Core Team. 2005. R: A Language and Environment for Statistical Computing. http://www.R-project.org [accessed on February 2009]

The MathWorks, Inc. 2005a. Image Processing Toolbox ver. 5.1. http://www.mathworks.com/products/image/ [accessed on February 2009]

The MathWorks, Inc. 2005b. Matlab ver. 7.1. http://www.mathworks.com [accessed on February 2009]

Weiler, M. & Flühler, H. 2004. Inferring flow types from dye patterns in macroporous soils. Geoderma, 120, 137-153.

Wilcke, W., Yasin, S., Fleischbein, K., Goller, R., Boy, J., Knuth, J., Valarezo, C. & Zech, W. 2008. Water Relations. In: Gradients in a Tropical Mountain Ecosystem of Ecuador, pp. 193-201.

Ziegler, A.D., Giambelluca, T.W., Tran, L.T., Vana, T.T., Nullet, M.A., Fox, J., Vien, T.D., Pinthong, J., Maxwell, J.F. & Evett, S. 2004. Hydrological consequences of landscape fragmentation in mountainous northern Vietnam: evidence of accelerated overland flow generation. Journal of Hydrology, 287, 124-146.

## 2.3 Relating flow patterns and processes on soil surface in an agricultural soil as affected by tillage

BOGNER C[a], MIRZAEI M[b], RUY S[b], HUWE B[a]

[a]*Soil Physics Group, University of Bayreuth, 95440 Bayreuth, German and* [b]*UMR EMMAH, INRA, Avignon, France*

Corresponding author: C. Bogner. E-mail: christina.bogner@uni-bayreuth.de.

In preparation

## Summary

Tillage influences processes on soil surface and affects infiltration of water into subsoil. While many analyses of soil surface processes or flow patterns in soils exist, studies relating both are rare in literature. We did two tracer experiments with Brilliant Blue FCF on a tilled and compacted plot and a non-tilled one with a well developed system of open cracks to (i) investigate processes on soil surface during simulated rainfall; (ii) analyse the resulting flow patterns and (iii) relate both to tillage treatment. On the non-tilled plot soil surface roughness increased after tracer application probably due to surface consolidation. However, on the tilled plot it did not change because the applied simulated rainfall had not enough energy to induce significant erosion. Despite tillage preferential flow along cracks occurred on both plots and macropores buried below tillage pan functioned as preferential flow paths. Our results show that the studied soil is susceptible to deep vertical solute propagation at dry no till conditions when cracks are open, connected down to the subsoil and infiltration rates through macropores are important. Although tillage destroys macroporosity in the topsoil it does not prevent macropore flow. Keeping in mind the shallow water table at this site, the risk for potential groundwater contamination is high.

## 2.3.1 Introduction

Tillage is known to influence processes on soil surface, is important for generation of surface runoff and affects soil erodibility and depression storage. Rough soil surfaces tend to store more water thus reducing runoff and increasing infiltration. Different tillage treatments induce different surface roughness conditions. However, soil roughness usually decreases with increasing rainfall regardless tillage treatment and canopy cover (Eltz & Norton, 1997). Many techniques have been used to measure soil surface roughness, ranging from a simple roller chain to a laser scanner. Sterephotogrammetry is a relatively recent method that allows rapid data acquisition under field conditions. An overview of various techniques and their applicability is given in Jester & Klik (2005). To describe surface roughness statistical, geostatistical and fractal indices can be used (Paz-Ferreiro *et al.*, 2008). Huang & Bradford (1992) demonstrated that two

fractal indices – crossover length *l* and fractal dimension *D* – are necessary to describe soil surface roughness.

Tillage does not only influence processes on soil surface, but also flow processes in subsoil, especially because it is often associated with mechanical compaction by employment of heavy machinery. In order to visualise flow processes in soil the food dye Brilliant Blue FCF is frequently used (Pickering *et al.*, 1988; Flury *et al.*, 1994). It has the advantage of good visibility against most soil colours and acceptable toxicological properties for environmental use (Flury & Flühler, 1994). With this staining technique Kulli *et al*. (2003) studied the effect of soil compaction on soil porosity, bulk density and on the water infiltration regime under field conditions. Their results demonstrated that soil compaction modified the pore system by degrading or destroying the soil structure. As a consequence, soil mechanical parameters like bulk density were changed and the transport properties of the pore system were modified.

While many analyses of soil surface processes (Vidal Vázquez *et al.*, 2006; Moreno *et al.*, 2008; Vidal Vàzquez *et al.*, 2008) or flow patterns in soils exist (Forrer *et al.*, 2000; Weiler & Flühler, 2004), studies relating both are rare in literature. We thus propose to combine stereophotogrammetry and dye staining techniques to (i) investigate processes at soil surface during simulated rainfall; (ii) analyse the resulting flow patterns and (iii) relate both to tillage treatment. A special emphasise will be given to connectivity of processes above and below the plough pan and impacts on groundwater vulnerability.

## 2.3.2 Materials and methods

### *Dye tracer experiments*

The research area is located on an alluvial plane at about 30 m above sea level at the confluence of Rhone and Durance rivers whose deposits constitute an important aquifer with a shallow water table rising up to 3 m below the surface. The mean annual precipitation is 722.4 mm (Meteo France, 2009). The study site is an agricultural field planted with wheat situated at the experimental domain of INRA Avignon, France. The soil is a highly structured Calcisol (IUSS Working Group WRB, 2007) with a pH (in $H_2O$) of 8.5. Due to the important content of clay (>35%) it develops deep cracks during desiccation. Table 2.3.1 summarizes the distribution of the soil fine fraction.

Table 2.3.1: Soil characteristics.

| Plot | Depth /cm | Sand /% | Silt /% | Clay /% |
|---|---|---|---|---|
| | 10-26 | 6.37 | 47.70 | 45.94 |
| I | 26-56 | 5.00 | 47.59 | 47.41 |
| | 56-100 | 15.98 | 46.36 | 37.66 |
| | 0-30 | 5.18 | 48.35 | 46.47 |
| II | 30-60 | 4.94 | 48.14 | 46.93 |
| | 60-100 | 15.53 | 47.63 | 36.85 |

We did two rainfall simulation experiments with a tracer solution consisting of 5 g $l^{-1}$ Brilliant Blue FCF and 5 g $l^{-1}$ potassium iodide. It was applied at 64 mm $hour^{-1}$ on a surface of about 1 $m^2$ during one hour with an automated sprinkler similar to that described by Ghodrati *et al*. (1990). Potassium iodide was used as a reference tracer because Brilliant Blue is sorbed by soil particles and thus might be considerably retarded with regard to infiltrating water (Ketelsen & Meyer-Windel, 1999; German-Heins & Flury, 2000; Kasteel *et al.*, 2002). In order to keep the infiltrated volume of tracer solution comparable between plots, a metallic frame was installed around the plot to avoid large run-off. During tracer application processes on soil surface and soil moisture were monitored.

To observe processes on soil surface a stereo system consisting of two Nikon D100 cameras was installed in approximately 3 m height (Mirzaei, 2008). For the first experiment the wheat was harvested and the soil surface was tilled with a plough share followed by a circular spike harrow combined with a cultipacker. The middle part of the plot was compacted by repeated passage of a tractor wheel. The day before tracer application we did a preliminary study on this plot with 56 mm of water without tracer to test the stereo system. The surface of plot two was not tilled, but the wheat was cut before tracer application. We have chosen plot two for its well developed open cracks.

The day after tracer application vertical soil profiles of 1 m × 1 m (10 on plot one and 11 on plot two) were prepared and photographed with a digital CCD camera in RAW format. A frame of 1 $m^2$ and two Kodak reference gray scales for geometric and colour corrections were photographed with the profiles. To control whether Brilliant Blue was retarded with regard to infiltrating water two profiles

per plot were treated with an indicator solution of iron(III) nitrate and starch (Lu & Wu, 2003) to visualize iodide infiltration patterns. However, we did not obtain any colour reaction indicating the creation of the iodine-starch complex possibly due to large pH values in the soil that hindered the reduction of iodide.

## *Image analysis*

### Soil surface

Stereoimages of the soil surface were used to calculate a digital elevation model (DEM) by point matching between photographs obtained from the left camera and those from the right one (e.g. Jeschke, 1990; Warner, 1995). This method is based on the so called disparity, the difference between locations of the same image point in photographs of the left camera and of the right one. In a calibration procedure, camera parameters defining a model for mapping of 3D points of the real world to 2D images generated by the camera were derived. Then, disparities were transformed to real world 3D coordinates using the focal length of the calibrated camera system and the distance between the cameras (MVTec Software GmbH, 2007c). Calculations were performed using the software HALCON (MVTec Software GmbH, 2007a). The matching algorithm is described in details by Mirzaei (2008).

Thereafter, on plot one the area occupied by puddles larger than 22 $cm^2$ (thereafter called 'large puddles') and the volume of water stored there were calculated. Puddles were identified by texture differences between smooth water filled regions and rough soil surface. Considering the water surface as flat, the stored volume was calculated by summing up the differences between the bottom of puddles and the mean elevation of puddles' borderline in the initial DEM (Mirzaei, 2008). On plot two the stored volume was negligible and was not calculated.

### Soil profiles

We photographed the soil profiles in the lossless RAW-format (unprocessed data of the CCD camera sensor) and converted the images to JPEG after white balance adjustment using Photoshop CS2 (Adobe, 2005). Perspective and radial distortion was corrected in HALCON (MVTec Software GmbH, 2007a), i.e. the photographs were transformed such that they corresponded to images captured by

an ideal camera that produces no radial distortion and looks perpendicularly to the soil profile. Perspective transformation of the images was obtained via a homogenous transformation matrix (a matrix that combines information about both rotation and translation in one matrix) and radial distortion was modelled by:

$$\begin{pmatrix} u \\ v \end{pmatrix} = \frac{1}{1+\kappa\left(\tilde{u}^2+\tilde{v}^2\right)}\begin{pmatrix} \tilde{u} \\ \tilde{v} \end{pmatrix}, \tag{2.3.1}$$

where the parameter κ is the magnitude of the radial distortion, $(\tilde{u}, \tilde{v})^T$ are coordinates of a point in the original image and $(u, v)^T$ are coordinates in the corrected one. If κ is negative, the distortion is barrel-shaped, while for positive κ it is pincushion-shaped. (MVTec Software GmbH, 2007c).

Then the corrected images were transformed from RGB to HIS (hue saturation intensity) colour space and classified in Brilliant Blue stained and non-stained areas. The HIS colour space is more suitable for robust separation of colours independently of their intensities under varying illumination (MVTec Software GmbH, 2007b). All gray values in an interval [low threshold, upper threshold] were classified as stained pixels and misclassifications due to the shadow of the metallic frame or large stones were corrected manually. The classified photographs were converted to binary images (stained parts were coded black and non-stained areas white) and were used to calculate the dye coverage function (number of stained pixels per depth) using MATLAB (The MathWorks, 2005).

### *Soil surface roughness*

To describe possible changes in soil surface roughness due to the simulated rainfall, we calculated the fractal dimension *D* and the crossover length *l* before and after the tracer application by analysing the semivariogram γ(*c*) (e.g. Huang & Bradford, 1992). The empirical semivariance can be calculated from the DEM as:

$$\gamma(c) = \frac{1}{2n}\sum_{i=1}^{n}\left[Z\left(x_i+c\right)-Z\left(x_i\right)\right]^2, \tag{2.3.2}$$

where *n* is the number of point pairs, Z are elevations at positions $x_i$ separated by a horizontal distance *c*. We used the add-on package RandomFields (Schlather, 2008) in R (R Development Core Team, 2008) to calculate the empirical variogram. If the variation of elevations follows the fractional Brownian motion model (fBm) (Mandelbrot & Van Ness, 1968), then:

$$\gamma(c) \propto c^{2h}, \tag{2.3.3}$$

where the parameter $h$ is related to fractal dimension $D$ by:

$$D = 3 - h, \tag{2.3.4}$$

Plotting the log of semivariance versus the log of $c$ the parameter $h$ can be calculated by determining the slope of the straight part of the variogram. Huang & Bradford (Huang & Bradford, 1992) proposed to use the fractal dimension $D$ together with the crossover length $l$ to describe soil roughness:

$$l = \exp\left(\frac{a}{2-2h}\right), \tag{2.3.5}$$

where $a$ is the intercept of the straight part of the semivaraince on the log-log plot. As stated by Huang & Bradford (1992) and Eltz & Norton (1997) higher values of $D$ are associated with a higher soil surface roughness. For the same value of fractal dimension $D$ soil roughness increases with increasing crossover length $l$. While $D$ is a relative measure of spatial distribution of different structural elements on soil surface, $l$ is related to the vertical component of soil surface roughness (Huang, 1998; Vidal Vázquez *et al.*, 2006).

### *Flow patterns*

Schlather & Huwe (2005) proposed the form parameter of the generalized Pareto distribution as a risk index for vulnerability of groundwater to pollutants. They fitted the distribution $1 - H$ to the dye coverage function $p(d)$ (number of stained pixels per depth), $H$ being the generalized Pareto distribution:

$$H(d, \xi_r, s) = 1 - \left\{1 + \frac{\xi_r \left(d - d^*\right)}{s}\right\}^{-\frac{1}{\xi_r}}, \tag{2.3.6}$$

where $d^*$ is the threshold depth beyond which the data are assumed to follow closely the Pareto distribution, $d$ is the profile depth ($d$ and $d^*$ are measured in pixels on a photograph, $d > d^*$), $\xi_r$ is the form parameter ($\xi_r \in \mathbb{R}$) and $s$ is the scale parameter ($s > 0$), such that $(1 + \xi_r (d - d^*) / s) > 0$. The distribution $1 - H$ describes the conditional probability that a path is still stained to a depth $d$, given that it is stained to the depth $d^*$ (for $d > d^*$). Bogner *et al*. (2008) analyzed the behaviour of the risk index under changing experimental (different irrigation rates) and initial conditions (variable initial soil moisture). They proposed to

combine the form and the shape parameters of the generalized Pareto distribution and interpret the form parameter as a risk index for vertical solute propagation.

The form parameter determines the behaviour of the tail of the generalized Pareto distribution. A negative form parameter indicates that the distribution has an upper end point, i.e. in a certain depth the probability to find the tracer is zero. The distribution has no upper end point for positive form parameters, it decreases slowly and does not reach zero. A form parameter of zero describes an exponential decrease of the generalized Pareto distribution. The shape parameter *s* 'stretches' the distribution (Figure 2.3.1). For negative form parameters *s* depends monotonically on the maximum infiltration depth, and for positive form factors it reflects the amount of stained area in a certain depth.

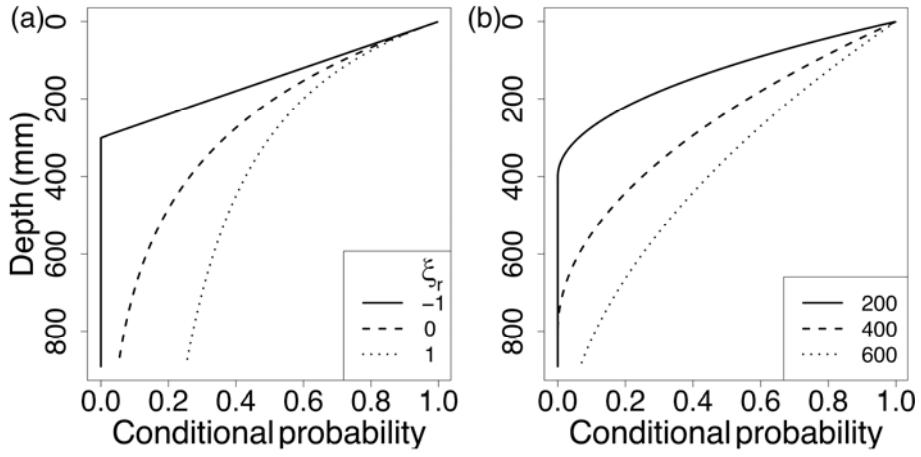



Figure 2.3.1: Distribution $1 - H$ for different parameters. (a) variation of the form parameter $\xi_r$ for $s = 300$ and $d^* = 0$; (b) variation of the scale parameter $s$ for $\xi_r = -0.5$ and $d^* = 0$.

We fitted the distribution $1 - H$ to the dye coverage function of superposed profiles. The latter were calculated by summing all dye coverage functions of the respective experimental plot and normalizing them by the width of the profiles. The distribution $1 - H$ indicates the probability to find the dye tracer in a certain depth and thus the susceptibility of the soil to vertical solute propagation.

## 2.3.3 Results and discussion

### *Initial conditions*

Initial gravimetric soil moisture varied between 19 and 41% on plot one before the preliminary application of 54 mm of water and between 16 and 19% on plot two before tracer application (Figure 2.3.2a). The multiple passage of the tractor

wheel on plot one caused an increase in soil bulk density in the top 10 cm (Figure 2.3.2b) compared to the tilled part. The soil bulk density profile reflects well the tillage of the surface. It varies little in the top 20 cm that correspond to the tilled part, increases rapidly in the tillage pan situated in about 25 cm depth and remains stable underneath. These results accords well with findings by Kulli et al. (2003) who reported a significant increase in bulk density in the top 15 cm of the soil and in 55 cm depth after single and multiple passages of a sugar beet harvester. However, between 15 and 55 cm depth no significant differences were found.

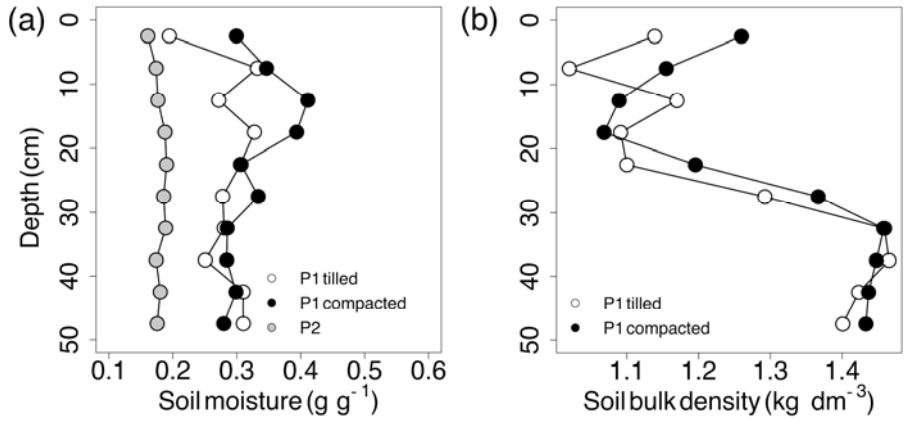



Figure 2.3.2: (a) Initial gravimetric soil moisture on plot one (P1) and plot two (P2); (b) soil bulk density of compacted and tilled parts of plot one (P1) measured by gamma ray attenuation.

## *Processes on soil surface*

### Digital elevation models

The resolution of the digital elevation models was 2 mm. The point matching process produced noisy data close to the metallic frame and in areas shaded by soil clods due to poor contrast. Furthermore, wheat residues on the surface of plot two resulted in particularly large elevations that were not due to the structure of the soil surface. In contrast, cracks produced missing values or particularly small elevations. These erroneous values were eliminated and the remaining values were corrected for linear slope prior to calculations of the empirical variogram. To identify erroneous values we used the add-on package robfilter (Fried & Schettlinger, 2008) in R (R Development Core Team, 2008) that extracts the relevant signal separating it from noise by robust filtering using moving windows. Points with the 5% largest and 5% smallest differences between the original elevations and the robust estimates were eliminated. Additionally, to discard high

elevation due to wheat residues and small values due to cracks on plot two, the 5% of the largest and the smallest values were deleted. Linear trend was removed by calculating a plane in R (Team, 2008) by least-squares and missing values were not interpolated. After the clean up 188168 and 173064 points were available on plot one in the DEM before and after the tracer application respectively; on plot two 179575 and 125716 points were left.

Soil surface roughness

The empirical variogram on both plots did fit the fBm model, but the slope on the log-log graph changed slightly at 10 and at 100 mm. Therefore, we only used distances between 10 and 100 mm to estimate the fractal dimension $D$ and the crossover length $l$. We calculated the fractal dimension and the crossover length for the tilled and the compacted parts on plot one separately. As expected $D$ and $l$ were smaller on the compacted part than on the tilled one: 2.43 and 0.20 mm versus 2.60 and 1.12 mm. Both parameters remained nearly unchanged after tracer application: $D$ = 2.45, $l$ = 0.17 mm on the compacted part and $D$ = 2.60, $l$ = 1.13 mm on the tilled one. This was not the case on plot two where $D$ increased slightly from 2.73 to 2.81 and $l$ from 1.28 mm to 2.64 mm.

Rainfall generally decreases soil roughness, but it also may increase it. Eltz & Norton (1997) mentioned in their study that fractal dimension increased after light rainfalls with low energy. They attributed this increase in soil roughness to consolidation of the surface and rearrangement of soil clods and aggregates after the soil had been moistened. Light rainfall did not significantly affect large clods, but the redistribution of fine particles created a denser surface with greater roughness. In our study, this consolidation might have happen on plot two where the dye tracer was applied on a dry surface after a long period without precipitations. Indeed, raindrops produced by the sprinkler are very small and have not enough energy to induce significant erosion. This effect of soil roughness increase was not observed on plot one probably because of the preliminary application of 56 mm of water before the tracer experiment.

Water storage

Due to poor quality of the DEM along frame walls, the stored volume and the area covered by large puddles were calculated on a reduced total surface of 0.98 $m^2$

and puddles along frame walls were not considered. Missing values in the DEM were filled by gray-level opening (MVTec Software GmbH, 2007b) to avoid under prediction of the stored volume (Figure 2.3.3).

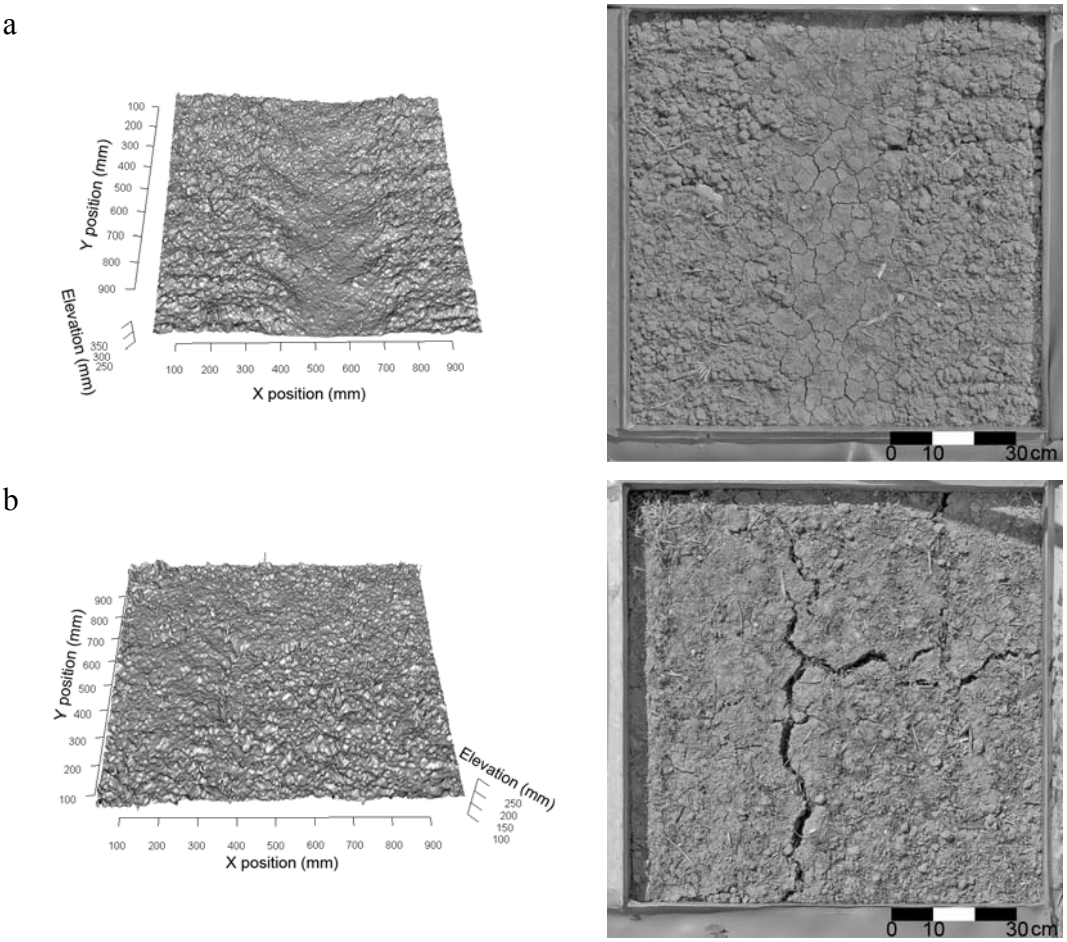



Figure 2.3.3: DEM and soil surface before dye tracer application on plot one with a compacted middle part (a); and on plot two with a well developed system of open cracks (b).

