LARGE SCALE APPLICATION OF AN ENVIRONMENTAL TRACER APPROACH: SPATIO-TEMPORAL PATTERNS OF HYDROCHEMISTRY IN A SEMI-ARID GRASSLAND

FRAUKE KATRIN BARTHOLD

A dissertation submitted to the Department of Natural Sciences and prepared at the Department of Agricultural Sciences, Nutritional Sciences and Environmental Management of the

Justus-Liebig-Universität Gießen, Germany

for the degree of Doctor of natural Sciences (Doctor rerum naturalium).

Submitted	May 18th, 2010
Defended	July 9th, 2010

Referees

Prof. Hans-Georg Frede	Justus-Liebig-Universität Gießen
Prof. Kellie B. Vaché	Oregon State University
Prof. Lorenz King	Justus-Liebig-Universität Gießen
Prof. Peter Felix-Henningsen	Justus-Liebig-Universität Gießen

The photograph on the front cover is used with kind permission of M. Wiesmeier.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

1 SYNOPSIS

1.1 INTRODUCTION

Effects of land use change on the hydrological cycle are currently a major research issue. This is particularly the case in rangelands which are situated in semi-arid or arid regions of the temperate zone and where water availability is already limited by definition. Much of the research focus has been laid on the rangelands of midwestern North America which have a long history of land use change that profoundly alters ecosystem processes. In rangelands, such changes, which were currently termed *transformative change* by *Wilcox* [2010], are induced by overgrazing [*Asner et al.*, 2004; *Reynolds et al.*, 2007], invasion by non-native species [*Masters and Sheley*, 2001; *Nagler et al.*, 2008; *Stromberg et al.*, 2009; *Wilcox*, 2007] and atmospheric change [*Polley et al.*, 2006] and result in three types of degradation: woody plant encroachment, invasive species and desertification [*Wilcox and Thurow*, 2006]. The effects of rangeland degradation are significant enough to describe it as a major component of global change [*Asner et al.*, 2004], however, its impacts on the water cycle on a landscape scale are still poorly understood [*Wilcox*, 2007].

There exist various approaches to evaluate the influence of land use change on hydrological processes. These include field studies that are designed as either paired site studies or single catchment experiments [*Bosch and Hewlett*, 1982; *Brown et al.*, 2005; *Stednick*, 1996]. However, most of these studies are performed in small catchments with sizes < 2 km^2, and focused on the hydrologic effects of forest vegetation management, including both deforestation as well as regeneration. At larger scales, manipulative field experiments are of limited value because an interpretation of the hydrologic effects is not feasible due to the wide range in natural variability [*Andreassian*, 2004]. Another approach, that is better suited for larger catchments, is the application of hydrological models.

Hydrological modeling of land use changes has been conducted since the 1970s [*Onstad and Jamieson*, 1970]. There exists a wide range of models that were developed to describe catchment hydrology on a regional scale. These include perceptual models ranging from process-based and spatially very detailed to conceptual-based with a rather coarse spatial representation. It is commonly argued that process-based fully distributed

models are best suited to simulate land use change effects [*Binley et al.*, 1991]. It is further argued that in most cases only parts of the land use in a catchment changes and hence, that spatially distributed models such as MIKE-SHE, DHSVM, RHESSys, TOPLATS, WASIM, SHETRAN, are able to capture these changes more precisely than lumped modeling approaches [*Beven*, 2001]. An alternative are physically-based but less complex in terms of spatial representation, semi distributed models, e.g. SWAT, SWIM, SLURP or HYLUC. This group of models simulates all hydrological processes within spatially non-explicit Hydrological Response Units (HRUs). The results for each HRU are lumped within subcatchments and then routed downstream. A further simplification is achieved if hydrological fluxes are simulated within the subcatchment scales as the smallest spatial unit (HBV, LASCAM) or even for the whole catchment (IHACRES).

One key assumption for hydrological modeling of land use changes is that the model captures all relevant hydrological and biogeochemical fluxes. The degree to which this assumption is correct is a key factor determining the effectiveness of the model results. One mechanism to establish this is model uncertainty analysis through comparisons between model outputs and measured system properties. This requires the existence of sufficiently large data sets that not only include model input data such as climate data or soil physical properties, but also data for model evaluation, e.g. discharge and groundwater level, amongst many others. However, these data are very often not available and this poses a general problem that has recently been addressed by the Predictions in Ungauged Basins (PUB) initiative of the International Association of Hydrological Sciences (IAHS). PUB prompts research in (1) a better understanding of rainfall-runoff processes, (2) new ideas of model conceptualization and (3) improved ways of model initialization and verification [*Sivapalan et al.*, 2003].

One contribution to this calling is the suggestion of soft data application in model calibration by *Seibert and McDonnell* [2002]. They propose to increase the dialog between the experimentalist and the modeler and implement more of the already existing process understanding into model concepts. They were able to show that additional information, the so called soft data, has the potential to go beyond the classical way of model evaluation that so far focuses on discharge measurements and groundwater levels. In a recently published commentary, *Sivapalan* [2009] carries this

idea even a bit further and suggests moving away from hydrograph fitting and turning towards creative analysis of available data in order to gain insight into the functioning of the catchment. Environmental tracers, such as water isotopes and water chemical constituents, offer a great means to deliver such soft data and to improve the understanding of runoff contributing processes in catchments. For example, water isotopes can be used to estimate Mean Travel Time (MTT) of waters [*Burns*, 2002] and water chemical constituents are commonly used in end member mixing analysis (EMMA) for hydrograph separation [*Christophersen and Hooper*, 1992].

In a previous study [*Schäfer*, 2009], SWAT [*Arnold et al.*, 1998] was used to evaluate land use change effects on the hydrology of a large, ungauged catchment in the grasslands of Inner Mongolia. The results showed poor model effectiveness as indicated by common model evaluation measures (Figure 1-1) which proofed that the model did not capture the relevant hydrologic processes. A better understanding of the catchment relevant hydrologic processes is crucial for further land use change related hydrological modeling.

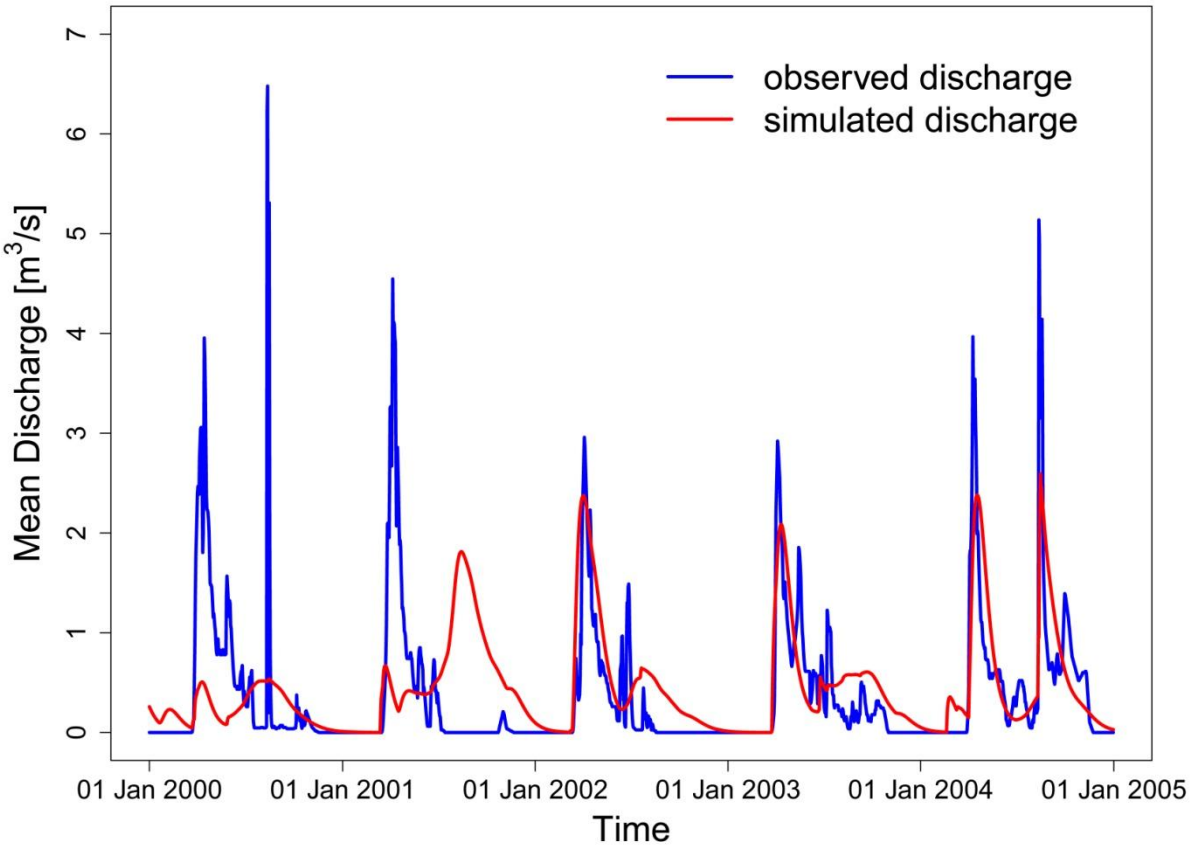

Figure 1-1. The lines represent observed and simulated discharge with SWAT (Nash-Sutcliffe-Efficiency (NSE) = 0.22 and R^2 = 0.16) [after *Schäfer*, 2009].

1.2 GENERAL OBJECTIVE

The primary objective of this dissertation was to generate a set of soft data for use in hydrological model evaluation. This research uses a multi tracer approach, including both isotopes and water chemical constituents, to improve the understanding of catchment relevant hydrological processes. The tracers are applied in various methodologies, although a specific focus here was placed on the determination and quantification of runoff contributing sources and processes on a catchment scale via end member mixing analysis.

1.3 STUDY AREA

This dissertation was embedded into the sino-german research group MAGIM (Matter fluxes in grasslands of Inner Mongolia as induced by stocking rate) which conducted research in Inner Mongolia. The principal objective of the research group was to clarify the interactions between grazing of steppe ecosystems and carbon, nitrogen and water cycling on site and regional scales with the aim to develop concepts for improved rangeland management.

The semi-arid steppe region in Inner Mongolia has undergone tremendous environmental change during the past 6 decades due to a number of political decisions. The major problem in large areas is desertification which results from overgrazing. In the 1950ies and 1960ies local stock farmers were forced to give up their nomadic way of living and settle in small villages or individual farms [*Sneath*, 1998]. In combination with a moderate increase of livestock numbers this policy has increased grazing pressure around the newly established settlements and ended in the large scale pastoral movement between seasonal pastures. In the 1980ies, a new economic policy was established that allowed individuals to profit directly from increased meat or wool production. This sharply fostered the pressure on the land resources and led to intensified land use and large scale overgrazing [*Jian and Meurer*, 2001]. In large parts of Inner Mongolia, overgrazing has already led to severe losses of soil organic matter, depletion of nutrients [*Steffens et al.*, 2008] and a decrease in steppe primary productivity [*White et al.*, 2000]. There exist several studies on steppe ecosystems in the prairies in North America which showed that

grazing affects heterogeneity of plant communities [*Burke et al.*, 1998; *Evans et al.*, 2001; *Vinton and Burke*, 1997], reduces above and belowground productivity [*Burke et al.*, 1998] and accelerates ecosystem C and N turnover and hence the biosphere-atmosphere exchange of greenhouse gases [*Epstein et al.*, 1998; *Mosier et al.*, 1997; *Watson and Mills*, 1998]. An ecological sustainable land use system of grasslands must be based on a precise knowledge of the interaction of stocking rate and herd management with cycles of water, carbon, nitrogen and soil erosion. Such information is scarce for colder climates in general and in Inner Mongolia in particular. Due to the marked differences in climate and vegetation [*X. R. Li*, 2001; *Pyankov et al.*, 2000; *White et al.*, 2000; *W. Zhang*, 1998] such concepts cannot simply be deduced from results obtained in other regions, e.g. the prairies of North America. The purpose of MAGIM was hence to understand how grazing of steppe ecosystem feedbacks on water, nitrogen and carbon fluxes and on soil erosion on site and regional scales in Inner Mongolia.

The Xilin River Basin represents the characteristic features of the grasslands in Inner Mongolia and was hence chosen as the principal study area of MAGIM. It is located in the Xilingol League in central Inner Mongolia Autonomous Province, China (Figure 1-2a) at approximately 43°24’ to 44°40’ N and 115°20’ to 117°13’ E and comprises about 10,000 km^2. The research was conducted in a subcatchment that is defined by the Xilinhot gauging station and comprises about 3600 km^2 (Figure 1-2b).

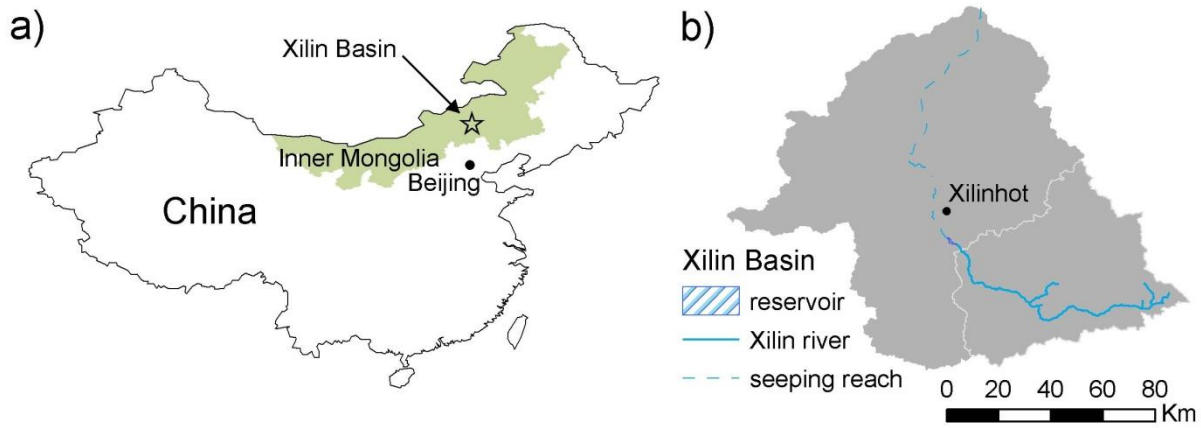

Figure 1-2. The maps show (a) the location of the Xilin River Basin in Inner Mongolia, China and (b) the extent of the Xilin River Basin and location of the subcatchment that was chosen as principal study area.

The Xilin River is an endorheic river system. It is characterized by a semi-arid continental climate with cold, dry winters and warm, relatively wet summers (Figure 1-3). The mean annual precipitation is 350 mm but is highly variable in space and time due to prevailing convective weather conditions. *Chen* [1988] reported annual ranges of 150 and 500 mm, with 60 to 80 % falling between June and August. The mean annual temperature is 2°C, with a January average of -23°C and a July average of 18°C [*Z. Chen*, 1988]. Due to its central location, the study area represents average climate conditions of Inner Mongolia where a west to east trending gradient with increasing precipitation rates and temperature characterizes the climate. The elevation ranges between 1010 and 1609 m.

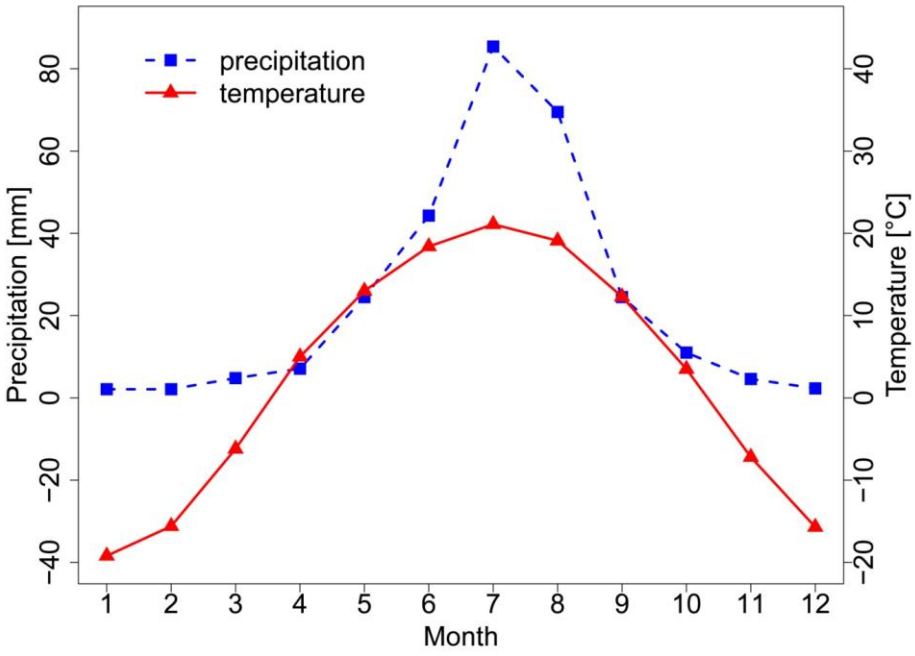

Figure 1-3. Mean monthly precipitation and temperature (1975-2003), Xilinhot, China.

Figure 1-4a provides a classification of the land use as derived from a Landsat image. The land use can be divided into 7 main ecological units. The map shows that 72.5 % of the study area is covered with steppe. The steppe is characterized by grassland species, dominated by *Leymus chinensis* and *Stipa grandis*. *Cleistogenes squarrosa* and *Artemisia frigida* dominate in degraded areas of the steppe regions [*Tong et al.*, 2004]. In higher elevated areas in the east and northeast of the basin, the vegetation changes from steppe to mountain meadows. The dominant species of the meadow grassland include *Bromus inermus*, *Agrostis gigantean*, *Carex pediformis, Stipa baicalensis* and *Calamagrostis*

epigeios. Marshland dominates in the vicinity of the river and its tributaries. The dominant species are *Phragmites australis*, *Carex appendiculata*, *Iris lacteal var. chinensis* and *Hippuris vulgaris*. Most of these grassland areas are used for agricultural purposes, mainly grazing of cattle, sheep and goat herds. Even sparsely vegetated areas, e.g. the paleo sand dunes that stretch through the center of the subcatchment in a northwest-southeast direction and branching out northwards (Figure 1-4a), are used for grazing. The sand dunes are populated with dense vegetation, particularly *Ulmus pumila* and other tree genii, i.e. *Betulus spp.*, *Malus spp.*, *Prunus spp.* and *Populus spp.* on north to northwest facing slopes or in depressions, while south to southeast facing slopes have generally sparser shrubs and grassland vegetation. A substantial and considerably increasing part of the area is also used for cultivation of crops such as maize, wheat and rapeseed [*Guo et al.*, 2004] despite unfavorable climate conditions and high risk for wind erosion.

In a previous study [*Barthold et al.*, under review] a soil map of the catchment was generated. The map delineates 8 Reference Soil Groups of the World Reference Base (WRB) [*IUSS Working Group WRB*, 2007] (Figure 1-4b). *Phaeozems* are by far the most extensive soils, with 1888.2 km^2 covering 51.4 % of the area. They occur mainly in more elevated areas throughout the whole area, covering large parts in the east and north east. *Arenosols*, *Chernozems* and *Gleysols* are almost equally extensive covering 16.4, 14.6 and 14.4 %, respectively. *Arenosols* mainly cover the sand dune area (Figures 1-4a and b). *Gleysols* are markedly developed along the river network and in groundwater dominated areas which are indicated by marshland vegetation such as *Phragmites australis*, *Carex appendiculata*, *Iris lacteal var. chinensis* and *Hippuris vulgaris* (personal field observations). *Chernozems* mainly occur in close vicinity with *Gleysols* and *Phaeozems*. They appear to fill the spatial gaps between the *Phaeozems* and the *Gleysols*. Very small patches of *Calcisols* and *Regosols* occur in areas underlain by sediments in the west of the catchment, northwards and southwards of the river. In total, they cover 0.3 % and 0.2 % of the study area, respectively. Some *Regosols* also occur in the east of the catchment. *Kastanozems* also cover just small parts of the area (1.1 %), mainly occurring on the basaltic plateaus in the southwestern part of the catchment (Figures 1-4b and c). The few sites that are mapped as *Cryosols* are linked with wetter regions in the vicinity of the river

network, e.g. especially in the wetland area located in the eastern part of the catchment and in a plain-like area in the north. They cover only 0.1 % of the area.

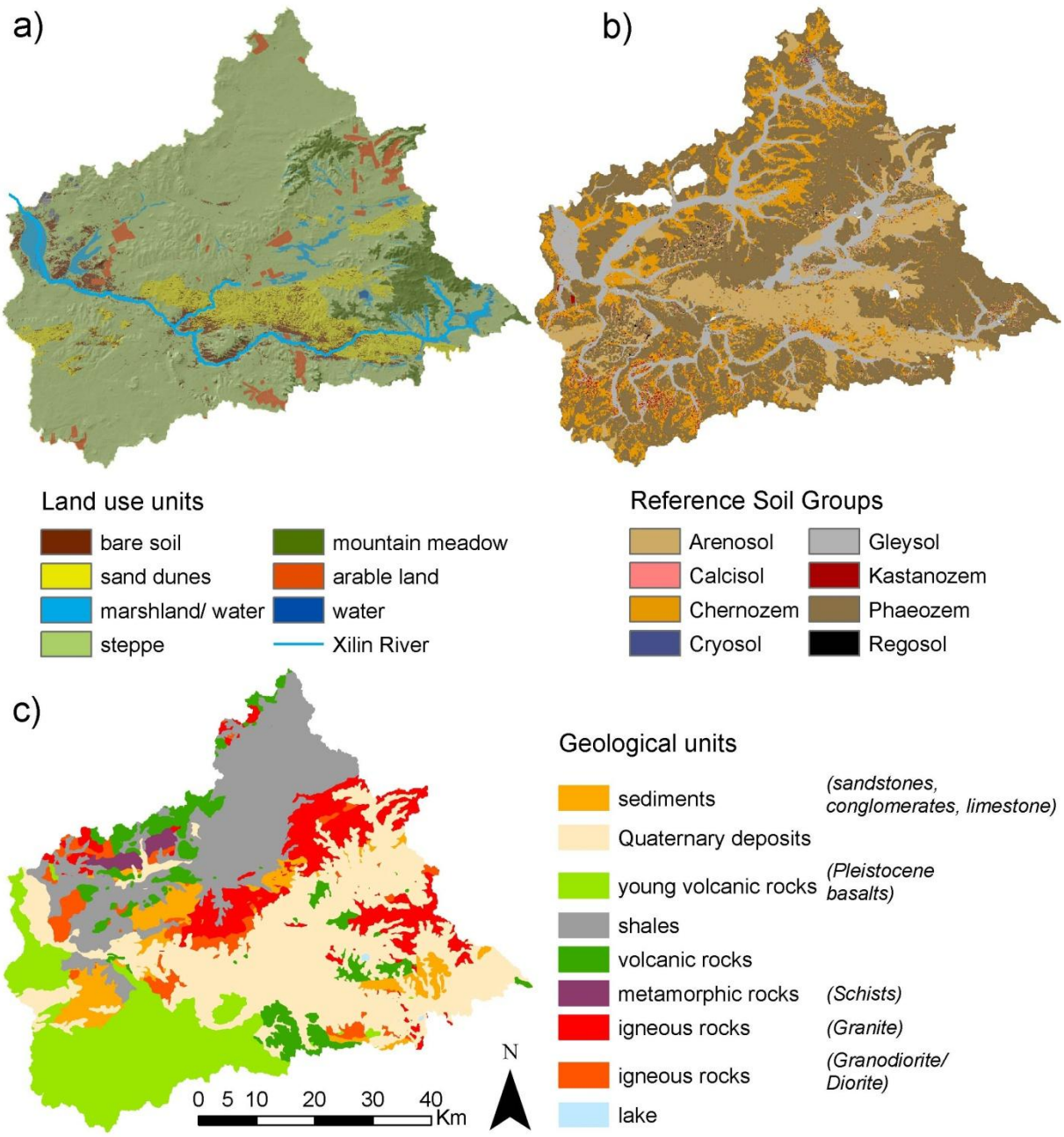

Figure 1-4. The maps show (a) hillshade of the Xilin River subcatchment with superimposed land use, (b) the soil map based on World Reference Base (WRB) Reference Soil Groups (RSGs) from [*Barthold et al.*, under review] and (c) geology of the study area. Landuse was classified based on a Landsat TM7 image from August 17th, 2005. The 1:200,000 geological map of the Inner Mongolian Bureau of Geology 1973 was modified and information was lumped into 9 new geological map units based on formation processes and age.

The geological setting of the study area is quite diverse and spans a large range of rocks (Figure 1-4c). The oldest systems which form the basis of the area originate from the Paleozoic and comprise a variety of igneous rocks (Granites, Granodiorites and Diorites), sediments and metamorphic rocks (Schists). Parts of them outcrop on the northwestern boundary of the catchment and in the east, as well as in a strip that stretches in a northeast - southwest direction through the center of the catchment. Some of these oldest areas are capped by volcanic rocks from the Mesozoic and together they form low mountains in the present-day landscape [*Z. Chen*, 1988; *Tong et al.*, 2004]. They also mark a borderline between an old sedimentation plain consisting of shales in the west and a large area of Quaternary deposits in the east. The Quaternary deposits are distributed along the river and take up most of the area in the east. Landforms such as sandy lands [*Z. Chen*, 1988; *Tong et al.*, 2004] and hills characterize their present-day landscape. Almost equally spatially dominant are Pleistocene basalts covering the southwest of the study area and forming lava table lands and high plateaus [*Z. Chen*, 1988; *Tong et al.*, 2004]. A more detailed description of rocktypes is given in *Barthold et al.* (under review).

1.4 THESIS OUTLINE

The thesis consists of a series of three closely related papers which address both the determination of runoff producing sources on a catchment scale and also some principal issues regarding EMMA. The former is presented in chapters 2 and 3. Dealing with the latter arose from the overwhelming wealth of environmental tracers that were available for use in the analysis. One basic assumption of EMMA is that the tracers are conservative, i.e. that they are neither involved in any biological nor chemical reactions and that they are time invariant. I struggled a long time deciding which and how many tracers to include. The selection of the right tracer set size and combination seemed crucial for the conceptual model that resulted from the analysis. After I spent a considerable amount of time on this aspect, I decided to address this problem in a separate paper and to estimate the uncertainty of the resulting conceptual model depending on tracer set size and composition. This study is presented in chapter 4 of this thesis. All of the research conducted for this dissertation was performed in the upper watershed of Xilin River catchment (475 km^2).

1.5 SUMMARY OF RESULTS

One of the grand challenges in large, ungauged catchments is to quickly and at low costs gain insight into the catchment relevant hydrologic processes. These are necessary in land use change oriented hydrological modeling to evaluate if the model of interest captures all relevant hydrological and biogeochemical processes. The research presented in this dissertation makes use of several methodologies to improve catchment understanding in a large, ungauged basin in the grasslands of Inner Mongolia (Figure 1-5).

Chapter 2 describes the attempt to quickly diagnose the dominant runoff producing sources and processes in a large ungauged basin. It uses a top-down approach from first field reconnaissance and data collection over perceptual model development, reservoir model conceptualization, evaluation with hydrochemical data to completion of the field campaign to assess the main runoff contributing sources of the Xilin river catchment. The specific objectives were (1) to use geochemistry-based field reconnaissance to assess dominant processes of water cycling within the mesoscale watershed, (2) to develop a conceptual reservoir model and (3) to evaluate and reject the model in order to yield new insight into catchment functioning that may guide a more complete field work.

Our findings indicate that a tributary, a shallow groundwater source (4 m deep) and the headwater source area are the dominant runoff producing sources based on geochemical analysis. We were able to develop a simple conceptual reservoir model with the identified sources as main storages. The model captured the general trend of the stream hydrograph but showed a general overestimation throughout the growing season. The simulated hydrographs of the three storages were also consistent with our general understanding of the runoff producing processes. The evaluation of the model which was based on simulations of the stream chemistry indicated that our three selected sources were not sufficient to explain the runoff in the catchment. The simulated geochemistry of the stream and the end members rather suggested that at least two important end members, of which one is characterized with low and the other one is characterized with high solute concentrations, are still missing. However, the approach which had been implemented in an effort to learn from wrong predictions at an early stage of catchment gauging proofed to be very valuable while leading from almost no catchment understanding to a preliminary identification of control variables on flow and an

assessment of the effectiveness of the sampling design. The findings served as a guide for further field work where a greater emphasis was placed on sampling more dilute end members such as rain and snowmelt and also more concentrated end members such as deeper groundwater aquifers.

Chapter 3 presents the use of a larger geochemical tracer set to determine the runoff contributing sources of the upper Xilin watershed. I used the data set from the previous chapter that was extended in time and space and also included a larger variety of tracers to further improve the process understanding. The additional data were collected in two consecutive field campaigns which followed the suggestions in chapter 2 to include a larger variety of end members. The data were used in a more advanced technique called end member mixing analysis (EMMA) to identify and quantify the runoff contributing sources. The basic assumption of EMMA is that the stream water is a mixture of sources with fixed chemical compositions, the mixing chemistry is a linear process and the tracers are conservative. EMMA includes the performance of principal component analysis (PCA) to determine the amount of end members needed to explain runoff in the system and identify the end members among potentially sampled source waters. An over determined set of equations based on a least squares procedure is then solved to calculate the contributions of the respective end members. From a suite of 33 possible tracers including cations, anions, pH and EC, and the isotopes δD and $\delta^{18}O$, 7 conservative tracers were selected for further analysis. The 7 conservative tracers were Li, Rb, Sr, Na, Mg, Cl^- and EC. The isotopic signatures of end member waters were used as an additional tool to characterize isolated water storages. The specific objectives were (1) to identify runoff producing sources in the Xilin river catchment, (2) to conceptualize catchment processes in a semi-arid catchment and (3) to investigate the applicability of EMMA in semi-arid catchments. Our findings indicate that the runoff can be explained by three interannually changing end members. In wetter years, such as 2006 and 2008, the runoff is mainly produced by shallow groundwater sources. Dryer years exhibit a dominance of deeper groundwater aquifers as sources of stream flow. Rain is in all years only of minor importance. However, the 3-year measurement programme lacks the documentation of the full spectrum of hydrologic conditions, e.g. very wet years and very dry years and hence prompts the continued development of long-term measurement programmes.

