Approaches to Hazard-oriented Groundwater Management Based on Multivariate Analysis of Groundwater Quality

Inauguraldissertation

zur
Erlangung der Würde eines Doktors der Philosophie
vorgelegt der
Philosophisch-Naturwissenschaftlichen Fakultät
der Universität Basel
von

Rebecca Mary Page
aus Gateshead, Tyne and Wear (UK) und Basel (CH)

Basel, 2011

Genehmigt von der Philosophisch-Naturwissenschaftlichen Fakultät der
Universität Basel auf Auftrag von

Prof. Dr. Peter Huggenberger
Angewandte und Umweltgeologie
Institut für Geologie und Paläontologie
Departement Umweltwissenschaften
Universität Basel
Schweiz

und

Prof. Dr. Gunnar Lischeid
Institut für Landschaftswasserhaushalt
Leibniz-Zentrum für Agrarlandschaftsforschung e.V.
Müncheberg
Deutschland

Basel, 21. Juni 2011

Prof. Dr. Martin Spiess
Dekan

1

Acknowledgements

Contents

List of Tables

List of Figures

1   Summary

Drinking water extracted near rivers in alluvial aquifers is subject to potential microbial contamination due to rapidly infiltrating river water during high discharge events. The heterogeneity of river-groundwater interaction and hydrogeological characteristics of the aquifer renders a complex pattern of groundwater quality. The quality of the extracted drinking water can be managed using decision support and HACCP (Hazard Analysis and Critical Control Point) systems, but the detection of potential contamination remains a complex task to master. The methodology proposed herein uses a combination of high-resolution measurements and multivariate statistical analyses to characterise actual groundwater quality and detect potential contamination. The aim of this project was to improve the protection of riverine groundwater extraction wells and to increase the degrees of freedom available to the management of fluvial planes with drinking-water production and aquifer recharge by river-groundwater interaction.

The monitoring network was set up in the *Reinacherheide* in North-west Switzerland and encompassed the depth-oriented installation of multiparameter instruments, a surface-water monitoring station and a flow-through cell with an automated sampler and high-precision measurement instruments. The parameters recorded included temperature, electrical conductivity, spectral absorption coefficient, particle density and turbidity. Two of the observation wells were equipped with a telemetry system and the flow cell could be controlled remotely. The well-field encompassed eight groundwater extraction wells.

The optimal choice of observation wells and indicator parameters was assessed using principal component analysis of groundwater head, temperature and electrical conductivity time-series to detect the influence of, for example, river-water infiltration or river-stage fluctuations on the time-series recorded in the groundwater observation wells. Groundwater head was susceptible to pressure waves induced by both river-stage fluctuations and groundwater extraction. Temperature time-series showed only weak responses to high discharge events. Electrical conductivity, however, showed a distance-driven response pattern to high discharge events. To further assess the representative strength of individual groundwater quality indicator parameters for identifying microbial contamination, a bi-weekly and a high-resolution sampling campaign were carried out. The results showed high faecal-indicator bacteria densities (*E. coli* and *Enterococcus* sp.) at the beginning of high discharge events, followed by a rapid decrease, leading to a strong *hit-and-miss* characteristic in the bi-weekly sampling campaign. The third approach

applied used the neural network-based combination of self-organizing maps and Sammon's projection (SOM-SM) to detect shifts in groundwater quality system states. The nonlinear analysis was carried out with groundwater head, temperature and electrical conductivity time-series from six observation wells. The subsequent shading of the projected trajectory of system states with independent time-series (spectral absorption coefficient and particle density) allowed the identification of critical system states, when actual groundwater quality decreased and contamination of the extraction wells was imminent. The time at which the changes in system state occurred and were detected were used as potential warning indicators for the water supplier. The effects of altered groundwater extraction (as a consequence of the SOM-SM warning) were then simulated using a groundwater flow model. The outcome of the SOM-SM analysis is, thus, proposed as an interface between the monitoring system and extraction-well management system.

The proposed approach incorporates hydrogeological knowledge and the analysis of prevalent conditions concerning river-groundwater interaction with real-time telemetric data transfer, data-base management and nonlinear statistical analysis to detect deterioration in actual groundwater quality due to rapidly infiltrating river water. As the SOM-SM is not based on threshold values and independent of indicator parameters, the approach can be transferred to other sites with similar characteristics.

2 Introduction

Drinking-water quality has become a frequent topic of discussion in both scientific media and the popular press (Furlow, 2005; Schwarzenbach et al., 2006; Hooper, 2009; UNESCO, 2009; Duhigg, 2009; BBC, 2011; BAZ, 2011). Two topics regularly featured in the discussion are: a) water quality and pollution, and b) water quantity and availability. Water quality can be impaired by a range of pollutants, including thermal, persistent and non-persistent contaminants from anthropogenic or geogenic sources. Anthropogenic contaminants encompass a range of microorganisms and chemical substances that can seriously harm consumer health. Drinking water treatment is necessary when contamination cannot be avoided. However, transient sources of pollution can be avoided if suitable management measures are taken and water quality is otherwise good. For example, river-groundwater interaction is strongly heterogenous and can give rise to microbial contamination of groundwater extraction wells near rivers. The selection of management measures and the time at which to deploy them are thus critical for securing high quality-standards for drinking water whilst maintaining the neccessary quantity to satisfy demand.

The increasing demands and rising standards for drinking water, together with the diffuse pressure on groundwater resources due to urban development and agricultural practises, makes the management of many drinking-water supply systems a complex interplay between economic, social and technical factors (Epting et al., 2008; Peters, 2009). The maintenance and management of drinking-water supply systems thus need concepts and methods that are able to integrate the different requirements, while maintaining the integrity and functioning of the ecosystem. Control and decision support systems, already in use in many areas, will have to deal with an increasingly diverse and challenging set of measurements, standards and goals concerning the availability and quality of drinking water. Decision support systems can be used to help find the optimal management measure under complex circumstances and when time is restricted and decisions have to be made quickly.

Besides the operational aspects of industrial and technical water supply systems, the heterogeneity of the groundwater resource, both in space and time, adds to the complex task of managing drinking-water supplies. Drinking-water extraction in many alluvial aquifers has traditionally been located close to rivers, which supply the aquifer with water. Rivers are, however, a source of contamination as they drain the catchment area, extracting and concentrating many potentially harmful substances. River-

ine contamination ranges from agricultural land run-off, sewage treatment plant overflow to industrial and municipal waste water, including many non-degrabable chemicals and a wide range of faecal pathogens. The management of drinking-water extraction from wells near rivers thus has to find a compromise between quantity and quality of extracted water. While water supply companies in some countries are legally required to treat all groundwater before distributing it, Switzerland and other countries have adopted a differentiated approach based on travel times and well-head protection areas and do not adhere to treating drinking water prior to distribution. The quality of water extracted from these wells is thus a function of the quality of the water recharging the aquifer and the filter efficiency of the aquifer material (Frind et al., 2006). The endangerment of drinking-water quality and, with it, human health does not often occur during typical hydrological situations, but frequently follows heavy precipitation and high discharge events (Wilkinson et al., 2006; Grisey et al., 2010). It is important to rapidly identify and assess such situations, so that the control systems of drinking-water supply plants can be managed accordingly. In environments where river-groundwater interaction plays a significant role in determining water quality, knowledge on process dynamics and on how to quantify potential contamination is essential but difficult to obtain due to the complex nature of river-groundwater interaction and groundwater flow (Epting et al., 2008). Spatial heterogeneity due to aquifer and river bed composition leads to patchy patterns of river-groundwater interaction and groundwater flow. As river-bed morphology and hydrogeological characteristics associated with the river-groundwater interface are themselves frequently non-stationary, the pattern is even more complex and resulting measurements of many parameters, such as groundwater head, are a reflection of interactions occuring at the interface between the river and the aquifer and large-scale, catchment-related processes (Longuevergne et al., 2007). Beyond the river-groundwater interface, the presence of paleochannels, silt lenses or high-conductivity layers add to the complexity of potential flowpaths that the infiltrated river water can follow (Stanford and Ward, 1993). Adding to the heterogeneity, the hyporheic zone is dominated by biogeochemical processes resulting in nonlinear behaviour of many measured parameters as they pass through the zone (Allen et al., 2010). Biogeochemical reactions can change over time, but are also spatially variable, creating hotspots for nutrient cycling (Dahm et al., 1998). The temporal heterogeneity of river-groundwater interaction has received considerable attention, culminating in the development of high-resolution measurement techniques. For example, fibre-optic cables have been used to map temperature, from mm resolution to river-

15

stretch scale measurement campaigns (Conant, 2004; Selker et al., 2006; Vogt et al., 2010). Temperature can be used to estimate water fluxes between the river and the hyporheic zone. However, these methods rely on a detectable temperature signal, which is progressively attenuated as the water travels through the aquifer. Measurements have to be made in the river and at different depths in the aquifer. Such univariate proxy analysis methods rely on the relationship between the proxy indicator (measured parameter) and the variable of interest (in this case, river-water infiltration) being known, replicable and constant over space and time. With the advancement of multiparametric measurement instruments, the availability of information accessible for analysis has significantly increased. The resulting large environmental data sets require additional analysis to extract information on the intrinsic behaviour of the system under observation (Lischeid, 2009).

The approach applied in this project makes use of multivariate measurement and analysis techniques to extract inherent information on system state from large environmental data sets, so as to be less reliant on individual parameters and their strength as indicators for river-water infiltration. By extracting and analysing patterns derived from multivariate data sets, the complexity and dimensionality of data is reduced so that relevant processes and changes in groundwater quality can be more readily recognised. The analysis of a multiparametric data set using statistical methods thus gives insight into the factors driving the system and forms the basis of understanding required to develop decision support systems for drinking water suppliers with wells susceptible to hazards associated with river-water infiltration.

2.1 Aims and Objectives

The aim of this project was to improve the protection of riverine groundwater extraction wells and to increase the degrees of freedom available for the management of fluvial planes with multiple stakeholders, including nature reserves, energy and drinking water production and urban development. Conflicts of use will remain between the different parties claiming use of the area, however, the results from this study will allow a differentiated analysis of the actual endangerment to drinking-water quality and provide a basis for discussion.

The overall objective of this project was to develop an approach to identifying and assessing actual endangerment of drinking-water extraction wells based on real-time measurements as proxy indicators for the hazard of river-borne microbial contamination.

2.2 Approach

The project environment was defined by boundary conditions given by legal, scientific, physical (measurements) and political or social constraints. The boundaries and individual work packages of the project are presented in Fig. 1. The legal, political and social aspects are essential to the success of such a project, they are however, not dealt with in detail in this thesis, where the focus lies on a practical solution that can be implemented in the management system of a small to medium sized drinking-water supplier susceptible to contamination due to river-water infiltration. The approach adopted in this project combined (a) hydrogeological process knowledge, (b) high-performance (resolution and precision) measurement techniques, and (c) multivariate statistical analyses.

(a) Hydrogeological process knowledge was largely based on previous studies, including regional groundwater-flow modelling of the Birs Valley and catchment area studies carried out to assess nitrate levels linked to agricultural activities (Huggenberger et al., 2006). With many of the dominating processes explored at the regional scale, the study area (Chapter 3) provided an ideal setting to investigate local-scale processes, such as river-groundwater interaction, in greater depth, both in time and space. In addition to groundwater-flow modelling, geophysical investigation methods have been used to study the aquifer structure. Georadar and geoelectrical experiments were carried out to study the structure of the aquifer and to identify potential infiltration hot-spots along the river bank, (Glur, 2008; Wüest, 2010). Further hydrogeological information was gained from a high-resolution map of the aquitard, a layer of *Elsässer Molasse*, modelled by H. Dresmann (Applied and Environmental Geology, University of Basel). The surface of the aquitard showed the shape of the basin, as formed by previous river-bed positions (Schmassmann, 1981).

(b) High-performance measurement techniques included a depth-oriented instrument installation in groundwater observation-well clusters. The existing measurement network was extended to include six further observation wells and an automated flow-through system recording parameters, such as particle density, otherwise not easily measurable in groundwater observation wells. The industrial partner for this project, Endress+Hauser Metso AG, used this study site to test new instrument developments. The robustness of the measurement instruments in the river proved to be a significant challenge in the face of mechanical damage due to high-discharge

events. The combination of distributed measurements, in two dimensions, and high temporal resolution (max. 1 measurement per minute) provided an extensive data set. A database was constructed to manage the time-series, which were variable in terms of temporal resolution and format. In addition to the continuous groundwater time-series, short-term microbial monitoring experiments were carried out. As part of her MSc thesis, Polat (2010) investigated bacterial densities in the river and groundwater based on two monitoring strategies: 1) bi-weekly monitoring for a longer period of time, and 2) event-based sampling during a high discharge event. The results are presented in Chapter 7.

(c) Multiparametric statistical analysis of the monitoring data set was based on two methods: (a) principal component analysis (PCA), and (b) a combination of self-organizing and Sammon's mapping techniques (SOM-SM). The dual approach included a linear method to identify observation wells and parameters that can be used to monitor river-water infiltration (PCA), and a nonlinear methodology to identify critical system states, i.e. circumstances when river-water infiltration has been detected in the aquifer and poses a hazard to drinking water quality if extraction is continued (SOM-SM). The second method aims to provide a model-free solution to data mining, by extracting 'hidden' information from a wealth of data, which can be translated into an integrated part of a drinking-water supply control-system. The identification of critical system states, was further checked using a groundwater flow model, the results of which are presented in Chapter 9.

These three areas build the basis for a process-based management tool for groundwater extraction wells susceptible to microbial contamination due to river-water infiltration. The following chapters describe the work carried out on each of these three aspects. **Chapter 3** is dedicated to the measurement network, instrumentation and data transfer. **Chapter 4** provides information on decision support systems, as they are often, or increasingly the mechanism by which natural-resource management occurs and include important aspects of food safety, such as HACCP (Hazard Analysis and Critical Control Point) approaches. Chapter 4 also includes time-series, proxy and indicator analysis. **Chapter 5** is based on a chapter in the book *Urban Geology* (Page and Simovic, 2011) and aims to provide some information on the motivation for multiparametric statistical analysis. **Chapter 6** comprises a manuscript submitted to the Journal of Hydrology (29.03.2011) and describes the application of multivariate statistical principal component
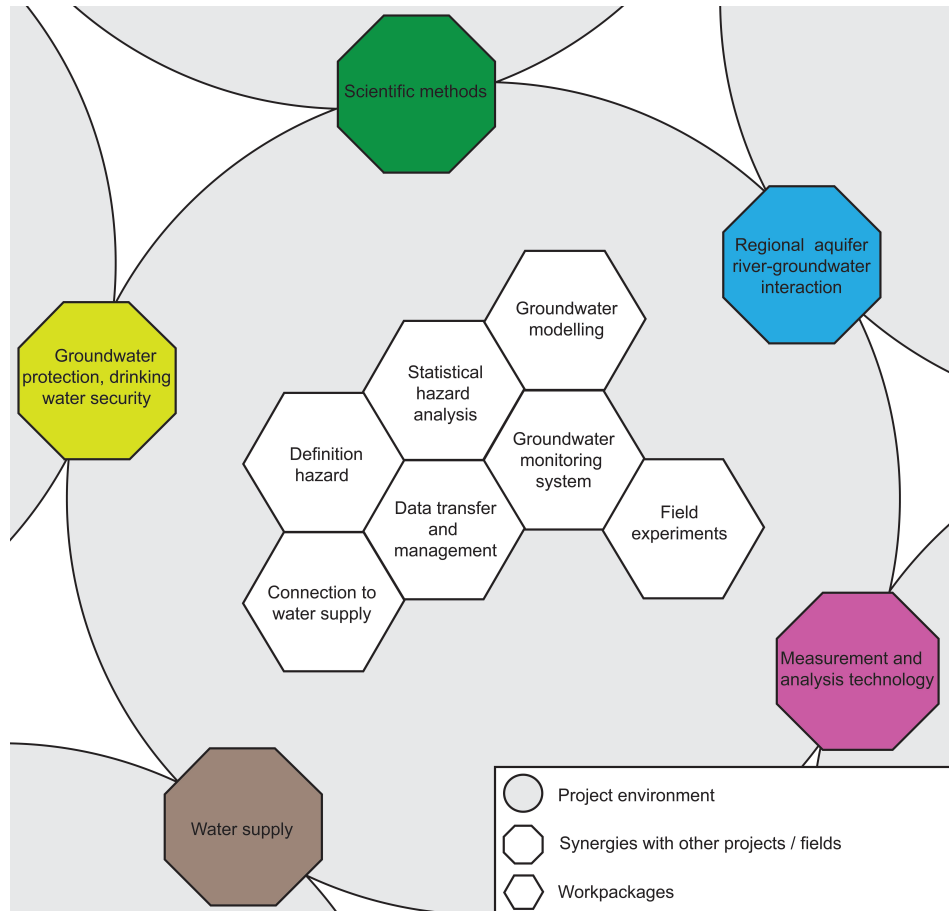
Figure 1: Schematic representation of the project environment, including connections to other projects and fields and the individual work packages involved in this project.

analysis to identifying indicator observation wells and parameters that can be used to detect river-water infiltration from a large data set. A further application of principal component analysis is presented in the appendix (Chapter 14, Page et al. (2011)). **Chapter 7** comprises a manuscript submitted to *Water Resources Management* (03.06.2011) and deals with monitoring and sampling strategies for faecal indicator bacteria. **Chapter 8** is in preparation to be submitted and describes the use of self-organizing maps, a nonlinear multiparametric neural-network based method, to identify critical situations for drinking water suppliers. **Chapter 9** shows the groundwater-flow model setup and scenario analyses, used in conjunction with the results from Chapter 8. **Chapter 10** describes the combined results from this study, and is an application-oriented section. **Chapter 11** describes the conclusions that can be drawn from this project and further steps and areas of research that require attention.

Chapters 3, 4 and 9 cover parts of the project not included, or only briefly mentioned in the peer-reviewed publications, but are essential to the development of the hazard-based groundwater-water-qualilty assessement method.