During simulated rainfall the large puddles occurred mainly in the compacted part of plot one. This indicates a decrease in infiltrability compared to the tilled part and accords well with results by Kulli et al. (2003). The volume stored increased from 0.7 mm after 6 min of tracer application to 6.4 mm after 46 min (Figure 2.3.4). The maximum area covered by large puddles was 26% after 46 min of rainfall. After rainfall stop the stored volume decreased rapidly. The large volume stored at the soil surface (especially on the compacted middle part) is mainly due to the metallic frame that surrounded the plot. Therefore, it does not correspond to a natural storage, but would have probably contributed to the surface runoff. On plot two we observed a visible closing of cracks on the surface during tracer application, but the storage of water was negligible.

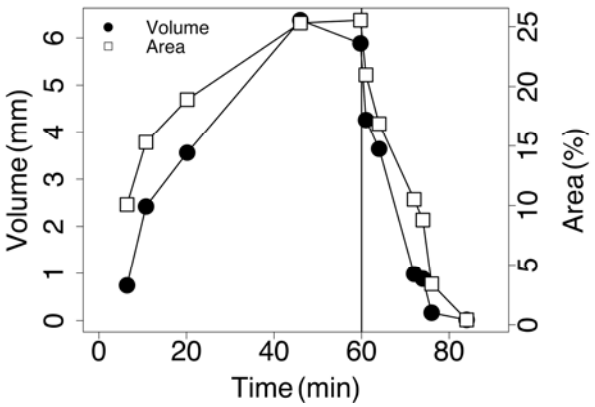



Figure 2.3.4: Volume stored on the soil surface during tracer application and area covered by puddles larger than 22 $cm^2$ on plot one.

## *Infiltration and flow patterns*

In the tilled zone of plot one, volumetric water content raised from 33 to 37% in 8 and 18 cm depth after about 10 min of rainfall (Figure 2.3.5a). After the end of irrigation rapid drainage was observed, especially in 8 cm depth. This fast increase in soil moisture as well as its quick decrease indicates the occurrence of preferential flow probably through inter-aggregate voids created by tillage. TDR probes situated in the compacted zone responded only slowly and the relative increase in water content was smaller. However, this shows that the compacted zone was not completely bypassed by the flow. The offset recorded at the beginning of irrigation is probably an artefact. Soil moisture continued to increase after the end of tracer application, probably due to the infiltration of water stored in large puddles.

On plot two, TDR probes situated in cracks and in the soil matrix in 8 cm depths responded at about the same time (Figure 2.3.5b). Deeper in the soil, however, no changes in soil water content were recorded in the soil matrix. In contrast, water content raised quickly after the onset of irrigation in cracks. After the end of irrigation, the TDR probe situated in 30 cm depth in a crack showed a very rapid decrease in soil moisture probably indicating a fast drainage of the macropore.

These TDR measurements accord very well with the heterogeneous dye patterns. On plot one two distinct zones could be identified: large stained areas above the plough pan and narrow stained objects below it. In the tilled part water flowed preferentially through inter-aggregate spaces created by tillage. The compacted zone is clearly visible and characterised by shallower infiltration, but it

was not completely bypassed by the flow (Figure 2.3.6a). Above the plough pan, the amount of stained area varied considerably more than below. The plough pan situated in about 25 cm depth constituted a discontinuity for the water flow and we observed lateral funnelling of the tracer solution. However, macropores below the plough pan conducted water flow. This is in accordance with Logsdon (1995) who found in column studies that buried macropores below the tilled layer were not harmed and could function as preferential flow pathways.

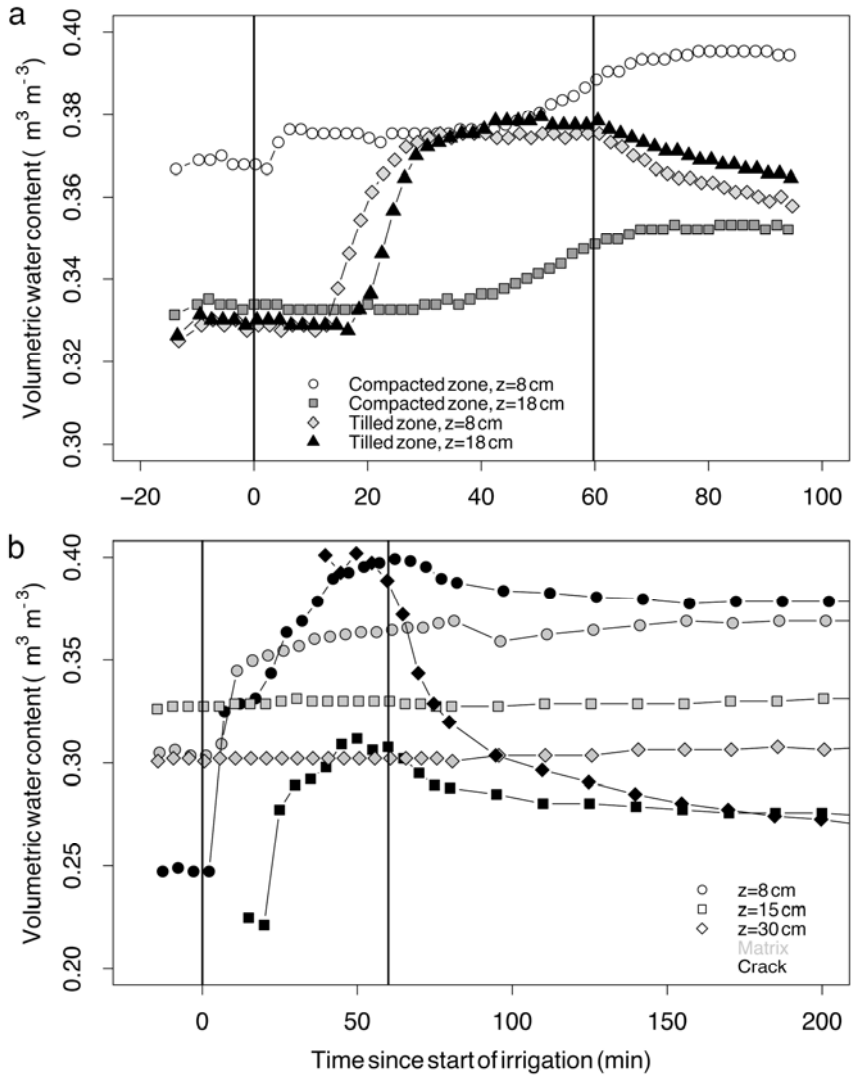



Figure 2.3.5: Volumetric water content in different depths during infiltration measured by TDR on plot one (a) and plot two (b).

In contrast, on plot two no visible plough pan was detected and cracks continued from soil surface down to the subsoil. The top 5 cm of the soil were homogenously stained indicating the dominance of the matrix flow. Below, the flow was funnelled to the cracks where the tracer was transported predominantly.

Indeed, due to low permeability of the soil matrix lateral spreading of the tracer from preferential pathways into the matrix was low. This resulted in narrow stained objects ranging from the top to the bottom of the soil profiles (Figure 2.3.6b). As described by Jarvis (2007) and the references given therein, for large macropores to conduct water flow water pressures must reach close to saturation (> –10 cm). It does not mean that the entire soil profile must wet up. Near-saturation conditions need only occur locally (e.g. in millimetre thick layers around cracks or above compacted soil zones at plough pan).

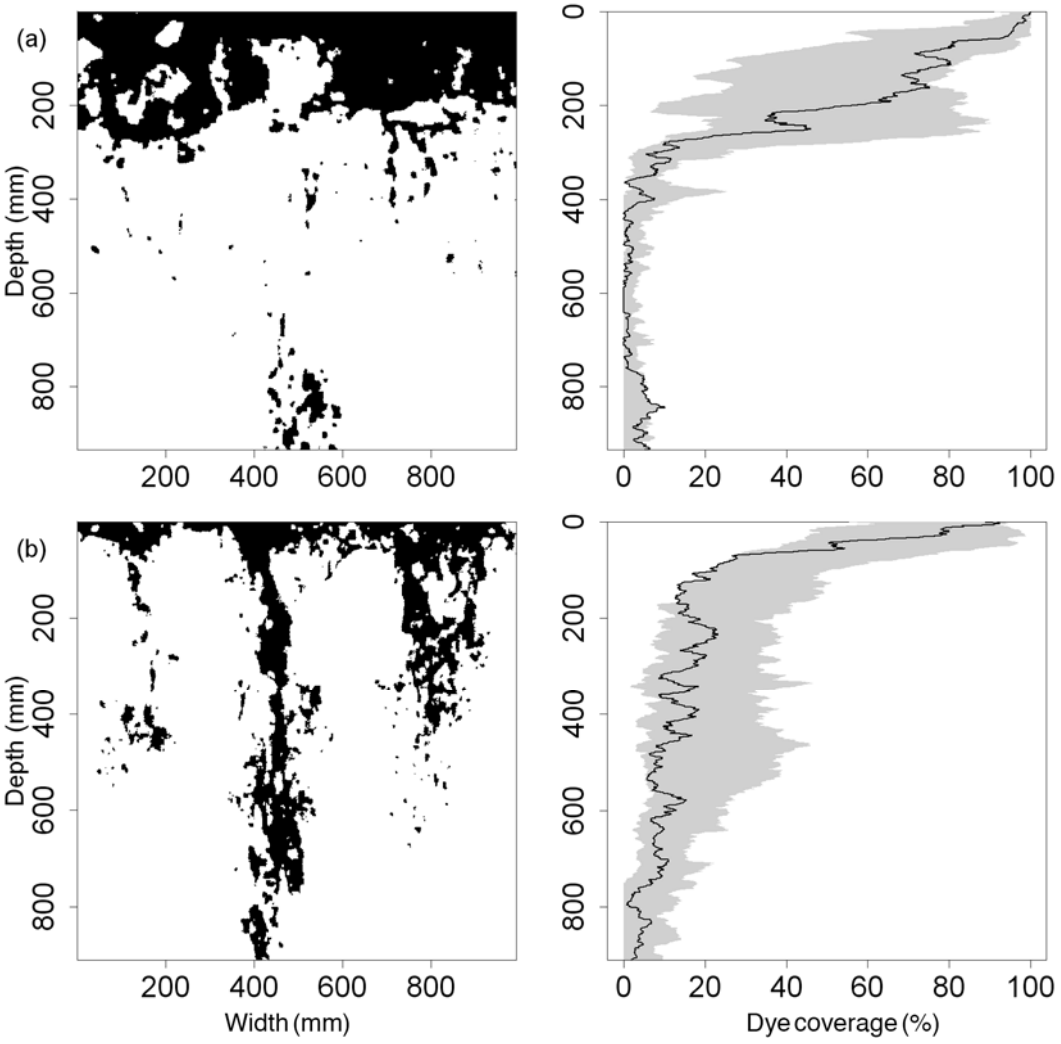



Figure 2.3.6: Example images of stained patterns on plot one with the compacted part between 300 and 700 mm (a) and on plot two (b).

On plot one, the amount of tracer infiltrated into the subsoil was smaller than on plot two. This was clearly reflected by the dye coverage function and the fitted distribution $1 - H$ (Figure 2.3.7). The plough pan is well visible in the dye coverage function of plot one and highlights the slope change in about 25 cm

depth. In contrast, the change of the slope on plot one is due to transition form matrix to macropore flow. The adjusted parameters of the distribution $1 - H$ are given in Table 2.3.2. Although on plot two the risk index is negative, combined with the scale parameter $s$ the risk of vertical solute propagation is higher. However, on both plots the probability to find the tracer at the bottom of the soil profiles (in 1 m depth) is not zero.

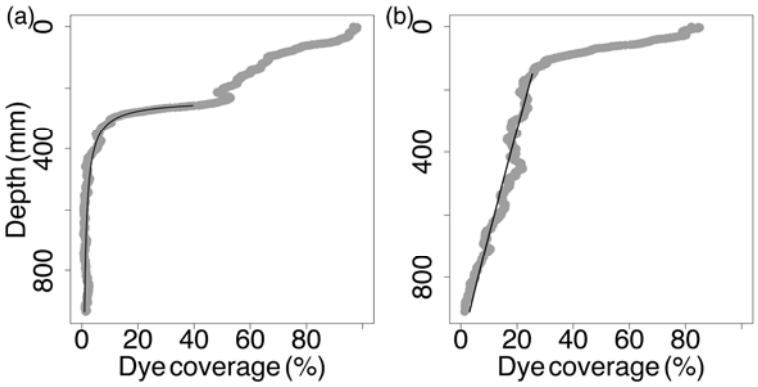



Figure 2.3.7: Dye coverage functions of superposed profiles (dots) and fitted distribution $1 - H$ (line) on plot one (a) and plot two (b).

Table 2.3.2 Calculated parameters of the distribution $1 - H$

| Plot | $\xi_r$ | $s$ | $R^2$ |
|---|---|---|---|
| 1 | 1.0 | 18 | 0.97 |
| 2 | −0.9 | 827 | 0.93 |

*Relationship between surface processes, flow patterns and tillage*

Agricultural cropping systems are highly dynamic and characterised by frequent disturbances of the soil surface (tillage). Our results showed that after tillage transport processes underneath the plough pan are partly disconnected from processes at soil surface and in the tilled horizon due to burying of macropores and funnelling of water along the plough pan. However, cracks under plough pan remained stable and still functioned as preferential flow paths. After draught periods and at no till conditions infiltration at our study site is dominated by macropore flow through open cracks, and transport systems (soil surface, topsoil and subsoil) are interrelated. Indeed, an existing old plough pan did not disturb the flow. Previous studies conducted at this site showed that during moist periods cracks were closed and macroporosity reduced. In this case matrix flow dominates

and Brilliant Blue does not penetrate below the plough pan (Bogner, unpublished data).

As a conclusion, our results show that the studied soil is susceptible to deep vertical solute propagation at dry no till conditions when cracks are open, connected down to the subsoil and infiltration rates through macropores are important. Keeping in mind the shallow water table at this site, the risk for potential groundwater contamination is high. Tillage destroys macroporosity in the topsoil and disconnects it from the subsoil, thus reducing the amount of solutes infiltrating into greater depths. However, it does not prevent macropore flow so that buried macropores could still function as preferred transport paths and allow solute and pesticide leaching down to the groundwater.

### 2.3.4 Conclusions

The interplay between drying/rewetting periods and switching between homogenous and preferential flow and transport regimes might induce large gradients and high dynamics (e.g. aerobic/anaerobic conditions, heterogeneous distribution of soil moisture) in contact regions between mobile and immobile domains. This affects chemical reactions, distribution of nutrients and solutes and thus their availability to plants. Further studies could consider analysis of soil chemical properties of cracks and soil matrix separately to detect eventual differences between these compartments. Buried cracks conducting water flow suggest their stability over a certain period of time. Isotope signatures of cracks and soil matrix might help to determine whether desiccation cracks develop at the same location.

Stereophotogrammetry allows monitoring processes on soil surface with high spatial and temporal resolution. In contrast, subsurface processes are usually considered by means of flow patterns, i.e. only with a high spatial resolution, because the dynamic of infiltration can be captured by point measurements with TDR probes or tensiometers only partially. Recent developments in geoelectrical methods like 3D Electrical Resistivity Tomography (ERT) are promising techniques that could be adapted to monitor heterogeneous infiltration of water into soils.

### 2.3.5 Acknowledgements

The authors are grateful to the technical stuff of INRA Avignon for their help during field work and to Prof. M. Schlather for his assistance on data analysis. The first author was financially supported by the Deutsche Forschungsgemeinschaft (DFG FOR 562) and the Deutsch-Französische Hochschule.

### 2.3.6 References

Adobe. 2005. Photoshop ver. CS2. (At: http://www.adobe.com. Accessed: 04/03/2009)

Bogner, C., Wolf, B., Schlather, M. & Huwe, B. 2008. Analysing flow patterns from dye tracer experiments in a forest soil using extreme value statistics. *European Journal of Soil Science*, **59**, 103-113.

Eltz, F.L.F. & Norton, L.D. 1997. Surface roughness changes as affected by rainfall erosivity, tillage, and canopy cover. *Soil Science Society of America Journal*, **61**, 1746-1755.

Flury, M. & Flühler, H. 1994. Brilliant Blue FCF as a dye tracer for solute transport studies - a toxicological overview. *Journal of Environmental Quality*, **23**, 1108-1112.

Flury, M., Flühler, H., Jury, W.A. & Leuenberger, J. 1994. Susceptibility of soils to preferential flow of water: a field study. *Water Resources Research*, **30**, 1945-1954.

Forrer, I.E., Papritz, A., Kasteel, R., Flühler, H. & Luca, D. 2000. Quantifying dye tracers in soil profiles by image processing. *European Journal of Soil Science*, **51**, 313-322.

Meteo France. 2009. http://france.meteofrance.com. Accessed: 15/04/2009

Fried, R. & Schettlinger, K. 2008. robfilter: A set of functions to filter time series based on concepts from robust statistics. R package version 2.4.

German-Heins, J. & Flury, M. 2000. Sorption of Brilliant Blue FCF in soils as affected by pH and ionic strength. *Geoderma*, **97**, 87-101.

Ghodrati, M., Ernst, F.F. & Jury, W.A. 1990. Automated spray system for application of solutes to small field plots. *Soil Science Society of America Journal*, **54**, 287-290.

MVTec Software GmbH. 2007a. Halcon ver. 8.0. (At: http://www.mvtec.com/halcon/. Accessed: 04/03/2009)

MVTec Software GmbH. 2007b. Solution Guide I. Basics. Halcon Technical Manual.

MVTec Software GmbH. 2007c. Solution Guide II-F. 3D Machine Vision. Halcon Technical Manual.

Huang, C.H. 1998. Quantification of soil microtopography and surface roughness. In: Fractals in Soil Science (eds. Baveye, P., Parlange, J.Y. & Stewart, B.A.). CRC Press Boca Raton, Florida.

Huang, C.H. & Bradford, J.M. 1992. Application of a laser scanner to quantify soil microtopography. *Soil Science Society of America Journal*, **56**, 14-21.

Jarvis, N.J. 2007. A review of non-equilibrium water flow and solute transport in soil macropores: principles, controlling factors and consequences for water quality. *European Journal of Soil Science*, **58**, 523-546.

1990. Digital close-range photogrammetry for surface measurement.

Jester, W. & Klik, A. 2005. Soil surface roughness measurement - methods, applicability, and surface representation. *Catena*, **64**, 174-192.

Kasteel, R., Vogel, H.J. & Roth, K. 2002. Effect of non-linear adsorption on the transport behaviour of Brilliant Blue in a field soil. *European Journal of Soil Science*, **53**, 231-240.

Ketelsen, H. & Meyer-Windel, S. 1999. Adsorption of Brilliant Blue FCF by soils. *Geoderma*, **90**, 131-145.

Kulli, B., Gysi, M. & Flühler, H. 2003. Visualizing soil compaction based on flow pattern analysis. *Soil & Tillage Research*, **70**, 29-40.

Logsdon, S.D. 1995. Flow mechanisms through continuous and buried macropores. *Soil Science*, **160**, 237-242.

Lu, J. & Wu, L. 2003. Visualizing bromide and iodide water tracer in soil profiles by spray methods. *Journal of Environmental Quality*, **32**, 363-367.

Mandelbrot, B.B. & Van Ness, J.W. 1968. Fractional Brownian motions, fractional noises and applications. *SIAM review*, 422-437.

Mirzaei, M.R. 2008. Effet de la variabilité spatiale des états de surface du sol sur le partage infiltration-ruissellement: Caractérisation expérimentale par photogrammétrie et modélisation. Ph.D. thesis, University of Avignon, Avignon.

Moreno, R.G., Alvarez, M.C.D., Requejo, A.S. & Tarquis, A.M. 2008. Multifractal analysis of soil surface roughness. *Vadose Zone Journal*, **7**, 512-520.

Paz-Ferreiro, J., Bertol, I. & Vazquez, E.V. 2008. Quantification of tillage, plant cover, and cumulative rainfall effects on soil surface microrelief by statistical, geostatistical and fractal indices. *Nonlinear Processes in Geophysics*, **15**, 575-590.

Pickering, N.B., Richard, T.L., Parlange, J.Y., Waltman, W.J. & Kromm, L. 1988. Dye tracer analysis of preferential flow to groundwater. Department of Agricultural and Biological Engineering, Staff Report 88-2. Cornell University, Ithaca, N. Y.

Schlather, M. 2008. RandomFields. R package version 1.3.37.

Schlather, M. & Huwe, B. 2005. A risk index for characterising flow pattern in soils using dye tracer distributions. *Journal of Contaminant Hydrology*, **79**, 25-44.

R Development Core Team. 2008. R: A Language and Environment for Statistical Computing. (At: http://www.R-project.org. Accessed: 04/03/2009)

The MathWorks, Inc. 2005. Matlab ver. 7.1. (At: http://www.mathworks.com. Accessed: 04/03/2009)

Vidal Vázquez, E., Miranda, J.G.V., Alves, M.C. & Paz González, A. 2006. Effect of tillage on fractal indices describing soil surface microrelief of a Brazilian Alfisol. *Geoderma*, **134**, 428-439.

Vidal Vàzquez, E.V., Moreno, R.G., Miranda, J.G.V., Diaz, M.C., Requejo, A.S., Ferreiro, J.P. & Tarquis, A.M. 2008. Assessing soil surface roughness decay during simulated rainfall by multifractal analysis. *Nonlinear Processes in Geophysics*, **15**, 457-468.

Warner, W.S. 1995. Mapping a 3-dimensional soil surface with hand-held 35 mm photography. Soil & Tillage Research, 34, 187-197.

Weiler, M. & Flühler, H. 2004. Inferring flow types from dye patterns in macroporous soils. *Geoderma*, **120**, 137-153.

IUSS Working Group WRB. 2007. World reference base for soil resources 2006, first update 2007. World Soil Resources Reports No. 103.

# Chapter 3

# Mechanisms and impacts of preferential flow

## *3.1 Investigating flow mechanisms in a forest soil by mixed-effects modelling*

BOGNER C[a,b], GAUL D[c], KOLB A[a], SCHMIEDINGER I[a], HUWE B[a]

[a]*Soil Physics Group, University of Bayreuth, 95440 Bayreuth, Germany,*

[b]*Ecological Modelling, University of Bayreuth, Dr.-Hans-Frisch-Straße 1-3, 95448 Bayreuth, Germany and* [c]*Plant Ecology, Albrecht-von-Haller Institute of Plant Sciences, University of Göttingen, Untere Karspüle 2, 37073 Göttingen, Germany*

Correspondence: C. Bogner. E-mail: christina.bogner@uni-bayreuth.de.

Short title:

*Flow mechanisms in a forest soil*

Submitted to the *European Journal of Soil Science*

## Summary

Uniform and preferential flow produce typical infiltration patterns. We did three tracer experiments in a Norway spruce forest soil and qualitatively identified the dominant flow regime based on stained patterns. We analysed soil texture, fine root density and soil bulk density from preferential flow paths and the soil matrix by means of linear mixed-effects models. They can account for dependencies in the data structure due to hierarchical sampling and can deal with missing values. There were between 44% (topsoil) and 76% (subsoil) higher root densities in preferential flow paths than in the soil matrix. The content of sand decreased and silt increased with depth, but there were no significant differences between flow domains. The bulk density was higher in the soil matrix by 0.12 g $cm^{-3}$, which is probably due to a higher soil organic matter content of preferential flow paths. Based on flow patterns and model results we identified the dominant flow mechanisms. Roots constituted main preferential flow paths and induced macropore flow, especially in the topsoil. In the subsoil root density decreased and inhomogeneous infiltration from preferential flow paths into the soil matrix caused unstable flow. Decreasing sand content might facilitate the lateral dispersion of the flow.