Despite the limitations of EMMA, which are the assumptions of temporal and spatial invariance of the end members and which probably contribute a large part to the uncertainty in our models, this work is one step forward in improving catchment understanding in the Xilin river basin.

Chapter 4 addressed the problem of which and how many tracers to include into an EMMA and aimed at assessing the uncertainty of the model structure that results from an EMMA. This work benefitted from the large number of chemical constituents that could be easily analyzed with three different instruments, namely the Inductively Coupled Plasma - Mass Spectrometer (ICP-MS), Ion Chromatograph (ICP) and DLT 100 (Off-Axis Integrated Cavity Output Spectroscopy (OA-ICOS)), and served as tracers. The broad range of potential end members that were collected over the years were also of great value for this study. The specific objectives were (1) to find a tracer set threshold beyond which the acquisition of insight in terms of system dimensionality becomes negligible, (2) to evaluate the influence of the tracer set size and composition on the model concept that results from an EMMA (3) to analyze the compositions of tracer sets that produce acceptable results as defined by mixing model theory. We implemented an automatic procedure that iteratively conducted EMMA, varying the tracer set size and composition starting with 2 tracers up to 14 tracers. The 14 tracers were Li, Fe, Rb, Sr, U, Na, Mg, K, Ca, Cl^-, SO_4^{2-}, δD, $\delta^{18}O$ and EC. The procedure applied three criteria that were based on mixing model theory to identify those tracer sets that produce plausible results regarding end member combinations and contributions. With a total of 16.369 tracer set possibilities, we were able to show that the number of identified end members increases with increasing tracer set size. We found no evidence that there exists a tracer set threshold beyond which the acquisition of insight becomes negligible. Our findings indicate that the resulting model concept from an EMMA is highly sensitive to the tracers set size and composition in terms of system dimensionality (end members needed to explain flow in a system), end member combination and contribution to stream flow. Only 23 tracer sets produced plausible results. A certain consent exists concerning the dimensionality of a system with all remaining tracer sets exhibiting a dimensionality of 3 which is also in line with the results in chapter 3. However, findings were ambiguous in terms of identified end members: for 2006, the most often selected end members were a

shallow and a deep groundwater aquifer (G4 and G5) as well as either the headwater source area (H) or a tributary (T1). The large differences in end member contributions that were produced among the different tracer sets demonstrated that we are still missing at least one important end member. They also indicate that the major elements that are preferably used in many other studies are not always the most useful tracers. This study proofed that larger tracer sets that include also minor trace elements and that are employed in an iterative EMMA approach with varying tracer set sizes and compositions have a high potential to avoid false conclusions about catchment functioning. The proposed approach is highly complementary to the traditional EMMA approach.

(Please find Figure 1 – 5 on p. 14)

Figure 1 – 5. View across the study area near the upper watershed outlet with basaltic plateau in the background (a). The Chinese Academy of Sciences (CAS) maintains the Inner Mongolia Grassland Experimental Station (IMGERS) which is located in the left half of image (a). It belongs to the Chinese Ecological Research Network (CERN) and is managed by the Institute of Botany (IB-CAS). IMGERS is dedicated to research and education and holds an indispensable data set on ecological, meteorological and hydrological parameters. Image (b) is a view of the riparian area near the watershed outlet during winter and shows wind redistribution of snow in the vicinity of the river which is responsible for the spring discharge peak. Photographs (c) and (d) were taken during early and late vegetation period, respectively, highlighting the different hydrologic conditions expressed through the vitality of the vegetation. Image (e) shows the sparsely vegetated sand dunes. A rainfall collector is presented in image (f). The following images present the potential end members: (g) ephemeral tributary (T1) during no-flow conditions, (h) shallow groundwater well (G2) and (i) deep groundwater well (G3). Photographs (a), (d) and (h) are used with kind permission of M. Wiesmeier and photographs (g) and (i) are provided by K. Schneider.

Figure 1-5. Impressions of the Xilin river catchment. (Please see detailed figure caption on p.13.)

1.6 LIMITATIONS OF THIS STUDY

The major limitations of this research relate to the general conceptual approach. We used end member mixing analysis to identify and quantify the major runoff contributing sources and processes. One of the basic assumptions of EMMA is that the stream water is a mixture of its sources. The chemical fingerprint of a source is characterized by its immediate environment from which the elements leach into the water. By applying EMMA at a large scale we ran the risk that the chemical fingerprint of a distant end member is obscured in the stream water due to the long distance the water has to travel. On its way to the stream the contact time to the medium through which the water travels increases and with it the potential that more leaching occurs. This may change the unique chemical composition of the end member and destroy important information of the tracer signals.

Another basic assumption of EMMA is that the chemical composition of the stream sources is time invariant. Due to the large extent of the area and the limited resources of manpower some of the end members were sampled less frequently than others. The resulting relatively small sample size of certain end members hampered me from assessing their temporal variability. Hence, I was unable to evaluate if these end members support the assumption of temporal invariance.

Due to the remoteness of the study area and limited resources we had to deal with some technical constraints. Snow water chemistry could not be sampled for the inclusion of this end member into the EMMA. The EMMA results are therefore biased towards processes that occur after the onset of the long term snow discharge peak.

1.7 FUTURE RESEARCH

The dissertation research provided a glimpse of the complex hydrologic dynamics in a large semi-arid catchment. It also presents the possibilities environmental tracer data offer to improve process understanding in such large ungauged basins and suggests an automatic procedure as valuable complement to the traditional EMMA approach. Several perspectives exist on how to continue research in this field. The two main future research topics include (1) evaluating the significance of these results for hydrologic modeling and (2) application of this potential model to simulate the effects of land use change on the hydrology of the Xilin catchment.

The first perspective can be accomplished by using the newly obtained information to guide conceptual model development in a "soft data" manner as described by *Seibert and McDonnell* [2002]. In addition, it can be applied in a process-based rejectionist framework as *a posteriori* model calibration criteria [*Vaché and McDonnell*, 2006; *Vaché et al.*, 2004]. The latter builds upon the work of *Seibert and McDonnell* [2002] and suggests the use of soft data in order to calibrate the model and evaluate the acceptability of the model structure.

The information about geographic source contributions to stream flow that was accumulated in this dissertation may guide the development of a physically based reservoir model using the Catchment Modeling Framework (CMF) presented by *Kraft et al.* (under review). CMF is a model tool kit that is based on the rejectionist framework as described by *Vaché and McDonnell* [2006]. It offers the possibility to facilitate the design of different model layouts to test various hypotheses about runoff generation processes. Such an approach is able to support hypothesis testing to evaluate the newly acquired information and to gain more insight into the processes of the catchment. The hypotheses that arose from this dissertation and that should be tested for the upper Xilin watershed may be the following:

(1) The grassland zone is not connected with the groundwater aquifers which are responsible for stream runoff.
(2) The Xilin river catchment is not an open catchment, i.e. the topographic catchment area equals the effective catchment area.

(3) Flow between the various groundwater aquifers is not connected.

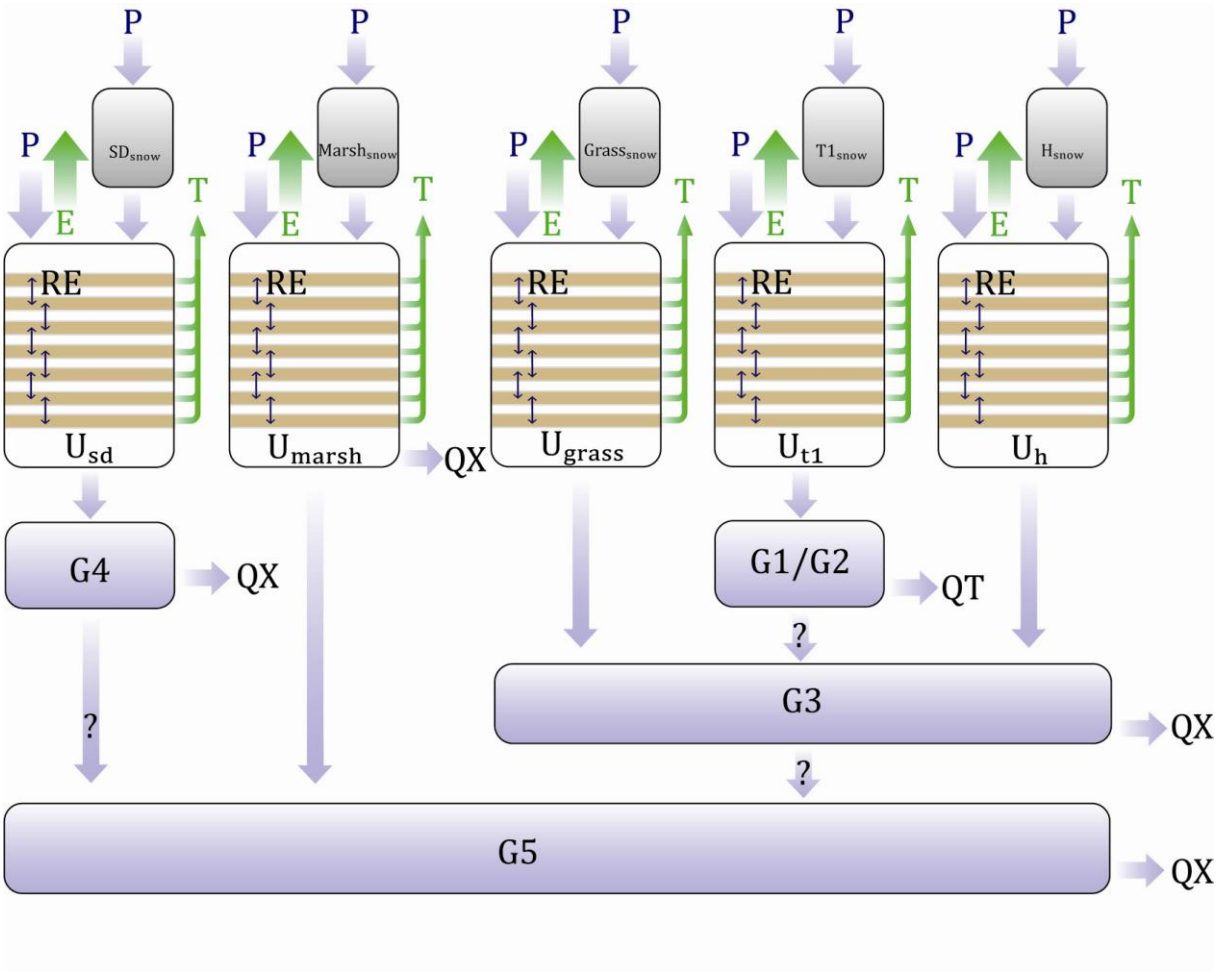

Figure 1-6. A conceptual reservoir model of the upper Xilin watershed consisting of 5 zones: SD = sand dunes, Marsh = marshland, Grass = grassland, T1 = tributary and H = headwater. P, E and T depict precipitation, evaporation and transpiration, respectively. Precipitation, that falls as snow (e.g. Marsh$_{snow}$), is modeled with an energy balance model. The brown boxes represent multiple layers of the unsaturated zone (U) which is modeled with the Richards Equation (RE). The arrows represent water flow between the various storages (e.g. G1-5 = groundwater storages 1-5), to the Xilin river (QX) and to the tributary (QT). The question marks (?) highlight the model processes to be tested as hypotheses.

Figure 1-6 depicts a potential conceptual reservoir model that was developed based on the information from this dissertation. It suggests dividing the watershed into 5 zones representing the headwater area, the tributary, marshland, grassland and sand dune zone. The precipitation that falls onto these zones dissipates through an unsaturated zone

that may consist of multiple layers, each layer representing a different depth. Each of the 5 zones then drains into either one of the groundwater aquifers or directly into the river where it produces the discharge. The discharge from groundwater storages is modeled as a linear function of the stored water. Snow melt and sublimation will be taken care of with an energy balance model. Evaporation and transpiration processes will be modeled with the Hargreaves method as suggested by *Schneider et al.* [2007]. Processes in the unsaturated zone will be treated with a Richards-Equations-approach.

This model then serves as a basis for testing the above mentioned hypotheses by slightly varying the model structure. The acquired information presented in this thesis can then be used as soft data during the calibration process by evaluating the simulated outflow from the various sources and compare these with the source contributions to stream flow that were calculated based on the tracer based approach. Furthermore, our knowledge about catchment processes will then serve as compositional criteria that may guide the decision on model rejection or acceptance.

Since both, the environmental tracer data and the selected analyses provide only a small piece of the puzzle which has to be combined with others to reveal a picture, future research should focus on field work that provides information on various other hydrologic aspects. This includes data collection that sheds light on the contributions of snow fall and melt water, on processes such as wind redistribution of snow and groundwater recharge, and is able to provide information about the effective catchment area. The already existing tracer data can be recycled once more in approaches to estimate the Mean Travel Time (MTT) or to establish stream chemical profiles in order to determine where the stream picks up significant amounts of discharge.

Once an acceptable model structure is found, the second major future research topic, simulating land use change effects on the hydrology of the Xilin catchment in order to provide information for land management decisions, may be pursued. Potential scenarios have been developed by *Schäfer* [2009] for application in the SWAT model. They include a base scenario, a production scenario and a sustainability scenario. The base and the production scenarios are primarily defined by the land use and the grazing intensities which are presented in Figure 1-7. Due to limited data, the sustainability scenario is only defined by grazing intensities and equals the base scenario in all other conditions. These

three scenarios need to be converted into measures that are of use for the conceptual reservoir model which is suggested in Figure 1-6. The model may then be used to simulate land use changes and their impacts on the hydrology of the system.

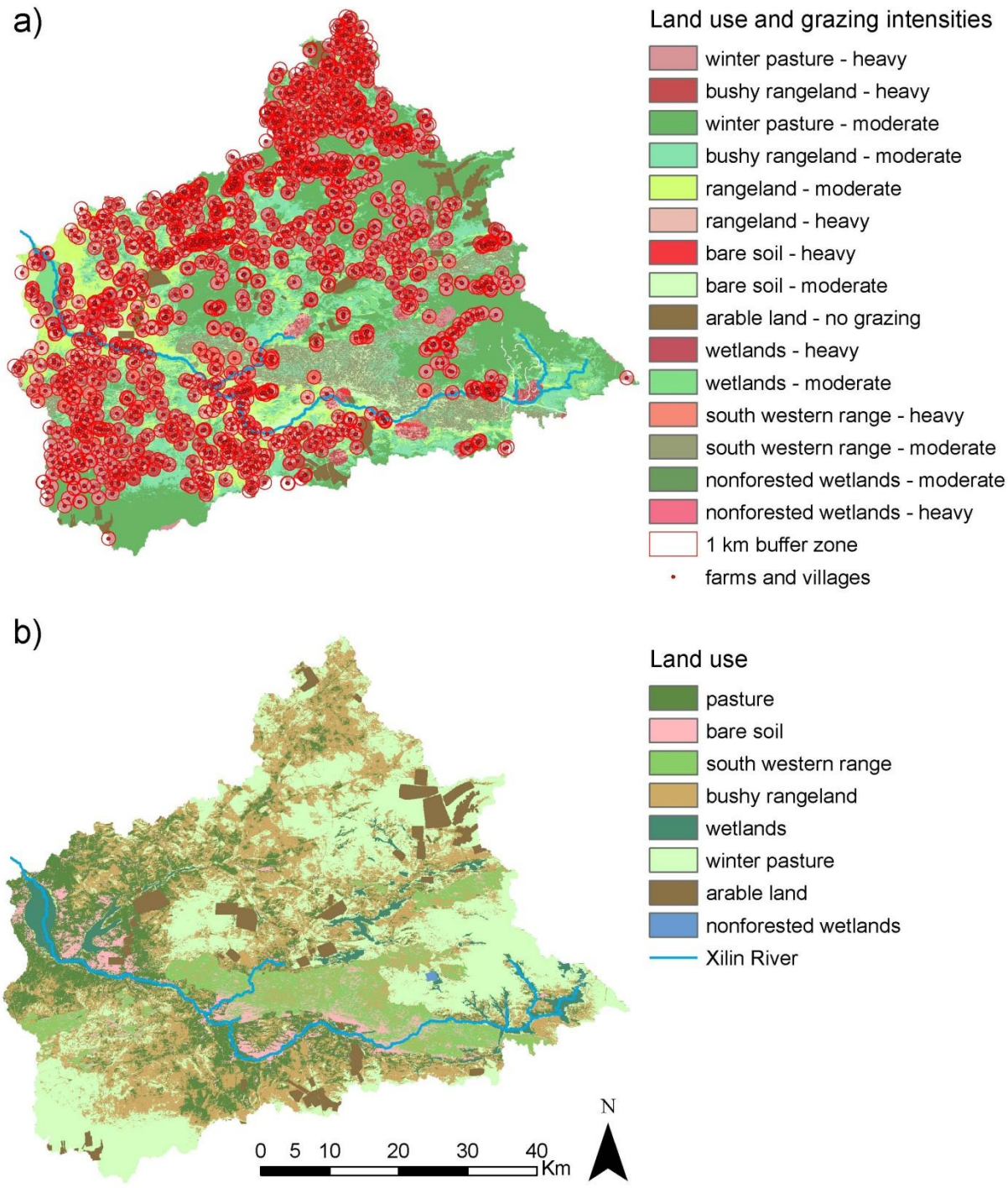

Figure 1-7. Land use and grazing intensities (a) of the base scenario and (b) of the production scenario [after *Schäfer*, 2009].

Last but not least, the results concerning the tracer based approach, i.e. the traditional EMMA approach, open another path for future perspectives. The uncertainty estimation of the model concept resulting from EMMA suggests conducting a simulation study with known end members and concentrations. Such a backward approach is likely to improve confidence about the size and the composition of the tracer sets.

2 GAUGING THE UNGAUGED BASIN: A TOP-DOWN APPROACH IN A LARGE SEMIARID WATERSHED IN CHINA

Abstract

A major research challenge in ungauged basins is to quickly assess the dominant hydrological processes of watersheds. In this paper we present a top-down approach from first field reconnaissance to perceptual model development, model conceptualization, evaluation, rejection and eventually, to a more substantial field campaign to build upon the initial modeling. This approach led us from an initial state where very little was known about catchment behavior towards a more complete view of catchment hydrological processes, including the preliminary identification of water sources and an assessment of the effectiveness of our sampling design.

With T. Sayama, K. Schneider, L. Breuer, K.B. Vaché, H.-G. Frede and J.J. McDonnell published in *Advances in Geosciences*, 18, 3-8, 2008.

2.1 INTRODUCTION

Gauging the ungauged basin is a major research challenge, in part so because we currently lack tools to quickly diagnose the dominant processes of watersheds for use in conceptual model development and prediction [*Sivapalan et al.*, 2003]. Usually, long-term datasets of hydrometric and hydrochemical information are needed to begin an evaluation of dominant runoff producing processes, however, in many catchments, these data are not available.

Conceptual models have a set of traits that can be advantageous in the rapid identification of runoff generating processes. They (1) can capture processes at the catchment scale (2) offer the potential of model development using information on age, origin and pathway of surface and subsurface storm flow, (3) can utilize known physical parameters (4) allow identification of parameters values through calibration against runoff [*Seibert and McDonnell*, 2002]. Conceptual models that treat volume-based mixing and mass balance are useful for streamflow modeling with geochemical dimension. Combining conceptual models and data from first field reconnaissance is an approach to begin the development of a more complete understanding of hydrological processes in previously ungauged basins.

In this study we follow a road map (Figure 2-1) for how one might start the process in a large, ungauged basin in Inner Mongolia, with identification of simple hydrological patterns as a first step in predicting flow in this data sparse region. We designed this process as a top-down approach, as initially introduced by *Klemeš* [1983], consisting of the following five steps:

(1) Field reconnaissance and data collection,

(2) Perceptual model development,

(3) Reservoir model conceptualization,

(4) Evaluation with hydrochemical data and

(5) Complete field campaign (Figure 2-1).

The objective of this study is to use geochemistry-based field reconnaissance to assess dominant processes of water cycling within the mesoscale watershed, develop a conceptual reservoir model, evaluate and reject the model in order to yield new insight into catchment functioning that may guide a more complete field work. The approach may be adopted as a road map in other ungauged or data sparse regions where a quick assessment of streamflow generating processes is necessary.

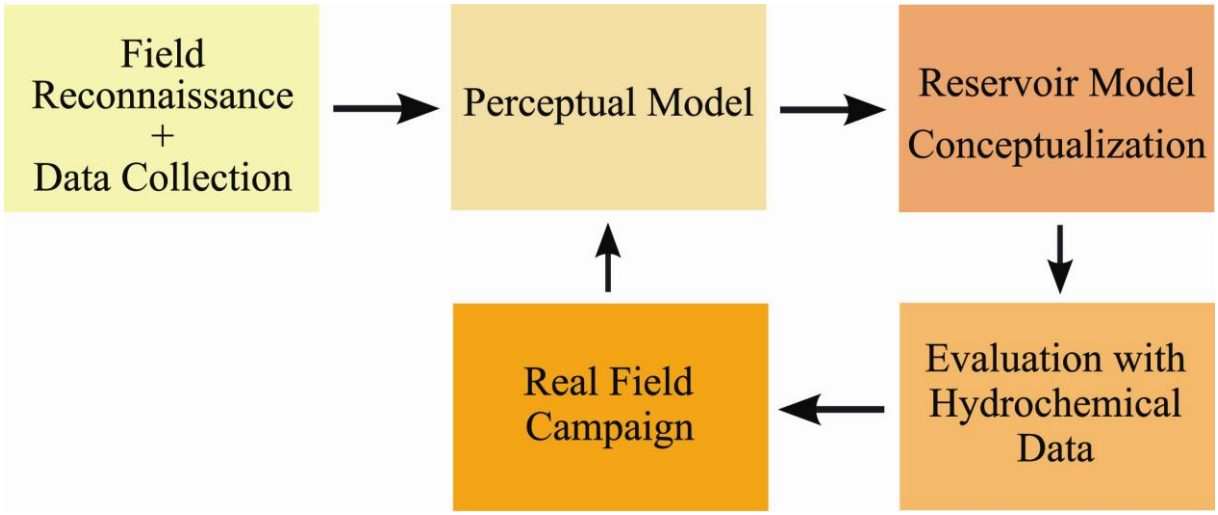

Figure 2-1. Flow chart of top-down approach.

2.2 Material and Methods

2.2.1 Study area

The study catchment is a 475 km^2 sized subbasin of the Xilin river watershed which is located at approximately 43° N and 116° E in the autonomous province of Inner Mongolia, China (Figure 2-2). The Xilin river watershed has a size of about 10,000 km^2 [*Tong et al.*, 2004] and is an inland river basin which is characterized by semiarid conditions with cold, dry winters and warm, wet summers. The mean annual precipitation is 350 mm but is highly variable in space and time. *Z. Chen* [1988] reported annual ranges of 180 to 500 mm, with 60 to 80 % of the annual precipitation falling between June and August. The mean annual air temperature is 2°C with a July average of 18°C and a January average of -23°C [*Z. Chen*, 1988]. Mean actual evapotranspiration during the vegetation period is

larger than 90 % of the precipitation [*Wen, Y.*, personal communication] and lower throughout the rest of the year.

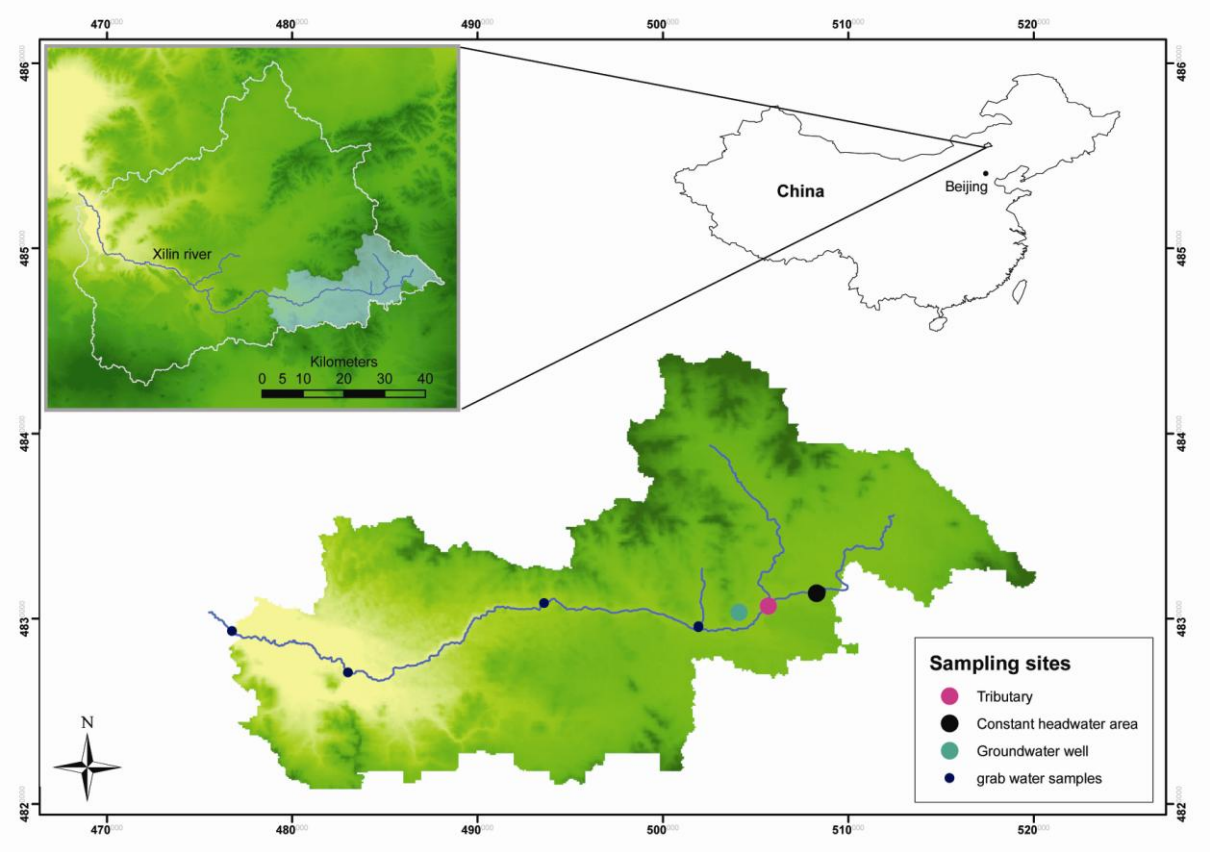

Figure 2-2. Digital elevation model of the Xilin river watershed and subbasin. Inset: Xilin river location in China.

2.2.2 FIELD DATA COLLECTION

Snapshot sampling [*Grayson et al.*, 1997] was used to collect hydrometric and chemical data. Grab samples of water were collected from a variety of locations at approximately the same time, and analyzed with an inductively coupled plasma – mass spectrometer (ICP-MS) for a suite of 20 major and minor chemical tracers (Li, B, Al, V, Cr, Mn, Fe, Co, Ni, Cu, Rb, Sr, Mo, Cd, Ba, U, Na, Mg, Ca, K). Field data collection in the Xilin river watershed is limited to the time period between April and late October, while the water of the river is unfrozen. Therefore, all datasets are limited to this time period. Although initial field reconnaissance of the watershed started in 2004, we exclusively used the most detailed data set collected during the field period of 2006 for this study. We applied simple explorative data analysis to identify possible end members, as common in mixing analysis. End members are defined as source solutions that have more extreme chemical

concentrations than the stream water [*Christophersen and Hooper*, 1992]. In theory, their mixture yields the hydrochemical concentration of the water in the main stream of the catchment. The identification of end members can be used to contribute to a more complete understanding of the different processes which contribute to the generation of streamflow.

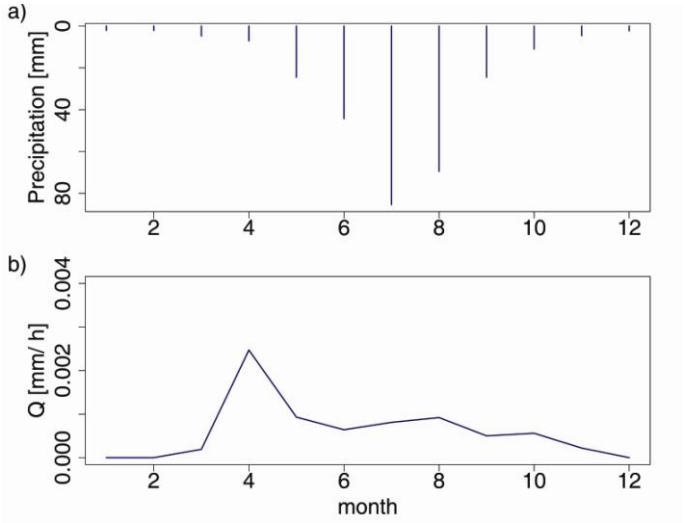

Figure 2-3. Mean monthly precipitation (a) and mean monthly discharge (b) for the whole Xilin river watershed (1954-2004).