3 Experimental Field Site and Measurement Setup

The study area is located in the lower Birs Valley of North-west Switzerland (Fig.2). It is an alluvial system with a shallow, coarse, gravelly aquifer. The River Birs was canalized at the end of the 19th century and since then, the river bed has incised several metres into the former floodplain, from which it was subsequently disconnected. The Birs runs 75 km through the Swiss Jura and joins the River Rhine in Basel, creating a catchment area of 866 km^2. The mean annual flow near the confluence zone of the tributary Birs with the Rhine is 15.4 m^3 s^{-1} and storm flows can reach up to 383 m^3 s^{-1}.
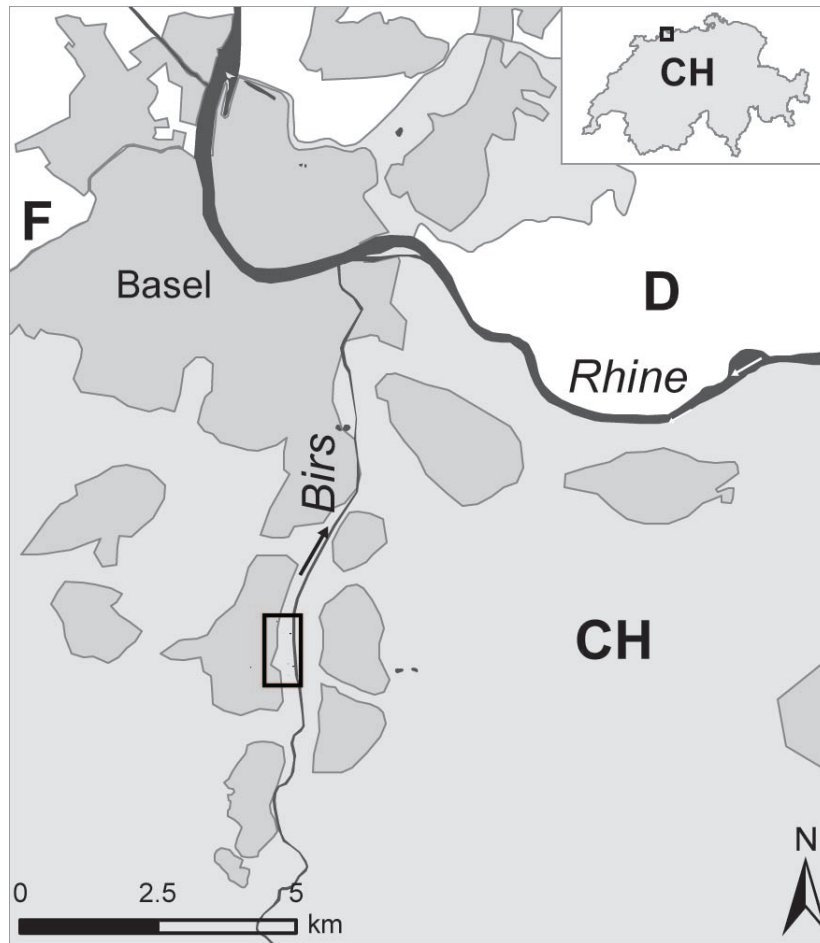
Figure 2: Regional setting of the experimental field site. The black box shows the location of the experimental field site.

3.1 Aquifer Heterogeneity

The aquifer material mainly consists of carbonate gravel, which is often well-rounded, with variable sediment sorting, and few intercalations of clay or silt lenses, resulting in a large variance in hydraulic properties. The lenses can be extensive overbank deposits, mostly consisting of fine fraction sediments (<0.063 mm) with variable gravel and limestone content. The carbonate gravel components of the aquifer are of Triassic to Jurassic origin with hydraulic conductivities between 3×10^{-3} and 10^{-2} ms^{-1}. The aquifer bottom represents an aquitard and consists of Tertiary deposits (*Elsässer Molasse*), generally with very low hydraulic conductivities (10^{-8} to 10^{-5} ms^{-1}). The maximal aquifer thickness is 35 m and the thickness of the saturated zone varies between 0.6 m and 10 m (Fig. 3).
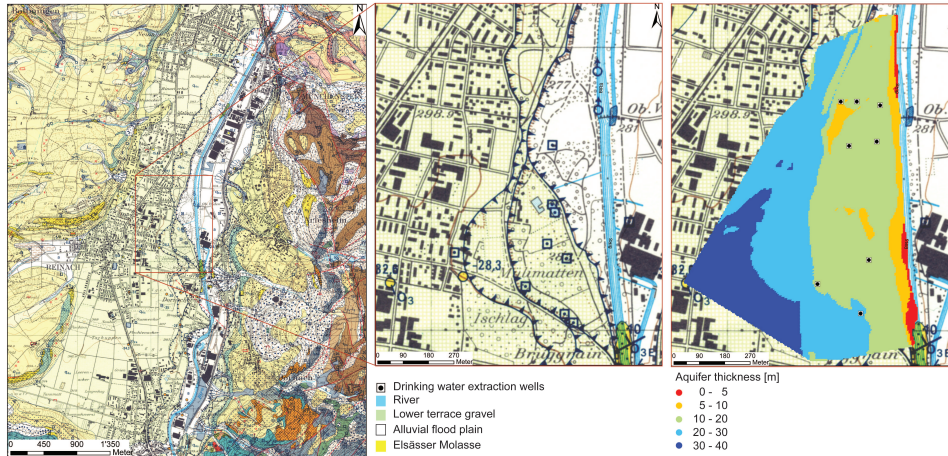
Figure 3: Geological setting of the study area showing terrace structures and aquifer thickness (background map based on Bitterli-Brunner et al. (1984)).

Most of the water is supplied to the aquifer from river-water infiltration, lateral inflow from the local catchment area, and intermittent artificial recharge. Eight active production wells in the study area (Fig. 4) supply drinking water to six communities, approximately 51'000 people, amounting to 5×10^6 m^3 drinking water per year.

3.2 Time-series

The experimental field site for river-groundwater interaction was set up in the gravel aquifer in the vicinity of the river bank (Fig.4). The installation of

the regional groundwater observation network occurred stepwise throughout previous studies focussing on land-use management, well-head protection areas and river-groundwater interaction (Huggenberger et al., 2006; Affolter et al., 2010). The existing observation-well network was expanded by six additional observation wells. Together with three existing wells, the additional wells formed three clusters of three observation wells (Fig. 4, panels B.1 - B.3). The filter stretch of each observation well (1 m) was positioned in a different depth in the aquifer (Fig. 4, panel C).
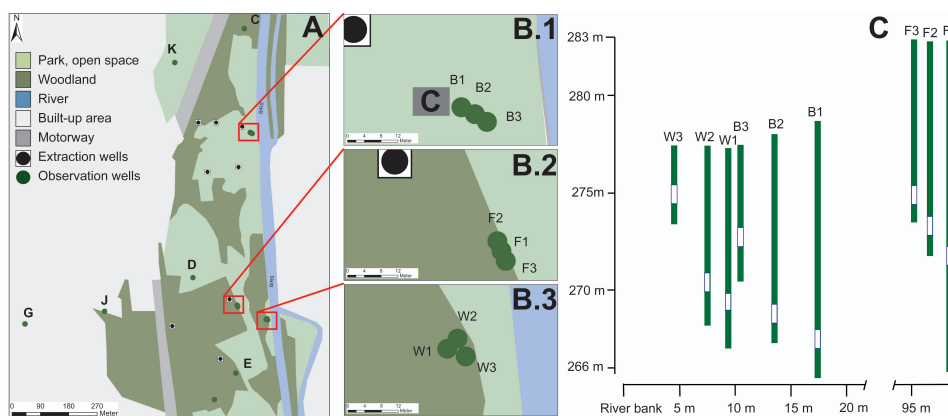
Figure 4: Measurement setup in the *Reinacherheide*, Switzerland. Panel A shows the location of the three observation-well clusters with respect to the river (blue) and the extraction wells (black circle and white square). Panels B.1 to B.3 show the location of the observation wells in each cluster and B.1 (*Heidebruggli* cluster) also indicates the position of the flow-cell (grey box with 'C' for *container*). The top (= ground surface) of the observation-well boreholes and the filter stretches (hollow sections) are shown in relation to distance from the river (panel C).

The three observation well clusters all recorded pressure (for groundwater head), temperature, electrical conductivity and turbidity. Two of the clusters (F and W) were equipped with multiparameter instruments (YSI 600 OMS V2). Individual measurement instruments for pressure, electrical conductivity (including temperature) and turbidity were installed in the third cluster (B, Endress+Hauser high-resolution equipment). The two YSI-clusters were battery-driven, while the Endress+Hauser cluster was connected to a mains power supply in the *container* (see below).

One groundwater extraction well in the study area was particularly susceptible to microbial contamination during high discharge events (panel B.1,

Fig. 4, cf. Chapter 9). To be able to closely monitor the critical system states, one cluster of observation wells was additionally equipped with an automated flow-cell (the *container*, Fig. 5 and 8), permitting the acquisition of parameter information not usually recordable in piezometers. In addition to the measurement instruments installed in observation wells B1, B2 and B3, groundwater was automatically extracted for 13.5 minutes every 45 minutes so that each observation well was regularly sampled. Between sampling each observation well, the flow-cell was flushed before the values were was recorded. The schematic setup of the *Heidebruggli* (B-cluster) installation is shown in Fig. 5. Alongside temperature, electrical conductivity and turbidity, pH, particle count (2-10 μm) and SAC (spectral absorption coefficient at 254 nm) were recorded in the *container*. A cooled (to 4 °C) autosampler enabled high-resolution microbial sampling (Endress+Hauser ASP 2000).

The river was also sampled in the proximity of the *Heidebruggli* cluster (B1, B2 and B3). In the orginal setup, the river water was included in the B-cluster flow-cell cycle, however, the riverine levels of SAC, turbidity and particle density were significantly higher than in the groundwater. As a consequence, the measurements fluctuated strongly and flow through some of the pipes was reduced due to deposition of riverine material. A bypass was constructed for the river water, so that microbial sampling could include river water, without distrupting the other measurements (Fig. 5, panel B). Pressure (for river stage), temperature, electrical conductivity, turbidity and pH were measured in the river water. The heterogenous environmental conditions, especially hydrological, had a strong impact on the river measurement installation. Although the installation site in the river was partially protected by a boulder, mechanical stress due to debris carried by the river (for example, trees) and the flow, severely damaged a number of instruments.

A further six observation wells had been installed and equipped during previous studies. The parameter data sets available from these observation wells were limited to groundwater head, some also had temperature, and one also had electrical conductivity measurements. A summary of the available parameter time-series and resolution is given in Tab. 1. The locations of the observation wells are shown in Fig. 6.
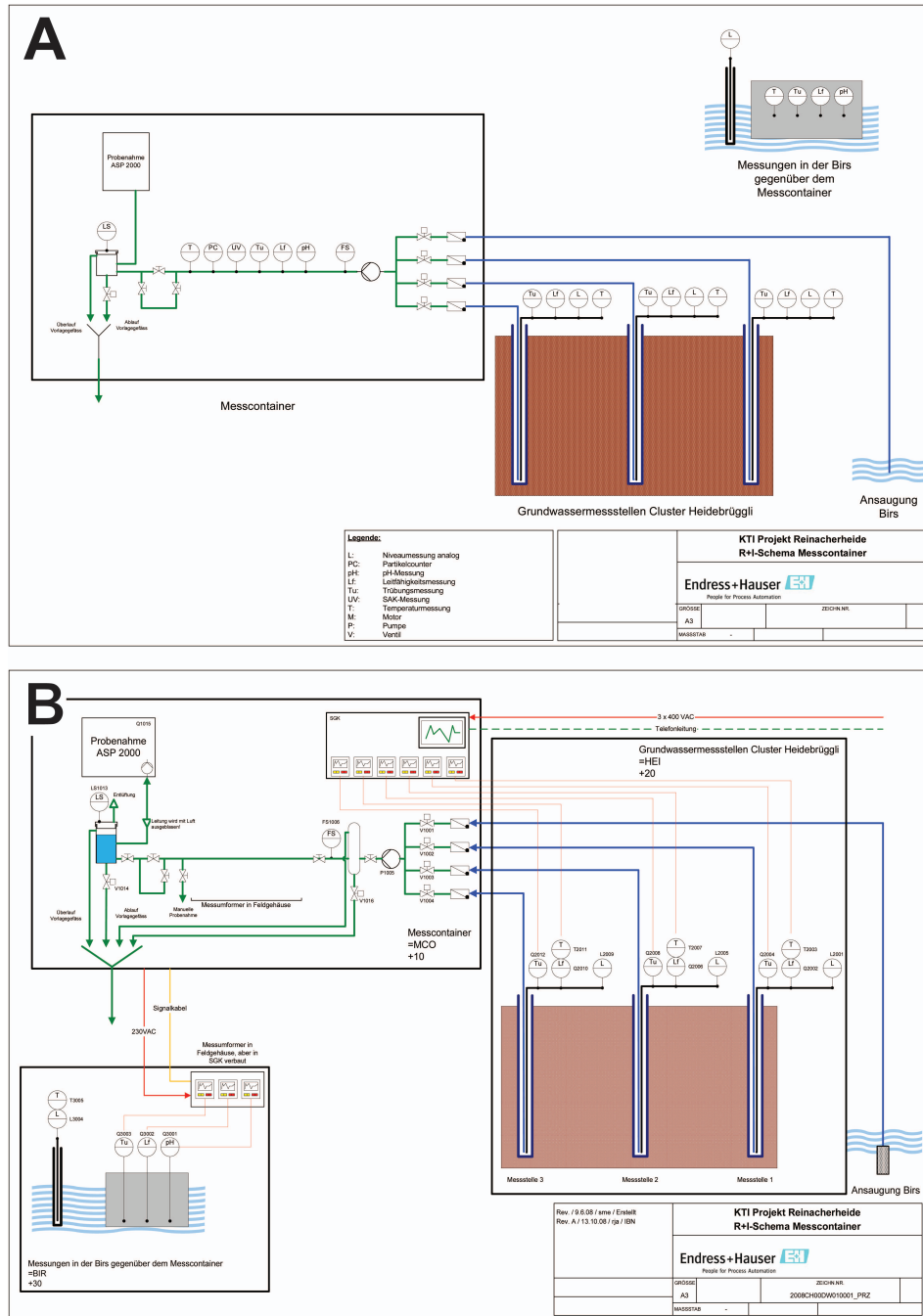
Figure 5: Measurement setup at the *Heidebruggli* (B-cluster) site (D. Waldmann, Endress+Hauser). Panel A shows the original setup, and panel B the optimized setup with a bypass system for river water.

Table 1: Borehole and observation well names (cf. Fig. 6) with corresponding parameters measured and the resolution of the measurements. The time-series recorded in the *container* are labelled with the prefix *Co.* H: groundwater head/river stage; T: temperature; EC: electrical conductivity; TR: turbidity; SAC: spectral absorption coefficient at 254 nm

Borehole	Obs. well	Parameter	Resolution [h^{-1}]
BL24J1	G	H, T	1
BL24J6	J	H, T	1
BL24J14	K	H	1
BL24J20	W1	H, T, EC, TR	1/2
BL24J21	B1	H, T, EC, TR	1/60
BL24J22	F1	H, T, EC, TR	1/2
BL24J23	F2	H, T, EC, TR	1/2
BL24J24	F3	H, T, EC, TR	1/2
BL24J25	W2	H, T, EC, TR	1/2
BL24J26	W3	H, T, EC, TR	1/2
BL24J27	B2	H, T, EC, TR	1/60
BL24J28	B3	H, T, EC, TR	1/60
BL24A4	E	H, T	1
BL24A7	D	H, T	1
BL24C12	C	H, T, EC	1
-	River	H, T, EC, TR, pH	1/60
-	Co-B1	T, EC, TR, pH, SAC, Particle count	3/4
-	Co-B2	T, EC, TR, pH, SAC, Particle count	3/4
-	Co-B3	T, EC, TR, pH, SAC, Particle count	3/4
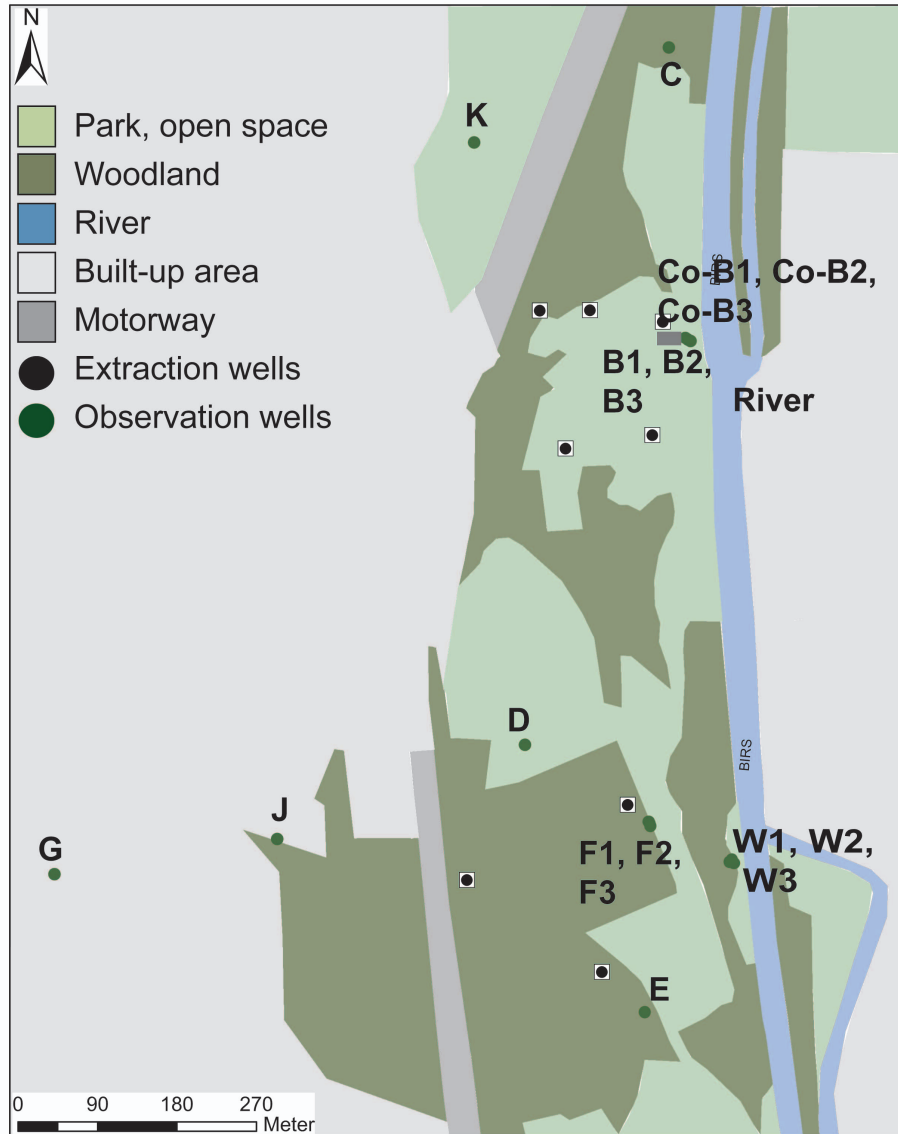
Figure 6: Location of the observation wells listed in Tab. 1

3.3   Remote Access and Telemetry

Remote access and data transfer are key points in decision support systems. The information required for decision support analysis is often collected in a decentralized system and may incorporate different technologies and data formats. Telemetry summarises the technology involved in remote measurement, control and information access. In this project three telemetry systems were used: 1) *Container* with access to mains power, but no landline connection. The main task for this system was to remotely manage a control and alarm system; 2) Stand-alone multiparameter instruments, where the telemetry system was used to transfer measurement data; 3) Water supply with access to mains power and a landline internet connection, however a series of safety issues and access rights had to be defined before any operational information on groundwater extraction and storage could be exchanged. Elements of systems 1) and 2) are shown in Fig. 8.