### 3.1.1 Introduction

Water flow in soils can be classified in two categories – uniform flow (Green & Ampt, 1911) and non-uniform i.e. preferential flow (Lawes *et al.*, 1882). The latter one comprises all phenomena where water flows through localised pathways bypassing a portion of the soil matrix (Hendrickx & Flury, 2001). Thus non-uniform flow results in rapid water movement and solute transport to greater depths than predicted by the classical concept of flow through unsaturated soils (Richards' equation).

Preferential flow can further be classified in macropore flow, unstable flow and funnel flow. Macropore flow describes water movement along root channels, earthworm burrows, soil fissures, cracks or large inter-aggregate voids in cultivated soils. The reader is referred to the review by Jarvis (2007) for more details on preferential flow through macropores. Unstable flow is common in coarse-textured soils and can be induced by variations in texture, water repellency, air entrapment or continuous non-ponding infiltration. Lateral redirection of water

caused by textural boundaries is referred to as funnel flow (Hendrickx & Flury, 2001). The impact on groundwater quality might be considerable if preferential flow – especially through continuous macropores – becomes a dominant mechanism of transport of agrochemicals to greater depths (Jarvis, 2007).

The three types of non-uniform flow lead to different flow patterns that can be visualised by tracer experiments (e.g. Flury *et al.*, 1994; Forrer *et al.*, 2000). Weiler & Flühler (2004) for instance showed typical dye patterns resulting from different flow regimes and classified them based on the width of the stained objects. They used the classification to infer flow types from dye patterns. Kulli *et al*. (2003) performed a hierarchical clustering of dye patterns to distinguish between zones of homogeneous infiltration and zones of preferential flow.

In our study we used Brilliant Blue to trace water flow and define preferential flow paths as stained areas. The goal of our research is (i) to qualitatively identify the dominant flow regime based on dye patterns and (ii) to test whether soil texture, root density and soil bulk density differed significantly between regions of preferential flow and soil matrix and could give insight into mechanisms of water flow.

### 3.1.2 Materials and methods

*Tracer experiments*

The study site – a Norway spruce forest – is situated in the Fichtelgebirge (770 m above sea level) in southeast Germany. The mean annual precipitation at the study site is about 1160 mm, and the mean annual air temperature about 5.3°C (Foken, 2003). The soil is classified as Haplic Podzol (IUSS Working Group WRB, 2007) and has a sandy to loamy texture with a mor-like organic layers of up to 10 cm thickness. Five mineral horizons – Ea, Bsh, Bs, Bw and Bw/C – can be identified. Hentschel et al. (2008) give a detailed overview about the soil chemistry.

We did three tracer experiments with Brilliant Blue FCF and potassium iodide. Brilliant Blue is a popular tracer in vadose zone hydrology because its colour is readily seen against most soil colours, and it is toxicologically acceptable for environmental use (Flury & Flühler, 1994; Mon *et al.*, 2006). Iodide was used as reference tracer because Brilliant Blue might be strongly sorbed by soil particles and thus considerably retarded with regard to water flow (Ketelsen & Meyer-Windel, 1999; German-Heins & Flury, 2000; Kasteel *et al.*,

2002). The tracer solution was applied on plots of about 2 $m^2$ with a sprinkler similar to that described by Ghodrati *et al.* (1990). The matric potential before irrigation on plot one was −254 hPa in 0.2 m, −216 hPa in 0.4 m and −157 hPa in 0.9 m. On plot two we measured −75 hPa in 0.2 m, −37 hPa in 0.4 m and −20 hPa in 0.9 m and on plot three −35 hPa in 0.2 m, −30 hPa in 0.4 m and −9 hPa in 0.9 m. We did not remove the litter, but took away spruce cones as they covered a large portion of the soil surface. The irrigation rate was 64 mm $hour^{-1}$ and the concentration of the tracers 5 g $l^{-1}$ respectively. Our study is part of a project dealing with impacts of extreme meteorological events on soil processes, thus the high irrigation rate. For comparison, the maximum 10-minutes intensity measured at the study site between 1999 and 2006 was 22 mm and the maximum one-hour intensity was 54 mm. So the irrigation rate was fairly high but not unrealistic.

The day after irrigation, 11 vertical soil profiles of 1 m × 1 m were prepared and photographed with a digital CCD camera in RAW format. We used a frame of 1 $m^2$ and two Kodak reference gray scales for geometric correction and white balance adjustment of the photographs. Two profiles per plot were treated with an indicator solution of iron(III) nitrate and starch (Lu & Wu, 2003) to visualize iodide infiltration patterns.

### *Soil sampling and laboratory analyses*

Profiles not treated with iron(III) nitrate and starch were sampled for laboratory analysis. In three profiles per plot we took bulk samples from Brilliant Blue stained and non-stained parts per soil horizon for analysis of soil texture, exchangeable cations and total C and N contents. Some thin horizons were largely bypassed by the flow and could not be sampled. Thus the number of samples differed from plot to plot resulting in 58 samples altogether. Three other profiles per plot were sampled for root density with a soil corer (diameter = 34 mm and height = 50 mm) in –10-0 cm (organic horizons), 0-20 cm, 20-40 cm and 40-60 cm depth. We could not sample every soil horizon because some of them were thinner than the diameter of the corer. We took three samples from stained areas and three from non-stained ones per profile. In some depths, there was no staining so that the total number of samples was 210. A mean value of root density was calculated per depth and sampled profile resulting in 71 values. In three other

profiles we took 96 samples for soil bulk density with a soil corer (diameter = 23 mm and height = 10 mm) in stained and non-stained areas in different depths.

Soil samples for analysis of root density were stored for 14 days at 4 °C. After washing them in a sieve (mesh size = 0.63 mm), fine roots (diameter < 2 mm, length > 10 mm) were dried at 70 °C for 48 h and weighed. Soil texture was analyzed by laser diffraction using a Mastersizer S (Malvern). To determine the soil bulk density, soil samples were dried at 105 °C for 24 h and weighed. Additionally, soil texture was analysed as described above. Analysis of exchangeable cations and total C and N content are described in Bogner *et al*. (2009). Table 3.1.1 summarizes the analyzed data.

Table 3.1.1: Summary of the sampled data

| Analyzed parameters | Sampling Method | Meta information | Number of samples |
|---|---|---|---|
| Soil texture | Bulk samples | Plot, profile, horizon, flow region | 58 |
| Bulk density | Metallic cylinders (diameter = 23 mm and height = 10 mm) | Plot, profile, depth, soil texture, flow region | 96 |
| Root density | Metallic cylinder (diameter = 34 mm and length = 50 mm) | Plot, profile, depth, flow region | 71 |

## *Image Analysis*

Photographs of soil profiles were taken in the lossless RAW-format (unprocessed data of the CCD camera sensor), adjusted for white balance and converted to JPEG in Photoshop CS2 (Adobe, 2005). Then the images were corrected in HALCON (MVTec Software GmbH, 2007a) for perspective and radial distortion. Perspective transformation of the images was obtained via a homogenous transformation matrix (a matrix that combines information about both rotation and translation in one matrix) and radial distortion was modelled by:

$$\begin{pmatrix} u \\ v \end{pmatrix} = \frac{1}{1+\kappa\left(\tilde{u}^2+\tilde{v}^2\right)}\begin{pmatrix} \tilde{u} \\ \tilde{v} \end{pmatrix}, \tag{3.1.1}$$

where the parameter $\kappa$ is the magnitude of the radial distortion, $(\tilde{u}, \tilde{v})^T$ are coordinates of a point in the original image and $(u, v)^T$ are coordinates in the corrected one. If $\kappa$ is negative, the distortion is barrel-shaped, while for positive $\kappa$ it is pincushion-shaped (MVTec Software 2007c).

To classify the images in Brilliant Blue stained and non-stained areas we transformed them from RGB to HSI (hue saturation intensity) colour space where colour segmentations are robust under varying illumination (MVTec Software GmbH, 2007b). Misclassifications due to the shadow of the metallic frame or large stones were corrected manually. The classified binary images (stained parts were coded black and non stained areas white) were used to calculate the dye coverage function (number of stained pixels per depth) using MATLAB (The MathWorks, 2005). The binary images were used to describe the dominant flow regime qualitatively.

## *Data Analysis*

### Mixed-effects models

Our data were not sampled randomly but are grouped on three levels: on the plot, the profile and the depth/horizon levels. Such grouped structure might induce dependencies between data, so for instance data sampled on one plot might have higher root densities than data from another plot. If we were interested in differences between exactly those three chosen experimental plots or exactly those profiles that were sampled, we could add a plot effect and a profile effect to our model. But in our study the plots were chosen at random. Although the profiles have a fixed spacing and are parallel to each other the exact position of the plot and thus the orientation of the profiles in space was chosen at random as well. The depth/horizon level is quite different. We are indeed interested in systematic variations of soil physical parameters and root densities between exactly those depths/horizons that were sampled. But there might be a random fluctuation between the same horizons on different plots, i.e. horizon Ea on plot one might have a higher sand content than on plot two. Therefore, we should consider a possible dependence induced on the depth/horizon level as well.

Additionally to probable dependences within the data, we have to deal with missing values because some horizons were largely bypassed by the flow and

could not be sampled. A good solution for the problem of grouped data and missing values is to use a mixed-effects model. It can account for fixed-effects representing parameters of the entire population or certain repeatable levels of experimental factors (like horizon) and for random-effects associated with individual experimental units drew at random from a population (like plots or profiles). Furthermore, it is robust against missing values (Pinheiro & Bates, 2000). According to the formulation proposed by Laird and Ware (1982) we can write a linear mixed-effects model for a single level of grouping as:

$$\begin{aligned} &\mathbf{y}_i = \mathbf{X}_i\boldsymbol{\beta} + \mathbf{Z}_i\mathbf{b}_i + \boldsymbol{\varepsilon}_i, \\ &\mathbf{b}_i \sim \mathrm{N}(0, \boldsymbol{\Psi}), \qquad \boldsymbol{\varepsilon}_i \sim \mathrm{N}(0, \sigma^2\boldsymbol{\Lambda}_i), \qquad i = 1, \ldots, M, \end{aligned} \tag{3.1.2}$$

where $\mathbf{y}_i$ is the $n_i$-dimensional response vector for the $i$th group, $\boldsymbol{\beta}$ is the $p$-dimensional vector of fixed-effects, $\mathbf{b}_i$ is the $q$-dimensional vector of random-effects, $\mathbf{X}_i$ and $\mathbf{Z}_i$ are the known fixed-effects and random-effects regressor matrices (also called design matrices) of size $n_i \times p$ and $n_i \times q$ respectively and $\boldsymbol{\Lambda}_i$ are positive-definite matrices. The single-level mixed-effects model can be extended to multiple, nested or crossed levels of random-effects (Pinheiro & Bates, 2000). The calculations were done in R (R Development Core Team, 2008) using the package nlme (Pinheiro *et al.*, 2008) that estimates parameters of the mixed-effects model by the restricted (or residual) likelihood (REML).

Model building strategy

We modelled soil fine texture (bulk samples), root density and soil bulk density. In all models we used flow region as a fixed-effect because our primary goal was to know whether there are significant differences between preferential flow paths and soil matrix. Flow region is a factor with levels ‘soil matrix’ and ‘preferential flow path’ indicating the origin of samples. Additionally, we included the depth for root density and soil bulk density and the horizon for texture as fixed-effect. The depth in the root density model is a factor with levels ‘5’, ‘20’, ‘40’ and ‘60’ indicating the middle of the sampled section in cm (taking the soil surface as zero). In the bulk density model the depth is a continuous variable. The horizon is a factor with levels ‘Ea’, ‘Bsh’, ‘Bs’, ‘Bw’ and ‘Bw/C’. The model for the bulk density was improved by taking the content of sand – a continuous covariate – into account. Shieh & Fouladi (2003) stated in their paper that collinearity in a

single-level mixed-effects model may inflate the estimates of variance and covariance components. Therefore we used horizon Bw as reference level because it had the largest number of observations. Using any other horizon with fewer observations as reference level increased the standard errors of the estimated fixed-effects due to collinearities with the intercept term (Wissmann & Toutenburg, 2007). We checked for further collinearities using the diagnostics proposed by Belsley *et al*. (1980). It calculates variance decomposition proportions based on a singular-value decomposition of the design matrix of the fixed-effects. A good tutorial on the implementation of the method can be found in Belsley (1991).

To find adequate models for our data we first built a full mixed-effects model incorporating all fixed-effects described above and random-effects for the plot, the profile within plot and the depth/horizon within plot. Then, the significance of random-effects was tested against a nested restricted model i.e. a model with a smaller number of random-effects but the same fixed-effects and the same variance-covariance structure of the within-group errors using a likelihood ratio test (LRT). The LRT statistic is defined as:

$$2\log(L_2/L_1) = 2[\log(L_2) - \log(L_1)] , \qquad (3.1.3)$$

where $L_2$ is the likelihood of the full model and $L_1$ the likelihood of the restricted model (with $L_2 > L_1$). If $k_2$ is the number of parameters estimated in the full model and $k_1$ the number of parameters of the restricted model, then the "large sample" or asymptotic distribution of the LRT statistic is a $\chi^2$ distribution with $k_2 - k_1$ degrees of freedom. And the null hypothesis tested is that the restricted model fits the data as well as the full model. This test can also be performed for models fit by REML if both models were fit by REML and if the structure of the fixed-effects remains the same (Pinheiro & Bates, 2000). Finally, if appropriate, the significance of parameters describing the variance-covariance structure of the within-group errors was assessed by the likelihood ratio test against a homoscedastic model.

Before making inference about the fixed-effects we examined whether the assumptions underlying our models were met. The first assumption is that the within-group errors that are allowed to be heteroscedastic and correlated are independent for different groups and independent of the random-effects. The

within-group residuals provide good surrogates of the within-group errors and can be used in diagnostics plots to check this assumption qualitatively (Pinheiro & Bates, 2000). We utilized QQ-plots of residuals to verify the normality, scatter plots of residuals versus fitted values or other covariates by group to check whether the residuals are centred at 0 and have equal variances across groups and are independent of the group levels. The second assumption is that the random effects are normally distributed, with mean 0 and covariance matrix **Ψ** (not depending on the group) and are independent for different groups. As in our models only simple random intercepts were used QQ-plots of random-effects were sufficient to assess this assumption.

Variable transformation

The distribution of the root density data was non-normal and we could not model the heteroscedasticity of the within-group errors because of its complex structure. Thus, root density data were transformed using the scaled Box-Cox transformation (Box & Cox, 1964) to achieve (an approximate) normality and stabilise the variance:

$$y^{(\lambda)} = \begin{cases} \left(y^{\lambda} - 1\right) / \lambda \tilde{y}^{\lambda - 1} & \lambda \neq 0 \\ \tilde{y} \log(y) & \lambda = 0 \end{cases}, \tag{3.1.4}$$

where $y^{(\lambda)}$ is the transformed value, $\tilde{y}$ is the geometric mean of the data and $\lambda$ the transformation constant. This transformation requires the data being positive. Hence, we added a constant $c$ to the data because their range included 0. We tried 0.01, 0.1 and 1 as constants and the best fit was achieved with 0.1.

An appropriate value for $\lambda$ was determined in a grid search, i.e. we calculated the mixed-effects model for a range of possible values for $\lambda$. Gurka *et al*. (2006) stated in their paper that using the scaled Box-Cox transformation instead of the simple Box-Cox transformation allowed employing the existing mixed-effects model estimation procedures in the grid search. Indeed, the residual likelihood of the scaled Box-Cox transformed model is that of a standard mixed-effects model. The estimate of the transformation constant $\hat{\lambda}$ is located by plotting the residual likelihood values against $\lambda$ as the value that corresponds to the highest residual likelihood (Gurka *et al.*, 2006). A jackknife estimate of the variance of $\hat{\lambda}$ was calculated as proposed by Lipsitz *et al*. (2000):

$$\mathrm{var}(\hat{\lambda}) = \sum_{i=1}^{N} \frac{N-1}{N}\left(\hat{\lambda}_{-i} - \hat{\lambda}\right)\left(\hat{\lambda}_{-i} - \hat{\lambda}\right)^{T}, \qquad (3.1.5)$$

where $\hat{\lambda}$ is the estimate of $\lambda$ using the whole data set, $\hat{\lambda}_{-i}$ is the estimated value obtained deleting the *i*th observation, *N* is the number of observations. The variance of $\hat{\lambda}$ indicates whether $\hat{\lambda}$ can be rounded to a convenient value to obtain a model that is easier to interpret like for instance the log-model. Interpretability is a problem for transformed data and it is even a bigger problem when using the scaled Box-Cox transformation. Therefore, for the final model we used the simple Box-Cox transformation as $\hat{\lambda}$ estimated under the scaled transformation holds for the simple Box-Cox transformed model as well. Dropping the scaling factor ỹ from the model introduce a slight bias in the estimation of fixed-effects (Gurka *et al.*, 2006). To assess the sensitivity of the estimated fixed-effects to the transformation constant $\lambda$ we followed the advice of Gurka *et al*. (2006) and Lipsitz *et al*. (2000) and compared the inference about the fixed-effects between the optimal model with $\hat{\lambda}$ that was found by the grid search, the model with $\lambda_{\min}$ equal to $\hat{\lambda} - \sqrt{\mathrm{var}(\hat{\lambda})}$, with $\lambda_{\max}$ equal to $\hat{\lambda} + \sqrt{\mathrm{var}(\hat{\lambda})}$ and the convenient model with $\lambda$ rounded appropriately with and without the scaling parameter $\tilde{y}$. Our concern is about interpretability and we prefer using a convenient value for $\lambda$ that might introduce a slight bias to more exact estimates of fixed-effects that are physically meaningless. In a last transformation step we optimised the constant *c* in a grid search.

Goodness-of-fit and multiple comparisons

The parameters of the mixed-effects models are estimated by REML. Thus, the measure of goodness of fit $R^2$ used for least squares estimates is not defined. To assess the overall goodness of fit of a mixed-effects model we can use the correlation between the fitted and the observed values (J. Pinheiro, personal communication). Furthermore, we can calculate the percentage of variance explained by fixed-effects by comparing the full model with an empty model (with the intercept as the only fixed-effect) (Edwards *et al.*, 2008):

$$R_{\beta}^{2} = \frac{(q-1)v^{-1}F\left(\hat{\boldsymbol{\beta}}, \hat{\boldsymbol{\Psi}}\right)}{1 + (q-1)v^{-1}F\left(\hat{\boldsymbol{\beta}}, \hat{\boldsymbol{\Psi}}\right)} \cdot 100\%, \qquad (3.1.6)$$

where $q$ is the number of estimated parameters in the full model, $v$ are the residual degrees of freedom ($v = n - \text{rank}(\mathbf{X})$, where n is the number of observations and $\mathbf{X}$ the regressor matrix of fixed-effects), $F\left(\hat{\boldsymbol{\beta}}, \hat{\boldsymbol{\Psi}}\right)$ is the $F$ statistics for a Wald test of the null hypothesis $H_0$: $\beta_1 = \beta_2 = \cdots = \beta_{q-1} = 0$ i.e. that all parameters except the intercept are zero. This definition allows calculating the explained variance using a single model fit and thus identical structure of ransom-effects.

After fitting the models, we compared different depths/horizons where it was appropriate. Hothorn et al. (2008) extended the theory of multiple comparisons to mixed-effects models. Different adjustments of p-values are possible and we used a multiple testing procedures under free combinations (Westfall, 1999). The calculations were done using the add-in package multcomp in R (R Development Core Team, 2008).

### 3.1.3 Results

*Flow regime*

Following the definition by Flühler et al. (1996) the expression regime is used to qualify the phenomenology of the flow (preferential or uniform) that can be dominated by one or more physical mechanisms. Iodide patterns (not shown here) did not indicate any serious retardation of Brilliant Blue with regard to water. While the first few centimetres of the litter layer were homogenously stained, in the organic soil horizons the flow rapidly condensed into distinct flow paths (Figure 3.1.1). In the upper 20 cm of the profile the flow was dominantly preferential and a large portion of the matrix was bypassed. This region was densely rooted and we observed many blue stained roots. Deeper, the dye patterns differed considerably among the three experimental plots. On plot one the dye built distinct spots along the whole profile while on plot two large areas were stained. On plot three, we found a dense stone layer in about 40 cm depth that formed an obstacle for the flow. As indicated by the grey painted area in Figure 3.1.1, the amount of staining varied little on this plot and had its maximum in the stone layer. The highest variation of staining was observed on plot two. Despite the differences in patterns the main common feature of the flow regime in the subsoil is the dispersion of the flow from the preferential flow paths into the matrix.

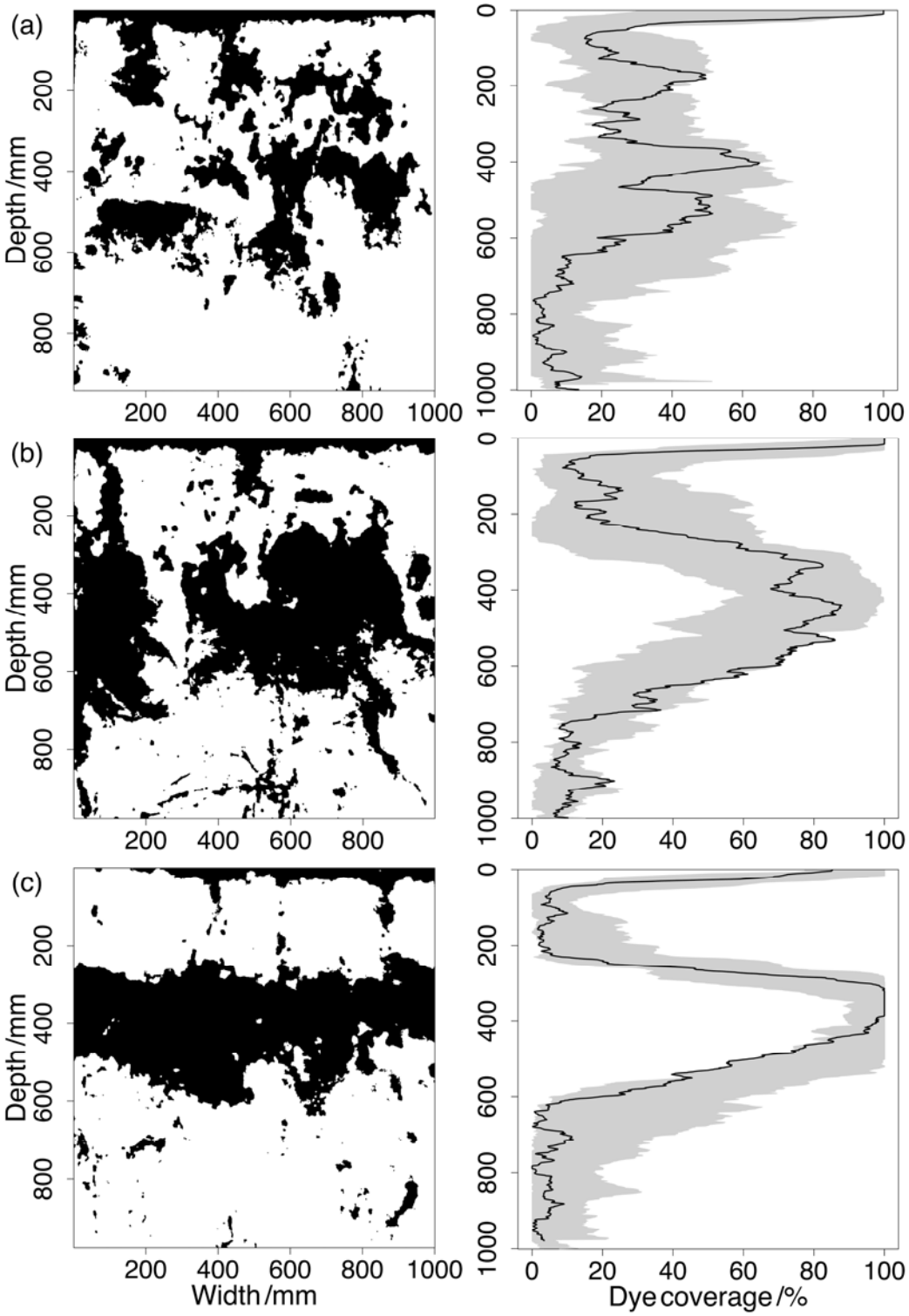



Figure 3.1.1: Example images of stained patterns (left) and dye coverage functions (right). From (a) to (c): plot one to three. The grey painted area indicates the variation between the minimum and the maximum staining calculated from all 11 profiles of a plot and the black line represents the dye coverage function of the picture on the left.