Precipitation was measured with a tipping bucket rain gauge (RM Young, Traverse City, Michigan, USA) and recorded in 10 min time steps [*Ketzer et al.*, 2008]. Evapotranspiration was assumed constant during the vegetation period and estimated as 90 % of the 2006 vegetation season precipitation [*Wen, Y.*, personal communication]. The hydrograph does not reflect the annual precipitation pattern: despite low snow rates during the winter months (November through March) vernal discharge reaches highest values during the melting period, whereas the precipitation peak in June and July only results in a secondary, only minor discharge peak towards the end of the vegetation period in August (Figure 2-3).

2.2.3 MODEL DESCRIPTION

We developed a simple lumped reservoir model based on the Tank Model, to represent the different water stores [*Suguwara*, 1961, 1995]. As in the Tank Model, each reservoir represents a different component of the catchment. The storage depth of each reservoir

is calculated as the difference between precipitation input (R in [mm/h]) and loss via evapotranspiration (E in [mm/h]) and calculated discharge (Q in [mm/h]). The output of each reservoir is the calculated runoff. Our lumped model consists of three reservoirs: S1, S2, and S3 (Figure 2-4). The reservoir model assumes a linear relationship between storage depth and output for the headwater source and a non-linear relationship for the tributary and the groundwater source. The various discharge components are calculated as follows:

Discharge to stream from headwater source (Q1)

(1) $$Q1 = a1 * (S1 - h1)^{m1}$$

Groundwater recharge from headwater source (Q13)

(2) $$Q13 = a13 * (S1 - h13)^{m13}$$

Discharge to stream from tributary (Q2)

(3) $$Q2 = a2 * (S2 - h2)^{m2}$$

Groundwater recharge from tributary (Q23)

(4) $$Q23 = a23 * (S2 - h23)^{m23}$$

Discharge to stream from groundwater (Q3)

(5) $$Q3 = a3 * (S3 - h3)^{m3}$$

where a and m are two calibrated conceptual model parameters which remain constant, S is the storage depth of the reservoir [mm], h is the threshold height for water to start flowing [mm] and 1, 2 and 3 indicate the different reservoirs. Total stream discharge is calculated as the sum of the discharge from the headwater source area, the tributary and

the groundwater, each weighted by contributing area. The area weight representing the headwater and tributary sources was assumed to be equivalent to the land surface areas of each (0.18 and 0.82 of the total area, respectively). We further assumed that the contributing area of the groundwater source was equivalent to the entire catchment area, giving it a weight of 1.0.

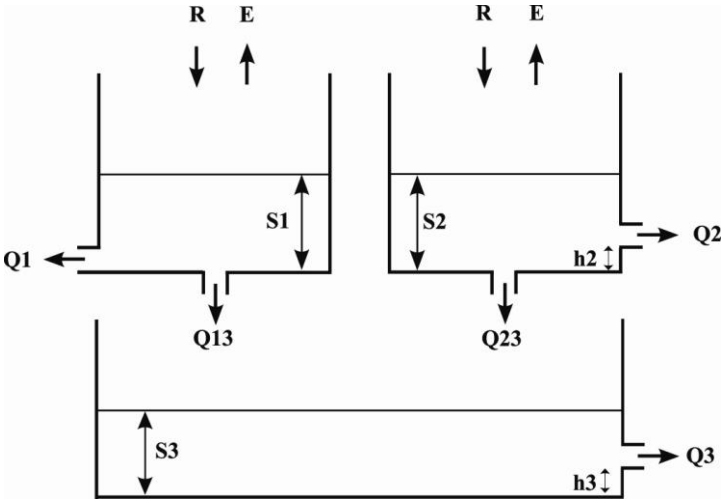

Figure 2-4. Lumped conceptual model of the Xilin river catchment with R=rainfall, E=evapotranspiration, Q=discharge, S=storage depth and h=threshold height for water to start flowing. The discharge indices refer to the type of discharge: single number means discharge to stream, composite number means groundwater recharge.

2.3 TOP-DOWN APPROACH

In the following we describe our top-down approach that consists of five steps (Figure 2-1). This approach generally follows the "downward route" - concept as introduced by *Klemeš* [1983], from field data to conceptual model development and testing.

2.3.1 STEP 1: FIELD RECONNAISSANCE AND DATA COLLECTION

Initial field reconnaissance of the Xilin River Basin started in 2004 and was focused on discharge measurements and grab water collection after the snapshot sampling method [*Grayson et al.*, 1997]. Specifications for data collection are given in section 2.2. A detailed, for our purposes useful dataset consisting of discharge and precipitation measurements as well as hydrochemical information of various water sources, could be established during the field season of 2006.

2.3.2 STEP 2: PERCEPTUAL MODEL DEVELOPMENT

Bivariate plots of Na and K concentration were used to identify potential end members of streamflow (Figure 2-5). Na and K have widely been used for hydrograph separation [e.g. *Elsenbeer et al.*, 1995; *Hill*, 1993; *Neal et al.*, 1997] and were chosen for these plots because they most clearly identified water sources. Weekly streamflow, a 4 m deep groundwater well, a tributary and streamwater at the headwater source area taken during the field season 2006 are presented in Figure 2-5.

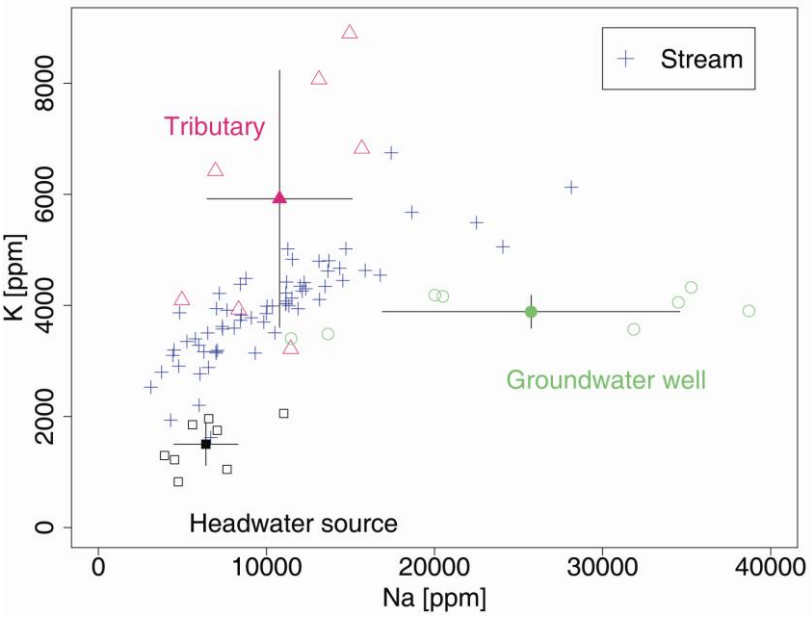

Figure 2-5. Bivariate plot of Na and K concentration of water samples of end members and the main stream, bold symbols represent the median and 95 % confidence interval of the end members.

The headwater source area, the groundwater well and the tributary were identified as end members which serve as anchor points for construction of the conceptual reservoir model (Figure 2-4). The stream concentrations are low at the beginning of the season, increase over the course of the season, peak in August and slightly decrease in September (Figure 2-5 and Figure 2-7). These data suggest a perceptual model of discharge development where stream flow is initially dominated by the headwater source. As the vegetation growth season progresses with increased rainfall and evapotranspiration contributions of water from the ephemeral tributary and the groundwater source begin

to increase, until in September when the tributary runs dry and the wetland area at the headwater source is saturated with water and again turns into the dominant contributor.

2.3.3 STEP 3: RESERVOIR MODEL CONCEPTUALIZATION

Based on the end members, the conceptual reservoir model was developed to capture the evolution of stream water sources as consistent with our perceptual model, and the data we used to develop it. The model was applied to simulate total stream discharge as well as the discharges of each of the end members (Figure 2-6). Calibration was conducted on a visual basis of simulated and observed discharges in the main stream (Figure 2-6b). The parameter values resulting from the calibration exercise are listed in Table 2-1.

	Unit	Q1	Q13	Q2	Q23	Q3
a	-	0.007	0.0002	0.00001	0.00002	0.001
Initial S	mm	4.5	4.5	0	0	11.5
h	mm	0	0	10	0	2
m	-	1	1	1.5	1	2

Table 2-1. Values of model parameters for discharge components of the conceptual reservoir model: a and m are two conceptual model parameters, S is storage depth of the reservoir and h is the threshold height for water to start flowing. Values were derived from manual model calibration.

The model captured the downward trend of the discharge at the beginning of the summer season and the major flow peaks, although the presented calibration includes a general overestimation of discharge throughout the growing season. Despite the overestimation, the simulated storm hydrographs of the end members clearly show the trend of the observed hydrographs: increase of the constant headwater source throughout the season and start of flow of the ephemeral tributary at the beginning of June (Figure 2-6c). The poor performance in the recession of the hydrographs is likely attributable to the linear storage concept. The model was able to produce a better fit for the discharge during the growing season, but this improvement resulted in deterioration in the fit during the snowmelt recession. While accomplished without the aid of an automatic multi-criteria calibration procedure, this inability to capture both the snowmelt

recession and the growing season discharge suggests either a structural flaw or the presence of an unidentified end member.

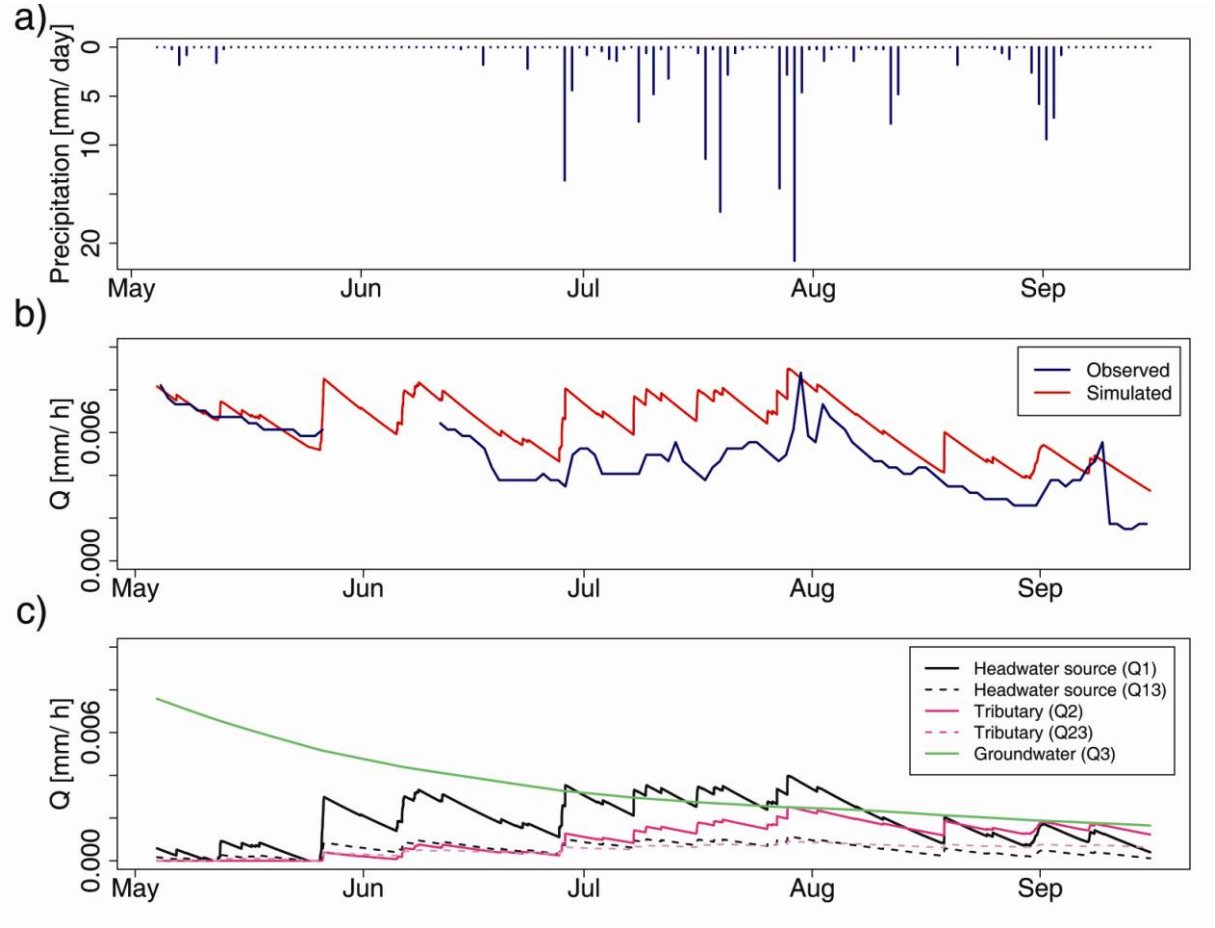

Figure 2-6. Precipitation (a), observed and simulated hydrographs of the stream (b) and simulated hydrographs of the end members (c) of the 2006 vegetation period.

2.3.4 Step 4: Evaluation using hydrochemical data and rejection of initial model

The key interest in this study lies in identifying the dominant runoff producing processes. Therefore, the following questions arise: Why does the model work during the snowmelt recession and not during the rest of the season? Is the groundwater really the source for the receding hydrograph during the spring period and does the ephemeral tributary account for the minor discharge peak in the late vegetation season (Figure 2-6c)?

In this top down approach, where quick diagnosis of catchment functioning is the focal point, we view the step of model evaluation and rejection as a crucial step for any further model development. At first glance, the most compelling reason for the inability to

capture the discharge during the vegetation period is a structural flaw resulting in an underestimation of actual evapotranspiration during the vegetation period (Figure 2-6b) since simulated and observed discharge match during the period of snowmelt at the start of the summer season but reveal overestimation of simulated discharge during the rest of the summer season. However, the simulated discharge in Figure 2-6b is the sum of the discharges of the three different end members headwater source, tributary and groundwater in Figure 2-6c. The apportionment of the discharges of our three different end members in the latter figure shows that the simulated discharge during spring is mainly controlled by the groundwater hydrograph. The time series of the K and Na concentrations (Figure 2-5 and Figure 2-7), however, suggest a dominant discharge contributing role of the headwater source at the beginning of the season. It is this contradiction that leads us to the assumption that our model is structurally correct, but that our data do not capture the true end members of the system.

	unit	Headwater source	Tributary	Groundwater	Modified groundwater
Na	ppm	5601.5	10780.7	22883.6	1000.0
K	ppm	1361.2	5919.8	3719.8	1000.0

Table 2-2. Mean concentrations of sodium (Na) and potassium (K) in waters of the three end members which were used as input parameters for model evaluation, and modified values of the groundwater source as input for the virtual experiment based on the model reevaluation.

We then utilized chemical data for evaluation of our conceptual reservoir model. The seasonal means of Na and K of each end member were used as inputs (Table 2-2). The K and Na concentrations of the stream were simulated as the sum of the area weighted discharges, as defined previously, of the end members over the course of the vegetation period. Figure 2-7 presents the simulated seasonal stream concentration which shows the converse trend of the observed bivariate plot with high concentrations at the beginning of the season, decreasing concentration throughout the season and again increasing concentrations at the end of the season.

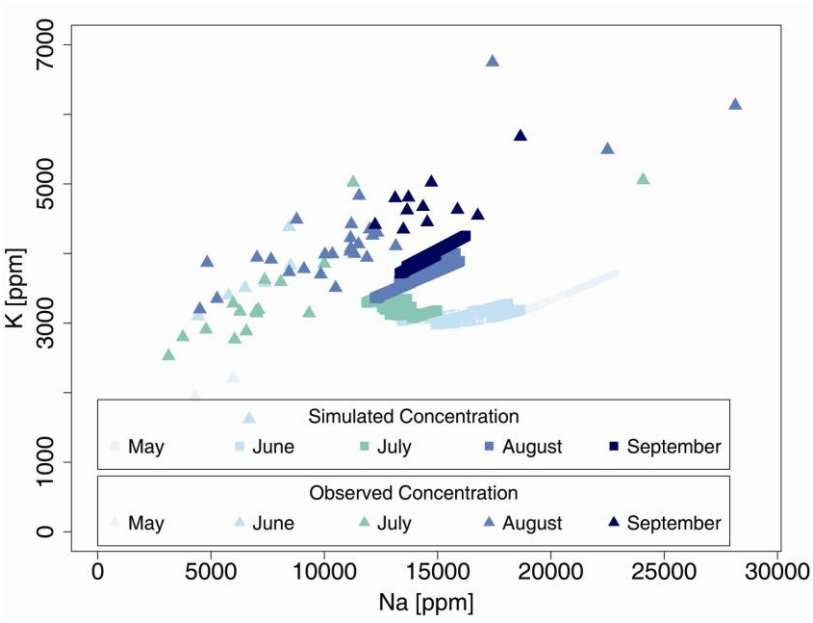

Figure 2-7. Bivariate plot of observed and simulated Na and K concentration of the stream in 2006 following the initial model.

Since the model does not reflect the expected chemical behavior we have to reject our initial model perception. We assume that our model does not capture the important end members that contribute to stream flow. Virtual experiments were used in an initial exploration of these missing links in our conceptual model.

As a first step, it is important to recognize that our conceptual reservoir model relies upon the assumption that the mean values of Na and K were representative of the end member concentrations. This simplification was initially used because we lacked any compelling evidence for another choice, but it is clear that the end member chemistry will, in fact, vary throughout the measured range. As an initial re-assessment of the model we focus upon this range of variation, rather than the model structure itself, and evaluate how changing the end member concentrations, within measured ranges, might result in changes to the mixture chemistry. Specifically, we adjusted the groundwater end member positions at the beginning of the season to smaller values to account for poorly defined water sources, but again did not change the model structure or parameterization (Table 2-2). The adjustment of the end member input concentrations yields the possible chemical characteristic of a newly defined end member because the simulation of modified end-members does capture the trend of the observed stream data with low

values at the beginning of the season with increasing values over the course of the season (Figure 2-8). The modification suggests the existence of a new end member with low concentrations. The most likely candidate for this additional end member is snow melt, which will now guide additional field work in this iterative approach to watershed diagnosis.

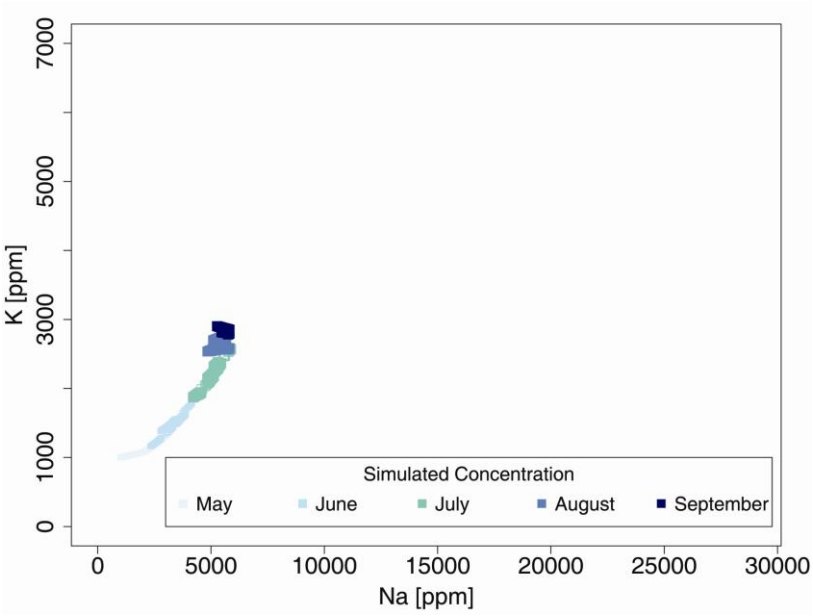

Figure 2-8. Bivariate plot of simulated Na and K concentration of the stream in 2006 with modified groundwater concentration input.

2.3.5 STEP 5: REAL FIELD CAMPAIGN

Our results of the virtual experiments lead us to the conclusion that a water source with low K and Na concentrations is a missing end member in our conceptual model as well as another end member with high concentrations. Snowmelt and an additional groundwater source would meet these requirements. These theories will guide further field campaigns where we will include snowmelt and deeper groundwater sources in our sampling scheme.

2.4 CONCLUSIONS

Large, poorly gauged basins present a major research challenge for understanding catchment processes [*Sivapalan et al.*, 2003]. Our study area, the Xilin river catchment, is such a poorly gauged basin. Initial field reconnaissance of the watershed started in 2004 and more detailed data collection during the vegetation periods focusing on identification of end members that control stream generation processes followed in 2006. Attempts to use these data in hydrologic model development have failed so far [*Schneider et al.*, 2007]. This is likely due to the fact that the major discharge peak is created by snowmelt during spring and only a secondary, minor discharge peak occurs during the vegetation period in the summer months. The top-down approach presented in this study, starting at first field reconnaissance, to the perceptual model and conceptualized model to rejection and revision, has been implemented in an effort to learn from wrong predictions at an early stage of catchment gauging and understanding as demanded by *Sivapalan et al.* [2003]. It led us from almost no catchment understanding to the preliminary identification of control variables on flow and an assessment of the effectiveness of our sampling design. Our next steps are to conduct new sampling to define a likely snowmelt and deeper groundwater aquifer end members responsible for flow generation.

3 IDENTIFICATION OF GEOGRAPHIC RUNOFF SOURCES IN A DATA SPARSE REGION: HYDROLOGICAL PROCESSES AND THE LIMITATIONS OF TRACER BASED APPROACHES

Abstract

There exists a growing need to improve our understanding of catchment processes on larger scales. This need is especially enhanced in rapidly developing countries such as China where the turnover of traditionally used land to urbanized and industrialized areas influences water resources. In this study, we chose a multi tracer approach in a large, ungauged basin in Inner Mongolia to elucidate the processes that generate stream flow. We applied End Member Mixing Analysis (EMMA) to identify and quantify the major runoff generating sources in a three end member system. Stream water and a set of 8 possible end members were sampled throughout 3 consecutive vegetation periods. Samples were analysed with an inductively coupled plasma-mass spectrometry (ICP-MS) and an ion chromatograph (IC) for a suite of 33 solutes including 24 cations and 7 anions. We determined that 7 tracers exhibited conservative behavior, they were Li, Rb, Sr, Na, Mg, Cl and electrical conductivity (EC). Stable water isotopes (δD and $\delta^{18}O$) highlight differences between the end members. Our results indicate strong interannual variability of end member composition and contribution between the three years. We were able to identify shallow ground water aquifers as important runoff generating sources in some years and deeper ground water aquifers in other years which vary in rainfall and discharge. A shallow sand dune aquifer which covers a significant part of the area plays an important role in storing and contributing water to the river. Our results also suggest that the major runoff generating source in the dryer year has not yet been identified. Our results prompt us to focus future work on understanding interannual changes in end member contribution especially in semi-arid regions.

With J. Wu, K.B. Vaché, K. Schneider, H.-G. Frede and L. Breuer in press in *Hydrological Processes*, 2010.

3.1 INTRODUCTION

Prediction in large ungauged basins is a major research area in the hydrological sciences, with much of the focus on better understanding how landscape heterogeneity modifies hydrological processes [*McDonnell et al.*, 2007] and how relationships between heterogeneity and process might be generalized for use in ungauged basins [*Tetzlaff and Soulsby*, 2008]. This focus is particularly relevant in larger basins, where the simple act of increasing the area of study is likely to both increase the heterogeneity of the system and at the same time decrease density of measurement points available to characterize it. In spite of this challenge, there is a clear need to develop predictive capabilities related to the identification of runoff generating sources and processes in large ungauged basins, particularly in emergent countries such as China [*D. Yang et al.*, 2005]. In part this need is driven by growth: the developing industry needed to support the growth rapidly alters the traditional use of land. In China, extensive natural landscapes and grasslands have been modified by the expansion of cultivated as well as urbanized and industrialized land [*Lin and Ho*, 2003]. Large scale landscape alteration has the potential to significantly influence the quality and availability of water resources upon which local populations depend [*S. Zhao et al.*, 2006], and understanding hydrological processes can contribute significantly to our ability to evaluate potential tradeoffs between development and the availability of water.

This need was recognized by the International Association of Hydrological Sciences (IAHS) and acknowledged by launching the Predictions for Ungauged Basins (PUB) Initiative with the aim to develop frameworks, methods and tools for application in such situations [*Sivapalan et al.*, 2003]. For larger catchments, one promising approach is to conceptualize catchment processes and to then use a generalized version of the resulting conceptual model of catchment process as a projection to provide prediction in ungauged basins [*Tetzlaff and Soulsby*, 2008]. Predictive uncertainty is the primary impediment to progress in this area, but continued progress is being made to more fully quantify uncertainty and more fully explore its implications.

One common tool to investigate and understand catchment functioning is End Member Mixing Analysis (EMMA) [*Christophersen and Hooper*, 1992]. EMMA is a widely accepted approach to identify runoff sources and flux components and calculate their contributions

to the stream discharge. It has been applied in many studies to identify end members at small catchment scales that describe the vertical sequence of water storages to flow contribution. These vertical end members are, for example, rain, soil water and groundwater [*Chaves et al.*, 2008], or overland flow, soil water and hillslope water [*Elsenbeer et al.*, 1995]. Fewer studies have applied EMMA at larger catchment scales of hundreds or even thousands of square kilometers [*Acuna and Dahm*, 2007; *Soulsby, Rodgers, et al.*, 2003; *Wade et al.*, 1999]. The challenge on larger scales is that landscape heterogeneity and catchment complexity both increase with size which may complicate the identification of well defined end members in the mixture [*Fröhlich et al.*, 2008]. Less studies have also tried to identify the horizontal contribution of end members to runoff generation in the sense of water provenance [*Acuna and Dahm*, 2007; *Fröhlich et al.*, 2007; *Ladouche et al.*, 2001; *Soulsby, Rodgers, et al.*, 2003]. However, this latter approach has potential to make significant contributions to a more complete understanding of catchment processes. The main assumption of EMMA is that the stream water is a mixture of compositionally distinct sources. The main constraints of the method are two additional assumptions that are required to develop the mathematics: 1) that the mixing processes are linear and 2) that the tracers are conservative, i.e. they do not participate in any adsorption nor biological processes and they are time invariant [*Hooper*, 2001]. End member mixing analysis techniques have been applied in a variety of studies, at both the plot and the catchment scales. These studies were primarily developed at sites in the temperate zone [*James and Roulet*, 2006; *Katsuyama et al.*, 2009; *Soulsby, Petry, et al.*, 2003], with a smaller number representing tropical regions [*Chaves et al.*, 2008; *Elsenbeer et al.*, 1995; *Hooper*, 2003]. Only few studies have applied EMMA in arid or semi-arid regions [*F. J. Liu et al.*, 2008].

In arid or semi-arid regions, water is a key limiting factor for biomass production. A better understanding of runoff-generation processes as well as catchment functioning is a key for improved water resources management. EMMA is one promising tool to obtain this process understanding. In Inner Mongolia, as in many other semi-arid and arid steppe environments of the world, a combination of the extreme climatic conditions, highly variable precipitation [*S. Q. Chen et al.*, 2002], and overgrazing, leads to a high risk of desertification [*Zheng et al.*, 2006]. The research area of this study, the Xilin river basin, is

located in the central part of Inner Mongolia. Here, more than 70% of the natural steppe has been reported as degraded [*Tong et al.*, 2004]. Apart from the pressure associated with overgrazing, recent land use development strategies have focused on the conversion of vast areas of grassland into cropland [*S. Q. Chen et al.*, 2002]. A future increase in the extent of cropland in the Xilin river catchment is likely, and will include additional pressure on limited water resources as irrigation requirements increase as has been observed elsewhere in Inner Mongolia [*Brogaard and X. Y. Zhao*, 2002; *Runnström*, 2003]. Hence, an improved understanding of the regional water budget is of utmost importance for a sustainable management of water resources.

In a first attempt to more fully quantify the water budget in the region, we investigate runoff generation and the spatial distribution of water source areas in a subcatchment of the Xilin river basin. Our methodology focuses on the development of a cost-effective and portable system of basic hydrologic measurements, which includes spatially distributed stream discharge measurements as well as isotopic and geochemical sampling and analysis, developed over a multi-year period. This dataset, representing the measured variability across multiple years, served as the basis for EMMA.

The objectives of this study are (1) to identify runoff producing sources in the Xilin river catchment, (2) to conceptualize catchment processes in a semi-arid catchment and (3) to investigate the applicability of EMMA in semi-arid catchments. We started to use EMMA year per year to test whether the observed process understanding is transferable in time. Hence, we applied EMMA on separate data sets and checked the comparability of findings. These results are designed to facilitate the development of more complete understanding of the processes producing runoff in the basin.

3.2 STUDY AREA

3.2.1 PHYSICAL CHARACTERIZATION

The study area is a subcatchment of the 10,000 km^2 large Xilin River Basin (43°24’ to 44°40’ N and 115°20’ to 117°13’ E) which is located in central Inner Mongolia A.P., China (Figure 3-1a and b). The subcatchment (Figure 3-1c) comprises about 475 km^2. The elevation ranges between 1175 m and 1500 m.

Inner Mongolia belongs to the Eurasian steppe. Its primary natural resources are grasslands which cover more than 70 % of the area (870,000 km^2) [*Kobayashi et al.*, 1994]. The Xilin River Basin represents the characteristic features of the grasslands in Inner Mongolia. Figure 3-2a shows that 40.5 % of our study area is covered with steppe. In higher elevated areas in the north and east of the catchment, the vegetation changes from steppe to mountain meadows. Marshland dominates in the vicinity of the river and its tributaries. Striking features of the catchment are paleo sand dunes that stretch through the center of the catchment and cover about 19.4 % of the area (Figure 3-2a). A substantial and considerably increasing part of the area is also used for cultivation of crops such as maize, wheat and rapeseed [*Guo et al.*, 2004]. This is due to the generally good soil conditions, despite unfavorable climatic conditions and high risk for wind erosion.