1) Container A GSM connection allowed remote access to the control system of the *container*. Besides basic operations, such as checking the status of the control system and parameter measurements, setting threshold values for alarms and initiating the autosampler could thus be managed remotely. Two categories of alarms were issued: 1) breaching of threshold values (for example, max. turbidity in the river), and 2) failure of the hydraulic pump, flow monitor, or overflow reservoir (sonic monitor). Category 1 only dispatched an alarm-text message if instructed to do so in the settings (non-obligatory) and the alarm could be cancelled via remote access. Category 2 alarms dispatched three text messages and a telephone alert. Category 2 alarms could not be remotely alleviated and required acknowledgement (blank text message) and on-site maintenance.

2) Stand-alone multiparameter instruments Two observation wells (W1 and W3) were equipped with a telemetry system. The system was set up in collaboration with Endress+Hauser and YSI Hydrodata Ltd and encompassed a datalogger (Campbell Scientific), antenna and a 12 V battery. The datalogger was configured to store and transmit data from two YSI multiparameter instruments. The data could be collected via a GSM connection and Loggernet Software (Campbell Scientific).

3) Water supply The information obtained from the water suppliers encompassed operational data, such as extraction volumes per well and storage volumes in reservoirs. This information is used to assess the current situation of the extraction wells and therefore the threat due to river-water infiltration. The data transfer from the water supply to the University network is based on a secure ftp account (University of Basel) and set up in

collaboration with Chestonag AG, the company managing the water supply's control system. The data files (1 data set per minute) are transferred from the operational system to the sftp server every 30 minutes (pathway 1, Fig. 7). After the analysis has taken place and a potential hazard detected, the water suppliers will be issued with a warning (pathway 2, Fig. 7). As the network of the water suppliers is a closed system, a precautionary step had to be introduced to avoid system failure due to false information. The warning is therefore issued to a member of the water supply staff, who then manually operates the supply system (pathway 3, Fig. 7).
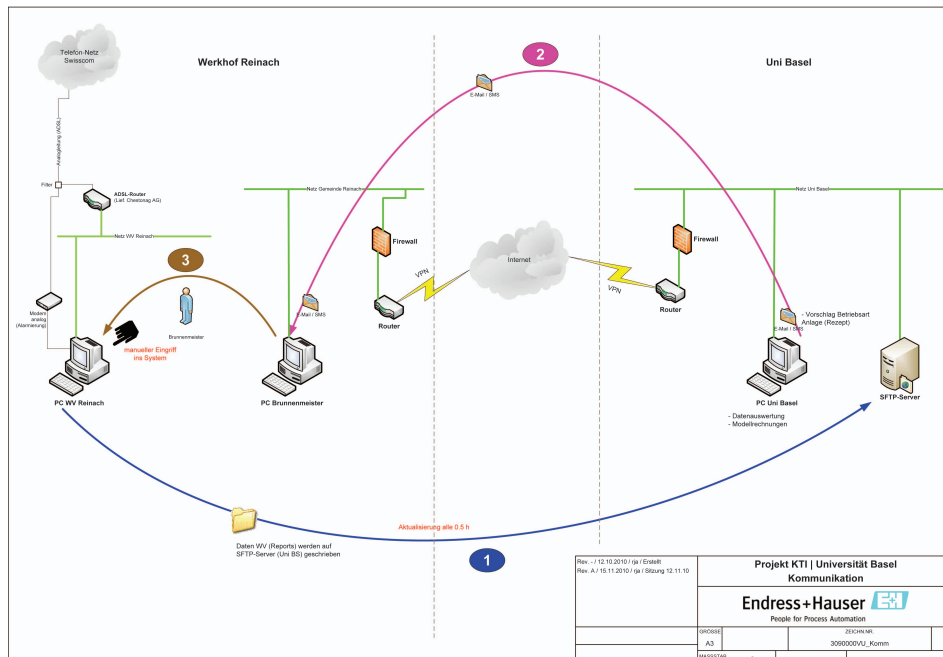
Figure 7: The setup for data file transfer between the water supply (*Werkhof Reinach*) and the University of Basel (R. Jaggi, Endress+Hauser). Three pathways are defined: 1) File transfer via sftp server and internal *Uni Basel* network. 2) Alarm or system state information via text message and email. 3) Manual operation by drinking-water supply staff.

Figure 8: Two telemetry stations. Panels A and B show the stand-alone installation for data transmission from observation wells W1 and W3. Panels C and D show the *container* and the touch-screen of the control system (same as remote access view).

4 Hazard Analysis and Critical Control Points, Decision Support Systems and Proxy Indicators

Environmental-health increases with per capita gross domestic product (GDP, measure of development, Esty et al. (2006)), accordingly, the availability of good quality drinking water increases as countries continue to develop. With rising pressure on water resources, the need for adaptive and integrated groundwater management becomes ever more important, as multiple stakeholder interests need to be incorporated into strategies to ensure the sustainability of the resource. The sustainable use of groundwater resources cannot be guaranteed without considering catchment-scale processes, such as nutrient fluxes and water mass-balances over different boundaries. Progressive urbanisation further increases the pressure on groundwater resources, including both quantitative and qualitative aspects. The quantity demand is linked to the consumers, for example, private households, industrial or agricultural users. Qualitative pressures on groundwater include traffic by-products and residual contamination, for example microorganisms from sewage outlets or persistent and non-persistent anthropogenic substances. Current practise in groundwater protection is often based on well-head protection areas (WHPA). The WHPA are derived from 1D or 2D advection calculations and seldom account for dispersion or dilution effects in the aquifer or following extraction. More sophisticated approaches include attenuation and dispersion effects alongside water treatment facilities (*multiple-barrier concept*, Frind et al. (2006)). This approach, based on well vulnerability, can be used to assess the optimal location of extraction wells with a known source of contamination. Time-variant sources of contamination, such as rapid river-water infiltration during high-discharge events, however may lead to an under- or overestimation of the potential hazard at a given well. One approach to adaptive monitoring and management of drinking-water extraction wells was suggested by Epting et al. (2008). The approach can be implemented with scenario development methods to augment current protection measures and improve the assessment of impacts on groundwater and the availability of safe drinking water. The approach is based on the development of simple, but robust instruments that enable the appraisal of impacts due to cumulative stresses on groundwater quality and flow regimes. The cumulative stresses are however often difficult to recognize, assess and manage. Even with an adaptive approach, the water-resource manager requires an assessment of actual quality and knowlegde of system dynamics upon which the management action/decision can be based. This chapter deals with three

concepts used in monitoring actual water quality and the decision-finding process: *Hazard Analysis and Critical Control Points (HACCP), Decision Support Systems (DSS)* and *Proxies and Indicators.*

Hazard Analysis and Critical Control Points (HACCP) Drinking water is a commercial product with certain obligations. It is considered a comestible/food and Swiss law has defined quality standards for untreated, treated and water in the distribution network that have to be met at all times. Drinking-water suppliers underlie Water Safety Plans (WSP) as determined by the legally-bound self-control (Art. 49 of the LGV, Swiss Federal Legislation (2005)), setting the boundaries for *Good Manufacturing Practise* (GMP). WSPs can include some or all steps in drinking-water production and provide a framework for managing water supplies (Hamilton et al., 2006). As part of GMP, quality assurance systems often involve a HACCP approach. As in many other branches of food production, a HACCP approach is applied to safeguard drinking-water quality throughout production (Havelaar, 1994; Hamilton et al., 2006). The analysis identifies weak points in production processes, which require monitoring or alleviating through technological solutions. Critical Control Points (CCPs) describe the points in a manufacturing chain, where action can be taken to protect the quality of the goods (in this case drinking water) and avoid hazards. HACCP and CCPs are further discussed in Chapter 8. CCPs can be applied to all aspects of drinking-water production and distribution, however, in this thesis, they will only be discussed in terms of identifying potential contamination due to infiltrating river water, i.e. prior to groundwater extraction. Further CCPs for drinking-water supply systems can be found in Havelaar (1994).

Decision Support Systems (DSS) are designed to assist in the solution-finding process when confronted with complex management problems (Lautenbach et al., 2009). There are different ways to approach this: (a) design a set of rules based on prior studies (simulations or including vulnerability areas); (b) include a simulation model in the DSS; (c) use real-time measurements. Approach (a) often lacks the ability to include a range of possible situations without losing detail. Approach (b) can include the outcome of management decisions and is often used for large areas and offline analysis. They may include considerations such as customer *willingness-to-pay* or long-term management options (Rosenberg et al., 2007). This category includes simulations based on flow and transport codes as well as artificial neural networks (ANN), as described by Rogers and Dowla (1994). Bauser et al. (2010) describe an application based on real-time data analysis and groundwater flow simulations. The approach adopted in this project be-

longs to category (c). The case at hand is simple in terms of management actions, as the pumps only have limited ranges: mostly *ON/OFF* and the extraction in two wells can be manually throttled. While the management options remain simple, the complexity of the studied problem arises due to the dominance of river-groundwater interaction in the study area, which is in itself heterogenous in time and space, groundwater extraction and seasonal fluctuations of the measured parameters.

Proxies and Indicators Many of the hazards to drinking-water quality are not directly measurable, for example, river-water infiltration, or require complex and lengthy analytical procedures, for example, waterborne microbial pathogens. Thus, the need for suitable indicators and proxies remains, as these can provide the basis for a qualitative and sometimes quantitative estimation of the parameter of interest, such as river-water infiltration or waterborne pathogens. The concept of indicators has been widely applied, not least in connection with climate change. For example, Cassin's auklet distribution and population dynamics is used as an indicator of ocean climate change in the California Current Ecosystem (Wolf et al., 2010). The understanding of river-groundwater interaction has also profited from the use of proxies, such as temperature-signal analyses (Selker et al., 2006; Keery et al., 2007; Vogt et al., 2010). The use of indicators is also widely applied to assess waterborne pathogens. Indicator organisms, such as the faecal indicator bacteria *Escherichia coli* and *Enterococcus* sp., are used to assess the presence of other pathogenic bacteria, also associated with faecal contamination, but require substantially more complex analytical procedures (Payment et al., 2003; Grisey et al., 2010). Heterotrophic plate counts (total aerobic bacteria) were also studied in this project, as they are the third kind of microbial indicator used in monitoring water quality in Switzerland. They are mostly used as indicators of water quality in distribution networks (Figueras and Borrego, 2010), but their application to monitoring groundwater quality is limited, as they do not necessarily relate to faecal contamination (Payment et al., 2003). As even the analysis for the faecal indicator bacteria requires time, other indicators or proxies that provide real-time information are required to assess actual groundwater quality (Brookes et al., 2005; Pronk et al., 2007). Hrudey et al. (2006) provide a list of limitations to monitoring drinking-water quality hazards, emphasizing problems in real-time monitoring of waterborne pathogens. Real-time identification and quantification of waterborne pathogens is currently in development. One of the methods used to assess the microbial content of water samples is based on flowcytometry and staining bacteria (Gruden et al., 2004). However, these systems are not readily available to many water suppliers as they are cost-intensive and

require trained staff and controlled conditions to operate in. The definition of proxies and indicators is variable. In the following chapters *indicators* (or *proxy indicators*) will be direct measurements of parameters, while *proxy* describes the outcome of an analysis of the indicator parameters. The result is a set of indicators, or proxies, that can be used to assess faecal indicator bacteria, which in turn represent other organisms associated with faecal contamination.

5   Multivariate Statistical Analysis

A significant part of our environment, especially in urban areas, is subject to compliance with legal limits and quality standards. The compliance is measured by a substantial set of networks to monitor air, surface and groundwater quality. Mostly, guidelines are set for ideal and critical threshold values. The measurements are required to remain within the boundaries set by the legislation. If the thresholds are breached, management actions are designed to return the system to the ideal state, for example, groundwater extraction is stopped to protect the well. However, it may often be too late, as demonstrated by the collapse of the Northwest Atlantic cod fishery. As the cod stocks were severely depleted, the ban placed on the fishery was not sufficient for the stocks to recover. The demise in the system was recognized too late for the management actions to prevent a regime shift (Savenkoff et al., 2007). Besides the selection of an appropriate management action, the time at which it is employed is crucial. Although not completely deterministic, environmental time-series are not independent of predecessor or successor measurements. The development of a measured parameter over time is also strongly dependent on the dynamics of the system. Small, yet steady changes can easily go undetected. They can however, tell us a lot about system dynamics and indicate future development. The time-scales at which these system-dynamics driven changes occur depend on the system under observation and the processes giving rise to the change.

Parameter monitoring and data collection are an essential part of understanding processes and ecosystem functioning. As we collect increasing amounts of data, we need an efficient mechanism to recognize the situations that are of interest to us. These include spatial components, such observation points that no longer function correctly, and temporal variation that indicates aberration from normal, or desired situations, for example, during industrial processes. Seldom is only one parameter, one measurement or one location monitored, the result is a large and complex data set. Although

there has been extensive environmental data collection in many areas, it is still under-exploited when it comes to the management of natural resources. To obtain sufficient information for decision making and the definition of management actions, further analysis is required to extract the essential information from the extensive data sets. The information content of the data set can however provide important guidelines for decision makers by using further analysis and data visualization methods (Fuertes et al., 2010).

Although questions or problems involving large, multiparametric data sets may differ in spatial and temporal scales, the answers require information on the development of the measured parameter over time in connection with observation points. In all cases, the preliminary analysis includes the assessment of variance in the data set. In this context, the sources of variance are sought and the reaction of different observation points compared. Further analysis looks at the development over time, manifesting the pattern arising due to the variance observed. This aids the identification of situations, where, for example, river-groundwater interaction poses a threat to drinking-water extraction, so, where a system moves away from a normal to a hazardous state: a critical system state.

Two methods, principal component analysis and a combination of self-organizing and Sammon's mapping, were applied to analyze large environmental data sets to obtain the necessary information for managing natural resources. These methods are readily applied to continuous time-series and offer considerable potential for the recognition of changes in hydrogeological processes or the evaluation of hazards in environmental systems.

5.1   Principal Components of a Multivariate Data Set

A principal component analysis (PCA), also called a Karhunen-Loève transform, is a non-parametric method of extracting information from complex data sets (Lischeid and Bittersohl, 2008). It can be used to reduce the dimension of a data set to reveal the sometimes hidden, inherent patterns. The method is a linear transform based on correlation coefficients of the data matrix. The principal components (PCs) are uncorrelated (are orthogonal to each other) and represent the joint variance observed in the data set (Gerbrands, 1981). A PCA aims at representing a maximum fraction of variance by a small number of components (Lischeid, 2009). In simple systems, one to two components may be sufficient to summarize the major sources of variance. The more complex a system is, the more components are required to explain the observed patterns in the data. There are different criteria for the optimal number of components to interpret (Kaiser, 1960; Cliff, 1988).

The loading of an input vector (original data) with a component provides an estimation of the importance of the source of variation (given by the component) at each observation point or for each measurement (input vector). A PCA can thus be carried out on continuous time-series or discrete samples. Helena et al. (2000) studied the influence of temporal variations, for example, caused by precipitation or agricultural activity, on groundwater quality, by applying a PCA to a large number of discrete samples. Lewandowski et al. (2009), on the other hand, carried out a PCA on continuous groundwater head measurements to study the relationship between groundwater head and river stage over time. The examples provided in Chapters 6 and 14 include the analysis of groundwater temperatures in an area influenced by river-groundwater interaction and artificial recharge, whereas Page and Simovic (2011) looked into the groundwater temperature distribution in an urban environment. The degree of influence of the two major factors on groundwater temperature is given by individual observation wells, which can be characterized by their sources of influence in terms of susceptibility to surface-water fluctuations or longer term, seasonal variation.

Many data sets in the field of urban hydrogeology are heterogeneous. They often derive from specific monitoring networks associated with construction or maintenance of urban infrastructure and are therefore temporally or spatially patchy and closely related to the problem at hand, for example, temperature plumes around buildings or electrical conductivity measurements close to construction sites. While a PCA can provide a means of comparing multiple parameters simultaneously, it is applied to a finite data set and can be considered a stationary method. Each analysis is dependent on the input data set and small changes will potentially result in a different outcome.

5.2   Pattern Analysis by Self-organization

The automatic detection of events, such as hazardous states, thus requires a time-sensitive approach. PCA provides a means for identification and indication of the distribution of influence of individual processes, however the questions related to event detection require a methodology that can capture the development of the whole system. Artificial neural networks (ANN) are increasingly used to classify data based on similarity and have the ability to "learn" from the data. The networks consist of a series of nodes that are based on the input data set and functions describing the relationship between the nodes, input and output data. ANNs are able to extract the inherent structure, the underlying patterns, directly from a data set without

an explicit physical model by resolving nonlinear input-output relationships (functions) in complex systems. They have a broad spectrum of applications, ranging from speech recognition over image analysis to anomaly detection. ANNs can be used for modelling purposes, for example, the impact of groundwater extraction on rivers or predicting microbial water quality (Lin et al., 2008). However, the focus of this overview lies on the application of the methods to time-series analysis and fault detection. The potential in the domain of resource management and environmental monitoring is very high, while the method is increasingly applied to survey industrial processes.

The self-organizing map (SOM, Kohonen (2001)) is an ANN method based on competitive, unsupervised learning. SOMs have found use in many engineering applications, for example, for monitoring industrial process states or drinking-water quality in distribution networks (Dominguez et al., 2007; Mustonen et al., 2008; Corona et al., 2010). SOMs are based on vector quantization where an approximation to the distribution of input data vector is made by using a set of alternative vectors (also called codebook, reference, model or weight vectors), which share characteristics with the original data set (for example, probability density function), but with a smaller dimension. The codebook vectors are associated with nodes in a regular grid, a 2- or 3-D output space, where their position is based on their similarity. Similar situations, or process states, are thus located close to each other in the output space. This leads to clustering of the data into different groups, which can then be classified, for example, as *normal* or *hazardous*. As the SOM is a visual data-mining approach (Dominguez et al., 2007), it also allows tracking of the process state by visualizing trajectories projected to the output space. This occurs by selecting the node with the least discrepancy between the input data and the codebook vector, the best-matching unit (BMU), for each measurement time. The succession of the BMUs becomes the trajectory. The result is a set of vectors describing the system state at each point in time used in the analysis (Mustonen et al., 2008). The dimension of the data set is reduced and now describes the internal structure of the data matrix.