The dye patterns suggest that at our study site roots constitute main preferential flow paths and induce macropore flow, especially in the topsoil. In the subsoil root density decreases and inhomogeneous infiltration from preferential flow paths into soil matrix causes unstable flow. There is no sharp boundary between the region where the flow is dominantly preferential and the zone of dispersion into the soil matrix. Therefore texture differences between soil horizons do probably not influence the flow regime.

## *Modelling results*

Content of sand, silt and clay

Globally the content of sand decreased, the content of silt increased and the amount of clay varied considerably with depth suggesting a horizon within plot random effect (Figure 3.1.2):

$$y_{ijk} = \beta_0 + \beta_1 \cdot soil\ horizon + \beta_2 \cdot flow\ region + b_{ij} + \varepsilon_{ijk},$$
$$i=1, 2, 3, \qquad j=1,\ldots, 4, \qquad k=1,\ldots n_{ij}, \qquad (3.1.7)$$
$$b_{ij} \sim N(0, \sigma_1^2), \qquad \varepsilon_{ijk} \sim N(0, \sigma^2),$$

where $y_{ijk}$ is sand or silt in percent on plot $i$ in horizon $j$, $\beta_i$ are the fixed-effects, $b_{ij}$ is the horizon within plot random-effect, $\varepsilon_{ijk}$ is the error term and $n_{ij}$ the number of observations on plot $i$ in horizon $j$.

Figure 3.1.3a indicates a good agreement between fitted and observed values. As an example, we show diagnostic plots for the sand model. The same procedure was applied to all other models to ensure that the model assumptions were met and that points with high residuals did not influence the inference. There is no evidence of heteroscedasticity (Figure 3.1.3b) or departure from the normality assumption (Figure 3.1.3c and Figure 3.1.3d). Boxplots of standardized residuals versus horizon or flow region by plot (not shown here), did not suggest any violation, either. Observation 20 has the highest residual, but excluding it from the model did not change the inference about the fixed effects.

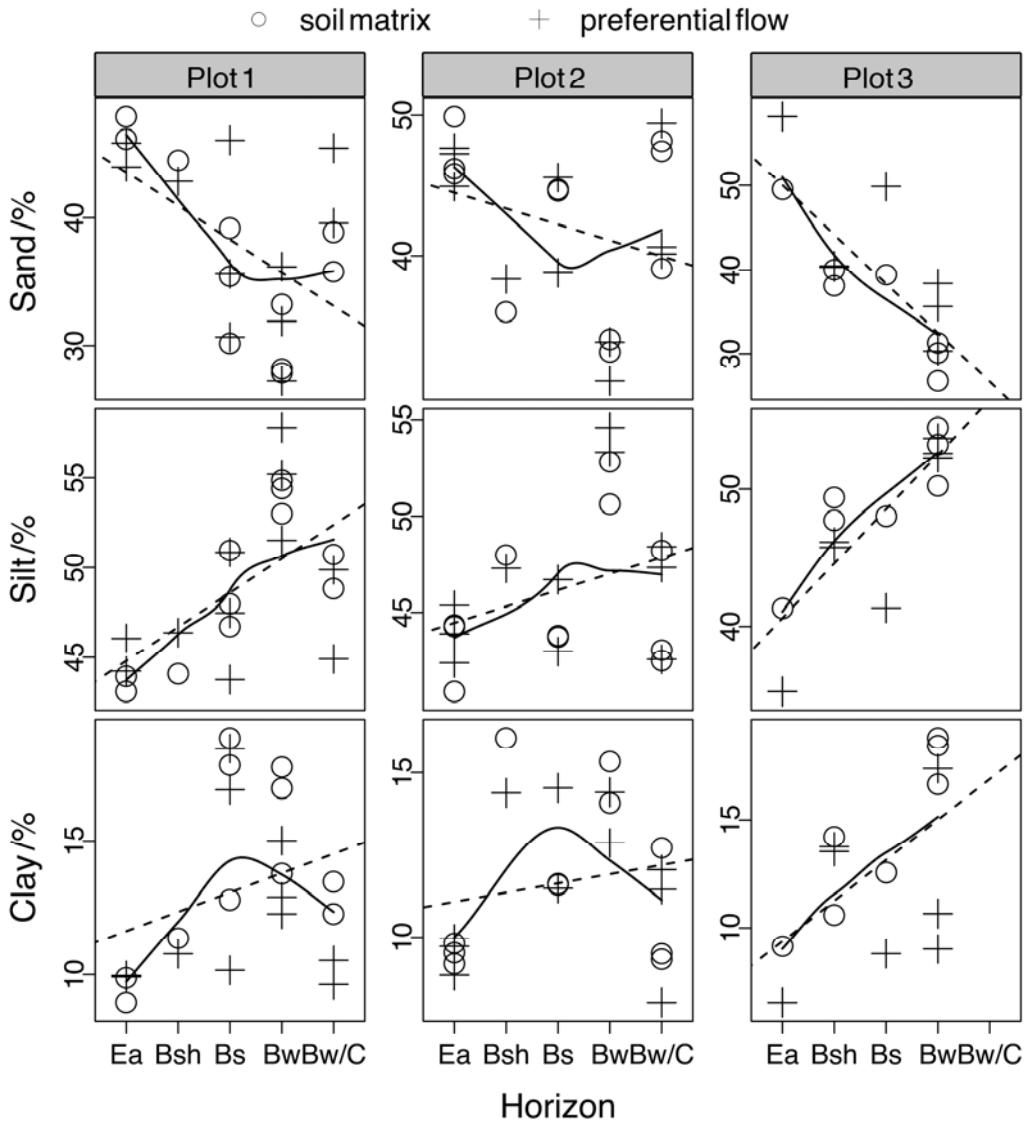



Figure 3.1.2: Particle size distribution of the soil fine fraction with depth (from left to right: plot one to three). Sand fraction is defined as 2000 - 63 µm, silt 63 - 2 µm and clay < 2 µm. Note different scaling of the y-axis. The solid line is a smooth curve computed by loess (with smoothness parameter 1.01) and the dotted line shows a linear regression.

The approximate 95% confidence intervals for the random-effects and the within-group errors are displayed in Table 3.1.2. The correlation between the observed and the fitted values was 0.78 (Figure 3.1.3a) and 0.81 (Figure 3.1.4a) and the variance explained by fixed-effects equalled 39% and 41% for the sand and the silt model respectively. There are no significant differences between the preferential flow paths and the soil matrix neither in the content of sand nor of silt (Figure 3.1.3e and Figure 3.1.4b).

Table 3.1.2: Estimated standard deviations and their approximate 95% confidence intervals for random-effects and within-group errors.

| | 95% lower | Estimate | 95% upper | | 95% lower | Estimate | 95% upper |
|---|---|---|---|---|---|---|---|
| | Sand /% | | | | Silt /% | | |
| Random-effect (horizon) | 1.6 | 3.1 | 6.1 | Random-effect (horizon) | 1.0 | 1.9 | 3.7 |
| Within-group error | 3.0 | 3.8 | 4.7 | Within-group error | 1.8 | 2.3 | 2.8 |
| | Clay /% | | | | Root abundance /log(g $l^{-1}$) | | |
| Random-effect (horizon) | 1.1 | 2.0 | 3.4 | Random-effect (horizon) | 0.1 | 0.3 | 0.5 |
| Random-effect (profile) | 0.4 | 0.9 | 2.0 | Within-group error | 0.3 | 0.4 | 0.5 |
| Within-group error | 1.1 | 1.4 | 1.8 | | | | |

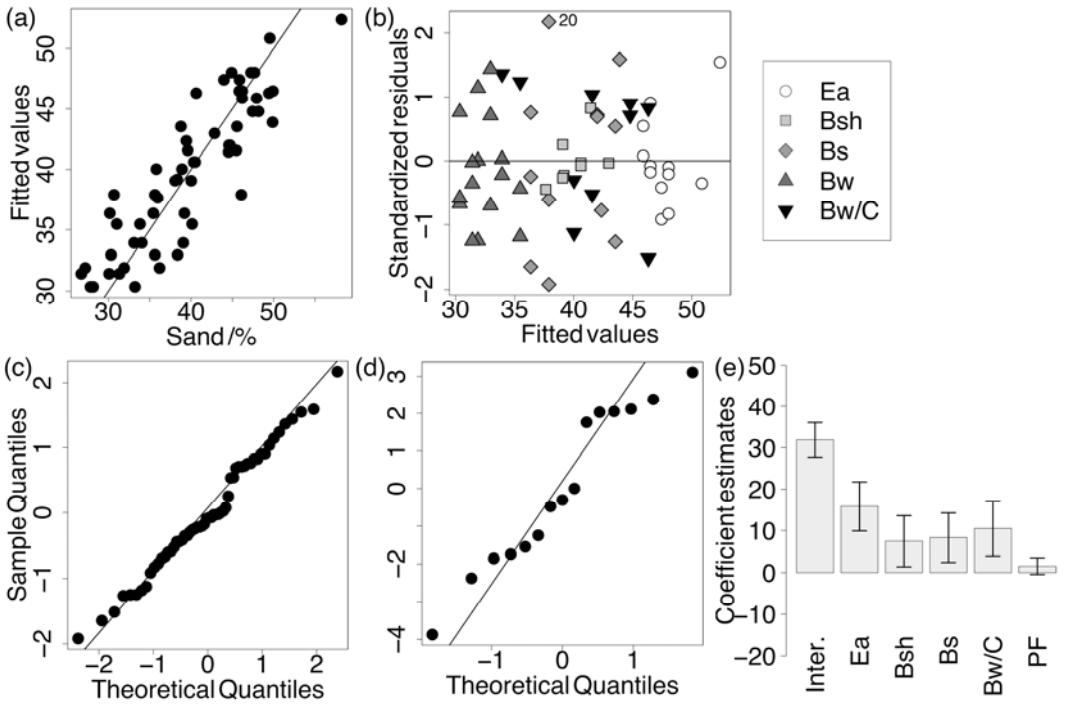



Figure 3.1.3: Diagnostic plots for the sand model. (a) fitted versus observed values (a 1:1 line was added for better visibility); (b) standardized residuals versus fitted values per horizon; (c) QQ-plot of the standardized residuals; (d) QQ-plot of the estimates of the horizons within plot random-effect; (e) estimated fixed-effects with their standard errors (Inter. = intercept, PF = preferential flow).

The content of clay was modelled as:

$$y_{ijkl} = \beta_0 + \beta_1 \cdot soil\ horizon + \beta_2 : flow\ region + b_{ij} + b_{ik} + \varepsilon_{ijkl},$$

$$i=1, 2, 3, \quad j=1,\ldots, 4, \quad k=1,\ldots, 4, \quad l=1,\ldots n_{ijk}, \qquad (3.1.8)$$

$$b_{ij} \sim N(0, \sigma_1^2), \quad b_{ik} \sim N(0, \sigma_2^2), \quad \varepsilon_{ijkl} \sim N(0, \sigma^2),$$

where $y_{ijkl}$ is the clay content in percent on plot $i$ in profile $j$ in horizon $k$, $\beta_i$ are the fixed-effects, $b_{ij}$ is the profile within plot random-effect, $b_{ik}$ is the horizon within plot random-effect, $\varepsilon_{ijk}$ is the error term and $n_{ijk}$ the number of observations on plot $i$ in profile $j$ in horizon $k$. We excluded three influential points from the fit which increased the correlation between the observed and the fitted values from 0.63 to 0.86 without changing the inference about the fixed-effects (Figure 3.1.4c). The percentage of variance explained by fixed-effects equalled 28%. The content of clay did not differ significantly between preferential flow paths and soil matrix (Figure 3.1.4d).

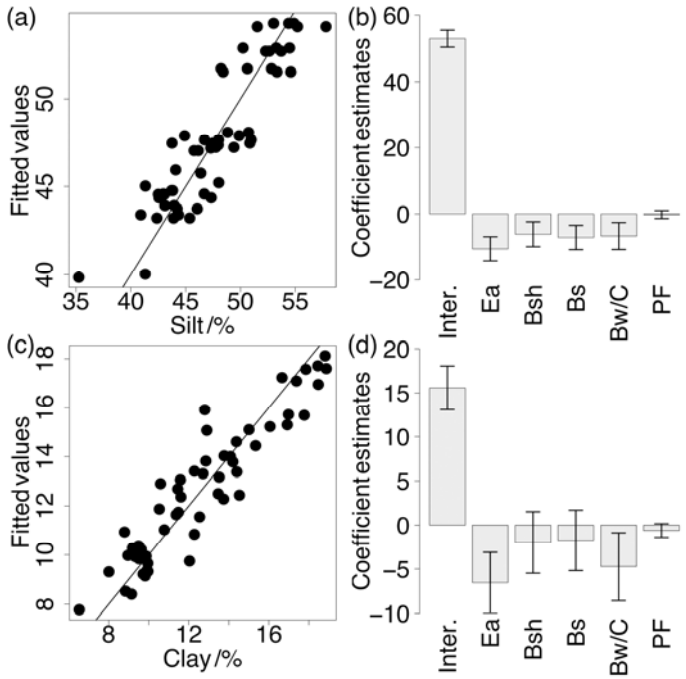



Figure 3.1.4: Results of the silt and clay models. (a), (c): fitted versus observed values a 1:1 line was added for better visibility); (b), (d): estimated fixed-effects effects with their standard errors (Inter. = intercept, PF = preferential flow paths).

Figure 3.1.2 suggests comparing the texture of successive horizons. We tested whether the content of sand decreased from the Ea to the Bw and increased from the Bw to the Bw/C and the contrary for the content of silt. As there was no clear trend for the content of clay, we only tested whether successive horizons were different. We could reject the hypotheses of decreasing sand content and increasing silt content between the Bsh and the Bs and of different clay content between the Bsh and the Bs and the Bs and the Bw (Table 3.1.3).

Root density

Root density decreased sharply with depth, but varied little between plots (Figure 3.1.5a). The grid search for the Box-Cox transformation resulted in a transformation constant $\hat{\lambda}$ of 0.16 with a jackknife variance of 0.009. The grid search for the constant $c$ gave 0.156. The comparison between models with $\lambda$ equal to 0.066 ($\lambda_{min}$), 0.254 ($\lambda_{max}$), and 0 (convenient value) led to similar estimates of the fixed-effects and we kept the convenient model:

$$\log(y_{ijk} + 0.156) = \beta_0 + \beta_1 \cdot depth + \beta_2 \cdot flow\ region + b_{ij} + \varepsilon_{ijk},$$
$$i=1, 2, 3, \quad j=1,\ldots, 4, \quad k=1,\ldots n_{ijk}, \qquad (3.1.9)$$
$$b_{ij} \sim N(0, \sigma_1^2), \quad \varepsilon_{ijk} \sim N(0, \sigma^2),$$

where $y_{ijk}$ is the root abundance in g $l^{-1}$on plot $i$ in depth $j$, $\beta_i$ are the fixed-effects, $b_{ij}$ is the depth within plot random-effect, $\varepsilon_{ijk}$ is the error term and $n_{ijk}$ the number of observation on plot $i$ in depth $j$. The correlation between the observed and the fitted values was 0.83 and the variance explained by the fixed-effects equalled 57% (Figure 3.1.5b). The factor flow region was significant indicating that the root density in preferential flow paths is higher than in the soil matrix (Figure 3.1.5c).

Table 3.1.3: Summary of multiple comparisons of the soil texture.

| | Linear hypothesis | Estimate | Standard error | p value |
|---|---|---|---|---|
| Sand | Ea − Bsh ≤ 0 | 8.366 | 3.128 | 0.010 |
| | Bsh − Bs ≤ 0 | −0.855 | 3.128 | 0.607 |
| | Bs − Bw ≤ 0 | 8.334 | 2.944 | 0.010 |
| | Bw/C − Bw ≤ 0 | 10.507 | 3.273 | 0.005 |
| Silt | Ea − Bsh ≥ 0 | −4.391 | 1.920 | 0.048 |
| | Bsh − Bs ≥ 0 | 0.967 | 1.920 | 0.617 |
| | Bs − Bw ≥ 0 | −7.070 | 1.812 | 0.001 |
| | Bw/C − Bw ≥ 0 | −6.631 | 2.013 | 0.005 |
| Clay | Ea − Bsh = 0 | −4.566 | 1.765 | 0.045 |
| | Bsh − Bs = 0 | −0.238 | 1.782 | 0.894 |
| | Bs − Bw = 0 | −1.740 | 1.710 | 0.495 |
| | Bw/C − Bw = 0 | −4.727 | 1.909 | 0.046 |

Using the convenient log-model has the advantage that we can retransform our results to the original scale quite easily. The mean of a log-normal variable is the exponential of the log mean plus one half of the log variance (Johnson *et al.*, 1994; Manning, 1998; Mullahy, 1998). As the residuals are normally distributed and homoscedastic, the retransformation on the plot level is straightforward:

$$y_{ijk} = \exp(\beta_0 + \beta_1 \cdot depth + \beta_2 \cdot flow\ region + \hat{b}_{ij} + 0.5 \cdot \hat{\sigma}^2) - 0.156, \qquad (3.1.10)$$

where $y_{ijk}$ is the root density on the original scale in g $l^{-1}$, $\hat{b}_{ij}$ is the estimate of the depth within plot random-effect, $\hat{\sigma}^2$ is the estimated within-group error variance and the other elements are the same as in equation (3.1.9). To obtain retransformed values on the population level, we replace the estimate of the random-effect by $0.5 \cdot \hat{\sigma}_1^2$, the estimate of the random-effect variance. The latter show a difference of about 44% in the organic horizons (5 cm depth) and even about 76% in the Bw/C horizon (60 cm depth), but there the absolute values are small (Figure 3.1.5d). Multiple comparisons between depths showed that the root density decreases significantly between 5 cm and 20cm and 40 cm and 60 cm (Table 3.1.4).

Table 3.1.4: Summary of multiple comparisons of the root density.

| Linear hypothesis | Estimate | Standard error | p value |
|---|---|---|---|
| D5 − D20 ≤ 0 | 1.046 | 0.246 | <0.000 |
| D20 − D40 ≤ 0 | 0.379 | 0.246 | 0.064 |
| D40 − D60 ≤ 0 | 0.651 | 0.247 | 0.011 |

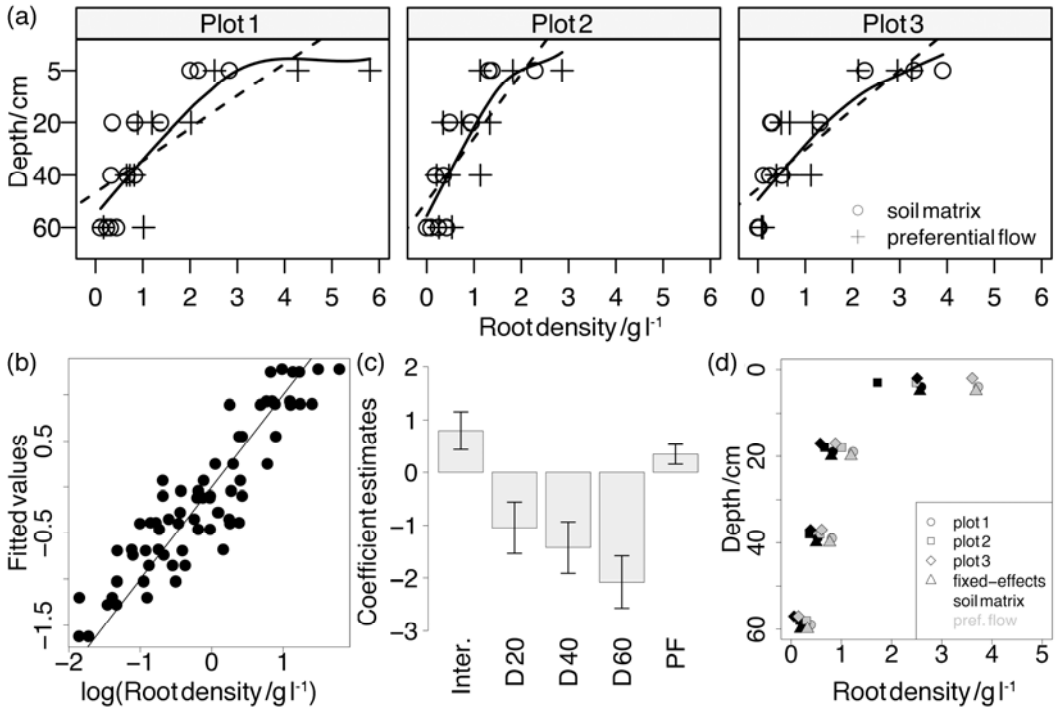



Figure 3.1.5: Variation of the root density with depth (a) and the results of the root abundance model: (b) fitted versus observed values (a 1:1 line was added for better visibility); (c) estimated fixed-effects effects with their standard

errors (Inter. = intercept, D = depth, PF = preferential flow paths); (d) model coefficients retransformed to original scale.

Bulk density

As expected, bulk densities increased with depth. The highest variability was encountered on plot two (Figure 3.1.6a). On plot three the soil was hardly stained below the stone layer and we sampled less and mostly in the topsoil where the bulk densities were smaller. We tested several models and the best results were achieved with the following linear model:

$$y_i = \beta_0 + \beta_1 \cdot content\ of\ sand + \beta_2 \cdot depth + \beta_3 \cdot flow\ region + \varepsilon_i,$$
$$\varepsilon_i \sim N(0, \sigma^2), \qquad (3.1.11)$$

where $y_i$ is the bulk density, $\beta_i$ are the regression parameters and $\varepsilon_i$ is the error term. We scaled the content of sand by 0.1 to decrease the condition number of the model matrix and thus increase numerical stability. We found collinearities between the intercept term, the depth and the sand content. Mean centring the latter two variables (i.e. subtracting the mean) resolved the problem.

The fitted values scatter considerably indicating that there might be some other explanatory variables not included in the model (Figure 3.1.6b). As there are no random-effects in the model, the only assumption to be verified is that the errors are normally distributed with mean 0 and a constant variance and the diagnostic plots (not shown) did not indicate any violation. The adjusted $R^2$ of model (3.1.11) equalled 0.53 and the residual standard error 0.25. All the estimated coefficients are significant and the bulk densities in the preferential flow paths are 0.12 g cm$^{-3}$ smaller than in the matrix (Figure 3.1.6c).

## 3.1.4 Discussion

### *Goodness-of-fit and model parameters*

The correlations between fitted and observed values for all mixed-effects models are larger than 0.7 and the percentage of variance explained by fixed-effects ranges between 28% and 57%. The estimated random-effects and the within group errors vary by factor 5 and by factor 2 respectively (Table 3.1.2). The plots of fitted versus observed values indicate some scatter (i.e. for the root density), but globally the models represent the data well. The bulk density model can be improved considerably by incorporating the content of organic matter as predictor. But preliminary results (not included here) showed that there was a

significantly higher content of organic matter in preferential flow paths than in the soil matrix. Therefore this predictor already incorporates differences between the flow domains and including it in the model changes the inference about the factor flow region and makes it not significant.