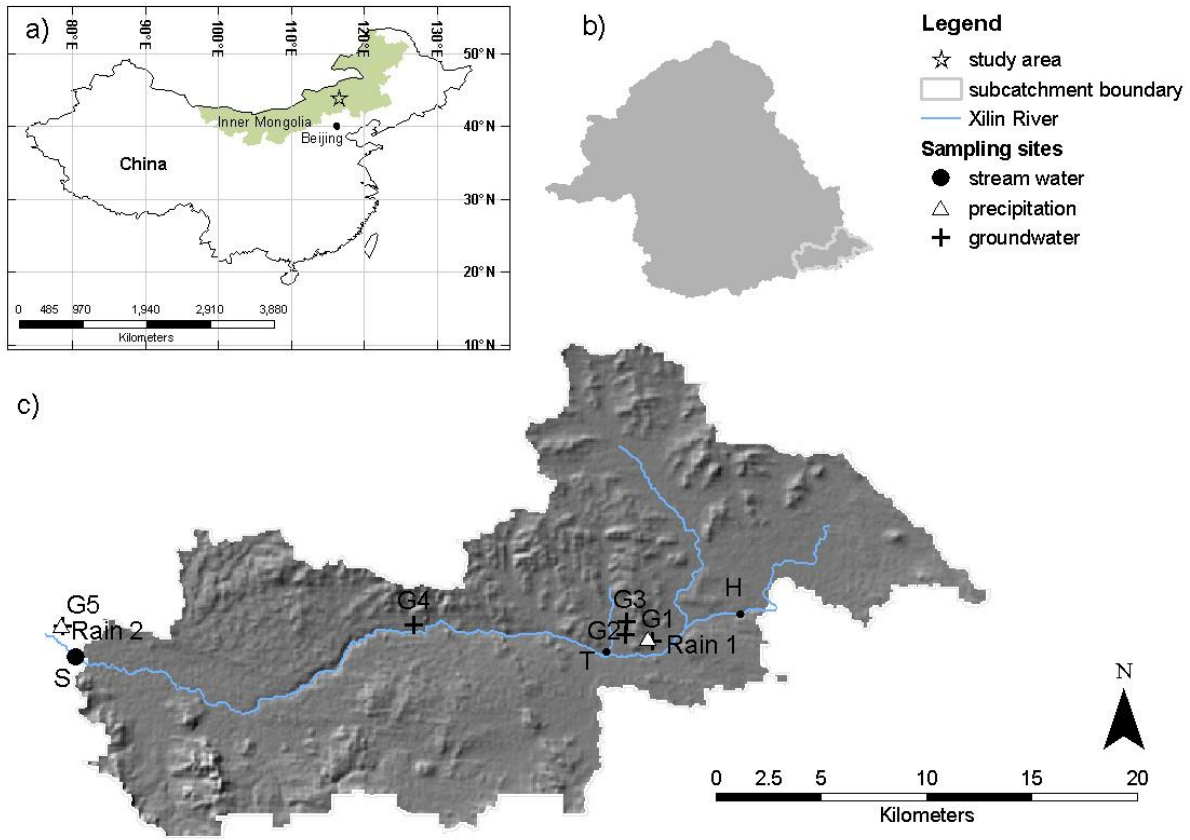

Figure 3-1. The maps show (a) the location of the Xilin River Basin in Inner Mongolia, China, (b) the extent of the Xilin River Basin and the location of the study area and (c) hillshade of the subcatchment derived from the SRTM digital elevation model. The Xilin river and the locations of the sampling sites are superimposed (G1-G5 = groundwater sites, H = headwater, T = tributary, Rain1 and Rain2 = precipitation sites and S = stream at outlet).

Figure 3-2b shows that the dominant soil types of the study area are the typical steppe soils Phaeozems, Chernozems and Kastanozems [*IUSS Working Group WRB*, 2007]. Arenosols cover the sand dune area and Gleysols are developed in wet areas near the river and its tributaries. Calcisols, Cryosols and Regosols occur mainly in the vicinity of the Gleysols. A more detailed description of the soils is given in *Barthold et al.* (under review).

The geological setting of the catchment comprises Quaternary deposits, sediments as well as igneous and volcanic rocks of different ages (Figure 3-2c). The oldest systems which form the basis of the area originate from the Paleozoic and comprise a variety of igneous rocks (Granites, Granodiorites and Diorites) and sediments. Some of these oldest areas are capped by volcanic rocks from the Mesozoic and together they form low mountains in the present-day landscape [*Z. Chen*, 1988; *Tong et al.*, 2004]. They mainly outcrop on the northern and southern borders of the catchment while the largest area is covered by the Quaternary deposits which are distributed along the river. Landforms such as sandy lands [*Z. Chen*, 1988; *Tong et al.*, 2004] and hills characterize the present-day landscape. A more detailed description of the geology is given in *Barthold et al.* (under review).

Year	Time of sampling	Number of days	Daily mean runoff [m^3/s]	Number of days with rain	Total precipitation [mm]	Total precipitation[v] [mm]
2006	May 4th – Sep 9th	128	0.48	53	255.5	256.3
2007	Jun 6th – Sep 20th	104	0.25	34	146.0	190.4
2008	Apr 26th – Sep 8th	135	0.61	54	291.7	294.3

Table 3-1. Precipitation and runoff characteristics during the sampling seasons and total precipitation of the vegetation period (v = May 1st until September 15th).

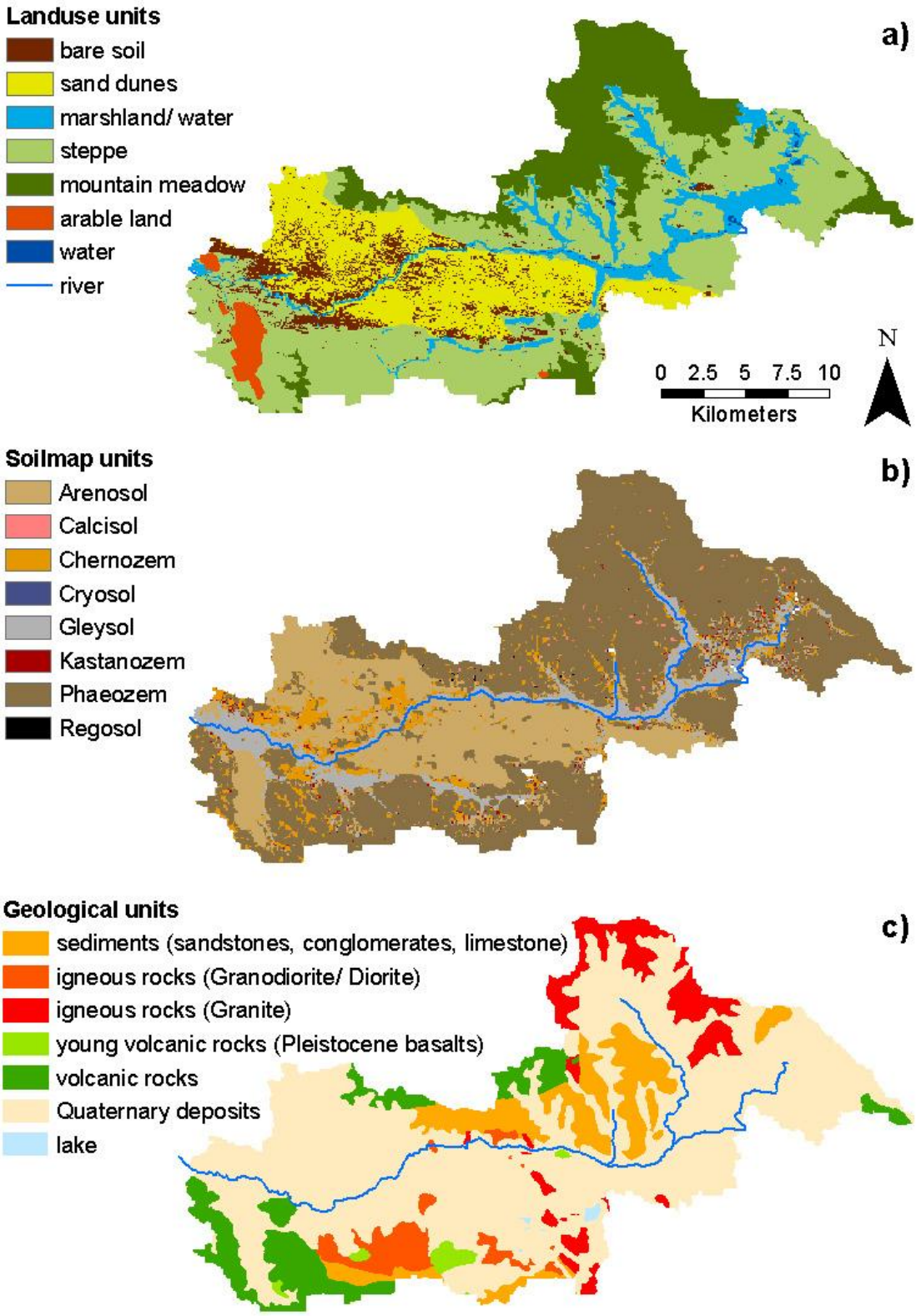

Figure 3-2. Landuse (a), soils (b) and geology (c) in the Xilin subcachtment. Landuse was classified based on a Landsat TM7 image from August 17th, 2005. The distribution of soils is a section from a digital soil map [*Barthold et al.,* in review]. The 1:200,000 geological map of the Inner Mongolian Bureau of Geology 1973 was modified and information was lumped into 9 new geological map units based on formation processes and age.

3.2.2 HYDROLOGY

The Xilin River is an endorheic river system. It is characterized by a semi-arid continental climate with cold, dry winters and warm, relatively wet summers. The mean annual precipitation is 350 mm but is highly variable in space and time due to prevailing convective weather conditions. Chen [1988] reported annual ranges of precipitation between 150 and 500 mm, with 60 to 80 % falling between June and August. The mean annual temperature is 2°C, with a January average of -23°C and a July average of 18°C [*Z. Chen*, 1988]. Mean actual evapotranspiration during the vegetation period [*Bai et al.*, 2004] is larger than 90 % of the precipitation [*Wen, Y.*, personal communication] and lower throughout the rest of the year. Due to its central location, the study area represents average climatic conditions of Inner Mongolia where a west to east trending gradient with increasing precipitation rates and temperature characterizes the climate.

A 50-year rainfall-runoff time series between 1954 and 2004 shows, that the major discharge peak occurs during the snowmelt period in spring and that a secondary minor discharge peak occurs during the vegetation period in summer when most of the precipitation occurs [*Barthold et al.*, 2008, chapter 2]. Low winter temperatures lead to complete freezing of soil water and of the river itself. The time that the river requires to melt and begin flowing is much shorter than the time necessary for complete thawing of the soil system. In fact, frozen soil layers at a depth of 1 m have been observed into the month of July in the upstream wetland areas [*Barthold et al.*, under review]. Over the course of the vegetation period, softening of the wetland areas has been observed.

A significant characteristic of the study area are the paleo sand dunes. These sand dunes cover about 20 % of the study area and appear to act as important groundwater storage. From a surface perspective, they are populated with dense vegetation, particularly *Ulmus pumila* and other tree genii, i.e. *Betulus* spp., *Malus* spp., *Prunus* spp. and *Populus* spp. on north to northwest facing slopes or in depressions, while south to southeast facing slopes have generally sparser shrubs and grassland vegetation. The vegetated and uncovered areas are characterized by extensive evapotranspiration. However, it has also been reported that sand dune systems, with their relatively high permeabilities and large storage capacities act as large near surface aquifers that can contribute significant to

groundwater recharge depending on rainfall intensity [*Dinçer et al.*, 1974; *Rimon et al.*, 2007; *Wang et al.*, 2004].

End member	Description	Well depth [m]	n	Year
Rain	Composite of rain collected at two sites	-	-	2006
			9	2007
			14	2008
Headwater source	Stream water near the origin of the stream	-	9	2006
			6	2007
			4	2008
Tributary	Ephemeral tributary	-	6	2006
			-	2007
			-	2008
G1	Shallow ground water well in upstream area	4	9	2006
			7	2007
			-	2008
G2	Shallow ground water well in upstream area	5	-	2006
			7	2007
			5	2008
G3	Deep ground water well in upstream area	25	-	2006
			7	2007
			5	2008
G4	Ground water well located in sand dunes	8	-	2006
			-	2007
			2	2008
G5	Deep ground water well near outlet	70-80	-	2006
			2	2007
			3	2008

Table 3-2. Detailed description of end members including names of end member, short description, well depths for ground water end members, number of samples (n) and year of sampling.

3.3 MATERIAL AND METHODS

3.3.1 DATA COLLECTION

Field work was conducted during the vegetation periods (May through September) in 2006, 2007 and 2008 (Table 3-1). Discharge was measured at the stream outlet using a level logger for continuously logging stream stage (Pegel-Logger PDLR70(Relativ), ecoTech Umweltmeßsysteme GmbH, Bonn, Germany) and a current meter (Ott Meßtechnik, Kempten, Germany) for intermittently measuring stream discharge. Rating curves were then applied to derive discharge time series on a daily basis. Daily precipitation was measured by the Institute of Botany, Chinese Academy of Sciences.

Grab water samples of the stream and of potential end members were collected on a biweekly basis. Source waters that were considered to act as potential end members are rain, the headwater area, groundwater and a tributary. The end member that we call "headwater area" is a stream water sampling point located at a point of the stream that belongs to the most distant portion of the river, most distant to the stream outlet. Two rain collectors were installed, one in the headwater area (Rain 1) and one near the stream outlet (Rain 2), to capture some of the spatial variability in rain water composition. A series of 5 previously existing groundwater wells with different depths also served as sampling points. Table 3-2 gives a more detailed description of the end members, including year of sampling, number of samples and depth of groundwater wells.

One of the major challenges of this study is data scarcity which is deeply rooted in the logistical constraints in the study area which defined the work in this remote study region. Given these constraints, not all end members were sampled throughout all three years. In particular, we were unable to sample the site HT during the years 2007 and 2008 because the tributary was not flowing. Sampling of groundwater wells among years was restricted by accessibility (G1) and existence of wells (G2, G3, G4 and G5). Explorative data analyses were used to test if the concentrations of that one end member, where a comparison between all three years was possible, vary significantly between the years. If end member concentrations show minimal between year variation, then one of the main assumptions of EMMA is confirmed: that the end members are time invariant. Significance tests (95 % CI) showed that in 13 of 14 solutes the concentrations of the end member H in 2006 do

not vary significantly from the two other years. Given the invariance of the end member concentrations between 2006 and 2008, we accept the standard EMMA assumption that end members chemistries are invariant, and utilize those periods for which we were able to collect samples as representative of end members across the study period.

Three kinds of samples were collected: isotope, ICP-MS (Inductively Coupled Plasma – Mass Spectrometer) and IC (Ion chromatography) samples. Isotope samples were filtered through 0.45 µm polypropylene membrane filters (Puradisc^{TM} 25PP Whatman Inc., Clifton, NJ, USA) and stored in 2 ml brown glass vials. In order to avoid any evaporation, the vials were sealed with a screw cap and an additional layer of paraffin. Samples were then transferred in a dark and cold container to avoid direct sunlight prior to analysis. ICP-MS samples were also filtered through 0.45 µm filters and acidified to pH < 2 with purified nitric acid to avoid trace metal precipitation and adsorption during storage. ICP-MS and IC samples were stored in polypropylene bottles and cooled to 4°C immediately during the field collection. At the end of each collection day, ICP-MS and IC samples were frozen until the day of analysis. Electrical conductivity (EC) and pH were measured in situ using a handheld multiparameter instrument (pH/Cond 340i, WTW, Weilheim, Germany).

3.3.2 LABORATORY ANALYSES

Stable isotopic analysis of water samples was conducted with the Liquid-Water Isotope Analyzer (DLT 100, Los Gatos, USA) based on off-axis integrated cavity output spectroscopy (OA-ICOS) at the Institute of Landscape Ecology and Resources Management, Justus-Liebig-University (JLU) Gießen, Germany. OA-ICOS measures δD and $\delta^{18}O$ directly and has recently been proven to provide results in the same or even better range than conventional stable isotope-ratio mass spectrometers [*Berman et al.*, 2009; *Lis et al.*, 2008]. Our analytical procedure follows the one described by *Lyon et al.* [2009]. The glass vials in which the samples were stored are relieved from the layer of paraffin and inserted into an auto sampler. A glass syringe then draws 0.8 µl of the sample to inject into a heated port where the absorption spectra for each isotope are determined. Each sample is injected six times to avoid memory effect between samples. Two samples are always bracketed by 5 standards. All values are reported in per mil (‰) units relative to

Vienna Standard Mean Ocean Water (VSMOW). Repeatability of δD and $\delta^{18}O$ were ± 0.6 and ± 0.2 ‰, respectively.

The element concentrations of a suite of 24 elements (Al, As, B, Ba, Ca, Cd, Co, Cr, Cu, Fe, K, Li, Mg, Mn, Mo, Na, Ni, Pb, Rb, Se, Sr, U, V, Zn) were determined via inductively coupled plasma-mass spectrometry (ICP-MS, Agilent 7500ce, Agilent Technologies, Waldbronn, Germany). The quality of the results of the ICP-MS measurements was controlled by certified reference material (NIST 1643e and NRC-SLRS4) and via additional internal calibration standards. Analytical results for the reference materials are provided by *Fröhlich et al.* [2007].

Concentration of anions (Br^-, Cl^-, Fl^-, NO_3^-, NO_2^-, PO_4^{3-}, SO_4^{2-}) were analyzed by ion chromatography (DX-120, Dionex GmbH, Idstein, Germany). Every sample value represents the mean of two consecutive measurements.

3.3.3 DATA ANALYSIS

We conducted EMMA following the procedures suggested by *Christophersen and Hooper* [1992]. These include (1) the determination of conservative tracers, (2) performance of principal component analysis (PCA) and eigenvector analysis to determine the dimensionality of the hydrologic system and to identify the end members, and finally, (3) the calculation of each end member`s contribution to stream flow.

From the suite of 33 possible tracers including cations, anions, pH and EC, those tracers that exhibit conservative behaviour were selected and used for further analysis. The determination of conservative tracers was based on the chemical characteristics of the solutes. The 7 conservative tracers are Li, Rb, Sr, Na, Mg, Cl and EC. Outliers were identified as those points that are > 1.5 times away from the interquartile range and were removed prior to analysis. Potential end members were identified by screening their elemental concentrations for significant differences to the stream water. The 95 % confidence interval (CI) was used as a measure of uncertainty in this analysis. Isotopes were used as an additional indicator to highlight potential end members.

PCA was applied to each year's dataset independently assuming that catchment processes might change among the years due to different weather conditions. PCA is a

commonly applied technique that is used to explore and describe the variability of a multivariate dataset. The purpose of PCA is to find a lower dimensional space, U space, in which the stream water observations lie and which describe the variability of the dataset. The dimensionality of U space is determined by the number of eigenvectors, or principal components, that are retained from the PCA. We applied the rule of one [*Joreskog et al.*, 1976] to determine the number of principal components to retain. The number of principal components plus one is then the number of end members that are needed to describe the system. With the rule of one, the last principal component retained in the dataset needs to explain 1/nth of the variance where n is the number of solutes used as tracers. The medians of all possible end members were projected into the U space of the stream water. Those end members that best bounded the stream water were identified. The following set of equations [*Christophersen et al.*, 1990] was solved to calculate the contributions of each end member to stream flow for a three end member system:

(1) $1 = x + y + z$

(2) $SW_{U1} = xEM_{1\,U1} + yEM_{2\,U1} + zEM_{3\,U1}$

(3) $SW_{U2} = xEM_{1\,U2} + yEM_{2\,U2} + zEM_{3\,U2}$

where x, y and z are the fractions of each end member, SW_{U1} and SW_{U2} are the projected stream water observation in U space coordinates and $EM_{n\ U1}$ and $EM_{n\ U2}$ are the coefficients of the in U space projected nth end member. All calculations were done using the software R version 2.8.1 [*R Development Core Team*, 2008].

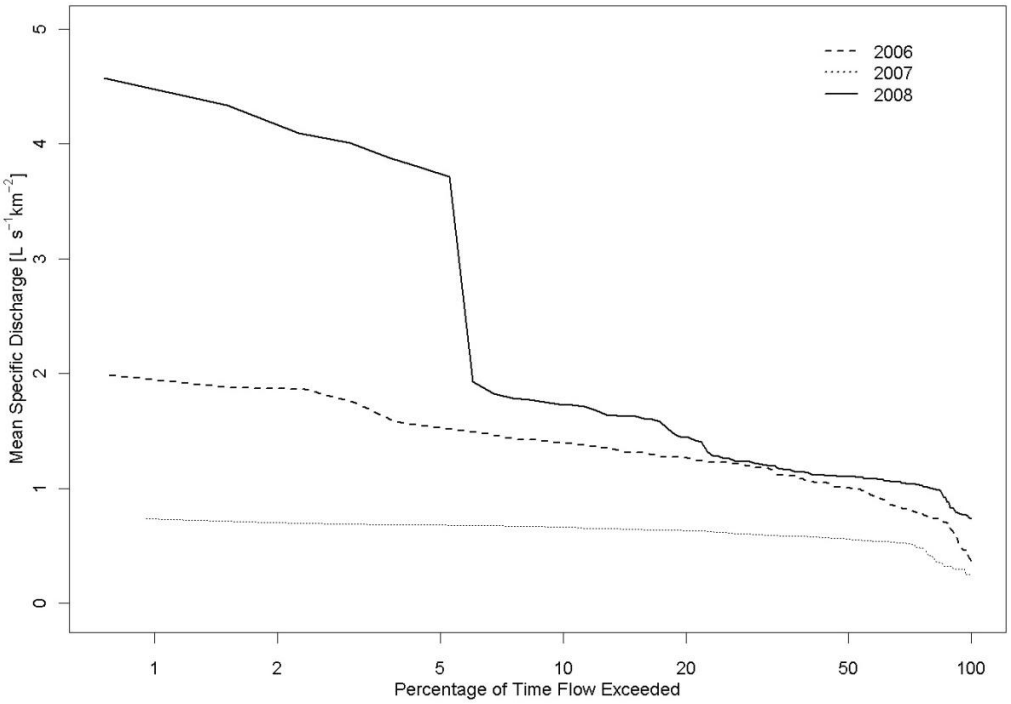

Figure 3-3. Flow Duration Curves for all three years.

3.4 RESULTS

3.4.1 TEMPORAL VARIATION OF RAINFALL AND RUNOFF

The rainfall-runoff characteristics between 2006 and 2008 vary greatly (Table 3-1). Since the sampling seasons slightly vary in length so that a direct comparison of total precipitation is not possible we present the total precipitation of the vegetation period (May 1st until September 15th). Table 3-1 indicates that 2007 was by far the driest year. The years 2006 and 2008 were much wetter but still represent moderate conditions as compared to *Z. Chen* [1988] and *Z. Chen et al.* [2002] who reported annual precipitation of more than 500 mm. The daily mean runoff for the sampling seasons reflect the rainfall characteristics of the vegetation periods with the lowest value for 2007 and higher values for 2006 and 2008 (Table 3-1).

Figure 3-3 displays the flow duration curves for all three years. The curves for 2006 and 2007 are flat which indicate the large groundwater capacities which sustain the baseflow of the stream. The flow duration curve for 2008 is in general also flat indicating significant groundwater contribution to total stream flow (Figure 3-3). However, there is a steep rise at around 5 % resulting from the rain-caused floods that occurred in the wet season.

	Stream	Rain	H	T	G1	G2	G3	G4	G5
Li									
2006	1.44 (0.08)	Na	1.01 (0.18)	0.74 (0.51)	2.32 (1.20)	NA	NA	NA	NA
2007	1.34 (0.09)	0.036 (0.08)	1.14 (0.15)	NA	1.80 (0.35)	2.63 (0.25)	2.72 (0.2)	NA	0.47 (0.001)
2008	1.63 (0.28)	0.066 (0.04)	1.00 (0.22)	NA	NA	2.55 (0.36)	2.79 (0.21)	0.94 (0.22)	0.92 (0.25)
Rb									
2006	0.0063[a] (0.0004)	NA	0.0046 (0.0015)	0.0120 (0.0072)	0.0049 (0.0025)	NA	NA	NA	NA
2007	0.0074 (0.0006)	0.013 (0.02)	0.0054 (0.0015)	NA	0.0071 (0.0009)	0.0012 (0.0002)	0.015 (0.0015)	NA	0.0004 (0.001)
2008	0.0061 (0.0014)	0.009 (0.007)	0.0062 (0.0033)	NA	NA	0.0013 (0.0006)	0.01 (0.0009)	0.003 (0.0024)	0.0014 (0.0004)
Sr									
2006	1.84[a] (0.09)	NA	1.25 (0.27)	2.27 (0.48)	3.91 (2.79)	NA	NA	NA	NA
2007	1.58 (0.14)	0.102 (0.09)	1.34 (0.11)	NA	7.46 (1.51)	3.79 (0.43)	1.75 (0.18)	NA	0.92 (0.05)
2008	1.56 (0.29)	0.14 (0.15)	1.01 (0.48)	NA	NA	3.62 (0.67)	2.13 (0.17)	1.79 (0.02)	1.62 (0.5)
Na									
2006	351.5[a] (53.1)	NA	243.6[a] (99.01)	430.3 (199.18)	891.2 (436.4)	NA	NA	NA	NA
2007	445.4 (24.8)	6.2 (10.9)	411.0 (33.99)	NA	931.7 (152.3)	592.9 (67.98)	330.56 (40.02)	NA	178.7 (50.43)
2008	469.3 (94.8)	8.5 (18.4)	381.1 (104.33)	NA	NA	550.24 (70.16)	369.62 (31.86)	311.07 (50.85)	297.0 (99.73)
Mg									
2006	301.1[a,b] (20.7)	NA	188.2 (48.22)	371.0 (104.8)	596.9 (398.24)	NA	NA	NA	NA
2007	208.9 (13)	9.48 (10.57)	173.1 (10.94)	NA	942.62 (163.6)	417.32 (55.26)	192.1 (19.19)	NA	124.53 (8.76)
2008	220 (40.1)	16.16 (8.88)	125.6 (82.23)	NA	NA	373.45 (41.44)	216.8 (15.42)	287.9 (5.1)	215.3 (63.76)
Cl									
2006	141.6[a] (5.1)	NA	109.2 (16.88)	100.1 (49.04)	988.4 (566.4)	NA	NA	NA	NA
2007	123.8 (8.9)	32.0 (30.42)	100.8 (33.03)	NA	1791.3 (309.6)	736.2 (116.5)	134.6 (21.43)	NA	144.01 (265.2)
2008	145.3 (8.6)	32.7 (91.57)	109.0 (59.03)	NA	NA	665.0 (296.68)	141.04 (89.49)	169.3 (220.4)	150.2 (2208.2)
EC									
2006	239 (7.8)	NA	184 (32.38)	317 (64.06)	737 (321.2)	NA	NA	NA	NA
2007	233 (7.5)	142 (57.3)	226 (27.21)	NA	1145 (938.2)	590 (25.5)	285 (9.7)	NA	239.5 (44.47)
2008	271 (26)	76 (39.4)	194 (77.77)	NA	NA	548 (54.4)	294 (10.74)	310 (127.1)	207 (5.30)

Table 3-3. Median (± 95 % CI) of solute concentration [µmol L-1] (NA = not available). Significances are notated with a when different from = not available). Significances are notated with a when different from 2007 and b when different from 2008, only for Stream and end member H elements.

3.4.2 ISOTOPIC COMPOSITION OF GROUNDWATER

The isotopic composition of the ground water end members varies considerably, and suggests significant differences in age (Figure 3-4). While we have not yet developed isotopic ages through one of the common methods such as e.g. convolution or any of the other methods reviewed by *McGuire and McDonnell* [2006], the differences in isotopic composition appear to confirm the hypothesis that the sand dunes (G4), the shallow aquifers (G1 and G2), and the deeper groundwater (G3 and G5) represent rather isolated volumes of water that may represent clear end members.