The first step in creating a SOM is to determine the grid. The number of units, or nodes, in the 2-D representation of the map was given by the default function in the toolbox used in this study (Vesanto et al., 2000):

$$nodes = 5 * IM^{0.54321} \tag{1}$$

where IM is the input data matrix.
The dimension of the map is then based on the ratio between the two largest

eigenvalues of the autocorrelation matrix of the IM. By using the ratio of the two largest eigenvalues, the lattice spacing in the map/grid becomes approximately uniform (Kohonen, 2001). Fig. 9 shows an example of normalized data used to constuct and train the SOM.
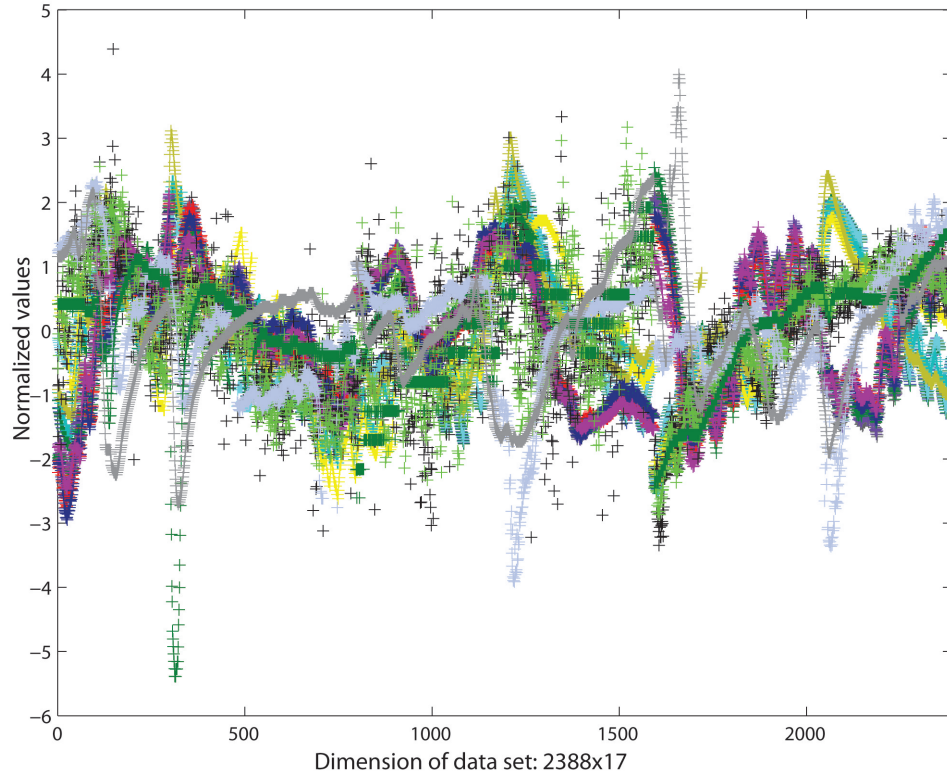
Figure 9: An example of input data used for the SOM analysis, showing the normalized time-series of 17 parameters. Groundwater head, temperature and electrical conductivity from six groundwater observation wells taken from three time periods and combined to give one data set with the dimension 2388 x 17. The analysis in Chapter 8 was carried out with this data set.

The codebook vectors can be linearly or randomly intialized. An example of random initialization is given in Fig. 11. The analysis in Chapter 8 was carried out using linear initialization. Linear initialization is based on the two eigenvectors with the largest eigenvalues, which provide initial axes. The eigenvectors are normalized and multiplied with the squareroots of the corresponding eigenvalues to give vectors (*evv*) with the same dimension as

the number of eigenvalues (and number of columns in the input matrix). The codebook vectors are a product of the *evv* vectors and the number of nodes in the grid (cf. Fig. 10, 'Grid without data'), giving rise to a n x d matrix, where n is derived from the number of nodes and d from the number of eigenvalues. This is part of the ordering process resulting in a grid with preliminary values derived from the original data set (Fig. 11).
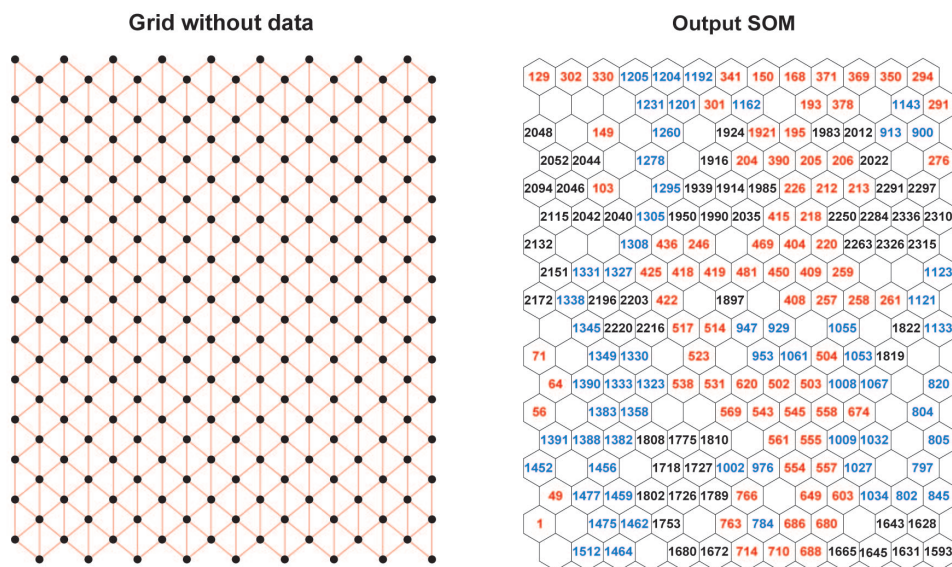
Figure 10: The left panel shows the grid structure of the SOM based on the ratio between the two largest eigenvalues. In this case there are 13x18 nodes in the map. The right-hand panel shows the graphical representation of the SOM. The numbers correspond to time steps (1:2388) and the colours to the time periods used in the analysis (red: Jan.'09, blue: Dec.'09, black: May'10, cf. Chapter 8).

The steps that are made during the construction, initialization and training, as well as visualization of the self-organizing map are summarized in Figures 10 and 11. During training, the best-matching units (BMUs) are calculated for each time step or line in the data matrix (dimension BMU: length input data set x 1). The BMUs of each time step correspond to the nodes with the greatest similarity between the input data and the map, thus the smallest error (Euclidean distance). Subsequent training of the map is based on a gaussian neighborhood function, which involves a learning rate factor and can be decribed as a relaxation process (Kohonen, 2001). Hereby,

the nodes surrounding the BMUs are updated and modified based on their distance to the BMU, previous state, predefined learning rate and input data vectors, a more detailed account is given in Chapter 8. In this sense, this methodology is essentially an optimization problem based on minimizing the error of a representation of input data in alternative dimensions (Kohonen, 2001).
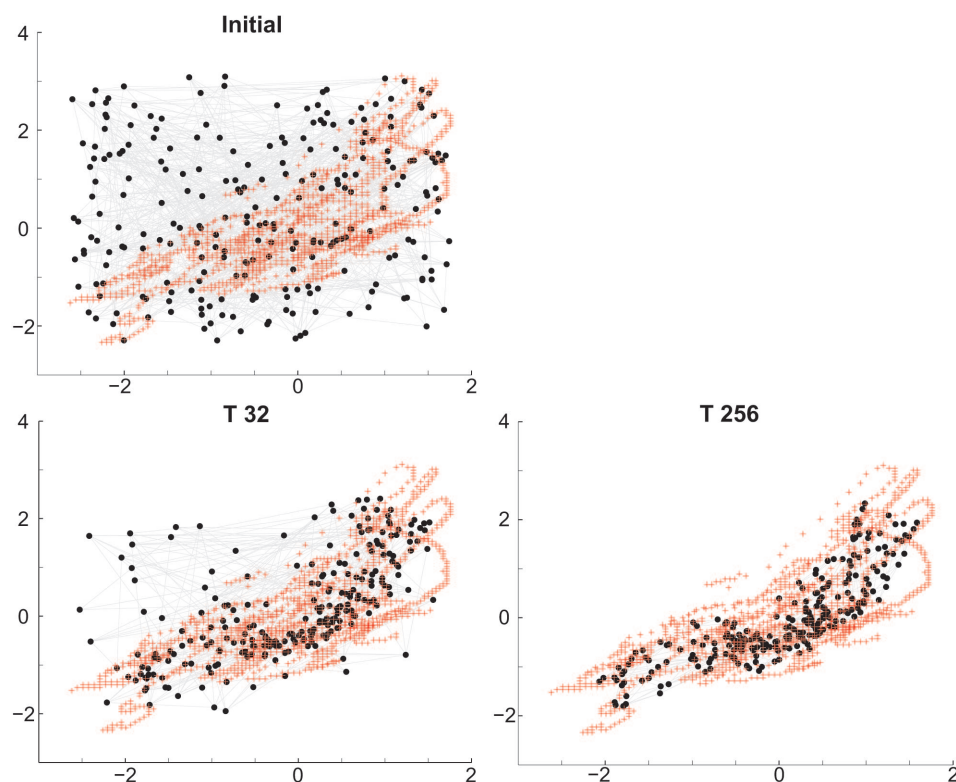
Figure 11: The red crosses represent a projection of two input data sets (normalized groundwater head measurements from two observation wells) and the black dots, the representation of the data in the normalized input space (nodes). The black dots are iteratively adapted to match the pattern prescribed by the input data. The three panels show the 'representation ability' of the SOMs at three stages during training (initial distribution of the nodes and input data, after 32 time steps and after 256 time steps).

The graphical representation of the SOM output is shown in Fig. 10. The sheet with hexagonal elements (polygons) and numbers represents the inherent patterns in the input data matrix. The numbers show the time

steps with the smallest error for each hexagon in the map. Although the overall structure of the data set can be analysed and visualized, the development over time is not evident. Trajectories can be used to visualize the data of selected observations and follow the succession of BMUs through time (Corona et al., 2010). However, the distance between polygons in the visualization is not proportional to the distance between the system states represented by the polygons. In addition, the BMUs often represent more than one line of input data. Some authors have used a combination of SOMs and nonlinear projection methods to capture the temporal aspect of system state development (Bernataviciene et al., 2006; Mustonen et al., 2008; Lischeid, 2009). Sammon's mapping (SM) is a nonlinear mapping algorithm aimed at preserving the distances in the measurement vector in a 2-D projection (Sammon, 1969). Sammon's mapping is very useful in determining the shape and density of clusters and the relative differences between these clusters (Kolehmainen et al., 2003), which cannot be represented by a SOM.

In Sammon's mapping, the cells (points) are located so that the distance between the cells represents dissimilarities, which give information about the change of a system under consideration over time (Fig. 12). The coordinates of the BMU calculated for the SOM for each time step are used as initialization coordinates for the SM, the combination of SOMs and SM is referred to as SOM-SM. Each row of data matrix thus receives xy coordinates (for the 2-D visualization). As the coordinates are based on the BMUs, the resulting graphical representation (Fig. 12) resembles the development of the system state through time. The plot is organized like a trajectory and similar situations are located in a similar position on the map (Fig. 12). The shading of the points in the map can be chosen to match the values of a particular parameter (such as groundwater head in observation well W3, as in Fig. 12), or a proxy indicator, such as faecal indicator bacteria density. Independent parameters, not used for the analysis, can thus be used to highlight particular situations or system states. The main mathematical steps of the SOM-SM analysis are described in Chapter 8.

The following two Chapters (6 and 8) describe the application of the PCA and the SOM-SM analysis methods to a multivariate data set.
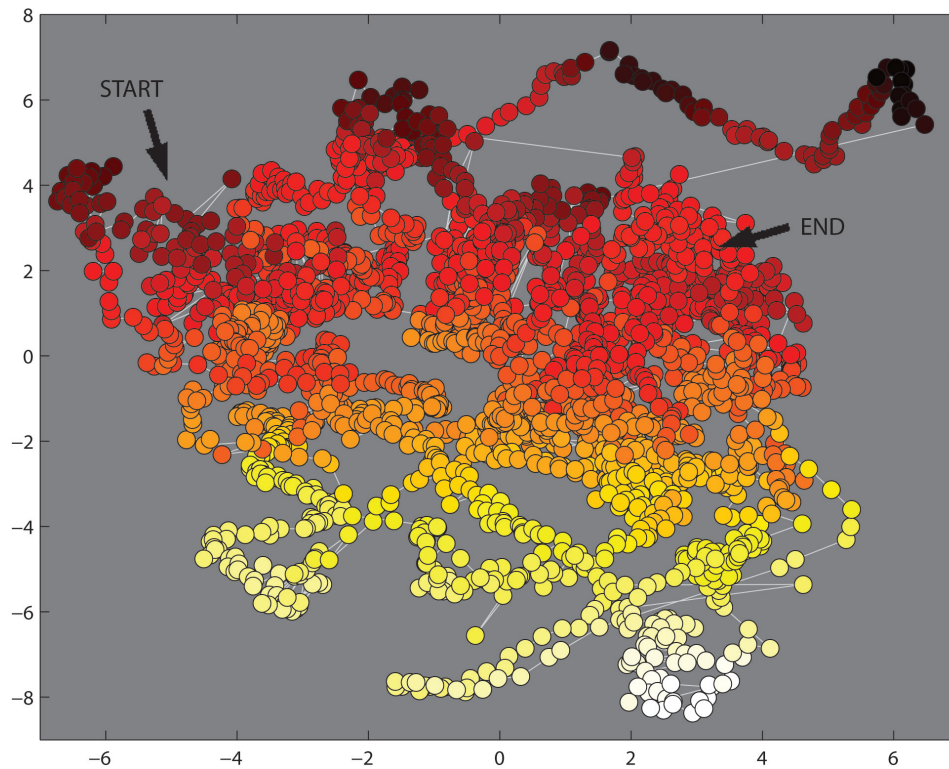
Figure 12: The output of the SOM-SM analysis using the 17 variables shown in Fig. 9. The shading of the points represents the normalized groundwater head measured in observation well W3. Each point (circle, spot) stands for one observation time (1:2388). Further discussion of this visualization is given in Chapter 8.

6 Principal Component Analysis of Time-series for Identifying Indicator Variables for Riverine-groundwater Extraction Management

Rebecca M. Page [a], Gunnar Lischeid [b], Jannis Epting [a], Peter Huggenberger [a]

[a]:Applied and Environmental Geology, Institute of Geology and Paleontology
   Department of Environmental Sciences, University of Basel
   Bernoullistrasse 32, CH-4056 Basel, Switzerland
[b]:Institute of Landscape Hydrology, Leibniz Centre for Agricultural Landscape Research
   Eberswalder Strasse 84, D-15374 Müncheberg, Germany

E-mail addresses:
rebecca.page@unibas.ch
lischeid@zalf.de
jannis.epting@unibas.ch
peter.huggenberger@unibas.ch

Corresponding author:
Rebecca M. Page
rebecca.page@unibas.ch
Applied and Environmental Geology, Institute of Geology and Paleontology,
Department of Environmental Sciences, University of Basel. Bernoullistrasse
32, CH-4056 Basel, Switzerland
Tel.: +41 61 2673447
Fax: +41 61 2672998

6.1 Abstract

Although alluvial aquifers connected to rivers can be a rich source of drinking water, they are susceptible to contamination by infiltrating river water.

The processes governing river-groundwater interaction are variable in time and space. Natural filtration mechanisms are often not sufficient during high discharge events in the river. To capture the dynamics of river-groundwater interaction, indicator parameters that can serve as proxies for river water infiltration need to be derived. Principal component analysis was used to identify indicator wells and derive indicator parameters for a study area in NW Switzerland. The results showed different sources of variation in the parameters, including river stage fluctuations. The multivariate approach highlighted differences between observation wells based on the response of the measured parameters to effects of damping and delay of the input signals. Three observation wells were shown to be representative of river-groundwater interaction dynamics in the study area. Of the three parameters analyzed, groundwater head and electrical conductivity are recommended as a combined proxy for river water infiltration in the study area. In contrast, temperature proved not to be a reliable indicator.

6.2 Introduction

Alluvial deposits are often rich sources of drinking water owing to the high permeabilities in the proximity of rivers (Su et al., 2007). In the last few centuries, many groundwater extraction wells were constructed in the proximity of rivers and water suppliers rely on natural groundwater recharge from the river to maintain the availability of extractable water (Sheets et al., 2002). River-groundwater interaction is heterogeneous both in time and space and depends on multiple factors, including river discharge, the morphologic and sedimentologic character of the river bed and the heterogeneity of the hydraulic properties of the boundary layer (Sophocleous, 2002). For example, the conductance of the river bed is a major factor in determining river-groundwater exchange and is a function of river discharge. The heterogeneity of hydrogeological zonation is scale dependent and non-stationary. The focus of river-groundwater interaction has frequently been on quantity,

however, the exchange of water between the river and the aquifer is one of the most important factors for groundwater quality (Wroblicky et al., 1998). Microbial contamination of riverine origin is one of the major concerns for drinking water suppliers extracting riverine groundwater (Sheets et al., 2002; Regli et al., 2003). Under average hydrological conditions, the natural filtration capacity of the subsurface matrix is sufficient to remove many physical and biological components from the infiltrating river water (McDowell-Boyer et al., 1986; Taylor et al., 2004; Gupta et al., 2009). However, high discharge events often result in a degradation of water quality (Wilkinson et al., 2006). (McKergow and Davies-Colley, 2010) have shown that during storm events, the concentration of *E. coli* in river water can be up to 30 times higher than during base flow conditions. As a consequence, the concentration of microorganisms in the adjacent groundwater can also increase by several orders of magnitude (Regli et al., 2003). Taking this into account, and that the first barrier to colloids and contaminants entering the aquifer - the clogging of the river bed - is weakened during high discharge events (Sophocleous, 2002), the potential for river-borne microbial contamination of extraction wells near the river can be significantly elevated (Regli et al., 2003). Continuous monitoring of the microbial load in groundwater is not suitable for applications in the field, as the cost and time investments are considerable. There have been recent advancements in real-time microbial monitoring methods, these are however not yet readily available for many applications (Martinez et al., 2010). Thus, the question arises whether proxies can be used to continuously assess the contamination hazard of riverine groundwater extraction wells. The concept foresees proxies to identify situations when river water infiltration is sufficient to pose a threat to the quality of the extracted drinking water. In addition, the identification of the possible infiltration hot spots along a river reach and areas within the aquifer strongly influenced by infiltration is an important step towards the adaptive and sustainable management of the water resource (Huggenberger et al., 1998; Epting et al., 2008). The optimal measurement parameter as a proxy for infiltration is dependent on the local and regional hydrogeological settings. Some authors have used temperature as a natural tracer

(Conant, 2004; Hoehn and Cirpka, 2006; Keery et al., 2007), while others use electrical conductivity to track river water infiltration (Sheets et al., 2002; Cirpka et al., 2007). In this study we have investigated time series of three parameters (groundwater head, electrical conductivity and temperature) at different wells for their usefulness as proxies for infiltrating river water. The usefulness was assessed using principal component analysis (PCA) to determine the factors dominating each parameter. PCA can be used to extract information on dominating influences or processes and patterns from large environmental data sets (Gerbrands, 1981; Longuevergne et al., 2007; Lischeid, 2009). Helena et al. (2000) studied the influence of temporal variations caused by precipitation or agricultural activity, on groundwater quality using a PCA on a large number of discrete samples. Based on the assumption that many physico-chemical parameters are representative of processes affecting and influencing an aquifer, Sanchez-Martos et al. (2001) studied a large number of parameters, including temperature and electrical conductivity. Both studies were able to identify the dominating processes and define a zonation pattern of influence for various sources. Lewandowski et al. (2009) applied PCA to continuous groundwater head measurements to study the relationship between groundwater head and river stage over time and were able to deduce infiltration patterns.