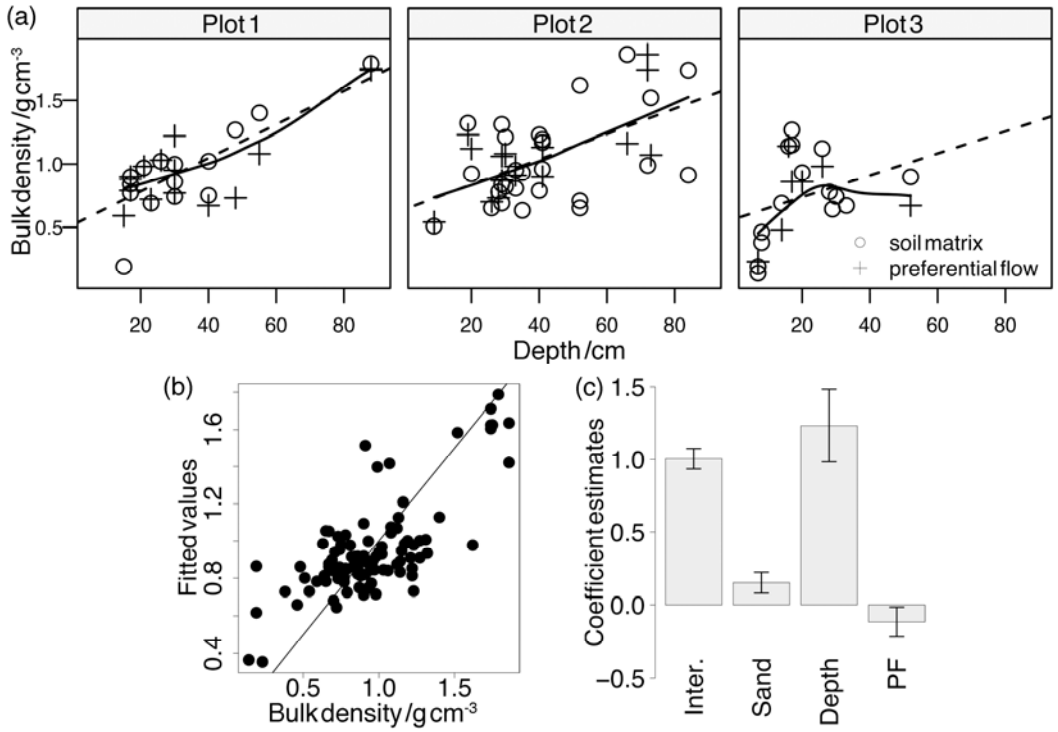



Figure 3.1.6: Variation of the soil bulk density with depth (a) and the results of the bulk density model: (b) fitted versus observed values (a 1:1 line was added for better visibility); (c) estimated coefficients (Inter. = intercept, PF = preferential flow paths).

The parameter intercept in the sand, silt and clay models indicates the content of the respective soil fine fraction in the horizon Bw in the soil matrix (reference level). The parameters Ea, Bsh, Bs and Bw/C are differences with regard to this reference level and the parameter preferential flow shows the difference between soil matrix and preferential flow paths. As an example consider the content of sand in the preferential flow paths of the horizon Ea (Figure 3.1.3e). As the flow region is not significant we only have to sum up the intercept (31.90%) and the parameter Ea (15.85%) and obtain 47.75% on the population level. If we need estimates on the plot level, we have to add the random-effects. The reference level for root abundance model is the depth 5 cm. As for the sand, silt and clay models the other parameters are differences to this reference level. The reference level for the bulk density is more complicated as the predictors have been centred. So the

intercept represents the bulk density in the soil matrix in the mean depth for the mean content of sand. The regression parameters for the sand and the depth are slopes and for the flow region differences to the preferential flow paths.

### *Flow mechanisms*

According to the description by Flühler *et al*. (1996) the homogenously stained top of the soil profile corresponds to the distribution flow in the attractor zone where water is funnelled towards "isolated entry ports". The results of the statistical analysis support our assumption on the dominant role of the root system for the occurrence of preferential flow. The Norway spruce has a shallow root system. A study at the same site by Gaul *et al*. (2008) mentioned that the first 30 cm of the soil profile contained 80% of the cumulative fine root biomass. The root density decreases significantly with depth and differs between the soil matrix and preferential flow paths. In the transmission zone (the horizons Ea, Bsh and partly Bs) water moves preferentially along root channels. These horizons have a high content of sand and thus a high conductivity which explains the lateral spreading of the dye from preferential flow paths into the soil matrix. This interaction between macropores and the matrix results in relatively large stained objects. The region of small root density coincides with the dispersion zone where water is forced to infiltrate into the soil matrix. The decreasing sand content might facilitate the lateral dispersion of the flow. As the root system is heterogeneous the upper limit of this zone varies in space.

The main differences between the plots are the mean root density and the stone content. Plot one has the highest root density (Figure 3.1.5d) and is characterized by isolated stained spots along the whole profile. Here preferential flow is still important in the subsoil horizons and we observed larger roots growing down to the horizon Bw/C. On plot three large stones induce funnel flow along stone walls and infiltration beyond the stone layer is hindered. So the dispersion here is mainly due to the presence of stones and not to the decrease of the root density. Stones are known to increase flow velocities by decreasing the cross section available for water flow (Schulin *et al.*, 1987; Hendrickx *et al.*, 1991; Sauer & Logsdon, 2002; Bogner *et al.*, 2008). At our study sites stones play an important role. According to Gerstberger *et al*. (2004) the content of rock fragments ranges between few percent in the topsoil and up to 75 vol% in the

subsoil. Additionally to the impact on water movement, stones might influence the soil chemistry as sources of cation exchange capacity (Ugolini *et al.*, 1996; Corti *et al.*, 2002) and constitute an additional reservoir of nutrients (Heisner *et al.*, 2004).

### *Ecological implications*

A significant difference between flow regions does not imply ecological relevance. The higher root densities in the stained areas in the topsoil are ecologically relevant, but the differences in the less rooted lower horizons are probably not. Roots constitute preferential paths for water flow which leads to a better water and nutrients supply than in the soil matrix and might amplify root production – a positive feedback. Further studies should consider both root density and root turnover to elucidate whether root production is higher in preferential flow paths. The difference in bulk density of 0.12 g $cm^{-3}$ is small. It is difficult to say whether it is important, but it might cause a higher porosity in preferential flow areas. Smaller bulk density is probably not the cause for but an effect of preferential flow. Root turnover is an important source of soil carbon and decomposition of dead roots is a major input to soil organic matter (Tate *et al.*, 1993; Guo *et al.*, 2005). Soil bulk density is known to decrease with increasing content of organic matter (e.g. Balland *et al.*, 2008). Furthermore, there might be a preferential transport of organic matter from the soil surface along macropores. Both higher root production and transport of organic matter might decrease the bulk density by increasing the soil organic matter content in the preferential flow paths.

## 3.1.5 Conclusions

The dye patterns suggest that the organic horizons play an important role in initiating preferential flow in the topsoil. This might be due to the structure of the litter layer or to chemical properties of the organic horizons such as water repellency. A detailed study of this part of the soil could help to understand which mechanisms force the flow to converge to certain localized flow paths. Roots constituted main preferential flow paths and induced macropore flow, especially in the topsoil. In the subsoil root density decreased and inhomogeneous infiltration from preferential flow paths into the soil matrix caused unstable flow.

Root macropores promote preferential transport of solutes from the organic horizons to the subsoil. Furthermore, roots are known to strongly influence their immediate environment, the rhizosphere, by exudation of organic compounds. As a consequence distinct chemical compartments might develop with gradients in the transition zone between the soil matrix and preferential flow paths. Further studies should consider possible impacts of preferential flow on soil chemistry.

### 3.1.6 Acknowledgements

The authors are grateful to the R and S-plus community, especially to Prof. John Fox, Prof. Douglas Bates, Dr. José Pinheiro and Dr. Kyle Roberts for their help with statistical issues and for providing various R-scripts. MVTec Software GmbH has granted a research license for the software HALCON during the first author's PhD. This project was funded by the Deutsche Forschungsgemeinschaft (DFG FOR 562).

### 3.1.7 References

Adobe. 2005. Photoshop ver. CS2. (At: http://www.adobe.com. Accessed: 04/03/2009)

Balland, V., Pollacco, J.A.P. & Arp, P.A. 2008. Modeling soil hydraulic properties for a wide range of soil conditions. *Ecological Modelling*, **219,** 300-316.

Belsley, D.A. 1991. A guide to using the collinearity diagnostics. *Computer Science in Economics and Management*, **4,** 33-50.

Belsley, D.A., Kuh, E. & Welsch, R.E. 1980. *Regression diagnostics: Identifying influential data and sources of collinearity*. Wiley Series in Probability and Mathematical Statistics. Wiley, New York.

Bogner, C., Borken, W. & Huwe, B. 2009. Impact of preferential flow on soil chemistry in a forest soil. *submitted to the European Journal of Soil Science*.

Bogner, C., Engelhardt, S., Zeilinger, J. & Huwe, B. 2008. Visualization and analysis of flow patterns and water flow simulations in disturbed and undisturbed tropical soils. In: *Gradients in a Tropical Mountain Ecosystem of Ecuador* (eds. Beck, E., Bendix, J., Kottke, I., Makeschin, F. & Mosandl, R.), pp. 387-396. Springer Verlag, Berlin Heidelberg.

Box, G.E.P. & Cox, D.R. 1964. An analysis of transformations (with discussion). *Journal of the Royal Statistical Society, Series B*, **26,** 211-252.

Corti, G., Ugolini, F.C., Agnelli, A., Certini, G., Cuniglio, R., Berna, F. & Sanjurjo, M.J.F. 2002. The soil skeleton, a forgotten pool of carbon and nitrogen in soil. *European Journal of Soil Science*, **53,** 283-298.

Edwards, L.J., Muller, K.E., Wolfinger, R.D., Qaqish, B.F. & Schabenberger, O. 2008. An R-2 statistic for fixed effects in the linear mixed model. *Statistics in Medicine*, **27,** 6137-6157.

Flühler, H., Durner, W. & Flury, M. 1996. Lateral solute mixing processes - A key for understanding field-scale transport of water and solutes. *Geoderma*, **70,** 165-183.

Flury, M. & Flühler, H. 1994. Brilliant Blue FCF as a dye tracer for solute transport studies - a toxicological overview. *Journal of Environmental Quality*, **23,** 1108-1112.

Flury, M., Flühler, H., Jury, W.A. & Leuenberger, J. 1994. Susceptibility of soils to preferential flow of water: a field study. *Water Resources Research*, **30,** 1945-1954.

Foken, T. 2003. *Lufthygienisch-Bioklimatische Kennzeichnung des oberen Egertales*. Bayreuther Forum Ökologie. Vol. 100, 69+XLVIII.

Forrer, I.E., Papritz, A., Kasteel, R., Flühler, H. & Luca, D. 2000. Quantifying dye tracers in soil profiles by image processing. *European Journal of Soil Science*, **51,** 313-322.

Gaul, D., Hertel, D. & Leuschner, C. 2008. Effects of experimental soil frost on the fine-root system of mature Norway spruce. *Journal of Plant Nutrition and Soil Science*, **171,** 690-698.

German-Heins, J. & Flury, M. 2000. Sorption of Brilliant Blue FCF in soils as affected by pH and ionic strength. *Geoderma*, **97,** 87-101.

Gerstberger, P., Foken, T. & Kalbitz, K. 2004. The Lehstenbach and Steinkreuz catchments in NE Bavaria, Germany. In: *Biogeochemistry of Forested Catchments in a Changing Environment* (ed. Matzner, E.), pp. 15-44. Springer Verlag, Berlin Heidelberg.

MVTec Software GmbH. 2007a. Halcon ver. 8.0. (At: http://www.mvtec.com/halcon/. Accessed: 04/03/2009)

MVTec Software GmbH. 2007b. Solution Guide I. Basics. Halcon Technical Manual.

MVTec Software GmbH. 2007c. Solution Guide II-F. 3D Machine Vision. Halcon Technical Manual.

Green, W.H. & Ampt, G.A. 1911. Studies on soil physics, 1. The flow of air and water through soils. *The Journal of Agricultural Science*, **4,** 1-24.

Guo, L.B., Halliday, M.J., Siakimotu, S.J.M. & Gifford, R.M. 2005. Fine root production and litter input: Its effects on soil carbon. *Plant and Soil*, **272,** 1-10.

Gurka, M.J., Edwards, L.J., Muller, K.E. & Kupper, L.L. 2006. Extending the Box-Cox transformation to the linear mixed model. *Journal of the Royal Statistical Society Series A-Statistics in Society*, **169,** 273-288.

Heisner, U., Raber, B. & Hildebrand, E.E. 2004. The importance of the soil skeleton for plant-available nutrients in sites of the Southern Black Forest, Germany. *European Journal of Forest Research*, **123,** 249-257.

Hendrickx, J.M.H. & Flury, M. 2001. Uniform and Preferential Flow Mechanisms in the Vadose Zone. In: *Conceptual Models of Flow and Transport in the Fractured Vadose Zone* (ed. Council, N.R.), pp. 149-187. National Academy Press, Washington, DC.

Hendrickx, J.M.H., Khan, A.S., Bannink, M.H., Birch, D. & Kidd, C. 1991. Numerical analysis of groundwater recharge through stony soils using limited data. *Journal of Hydrology*, **127,** 173-192.

Hentschel, K., Borken, W. & Matzner, E. 2008. Repeated freeze-thaw events affect leaching losses of nitrogen and dissolved organic matter in a forest soil. *Journal of Plant Nutrition and Soil Science*, **171,** 699-706.

Hothorn, T., Bretz, F. & Westfall, P. 2008. Simultaneous inference in general parametric models. *Biometrical Journal*, **50,** 346-363.

Jarvis, N.J. 2007. A review of non-equilibrium water flow and solute transport in soil macropores: principles, controlling factors and consequences for water quality. *European Journal of Soil Science*, **58,** 523-546.

Johnson, N.L., Kotz, S. & Balakrishnan, N. 1994. *Continuous univariate distributions*. Vol. 1, 2nd edn., Wiley, New York.

Kasteel, R., Vogel, H.J. & Roth, K. 2002. Effect of non-linear adsorption on the transport behaviour of Brilliant Blue in a field soil. *European Journal of Soil Science*, **53,** 231-240.

Ketelsen, H. & Meyer-Windel, S. 1999. Adsorption of Brilliant Blue FCF by soils. *Geoderma*, **90,** 131-145.

Kulli, B., Stamm, C., Papritz, A. & Fluhler, H. 2003. Discrimination of flow regions on the basis of stained infiltration patterns in soil profiles. *Vadose Zone Journal*, **2,** 338-348.

Laird, N.M. & Ware, J.H. 1982. Random-effects models for longitudinal data. *Biometrics*, **38,** 963-974.

Lawes, J.B., Gilbert, J.H. & Warrington, R. 1882. *On the amount and composition of the rain and drainage water collected at Rothamsted*. Williams, Clowes & Sons, London.

Lipsitz, S.R., Ibrahim, J. & Molenberghs, G. 2000. Using a Box-Cox transformation in the analysis of longitudinal data with incomplete responses. *Journal of the Royal Statistical Society Series C-Applied Statistics*, **49,** 287-296.

Lu, J. & Wu, L. 2003. Visualizing bromide and iodide water tracer in soil profiles by spray methods. *Journal of Environmental Quality*, **32,** 363-367.

Manning, W.G. 1998. The logged dependent variable, heteroscedasticity, and the retransformation problem. *Journal of Health Economics*, **17,** 283-295.

Mon, J., Flury, M. & Harsh, J.B. 2006. Sorption of four triarylmethane dyes in a sandy soil determined by batch and column experiments. *Geoderma*, **133,** 217-224.

Mullahy, J. 1998. Much ado about two: reconsidering retransformation and the two-part model in health econometrics. *Journal of Health Economics*, **17,** 247-281.

Pinheiro, J., Bates, D., DebRoy, S., Sarkar, D. & The R Core team. 2008. nlme: Linear and Nonlinear Mixed Effects Models. R package version 3.1-89.

Pinheiro, J.C. & Bates, D.M. 2000. *Mixed-Effects Models in S and S-Plus*. Springer, New York.

Sauer, T.J. & Logsdon, S.D. 2002. Hydraulic and physical properties of stony soils in a small watershed. *Soil Science Society of America Journal*, **66,** 1947-1956.

Schulin, R., Wierenga, P.J., Flühler, H. & Leuenberger, J. 1987. Solute Transport through a Stony Soil. *Soil Science Society of America Journal*, **51,** 36-42.

Shieh, Y.-Y. & Fouladi, R.T. 2003. The Effect of Multicollinearity on Multilevel Modeling Parameter Estimates and Standard Errors. *Educational and Psychological Measurement*, **63,** 951-985.

Tate, K.R., Ross, D.J., Obrien, B.J. & Kelliher, F.M. 1993. Carbon storage and turnover, and respiratory activity, in the litter and soil of an old-growth southern beech (Nothofagus) forest. *Soil Biology & Biochemistry*, **25,** 1601-1612.

R Development Core Team. 2008. R: A Language and Environment for Statistical Computing. http://www.R-project.org [accessed on February 2009]

The MathWorks, Inc. 2005. Matlab ver. 7.1. (At: http://www.mathworks.com. Accessed: 04/03/2009)

Ugolini, F.C., Corti, G., Agnelli, A. & Piccardi, F. 1996. Mineralogical, physical, and chemical properties of rock fragments in soil. *Soil Science*, **161,** 521-542.

Weiler, M. & Flühler, H. 2004. Inferring flow types from dye patterns in macroporous soils. *Geoderma*, **120,** 137-153.

Westfall, P.H. 1999. *Multiple comparisons and multiple tests*. Cary, NC: SAS Institute Inc.

Wissmann, M. & Toutenburg, H. 2007. *Role of Categorical Variables in Multicollinearity in the Linear Regression Model*. University of Munich. Department of Statistics: Technical Reports, No.8.

IUSS Working Group WRB. 2007. World reference base for soil resources 2006, first update 2007. World Soil Resources Reports No. 103.

## 3.2 Impact of preferential flow on soil chemistry in a forest soil

BOGNER C[a,b], BORKEN W[c], HUWE B[a]

[a]*Soil Physics Group, University of Bayreuth, 95440 Bayreuth, Germany,*

[b]*Ecological Modelling, University of Bayreuth, Dr.-Hans-Frisch-Straße 1-3,*

*95448 Bayreuth, Germany and* [c]*Soil Ecology, University of Bayreuth, 95440*

*Bayreuth, Germany*

Correspondence: C. Bogner. E-mail: christina.bogner@uni-bayreuth.de.

Short title:

*Soil chemistry in preferential flow paths*

Submitted to the *European Journal of Soil Science*

## Summary

Main flow paths in forest soils can be stable for decades favouring the development of distinct chemical properties. At our study site, roots constituted preferential flow paths and induced macropore flow, especially in the topsoil. We did three tracer experiments in a Norway spruce forest soil and analysed soil chemical parameters (exchangeable cations, pH, total C and total N) of preferential flow paths and soil matrix by means of mixed-effects models that take the hierarchical nature of sampled data into account. We found smaller pH values (0.11-0.34 pH units), 32% more Ca, 51% more Mg, 11-30% more C and 13-24% more N in preferential flow paths. We attribute the increase of Ca and Mg to their transport via preferential flow paths after the application of lime some years ago. Compared to the adjacent soil matrix, 27% more aluminium and 192% more iron (but small absolute amounts) were found in the horizon Bw/C (subsoil) where macropore flow along root channels decreases and heterogeneous matrix flow dominates. These distinct chemical properties are probably due to root exudates and translocation of solutes and DOM along preferential paths. Our results show that preferential flow has considerable effects on chemical processes in the subsoil and could promote C sequestration in subsoil horizons.

### 3.2.1 Introduction

Preferential flow describes all phenomena where water flows through localised pathways bypassing a portion of the soil matrix (Hendrickx & Flury, 2001). Main transport mechanisms along preferential flow paths are inhomogeneous matrix flow and macropore flow. The first mentioned usually occurs in coarse-textured soils induced by, for instance, textural differences or water repellency. Macorpore flow is typical for highly structured soil or soils containing biopores like root channels or earthworm burrows.

Several studies showed that predominant flow paths could be stable for decades. Ritsema & Dekker (2000) observed that preferential flow paths recurred at the same locations during successive rain events in a water repellent sandy soil. They stated that these preferential flow paths might remain stable for an unlimited period of time, except that human influence (e.g. tillage) might change the pattern. Although tillage destroys stable continuous macropores, macropores burrier

below the tilled layer could still function as preferential flow paths (Logsdon, 1995). Hagedorn & Bundt (2002) estimated an age of about 40 years for preferential flow paths in a forest soil by measuring the distribution of radionuclides.

Due to temporal stability and reoccurrence of preferential flow at the same location, distinct biological and chemical environments in preferential flow paths and soil matrix are likely to develop. Bundt *et al*. (2001b) found higher cation exchange capacity, base saturation, organic C and organic N concentrations in preferential flow paths of a forest soil. There are indications that soil organic matter in preferential flow paths is younger and N cycling faster than in the soil matrix (Bundt *et al.*, 2001a). This is in accordance with Hagedorn *et al*. (1999) who reported an increased N transformation in preferential flow paths compared with the soil matrix. Along flow paths nitrate concentrations were higher; denitrification activity after rainfalls increased and net nitrification started earlier after drying and was enhanced.

In our study we first identified main flow mechanisms in a spruce forest by mixed-effects modelling (Bogner *et al.*, 2009a). We found that roots constituted main preferential flow paths and induced macropore flow, especially in the topsoil. In the subsoil, root density decreased and inhomogeneous infiltration from preferential flow paths into the soil matrix caused unstable matrix flow. Based on these results, the aim of this study is to determine how soil chemical properties (concentration of exchangeable cations, pH, total C and N) differ between preferential flow paths and soil matrix. Special emphasis will be given to the hierarchical nature of sampled data. Most studies that investigate differences between preferential flow paths and soil matrix use the paired t-test of mean values or its non-parametric equivalent and test different depths separately. We propose to employ mixed-effects models and to consider all plots and all depths in one single analysis.

## 3.2.2 Materials and methods

### *Field and laboratory work*

We did three tracer experiments with Brilliant Blue FCF in a Norway spruce forest situated in the Fichtelgebirge (770 m above see level) in southeast Germany. The soil is classified as Haplic Podzol (IUSS Working Group WRB,

2006) with five mineral soil horizons (Ea, Bsh, Bs, Bw and Bw/C) with a sandy to loamy texture and a 6-10 cm thick mor-type organic horizons (Oi, Oe and Oa). The pH ($H_2O$) ranges between 4.0 (Oa) and 4.5 (Bw). The C and N content of the Oa horizon is 18% and 1% respectively and decreases with depth to 0.4% C and < 0.05% N in Bw. The soil was limed in the past as indicated by the high base saturation especially in the organic horizons (Hentschel *et al.*, 2008).

Three profiles per plot were sampled for analysis of exchangeable cations (H, K, Na, Ca, Mg, Mn, Al, and Fe), content of total C and N, pH ($H_2O$) and soil texture. We took bulk samples from Brilliant Blue stained and non-stained parts per soil horizon. Some thin horizons were hardly stained and could not be sampled. Therefore, the number of samples differed from plot to plot resulting in 58 samples altogether. A detailed description of the tracer experiment and soil sampling is given in Bogner *et al*. (2009a).

Exchangeable cations were extracted with a 1M $NH_4Cl$ solution and analyzed by ICP-OES (**I**nductively **C**oupled **P**lasma **O**ptical **E**mission **S**pectrometry) (Jobin-Yvon Horiba Group, JY2000, U.S.A.). Total C and N in soil samples were determined using a CNS-Analyser (Vario EL, Elementar). As pH ($H_2O$) values of the soil ranged between 5.0 and 3.3, total carbon is equivalent to organic carbon. Brilliant Blue FCF ($C_{37}H_{34}N_2Na_2O_9S_3$, molar mass 792.9 g $mol^{-1}$) is an organic molecule consisting of 56% C and 4% N and its sorption on soil particles affects C and N contents of soil. We determined the content of Brilliant Blue in soil by visible diffuse reflectance spectroscopy (Bogner *et al.*, 2009b). The predicted concentrations ranged from -1.0 mg $g^{-1}$ to 5.1 mg $g^{-1}$. After transformation to C and N contents the positive predicted values were subtracted from the total C content and total N content, respectively. Samples with negative predicted values were not corrected. The maximum predicted value for additional carbon and nitrogen due to Brilliant Blue was 0.29% and 0.022%, respectively. Analysis of soil texture, root abundance and bulk density are presented by Bogner *et al*. (2009a).

### *Initial data analysis*

The tracer solution contained K and Na so that at least these compounds are influenced by the experiment and their measured concentrations do not reflect the natural background in soil. To find further possible disturbances due to tracer

application we performed a principal component analysis (PCA) on the log-transformed data set of exchangeable cations and pH. The latter was transformed to concentration of hydronium per g soil. The log-transformation was necessary as the PCA is sensitive to non-normality (especially in small data sets). In two samples Mg concentrations were below the detection limit of the ICP-OES and were set to 0.5·detection limit to avoid zeros. The same procedure was applied to eight other samples for Fe concentrations. C and N were not included in the PCA as they were corrected for the content of Brilliant Blue.