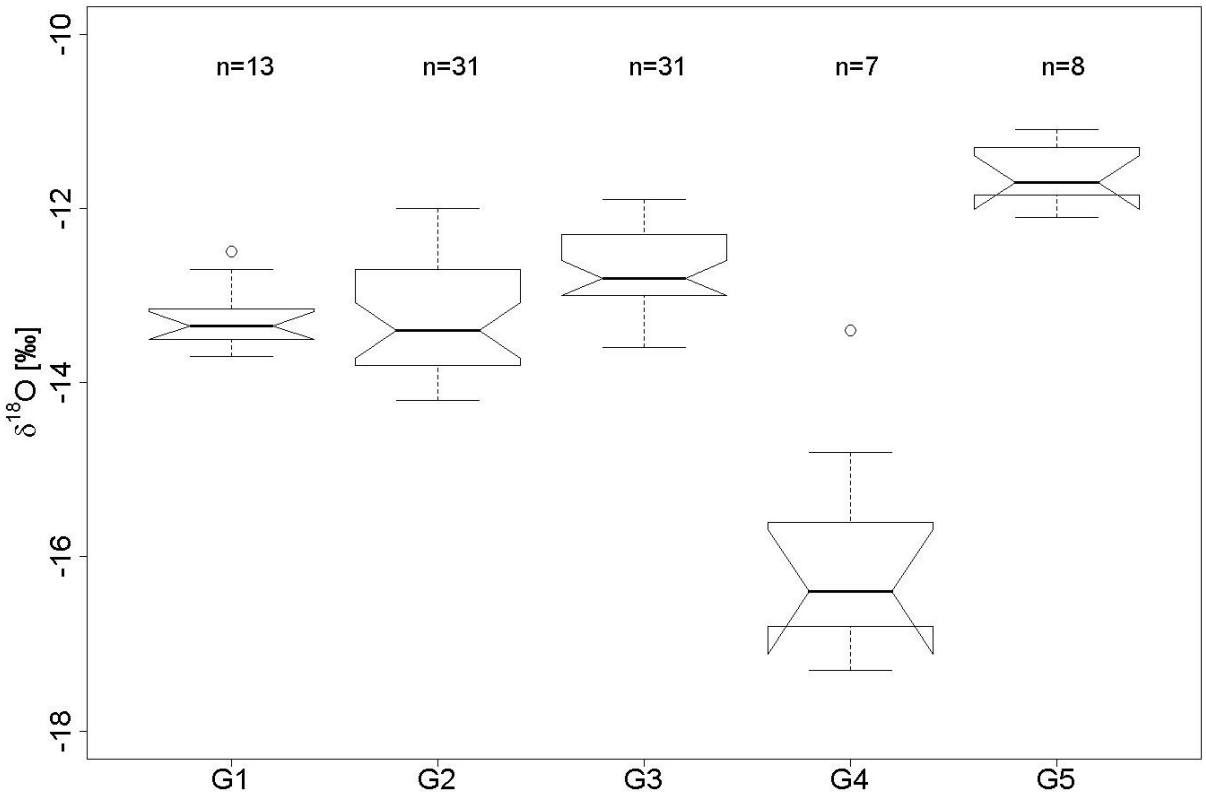

Figure 3-4. Box plots of isotopic $\delta^{18}O$ composition for sampled ground waters. The black bar in the box represents the sample median. Notches represent the 95 % confidence intervals around the median and the length of the box indicates the interquartile range. The fences are either marked by extremes if there are no outliers, or else by the largest and smallest observation that is not an outlier. The open circles are outliers that are >1.5 times the interquartile range away from the upper/lower quartile. G1 and G2 are shallow groundwater aquifers in the headwater area, G3 is a deep ground water aquifer in the headwater area, G4 a shallow ground water aquifer located in the sand dune area, and G5 is a deep groundwater aquifer near the catchment outlet.

3.4.3 IDENTIFICATION OF END MEMBERS

The basic assumption in EMMA is that the stream water is a discrete mixture of its sources. The sources must therefore be of sufficiently different concentrations than the stream water. In Table 3-3 we compare the geochemical signatures of the end members with the stream water among the 3 years. It shows that the stream water of the 3 years differ in some cases significantly (95 % CI) from each other: 2006 (n=41) is significantly different from 2007 (n=17) in the solutes Rb, Sr, Na, Mg and Cl. From 2008 (n=19) it differs only in one solute significantly (Mg). Rain water exhibits the most dilute concentrations for all elements except for Rb. The headwater source H also has significantly lower concentrations than the stream water but is not as extreme as the rain. The tributary is slightly higher than the stream water. The shallow ground water wells (Table 3-2), G1 and G2, which are located in the headwater area, have higher concentrations than the stream water (except for Rb). The deep ground water well G3 and the ground water well representing the sand dune area, G4, do not exhibit a clear trend of either high or low concentrations. They vary among the solutes. The deep groundwater well, G5, near the stream outlet has most of the times lower solute concentration than the stream water.

We conducted a PCA on the stream water data set of each year independently to determine the dimensionality of the system and to identify the end members. Table 3-4 lists the results of the PCAs including standard deviation, proportion of variance and cumulative proportion of every principal component (PC). We applied the rule of one to estimate the dimensionality of the system. With 7 solutes, the last PC needs to explain 14.3 % of the variance. This results in the retention of one PC for 2006 and two PCs for 2007 and 2008. Hence, a minimum of two end members for 2006 and a minimum of three end members for 2007 and 2008 are required. However, we hypothesize that the number of dimensions should be the same over all years. *James and Roulet* [2006] formulate a similar hypothesis to explain spatial variability of end members in a system. We applied this hypothesis to a temporal variability problem in this study and assume that all end members exist over all three years but change in their contribution. We therefore decided for a three end member system for all years.

	PC1	PC2	PC3	PC4	PC5	PC6	PC7
2006							
Standard deviation	2.24	0.9	0.76	0.61	0.38	0.25	0.09
Proportion of variance	0.717	0.116	0.083	0.053	0.021	0.009	0.001
Cumulative Proportion	0.717	0.833	0.916	0.969	0.99	0.999	1.000
2007							
Standard deviation	2.131	1.238	0.8103	0.3780	0.2680	0.18804	0.13381
Proportion of variance	0.649	0.219	0.0938	0.0204	0.0103	0.00505	0.00256
Cumulative Proportion	0.649	0.868	0.9617	0.9821	0.9924	0.99744	1
2008							
Standard deviation	2.251	1.134	0.6699	0.384	0.17248	0.14320	0.02507
Proportion of variance	0.724	0.184	0.0641	0.021	0.00425	0.00293	0.00009
Cumulative Proportion	0.724	0.908	0.9717	0.993	0.99698	0.99991	1

Table 3-4. Results of the Principal Component Analysis (PCA) which was conducted for each year independently. (PC = Principal Component)

It may be added that in PCA analyses, the decision on how many PCs to retain is often a subjective consideration based on the variance explained by each principal component, as is the rule of one [*Hooper*, 2003]. In our approach we used the rule of one as a decision tool to identify the total number of end members needed to describe the system over the years.

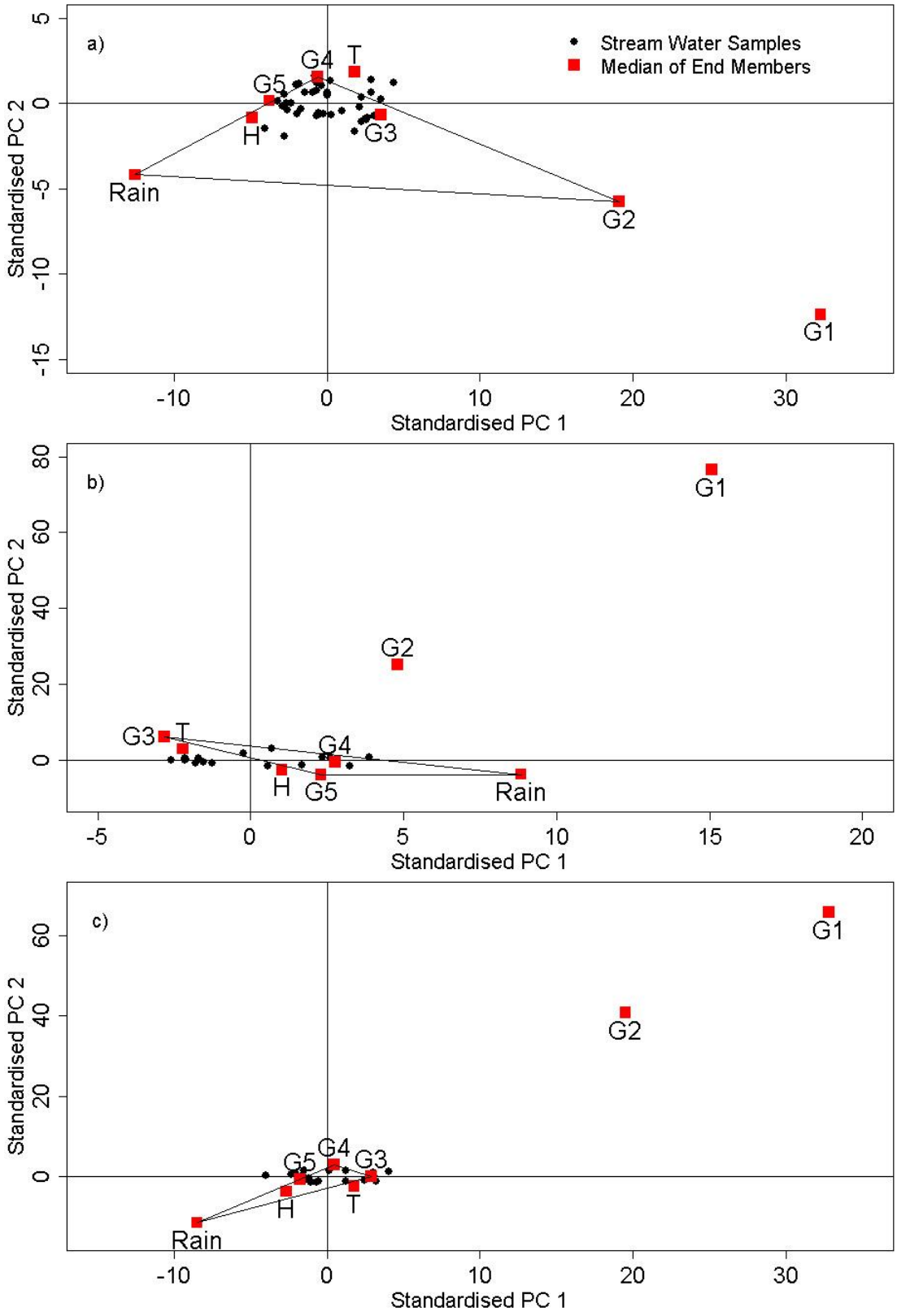

Figure 3-5. Stream water observations and medians of end members projected into U space of the stream water for all years.

We then projected the medians of the end members into the U Space of the stream water of each year (Figure 3-5). The most suitable three end members for 2006 are G2, G4 and rainfall water. Figure 3-5a shows that most of the stream water observations fall into the triangle that is spanned by these three end members. However, there exist some stream water observations that lie outside of the triangle. These types of outliers are typical in EMMA models [*Chaves et al.*, 2008; *Christophersen and Hooper*, 1992; *Elsenbeer et al.*, 1995; *F. J. Liu et al.*, 2008] and result from a number of factors including: (1) uncertainty in field sampling or laboratory analyses, (2) lack of temporal invariance of end members or (3) the expression of different end member in the mixture as water source areas change temporally. Overall, the result can lead to over- or underprediction of the contributions of each end member to the stream water, and should be understood as a source of uncertainty.

Figure 3-5b and c show the projected stream water observations and end members for 2007 and 2008, respectively. The location of the end members in these two years differ greatly from 2006 (Figure 3-5a). In 2007, the end members G3, G5 and rainfall best bound the stream water observations (Figure 3-5b). In 2008, the best possible end member combination to explain the stream water mixture includes G3, G4 and rainwater, similar to mixing conditions in 2006.

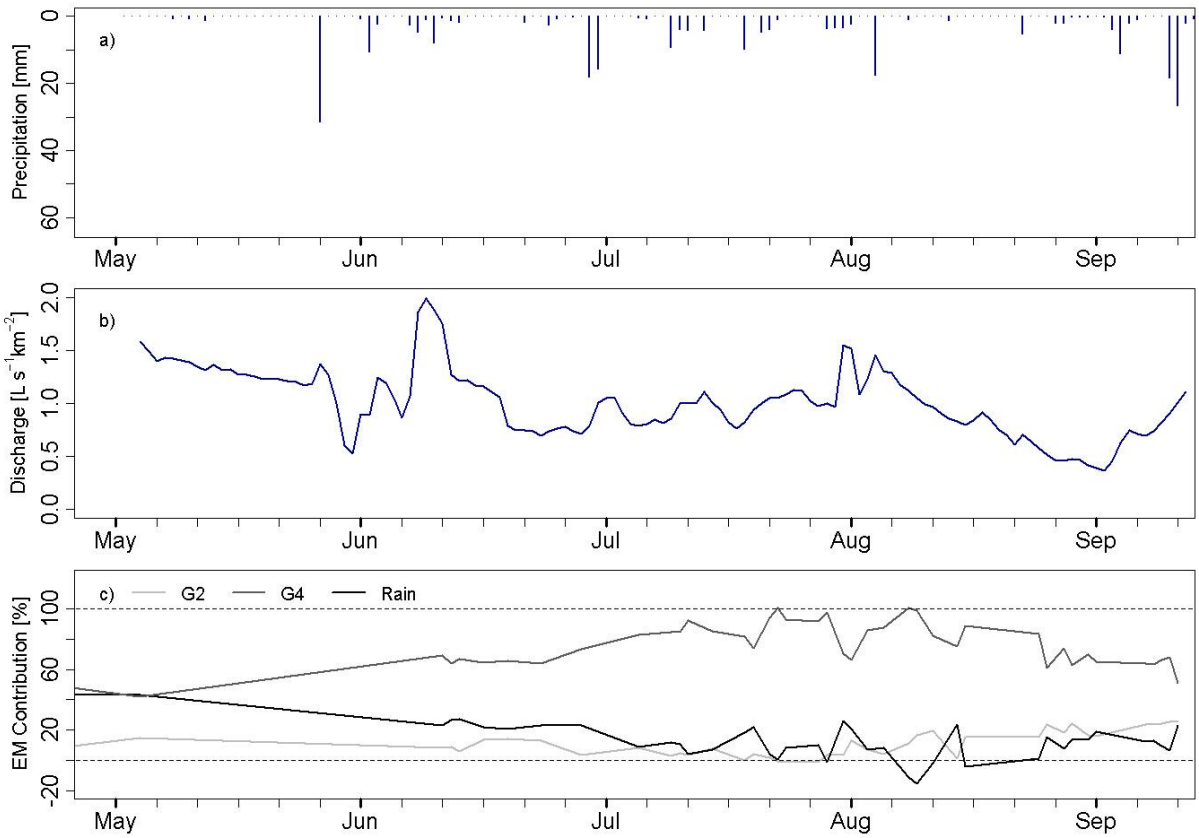

Figure 3-6. Rainfall (a), runoff (b) and contributions of end members to the hydrograph (c) in 2006.

3.4.4 CONTRIBUTION OF END MEMBERS TO RUNOFF

The rainfall-runoff graphs in Figure 3-6a and b show the recession of the spring discharge peak at the beginning of the sampling campaign. The peak that occurs in mid June as a result of 4 rainfall events is assumed to be the secondary minor peak of the season since the snowmelt peak usually occurs in April. The 2006 rainfall-runoff time series indicates that the year 2006 can be classified as an average year with respect to the hydrologic conditions following the same trend as the average 50-year hydrograph between 1954 and 2004 [*Barthold et al.*, 2008]. This average hydrograph is dominated by a snow melt peak in April followed by a steady decline in discharge over the vegetation period. During July and August, when most rainstorms occur, there is a slight increase in discharge, followed by an again declining hydrograph until October, when the river itself freezes.

The contributions of the end members G2, G4 and rainfall were calculated for 2006 by solving equations (1) to (3). The shallow ground water source G2 contributes a consistent though relatively minor amount (around 10 %) to the hydrograph over the vegetation

period (Figure 3-6c). The fraction that is contributed by rainfall decreases consistently throughout the season from 50 % until it reaches a minimum in August of approximately 10 %. The ground water that originates from the sand dune area (G4) contributes the most to the discharge. It increases over the course of the season until it reaches a maximum in mid July and mid August. It then slightly decreases to the end of the season (Figure 3-6c). It reacts with an inverted, negative peak to rainfall events.

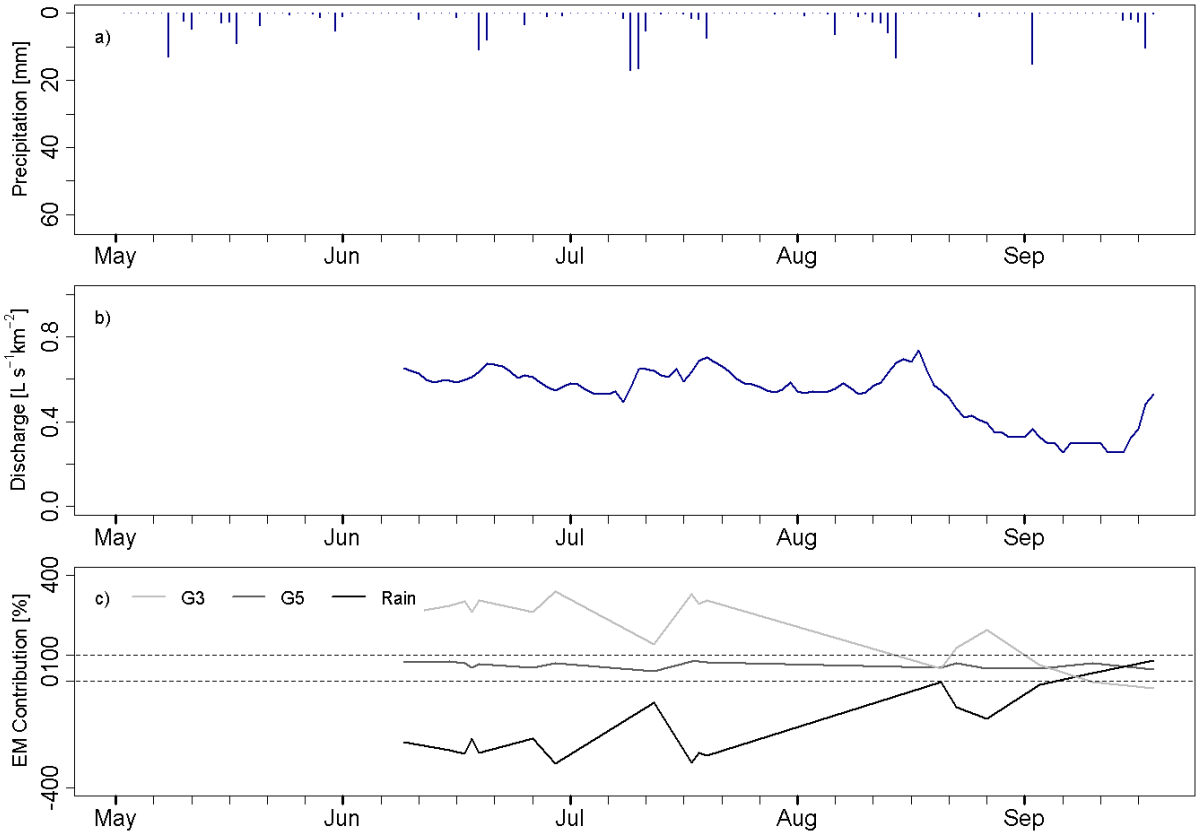

Figure 3-7. Rainfall (a), runoff (b) and contributions of end members to the hydrograph (c) in 2007.

The field campaign in 2007 began in June, later than in 2006, and after the snow melt period (Figure 3-7, Table 3-1). Total precipitation, the number of events and the magnitude of single rainfall events was less than in 2006. The discharge time series followed again the observed long-term average hydrograph. The runoff slightly declines throughout the vegetation period starting at approximately 0.6 L s^{-1} km^{-2} in June and ending with approximately 0.4 L s^{-1} km^{-2} in September. A slight peak occurred in mid August (Figure 3-7b). The contributions of the end members were calculated and plotted

in Figure 3-7c. The end members G5 and rainfall are largely over- and underpredicted, respectively. Only the contribution of the deep ground water well G3 falls into the plausible range between 0 and 100 % contribution. It contributes around 70 % to the stream discharge throughout the whole period.

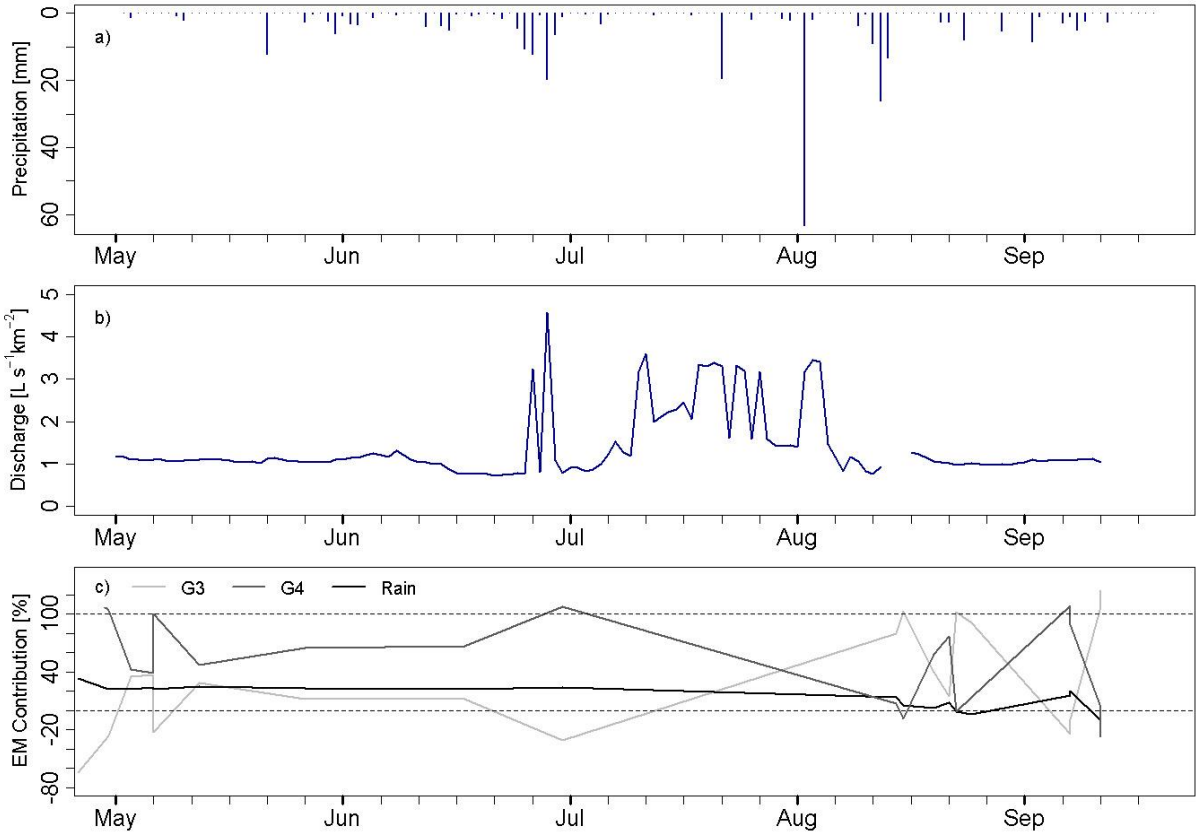

Figure 3-8. Rainfall (a), runoff (b) and contributions of end members to the hydrograph (c) in 2008.

Rainfall and runoff in 2008 exceed the values of the previous two years. The hydrograph reacts with high discharge peaks to rain events at the end of June, throughout July and at the beginning of August (Figure 3-8a and b). The highest discharge occurs at the end of June as a result of a rain event. In July and August, additional runoff peaks occur. It is striking that most of these peaks appear to react to the slightest rainfall events. This observation can be explained by the highly heterogeneous rainfall patterns that occur in this region as a result of convective rainfall [*Schneider et al.*, 2008]. Rainfall shown in Figure 3-6a, Figure 3-7a and Figure 3-8a was in the very eastern part of the catchment (Figure 3-1), suggesting that the sparse network of rain gauges simply missed some rather

large convective events, but that these events were evident in the stream discharge. It is likely that runoff generating rainfall in the headwaters of the Xilin might produce discharge peaks that are not reflected in the existing hyetographs. Another explanation is that the shallow runoff contributing storage were filled up by precipitation at the beginning of the season and now react to the slightest rainfall events, but this seems unlikely given that the cumulative total of events (294.3 mm) tends to be much smaller than storage. However, we should also consider that snowmelt might have filled up these storages. As it slowly moves through the system it may be expressed in the stream water some months later after it infiltrated into the soils.

In general, the observed hydrograph in 2008 does not follow the average hydrograph. While missing samples somewhat limit the approach for this year, we are able to identify the sand dune aquifer, G4, as a dominant runoff generating source. The deep groundwater aquifer, G3, which seemed to be an important runoff generating storage in 2007, appears as only a minor contributor in 2008. This implies a switch of runoff generating sources back to shallower groundwater storages such as in the sand dune area, consistent with observations from 2006. Despite high rainfall amounts, the precipitation remains only a minor runoff contributor given our analysis, steadily declining over the season from approximately 30 down to 10 %. Of course, the equal interval sampling strategy was designed to focus on longer term flow dynamics and would not capture events, particularly because of the fast temporal dynamics and significant spatial variation which defines them. If event water does contribute directly to steam discharge, as either Hortonian or saturation excess overland flow, it would not take on the shallow groundwater chemistry, and as such has the potential to be under represented. A more complete event-based sampling strategy could be devised to better evaluate the direct contribution of event water to the stream, but in remote ungauged basins such as the Xilin, a first step is to evaluate the efficacy of the methods, returning with more sophisticated sampling plans after some of the more general hydrology has been explored.

3.5 DISCUSSION

In 2006 and 2007 the slope of the entire flow duration curve was small, which can be used as an indicator for significant and sustained groundwater contribution [*Smakhtin*, 2001]. Even in 2008, a wetter year, the flow duration curve showed this behaviour. However, the slope increases at around Q5 which can be attributed to the high flows at the end of June and throughout July 2008. Overall, flow duration curve analysis indicates that the Xilin river catchment is a groundwater dominated system. This conclusion is supported by the hydrograph that shows a dampened effect on most rainfall events in all three years.

Our study indicates interannual variability among the runoff producing processes and sources. The simple analysis of the rainfall-runoff characteristics of the three years may give a first hint that the yearly climatic variation appears to have resulted in this variability. This variability has a pronounced influence on the EMMA approach, and suggests the possibility that some end members may only be sampled during certain years. Precipitation and discharge records from 2006 are similar to the 50 year average, suggesting it was a typical hydrologic year. Our PCA results and projections of end members into the U space of stream water observation identified two shallow ground water aquifers, G2 and G4, and rainfall as end members. The dominant runoff producing source in 2006 was the sand dune ground water aquifer, G4. The sand dunes are characterized with high infiltrability and high saturated hydraulic conductivities (K_{sat}). Average K_{sat} in 0 to 1 m depth is 31.5 cm/h. We assume that if enough precipitation occurs, water can percolate quickly and fill the groundwater aquifer which underlies the dunes. This observation has been reported from other sand dune dominated locations [*Dinçer et al.*, 1974; *Rimon et al.*, 2007; *Wang et al.*, 2004], and while the magnitude of this precipitation threshold remains unknown, it appears likely that the average conditions measured during 2006 exceeded it. Subsequently, the shallow sand dune aquifer became a significant source of water for the river, and could be geochemically differentiated as an end member in the EMMA approach. Our isotope analysis also supports this assumption with relatively low $\delta^{18}O$ values of the G4 water indicating that the water below the sand dunes is much different from the water making up the deeper groundwater aquifers (Figure 3-4). The shallow groundwater aquifer, G2, consists of

relatively older water than G4 since infiltrability and K_{sat} are lower in this area. This storage consistently contributes to runoff but is only a minor source for the river if enough rain falls to fill up the sand dune aquifer. These results suggests that for many catchments, a long term sampling strategy is needed to capture the long term climate (and hence runoff) variability. Some runoff sources may cease to contribute to stream water discharge during particularly dry years, and this fact is a key feature of our overall understanding of the variety of ways in which climate, hydrology, and the near surface environment contribute to stream discharge.

The triangle defined by the end members G2, G4 and rainfall bound more than 80 % of the stream water observations in 2006. However, some observations fall outside of the triangle. There are various explanations: (1) uncertainty in field and laboratory analyses, (2) non-conservative solute behaviour (3) unidentified end members and (4) temporal variability of end members. In many other studies that apply end member mixing analysis such a situation has been described [*Chaves et al.*, 2008; *Christophersen and Hooper*, 1992; *Hooper*, 2001; *James and Roulet*, 2006; *F. J. Liu et al.*, 2008, 2004]. These "outliers" usually cause over- and underprediction when calculating the contributions of the end members. To avoid this problem, these observations are often projected back into the mixing domain [*Chaves et al.*, 2008; *F. J. Liu et al.*, 2008, 2004]. However, this operation is somewhat unsatisfactory in that it does not solve the problem, but simply results in calculating plausible ranges of the end member fraction that contribute to stream flow. We elected to leave these data in the model and attribute slight variations of stream water observations from the end member mixing domain to uncertainty in field and laboratory analyses. In those cases where variations are more significant, we provide some additional thoughts to indicate potential missing end member and lack of time invariance of end member.

The location and distribution of outliers from 2006 suggest uncertainty in field sampling and/or lab analyses as the primary cause. As described above, they cause some over- and underprediction in the calculated contributions. However, the identification of end members in 2006 is supported by the chemical composition of end members (Table 3-3) which appear to bound the stream water quite well.

The conditions in 2007 varied greatly from those in 2006. Projections of the end members into the U space of the river water give a different picture of the dominant runoff producing sources than for 2006. In fact, the dominant runoff producing sources for 2006, G2 and G4, can be rejected as end members, based on analysis of the 2007 data. The end members G3, G5 and rainfall bracket the stream water observations and appear therefore as more likely end members in 2007. G3 and G5 are both deep groundwater aquifers (Table 3-2), one is located in the headwater area and one is located further downstream. This shift of end members to deeper groundwater storages suggests that shallow sources were depleted throughout the sampling season, due to lack of input or recharge by precipitation.

However, more than 50 % of the stream water observations lie outside the mixing triangle comprised by G3, G5 and rainfall. This large number of outliers suggests that sources of error beyond field and laboratory uncertainty existed in the 2007 data. The most likely source of error is a missing end member, suggesting that for dry years, additional sampling locations would need to be established to fully characterize the sources of water in the stream mixture.

The year 2008 was characterized by high rainfall rates, and is similar in that sense to 2006. Again, a shift in runoff producing sources occurred. The deep ground water aquifer in the headwater area, G3, appears to have remained an important source of stream flow. The rainfall rates were high enough to fill up the sand dune aquifer G4 which then expressed itself geochemically in the stream discharge. Figure 8c does indicate a decreasing rainfall contribution; a result somewhat contradictory to the general understanding of rainfall as an important runoff producing source. However, data are sparse for 2008 (n=19) and no sampling occurred during the main rainfall events, hence results for 2008 might be biased towards an overestimation of groundwater sources.