The aim of this study was to select an optimal set of parameters and observation wells that can be used as proxies for the potential threat to drinking water extraction due to infiltrating river water. This hazard-based approach to managing groundwater extraction included the analysis of multivariate time series of groundwater observation wells at different locations along a river section. The analysis was based on identifying the components influencing the measured parameters and characterising their association with infiltrating river water.

6.3   Study Area and Methods

Study area
The study area, approximately 0.9 km^2, is located in the lower Birs Valley

of Northwest Switzerland (Fig. 13). The alluvial system mainly consists of high-permeability coarse gravels. The river Birs was canalized at the end of the 19th century and since then, the river bed has incised several meters into the former floodplain, from which it was subsequently disconnected. The river Birs runs 75 km through the Swiss Jura and joins the Rhine in Basel, creating a catchment area of 866 km^2. The mean annual flow near the mouth (Münchenstein) is 15.4 m^3 s^{-1} and storm flows can reach up to 383 m^3 s^{-1}. The aquifer material mainly consists of carbonate gravel, which is often well rounded, with variable sediment sorting, and few intercalations of clay or silt layers, resulting in a large variance in hydraulic properties. The carbonate gravel components of the aquifer are of Triassic to Jurassic origin with hydraulic conductivities between $10^{-2}$and 3 10^{-3} m s^{-1}. The aquifer bottom consists of Tertiary deposits of the Rhinegraben Molasse, generally with very low hydraulic conductivities (kf 10-5 to 10-8 m s^{-1}, aquitard). The maximal depth to groundwater is 29 m and the thickness of the saturated zone varies between 0.6 m and 10 m. Most of the water is supplied to the aquifer from river infiltration, lateral inflow from the local catchment area, and intermittent artificial recharge. There were eight active production wells in the study area (Fig. 13) supplying water to six communities, approximately 51.000 people, amounting to 5 million m^3 drinking water per year. The experimental field site for river-groundwater interaction was set up in the gravel aquifer in the vicinity of the river bank (Fig. 13).

Measurement Set-up To address the scale and dynamics of the river-groundwater interaction, nine observation wells in the experimental field site were equipped to monitor multiple parameters (Fig. 13). In order to include depth-oriented measurements, the wells were grouped into three clusters of three observation wells, each with a different depth of filter screen. The diameter of the PE tubes lining the observation wells was 4.5". Each measurement installation monitored groundwater head (GH), electrical conductivity (EC) and temperature (T). Four further observation wells recorded GH and T and two additional observation wells provided only GH measurements. River stage and temperature were also recorded.

Data pre-processing
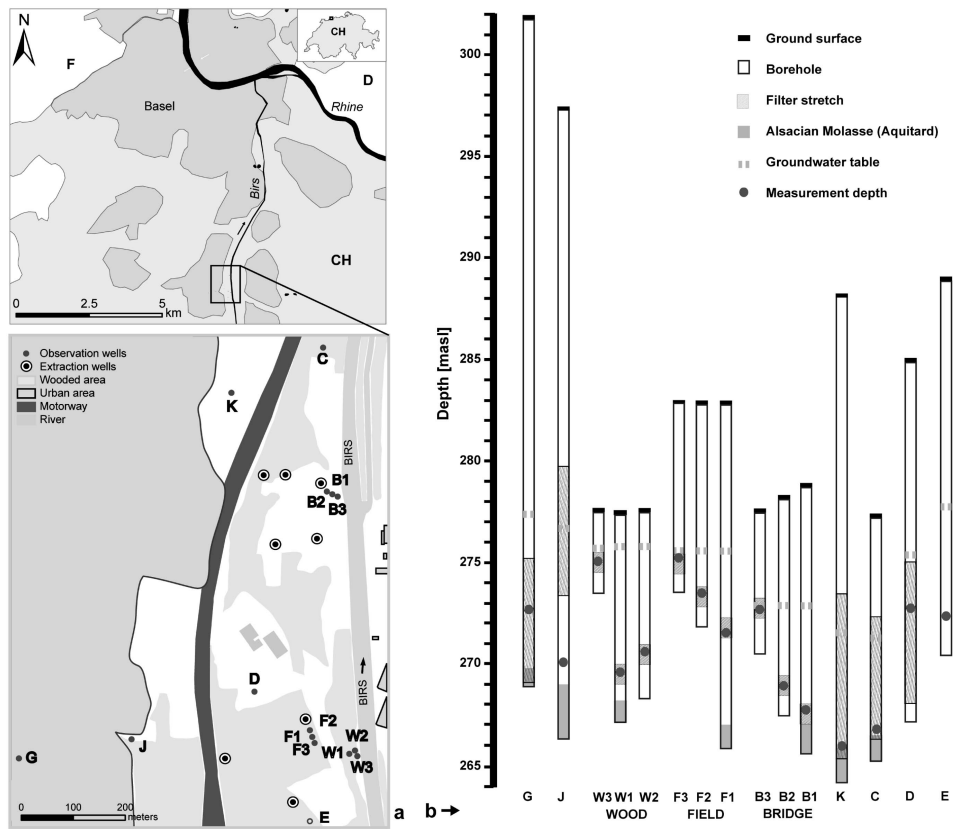
Figure 13: a) Location of study area, observation and extraction wells. b) depth of filter stretches and measurement instruments in observation wells.

The data were checked for outliers and irregularities caused during equipment maintenance and corrected where possible, or excluded from the data set. Some of the observation wells had to be omitted from the analysis, as the variance observed was smaller than the measurement error. The data sets were standardized to a zero mean and unit variance (Mustonen et al., 2008). The PC analysis was carried out with a 1-hour resolution.

Principal component analysis (PCA)

A PCA reduces the number of dimensions in the data set, while not losing detail or underlying patterns observed in some or all of the observation wells (Lischeid, 2009). This methodology is based on correlation coefficients of the data matrix. The principal components (PCs) are uncorrelated (are orthogonal to each other) and represent the joint variance observed in the data set (Gerbrands, 1981). The PCA was carried out for 16 head observation data sets, ten EC observation data sets and 15 temperature data sets to identify underlying patterns or signals in the data set. The PCA was carried out using the MATLAB princomp function. The derivation of the optimal number of PCs, the minimal number of nontrivial components describing the maximum amount of variance, has received considerable attention (e.g. Cliff (1988); Jackson (1993)). The Kaiser, or Kaiser-Guttman, criterion includes all PCs with an eigenvalue greater than one (Kaiser, 1960). The scree test (Cattell, 1966) is based on ranking the eigenvalues in descending order and identifying the break between the components representing the variance due to input signals and those due to ambient noise. This method can be difficult to apply, especially when there is no obvious break between reliable and interpretable components and components due to random noise (Jackson, 1993). The third method, also called communality rule, considers the amount of variance described by each component and the cumulative sum. The following PCA results considered the Kaiser-Guttman criterion, the scree test and the communality rule.

6.4 Results

The time series used for analysis were recorded in January 2009. Selected time series are shown in Fig. 14. During the measurement period, two high discharge events occurred (max. discharge: 80 m^3 s^{-1}, 19/20 and 24/25 January 2009). Decreases in T and EC were observed during the high river discharge events, as were increases in GH. Groundwater was continually extracted throughout the two high discharge events. Especially the GH data showed a strong relation to groundwater extraction. Observation wells close to an extraction well, e.g. F3 showed a stronger response than wells further away, e.g. J. Other potential sources of variation included air and river water temperature fluctuations.

The PCA was carried out separately for the water table (GH and river stage, 16 time series), T (13 time series) and EC (7 time series). The resulting patterns differed considerably and are discussed separately. Fig. 15 shows the representation of the GH, T and EC loadings for PC1 and PC2. The loadings were calculated as the correlation between the PC scores (projections of the input time series into eigenvalue space) and the input time series. The results of the PCA are displayed in relation to a unit circle and in dependency of PC1 and PC2 (Fig. 15). The unit circle gives the maximally possible loadings for time series where the variance is completely explained by the first two components. Each point represented the loadings of the data set from each observation well for PC1 and PC2. The smaller the distance from the unit circle, the more variance was explained by the first two PCs, and thus the less significant other sources of variation were in determining the observed fluctuations.

For all three parameters, only the eigenvalues of the first two PCs exceeded one. Both the scree test and the communality rule indicated the dominance of the first two components in determining the observed patterns in groundwater dynamics. The first two components were able to explain over 90 percent of the variance of most of the GH and EC time series. Three PCs were required to explain 90 percent of the variance observed in the T time series.
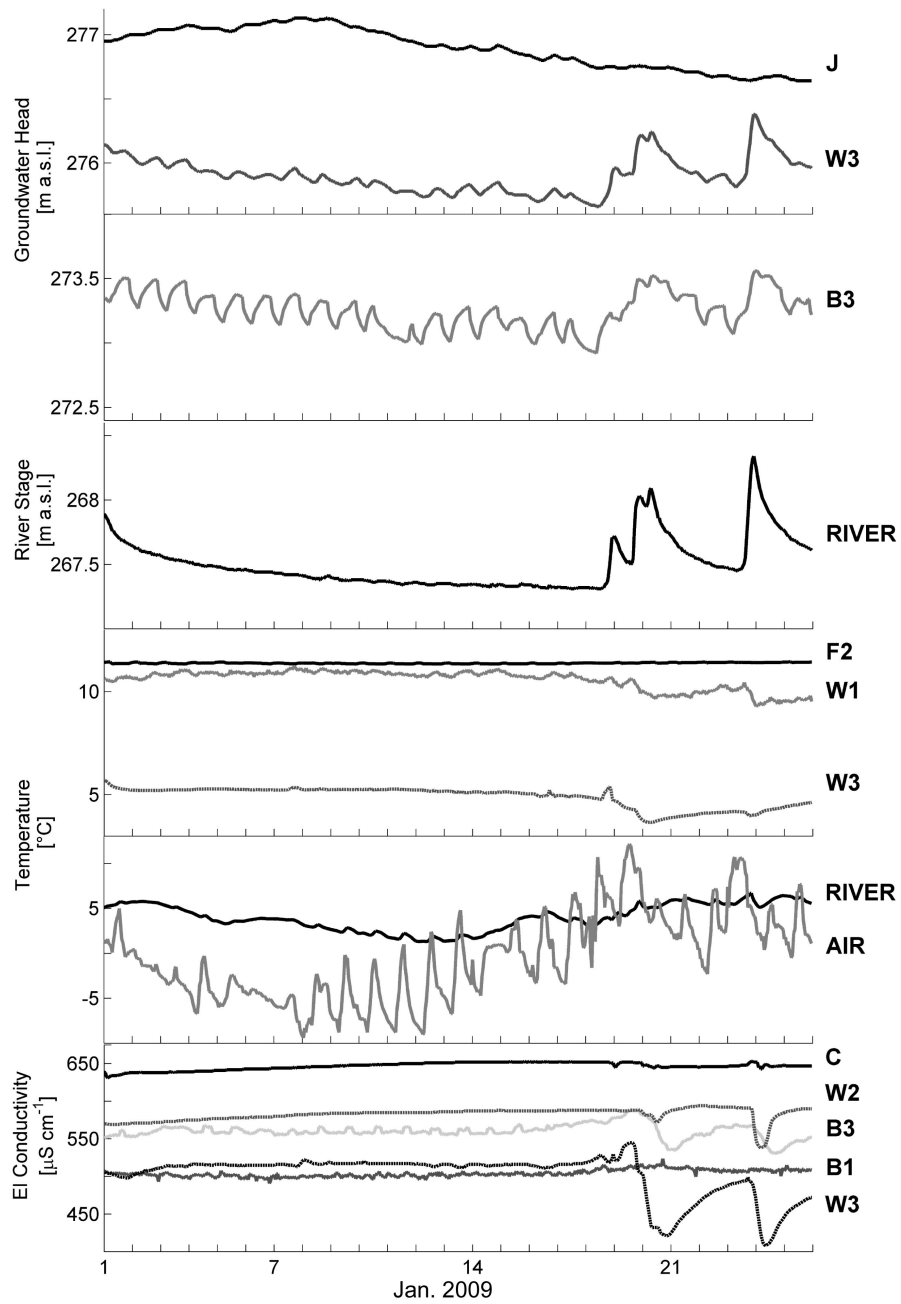
Figure 14: Selected time series used in the PCA during the observation period in January 2009. The parameters measured were: groundwater head, river stage, groundwater temperature, river and air temperature, and groundwater electrical conductivity
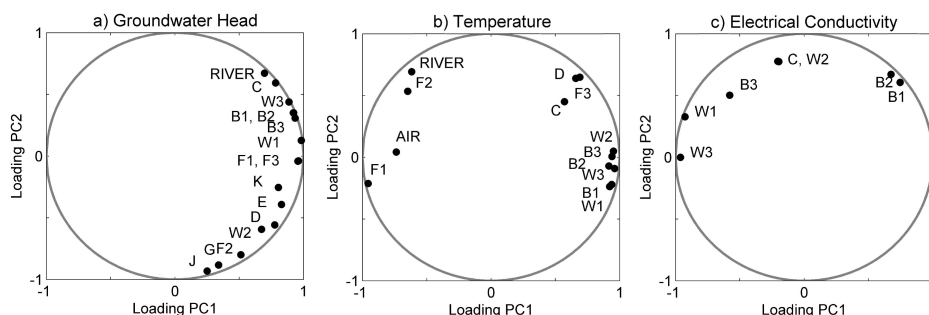
Figure 15: PC1 and PC2 loadings in relation to a unit circle. a) Groundwater head and river, b) temperature, c) electrical conductivity

Groundwater head With few exceptions, a good correlation was found between the counter-clockwise position of the individual loadings around the unit circle and the distance of the corresponding observation wells from the river. While the GH loadings for PC1 showed a linear relationship with distance from the river, the loadings for PC2 displayed logarithmic behaviour (Fig. 16). As both PCs showed a relationship with distance from the river, the two PCs are likely to represent two effects of similar origin, e.g. hydraulic fluctuations of the river stage and a damping of the fluctuations as they propagated through the aquifer (Fig. 14). Two exceptions from this observation were F2 and W2. Geophysical studies in the area between these two observation wells showed remnants of old river beds, which were responsible for the sedimentological composition of hydraulic conductivity of the aquifer. The zonation of hydrogeological properties is based on units of similar origin and composition. These two observation wells were located approximately 6 m (W2) and 96 m (F2) away from the river, on the edge of such an ancient river bed. The other observation well that showed uncharacteristic behaviour was K. The loading from this observation well was higher than would be expected considering its distance from the river. However, the saturated aquifer thickness around K was considerably lower than at equivalent distances elsewhere in the study area. Although, increased transmissivity leads to a faster response to river stage fluctuations (Serrano and Workman, 1998), the response to the river stage fluctuations around K

were more rapid than in areas with a greater saturated thickness. As transmissivity is dependent on saturation thickness and hydraulic conductivity, the results suggested that hydraulic conductivities in the northern area are larger than in the southern areas. Other observation wells in the northern area (B1 and B2) show a similar pattern, however not as explicitly as K. These results led to the conclusion that the dominant process in determining GH dynamics was closely associated with river stage fluctuations and that local heterogeneity of hydraulic conductivities played a secondary role.
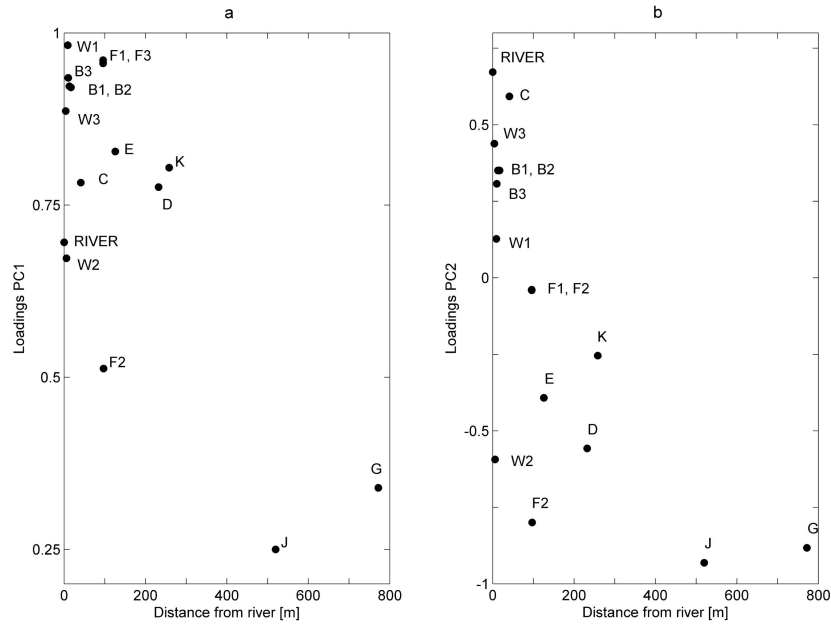
Figure 16: Groundwater head and river loadings for PC1 (a) and PC2 (b) in relation to distance from the river

Temperature

The distribution of the loadings around the unit circle yielded a more complex pattern than the GH analysis (Fig. 15). The river and air temperatures were both located on the opposite side of the circle to the T values for most of the groundwater observation wells. As the observation period was in January, the river and air temperature were lower than the groundwater temperature, except during the two high discharge events, when air temper-

ature increased. (Fig. 14). River and air temperature were closely linked, however, damping effects were observable in the river temperature maxima and minima, as well as in daily fluctuations, which were significantly smaller in the river (Fig. 14). F1 and F2 showed a similarity to air and river temperatures in their positions around the unit circle. The majority of the groundwater observation wells were clustered at the same level as the air temperature with respect to PC2 (Fig. 15). The effects of delay of temperature transmission were represented by PC1, as there was little delay between changes in air and river temperature, but a significant delay between changes in air and groundwater temperature, e.g. in B2. On the other hand, PC2 also represented effects of damping through other influences, e.g. river water temperature was also dependent on discharge. Due to the measurement setup, as the distance from the river increased, so did the installation depth of the sondes. The assignment of damping by heat transport through the substrate matrix was therefore not distinguishable from heat transport due to river water infiltration. The distance from both the river and from the ground surface influenced the amount of variance observed in the observation wells. The greater the distance, the greater percentage variance was explained by the first two PCs, however with three exceptions: F2, C and D (Fig. 17).