### *Mixed-effects models*

In the present study we have to deal with missing values due to bypassing of some horizons and with hierarchically sampled data that are grouped at three levels: plot, profile and horizon. Hierarchical sampling might induce dependencies in the data and therefore classical statistical techniques like the analysis of variance are not applicable. We used mixed-effects models that explicitly accounts for possible dependencies in data. They contain fixed-effects representing parameters of the entire population or certain repeatable levels of experimental factors and random-effects associated with individual experimental units drawn at random from a population. Furthermore, they are robust against missing values (Pinheiro & Bates, 2000).

We analysed exchangeable cations (H, Ca, Mg, Al and Fe), pH and corrected total C and N contents. As predictors we used the flow region because we were interested in differences between preferential flow paths and the soil matrix. This factor has two levels – 'soil matrix' and 'preferential flow paths'– indicating the origin of samples. Furthermore, we included soil horizon combining Bw and Bw/C into one level because they were not significantly different. So soil horizon is a factor with four levels: 'Ea', 'Bsh', 'Bs' and 'Bw/C'. For Ca, Mg, Al, C and N we added the content of silt – a continuous covariate – as predictor. As the concentration of H, Ca, Mg, Fe and pH were not normally distributed, the data were log-transformed. For details on model building strategy, variable transformation, diagnostic plots and goodness-of-fit statistics the reader is referred to Bogner *et al*. (2009a).

### 3.2.3 Results

*Principal component analysis*

We retained the first three principal axis based on the broken stick model (Frontier, 1976). It states that if the total variance in a multivariate data set was distributed at random across principal components, then the proportion of explained variance $b_k$ associated with the eigenvalue of the $k$th component would be:

$$b_k = \frac{1}{p}\sum_{i=k}^{i=p}\frac{1}{i}, \tag{3.2.1}$$

where $p$ is the number of principal components. If the eigenvalue of the $k$th principal component is larger than $b_k$, then it is non-trivial and should be retained. The percentage of variance explained by the first three principal components was 40.2 %, 25.7% and 15.1 % respectively. The first principal axis has negative scores for the topsoil (Ea, Bsh and Bs) and positive ones for the subsoil (Bw and Bw/C). Especially matrix samples from the subsoil form a distinct cluster in the left lower corner of the biplot (Figure 3.2.1 left).

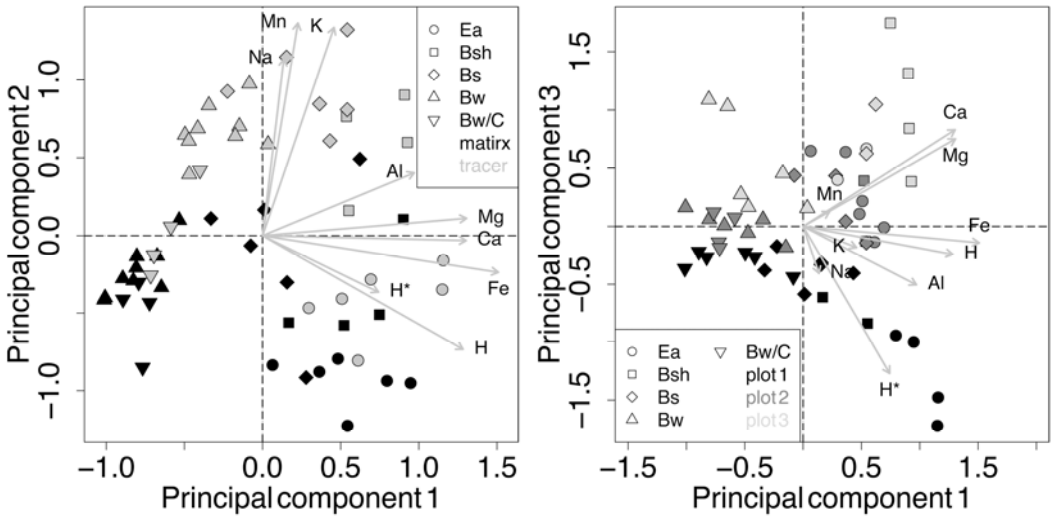



Figure 3.2.1: Biplots of the PCA with a scatter plot of component scores and eigenvector loadings as arrows. The variable H* denotes pH values transformed to hydronium concentrations per g soil.

The second principal component separates samples from preferential flow paths from matrix samples. K and Na were applied with the tracer and show high loadings on the second axis. Mn concentration was increased by the tracer, probably because it was reduced by iodide also present in the tracer solution. Therefore, we excluded K, Na and Mn from further analysis as they have been

influenced by the tracer application. The third principal component separates the three experimental plots (Figure 3.2.1l right). So the highest variation in the data is due to the depth gradient, followed by tracer influence and differences between plots.

## *Modelling results*

### Hydronium

Hydronium concentrations decreased with increasing depths and varied between the same horizons at different plots and between profiles within plots (Figure 3.2.2) suggesting a profile within plot and depth within plot random-effects:

$$\log(y_{ijkl}) = \beta_0 + \beta_1 \cdot horizon + \beta_2 \cdot flow\ region + b_{ij} + b_{ik} + \varepsilon_{ijkl},$$
$$i=1,\dots,3, \quad j=1,\dots,3, \quad k=1,\dots,4, \quad l=1,\dots,n_{ijk} \qquad (3.2.2)$$
$$b_{ij} \sim N(0,\sigma_1^2), \quad b_{ik} \sim N(0,\sigma_2^2), \quad \varepsilon_{ijkl} \sim N(0, \sigma^2\delta^2_k),$$

where $y_{ijkl}$ is the concentration of hydronium in mmol $kg^{-1}$, $\beta_i$ are the fixed-effects, $b_{ij}$ is the profile within plot random-effect, $b_{ik}$ is the depth within plot random-effect and $\varepsilon_{ijkl}$ the error term and $n_{ijk}$ the number of observation on plot $i$ in horizon $j$ and depth $k$. Despite the log-transformation the error variance was not constant and was allowed to differ between horizons.

Model (2.2.2) fits the data well (Figure 3.2.3a); the within-group errors and the random-effect are well estimated (Table 3.2.1). The variance explained by fixed-effects equals 79% and the correlation between the transformed observed values and the fitted ones is 0.95. There is no significant difference in hydronium concentrations between the soil matrix and preferential flow paths (Figure 3.2.4a).

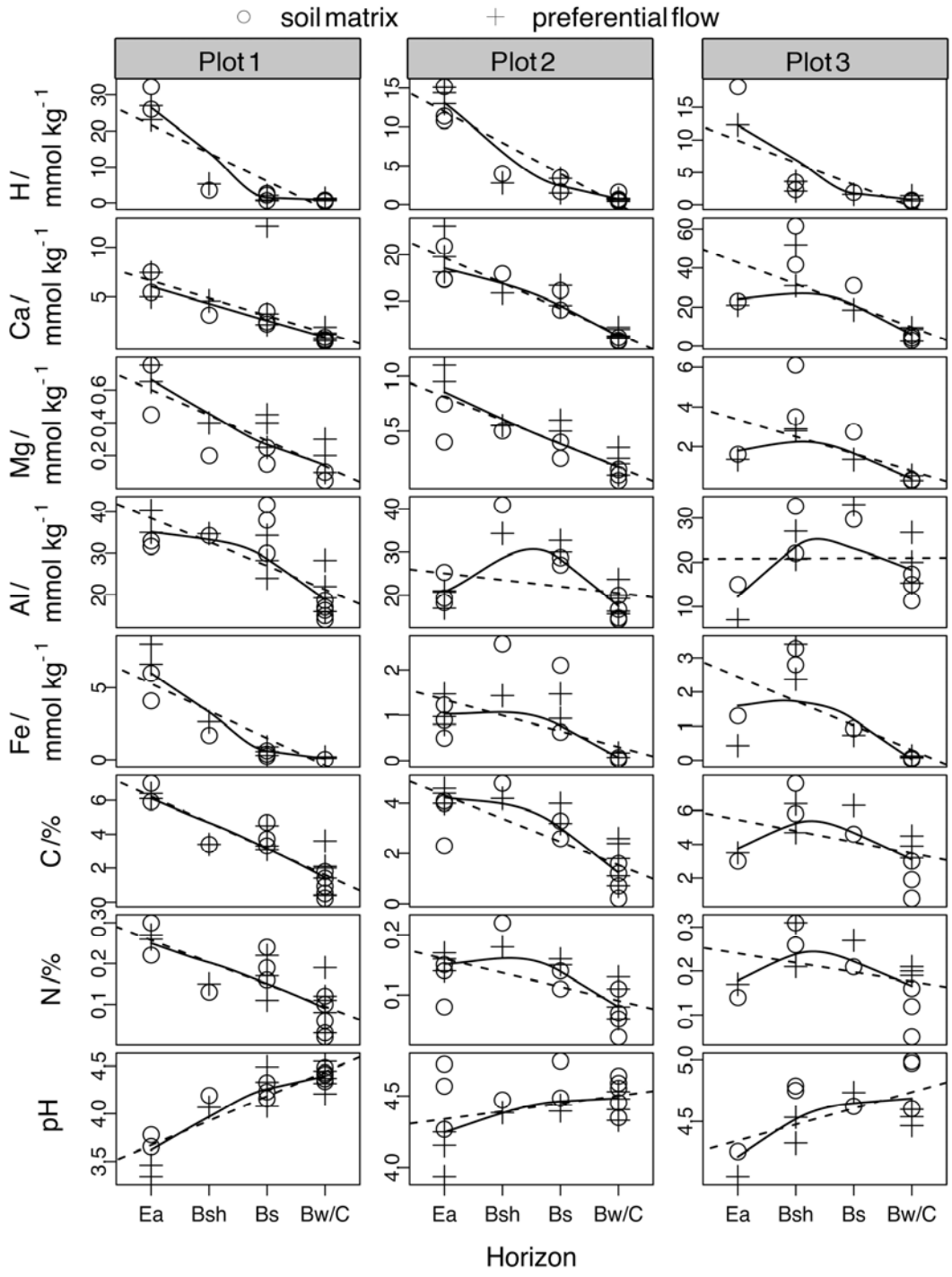



Figure 3.2.2: Variation of exchangeable cations, total C, total N and pH with depth. The solid line is a smooth curve computed by loess (with smoothness parameter 1.01) and the dotted line shows a linear regression.

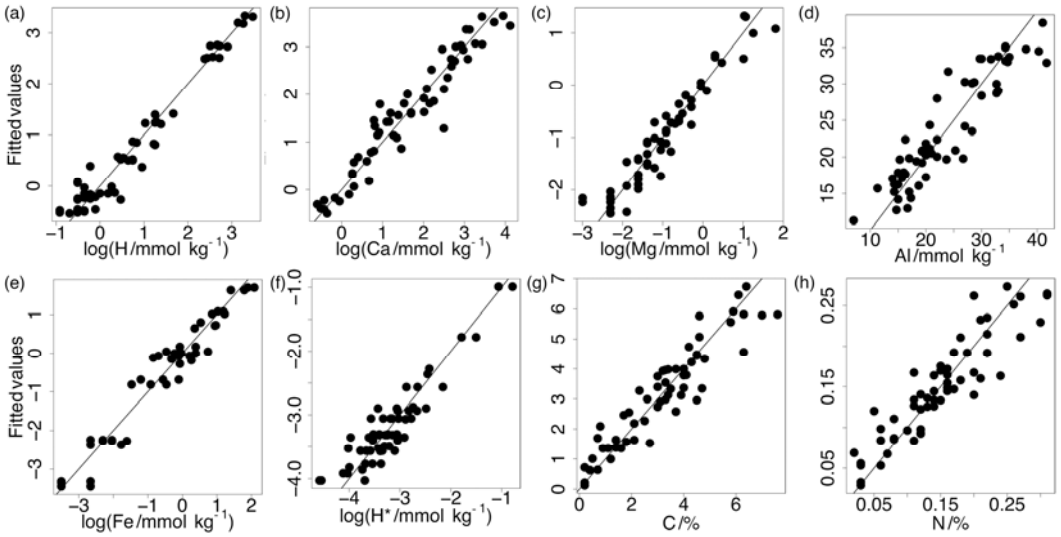



Figure 3.2.3: Fitted versus observed values of the (a) hydronium, (b) calcium, (c) magnesium, (d) aluminium, (e) iron, (f) pH, (g) carbon and (h) nitrogen models; (a straight line with slope 1 was added for better visibility).

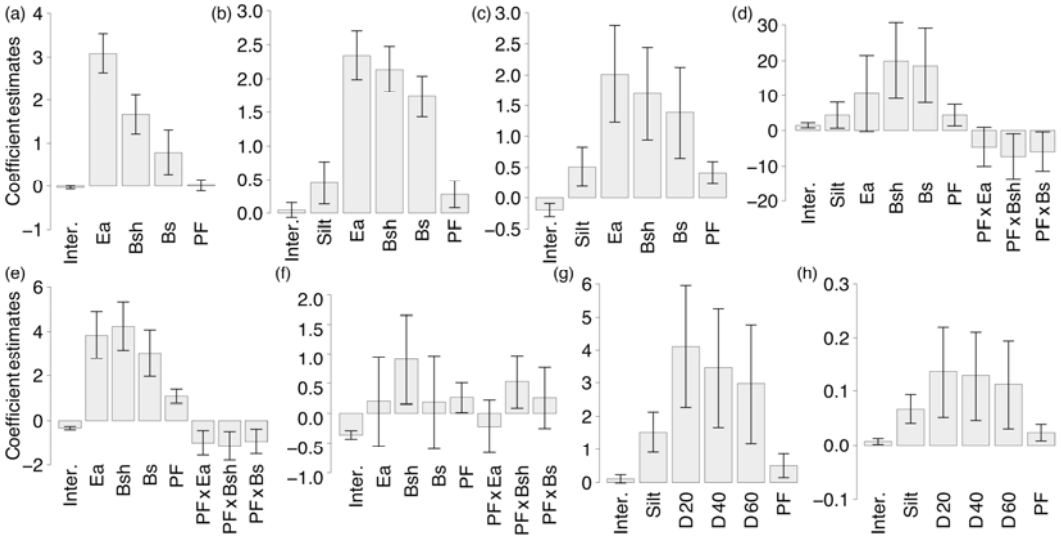



Figure 3.2.4: Estimated fixed-effects effects with standard errors (Inter. = intercept, PF = preferential flow paths) of the (a) hydronium, (b) calcium, (c) magnesium, (d) aluminium, (e) iron, (f) pH, (g) carbon and (h) nitrogen models.

Table 3.2.1: Parameter estimates and approximate 95% confidence intervals for random-effects and within-group errors.

| | 95% lower | Estimate | 95% upper |
|---|---|---|---|
| Hydronium /mmol $kg^{-1}$ | | | |
| Random-effect (profile) | 0.1 | 0.2 | 0.5 |
| Random-effect (horizon) | 0.1 | 0.2 | 0.5 |
| Within-group error (Ea and Bsh) | 0.1 | 0.2 | 0.3 |
| Within-group error (Bs) | 0.2 | 0.4 | 0.9 |
| Within-group error (Bw/C) | 0.2 | 0.3 | 0.5 |
| Magnesium /mmol $kg^{-1}$ | | | |
| Random-effect (plot) | 0.3 | 0.9 | 2.4 |
| Random-effect (horizon) | 0.2 | 0.34 | 0.7 |
| Within-group error | 0.3 | 0.3 | 0.4 |

| | 95% lower | Estimate | 95% upper |
|---|---|---|---|
| Calcium /mmol $kg^{-1}$ | | | |
| Random-effect (plot) | 0.3 | 0.9 | 2.6 |
| Within-group error | 0.3 | 0.4 | 0.5 |
| Aluminium /mmol $kg^{-1}$ | | | |
| Random-effect (horizon) | 3.3 | 5.9 | 10.4 |
| Within-group error | 3.1 | 3.9 | 4.8 |

| Iron /mmol $kg^{-1}$ | | | |
|---|---|---|---|
| Random-effect (horizon) | 0.3 | 0.6 | 1.0 |
| Within-group error | 0.3 | 0.4 | 0.5 |

| H*[a] /mmol $kg^{-1}$ | | | |
|---|---|---|---|
| Random-effect (plot) | 0.2 | 0.5 | 1.5 |
| Random-effect (horizon) | 0.2 | 0.3 | 0.6 |
| Within-group error | 0.3 | 0.3 | 0.4 |

| Carbon /% | | | |
|---|---|---|---|
| Random-effect (horizon) | 0.5 | 1.0 | 1.9 |
| Within-group error (D 5, D 20, D 60) | 0.5 | 0.6 | 0.8 |
| Within-group error (D 40) | 0.9 | 1.1 | 1.4 |

| Nitrogen /% | | | |
|---|---|---|---|
| Random-effect (horizon) | 0.03 | 0.05 | 0.09 |
| Within-group error (D 5, D 40, D 60) | 0.03 | 0.04 | 0.05 |
| Within-group error (D 20) | 0.01 | 0.02 | 0.03 |

[a]Concentration of hydronium per g soil calculated from $pH_{H2O}$

Calcium

Figure 3.2.2 shows decreasing calcium concentrations with depth and a high variability in mean values between the tree plots which suggests a plot random-effect:

$$\log(y_{ij}) = \beta_0 + \beta_1 \cdot silt + \beta_2 \cdot horizon + \beta_3 \cdot flow\ region + b_i + \varepsilon_{ij},$$
$$i=1, 2, 3, \qquad j=1,\ldots,n_i, \qquad (3.2.3)$$
$$b_i \sim N(0,\sigma_1^2), \qquad \varepsilon_{ij} \sim N(0, \sigma^2),$$

where $y_{ij}$ is the concentration of calcium in mmol $kg^{-1}$ soil, $\beta_i$ are the fixed-effects, $b_i$ is the plot random-effect, $\varepsilon_{ij}$ the error term and $n_i$ the number of observations on plot $i$. Figure 3.2.3b indicates that the model fits the data well and Figure 3.2.4b shows the estimated fixed-effects and their standard errors. The explained variance equals 87% and the correlation between fitted and transformed values is 0.92.

Using the convenient log-model has the advantage that we can calculate the ratio of concentrations in the soil matrix and those in preferential flow paths directly (Johnson *et al.*, 1994; Manning, 1998; Mullahy, 1998). As the residuals are normally distributed and homoscedastic, the ratio equals exp(*PF*), where *PF* is the effect of flow region in preferential flow paths (Figure 3.2.4). We found 32% more calcium in preferential flow paths than in soil matrix.

Magnesium

Similar to calcium concentrations of magnesium decreased with increasing depth and the variability of the data suggests a plot and a depth within plot random-effects (Figure 3.2.2):

$$\log(y_{ijk}) = \beta_0 + \beta_1 \cdot silt + \beta_2 \cdot horizon + \beta_3 \cdot flow\ region + b_i + b_{ij} + \varepsilon_{ijk},$$
$$i=1,2,3, \qquad j=1,\ldots 4, \qquad k=1,\ldots,n_{ij}, \qquad (3.2.4)$$
$$b_i \sim N(0,\sigma_1^2), \qquad b_{ij} \sim N(0, \sigma_2^2), \qquad \varepsilon_{ijk} \sim N(0, \sigma^2),$$

where $y_{ijk}$ is the concentration of magnesium in mmol $kg^{-1}$ soil, $\beta_i$ are the fixed-effects, $b_i$ is the plot random-effect, $b_{ij}$ the horizon within plot random-effect, $\varepsilon_{ijk}$ the error term and $n_{ij}$ the number of observation on plot $i$ in horizon $j$. The correlation between the fitted and the observed values is 0.92 (Figure 3.2.3c) and the explained variance 57%. Fixed-effects estimates and their standard errors are

shown in Figure 3.2.4c. We found 51% more magnesium in preferential flow paths than in soil matrix.

Aluminium

There is a high variability between the same horizons on different plots so we incorporated a horizon within plot random-effect (Figure 3.2.2). Concentrations of aluminium were not transformed, but modelled directly by:

$$y_{ijk} = \beta_0 + \beta_1 \cdot silt + \beta_2 \cdot horizon + \beta_3 \cdot flow\ region + \beta_4 \cdot horizon \cdot flow\ region + b_{ij} + \varepsilon_{ijk},$$

$$i=1,2,3, \quad j=1,2,3, \quad k=1,..,4, \quad l=1,\ldots,n_{ij}, \quad b_{ij} \sim N(0, \sigma_1^2), \quad \varepsilon_{ijk} \sim N(0, \sigma^2), \qquad (3.2.5)$$

where $y_{ijk}$ is the concentration of aluminium in mmol $kg^{-1}$ soil, $\beta_i$ are the fixed-effects, $b_{ij}$ is the horizon within plot random-effect, $\varepsilon_{ijk}$ the error term and $n_{ij}$ the number of observation on plot $i$ in horizon $j$. Model (3.2.5) explains 37% of the variance and the correlation between observed and fitted values is 0.84 (Figure 3.2.3d). There are 4.4 mmol more aluminium per kg soil (27%) in preferential flow paths of the horizon Bw/C. Figure 3.2.4d indicates that the interaction term 'PF×Ea' is not significant ($p = 0.099$) so there might be more aluminium in preferential flow paths than in soil matrix in the horizon Ea as well. But refitting the model with Ea as origin level clearly indicates no significant differences between preferential flow paths and soil matrix in this horizon. This problem is possibly due to a small number of observations and a high variability in horizon Ea.

Iron

Figure 3.2.2 illustrates the high variability of iron concentrations between horizons within plots so a horizon within plot random-effect might be appropriate:

$$\log(y_{ijk}) = \beta_0 + \beta_1 \cdot horizon + \beta_2 \cdot flow\ region + \beta_3 \cdot horizon \cdot flow\ region + b_{ij} + \varepsilon_{ijk},$$

$$i=1,2,3, \quad j=1,2,3, \quad k=1,\ldots,n_{ij}, \quad b_{ij} \sim N(0,\sigma_1^2), \quad \varepsilon_{ijk} \sim N(0, \sigma^2), \qquad (3.2.6)$$

where $y_{ijk}$ is the concentration of iron in mmol $kg^{-1}$ soil, $\beta_i$ are the fixed-effects, $b_{ij}$ is the horizon within plot random-effect, $\varepsilon_{ijk}$ the error term and $n_{ij}$ the number of observation on plot $i$ in horizon $j$. Model (3.2.6) explains 69% of the variance in the data and gives a correlation of 0.96 between fitted and transformed values

(Figure 3.2.3e). The estimated fixed-effects are shown in Figure 3.2.4e. There is 192% more iron in the preferential flow paths in the lower soil horizon Bw/C.

pH

pH values increased with decreasing depth and the highest variability in the data was between plots and horizons within plots. Prior to modelling pH values were converted into hydronium concentrations:

$$\log(y_{ijk}) = \beta_0 + \beta_1 \cdot horizon + \beta_2 \cdot flow\ region + \beta_3 \cdot horizon \cdot flow\ region + b_i + b_{ij} + \varepsilon_{ijk},$$

$$i=1,2,3, \qquad j=1,\ldots 4, \qquad k=1,\ldots,n_{ij}, \qquad (3.2.7)$$

$$b_i \sim N(0,\sigma_1^2), \qquad b_{ij} \sim N(0, \sigma_2^2), \qquad \varepsilon_{ijk} \sim N(0, \sigma^2),$$

where $y_{ijk}$ is the concentration of hydronium in mmol $kg^{-1}$ soil (denoted H*), $\beta_i$ are the fixed-effects, $b_i$ is the plot random-effect, $b_{ij}$ the horizon within plot random-effect, $\varepsilon_{ijk}$ the error term and $n_{ij}$ the number of observation on plot $i$ in horizon $j$. The model explains 49% of variance and the correlation between fitted and transformed values is 0.86 (Figure 3.2.3f). The pH values are between 0.11 (Bsh, Bs and Bw/C) and 0.34 units (Ea) lower in preferential flow paths than in soil matrix.