One shortcoming of this study is that the identification of end members is limited to data collected during the vegetation period which comprises only four months of the year. Even if this is the period with most of the precipitation, we may miss some important components of the system. For example, the average annual hydrograph [*Barthold et al.*, 2008] indicates that snowmelt produces a major discharge peak. Snowmelt water may also infiltrate into the soils, fill up the shallow ground water aquifers and, with some time

lag, may occur again in the stream water after it slowly moved through the system. In this case, which is likely to occur, our models do capture snow melt contribution as the water takes on the chemical signature of the shallow groundwater. However, our data limitations do prohibit the identification of snowmelt as an end member in and of itself.

Our study indicates that one end member mixing model which is able to explain the runoff in all three years does not exist. In all three years, EMMA fails to represent all the stream samples. The explanation that appears to most consistently explain this shortcoming is the temporal variability of end members. There are two types of temporal variability: (1) time variance of end members, i.e. the elemental concentration of end members changes over time due to e.g. leaching processes in the soil layers and (2) the combination of end members that contribute to the runoff changes over time, i.e. seasonal or interannual variation in end members caused by a depletion of water storages. Due to the relatively long period over which we apply EMMA both of these cases may influence our results, and clearly demonstrate the limitations of EMMA and its application in data sparse catchments. Time variance of end members where the concentrations change over time, has been pointed out before as a major limitation of EMMA [e.g. *Christophersen et al.*, 1990; *Elsenbeer et al.*, 1995]. However, temporal variability of end members where the composition of sources stay constant but the end member combination that is expressed in the stream mixture changes is a limitation that may contribute a large part to the uncertainty in our models. This uncertainty may be reduced by increasing the number of end members in the model. The underlying assumption then is that all end members are present in each year, although each end member could certainly be contributing to varying degrees among the three years. We have pointed out in section 4.3. that *James and Roulet* [2006] formulate a similar hypothesis to explain spatial variability of end members in a system. In this study, we apply this hypothesis to a temporal variability problem. The determination of an increased number of end members in a system is only possible with a larger tracer set since the calculations of the end member contributions are based on an analytical approach that solves an over determined set of equations based on a least squares procedure. In such a case, the value of large tracer sets becomes obvious.

The results clearly indicate that runoff producing sources in this region may vary among years that differ in climatic conditions. Such seasonal and interannual variability of end member contributions to stream flow has been reported by *Liu et al.* [2008]. They also hypothesize that end members might vary among the years, however, they did not apply EMMA on each year independently. *Bernal et al.* [2006] showed that one-fits-it-all EMMA did not work in intermittent Mediterranean streams. They investigated single storm events and showed that stream water during the transition between events was not bracketed by end members that worked well during the wet and dry periods. We are not aware of further detailed studies that investigated the interannual variability of end member mixing systems and suggest that more research is needed on this topic.

3.6 CONCLUSIONS

Geochemical and isotopic tracers and the application of EMMA proved to be useful tools in the process of identifying runoff generating sources at the catchment scale in a semi-arid watershed in Inner Mongolia. Despite sparse data, findings suggest that shallow groundwater aquifers are the primary sources of stream flow in average years, and that these sources of discharge are depleted during dryer years. During dryer years, EMMA results are consistent with the notion that deeper groundwater aquifers are the dominant sources of stream water. The interannual variability of end members that contribute to stream flow is crucial in many semi-arid watersheds when conceptualizing catchment hydrology at mesoscales.

While the results presented here help corroborate the dominance of groundwater as the primary contributor to streamflow, they also suggest the need for long term records to fully characterize the range of climatic variability that the catchment experiences. The 3 year measurement program appears to have characterized average and somewhat lower wetness years, but the response during wetter years remains undocumented. It seems reasonable to suggest that rainwater would increase its dominance as a source for stream water under these conditions, nevertheless sampling the geochemistry in a manner similar to that established in 2006, 2007, and 2008 would be necessary to complete the EMMA analysis across the range of typical conditions. This is an obvious challenge to the

application of EMMA in some regions, and suggests the need for the continued development of long term measurement programs.

The geochemical data and EMMA analysis, along with standard hydrologic measurements suggest a conceptual model of streamflow generation to describe much of the measured flow variability. We are working now on capturing this conceptual model as numerical model to further evaluate its utility across a wider range of climatic conditions. In addition, the development of numerical models which are consistent with known process dynamics have the potential to approach fundamental questions related to large scale landuse and management changes.

4 EMMA: ESTIMATING THE VALUE OF LARGE TRACER SETS VERSUS SMALL TRACER SETS

Abstract

End member mixing analysis (EMMA) is a commonly applied method to identify and quantify the dominant runoff producing sources of water. It employs isotopic and geochemical tracers to determine the dimensionality of the hydrologic system, i.e. the number of end members (EM) or components that are necessary to reproduce runoff in the catchment under investigation. This is typically done by applying Principle Component Analysis (PCA) and the Rule of 1 [*Joreskog et al.*, 1976] or the diagnostic tools of *Hooper* [2003]. Based on this information the contribution of each EM to runoff is quantified by solving a set of mass balance equations. Many studies have been conducted to identify and quantify runoff sources using 2 to 6 tracers, with the main tracers being Ca, K, Mg, Na, Cl^-, SO_4^{2-}, Si, the isotopes $\delta^{18}O$ and δD, acid neutralizing capacity (ANC), alkalinity and EC. Only few studies use larger tracer sets including also minor trace elements such as Li, Rb, Sr, Ba etc. None of the studies so far addressed the question of the tracer set size and composition, despite the fact that these determine which and how many EM will be identified. In this study we examine how tracer set size and composition affects the model concept that results from an EMMA. We developed an automatic procedure that iteratively changes tracer set size and composition and conducts EMMA for each tracer set possibility. We are using a set of 14 tracers and 9 EMs. The validity of resulting model concepts was investigated under the aspects of dimensionality (EMs needed to explain flow in a system), EM combinations and contributions to stream water. From the 16.369 tracer set possibilities only 23 delivered results that were plausible as defined by mixing model theory. This large reduction of plausible results can mainly be attributed to differences in tracer ratios of stream water and EM waters. Hence, the resulting model concepts are highly sensitive to the tracer set size and composition. However, a certain degree of consent exists among dimensionality and selected EM combinations. The moderate reproducibility of EM contributions indicates that more field work is necessary to identify a still missing EM. It also emphasizes that the major elements are not always the most useful tracers and that larger tracer sets that contain also minor trace elements have an enhanced capacity to avoid false conclusions about catchment functioning. The presented iterative EMMA approach is able to produce results that may not be apparent from the traditional approach and we stress the fact that this approach provides a very valuable complement.

With C. Tyralla, K. Schneider, K.B. Vaché, H.-G. Frede and L. Breuer submitted to *Water Resources Research*.

4.1 INTRODUCTION

Isotopic and geochemical tracers have widely been used in studies to quantify the dominant runoff producing components. The two standard techniques that govern tracer studies and that have evolved over the past five decades are hydrograph separation [*Pinder and Jones*, 1969] and end member mixing analysis (EMMA) [*Christophersen et al.*, 1990]. Hydrograph separation aims at quantifying temporal source components, i.e. separating the hydrograph into pre-event and event water, whereas EMMA targets a separation in geographical source components. Both techniques are predicated on the same assumptions: (1) the stream water is a mixture of source solutions with a fixed composition, (2) the mixing process is linear and relies completely on hydrodynamic mixing, (3) the solutes used as tracers are conservative and (4) the source solutions have extreme concentrations. However, the approaches differ in their mathematical accomplishments: hydrograph separation relies on a mass balance approach and EMMA is based on an analytical approach that solves an over determined set of equations based on a least squares procedure.

The most commonly applied tracers in hydrograph separation studies are Ca, Mg, K, Na, Fe, Cl^-, CO_3, NO_3, Si, the isotopes $\delta^{18}O$ and δD, and electrical conductivity (EC). In general, pre-event and event water is described by groundwater and precipitation in two-component models [*Sklash and Farvolden*, 1979] and by groundwater, soil water and precipitation in three-component models [*Bazemore et al.*, 1994; *Dewalle et al.*, 1988; *Kennedy et al.*, 1986; *Swistock et al.*, 1989]. Hence, only a subset of the tracers is needed to calculate the fractions: one tracer for two-component models and two tracers for three-component models. As opposed to hydrograph separation studies, the amount of tracers needed in EMMA studies is not restricted to one or two due to the mathematical nature of the approach. Usually, two to six tracers are used with the main tracers being Ca, K, Mg, Na, Cl^-, SO_4^{2-}, Si, the isotopes $\delta^{18}O$ and δD, acid neutralizing capacity (ANC), alkalinity and EC [*Burns et al.*, 2001; *Elsenbeer et al.*, 1995; *Katsuyama et al.*, 2009; *F. J. Liu et al.*, 2008; *Soulsby, Rodgers, et al.*, 2003]. Only few studies use larger tracer sets including also minor trace elements such as Li, Rb, Sr, Ba [*Fröhlich et al.*, 2008; *Ladouche et al.*, 2001]. But which and how many tracers are best? This question stands at the beginning of each tracer study since the selection of the tracer set size and composition

influences not only the costs of the study but may also have impacts on the outcome of the study. *Rice and Hornberger* [1998] provide a study where the reproducibility of hydrograph separation results among seven different pairs of tracers is tested. They concluded that the results of three-component separations were widely variable but that consistent patterns in the amount of subsurface water contributing to storm flow could be observed independently on the combination of tracers used.

However, none of the EMMA studies so far has addressed the question of the tracer set size and composition; despite the fact that these determine which and how many end members will be identified. For example, the EMMA approach also includes the determination of the dimension of the hydrologic system, i.e. the number of end members (EM) or components that are necessary to reproduce runoff in a catchment, and the identification of these end members. For these purposes, principal component analysis (PCA) is applied. Principal component analysis is generally used in problems where the variation in a set of many, correlated variables needs to be described by a new set of few, uncorrelated variables. The new set of variables, also called the components, is sorted by decreasing order of importance, i.e. by the amount of the variation that each component explains. The general hope of PCA is that only a few components can explain a substantial portion of the variation of the original variables and hence, that the components make up a lower dimensional space that explains most of the variation. The decision on how many components to retain remains subjective and is influenced by the amount of input variables. The two most common tests in EMMA studies on which this decision is based are (1) the Rule of One [*Joreskog et al.*, 1976] and (2) the diagnostic tools of *Hooper* [2003]. The Rule of One says that at least 1/nth of the variation must be explained by the last principal component, where n is the number of tracers used in the study. *Hooper's* [2003] diagnostic tools include a residual analysis where the differences between the original and the projected variables are calculated and screened for structure. A good mixing subspace is indicated by a random pattern of the residuals. Structure in the residuals means lack of fit in the mixing subspace indicating non-conservative behavior of solutes among other reasons. The dimensionality of the system is then determined by the smallest dimension possible such that the residuals exhibit no structure when plotted against observed concentration.

In the first case, the Rule of One, it is apparent that the amount of tracers used in the study influences the amount of end member determined. In the second case, the diagnostic tools of *Hooper*, the use of a larger tracer set holds the necessity of screening more scatter plots (one for each tracer) while at the same time the number of possible dimensions increases and with it again, the number of scatter plots to screen (one for each tracer and each dimensionality) which is time consuming and decreases clarity.

In this study we challenge the seemingly random selection of natural tracers in EMMA studies by addressing the degree to which additional tracers contribute to additional information. The objectives of our study are (1) to find a tracer set threshold beyond which the acquisition of insight in terms of system dimensionality becomes negligible, (2) to evaluate the influence of the tracer set size and composition on the model concept that results from an EMMA (3) to analyze the compositions of tracer sets that produce acceptable results as defined by mixing model theory.

4.2 Material and Methods

4.2.1 Study area

We are using a water sample data set from a subcatchment of the the Xilin River basin in Inner Mongolia, China (Figure 4-1). The subcatchment, which covers an area of 475 km^2, is located in central Inner Mongolia at approximately 43°33’ North and 116°45’ East and represents a typical steppe ecosystem. It is characterized by a semi-arid continental climate with dry, cold winters and warm, relatively wet summers. Mean annual precipitation is 150 to 500 mm with 60 to 80 % falling during the vegetation period between June and August [*Z. Chen*, 1988]. The mean annual temperature is 2°C with a January average of -23°C and a July average of 18°C [*Z. Chen*, 1988]. The elevation ranges between 1175 and 1500 m. The area is mainly covered with grasslands which shift to mountain meadow in higher elevated areas in the east of the catchment. Wetlands dominate in the vicinity of the river and a strip of paleo sand dunes stretches through the area in a northwest- southeast direction. These sand dunes take up about 19.4 % of the area. Industries that may potentially act as anthropogenic point sources into the river do not exist in the study area. The dominating soils in the study area are Phaeozems,

Arenosols and Gleysols (*Barthold et al.*, under review). Phaeozems are found in steppe areas, the less developed Arenosols cover the sand dunes and Gleysols dominate in the wetland areas. The underlying geology is quite diverse spanning a large range of rocks. The oldest systems which form the basis of the area originate from the Paleozoic and comprise a variety of igneous rocks (Granites, Granodiorites and Diorites), sediments and metamorphic rocks (Schists). Some of these oldest areas are capped by volcanic rocks from the Mesozoic and together they form low mountains in the present-day landscape [*Z. Chen*, 1988; *Tong et al.*, 2004]. Pleistocene basalts cover the south west of the area forming lava table lands and high plateaus. The largest part of the study area is dominated by quaternary deposits that are distributed along the river bed. A more detailed description of the land use, soils and geology are given in *Barthold et al.* (under review).

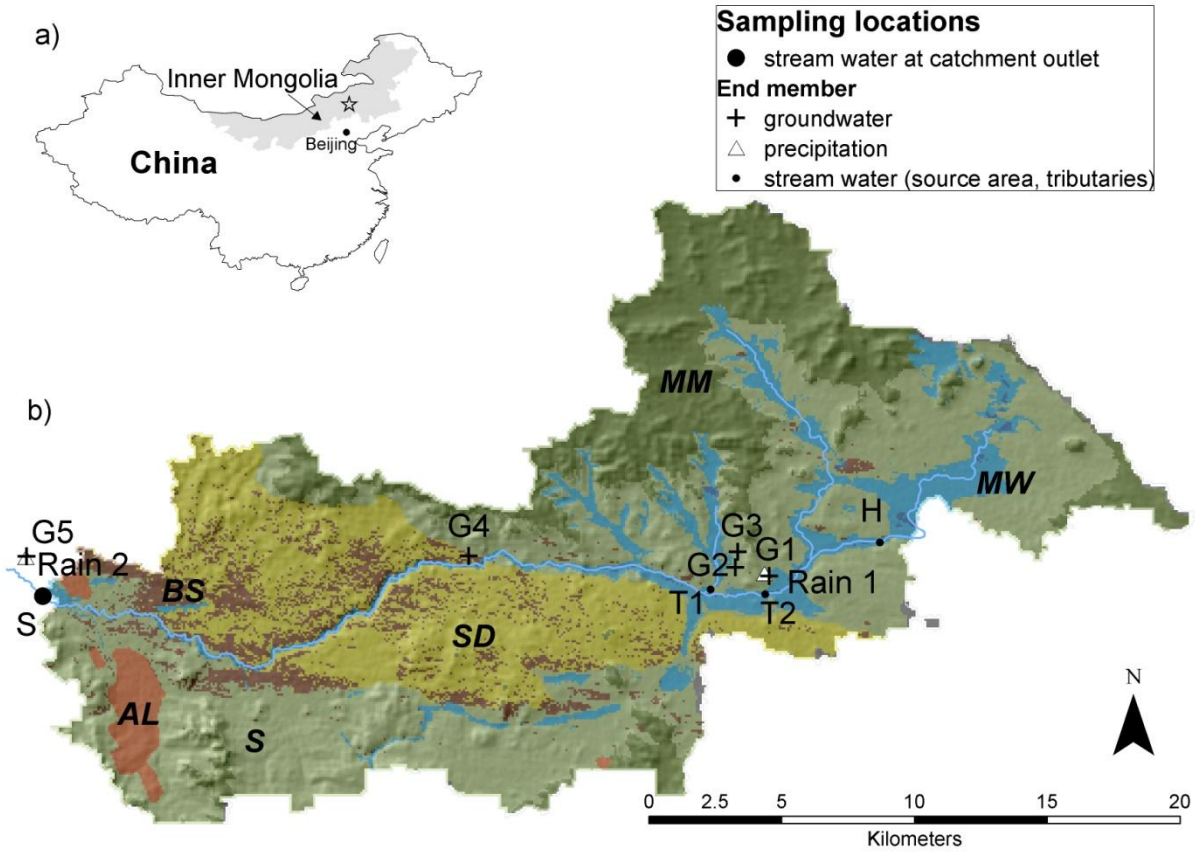

Figure 4-1. Location of study area in China (a) and land use units with locations of sampling sites superimposed on the hillshade of the Xilin river subcatchment (b). Bold italics indicate land use units: AL = arable land, BS = bare soil, MM = mountain meadow, MW= marshland/water, S = steppe and SD = sand dunes.

4.2.2 DATA SET

The data set consists of stream water samples collected at the outlet of the catchment (S) (Figure 4-1) and samples of various possible end members such as rain, stream water from 2 tributaries, stream water from the source area and 5 groundwater wells. Table 4-1 holds details about the potential end members. Rain was collected at two sites, one located in the head water area ($Rain_h$) and one located near the catchment outlet ($Rain_{out}$). Data from both collectors were lumped into one set and represent the rain end member. The stream data is a time series collected in the vegetation period of 2006 (April 15th until September 15th). Samples of end members exist for 2006, 2007 and 2008. However, not all end members were sampled in every year due to logistical restrictions. Since the years 2006 and 2008 were identified as hydrologically similar [*Barthold et al.*, 2010, chapter 2], we assumed time-invariance of end members among the two years. Table 4-1 indicates in which year the end members were sampled.

Grab water samples were collected of stream and end member waters and divided among three different storage containers: 2 ml brown glass vials for isotope samples and polypropylene bottles for inductively coupled plasma-mass spectrometer (ICP-MS) and ion chromatography (IC) samples. Isotope and ICP-MS samples were filtered through 0.45 µm polypropylene membrane filters (Puradisc^{TM} 25PP Whatman Inc., Clifton, NJ, USA). Isotope vials were sealed with a screw cap and an additional layer of paraffin to avoid any evaporation. ICP-MS samples were acidified to pH < 2 with purified nitric acid to avoid trace metal precipitation and adsorption during storage. ICP-MS and IC samples were cooled to 4°C immediately during the field collection and, at the end of each collection day, frozen until the day of analysis.

Samples were analyzed for a suite of 34 tracers composed of 24 elements (Al, As, B, Ba, Ca, Cd, Co, Cr, Cu, Fe, K, Li, Mg, Mn, Mo, Na, Ni, Pb, Rb, Se, Sr, U, V, Zn) determined by ICP-MS (ICP-MS, Agilent 7500ce, Agilent Technologies, Waldbronn, Germany), 7 elements (Br^-, Cl^-, Fl^-, NO_3^-, NO_2^-, PO_4^{3-}, SO_4^{2-}) determined by IC (DX-120, Dionex GmbH, Idstein, Germany), 2 stable isotopes (δ D, $\delta^{18}O$) and EC (measured in situ with a handheld multiparameter instrument (pH/Cond 340i, WTW, Weilheim, Germany)). Stable isotopic analysis of water samples was conducted with the Liquid-Water Isotope Analyzer (DLT 100, Los Gatos, USA) based on off-axis integrated cavity output spectroscopy (OA-ICOS).

All values are reported in per mil (‰) units relative to Vienna Standard Mean Ocean Water (VSMOW). Repeatability of δD and $\delta^{18}O$ were ± 0.6 ‰ and ± 0.2 ‰, respectively. A detailed description of lab methods as well as analytical methods and precision is given in *Barthold et al.* [2010] and *Fröhlich et al.* [2007]. From all available tracers, 14 were chosen which are commonly assumed to exhibit conservative behavior. Among them are Li, Fe, Rb, Sr, U, Na, Mg, K, Ca, Cl^-, SO_4^{2-}, δ D, $\delta^{18}O$ and EC.

Site name	Description	Year of sampling	n	Depth to GW aquifer [m]
$Rain_h$	Rain in the headwater area	2008	3	-
$Rain_{out}$	Rain near the catchment outlet	2008	2	-
H	stream water of headwater	2006	9	-
T1	Stream water of a tributary	2006	7	-
T2	Spring in the headwater area	2008	5	-
G1	Groundwater well 4 m	2006	9	4
G2	Groundwater well 5 m	2008	5	5
G3	Groundwater well 25 m	2008	5	25
G4	Groundwater in sand dunes	2008	2	8
G5	Groundwater station	2008	2	70-80

Table 4-1. Detailed description of end members.

4.2.3 END MEMBER MIXING ANALYSIS (EMMA) PROCEDURE

The core procedure of our approach includes end member mixing analysis as described in *Christophersen and Hooper* [1992]. It includes (1) performance of principal component analysis (PCA) and the Rule of One to determine the dimensionality of the hydrologic

system (2) the identification of end members, and (3) the calculation of each end member`s contribution to stream flow.

PCA was applied to the stream water time series at the catchment outlet. PCA is a commonly applied technique that is used to explore and describe the variability of a multivariate dataset. The purpose of PCA is to find a lower dimensional space, U space, in which the stream water observations lie and which describe the variability of the dataset. As opposed to U Space, S Space is defined as the vector space of the original solutes. The dimensionality of U space is determined by the number of eigenvectors, or principal components (PCs), that are retained from the PCA. We applied the Rule of One [*Joreskog et al.*, 1976] to determine the number of principal components to retain. With the Rule of One, the last principal component retained in the dataset needs to explain 1/nth of the variance where n is the number of solutes used as tracers. The number of principal components plus one is then the number of end members that are needed to describe the system. The medians of all possible end members were projected into the U space of the stream water. Those end members that best met the constraints of mixing model theory as described by *Christophersen and Hooper* [1992] were identified. The constraints include, (1) the distance between end member medians in U Space and S Space is relatively small, (2) the end members circumscribe the stream water data in U Space and (3) the distance between the stream water and the end member medians in U Space is relatively small.

The following set of equations [*Christophersen et al.*, 1990] was solved to calculate the contribution of each end member to stream flow:

(1) $1 = a_1 + a_2 + a_3$

(2) $SW_{U1} = a_1EM_{1\,U1} + a_2EM_{2\,U1} + a_3EM_{3\,U1}$

(3) $SW_{U2} = a_1EM_{1\,U2} + a_2EM_{2\,U2} + a_3EM_{3\,U2}$

where a_1, a_2 and a_3 are the fractions of each end member, SW_{U1} and SW_{U2} are the projected stream water observations in U space coordinates and $EM_{n\ U1}$ and $EM_{n\ U2}$ are the coefficients of the nth end member projected in U space. This set of equations is an example for a three end member problem. It can easily be extended or reduced for larger or smaller sized end member problems, respectively.

4.2.4 AUTOMATION

The objective of our study is to compare the outcomes of EMMA using different tracers in sets that vary in size and composition. The approach that we developed to address our above mentioned objectives includes an automation of the procedures that are involved in EMMA while iteratively varying the tracer set size and composition. We were using a set of 14 tracers and conducted EMMA for every possible tracer set size and combination beginning with a tracer set size of 2 tracers up to a set that included all 14 tracers. This resulted in 16.369 possibilities. For every tracer set possibility, the procedure in section 2.3 was conducted. Since it was impossible to screen every EMMA result, i.e. to look at every mixing space including the end member positions with reference to the stream water and at every end member hydrograph contribution in every end member combination in each tracer set possibility, we defined three criteria upon which an automatic decision about one best end member combination could be made. Based upon these criteria, we were able to determine automatically whether or not one plausible end member combination existed for each tracer set.

The three criteria are based on the constraints in mixing model theory which are presented by *Renner* [1993] and applied by *Christophersen and Hooper* [1992]. According to the mixing model theory a stream mixture is a convex combination of its end members and the concentrations of the end members in U space circumscribe the stream water data. In mathematical terms a convex combination is a geometric problem and describes the linear combination of points where all coefficients are none-negative and sum up to one. In equations (1) to (3) SW is identified as convex combination with the constraints:

(4) $a_i \geq 0$ and $\sum a_i = 1$

where a_i is the *i*th end member.

To meet these mixing model constraints our criteria included (1) the Euclidean distance of projected end members in U Space to observed end members in S Space, i.e. the percent difference between the original and the projected end members, (2) the amount of deviation of end member contribution to the hydrograph from the plausible range between 0 and 100 % and (3) the Euclidean distance between the median of the end member to the median of the stream water in U Space.

For the first criterion, the medians of the end members were projected into U Space of the stream water and their distances to the median of the original end members were calculated. The end members were destandardized and expressed in terms of solutes as opposed to in terms of PCs. The mean of the Euclidean differences was calculated and an acceptable fit was defined at 15 % difference over all solutes for every end member as suggested by *James and Roulet* [2006]. The first criterion intends to exclude those EM that are too far away from being able to explain the stream water as a convex combination. Hence, it reduces the number of possible end members in a combination excluding those end members whose projections plot too far away from the original values.

The second criterion is based on the assumption that over- and underprediction indicate poor circumscription of the stream water data. The larger the over- or underprediction of the contribution of an end member to the hydrograph is the more stream water data points lie outside of the space that is spanned by the end members and the less likely it is that the selected end members are plausible. Therefore, the contributions of every end member in every possible end member combination for every tracer set were calculated and the mean value of absolute over- and underpredictions was determined. That end member combination with the smallest deviation from the plausible range was selected. If two combinations resulted with the exact same amount of deviation, then criterion 3 was used for further evaluation. In addition, those combinations with relative deviations larger than 50 % were discarded. This decision is based upon the idea that results with such large deviations from the plausible range are regarded as not acceptable.

The third criterion is defined to ensure that the projected end members are similar to the stream water and hence may be considered as plausible. A previous study showed that some of the end members may plot too far away from the stream water indicating that

their concentrations are too extreme to explain the mixture of the stream water [*Barthold et al.*, 2010]. Hence, the smaller the distance between an end member and the stream water, the more plausible it is that this end member contributes significantly to the stream. Therefore, end members were projected into U Space and the Euclidean distance to the median of the projected stream water median was calculated. The distances of end members in each combination were averaged and the combination with the smallest distance was selected. All calculations were done using the software R version 2.8.1 [*R Development Core Team*, 2008].

We then took a closer look at (1) the number of end members identified through PCA and the Rule of One, (2) the dimensionality of remaining models, end member combinations and contributions resulting from tracer sets that produced plausible results and (3) the size and composition of remaining tracer sets.

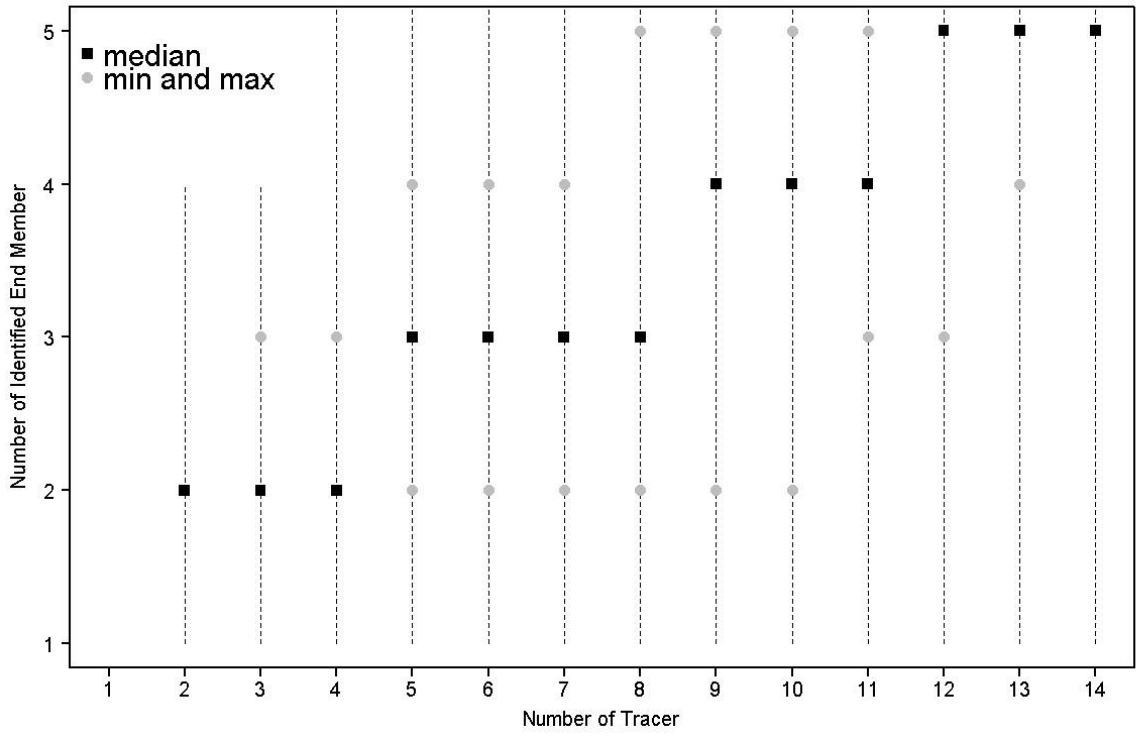

Figure 4-2. Summary of all possible results of the principal component analysis and the Rule of One where the number of identified end members is plotted as a function of the tracer set size.