Only two measurement points were influenced by PC3. Air temperature and groundwater temperature measured in observation well C. Air temperature and groundwater temperature in C increased during the precipitation event, as opposed to temperature in W3, which decreased gradually and delayed. Observation well C (53 percent variance explained by PC1 and PC2, a further 40 percent explained by PC3) was located 41 m away from the river in a shallow part of the aquifer. The river bed in this area was lowered due to erosion processes and groundwater exfiltrates to the river. The temperature variance pattern can be assigned to the brief periods of infiltration during high discharge events, when river water is able to infiltrate due to a rapid increase in river stage. The difference to the other observation wells subject to river water infiltration arose due to the changing direction of river-groundwater interaction in the proximity of observation
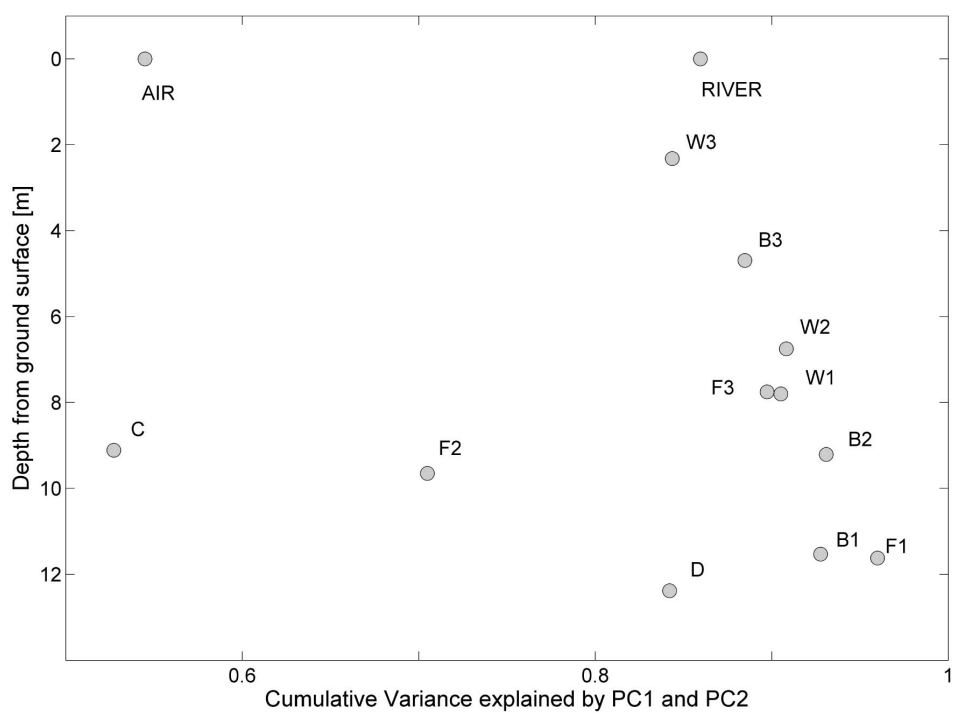
Figure 17: Cumulative variance explained by PC1 and PC2 for the temperature measurements.

well C. Groundwater temperature was mainly affected by the transmission of air temperature with damping and delay due to diffusive heat transport through the subsurface. Although W3 was located in an area where hydraulic exchange processes between the river and the groundwater were rapid, the temperature signal from the river was too weak to override the effect of damping and delay caused by the hydrogeological properties of the aquifer material in the study area.

Electrical conductivity

The position of the points around the unit circle was mainly attributed to their response during a high discharge event, as there was little variation during the rest of the observation period (Fig. 14). The observation well that showed the greatest response was W3, while B1 showed the weakest response. Groundwater EC was lower than the individual EC measurements in the river, thus decreases in groundwater EC were observed as recent riverine water reached the observation wells. As observed for the GH analysis, the counter-clockwise orientation of the observation wells around the unit circle corresponded to their distance from the river, with exception of C and W2. The high EC values observed in C suggested a longer residence time, than in other observation wells closer to the river (Fig. 14). During the two high discharge events, only small perturbations were visible (Fig. 14). This is in agreement with the results from the T analysis, suggesting exfiltrating conditions except for during high discharge events. The fluctuations during the high discharge were only minor, as most of the signal was lost during transport. As for the GH analysis, groundwater flow and pressure wave propagation in the proximity of W2 is likely to be hindered by ancient channel and river bed characteristics. During the two high discharge events two potential factors could influence EC. Infiltrating river water had lower EC than the groundwater and would effect a decrease in EC, as was the case for W3. Another factor is determined by rising groundwater head and pressure wave propagation. These two processes give rise to higher mineralised water, as previously deposited ions are entrapped and lead to an increase in EC, as was observed for B1. The observed EC pattern in B3 and C can be attributed to a combination of these two factors, which resulted in a pro-

longed initial increase and subsequent decrease in EC. PC1 represents river water infiltration and transport, whereas PC2 is a function of groundwater head and pressure wave propagation due to river stage fluctuations.

6.5   Discussion

The observation period (January 2009) included high and average discharge conditions. The parameter response to the two high discharge events was dependent on the observation well under consideration. The observation well closest to the river, W3, showed the strongest response to the two high discharge events. GH was strongly influenced by the diurnal groundwater extraction regime and fluctuating river stage. Of the two, the effect of groundwater extraction was evident throughout the study area and contributed a smaller amount of variation than the effect of fluctuating river stage. The results from the PCA showed that the main factor influencing GH dynamics was related to river stage fluctuations. The pressure wave induced by river stage fluctuations was damped as it propagated through the aquifer. The damping effects observed were related to travel distance from the river and to hydrogeological heterogeneity of the aquifer. A higher river stage with respect to groundwater head favours infiltration, however, mass transport is generally slower than the pressure wave propagation (Lewandowski et al., 2009). In comparison, EC is more frequently associated with mass transport driven by advection and dispersion. While the results support this observation, EC can be influenced by non-conservative behaviour in addition to advection. Factors leading to the non-conservative behaviour of EC include chemical reactions in the river bed and aquifer (Sheets et al., 2002; Cirpka et al., 2007). Although the interpretation of EC fluctuations is difficult without a conservative tracer (e.g. isotope analysis, Sklash et al. (1986)), the amount of infiltrating water was sufficient to induce a noticeable drop in EC a clear response to high discharge events in the river. The EC results corroborated higher hydraulic conductivities in the northern region of the study area derived from the GH analysis. The difference between B3 and W1 (similar distance from the river) during the high discharge events

can be assigned to higher hydraulic conductivities resulting in increased infiltration in the vicinity of B3. The EC pattern was related to the distance from the river, however, in contrast to GH dynamics, the effects of damping and delay, as caused by dispersion and physicochemical reactions, were strongly influenced by local characteristics of the aquifer material surrounding the observation wells. Measurements of GH, T and EC in W2 all showed similar patterns, indicating that water exchange processes in this area were reduced. In contrast, the results suggested hydrogeological conditions in the vicinity of B3 that delay the transport of water with lower EC (riverine water) without significant damping. The observed effects can be assigned to damping over distance in areas with similar hydrogeological properties and effects assigned to different hydrogeological properties (such as hydraulic conductivity). The T analysis showed that the main drivers for temperature variation in the study were not only linked to river water infiltration, but also to air temperature fluctuations. The damping and delay of the temperature signal was partly caused by transport through the subsurface and groundwater. While mapping streambed temperatures by direct measurements has successfully led to estimates of seepage across the river-groundwater interface (e.g. Conant (2004); Schmidt et al. (2007); Vogt et al. (2010)), temperature was not applicable as a tracer for recent river water infiltration in this study. When considering the selection of observation wells for monitoring purposes to use in an early warning system for extraction well management, the study area can be split into two sections. B3 would be a good observation well to monitor potential contamination in the northern region of the study area. This region is characterised by higher hydraulic conductivities, but also requires greater discharge events for significant infiltration. In the southern region, smaller discharge events in the river were able to evoke a stronger hydraulic gradient between the river and the groundwater, thus leading to a greater infiltration. The results of the analysis of the time series from the observation wells in the southern region suggested lower hydraulic conductivities, resulting in a higher filter efficiency of the aquifer matrix. Although W3 was the most responsive observation well in this area, W1 was more representative of the behaviour of

the aquifer and would thus provide a better monitoring well for potential contamination of the surrounding groundwater extraction wells.

6.6   Conclusion

The PC analysis was able to identify parameters, in this case EC and GH, that can be used as proxies for potential contamination due to infiltrating river water. GH measurements were rapid to respond to pressure changes due to river stage fluctuations, but did not necessarily correlate well with mass transport. Although EC is generally a non-conservative tracer, by considering the relative differences between multiple observation wells, local patterns of responses to high discharge events were derived. The application of PCA also allowed the identification of observation wells (W1, W3 and B3) suitable to serve as indicator observation wells for monitoring river water infiltration. The next step is to identify crucial states when hydraulic gradients between the river and the groundwater are large enough to lead to higher flow velocities in the region of B3. The identification of such states can then be used as an early warning tool for groundwater management purposes. The combined approach of using proxies for mass transport and pressure wave propagation allows to characterize the aquifer in terms of hazards for extraction wells and hot spots for infiltration. The combination of EC and GH measurements in groundwater observation wells within 10 m of the river provides sufficient information to use in an early warning system for the extraction wells in the study area.

6.7   Acknowledgements

7 Faecal Indicator Bacteria: Groundwater Dynamics and Transport following Precipitation and River Water Infiltration

Rebecca M. Page [a*], Stefan Scheidler [a], Elif Polat [b], Paul Svoboda [c], Peter Huggenberger [a]

[a]:Applied and Environmental Geology, Institute of Geology and Paleontology
   Department of Environmental Sciences, University of Basel
   Bernoullistrasse 32, CH-4056 Basel, Switzerland
[a*]Corresponding author
   Email: rebecca.page@unibas.ch
[b]:Swiss Tropical and Public Health Institute, University of Basel,
   Socinstrasse 57, CH-4051 Basel
[c]:Cantonal Laboratory Basel-Landschaft
   Gräubernstrasse 12, CH-4110 Liestal

7.1 Abstract

Faecal contamination of drinking water extracted from alluvial aquifers can lead to severe problems. River-water infiltration can be a hazard for extraction wells located nearby, especially during high discharge events. The high dimensionality of river-groundwater interaction and the many factors affecting bacterial survival and transport in groundwater make a simple assessment of actual water quality difficult. The identification of proxy indicators for river-water infiltration and bacterial contamination is an important step in managing groundwater resources and hazard assessment. A proxy-based approach requires in-depth knowledge of the relationship between the variable of interest, e.g. river-water infiltration, and the proxy indicator measured. In addition, the resolution of microbial monitoring studies is often too low to verify this relationship. To this means, continuously recorded

physico-chemical parameters (temperature, electrical conductivity, turbidity, spectral absorption coefficient (SAC), particle density) were analyzed in combination with two faecal indicator bacteria, *Escherichia coli* and *Enterococcus* sp.. Sampling for faecal indicator bacteria occurred on two temporal scales: (a) bi-weekly monitoring, and (b) three-day event-based sampling. Both sampling set-ups showed the highest bacterial densities in the river. *E. coli* and *Enterococcus* sp. densities decreased with time and flow path in the aquifer. The event-based sampling was able to demonstrate differences in bacterial inactivation between clusters of observation wells linked to aquifer composition. Although no individual proxy indicator for bacterial contamination could be established, it was shown that a combined approach based on time-series of physico-chemical parameters could be used to assess river-water infiltration as a hazard for drinking-water quality management. *Keywords: faecal indicator bacteria; groundwater; river-water infiltration; multiparametric time-series*

7.2   Introduction

Groundwater is an important source of drinking water in many countries, providing approximately 50 percent of global drinking water supply, as well a significant part of water required for industrial and agricultural purposes (Zektser and Everett, 2004). Industrial and urban development has increased the pressure on aquifers, leading to degradation of both the quality and the availability of groundwater (Foster and Chilton, 2003). Faecal contamination of groundwater is a major concern for municipal water suppliers with low treatment levels and can lead to severe disease outbreaks (Kukkula et al., 1997; Sheets et al., 2002; Regli et al., 2003; Nichols et al., 2009).

The river bank and bed act as a primary barrier to infiltrating river water and many river-borne contaminants. The filter capacity of the aquifer material further reduces the contaminant load as water passes through the subsurface matrix, so that water extracted from wells near rivers can satisfy the quality standards for drinking water during average flow and surface-water quality conditions (Regli et al., 2003; Taylor et al., 2004; Dash et al.,

2010). However, increased bed sheer stresses and related turbulence during high discharge events lead to the winnowing of fine sediments, mobilisation of the river bed and an increase in river-water infiltration (Mutiti and Levy, 2010). High discharge events alter the water depth in the river leading to a greater inundated area and gradients between the river and groundwater favouring infiltration. In addition to increased connectivity and infiltration, waterborne pathogenic bacteria densities in river water can be up to several orders of magnitude higher during storm events than during base flow conditions (Regli et al., 2003; McKergow and Davies-Colley, 2010). In consequence, groundwater quality is severely degraded so that drinking-water extraction cannot be continued and alternative sources need to be used or 'boil water' warnings issued.

Some drinking-water suppliers make use of decision support or simple surveillance systems to manage groundwater extraction. The arrival time of infiltrating water at the extraction wells is dependent on a number of hydrogeological and hydrological properties, rendering a complex pattern in space and time (Krause et al., 2007). Groundwater flow simulations can help to delineate well capture or well head protection areas (WHPA), upon which management decisions are then based. As bacteria can be transported faster than the average pore water velocity (Unc and Goss, 2003; Taylor et al., 2004), many WHPA may nevertheless be susceptible to faecal contamination originating from nearby rivers during high discharge events. In addition to the heterogeneity of river-groundwater interaction and the non-linear relationship between groundwater flow and microbial transport, there are but few microbial field studies during high discharge events that can be used to validate these models in terms of bacterial hazard for drinking water extraction wells.

Although there have been many studies on the processes influencing bacterial and viral transport and survival in porous material under laboratory conditions (John and Rose, 2005), to assess actual water quality, the behaviour of waterborne pathogens needs to be studied in the environment and methods developed to detect hazards such as riverine microbial contamination. The controlled conditions in laboratory studies are not able to

reproduce the natural heterogeneity of environmental sites, which is a decisive factor in determining potential hazards for drinking water extraction wells (Regli et al., 2003; Frind et al., 2006).

As the detection of many of the waterborne pathogens requires complex analytical procedures (Brookes et al., 2005), *E. coli* and *Enterococcus* sp. are often used as faecal indicator bacteria (FIB) for microbial pathogens in drinking water (Borchardt et al., 2004; Pronk et al., 2007). While useful for background studies, where obtaining information on bacterial densities is not a question of time, the delay between sampling and obtaining the results of the FIB analysis is still too long to be useful in assessing actual water quality (Lin et al., 2008). Although there have been recent advancements in real-time monitoring methods, they are not yet widely available for many applications (Marrone, 2009; Martinez et al., 2010). The application of flow-cytometric methods requires controlled environmental conditions, a substantial investment for drinking water suppliers and is therefore problematic for decentralized applications. In such cases, the use of further proxy indicators is required to assess actual water quality. Proxy indicators are parameters that are easier to measure or quantify and the relationship with the variable of interest, e.g. infiltrating river water, is well-established. River-water infiltration, for example, can be assessed by high-resolution monitoring of physical parameters, such as electrical conductivity or temperature, close to or in the river bed (Conant, 2004; Keery et al., 2007; Vogt et al., 2010). Other researchers have focussed on identifying proxy indicators for bacterial transport, including turbidity, particle density and size distribution, or the spectral absorption coefficient (SAC at 254nm) (Brookes et al., 2005; Pronk et al., 2007; Stadler et al., 2010). Many of these proxy indicators are applicable in karst systems (Pronk et al., 2007; Stadler et al., 2010), where the transport phenomena associated with precitipation events are to different to alluvial aquifers, or require measurements in the river bed (Keery et al., 2007; Vogt et al., 2010), increasing the susceptibility to damage due to mechanical stress.

While many studies have established the relationship between high discharge or precipitation events and elevated bacterial densities, the resolution

of the studies, bi-weekly or less frequent, is insufficient to show the temporal development of faecal indciator bacteria densities during an event (Grisey et al., 2010; Nnane et al., 2011). However, drinking-water extraction-well management requires information on actual water quality with a sufficiently small time resolution, so that management options, such as turning off extraction pumps, can be carried out. Therefore, the need arises to identify proxy indicators that can be used to assess actual water quality and describe the development of microbial contamination during a high discharge event.

The overall aim of this study was to apply a proxy-based approach to assessing actual groundwater quality by combining proxy indicators for river-water infiltration and short-term microbial monitoring studies in an alluvial aquifer. The specific objectives were: (a) to compare two sampling approaches: a bi-weekly monitoring and an event-based sampling experiment; (b) to investigate the relationship between high discharge events in a river and faecal indicator bacterial densities in an adjoining aquifer; and (c) to identify proxy indicators best used as surrogates for bacterial contamination in an alluvial aquifer.

7.3 Methods

The study area is located in the lower Birs Valley of Northwest Switzerland (Fig.18). It is an alluvial system with a shallow, coarse, gravelly aquifer. The River Birs was canalized at the end of the 19th century and since then, the river bed has incised several meters into the former floodplain, from which it was subsequently disconnected. The River Birs runs 75 km through the Swiss Jura and joins the River Rhine in Basel, creating a catchment area of 866 km^2. The mean annual flow near the confluence zone of the tributary Birs with the Rhine is 15.4 m^3 s^{-1} and storm flows can reach up to 383 m^3 s^{-1}.

The aquifer material mainly consists of carbonate gravel, which is often well rounded, with variable sediment sorting, and few intercalations of clay or silt layers, resulting in a large variance in hydraulic properties. The carbonate gravel components of the aquifer are of Triassic to Jurassic ori-

Figure 18: Panel A and B are schematic maps showing the location of the 6 observation wells, the river and the drinking-water extraction wells. Panel C shows the installation depths of the instruments and filter stretches in the observation wells relative to the surface topology and aquifer bottom (where reached).

gin with hydraulic conductivities between 3 x 10^{-3} and 10^{-2} ms^{-1}. The aquifer bottom represents an aquitard and consists of Tertiary deposits of the *Elsässer Molasse*, generally with very low hydraulic conductivities (kf 10^{-8} to 10^{-5} ms^{-1}). The maximal depth to groundwater is 29 m and the thickness of the saturated zone varies between 0.6 m and 10 m. Most of the water is supplied to the aquifer from river infiltration, lateral inflow from the local catchment area, and intermittent artificial recharge. Eight active extraction wells in the study area (Fig.18) supply drinking water to six communities, approximately 51'000 people, amounting to 5 x 10^{6} m^{3} drinking water per year.

The experimental field site for river-groundwater interaction was set up in the gravel aquifer in the vicinity of the river bank (Fig.18). Groundwater head, temperature, turbidity and electrical conductivity were recorded in the groundwater observation wells and in the river. The depth of the filter stretches in the aquifer increase with increasing distance from the river (Fig. 18). Measurements were taken automatically every 30 minutes using an multiparameter instrument (YSI 600 OMS V2). To address the scale

and dynamics of the river-groundwater interaction, three observation wells (B1, B2 and B3) in the experimental field site were equipped to monitor additional quality parameters (all Endress+Hauser single parameter instruments): particle density (2-10μm) and SAC (spectral absorption coefficient at 254 nm). The observation well cluster B1, B2 and B3 was also equipped with a automated sampler cooled to 4 °C (Endress+Hauser ASP 2000).