Carbon and nitrogen

The content of total carbon and total nitrogen was globally decreasing with depth with the highest variation between horizons within plots. Transformation was not necessary and we used the model:

$$y_{ijk} = \beta_0 + \beta_1 \cdot silt + \beta_2 \cdot depth + \beta_3 \cdot flow\ region + b_{ij} + \varepsilon_{ijk},$$

$$i=1,2,3, \qquad j=1,\ldots 4, \qquad k=1,\ldots,n_{ij}, \qquad (3.2.8)$$

$$b_{ij} \sim N(0, \sigma_1^2), \qquad \varepsilon_{ijk} \sim N(0, \sigma^2\delta^2_j),$$

where $y_{ijk}$ is the content of carbon or nitrogen in percent, $\beta_i$ are the fixed-effects, $b_{ij}$ random-effect for the horizon within plot, $\varepsilon_{ijk}$ the error term and $n_{ij}$ the number of observation on plot $i$ in horizon $j$. The error variance was not constant and was allowed to vary between horizons. Model (3.2.8) explains 49% and 48% of variance of total carbon and total nitrogen, respectively. The correlation between fitted and observed values is 0.87 (Figure 3.2.3g) and 0.84 (Figure 3.2.3h), respectively. The estimated fixed-effects indicate 11% (Ea and Bsh), 13% (Bs) and 30% (horizon Bw/C) more carbon and 13% (Ea and Bs), 12% (Bsh) and 24%

(horizon Bw/C) more nitrogen in preferential flow paths (Figure 3.2.4g and Figure 3.2.4h) for a mean content of silt. Correcting the content of C and N for presence of Brilliant Blue diminished the differences between preferential flow and soil matrix by about 3% for C and 5% for N.

## 3.2.4 Discussion

### *Proportion of preferential flow domain*

In our study, we defined preferential flow paths as dye stained areas. Predominant flow mechanisms were identified by Bogner *et al*. (2009a) based on stained patterns and statistical analysis. Accordingly, roots constituted main preferential flow paths and induced macropore flow, especially in the topsoil. In the subsoil, root density decreased and inhomogeneous infiltration from preferential flow paths into the soil matrix caused unstable flow. Due to the large sand content (i.e. high permeability) there were strong interactions between preferential flow paths and soil matrix resulting in large stained objects. The amount of stained area per depth (dye coverage) varied among the plots. Based on the mean dye coverage, the vol% of preferential flow were 8-25% in the Ea, 14-30% in the Bsh, 31-41% in the Bs, 31-76% in the Bw and 9-14% in the Bw/C horizons. These high proportions of stained area show the importance of preferential flow at this study site. Based on experimental evidence from other studies (Ritsema & Dekker, 2000; Hagedorn & Bundt, 2002) we postulate that preferential flow paths remain stable down to the subsoil at our study site. Indeed, as roots dominate the flow system in the topsoil, the upper boundary condition for underlying horizons remains unchanged for a long time (Hagedorn & Bundt, 2002). Therefore, the infiltration into the soil matrix in subsoil horizons is likely to occur at the same location.

### *Chemical properties of preferential flow paths*

Preferential flow paths have lower pH values and a higher content of calcium and magnesium through the whole soil profile. Additionally, higher contents of aluminium and iron are found in the subsoil where macropore flow along root channels decreases and heterogeneous matrix flow dominates. Though in this part of the soil absolute concentration of iron are low. There are hints that aluminium

content in preferential flow paths in the horizons Bsh and Bs is lower, but we need more data to verify.

These distinct chemical properties can be explained by root activity and translocation of solutes via preferential flow paths. Higher calcium and magnesium concentration in preferential flow paths are probably due to transport from the soil surface after liming. This is in agreement with Jardine *et al*. (1989) who investigated the transport of organic and inorganic natural tracers (magnesium, ammonium, bromide, nitrate and dissolved organic carbon (DOC)) in a forest soil, but at saturated flow conditions. They observed skewed break through curves indicating preferential flow and concluded that a significant quantity of magnesium, ammonium, and DOC moved through soil via preferential flow paths. The fact that at our study site increased concentrations of basic cations are found in preferential flow paths down to greater depth shows that soil amelioration measures like liming affect not only the topsoil, but could have effects on subsoil horizons.

Higher aluminium and iron concentration in the subsoil could be explained by release and translocation of these solutes during podzolisation. Thereby aluminium and iron are mobilised in the Ea horizon, eluviated downwards and sorbed mainly in the Bsh and Bs horizons. Along preferential flow paths these cations can be transported deeper to the horizon Bw/C. A study at the same site by Hentschel et al. (2008) mentioned the lowest pH values in organic horizons. Preferential transport of acid soil solution along root channels from the organic horizons could be responsible for lower pH values and possibly enhance podzolisation along preferential flow paths.

Rhizodeposition of organic compounds, decomposition of dead roots and transport of DOC from organic to mineral horizons are major sources of organic C input to the soil (e.g. Kuzyakov & Komansky (2000)). Higher root densities in preferential flow paths (see Bogner *et al*. (2009a) for details) lead to a higher C input through roots but also facilitate preferential transport of DOC. Indeed, there is strong experimental indication of transport of DOC via preferential flow paths at our study site. Schulze et al. (2009) investigated the dynamics of the radiocarbon signature of DOC in throughfall and soil solution beneath the Oa horizon and at 90 cm depth (horizon Bw/C). They found that DOC below organic horizons originated mainly from the Oa horizon. In contrast, DOC signature in 90

cm depth could not be explained by the Bs and Bw horizons as the only sources. Thus, preferential transport of DOC from Bsh and maybe from the Ea and Oa horizons is likely. DOC is strongly adsorbed in soils by Al and Fe oxides/hydroxides and clay minerals, particularly if the surface of the adsorbants has low pre-existing levels of adsorbed C like in subsoil horizons (Kalbitz *et al.*, 2000). During transport along preferential flow paths contact time between DOC and soil is reduced so that DOC is transported to greater depth where it potentially form organo-mineral associations. If this holds true, preferential flow is a mechanism that promotes C sequestration in subsoil.

### 3.2.5 Conclusions

At our study site, preferential flow paths constitute a chemical environment with properties that are distinct form those encountered in the soil matrix. Preferential flow does not only influence its immediate environment around paths, but also underlying horizons. Especially the impact on subsoil by preferential transport of basic cations and probably DOC has to be emphasized. Further studies should consider isotopic signatures of soil matrix and preferential flow paths separately to elucidate the relative contribution of root exudates and transported DOC to soil organic matter in preferential flow paths. Furthermore, as preferential flow transport oxygen rich soil solution from the topsoil to the subsoil the role of preferential flow in gas exchange process should be considered.

The measured chemical properties showed a high heterogeneity between plots and within horizons. A higher spatial resolution of measurements could reveal more details. A promising technique to acquire data in a higher spatial resolution is portable VIS-NIR diffuse reflectance spectroscopy. Its use for further tracer studies could address the question whether differences between preferential flow paths and soil matrix detected on bulk samples are the same on the scale of some $cm^2$.

### 3.2.6 Acknowledgements

We thank MVTec for providing a research-license for the software HALCON. This project was funded by the Deutsche Forschungsgemeinschaft (DFG FOR 562).

### 3.2.7 References

Bogner, C., Gaul, D., Kolb, A., Schmiedinger, I. & Huwe, B. 2009a. Investigating flow mechanisms in a forest soil by mixed-effects modelling. submitted to the European Journal of Soil Science.

Bogner, C., Schmiedinger, I. & Huwe, B. 2009b. Rapid prediction of Brilliant Blue concentrations in soil by diffuse reflectance spectroscopy. submitted to the European Journal of Soil Science.

Bundt, M., Jaggi, M., Blaser, P., Siegwolf, R. & Hagedorn, F. 2001a. Carbon and nitrogen dynamics in preferential flow paths and matrix of a forest soil. Soil Science Society of America Journal, 65, 1529-1538.

Bundt, M., Widmer, F., Pesaro, M., Zeyer, J. & Blaser, P. 2001b. Preferential flow paths: biological 'hot spots' in soils. Soil Biology & Biochemistry, 33, 729-738.

Frontier, S. 1976. Étude de la décroissance des valeurs propres dans une analyse en composantes principales: Comparaison avec le modèle du bâton brisé. Journal of Experimental Marine Biology and Ecology, 25, 67-75.

Hagedorn, F. & Bundt, M. 2002. The age of preferential flow paths. Geoderma, 108, 119-132.

Hagedorn, F., Mohn, J., Schleppi, P. & Flühler, H. 1999. The role of rapid flow paths for nitrogen transformation in a forest soil: A field study with micro suction cups. Soil Science Society of America Journal, 63, 1915-1923.

Hendrickx, J.M.H. & Flury, M. 2001. Uniform and Preferential Flow Mechanisms in the Vadose Zone. In: Conceptual Models of Flow and Transport in the Fractured Vadose Zone (ed. Council, N.R.), pp. 149-187. National Academy Press, Washington, DC.

Hentschel, K., Borken, W. & Matzner, E. 2008. Repeated freeze-thaw events affect leaching losses of nitrogen and dissolved organic matter in a forest soil. Journal of Plant Nutrition and Soil Science, 171, 699-706.

Jardine, P.M., Wilson, G.V., Luxmoore, R.J. & McCarthy, J.F. 1989. Transport of inorganic and natural organic tracers through an isolated pedon in a forest watershed. Soil Science Society of America Journal, 53, 317-323.

Johnson, N.L., Kotz, S. & Balakrishnan, N. 1994. Continuous univariate distributions. Vol. 1, 2nd edn., Wiley, New York.

Kalbitz, K., Solinger, S., Park, J.H., Michalzik, B. & Matzner, E. 2000. Controls on the dynamics of dissolved organic matter in soils: A review. Soil Science, 165, 277-304.

Kuzyakov, Y. & Domanski, G. 2000. Carbon input by plants into the soil. Review. Journal of Plant Nutrition and Soil Science - Zeitschrift für Pflanzenernährung und Bodenkunde, 163, 421-431.

Logsdon, S.D. 1995. Flow mechanisms through continuous and buried macropores. Soil Science, 160, 237-242.

Manning, W.G. 1998. The logged dependent variable, heteroscedasticity, and the retransformation problem. Journal of Health Economics, 17, 283-295.

Mullahy, J. 1998. Much ado about two: reconsidering retransformation and the two-part model in health econometrics. Journal of Health Economics, 17, 247-281.

Pinheiro, J.C. & Bates, D.M. 2000. Mixed-Effects Models in S and S-Plus. Springer, New York.

Ritsema, C.J. & Dekker, L.W. 2000. Preferential flow in water repellent sandy soils: principles and modeling implications. Journal of Hydrology, 231, 308-319.

Schulze, K., Borken, W. & Matzner, E. 2009. Dynamics of dissolved organic 14C signature in throughfall and soil solution of a Norway spruce forest. submitted to Biogeochemistry.

WRB, I.W.G. 2006. World reference base for soil resources 2006. World Soil Resources Reports No. 103. FAO, Rome.

# Chapter 4

# General conclusions

## *4.1 Flow pattern analysis by extreme value statistics*

This thesis began with the idea to summarize the main information from dye tracer experiments in one real-valued index, the risk index for vulnerability of groundwater to pollutants developed by Schlather & Huwe (2005). Based on our experimental data we modified the interpretation of the risk index and used it as a risk index for vertical solute propagation rather than for groundwater vulnerability. We could partly confirm the hypothesis that the risk index was to some degree invariant against small changes in boundary conditions. But this parameter alone is not sufficient to determine the risk of vertical solute propagation and we combined it with the second parameter of the generalised Pareto distribution to obtain a probability function of vertical solute propagation.

We could successfully apply the risk index to different soils under varying boundary conditions and characterise the risk of vertical displacement of solutes. The main difficulty in adjusting the parameters of the generalised Pareto distribution arises when the dye coverage function fluctuates or does not decrease monotonically. This might be due to tortuosity of paths, varying flow mechanism or changing physical properties (stratification). The opposite might also be true: simple shaped monotonically decreasing dye coverage functions are usually the result of homogenous flow processes (not necessarily uniform matrix flow) and low tortuosity of flow paths. Thus, in stratified soil, we restricted the analysis to the lowest part of the profile. Since the theory of the risk index is based on extreme values of vertical solute propagation it is the lowest part of the profile that is the most interesting.

Despite a certain resistance to changes of experimental conditions, the risk index is not an intrinsic soil parameter. Since the flow regime in the same soil can be dominated either by preferential flow or by uniform matrix flow (as at the

French site, for instance), the risk of vertical solute propagation will change. It is a physical reality and not a default in the risk index theory. The adjusted parameters of the Pareto distribution will capture the relevant flow regime as reflected by the employed tracer, i.e. they are sensitive enough to reflect differences in sorption behaviour of Brilliant Blue and iodide. Bearing in mind the boundary conditions of the tracer experiment like irrigation rate, the tracer employed, soil initial moisture or type of plantation (permanent or seasonal, deep rooted or shallow rooted) it is possible to compare different study sites or the same site at different boundary conditions and to access the risk of vertical solute propagation.

For practical application tracer experiments could be standardised and applied on different soils to determine their susceptibility to vertical solute propagation. Brilliant Blue could serve as a reference solute as it is easy to use and has acceptable toxicological properties. As long as preferential flow occurs, Brilliant Blue reflects well the water flow. For matrix flow dominated soils, laboratory studies on soil columns could compare the sorption behaviour of Brilliant Blue with other solutes of interest. The statistical analysis based on the risk index produces compact results that are easy to understand and could be used to establish water protection areas.

## 4.2 Flow processes and their impacts on soil chemistry

Pattern analysis allows identification of the dominant flow regime and the distribution $1 - H$ indicates the risk of vertical solute propagation for given boundary conditions. The mechanisms and consequences of water flow can be investigated by comparing soil physical and chemical parameters in preferential flow paths and soil matrix. Inferential statistics often imposes a certain number of constraints to analysed data, like independence or homogeneity of variance in the classical analysis of variance, for instance. These hypotheses are often violated by the sampling method since data is hardly ever acquired randomly. Especially when soil samples are taken during tracer studies grouping of data at several hierarchical levels (plot, profile and depth) is evident and dependencies in data might occur. Additional difficulties are missing data and complicated experimental designs with a mixture of crossed and nested fixed and random factors.

Often such data is analysed by classical statistical tests after separation in homogeneous groups, i.e. different plots or different depths are analysed separately. In this thesis we have successfully applied a particular type of statistical analysis, the mixed-effects model. It does not only reflect possible dependencies in data, but also improve the robustness of statistical analysis because all data can be included in the model without the need to investigate different depths or different plots separately. Conclusions about differences between preferential flow paths and soil matrix are not affected by between-plot or between-sample variability since it is taken into account by random-effects.

## *4.3 Fu rther investigations*

### 4.3.1 Experimental techniques

Successful application of VIS-NIR spectroscopy to determine the concentration of Brilliant Blue in laboratory should be extended to in situ measurements on soil profiles. Provided that robust calibration by combination of spectral and designed variables is possible, concentration maps of Brilliant Blue could be determined for study sites with varying soil colours. In situ measurements of Brilliant Blue would easily increase the number of samples for calibration against photographs of soil profiles.

Rainfall simulations on soil columns showed that Brilliant Blue might be substantially retarded with respect to reference tracers. Since the retardation is largest when matrix flow dominates a retardation index based on the classical CDE (convection dispersion equation) model could be developed. However, we need a better description of boundary conditions of the experimental design. Indeed, a suction applied directly to the bottom of the soil column without a ceramic plate could affect the gas phase and complicate modelling.

### 4.3.2 Dynamics of water flow at plot scale

We now can monitor processes on soil surface with high spatial and temporal resolution using stereophotography. But subsurface processes are considered by means of patterns, i.e. only with a high spatial resolution. The dynamic of infiltration can not be captured by point measurements with TDRs or tensiometers, as demonstrated during tracer experiments at the French site. Geoelectrical methods like 3D Electrical Resistivity Tomography (ERT) allow a

non-destructive monitoring of soil moisture (Samouëlian et al., 2003; Samouëlian et al., 2005; al Hagrey, 2007). However, this technique must be adapted to allow visualisation of distinct flow patterns at a spatial resolution of some centimetres.

### 4.3.3 Gradients of soil chemical properties

Analysis of soil bulk samples revealed distinct chemical properties in preferential flow paths and soil matrix. However, this sampling technique does not permit investigation of possible gradients in the transition zone. Spatial resolution of measurements could be increased using soil visible and near infrared diffuse spectroscopy (VIS-NIR spectroscopy). Small scale measurements (areas of some $cm^2$) of soil physical and chemical parameters could help to evaluate the extent of zones influenced by preferential flow.

### 4.3.4 Data analysis

In this thesis we acquired a large number of images of flow patterns from different soils. Images, even binary ones, are high dimensional objects. Recent developments in machine learning and artificial intelligence demonstrated how dimensionality reduction techniques could reveal low dimensional structure in high dimensional data (Weinberg & Saul, 2006). Applied to stained patterns this method could uncover zones of homogeneous flow within soil profiles.

## *4.4 References*

al Hagrey, S.A. 2007. Geophysical imaging of root-zone, trunk, and moisture heterogeneity. *Journal of Experimental Botany*, **58,** 839-854.

Samouëlian, A., Cousin, I., Richard, G., Tabbagh, A. & Bruand, A. 2003. Electrical resistivity imaging for detecting soil cracking at the centimetric scale. *Soil Science Society of America Journal*, **67,** 1319-1326.

Samouëlian, A., Cousin, I., Tabbagh, A., Bruand, A. & Richard, G. 2005. Electrical resistivity survey in soil science: a review. *Soil & Tillage Research*, **83,** 173-193.

Schlather, M. & Huwe, B. 2005. A risk index for characterising flow pattern in soils using dye tracer distributions. *Journal of Contaminant Hydrology*, **79,** 25-44.

Weinberger, K.Q. & Saul, L.K. 2006. Unsupervised learning of image manifolds by semidefinite programming. International Journal of Computer Vision, 70, 77-90.

# Appendices

# Appendix A

## Rapid prediction of Brilliant Blue concentrations in soil by diffuse reflectance spectroscopy

BOGNER C[a,b], SCHMIEDINGER I[a], HUWE B[a]

[a]*Soil Physics Group, University of Bayreuth, 95440 Bayreuth, Germany and*

[b]*Ecological Modelling, University of Bayreuth, Dr.-Hans-Frisch-Straße 1-3, 95448 Bayreuth, Germany*

Correspondence: C. Bogner. E-mail: christina.bogner@uni-bayreuth.de.

Short title:
*Prediction of Brilliant Blue by VIS-DRS*

## Summary

Brilliant Blue is often used to trace water movement in soils. Its concentration in soil samples is usually determined by extraction – a laborious procedure with varying accuracy. We show that Brilliant Blue can be measured directly by visible diffuse reflectance spectroscopy. We build a PLSR model for the concentration range of 0.1 to 15 mg Brilliant Blue per g soil with an RMSE of 1 mg $g^{-1}$ and an $R^2_{adj}$ of 0.9. As the method is based on visible spectra, prediction accuracy can be seriously affected by variation in soil colour between calibration and prediction data sets.

## Introduction

Brilliant Blue FCF is frequently used in vadose zone hydrology to trace water movement in soil and to study preferential flow (Flury *et al.*, 1994). Stained patterns are photographed and used to determine concentration maps of the dye, for example (Forrer *et al.*, 1999). Other studies deal with the impact of preferential flow on soil chemistry and biology (Hagedorn *et al.*, 1999; Bundt *et al.*, 2001a; Bundt *et al.*, 2001b). Thereby, the content of total and organic C and N in soil is analysed. Brilliant Blue ($C_{37}H_{34}N_2Na_2O_9S_3$, molar mass 792.9 g $mol^{-1}$) is an organic molecule consisting of 56% C and 4% N. Sorption of the dye on soil particles affects the C and N contents of soil and should be corrected.

Usually, Brilliant Blue concentrations are determined by extracting the dye with a water acetone solution or a 0.5 M $K_2SO_4$ (e.g. Bundt *et al.*, 2001a). This is a laborious procedure with changing accuracy due to varying mass recovery (Forrer *et al.*, 2000). We propose to determine the content of Brilliant Blue by visible diffuse reflectance spectroscopy (VIS-DRS) directly on soil samples without extraction.

## Materials and methods

### *Soils*

We sampled soils from two different sites: a Calcisol (IUSS Working Group WRB, WRB, 2007) on an agricultural field at the experimental site of INRA Avignon, France, and a Podzol (IUSS Working Group WRB, WRB, 2007) in a Norway spruce forest in the Fichtelgebirge in south-eastern Germany. While on the first site soil colour remained homogenously greyish yellow (Oyama &

Takehara, 1999) throughout the profile, it varied considerably from greyish yellow brown to light yellow in the forest soil (Table A.1).

Table A.1: Soil characteristics.

| Site | Horizon or depth /cm | Sand /% | Silt /% | Clay /% | Colour[a] | Colour name[a] |
|---|---|---|---|---|---|---|
| I | 0-30 | 5.18 | 48.35 | 46.47 | 2.5Y 6/2 | greyish yellow |
| | 30-60 | 4.94 | 48.14 | 46.93 | 2.5Y 6/2 | greyish yellow |
| | 60-100 | 15.53 | 47.63 | 36.85 | 2.5Y 6/2 | greyish yellow |
| II | Ea | 43.16 | 48.47 | 8.37 | 10YR 5/2 | greyish yellow brown |
| | Bsh | 38.63 | 51.48 | 9.90 | 10YR 4/2 | greyish yellow brown |
| | Bs | 33.94 | 50.38 | 15.69 | 10YR 4/4 | brown |
| | Bw | 31.61 | 54.53 | 13.87 | 2.5Y 7/4 | light yellow |
| | Bw/C | 36.62 | 49.71 | 13.67 | 2.5Y 7/4 | light yellow |

[a]determined on dry samples according to Oyama & Tekahara (1999)

### *Sample preparation and measurement*

We prepared 40 calibration samples from site one (calibration set one) and 65 from site two (calibration set two). Five g of sieved soil (< 2mm) were mixed with 5 ml of Brilliant Blue solution. The final concentration of the dye for every depth or horizon ranged from about 0.1 to about 15 mg $g^{-1}$ soil for calibration set one and from 0.1 to about 10 mg $g^{-1}$ soil for calibration set two. After mixing the soil with Brilliant Blue solution the samples were dried at 40°C and ground. Two independent data sets of 20 samples each from site one (validation set one) and from site two (validation set two) were prepared in the same way.

Visible diffuse reflectance spectra were collected using a Cricket accessory (Harrick Scientific Products) installed in a Cary 100 Conc UV-VIS spectrometer (Varian). An aliquot of a ground sample was scanned between 400 to 700 nm in 1 nm steps averaging 10 measurements per step. Then, a baseline correction procedure was applied:

$$\text{Corrected spectrum} = \frac{\text{Raw spectrum} - 0\%\text{Transmission}}{100\%\text{Transmission} - 0\%\text{Transmission}} \qquad \text{(A.1)}$$

where 0%Transmission is the zero reference baseline collected with the sample beam covered and the 100%Transmission is the spectrum collected on spectralon (100% reflectance reference). Spectra were recorded in reflectance units (R) and transformed to absorbance units, i.e. log (1/R).

### *Partial least squares regression*

We used the software ParLes (Viscarra Rossel, 2008) that allows multivariate calibration and prediction based on partial least squares regression combined with bootstrap aggregation (bagging-PLSR). PLSR is a multivariate linear regression technique that is particularly useful when predicting a dependent variable from a large number of independent variables. For an introduction to PLSR see for instance Geladi & Kowalski (1986), a tutorial on bootstrapping for chemical application can be found in Wehrens *et al*. (2000). During bagging, the calibration data set is repeatedly sampled with replacement and a PLSR model is calculated for each of these subsamples. Then, a mean predictor with confidence intervals can be derived. Viscarra Rossel (2007) showed that bagging improved the robustness of PLSR, was less susceptible to overfitting, improved the prediction accuracy and provided a measure for model uncertainty.

Prior to calibration, spectra were mean centred and smoothed with a wavelet filter at scale three. To choose the number of factors to use in the PLSR model, a leave-one-out cross validation was performed. The accuracy of the cross validation is measured by the root mean squared error of prediction (RMSE):

$$RMSE = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(y_i - \hat{y}_i)^2}\,, \qquad \text{(A.2)}$$

where $N$ is the number of samples $y_i$ the observed value, $\hat{y}_i$ the predicted one. The optimal number of factors, i.e. the parsimonious model is chosen based on the corrected Akaike's Information Criterion ($AIC_c$) (Sugiura, 1978):

$$AIC_c = N\ln(RMSE^2) + 2K + \frac{2K(K+1)}{N-K-1}\,, \qquad \text{(A.3)}$$

where $N$ is the number of samples and $K$ the number of factors. The model that represents an 'elbow' in the scree plot ($AIC_c$ vs number of factors curve) is

chosen. Some further measures to assess the goodness of the model are the bias (BIAS):

$$BIAS = \frac{1}{N}\sum_{i=1}^{N}(\hat{y}_i - y_i), \tag{A.4}$$

the standard deviation of the error distribution (SDE):

$$SDE = \sqrt{\frac{\sum_{i=1}^{N}((\hat{y}_i - y_i) - BIAS)^2}{N-1}}, \tag{A.5}$$

and RPD – the ratio of the standard deviation of the data to the RMSE.