4.3 RESULTS

4.3.1 PRINCIPAL COMPONENT ANALYSIS AND EIGENVECTOR ANALYSIS

Figure 4-2 presents a summary of the results of the PCA and the Rule of One. The summary includes the number of identified end members for each of the 16.369 tracer sets. This number is plotted as a function of the tracer set size. The medians of identified end member increase with increasing tracer set size. The smallest dimension of a system is 1, hence, 2 end members. One-dimensional systems can be identified with tracer sets that include 2 to up to 10 tracers. Larger tracer sets implicate a higher dimensionality. The highest number of end member that can be identified with 14 tracers is 5. But 5 end members can also be identified with tracer sets containing 8 until 13 tracers. The range of identified end member is largest for tracer sets containing 8, 9 and 10 tracer, i.e. 2 to 5 end members.

Table 4-2 presents the correlation matrix of all stream water solutes. It shows high values of the coefficients of determination for the bivalent alkaline earth metals Ca, Mg and Sr and the alkali metals Li and Na indicating that they are highly correlated with each other. In contrast, the coefficients of determination for U and SO_4 are below 0.25 indicating little correlation between these two tracers.

	Li	Fe	Rb	Sr	U	Na	Mg	K	Ca	Cl	SO_4	δD	$\delta^{18}O$	EC
Li	1.00	-0.42	0.65	0.76	-0.02	0.94	0.90	0.71	0.84	0.67	-0.07	0.59	0.51	0.60
Fe	-	1.00	-0.30	-0.18	0.06	-0.50	-0.41	-0.61	-0.39	-0.73	0.32	-0.24	-0.41	-0.43
Rb	-	-	1.00	0.51	0.36	0.65	0.61	0.56	0.58	0.52	-0.11	0.40	0.26	0.35
Sr	-	-	-	1.00	0.13	0.64	0.91	0.46	0.94	0.35	-0.40	0.48	0.45	0.72
U	-	-	-	-	1.00	-0.06	0.06	-0.17	0.08	-0.01	-0.07	-0.29	-0.38	-0.20
Na	-	-	-	-	-	1.00	0.88	0.79	0.81	0.73	-0.02	0.61	0.48	0.59
Mg	-	-	-	-	-	-	1.00	0.68	0.98	0.61	-0.30	0.57	0.50	0.74
K	-	-	-	-	-	-	-	1.00	0.65	0.67	-0.18	0.58	0.60	0.58
Ca	-	-	-	-	-	-	-	-	1.00	0.57	-0.37	0.57	0.50	0.78
Cl	-	-	-	-	-	-	-	-	-	1.00	-0.11	0.35	0.33	0.52
SO_4	-	-	-	-	-	-	-	-	-	-	1.00	-0.19	-0.41	-0.49
δD	-	-	-	-	-	-	-	-	-	-	-	1.00	0.71	0.79
$\delta^{18}O$	-	-	-	-	-	-	-	-	-	-	-	-	1.00	0.70
EC	-	-	-	-	-	-	-	-	-	-	-	-	-	1.00

Table 4-2. Correlation matrix of stream water concentrations.

4.3.2 EMMA - TRACER SET SIZES AND COMPOSITION

From the wealth of possible tracer sets, i.e. 16.369, only 23 tracer sets remain that deliver an end member combination which best meets the constraints of all three criteria (Table 4-3). The size of the remaining tracer sets varies between 3 tracers and 8 tracers. Mg and U occur in most of the tracer sets whereas K and Fe are never used (Table 4-4). All of the remaining tracer sets exclusively identify 3-EM systems, however, end member combinations vary.

End member combination	T1, G4, G5	H, G4, G5	Rain, H,G5	T1, G3, G5	T1, H, G3	H, G3, G4	T1, H, G5	G3, G4, G5
Frequency [%]	26.1	21.7	8.7	8.7	8.7	8.7	8.7	4.3
(n)	(6)	(5)	(2)	(2)	(2)	(2)	(2)	(1)

Table 4-3. Frequencies of end member combinations that produce plausible results as defined by the three criteria listed in descending order from left to right.

4.3.3 EMMA – DIMENSIONALITY AND END MEMBER COMBINATIONS

In this study, we wanted to investigate how reproducible the end member combinations are if the tracers used differ in nature and amount. With 9 possible end members (Table 4-1) there exist 372 combination possibilities when the size of end member sets varies between 2 and 5 as indicated by Figure 4-2. Based on the three criteria described in section 2.4, those combinations were screened for one best fit for every tracer set, i.e. one end member combination for every tracer set that passes all three criteria. Since all of the remaining tracer sets unexceptionally identify 3-EM systems, the number of possible end member combinations is reduced to 84. From the now 84 possible end member combinations only 9 met the constraints of all criteria (Table 4-3). The most often selected end member combinations are T1, G4, G5 and H, G4 and G5 (Table 4-3).

Tracer name	Mg	U	Sr	EC	SO_4^{2-}	$\delta^{18}O$	Cl^-	Na	δD	Ca	Li	Rb	K	Fe
Frequency [%]	65.2	52.2	47.8	43.5	39.1	39.1	39.1	34.8	34.8	21.7	17.4	8.7	0	0
(n)	(15)	(12)	(11)	(10)	(9)	(9)	(9)	(8)	(8)	(5)	(4)	(2)	(0)	(0)

Table 4-4. Frequencies of tracers occurring in tracer sets that identify end member combinations with plausible results listed in descending order from left to right.

We calculated the frequencies of each end member that was identified in the resulting end member combinations. Table 4-5 shows that the four most often identified end members are G4 and G5 occurring in 60.9 and 78.3 %, respectively, as well as H (the source area) and T1 which are equally often occurring with 56.5 %. This is outstandingly more frequent than all other possible end members whose frequencies range between 4.4 and 34.8 % (G3, rain and G1) (Table 4-5). G2 and T2 are never identified.

End member name	G5	G4	H	T1	G3	Rain	G1	G2	T1
Frequency [%]	78.3	60.9	56.5	56.5	34.8	8.7	4.4	0	0
(n)	(18)	(14)	(13)	(13)	(8)	(2)	(1)	(0)	(0)

Table 4-5. Frequencies of all end members that are identified in the remaining end member combinations listed in descending order from left to right.

4.3.4 EMMA – END MEMBER CONTRIBUTIONS

In addition to dimensionality and end member combination, this study aimed at investigating the reproducibility of the end member contributions to the hydrograph when varying tracer set size and composition. We therefore calculated the contributions of all end members as a result from different tracer set sizes and combinations. In the following, we present the different hydrograph contributions of the four most often selected end members G4, G5, T1 and H in the two most often selected end member combinations G4, G5, T1 (Figure 4-3) and G4, G5 and H (Figure 4-4) which resulted from different tracer sets after application of all criteria.

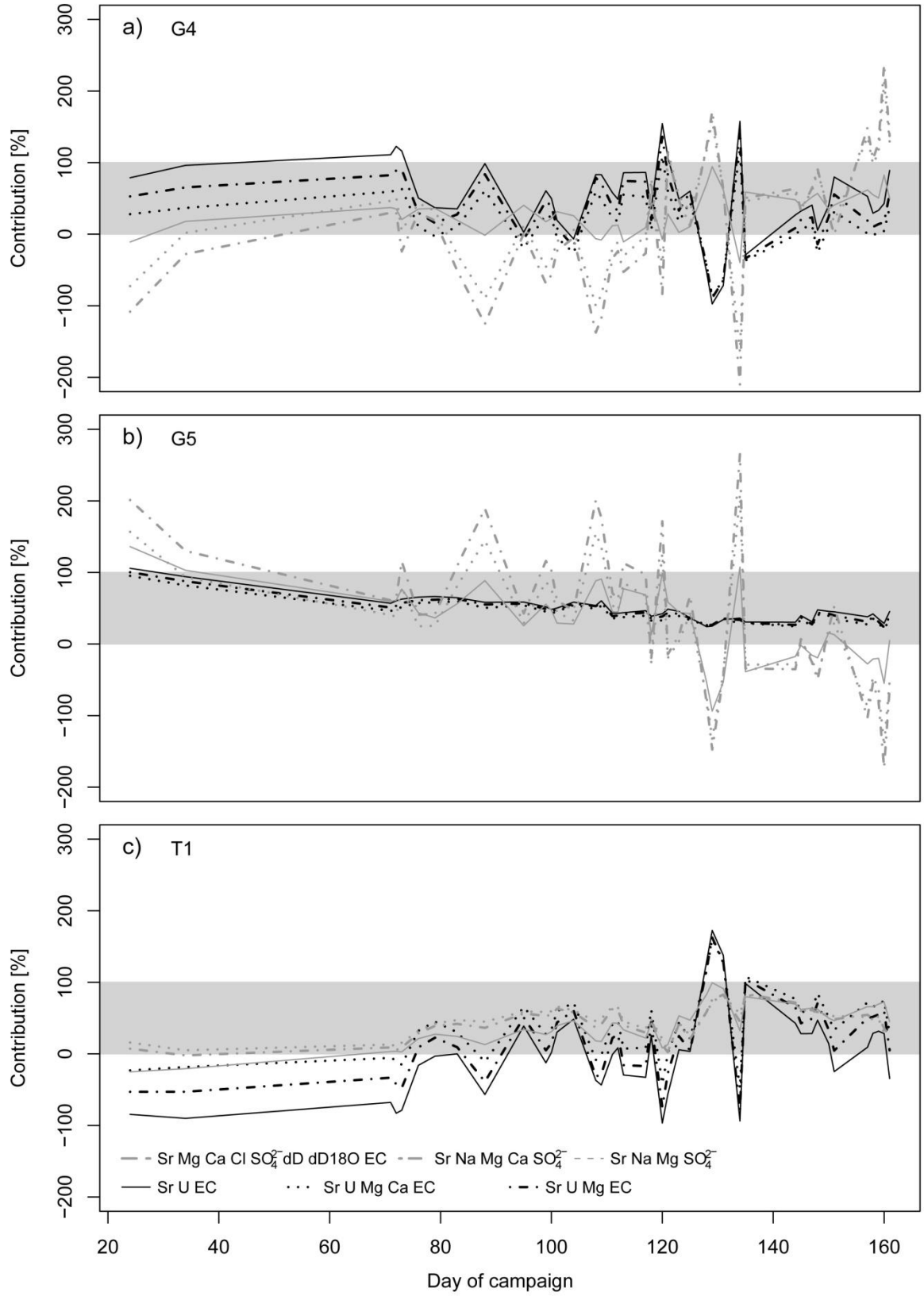

Figure 4-3. Hydrograph contributions of a) G4, b) G5 and c) T1 resulting from different tracer sets. Differences in color depict two tracer set groups defined by U and SO_4. The shaded area represents the plausible range of hydrograph contributions between 0 and 100%.

Six different tracer sets identify the end member combination G4, G5 and T1 (Figure 4-3). The contributions of these end members vary tremendously between different tracer sets and even show opposite behavior at certain time points. Figure 3a and b indicate that the largest differences between hydrograph contributions occur for G4 and G5, respectively. Opposite behavior of hydrograph contributions are most remarkable for G4 (Figure 4-3a). The contributions of T1 vary from each other but still follow the same trend with increasing contributions over the season (Figure 4-3c).

Figure 4-3a indicates that the oppositional reacting hydrograph contributions are produced by two groups of tracer sets that can be distinguished by certain tracers. The contributions that react with a negative peak on days 88, 109 and 134 are produced by tracer sets that contain SO_4 as a tracer. The other three contributions that react with a positive peak on these same days are a result of tracer sets that include U. In addition, it should be mentioned that tracer sets which contain SO_4 produce contributions that are largely underestimated at most times. The contributions of G5 display the same group patterns of tracer sets: sets that include SO_4 produce hydrograph contributions that are highly reactive over the season (Figure 4-3b). In contrast, tracer sets containing U produce contributions that stay well within the plausible range between 0 and 100 %, clearly decrease over the season and are little reactive. This separation of hydrograph contributions by tracer set combinations is less distinctive in the T1 results (Figure 4-3c). The T1 contributions all display increasing behavior over season. However, contributions of the SO_4 group stay well within the plausible range and are much less reactive than the contributions produced by the tracer sets containing U.

Five different tracer sets identify the end member combination G4, G5 and H (Figure 4-4). In comparison to the previously described end member combination, the contributions of each of these end members show much less variation. In fact, the contributions of the G5 end member lie very close together, all of them are indicating a decreasing trend over the course of the season (Figure 4-4b). G4 clearly exhibits increasing contributions over the season (Figure 4-4a). Variations between the contributions are slightly higher than for G5 (Figure 4-4a and b). Figure 4-4c shows that differences in contributions of end member H are the most pronounced in this end member combination. However, they are still less than those of T1 in the previously described end member combination (Figure 4-3c and 4-

4c). In comparison with the other two end members G4 and G5, H contributes only a small amount to the hydrograph which even decreases over the season. H seems to run dry towards the end of the season, however, contributions are largely underestimated. They fall beneath 0 % which is beyond the plausible range of hydrograph contribution.

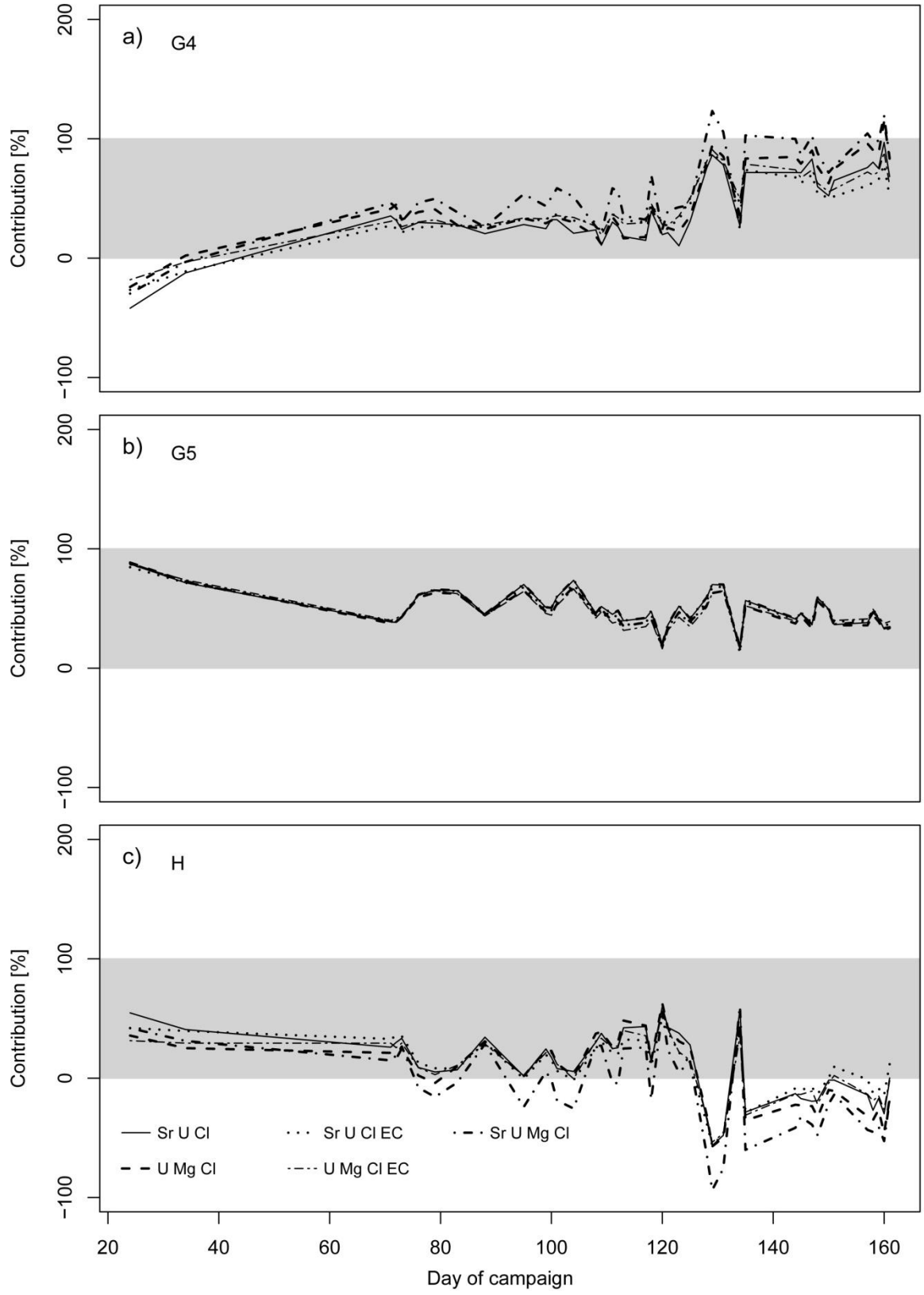

Figure 4-4. Hydrograph contributions of a) G4, b) G5 and c) H resulting from different tracer sets. Differences in color depict two tracer set groups defined by U and SO4. The shaded area represents the plausible range of hydrograph contributions between 0 and 100%.

4.4 DISCUSSION

4.4.1 TRACER SET THRESHOLD

We hypothesize that with increasing number of tracers used in an EMMA, the potential to identify a higher dimensionality of the system (i.e. a higher number of end members needed to explain flow in a system) also increases. This hypothesis is especially designed for larger catchments (e.g. > 100 km^2) and longer time periods (e.g. a whole season) which often exhibit increased spatial and temporal variability of tracer concentrations. This is based on the assumption that multiple end members exist whose contributions vary over space and time and that this multiplicity can be captured with an increased diversity of tracers used in the study. Hence, if all tracers that are able to capture this variability are included in the study then there should be a tracer set size threshold beyond which no more end members can be identified.

Our results show no indication that the increasing trend of identified dimensions with increasing number of tracers used in the analysis will level out within the available range of 14 tracers. It may be assumed that the number of dimensions will still increase when increasing the number of tracers. This assumption is based on the fact that the study area is very large and the period analyzed covers a whole season. Therefore, the inherent spatial and temporal variability of the data set might be explained by a number of sources that is larger than five. It is very reasonable that there exists a larger number of hydrological and geochemical independent sources that are expressed through the complete spectrum of tracer information in the stream water. Hence, a tracer set threshold beyond which the acquisition of insight in terms of end members that contribute flow to the stream becomes negligible could not be detected. However, it becomes clear that the size and the composition of the tracer set are crucial for the identification of dimensions in a system. Small tracer sets, in our example consisting of at least 3 or 4 tracers, are in most cases only able to identify three end members at most. If more tracers are determined to serve as input into the analysis, e.g. 13 or 14, three end members are not enough to explain the contributions to the discharge.

With an increased tracer set size the probability that three end members are identified increases. However, more tracers in a set do not necessarily mean that more end

members can be identified. This trend yet depends on the right selection of tracers. For example, with 12 tracers, there are still cases that identify a 3-EM system. This would be a "worst" case considering that most of the 12-tracer set cases are able to identify four or five end members which could be avoided by selecting a better suited tracer set.

The differences in resulting dimensionality between equal sizes of tracer sets can be explained with the coefficients of determination of certain tracers (Table 4-2). The high correlations between Ca, Mg and Sr as representatives of the alkali earth metals and Li and Na as representatives of the alkali metals indicate that they deliver similar information. This results from their chemical similarities, i.e. the bivalent nature of the elements in the former group and the monovalent nature of the elements in the latter group. Hence, the more of these tracer are included in a tracer set, the less variability can be explained and the less dimensions will be identified. Vice versa, U and SO_4 are not correlated with each other indicating that they deliver independent information. Consequently, a tracer set of a certain size is likely to indicate more significant PCs, and hence more end members, when U or SO_4 are included than when any one of the highly correlated tracers e.g. Mg is included.

4.4.2 VALIDITY OF MODEL CONCEPT

The second objective of this study aimed at investigating the effect of tracer set size and composition on the conceptual model that result from an EMMA. We thereby considered the validity of the resulting model concept under the aspects of the dimensionality of the system (the number of end members needed in a system), the end member combination and the contributions of end members to the hydrograph.

From the vast number of possible tracer sets only a small amount of them delivered an end member combination with an acceptable fit in terms of the criteria that were defined on the basis of mixing model theory. This reduction of tracer sets that lead to a result can mostly be attributed to the application of the first criterion. The majority of the end member combinations exhibit large deviations between the original and the projected end member medians which led to their exclusion as plausible end member. Hence, the

potential of obtaining highly reproducible results among varying tracer set sizes and combinations is already largely reduced.

At this point, it is important to emphasize that the definition of the three criteria in the automation procedure is really strict and does not allow for any deviations from the boundaries that we set. Such deviations may lead to different results if they occur in a single EMMA where decisions are based on a graphical approach and are guided by the mental model that the researcher already has of the watershed. For example, criterion 1 was first introduced by *Christophersen and Hooper* [1992] where it served as an additional measure to estimate the plausibility of the chosen end members at Birkenes and Panola Mountain. Based on this criterion, they concluded that the organic end member at Panola was not plausible and suggested to collect additional field data for refinement of the identified end members in the system. *James and Roulet* [2006] also included this criterion into their study in which they compared the applicability of end members across scale. They defined a 15 % threshold beyond which the fit was not considered acceptable any more. In their study, there is not one end member for which all of the solutes fall below this threshold. However, the presented end members serve as explanation of watershed processes even though some of them violate the 15 % - rule. But after all it remains unclear if they discard an end member based on this information. This indicates how much the outcomes of an EMMA depend on the preexisting understanding of processes in the area and how much this understanding influences the decision by a researcher in an EMMA procedure. In this study, these subjective decisions are excluded from any EMMA procedure and decisions completely rely on numbers.

Christophersen and Hooper [1992] and *James and Roulet* [2006] are the only studies we know that apply this criterion. In many other studies this criterion is not considered at all even though it refers to one of the main assumptions of mixing model theory. This indicates that in many other studies the application of this criterion might have delivered other results.

Dimensionality

The determination of the dimensionality of a system is the first step in an EMMA. From the two common methods in EMMA, the Rule of One [*Joreskog et al.*, 1976] and the Residual Analysis [*Hooper*, 2003], we chose the Rule of One due to the design of our methodology. The dimensionality of the system is the first aspect under which we consider the validity of the model.

All remaining models exclusively exhibit a dimensionality of three. This indicates a certain degree of reproducibility among the selected tracer sets and hence emphasizes the confidence into the validity of the model concept. The consent of the results concerning the dimensionality of the system is also in line with findings of a previous study [*Barthold et al.*, 2010] which identified the catchment as a 3-EM system.

End member combination

The second step of an EMMA includes the identification of the end members. The identified end members and their combinations in the resulting 3-EM model concepts served as the second aspect under which the validity of the model concepts could be assessed.

Among the end member combinations that were identified to deliver plausible results, four of the nine end members were noticeably often selected. The frequency with which they occur indicates a considerable reproducibility of results when varying tracer set size and composition. The repeatedly selection of the same four end members underlines the importance of these end members for the system and increases their validity as true end members. However, the fact that there are four end members that are considerably often selected in model concepts but that exclusively 3-EM systems result after the application of the criteria is already evidence for ambiguity in the results concerning the end member combinations.

The clear identification of the catchment as a 3 EM-system remarkably reduces the possible end member combinations to 84 from which only 12 combinations are repeatedly selected to best fit the system. The four most often selected end members occur in the 2 most often identified end member combinations which are T1, G4 and G5

and H, G4 and G5. These 2 end member combinations are selected in a noticeably higher amount of cases than all other end member combinations. But even though the reduction of possible model concepts to the two aforementioned end member combinations is quite an achievement these results indicate a reduced validity of model concepts. It also shows how highly sensitive the results are to the tracer set size and composition.

In addition, a certain fraction of tracer sets produces results that vary from the just mentioned majority. This indicates that the resulting model concepts are not only highly sensitive to the input tracer set but also inhere subjectivity and depend on the decisions of the researcher.

For example, in a previous study [*Barthold et al.*, 2010] we showed that the discharge of the Xilin river can be explained with three interannually changing end members if 7 tracers are used (Li, Rb, Sr, Na, Mg, Cl and EC). The study was based on a three year time series and the selection of tracers was restricted to data availability. The three end members that were identified in 2006 are rain, G2 and G4. This combination of end members differs from all results that were produced in this study. However, if we look at the other two years, identified end members were rain, G3 and G5 in 2007 and rain, G3 and G4 in 2008. These end member combinations vary significantly from the results that we produced in this study.

One explanation may be that the automatic procedure produces results that differ from results which are based on the mental model that the researcher has of the catchment. Even if our three criteria are designed to make decisions that strictly follow the mixing model theory there remains a fraction of decisions that maybe judged differently if done on a subjective basis: we already mentioned criterion 1 with respect to the decision if an end member is plausible or not. In addition, criterion 3 which calculates the Euclidean distance between the end member median and the stream water median completely relies on the distance but not on the position of the end member. The position of the end member in relation to the stream water is, however, considered by the amount of over- or underprediction. Hence, a subjective decision may lead to a different result.

A more likely explanation for the discrepancies between the previous study and this study is that we did not calculate the Euclidean distance between original and projected end members as implemented in our criterion 1 in the previous study [*Barthold et al.*, 2010].

Criterion 1 is responsible for the large reduction of selected tracer sets that deliver a plausible result. Not many studies apply such a criterion. We might have decided for a different end member combination if we had applied this criterion. This may also be the explanation for the fact that not one case exists with the same end member combination as in the previous study [*Barthold et al.*, 2010].

End member contributions

The last aspect under which we assess the validity of the model concepts is the contributions of the end members to the hydrograph. Overall, the reproducibility of hydrograph contributions for the end member combination H, G4 and G5 is superior to the end member combination T1, G4 and G5. Similar trends of contribution over the course of the season are captured for the G5 end member in both combinations. This increases the confidence into G5 as being a true end member.

However, the seasonal trend of the G4 end member varies between the two combinations. In addition, differences between contributions produced by different tracer sets are tremendous for the G4 and G5 end member in the combination with T1. The fact, that the differences in contributions can be attributed to two different tracer set groups that are defined by containing either U or SO_4, can be explained as follows.

The concentrations of U and SO_4 in the stream water do not correlate with each other (Table 4-2). This indicates that the two tracers deliver independent information and results in end member contributions that vary from each other. However, the oppositional behavior of the contributions cannot be explained by this alone. The oppositional behavior is rather a mathematical artifact that is caused by the fact that the selected end member does not fit to the stream water composition as defined by the U space of the two tracer set combinations. The first PCs of both tracer sets contain very similar information as is indicated by a high value of the coefficient of determination when regressing both components with each other ($R^2 = 0.82$). These PCs represent the general positive correlation between most solutes with the exception of U and SO_4. Tracer set combinations that contain U project G4 with a higher than average value on the first PC than do those tracer sets that contain SO_4. This leads to the oppositional

behavior of end member contributions as indicated in Figure 4-3a and b: stream water with predominantly high solute concentrations is interpreted to consist to a high proportion of G4 when the tracer set contains U and to a low proportion when the tracer set contains SO_4. Despite the rejection of EM combinations with U space - S space distances greater 15 %, dislocation due to projection seems to still be a large source of uncertainty for predictions of EM contributions. We therefore stress the importance of including this criterion into EMMA-based studies. Ways to express the uncertainty due to U space - S space distances of end members seems to be a necessary extension of the EMMA approach.

Our results show that the predicted relevance of end members varies with different tracer sets and that G4 is a poor selection. Since this proposed end member combination does not deliver explicit results we suggest that we are missing at least one other end member that is able to explain the stream water in combination with G5 and T1. We conclude that a higher number of tracers provides an enhanced capacity to avoid false conclusions that may arise from small, limited tracer sets.

4.4.3 SELECTION OF TRACER SET SIZE AND COMPOSITION

The primary objective of this study was to provide information that may guide decisions about tracer set size and composition during the design stage of a tracer study. Our results give evidence that some tracers are superior to other tracers in terms of producing plausible results, at least for the conditions in the Xilin catchment. For example, the major cations have superiority in the following order: Mg, Na, Ca and K. However, K never even appeared in a tracer set that produced plausible results. And only Mg was selected outstandingly often. In contrast, some of the minor cations, e.g. U, Sr and SO_4, occur noticeably often in tracer sets that produce plausible results. Frequencies of tracers in sets that produce plausible results do differ from those tracers that are most often used in tracer studies, these often being a selection of 2 to 6 from the following: Ca, K, Mg, Na, Cl^-, SO_4^{2-}, Si, the isotopes $\delta^{18}O$ and δD, acid neutralizing capacity (ANC), alkalinity and EC [*Burns et al.*, 2001; *Elsenbeer et al.*, 1995; *Katsuyama et al.*, 2009; *F. J. Liu et al.*, 2008; *Soulsby, Rodgers, et al.*, 2003].

Despite the fact that the selection of the tracer set always depends on the unique character of the study area, one should keep in mind that minor trace elements may contribute information that cannot be inferred from the major elements. However, our results do not allow us to suggest certain tracers over others, in part because the resulting model concepts are poorly reproducible and we lack important understanding of the study area. This may be overcome by conducting a simulation study which analyzes known, artificial solutions of end members and stream water with the methodology described in this study.

4.5 CONCLUSIONS

We investigated the value of large tracer sets versus small tracer sets in EMMA studies. To accomplish this, we developed an automated approach to conduct EMMA iteratively, varying tracer set sizes and compositions beginning with tracer set sizes of 2 up to tracer sets with 14 tracers. We were not able to identify a tracer set threshold beyond which the acquisition of insight became negligible in terms of system dimensionality. We could show that model concepts are only reproducible to a small degree. Hence, the uncertainty in model concept validity among different tracer set sizes and compositions is high. The large differences in end member contributions that were produced among the different tracer sets demonstrated that we are still missing at least one important end member. It especially suggests that the major elements are not always the most useful tracers. Larger tracer sets that include also minor trace elements and that are employed in an iterative EMMA approach with varying tracer set sizes and compositions have a high potential to avoid false conclusions about catchment functioning.