7.3.1 Bi-weekly Monitoring

To determine the behaviour and background densities of the faecal indicator bacteria species *Escherichia coli*, two samples a week were taken between 9. June and 23. July 2009. Two groundwater observation wells and the river were sampled manually using the procedure described in subsection 7.3.2. Samples were taken from B1, B3 and the river.

7.3.2 Event-based Sampling

Six groundwater observation wells, one drinking-water extraction well (EW, cf. Fig. 18) and the river were sampled for two common faecal pathogenic indicator bacteria, namely *Escherichia coli* and *Enterococcus* sp..

Hygenic water quality was assessed for all samples by standard methods complying with the Swiss Food Manual (BAG, 2004). Sterilised glass bottles (1,5 l) were used in the automated sampler. After the sampling, the glass bottles were sterilised at 180 °C for 4 hours. The river, the observation wells in the alluvial forest (W1, W2) and observation well F2 were sampled manually using a PVC tube and wet-pit pump. Before each sample was taken, the tube was flushed to replace the volume of water in the observation well. The samples from the extraction well were taken from the untreated groundwater in the water chamber. All samples were stored in polyethylene bottles containing thiosulfate and kept cool at 4 °C before being analysed within 8 hours of collection.

To monitor *E. coli*, sample volumes of 1 ml, 10 ml and 100 ml, were filtered through cellulose nitrate membrane with a pore size of 0.45 μm (Microsart CN-Filter 11406Z-50 SC). In accordance with the method described

in the Swiss Food Manual (BAG, 2004), the filters were first incubated on Tryptone Soya Agar (TSA) (Oxoid CM0131) for two hours at $37\,°C$ and then transferred to the selective T.B.X agar (Oxoid CM0945). The plates were incubated for 18-24 hours at $44\,°C$. The blue pigmented colonies were counted and normalized to *E. coli* colony forming units (CFU) 100 ml $^{-1}$.

To monitor *Enterococcus* sp., the analysis was carried out according to the Swiss Food Manual (BAG, 2004). Three volumes of 1 ml, 10 ml and 100 ml were filtered through a 0.45 μm cellulose nitrate membrane filter (Microsart CN-Filter 11406Z-50 SC). The filters were placed onto Chromocult Agar and incubated for 24 hours at $37\,°C$. The brown pigmented colonies were counted and normalized to enterococci colony forming units (CFU) 100 ml $^{-1}$.

7.4 Results

7.4.1 Bi-weekly Monitoring

The monitoring period in June/July 2009 included average discharge situations and two clusters of high discharge events with over 80 m ^{3} s $^{-1}$ peak discharge (Fig. 19). Groundwater head was strongly influenced by the pumping regime in nearby extraction wells: daily head fluctuations due to the diurnal extraction regime reached up to 0.2 m. The bacterial sampling results showed clear indication of elevated *E. coli* in the river during the high discharge events (Fig. 19), whereby bacterial densities increased by up to an order of magnitude during the high discharge events. Densities in the river ranged between 600 and 2000 CFU 100 ml $^{-1}$ during average discharge situations and reached 13'700 to 24'100 CFU 100 ml $^{-1}$ during the high discharge events. The bacterial densities counted in groundwater samples taken in B3, the observation well closest to the river (10 m), were increased by up to two orders of magnitude following high discharge events, reaching up to 300 CFU 100 ml $^{-1}$. Bacterial densities in the samples taken from B1 (17 m from the river), were low and comparable to densities in B3 before the two major high discharge events during the monitoring period. The low bacterial densities found in both groundwater observation wells

(zero to three colony forming units (CFU) 100 ml $^{-1}$) were classed as typical background levels of *E. coli* for these observation wells.

7.4.2 Event-based Sampling

The event-based experiment was carried out during a high discharge event in May 2010. Although the event was only small compared to the range of discharge events in 2010 (less than 40 m ^{3} s $^{-1}$ peak discharge, half of the discharge recorded in the bi-weekly monitoring study), it was the first one after a longer period of little precipitation and low river stage. In total, 86 water samples were taken from six groundwater observation wells, the river and a drinking-water extraction well. Figure 20 shows the sampling frequency and river stage. During the event, the maximum precipitation intensity was 40.8 mm h $^{-1}$, however only for approximately 30 minutes.

The highest density of pathogenic indicator bacteria was found in the water sampled from the river. No faecal indicator bacteria were found in the water sampled from the extraction well and no *E. coli* and only two enterococci CFU were found in the observation well 95 m away from the river (F2). The *E. coli* and *Enterococcus* sp. densities in the groundwater decreased with increasing distance of the sampled well from the river (Fig. 21). The standard deviations (Fig. 21) also decreased with distance from the river. Of the other five groundwater observation wells, W1 yielded the lowest bacterial densities and the smallest standard deviation. In comparison to the results from the monitoring period, bacterial densities in B1 were elevated during the high discharge event.

E. coli and *Enterococcus* sp. densities were reduced by approximately two log scales in the first of the observation wells in the B-cluster (B3, 10 m from the river bank) and one log scale in the W-cluster (W3, 4 m from the river bank), compared to river densities. A further log scale reduction was observed between B3 and B1 (7 m distance) and two log scale reductions between W3 and W1 (5 m distance) (Fig. 21).

Figure 22 shows the temporal development of *E. coli* densities during the sampling period. The highest densities were observed at the beginning
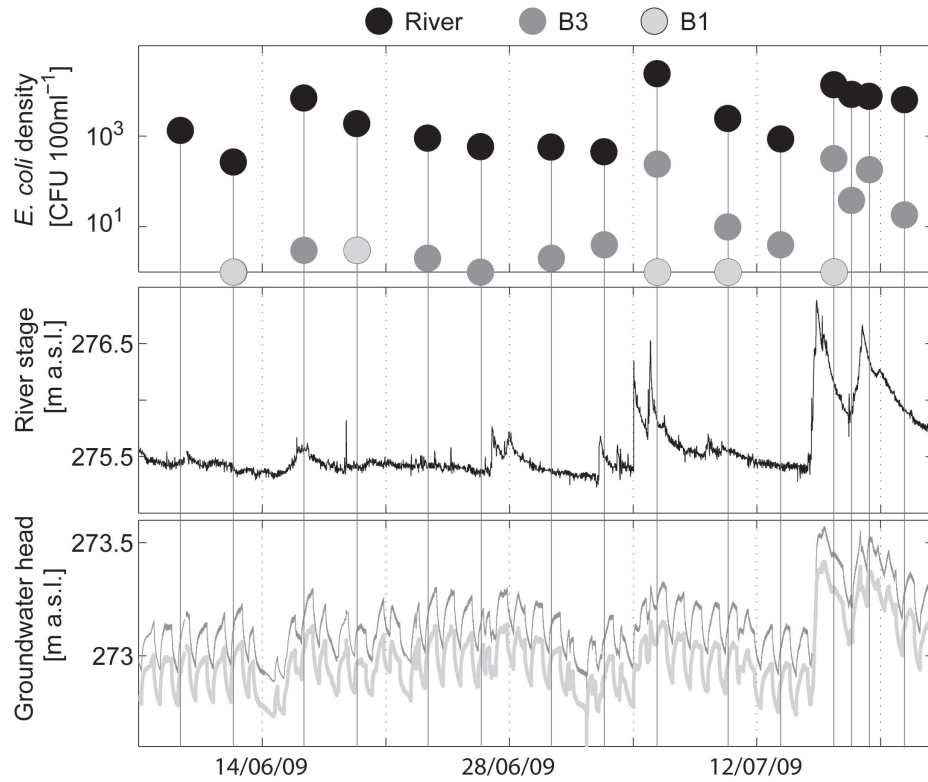
Figure 19: The top graph shows the *Escherichia coli* density results from the bi-weekly monitoring in the river and two groundwater observation wells in June/July 2009. The middle graph shows the river stage with the high discharge events in the last third of the sampling period. The groundwater head was heavily characterized by diurnal groundwater pumping regime in the nearby drinking water extraction wells (bottom graph).
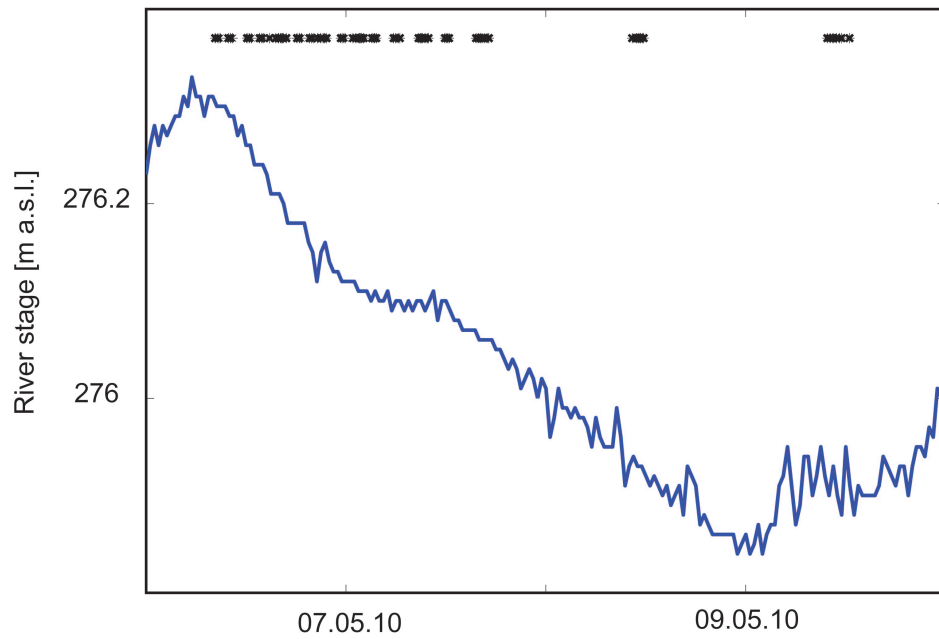
Figure 20: The black stars represent water samples from the river and groundwater during the event-based sampling period in May 2010. The water samples were taken at different intervalls, depending on the river stage. The highest sampling frequency coincided with the initial decrease in the river hydrograph.
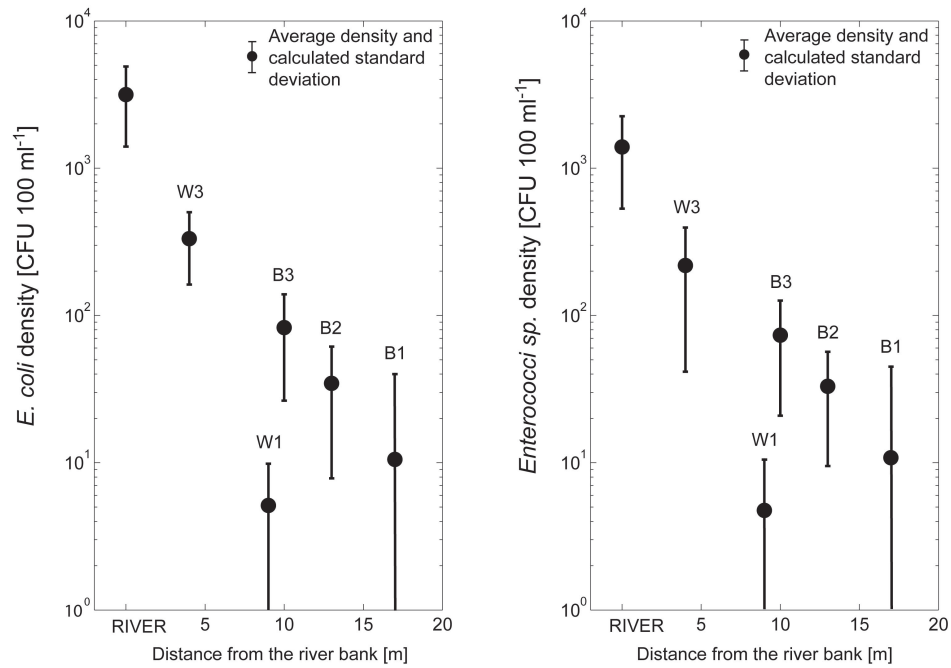
Figure 21: Faecal indicator bacteria (*E. coli* and *Enterococcus* sp.) densities in river- and groundwater. Average and standard deviation of samples decrease with increasing distance from the river. The overall reduction of densities amounts to approximately three log scales within 20m of flow path in the aquifer.
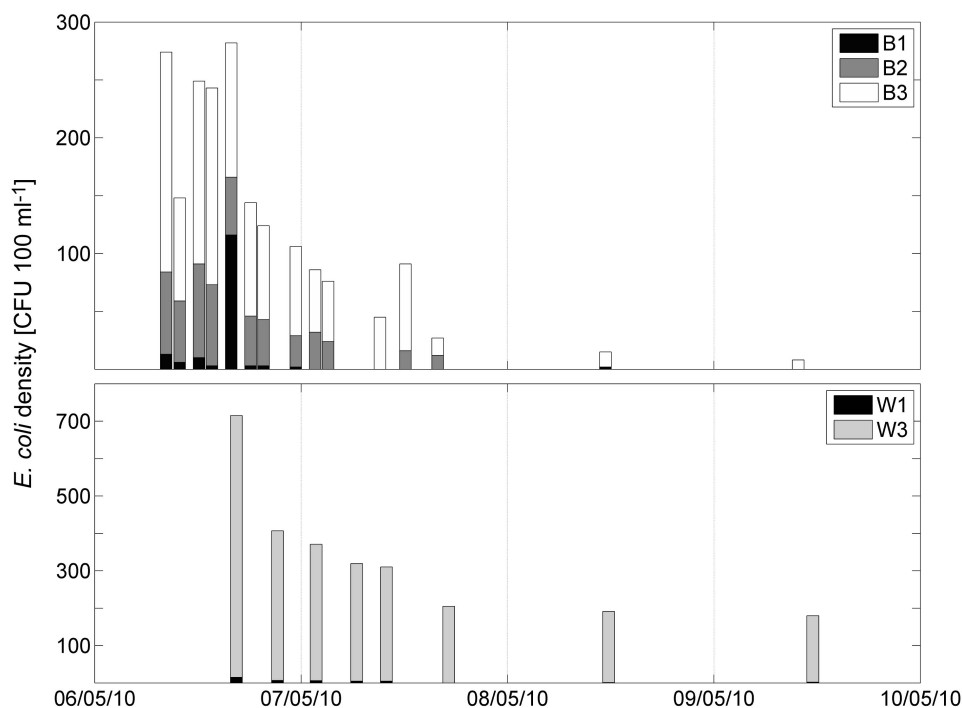
Figure 22: Temporal development of *E. coli* and *Enterococcus* sp. densities in the groundwater during the sampling period following precipitation and increased river discharge (6. - 9. May 2010).

of the sampling period, the lowest at the end. The densities in the B-cluster (B1, B2 and B3) were reduced to background levels, as observed during the monitoring period in June/July 2009, within three days, while densities of *E. coli* sampled from W3 remained elevated at 178 CFU 100 ml^{-1}. Background densities in B1, according to densities found during the bi-weekly monitoring, were reached approximately one day after sampling commenced. *E. coli* densities in the river also remained elevated (3200 CFU 100 ml^{-1}) compared to background densities observed in June/July 2009.

Six time-series of continuously measured parameters were used for the indicator analysis during the event-based sampling: groundwater head, temperature, electrical conductivity, turbidity, particle density, spectral absorption coefficient (SAC). All time-series used to assess the suitability of the

different parameters to serve as proxy-indicators for faecal contamination of groundwater were recorded in observation well B3 (Fig. 23). Extraction of groundwater was only stopped in the two observation wells closest to the river. The three peaks observed in groundwater head are characteristic of the pumping regime in the futher six groundwater extraction wells. Any increase in groundwater head was partly masked by the regular pumping regime, which resulted in approximately 0.15 m fluctuations in groundwater head and no clear peak associated with the elevated river stage. The changes measured in the other parameters during the high discharge event in the river were more easily detected: 1) groundwater temperature ranged between 10 and 12 °C and increased gradually after a small decline at the beginning of the sampling period, 2) electrical conductivity ranged between 480 and 570 μS cm^{-1}, as river water electrical conductivity was lower than groundwater el. conductivity, leading to a characteristic decrease indicating the arrival of rapidly infiltrating river water at the groundwater observation well, 3) turbidity increased slightly (from 0.55 to 0.75 NTU), 4) particle density increased four-fold, the highest densities were recorded towards the end of the sampling period, and 5) SAC increased slightly and showed a similar behaviour to turbidity. Peak densities of *E. coli* and *Enterococcus* sp. occured before the peaks in SAC and turbidity, and during the decreasing phase of the electrical conductivity (Fig. 23).