## Results and discussion

### *Calibration and validation*

The diffuse spectra show a maximum of absorption at about 630 nm which corresponds to the Brilliant Blue absorption maximum (e.g. Forrer (1997)). The second smaller peak at about 430 nm is hardly developed (Figure A.1).

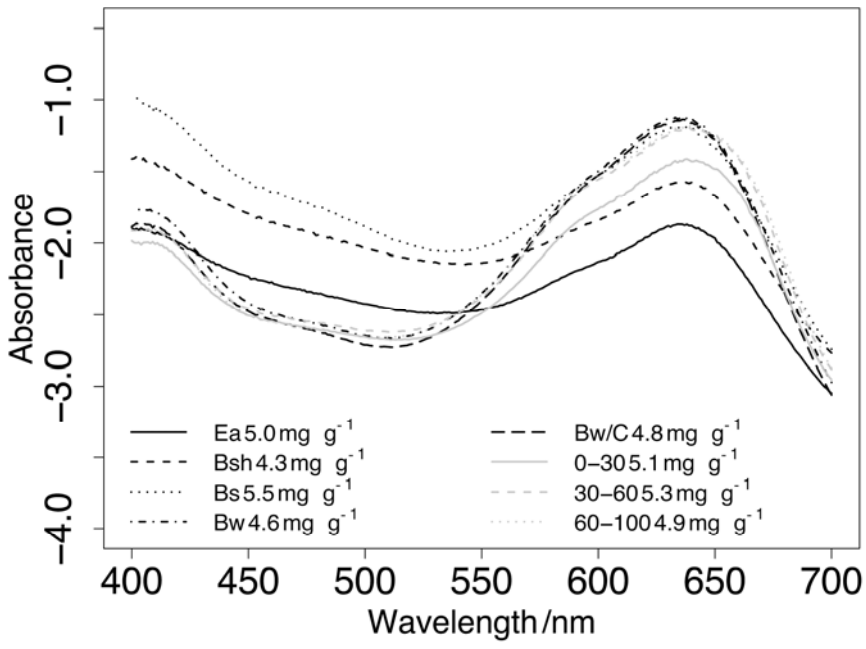



Figure A.1: Soil visible diffuse spectra with about 5 mg Brilliant Blue per g soil.

The leave-one-out cross validation indicated a model with 4 factors for calibration set one and a model with 6 factors for calibration set two (Figure A.2). The first 4 factors account for 95.61% of variation in the Brilliant Blue concentrations and 99.99% of variation in the spectra of calibration set one. In the calibration set two, the model explained 95.17% of variation in the Brilliant Blue concentrations and 99.99% of variation in the spectra. The leave-one-out validation statistics are comparable between the two calibration sets. The RMSE of about 1 mg $g^{-1}$ and an $R^2_{adj}$ of 0.94 indicate good models (Table A.2). The fitted versus observed values plots did not show any non linearity (Figure A.3). In the calibration set one, one point with a high residual was detected, but it did not deteriorate the RMSE of the validation.

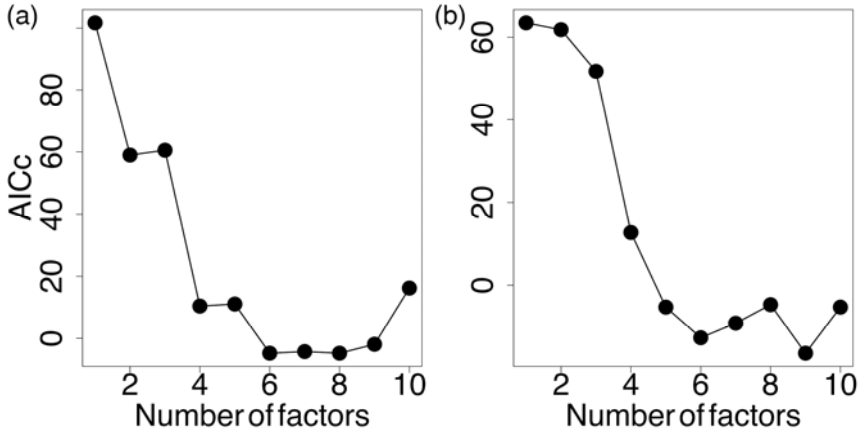



Figure A.2: Evolution of the corrected Akaike's Information Criterion ($AIC_c$) with increasing number of factors; (a) calibration set one and (b) calibration set two.

Table A.2: Statistics of the calibration procedure based on leave-one-out cross validation.

| Calibration set | RMSE /mg $g^{-1}$ | $R^2_{adj}$ | BIAS /mg $g^{-1}$ | SDE /mg $g^{-1}$ | RPD |
|---|---|---|---|---|---|
| One | 1.29 | 0.94 | −0.01 | 1.31 | 4.04 |
| Two | 0.82 | 0.94 | −0.01 | 0.82 | 4.10 |

The statistics of the leave-one-out cross validation and the independent validation are comparable (Table A.3). Validation set one gave an RMSE of 1.09 mg $g^{-1}$ and an $R^2_{adj}$ of 0.95. For validation of set two we obtained an RMSE of 0.98 mg $g^{-1}$ and an $R^2_{adj}$ of 0.87. All depths in validation set one and horizons in validation set two were equally well predicted (Figure A.4).

Table A.3: Statistics of the validation procedure.

| Calibration set | RMSE /mg $g^{-1}$ | $R^2_{adj}$ | BIAS /mg $g^{-1}$ | SDE /mg $g^{-1}$ | RPD |
|---|---|---|---|---|---|
| One | 1.09 (0.83, 1.59)[a] | 0.95 | −0.15 | 1.11 | 4.57 |
| Two | 0.98 (0.74, 1.43)[a] | 0.87 | −0.14 | 1.00 | 2.87 |

[a]95% confidence interval

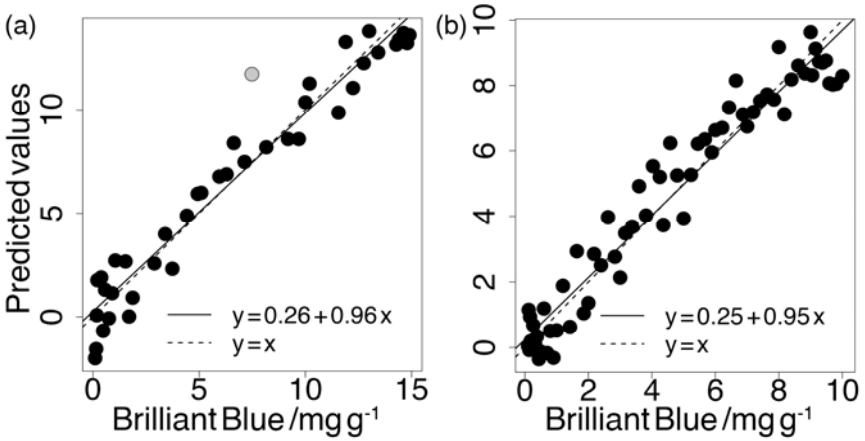



Figure A.3: Results of the leave-one-out cross validation; (a) calibration set one and (b) calibration set two.

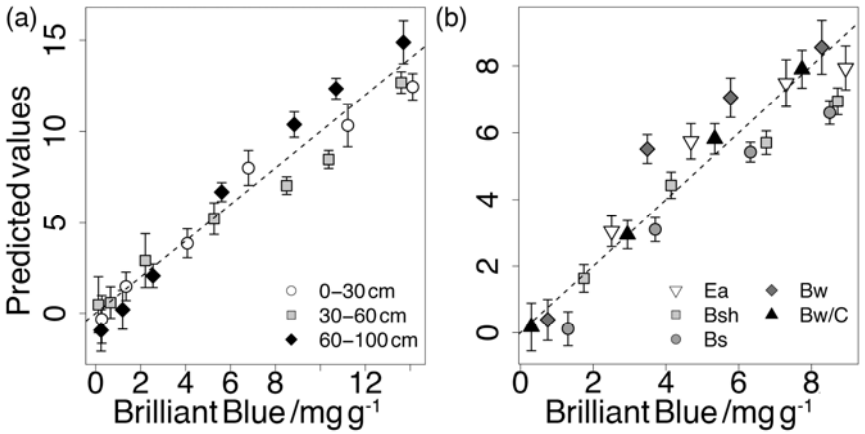



Figure A.4: Results of the independent validation; (a) calibration set one and (b) calibration set two.

## *Robustness*

To infer the robustness of the models, we calculated the concentrations of Brilliant Blue in validation set one using the calibration set two and vice versa. Using the same PLSR models as described above yielded an $R^2_{adj}$ of 0.22, an RMSE of 13.51 mg $g^{-1}$ and a BIAS of -10.66 mg $g^{-1}$ for validation set two. Prediction was better for calibration set one, leading an $R^2_{adj}$ of 0.87, an RMSE of

2.94 mg $g^{-1}$ and a BIAS of 1.65 mg $g^{-1}$. This deterioration of prediction ability is probably due to variations in soil colour, as the concentration range of Brilliant Blue is similar in both calibration sets. Soil colour must be considered as a nuisance parameter that changes the hue of samples with the same concentration of Brilliant Blue, thus affecting their VIS diffuse spectra.

To improve prediction we combined both calibration sets. The pre-processing of the spectra had to be adjusted to account for a more heterogeneous data set. In addition to wavelet smoothing and mean-centring we applied a wavelet detrending and calculated the first derivative. Wavelet detrending is often used to correct baseline shifts or to remove curvilinearity in the spectra (Viscarra Rossel *et al.*, 2007). The prediction of both validation sets together was satisfactory and gave an $R^2_{adj}$ of 0.88, an RMSE of 1.34 mg $g^{-1}$ and a BIAS of 0.11 mg $g^{-1}$.

## Conclusions

Rapid and accurate prediction of Brilliant Blue by diffuse reflectance spectroscopy is possible. But the accuracy of prediction can decrease when soil colour changes between calibration and prediction data sets. To deal with this problem a regression model with both spectral and designed variables including additional information like horizon or soil colour could be applied (Jørgensen *et al.*, 2004).

## Acknowledgements

This project was funded by the Deutsche Forschungsgemeinschaft (DFG FOR 562).

## References

Bundt, M., Jaggi, M., Blaser, P., Siegwolf, R. & Hagedorn, F. 2001a. Carbon and nitrogen dynamics in preferential flow paths and matrix of a forest soil. *Soil Science Society of America Journal*, **65,** 1529-1538.

Bundt, M., Widmer, F., Pesaro, M., Zeyer, J. & Blaser, P. 2001b. Preferential flow paths: biological 'hot spots' in soils. *Soil Biology & Biochemistry*, **33,** 729-738.

Flury, M., Flühler, H., Jury, W.A. & Leuenberger, J. 1994. Susceptibility of soils to preferential flow of water: a field study. *Water Resources Research*, **30,** 1945-1954.

Forrer, I.E. 1997. *Solute transport in an unsaturated field soil: visualisation and quantification of flow patterns using image analysis*. Ph.D. thesis, Swiss Federal Institute of Technology, Zurich.

Forrer, I.E., Kasteel, R., Flury, M. & Flühler, H. 1999. Longitudinal and lateral dispersion in an unsaturated field soil. *Water Resources Research*, **35,** 3049-3060.

Forrer, I.E., Papritz, A., Kasteel, R., Flühler, H. & Luca, D. 2000. Quantifying dye tracers in soil profiles by image processing. *European Journal of Soil Science*, **51,** 313-322.

Geladi, P. & Kowalski, B.R. 1986. Partial least-squares regression - a tutorial. *Analytica Chimica Acta*, **185,** 1-17.

Hagedorn, F., Mohn, J., Schleppi, P. & Flühler, H. 1999. The role of rapid flow paths for nitrogen transformation in a forest soil: A field study with micro suction cups. *Soil Science Society of America Journal*, **63,** 1915-1923.

IUSS Working Group WRB. 2007. World reference base for soil resources 2006, first update 2007. World Soil Resources Reports No. 103.

Jørgensen, K., Segtnan, V., Thyholt, K. & Næs, T. 2004. A comparison of methods for analysing regression models with both spectral and designed variables. *Journal of Chemometrics*, **18,** 451-464.

Oyama, M. & Takehara, H. 1999. *Revised standard soil color charts*. Eijkelkamp Agrisearch Equipment, Soil colour book 08.11, Netherlands.

Sugiura, N. 1978. Further analysts of the data by akaike's information criterion and the finite corrections. *Communications in Statistics-Theory and Methods*, **7,** 13-26.

Viscarra Rossel, R.A. 2007. Robust modelling of soil diffuse reflectance spectra by "bagging-partial least squares regression". *Journal of near Infrared Spectroscopy*, **15,** 39-47.

Viscarra Rossel, R.A. 2008. ParLeS: Software for chemometric analysis of spectroscopic data. *Chemometrics and Intelligent Laboratory Systems*, **90,** 72-83.

Viscarra Rossel, R.A., Taylor, H.J. & McBratney, A.B. 2007. Multivariate calibration of hyperspectral γ-ray energy spectra for proximal soil sensing. *European Journal of Soil Science*, **58,** 343-353.

Wehrens, R., Putter, H. & Buydens, L.M.C. 2000. The bootstrap: a tutorial. *Chemometrics and Intelligent Laboratory Systems*, **54,** 35-52.

# Appendix B

## Small scale dynamics of water flow under extreme thermal and soil hydrological boundary conditions

Published in: Mitteilungen der DBG 110(1):83-84

Original title: **Kleinskalige Dynamik und Fließmuster bei extremen thermischen und bodenhydrologischen Randbedingungen**

C. BOGNER[1], G. WIEDEMANN[1], B. HUWE[1], V. MARC[2], S. RUY[3], Y. TRAVI[2]

[1] *Soil Physics Group, University of Bayreuth, D-95440 Bayreuth, Germany*

[2] *Laboratoire d'Hydrogéologie, UFR Sciences, 33 rue Louis Pasteur, 84000 Avignon, France*

[3] *INRA - Unite Climat, Sol et Environnement, Domaine St Paul – Agroparc, 84914 Avignon Cedex 9, France*

*Corresponding author: christina.bogner@uni-bayreuth.de*

## Introduction

In future extreme events like drying/rewetting and freezing/thawing of soils will probably occur more often due to climate changes. As part of the DFG research group 562 "Dynamics of soil processes under extreme meteorological boundary conditions" we studied possible impacts of these phenomena on water flow, flow patterns and transport of nutrients and contaminants towards groundwater.

## Materials and methods

We did freezing/thawing and drying/rewetting experiments in a Norway spruce forest in Waldstein (southeast Germany). Thereafter, undisturbed soil columns (diameter 15 cm and height 30 cm) and soil material for disturbed soil columns were sampled on treatment and control plots. In this study we focus on dynamics of water flow and behaviour of Brilliant Blue in disturbed and undisturbed soil columns from control plots.

Soil columns were irrigated with a bromide solution at different rates (10, 32 und 64 mm $h^{-1}$). Figure B.1 shows the experimental setup. Brilliant Blue is a well known tracer for applications in vadose zone hydrology because of its low toxicity and good visibility against most soil colours (Flury and Flühler 1994). The sorption behaviour of Brilliant Blue, however, is non-linear and depends on soil properties (German-Heins and Flury 2000). In order to see a possible retardation of Brilliant Blue with regard to infiltrating water, we used KBr and $^{18}O$-enriched water as ideal tracers (analyses of $^{18}O$ samples have not been finished yet).

## First results

During rainfall simulations with a bromide solution at different irrigation rates (10, 32 und 64 mm $h^{-1}$) disturbed and undisturbed soil columns began to drain after a similar time of irrigation (not shown here). Bromide concentration in breakthrough curves (BTCs) started to rise earlier in undisturbed columns, but increased slowly. The time between breakthrough and a quasi steady state concentration was similar in disturbed and undisturbed soil columns at the same irrigation rate. The difference in breakthrough times between disturbed and undisturbed soil columns increased with increasing irrigation rate.

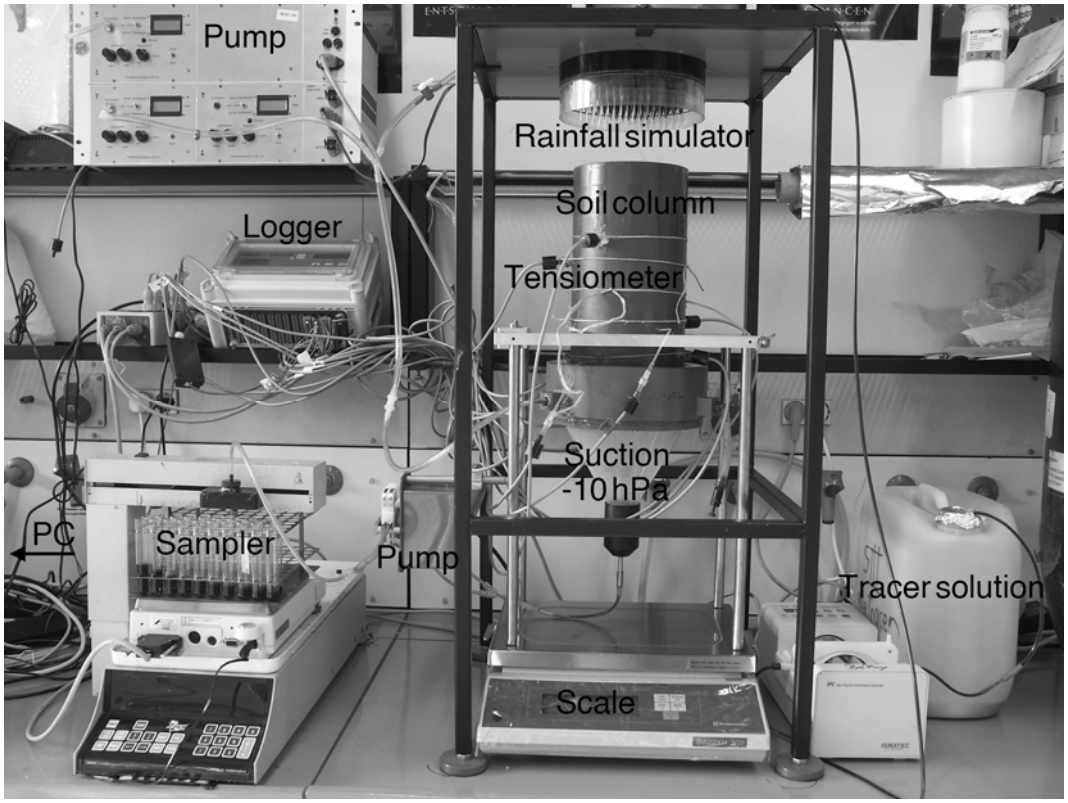



Figure B.1 Experimental setup for irrigation studies on soil columns

Figure B.2 illustrates results of the rainfall simulation with a solution containing Brilliant Blue, KBr, KI and $^{18}$O-enriched water (atmosphere pressure at lower end of the soil column). In disturbed soil columns BTCs of bromide and iodide began to rise earlier than that of Brilliant Blue. After about three hours of irrigation the concentrations of bromide and iodide in the drainage water was half their respective concentrations in the irrigation solution. Assuming that matrix flow dominated this is approximately the time the tracers needed to percolate through the columns. In undisturbed columns all three tracers broke through at about the same time and earlier than in disturbed columns. This is probably due to rapid flow paths.

## Conclusions

Although disturbed and undisturbed columns began to drain after a similar time of irrigation, bromide breakthrough times were longer for disturbed columns. This means that at the beginning old water left the columns, but that rapid flow paths were active in undisturbed soil columns. When matrix flow dominates Brilliant Blue is heavily retarded with regard to bromide and iodide as indicated by its late

breakthrough in disturbed soil columns. In such cases its transport behaviour differs from flow behaviour of infiltrating water.

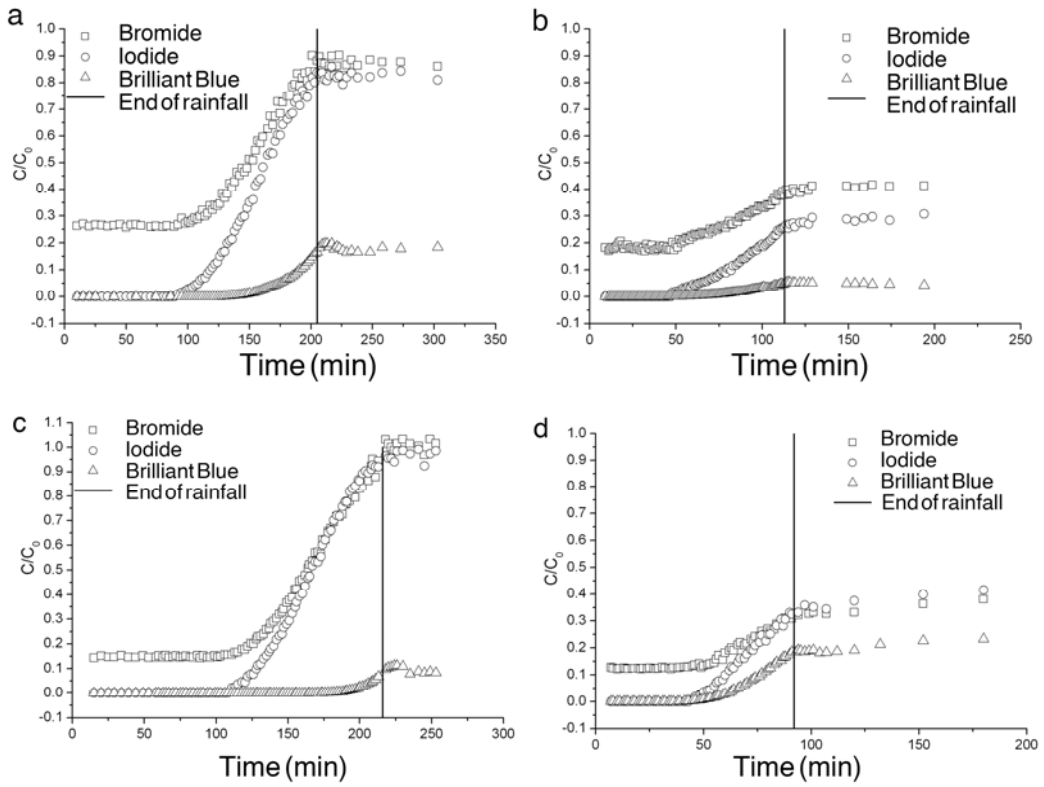



Figure B.2 Break through curves of bromide, iodide and Brilliant Blue in disturbed (a and c) and undisturbed (b and d) soil columns.

## Acknowledgements

We thank Andreas Kolb, Iris Schmiedinger, Folkert Bauer, Prof. Y. Kuzyakov, Prof. H. Flühler, the mecanics workshop GEO and the DFG.

## References

Flury, M. & Flühler, H. 1994. Brilliant Blue FCF as a dye tracer for solute transport studies - a toxicological overview. *Journal of Environmental Quality* **23**(5), 1108-1112.

German-Heins, J. & Flury, M. 2000. Sorption of Brilliant Blue FCF in soils as affected by pH and ionic strength. *Geoderma* **97**(1-2), 87-101.

# Appendix C

## List of other publications

Hentschel, K., Borken, W., Zuber, T., Bogner, C., Huwe, B. & Matzner, E. 2009. Effects of soil frost on nitrogen net mineralization, soil solution chemistry and seepage losses in a temperate forest soil. *Global Change Biology*, **15,** 825-836.

## DECLARATION OF AUTHENTICITY

To the best of my knowledge and belief, this thesis contains no material previously published or written by another person, except where due reference has been made.

This thesis contains no material which has been accepted or definitely rejected for the award of any other doctoral degree in any university.

## EHRENERKLÄRUNG

Hiermit erkläre ich, dass ich die Arbeit selbständig verfasst und keine anderen als die von mir angegebenen Quellen und Hilfsmittel benutzt habe.

Ferner erkläre ich, dass ich anderweitig mit oder ohne Erfolg nicht versucht habe, diese Dissertation einzureichen. Ich habe keine gleichartige Doktorprüfung an einer anderen Hochschule endgültig nicht bestanden.

Bayreuth, den 21. April 2009