At this point, we also need to emphasize that process understanding as a mechanism to pre-define end members is a key component of the vast majority of EMMA studies. However, our results prompt us to stress the fact that an iterative EMMA approach as presented in this study might be very complementary. Our iterative EMMA approach outlined some significant uncertainty in the model concept and also a missing end member that may not have been apparent from the traditional approach.

5 REFERENCES

Acuna, V., and C. N. Dahm (2007), Impact of monsoonal rains on spatial scaling patterns in water chemistry of a semiarid river network, *J. Geophys. Res.-Biogeosci.*, *112*(G4), 11, doi:G04009 10.1029/2007jg000493.

Andreassian, V. (2004), Waters and forests: from historical controversy to scientific debate, *J. Hydrol.*, *291*(1-2), 1-27, doi:10.1016/j.jhydrol.2003.12.015.

Arnold, J. G., R. Srinivasan, R. S. Muttiah, and J. R. Williams (1998), Large area hydrologic modeling and assessment - Part 1: Model development, *J. Am. Water Resour. Assoc.*, *34*(1), 73-89.

Asner, G. P., A. J. Elmore, L. P. Olander, R. E. Martin, and A. T. Harris (2004), Grazing systems, ecosystem responses, and global change, *Annu. Rev. Environ. Resour.*, *29*, 261-299.

Bai, Y., X. Han, J. Wu, Z. Chen, and L. Li (2004), Ecosystem stability and compensatory effects in the Inner Mongolia grassland, *Nature*, *431*(7005), 181-184, doi:10.1038/nature02850.

Barthold, F. K., T. Sayama, K. Schneider, L. Breuer, K. Vaché, H. G. Frede, and J. J. McDonnell (2008), Gauging the ungauged basin: a top-down approach in a large semiarid watershed in China, *Adv. Geosci.*, *18*, 3-8.

Barthold, F. K., J. Wu, K. Vaché, K. Schneider, H. G. Frede, and L. Breuer (2010), Identification of geographic runoff sources in a data sparse region: hydrological processes and the limitations of tracer-based approaches, *Hydrol. Process.*, doi:10.1002/hyp.7678.

Barthold, F. K., M. Wiesmeier, L. Breuer, H.-G. Frede, J. Wu, F.B. Blank (under review), Digital soil mapping with Random Forest in the grasslands of Inner Mongolia, submitted to *Eur. J. Soil Sci.*

Bazemore, D. E., K. N. Eshleman, and K. J. Hollenbeck (1994), The role of soil-water in stormflow generation in a forested headwater catchment - Synthesis of natural tracer and hydrometric evidence, *J. Hydrol.*, *162*(1-2), 47-75.

Berman, E. S. F., M. Gupta, C. Gabrielli, T. Garland, and J. J. McDonnell (2009), High-frequency field-deployable isotope analyzer for hydrological applications, *Water Resour. Res.*, *45*, 7, doi:W10201 10.1029/2009wr008265.

Bernal, S., A. Butturini, and F. Sabater (2006), Inferring nitrate sources through end member mixing analysis in an intermittent Mediterranean stream, *Biogeochemistry*, *81*(3), 269-289, doi:10.1007/s10533-006-9041-7.

Beven, K. (2001), *Rainfall-Runoff Modelling: The Primer*, Wiley, Chichester.

Binley, A. M., K. J. Beven, A. Calver, and L. G. Watts (1991), Changing responses in hydrology - assessing the uncertainty in physically based model predictions, *Water Resour. Res.*, *27*(6), 1253-1261.

Bosch, J. M., and J. D. Hewlett (1982), A review of catchment experiments to determine the effect of vegetation changes on water yield and evapotranspiration, *J. Hydrol.*, *55*(1-4), 3-23.

Brogaard, S., and X. Y. Zhao (2002), Rural reforms and changes in land management and attitudes: A case study from Inner Mongolia, China, *Ambio*, *31*(3), 219-225.

Brown, A. E., L. Zhang, T. A. McMahon, A. W. Western, and R. A. Vertessy (2005), A review of paired catchment studies for determining changes in water yield resulting from alterations in vegetation, *J. Hydrol.*, *310*(1-4), 28-61, doi:10.1016/j.jhydrol.2004.12.010.

Burke, I. C., W. K. Lauenroth, M. A. Vinton, P. B. Hook, R. H. Kelly, H. E. Epstein, M. R. Aguiar, M. D. Robles, M. O. Aguilera, K. L. Murphy, and R. A. Gill (1998), Plant-soil interactions in temperate grasslands, *Biogeochemistry*, *42*(1-2), 121-143.

Burns, D. A. (2002), Stormflow-hydrograph separation based on isotopes: the thrill is gone - what's next?, *Hydrol. Process.*, *16*(7), 1515-1517, doi:10.1002/hyp.5008.

Burns, D. A., J. J. McDonnell, R. P. Hooper, N. E. Peters, J. E. Freer, C. Kendall, and K. Beven (2001), Quantifying contributions to storm runoff through end-member mixing analysis and hydrologic measurements at the Panola Mountain Research Watershed (Georgia, USA), *Hydrol. Process.*, *15*(10), 1903-1924.

Chaves, J., C. Neill, S. Germer, S. G. Neto, A. Krusche, and H. Elsenbeer (2008), Land management impacts on runoff sources in small Amazon watersheds, *Hydrol. Process.*, *22*(12), 1766-1775, doi:10.1002/hyp.6803.

Chen, S. Q., X. M. Xiao, J. Y. Liu, and D. F. Zhuang (2002), Observation of land use/cover change of the Xilin River Basin, Inner Mongolia, using multi-temporal Landsat images, in *International Conference on Ecosystems Dynamics, Ecosystem-Society Interactions, and Remote Sensing Applications for Semi-Arid and Arid Land*, edited by X. L. Pan, W. Gao, M. H. Glantz, and Y. Honda, pp. 674-685, SPIE - Int Soc Optical Engineering, Hangzhou, Peoples R China.

Chen, Z. (1988), Topography and climate of the Xilin River basin, in *Inner Mongolia Research Station of Academia Sinica, Research on Grassland Ecosystem*, pp. 13-22, Science Press, Beijing.

Christophersen, N., and R. P. Hooper (1992), Multivariate analysis of stream water chemical data: the use of principal components analysis for the end member mixing problem, *Water Resour. Res.*, *28*(1), 99-107.

Christophersen, N., C. Neal, R. P. Hooper, R. D. Vogt, and S. Andersen (1990), Modelling streamwater chemistry as a mixture of soilwater end-members - A step towards second-generation acidification models, *J. Hydrol.*, *160*, 307-320.

Dewalle, D. R., B. R. Swistock, and W. E. Sharpe (1988), Three-component tracer model for stormflow on a small Appalachian forested catchment, *J. Hydrol.*, *104*(1-4), 301-310.

Dinçer, T., A. Al-Mugrin, and U. Zimmermann (1974), Study of the infiltration and recharge through the sand dunes in arid zones with special reference to the stable isotopes and thermonuclear tritium, *J. Hydrol.*, *23*(1-2), 79-109.

Elsenbeer, H., D. Lorieri, and M. Bonell (1995), Mixing model approaches to estimate storm flow sources in an overland flow-dominated tropical rain forest catchment, *Water Resour. Res.*, *31*(9), 2267-2278.

Epstein, H. E., I. C. Burke, A. R. Mosier, and G. L. Hutchinson (1998), Plant functional type effects on trace gas fluxes in the shortgrass steppe, *Biogeochemistry*, *42*(1-2), 145-168.

Evans, R. D., R. Rimer, L. Sperry, and J. Belnap (2001), Exotic plant invasion alters nitrogen dynamics in an arid grassland, *Ecol. Appl.*, *11*(5), 1301-1310.

Fröhlich, H. L., L. Breuer, H. G. Frede, J. A. Huisman, and K. B. Vache (2007), Water source characterization through spatiotemporal patterns of major, minor and trace element stream concentrations in a complex, mesoscale German catchment, *Hydrol. Process.*, *22*(12), 2028-2043, doi:10.1002/hyp.6804.

Fröhlich, H. L., L. Breuer, K. B. Vache, and H. G. Frede (2008), Inferring the effect of catchment complexity on mesoscale hydrologic response, *Water Resour. Res.*, *44*(9), 15, doi:W09414 10.1029/2007wr006207.

Grayson, R. B., C. J. Gippel, B. L. Finlayson, and B. T. Hart (1997), Catchment-wide impacts on water quality: the use of 'snapshot' sampling during stable flow, *J. Hydrol.*, *199*(1-2), 121-134.

Guo, J. K., J. Y. Liu, G. M. Huang, D. F. Zhuang, and H. M. Yan (2004), Cropland identification in Inner Mongolia, China with SPOT-4 VEGETATION imagery, in *IEEE International Geoscience and Remote Sensing Symposium*, pp. 3988-3991, IEEE, Anchorage, AK. [online] Available from: ://000227006901051

Hill, A. R. (1993), Base cation chemistry of storm runoff in a forested headwater wetland, *Water Resour. Res.*, *29*(8), 2663-2673.

Hooper, R. P. (2001), Applying the scientific method to small catchment studies: a review of the Panola Mountain experience, *Hydrol. Process.*, *15*(10), 2039-2050, doi:10.1002/hyp.255.

Hooper, R. P. (2003), Diagnostic tools for mixing models of stream water chemistry, *Water Resour. Res.*, *39*(3), 13, doi:1055 10.1029/2002wr001528.

IUSS Working Group WRB (2007), *World Reference Base for Soil Resources 2006, first update 2007*, FAO, Rome.

James, A. L., and N. T. Roulet (2006), Investigating the applicability of end-member mixing analysis (EMMA) across scale: A study of eight small, nested catchments in a temperate forested watershed, *Water Resour. Res.*, *42*(8), 17, doi:W08434 10.1029/2005wr004419.

Jian, Y., and M. Meurer (2001), Die Steppen Nordchinas und ihre Belastung durch weide- und landwirtschaftliche Nutzung, *Geographische Rundschau*, *53*, 48-52.

Joreskog, K., J. Klovan, and R. Reyment (1976), *Geological Factor Analysis*, Methods in Geomathematics, Elsevier, Amsterdam.

Katsuyama, M., N. Kabeya, and N. Ohte (2009), Elucidation of the relationship between geographic and time sources of stream water using a tracer approach in a headwater catchment, *Water Resour. Res.*, *45*, 13, doi:W06414 10.1029/2008wr007458.

Kennedy, V. C., C. Kendall, G. W. Zellweger, T. A. Wyerman, and R. J. Avanzino (1986), Determination of the components of stormflow using water chemistry and environmental isotopes, Mattole River Basin, California, *J. Hydrol.*, *84*(1-2), 107-140.

Ketzer, B., H. Z. Liu, and C. Bernhofer (2008), Surface characteristics of grasslands in Inner Mongolia as detected by micrometeorological measurements, *Int. J. Biometeorol.*, *52*(7), 563-574, doi:10.1007/s00484-008-0148-5.

Klemeš, V. (1983), Conceptualization and scale in hydrology, *J. Hydrol.*, *65*(1-3), 1-23.

Kobayashi, H., T. Sugiura, M. Uto, and S. Suzuki (1994), International study on the pasture improvement in the Inner Mongolia desert-steppe, China., in *Proceedings of the International Symposium on Grassland Ressources*, edited by B. Li, pp. 1007-1009, Agricultural Sciencetech Press, Beijing.

Ladouche, B., A. Probst, D. Viville, S. Idir, D. Baque, M. Loubet, J. L. Probst, and T. Bariac (2001), Hydrograph separation using isotopic, chemical and hydrological approaches (Strengbach catchment, France), *J. Hydrol.*, *242*(3-4), 255-274.

Li, X. R. (2001), Study on shrub community diversity of Ordos Plateau, Inner Mongolia, Northern China, *J. Arid. Environ.*, *47*(3), 271-279, doi:10.1006/jare.2000.0707.

Lin, G. C. S., and S. P. S. Ho (2003), China's land resources and land-use change: insights from the 1996 land survey, *Land Use Policy*, *20*(2), 87-107.

Lis, G., L. I. Wassenaar, and M. J. Hendry (2008), High-precision laser spectroscopy D/H and O-18/O-16 measurements of microliter natural water samples, *Anal. Chem.*, *80*(1), 287-293, doi:10.1021/ac701716q.

Liu, F. J., R. Parmenter, P. D. Brooks, M. H. Conklin, and R. C. Bales (2008), Seasonal and interannual variation of streamflow pathways and biogeochemical implications in semi-arid, forested catchments in Valles Caldera, New Mexico, *Ecohydrology*, *1*(3), 239-252, doi:10.1002/eco.22.

Liu, F. J., M. W. Williams, and N. Caine (2004), Source waters and flow paths in an alpine catchment, Colorado Front Range, United States, *Water Resour. Res.*, *40*(9), 16, doi:W09401 10.1029/2004wr003076.

Lyon, S. W., S. L. E. Desilets, and P. A. Troch (2009), A tale of two isotopes: differences in hydrograph separation for a runoff event when using delta D versus delta O-18, *Hydrol. Process.*, *23*(14), 2095-2101, doi:10.1002/hyp.7326.

Masters, R. A., and R. L. Sheley (2001), Principles and practices for managing rangeland invasive plants, *J. Range Manage.*, *54*(5), 502-517.

McDonnell, J. J., M. Sivapalan, K. Vache, S. Dunn, G. Grant, R. Haggerty, C. Hinz, R. Hooper, J. Kirchner, M. L. Roderick, J. Selker, and M. Weiler (2007), Moving beyond heterogeneity and process complexity: A new vision for watershed hydrology, *Water Resour. Res.*, *43*(7), 6, doi:W07301 10.1029/2006wr005467.

McGuire, K. J., and J. J. McDonnell (2006), A review and evaluation of catchment transit time modeling, *J. Hydrol.*, *330*(3-4), 543-563, doi:10.1016/j.jhydrol.2006.04.020.

Mosier, A. R., W. J. Parton, D. W. Valentine, D. S. Ojima, D. S. Schimel, and O. Heinemeyer (1997), CH4 and N2O fluxes in the Colorado shortgrass steppe .2. Long-term impact of land use change, *Glob. Biogeochem. Cycle*, *11*(1), 29-42.

Nagler, P. L., E. P. Glenn, O. Hinojosa-Huerta, F. Zamora, and K. Howard (2008), Riparian vegetation dynamics and evapotranspiration in the riparian corridor in the delta of the Colorado River, Mexico, *J. Environ. Manage.*, *88*(4), 864-874, doi:10.1016/j.jenvman.2007.04.010.

Neal, C., J. Wilkinson, M. Neal, M. Harrow, H. Wickham, L. Hill, and C. Morfitt (1997), The hydrochemistry of the headwaters of the River Severn, Plynlimon, *Hydrol. Earth Syst. Sci.*, *1*(3), 583-617.

Onstad, C. A., and D. G. Jamieson (1970), Modelling effect of land use modifications on runoff, *Water Resour. Res.*, *6*(5), 1287-&.

Pinder, G. F., and J. F. Jones (1969), Determination of ground-water component of peak discharge from chemistry of total runoff, *Water Resour. Res.*, *5*(2), 438-&.

Polley, H. W., C. R. Tischler, and H. B. Jobnson (2006), Elevated atmospheric CO2 magnifies intra-specific variation in seedling growth of honey mesquite: An assessment of relative growth rates, *Rangel. Ecol. Manag.*, *59*(2), 128-134.

Pyankov, V. I., P. D. Gunin, S. Tsoog, and C. C. Black (2000), C-4 plants in the vegetation of Mongolia: their natural occurrence and geographical distribution in relation to climate, *Oecologia*, *123*(1), 15-31.

R Development Core Team (2008), *R: A language and environment for statistical computing*, R Foundation for Statistical Computing, Vienna, Austria. [online] Available from: http://www.R-project.org

Renner, R. M. (1993), The resolution of a compositional data set into mixtures of fixed source compositions, *Appl. Stat.-J. R. Stat. Soc.*, *42*(4), 615-631.

Reynolds, J. F., D. M. Stafford Smith, E. F. Lambin, B. L. Turner, M. Mortimore, S. P. J. Batterbury, T. E. Downing, H. Dowlatabadi, R. J. Fernandez, J. E. Herrick, E. Huber-Sannwald, H. Jiang, R. Leemans, T. Lynam, F. T. Maestre, M. Ayarza, and B. Walker (2007), Global desertification: Building a science for dryland development, *Science*, *316*(5826), 847-851, doi:10.1126/science.1131634.

Rice, K. C., and G. M. Hornberger (1998), Comparison of hydrochemical tracers to estimate source contributions to peak flow in a small, forested, headwater catchment, *Water Resour. Res.*, *34*(7), 1755-1766.

Rimon, Y., O. Dahan, R. Nativ, and S. Geyer (2007), Water percolation through the deep vadose zone and groundwater recharge: Preliminary results based on a new vadose zone monitoring system, *Water Resour. Res.*, *43*(5), 12, doi:W05402 10.1029/2006wr004855.

Runnström, M. C. (2003), Rangeland development of the Mu Us Sandy Land in semiarid China: an analysis using Landsat and NOAA remote sensing data, *Land Degrad. Dev.*, *14*(2), 189-202.

Schäfer, J. (2009), Auswirkungen unterschiedlicher Landnutzungsszenarien auf den Wasserhaushalt im Einzugsgebiet des Xilin, Innere Mongolei, China, Diplomarbeit, Philipps-Universität Marburg.

Schneider, K., J. A. Huisman, L. Breuer, Y. Zhao, and H. G. Frede (2008), Temporal stability of soil moisture in various semi-arid steppe ecosystems and its application in remote sensing, *J. Hydrol.*, *359*(1-2), 16-29, doi:10.1016/j.jhydrol.2008.06.016.

Schneider, K., B. Ketzer, L. Breuer, K. Vaché, C. Bernhofer, and H. G. Frede (2007), Evaluation of evapotranspiration methods for model validation in a semi-arid watershed in northern China, *Adv. Geosci.*, *11*, 37-42.

Seibert, J., and J. J. McDonnell (2002), On the dialog between experimentalist and

modeler in catchment hydrology: Use of soft data for multicriteria model calibration, *Water Resour. Res.*, *38*(11), 14, doi:1241 10.1029/2001wr000978.

Sivapalan, M. (2009), The secret to 'doing better hydrological science': change the question!, *Hydrol. Process.*, *23*(9), 1391-1396, doi:10.1002/hyp.7242.

Sivapalan, M., K. Takeuchi, S. W. Franks, V. K. Gupta, H. Karambiri, V. Lakshmi, X. Liang, J. J. McDonnell, E. M. Mendiondo, P. E. O'Connell, T. Oki, J. W. Pomeroy, D. Schertzer, S. Uhlenbrook, and E. Zehe (2003), IAHS decade on Predictions in Ungauged Basins (PUB), 2003-2012: Shaping an exciting future for the hydrological sciences, *Hydrol. Sci. J.-J. Sci. Hydrol.*, *48*(6), 857-880.

Sklash, M. G., and R. N. Farvolden (1979), Role of groundwater in storm runoff, *J. Hydrol.*, *43*(1-4), 45-65.

Smakhtin, V. U. (2001), Low flow hydrology: a review, *J. Hydrol.*, *240*(3-4), 147-186.

Sneath, D. (1998), Ecology - State policy and pasture degradation in inner Asia, *Science*, *281*(5380), 1147-1148.

Soulsby, C., J. Petry, M. J. Brewer, S. M. Dunn, B. Ott, and I. A. Malcolm (2003), Identifying and assessing uncertainty in hydrological pathways: a novel approach to end member mixing in a Scottish agricultural catchment, *J. Hydrol.*, *274*(1-4), 109-128.

Soulsby, C., P. Rodgers, R. Smart, J. Dawson, and S. Dunn (2003), A tracer-based assessment of hydrological pathways at different spatial scales in a mesoscale Scottish catchment, *Hydrol. Process.*, *17*(4), 759-777, doi:10.1002/hyp.1163.

Stednick, J. D. (1996), Monitoring the effects of timber harvest on annual water yield, *J. Hydrol.*, *176*(1-4), 79-95.

Steffens, M., A. Kolbl, K. U. Totsche, and I. Kogel-Knabner (2008), Grazing effects on soil chemical and physical properties in a semiarid steppe of Inner Mongolia (PR China), *Geoderma*, *143*(1-2), 63-72, doi:10.1016/j.geoderma.2007.09.004.

Stromberg, J. C., M. K. Chew, P. L. Nagler, and E. P. Glenn (2009), Changing Perceptions of Change: The Role of Scientists in Tamarix and River Management, *Restor. Ecol.*, *17*(2), 177-186, doi:10.1111/j.1526-100X.2008.00514.x.

Suguwara, M. (1961), On the analysis of runoff structure about several Japanese rivers, *Japanese Journal of Geophysics*, *2*(4), 1-76.

Suguwara, M. (1995), Tank Model, in *Computer Models of Watershed Hydrology*, edited by V. Singh.

Swistock, B. R., D. R. Dewalle, and W. E. Sharpe (1989), Sources of acid storm flow in an appalachian headwater stream, *Water Resour. Res.*, *25*(10), 2139-2147.

Tetzlaff, D., and C. Soulsby (2008), Sources of baseflow in larger catchments - Using tracers to develop a holistic understanding of runoff generation, *J. Hydrol.*, *359*(3-4), 287-302, doi:10.1016/j.jhydrol.2008.07.008.

Tong, C., J. Wu, S. Yong, J. Yang, and W. Yong (2004), A landscape-scale assessment of steppe degradation in the Xilin River Basin, Inner Mongolia, China, *J. Arid. Environ.*, *59*(1), 133-149, doi:10.1016/j.jaridenv.2004.01.004.

Vaché, K. B., and J. J. McDonnell (2006), A process-based rejectionist framework for evaluating catchment runoff model structure, *Water Resour. Res.*, *42*(2), 15, doi:W02409 10.1029/2005wr004247.

Vaché, K. B., J. J. McDonnell, and J. Bolte (2004), On the use of multiple criteria for a posteriori model rejection: Soft data to characterize model performance, *Geophys. Res. Lett.*, *31*(21), 4, doi:L21504 10.1029/2004gl021577.

Vinton, M. A., and I. C. Burke (1997), Contingent effects of plant species on soils along a regional moisture gradient in the Great Plains, *Oecologia*, *110*(3), 393-402.

Wade, A. J., C. Neal, C. Soulsby, R. P. Smart, S. J. Langan, and M. S. Cresser (1999), Modelling streamwater quality under varying hydrological conditions at different spatial scales, *J. Hydrol.*, *217*(3-4), 266-283.

Wang, X. P., R. Berndtsson, X. R. Li, and E. S. Kang (2004), Water balance change for a re-vegetated xerophyte shrub area, *Hydrol. Sci. J.-J. Sci. Hydrol.*, *49*(2), 283-295.

Watson, C. J., and C. L. Mills (1998), Gross nitrogen transformations in grassland soils as affected by previous management intensity, *Soil Biol. Biochem.*, *30*(6), 743-753.

White, R., S. Murray, and M. Rohweder (2000), *Pilot analysis of global ecosystems - grassland ecosystems*, Washington, D.C. USA.

Wilcox, B. P. (2007), Does rangeland degradation have implications for global streamflow?, *Hydrol. Process.*, *21*(21), 2961-2964, doi:10.1002/hyp.6856.

Wilcox, B. P. (2010), Transformative ecosystem change and ecohydrology: ushering in a new era for watershed management, *Ecohydrology*, *3*(1), 126-130, doi:10.1002/eco.104.

Wilcox, B. P., and T. L. Thurow (2006), Emerging issues in rangeland ecohydrology: Vegetation change and the water cycle, *Rangel. Ecol. Manag.*, *59*(2), 220-224.

Yang, D., C. Li, B. Ye, H. Hu, H. Zhang, Z. Liu, and J. Xia (2005), Chinese Perspectives on PUB and the Working Group Initiative, *Predictions in ungauged basins: international perspectives on the state of the art and pathways forward*, *IAHS Publication 301*(301), 26-29.

Zhang, W. (1998), Changes in species diversity and canopy cover in steppe vegetation in Inner Mongolia under protection from grazing, *Biodivers. Conserv.*, *7*(10), 1365-1381.

Zhao, S., C. Peng, H. Jiang, D. Tian, X. Lei, and X. Zhou (2006), Land use change in Asia and the ecological consequences, *Ecol. Res.*, *21*(6), 890-896.

Zheng, Y. R., Z. X. Xie, C. Robert, L. H. Jiang, and H. Shimizu (2006), Did climate drive ecosystem change and induce desertification in Otindag sandy land, China, over the past 40 years?, *J. Arid. Environ.*, *64*(3), 523-541, doi:10.1016/jjaridenv.2005.06.007.

ACKNOWLEDGMENTS

When I first started applying for PhD positions, Gießen was definitely not on my mental map of potential places where I thought I would want to live. This was confirmed during my first visit to the city. I came to interview for one certain position, where at the end of the day, I was offered another position, the one that I took in the end, with field work in China. I had difficulties feeling enthusiastic about this offer. I had never been to Asia, not to mention China before. I was an “american” person, leave it be North, Central or South. For a whole decade, the destination of my long-distance trips was always that continent. For a high school exchange year (USA), voluntary civil service exchange (Brazil), a study abroad semester (USA), a study project (Brazil) and my master thesis (Panama); you name it. But China did not trigger any emotions. I felt absolutely halfhearted about that part of the world. So I took a long time to think about it. Fortunately, the project itself sounded interesting enough and the chemistry of the working group seemed just right. And I am so glad about Lutz`s patience to not give the position away (and for the maybe non-existence of other applicants). Now three years later, I know it was absolutely the right decision which opened up so many possibilities and broadened my horizon in so many ways. I can say, that every single day, I enjoyed going to the office thanks to the research I was working on and thanks to the incredible positive atmosphere at the ILR and especially within our working group. But not only the “desk work time” here at home in Germany made these three years so valuable. I absolutely do not want to miss my field trips to China which are, despite the many difficulties due to technical problems or cultural differences, unforgettable additions to my collection of life experiences. China really has grown dear to me and I already miss not frequently having to go there anymore.

This dissertation research would not be what it became without the help and the support of many people which I would now like to thank.

First of all, I would like to thank my advisors Hans-Georg Frede and Lutz Breuer: Hans-Georg for giving me the possibility to work where I worked and for creating this incredible positive working environment which is highly fruitful for dissertation research; and Lutz for his tremendous support, his enthusiasm and for his always open door.

I thank Kellie Vaché (Oregon State University) for reviewing my thesis, many helpful ideas along the path and his continuous help in paper preparation.

Thank you also to the third reviewer, Prof. Dr. Lorenz King (JLU Gießen) for reading and reviewing this thesis.

I deeply appreciate Jeff McDonnell's (Oregon State University) generous hospitality during my 3 months visit of his watershed hydrology lab at Oregon State University and thank him for fruitful discussions and his inspiration.

There are many people without whom the completion of this work would not have been possible. I especially have to thank Dorit Zörner and the technical staff at the Institute for Landscape Ecology and Resources Management, Heike Weller, Beate Lindenstruth and Nelly Weis, for all the lab analyses of the many water samples that I hauled back from the field. A huge thank you goes to Ruth Strittmatter and Melanie Kehl for putting up with the administrative difficulties which I created with all the traveling that was involved in this work. Friederike Gerschlauer and Johanna Schäfer are thanked for their helping hands in the field. Financial support was provided from the German Science Foundation (DFG) through the sino-german project MAGIM (Matter fluxes in grasslands of Inner Mongolia as induced by stocking rate) (FG 536). I thank the Institute of Botany, Chinese Academy of Sciences, and Christian Bernhofer, Bettina Ketzer and Wang Lei from TU Dresden for providing meteorological data. I deeply acknowledge the technical and logistical support from the Inner Mongolian Grassland Ecosystem Research Station (IMGERS) and the administrative support from Yuandi Zhu. I especially want to thank Yongsheng Jin, our driver, without whom many technical problems would have stayed unsolved.

I am endlessly glad about my PhD fellows in the MAGIM group Martin Wiesmeier, Nicole Fanselow, Katrin Müller, Philipp Schönbach, David Schaffrath, Max Wittmer, Benjamin Wolf, Stefan Metzger and Matthias Reiche. Among them it was possible to feel somewhat at home at a faraway place which is culturally worlds apart from ours.

My deep appreciation goes to Jinkui Wu, my Chinese PhD partner, for a very pleasant cooperation, his reliable availability and for always adding a huge amount of understanding of everything during my visits to China.

I would like to thank the past and current members of the resources management working group Amelie Bücker, Katrin Schneider, Philipp Kraft, Stefan Julich, Jean-François Exbrayat, Benjamin Blank and Christoph Tyralla. My discussion with this group went a long way towards my ideas that are outlined in this thesis.

Last but not least, I would like to thank my family and friends: my parents, Bernd and Jutta Barthold, who are my biggest fans and endlessly support and encourage me with everything that I do, Stefan and Natalie for having to document the tough times of this process, Carolin, Assi, Mone, Wiwi, und Inga for their friendship.

ERKLÄRUNG

Ich erkläre: Ich habe die vorgelegte Dissertation selbständig und ohne unerlaubte fremde Hilfe und nur mit den Hilfen angefertigt, die ich in der Dissertation angegeben habe. Alle Textstellen, die wörtlich oder sinngemäß aus veröffentlichten Schriften entnommen sind, und alle Angaben, die auf mündlichen Auskünften beruhen, sind als solche kenntlich gemacht. Bei den von mir durchgeführten und in der Dissertation erwähnten Untersuchungen habe ich die Grundsätze guter wissenschaftlicher Praxis, wie sie in der „Satzung der Justus-Liebig-Universität Gießen zur Sicherung guter wissenschaftlicher Praxis“ niedergelegt sind, eingehalten.

Frauke Barthold

Gießen, Mai 2010