7.5   Discussion

Bi-weekly monitoring vs event-based sampling During both sampling periods, the faecal indicator bacteria density in the river and groundwater were one to two orders of magnitude greater during and shortly after a high discharge event and subsequently decreased. The event-based sampling took place during a minor event with less than 40 m^3 s^{-1} peak discharge, but after a long period (one month) without discharge-effective precipitation. Without this natural source of recharge, both the river stage and the groundwater table were relatively low in the study area. Nichols et al. (2009) associated these conditions - low level rainfall for a few weeks prior to excess

73
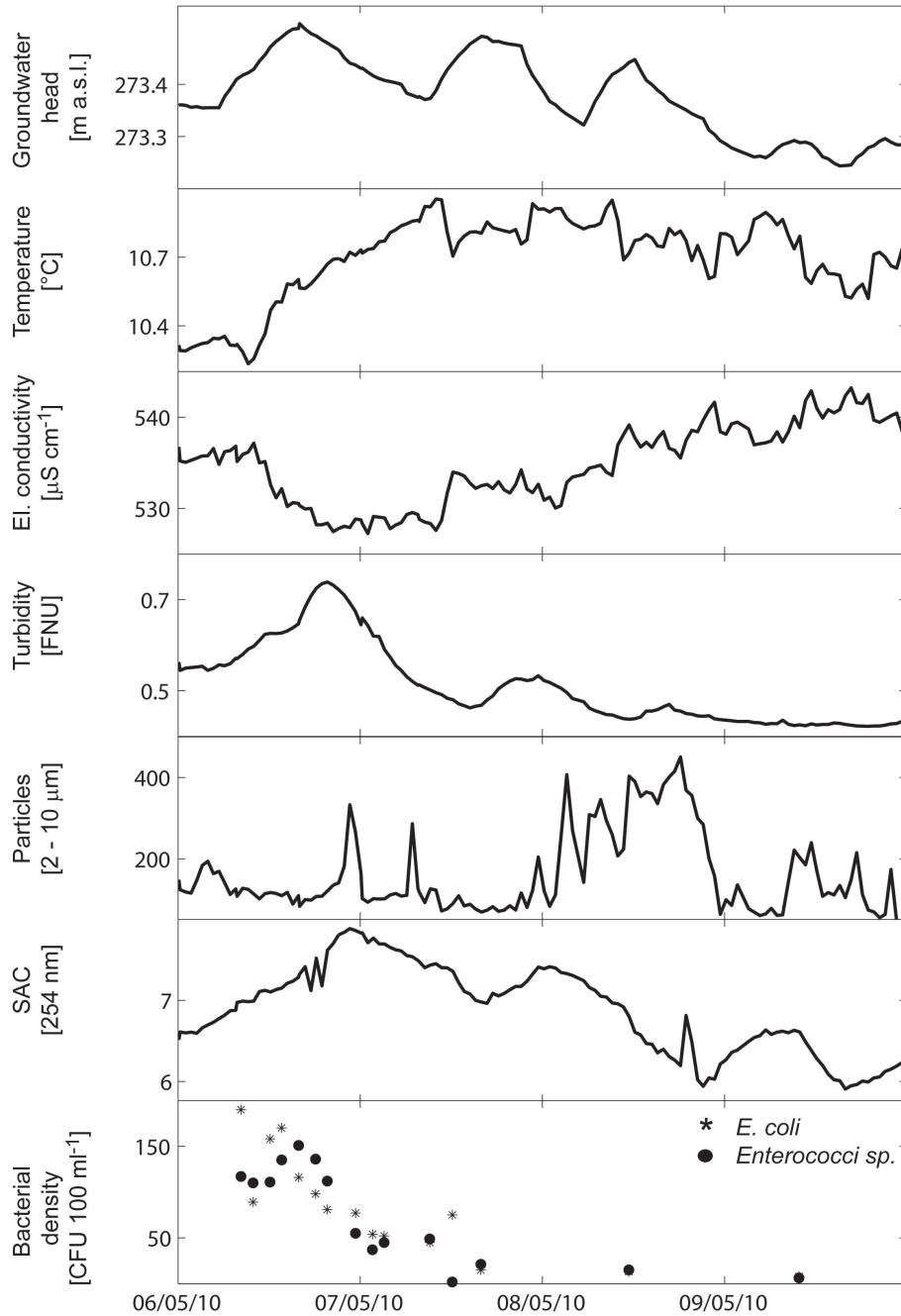
Figure 23: Comparison of *E. coli* and *Enterococcus* sp. densities with time-series of continually monitored groundwater parameters (turbidity, SAC, particle density, electrical conductivity, temperature and river stage) recorded in observation well B3.

rainfall - with an increase in waterborne disease outbreaks. While previous studies, e.g. Nichols et al. (2009); Grisey et al. (2010), describe the connection between rainfall and elevated levels of faecal indicator bacteria, the results from the event-based sampling focus on a high resolution analysis of faecal indicator bacteria densities in groundwater following a precipitation-induced increase in river stage.

Although the *E. coli* densities in the river were higher during the bi-weekly monitoring period than they were during the event-based experiment and peak river-discharge was double during the bi-weekly sampling, the bacterial densities observed in the groundwater samples taken during the bi-weekly monitoring study were comparable for B3 and low for B1. This discrepancy indicates that the bi-weekly monitoring was not able to detect high bacterial densities present at the beginning of an event, thus emphasizing the strong 'hit-and-miss' characteristic of this approach. The high-frequency, event-based experiment in May 2010 provided more information on the development of bacterial densities during a high discharge event and is therefore better suited for studying the relationship between river-groundwater interaction and faecal indicator bacteria densities, as well as for detecting potential contamination due to infiltrating river water during a high discharge event.

Temporal development during high discharge events The highest *E. coli* densities can be expected during the rising limb of a river discharge hydrograph, followed by a rapid return to concentrations similar to those prior to the event (Wilkinson et al., 1995). In addition to this 'first flush' phenomenon, i.e. high densities early in the high discharge event, the highest connectivity between the river and the adjoining aquifer (greatest conductance of the river bed) occurs during the rising limb and immediately after the peak river stage (Mutiti and Levy, 2010). In agreement, the highest bacterial densities in this study were found in the river and at the beginning of the event-based sampling experiment in May 2010. After 1-2 days, the faecal indicator bacteria densities in the B-cluster and W1 decreased to low, or background levels, as observed during the monitoring period in June/July 2009. The sustained higher densitites towards the end of the event-based

sampling in W3, suggest that the connectivity between the river and the aquifer remained elevated in the proximity of this observation well. Although *E. coli* and *Enterococcus* sp. can survive for a limited period of time in groundwater, their viability decreases by half in under 24 h (McFeters et al., 1974), thus the high bacterial densities towards the end of the sampling period are likely to orginate from 'fresh' river water. In contrast, the bacterial densities observed in samples taken from B1, B2 and B3 suggest the exchange is dependent on river stage elevation, rather than connectivity between the river and the groundwater as determined by clogging mechanisms. Thus the decrease observed in bacterial densities in the B-cluster is attributable to dispersion and die-off (including loss of viability) of bacterial cells.

The two clusters (B and W) showed different inactivation rates, i.e. removal or loss of faecal indicator bacteria from the water column (Gronewold et al., 2011). The greatest inactivation of *E. coli* and *Enterococcus* sp. was observed between the river and the observation wells closest to the river (B3 and W3), underlining the importance of the river bed and bank in removing a large amount of microbial load from the infiltrating river water, the mechanics characterizing river bank filtration systems (Dizer et al., 2004). A similar inactivation rate observed for W1, is likely to originate from the hydrogeological properties of the aquifer. An approximately 2 m thick layer of overbank deposits, dominated by fine fraction sediments (<0.063 mm) with variable gravel content, separated the upper and lower aquifers in the area around the W-cluster. The filter stretches of W1 and W3, representing the depth in the aquifer from where the water sample originated, were thus separated by an additional barrier to the faecal indicator bacteria. Fine sediments have a higher porosity than sand and an increased clay content can provide greater pore continuity, but the hydraulic conductivity is smaller (Unc and Goss, 2003). Transport of particles or bacteria can be hindered by e.g. pore size exclusion or clogging. Although both faecal indicator bacteria species can form biofilms as a means of protection, enterococci show a greater affinitiy to form biofilms than *E. coli* do (Fisher and Phillips, 2009; Reisner et al., 2006). The presence of biofilms would further decrease the

pore volume available for transport and enhance the filter efficiency. Subsequent detachment of biofilms would lead to peaks in bacterial densities, which were not observed in this study. The aquifer composition in the proximity of B1, B2 and B3 was more homogenous, with a lower fine sediment fraction than around W1 and W3. Based on the higher hydraulic conductivities and the lower removal rates in the proximity of B1, B2 and B3 pose a threat to the nearby drinking water extraction wells. Besides the differences observed in overall hydraulic conductivity between the B and W clusters, large local variances in hydraulic conductivity, as observed in the W-cluster, can give rise to fast flow paths (*preferential*), along which microorganisms can be transported over great distances. Although no bacterial colonies were found in the extraction well sampled, a greater or longer high discharge event may cause problems, especially considering potential preferential flow paths. Further studies should therefore include a range of hydrological situations, especially high discharge situations after long periods of dry weather.

Proxy indicators Besides the assessment of actual microbial contamination based on faecal indicator bacteria, the suitability of different parameters to serve as proxy indicators for microbial contamination was investigated. While many of the continuously measured parameters showed a response to the high discharge event in the river, the timing differed, i.e. the rate of increase or time to peak measurements. Peaks in some of the parameters frequently associated with microbial transport, e.g. SAC (Stadler et al., 2010), occured during the decreasing limb of microbial density. Particle densities were highest towards the end of the event-based sampling period and therefore not applicable as proxy indicators for potential microbial contamination in groundwater. The size range used in this study may have influenced the applicability, as it covers the top of the range of particle sizes useful as indicators for bacterial densities defined in other studies (Pronk et al., 2007; Goldscheider et al., 2010). Electrical conductivity and temperature are used as indicators for rapidly infiltrating river water (Sheets et al., 2002; Keery et al., 2007). Although both parameters showed a reaction to the high discharge event, a decrease in electrical conductivity and an increase in temperature, the development of these two parameters could

not be matched to bacterial density. In addition to the individual behaviour during the high discharge event, the reaction of each parameter is dependent on external factors, e.g. temperature shows seasonal variation and a strong relationship with river-water temperature, while electrical conductivity is related to physico-chemical solution and dilution rates. The occurence of pathogenic microorganisms also shows a non-linear relationship with river discharge and groundwater flow, e.g. Wilkinson et al. (1995) found the highest concentration of *E. coli* during the rising limb of a river discharge hydrograph and then a rapid return to concentrations similar to those prior to the event. No individual, continuously measured parameter was suitable for use as a proxy indicator for faecal indicator bacteria, which reflects the high-dimensionality of river-groundwater interaction, both in time and space. However, our investigations show that the onset of change in the time-series of each parameter showing a reaction to the high discharge event, may be used in a proxy analysis. A principal component analysis (PCA) based on variance observed in a multivariate data set (measured parameters and observation wells), can shed light on the use of individual observation wells and parameters to identify rapidly infiltrating river water (Page et al., 2011). Thus, the dominant factors influencing groundwater quality in the study area and individual observation wells can help assess and quantify the potential for microbial contamination due to river-water infiltration. While the PCA gives an indication of zonation of riverine influence and with it susceptibility to riverine contamination, it is a static linear method. A non-static and nonlinear proxy analysis can identify shifts in the time-series that indicate e.g. infiltrating river water, which would not be identifiable based on the analysis of an individual parameter. Multiparametric time-series are able to capture system state and therewith actual groundwater quality, but often result in large data sets with a high dimensionality, which are difficult to interpret without additional analysis. Such complex, multivariate data sets can be simplified and inherent patterns identified using a combination of a self-organizing map and Sammon's projection (SOM-SM), giving rise to a neural network-based, nonlinear method of statistical data analysis (Lischeid, 2009). This methodology has frequently be used in fault detection

systems of industrial processes and recently in a water-distribution network (Mustonen et al., 2008). Further research in this area includes using this data set to test a multivariate SOM-SM approach to identify critical system states, where groundwater quality is degraded due to infiltrating river water (Page and Simovic (2011), Page et al., in prep.). The application of a SOM-SM method thereby reduces the complexity of a multivariate data set and allows real-time assessment of groundwater quality. The analysis can be used to detect shifts in groundwater quality induced by infiltrating river water, as it is based on variance in the measured time-series.

7.6 Conclusion

The aim of this study was to assess actual groundwater quality using two proxy-based approaches. The first approach was based on faecal indicator bacteria, while the second focussed on continuous time-series of groundwater measurements.

The density of faecal indicator bacteria, as a proxy indicator for pathogenic faecal contamination, was investigated using two sampling strategies. The event-based sampling showed greater detail of the behaviour of *E. coli* and *Enterococcus* sp. densities in groundwater during a high discharge event than the bi-weekly monitoring did. Bacterial densities were highest at the beginning of the event and decreased to background levels within 1-2 days. Low-resolution monitoring following high discharge events, as is often the case in practise, is therefore likely to miss peak bacterial densities in some, or all of the sampling points. Furthermore, high bacterial densities early in events may pose a considerable threat to drinking-water quality, if the water suppliers are not informed in time to avoid extracting freshly infiltrating river water. In addition to transient bacterial densities in groundwater, the heterogeneity of hydraulic conductivities leads to complex flow and transport patterns, which cannot be interpreted based on only few microbial samples, hydrological situations, or observation well time-series. In addition to the event sampled, future event-based sampling experiments will therefore focus on high-resolution early-event sampling, i.e. during the rising limb of the

river hydrograph, and include pre-event investigations of bacterial densities.

The second proxy-analysis approach under investigation was based on the assessment of continuous time-series of six parameters. While the parameters all showed a reaction to infiltrating river water, none matched the behaviour of bacterial densities. Peak measurements (min. or max.) of individual parameters did also not correspond well with the other parameters. The authors thus propose a combined analysis of a set of indicator parameters, such as electrical conductivity, temperature and SAC, to detect rapidly infiltrating river water. A SOM-SM analysis of a multivariate data set, including multiple observation wells and multiple parameters, can identify shifts in groundwater quality based on observed variance in the data set and provide the basis for an early warning system for the water supplier.

The connection between precipitation, high discharge events in the river and increased faecal indicator bacteria densities is well-established. However, investigations concerning the dynamics between the onset of increased river-water infiltration and endangerment of drinking-water wells require high-frequency event-based microbial sampling to highlight the transient character of bacterial density during a high discharge event. This study showed the importance of sampling frequency and knowledge of aquifer composition when conducting microbial sampling studies. The proposed methodology based on high-resolution sampling in combination with a multiparametric monitoring network and multivariate analysis, removes the focus from a threshold-oriented, towards a process-oriented approach to groundwater-quality management for drinking-water extraction wells near rivers.

7.7 Acknowledgements

8 Multivariate Analysis of Groundwater-Quality Time-Series

Page R. M., Lischeid G., Huggenberger P.

8.1 Abstract

Groundwater extracted from alluvial aquifers close to rivers is vulnerable to faecal contamination due to infiltrating river water. Infiltration is often increased during high discharge events, when the levels of waterborne pathogens are also increased. Water suppliers with low-level treatment thus rely on procedures and information on system state to manage the resource and maintain drinking-water quality. The complexity of the processes at the river-groundwater interface needs to simplified to extract the essential information required for management purposes. In this study, a combination of self-organizing and Sammon's mapping methods (SOM-SM) was used as a proxy analysis of a multiparametric time-series to detect critical system states, i.e. where contamination of the extraction wells is imminent. Groundwater head, temperature and electrical conductivity time-series from six groundwater observation wells and three time periods were analysed using the SOM-SM proxy method. The critical system states could then be identified using independent measurements. The independent parameters used were spectral absorption coefficient and particle density, both strongly associated with river water and faecal bacteria contamination. The critical system states during all three time periods were located in the same area of the SOM-SM projection. This approach can thus be used to identify critical system states and provide critical control limits that can be integrated into a groundwater-management approach using real-time continuous measurements of groundwater head, temperature and electrical conductivity.

Keywords: Groundwater; time-series analysis; self-organizing map; Sammon's projection; drinking water

8.2   Introduction

Alluvial groundwater is an important source of drinking water in many countries, providing approximately 50 percent of the global drinking-water supply, as well as a significant part of water required for industrial and agricultural purposes (Zektser and Everett, 2004). Many drinking water extraction wells are located in the proximity of rivers, drawing water from high-permeability alluvial deposits in the river valley. A diverse riparian ecosystem is able to develop alongside groundwater extraction well-fields without competing for space or water, enabling multiple users to coexist (Huggenberger et al., 1998). Under average hydrological situations, the filter capacity of the river bank and alluvial aquifer is generally sufficient to remove river-borne contamination before it reaches groundwater extraction wells (Taylor et al., 2004; Dash et al., 2010). However, during high discharge events, the microbial load in the river water is elevated and the connectivity between the river and aquifer is increased (Wilkinson et al., 2006; McKergow and Davies-Colley, 2010; Mutiti and Levy, 2010; ten Veldhuis et al., 2010). As a consequence, high discharge events lead to an influx of microorganisms from surface water, many of which are pathogenic. Microbial contamination of groundwater poses a major concern for municipal water suppliers with low treatment levels (Kukkula et al., 1997; Sheets et al., 2002; Regli et al., 2003; Nichols et al., 2009). The failure to detect contaminated drinking water can cause severe outbreaks of disease (Kukkula et al., 1997; Joerin et al., 2010).

Faecal indicator bacteria (FIB) are used as proxy indicators for many waterborne pathogens, which require complex and lengthy analytical procedures (Borchardt et al., 2004; Brookes et al., 2005; Pronk et al., 2007). Although there have been recent advancements in real-time monitoring methods for microbial water quality, they are not yet readily available for many applications (Martinez et al., 2010). As even sampling for faecal indicator bacteria is both cost and time intensive, proxy indicators for river-water infiltration are used in addition to short-term microbial monitoring studies. Proxy indicators are parameters that can be easily measured and the relationship with the variable of interest, e.g. infiltrating river water, is

well-established. Proxy indicators for river-water infiltration are usually physical parameters, such as electrical conductivity or temperature, which can be used to quantify water exchange across the river-groundwater interface (Conant, 2004; Keery et al., 2007; Vogt et al., 2010). Such univariate proxy analyses rely on in-depth knowledge of the relationship between the variable of interest, e.g. river water infiltration, and the proxy indicator, e.g. temperature, and consistency of this relationship in time and space. However, hydrogeological heterogeneity (e.g. river bed morphology, biogenic stabilization of sediments, hydraulic conductivity of aquifer material) and hydrological variability (e.g. seasonality, precipitation and water table levels) result in a complex pattern of river-groundwater interaction, both in time and space (Arntzen et al., 2006; Gerbersdorf et al., 2008). The resulting heterogenous pattern makes predicting actual microbial water quality based on univariate proxy analysis a complex task for drinking-water extraction managers.

To reduce the complexity and aid safeguarding drinking-water quality, water suppliers make use of decision support and management tools, such as water safety plans. Originating in the food safety sector, Hazard Analysis and Critical Control Point (HACCP) procedures are an integral part of many drinking-water safety plans (Havelaar, 1994; Hamilton et al., 2006). One of the major tasks in applying a HACCP approach is to establish critical control points (CCPs). Process steps in drinking-water production, including extraction and distribution, can be used as CCPs if they comply with the *Codex Alimentarius* definition "A step at which control can be applied and is essential to prevent or eliminate a food safety hazard or reduce it to an acceptable level" (Codex, 1997). For example, disinfection is a CCP where the residual concentration of disinfectant can be monitored. For each CCP, critical control limits (CCL), e.g. maximum residual concentration of disinfectant, have to be defined and compliance monitored (Codex, 1997). The heterogeneity of system states makes the definition of CCLs more difficult for groundwater extraction. However, simple precautionary steps, such as extraction management, can prevent contaminated river water entering the further process steps in drinking-water production. Existing measures,

such as conventional well-head protection areas (WHPA) fall short of providing adequate protection, as many of the parameters upon which they are based, such as travel time of contaminants, are insufficiently characterized or transient (Frind et al., 2006).

In order to bypass the problem of changing boundary conditions for the hazard assessment, a multivariate proxy analysis based on actual measurements can be used to define CCLs for transient groundwater extraction. The challenge of identifying the point in time when management actions need to be taken to avoid contamination of extraction wells, is enhanced by the use of multiparametric time-series. For a multiparametric time-series data matrix to be useful in the operation of drinking-water extraction wells, the complexity has to be reduced to provide the decision-maker with a clear indication of potential hazard.

Each high discharge event may have a different effect on the measured time-series, so that a threshold approach cannot be applied. While a multivariate analysis can overcome the problem of individual responses of parameters strongly affected by, for example, seasonality, multiple time-series are difficult to interpret simultaneously, especially when rapid decisions need to be made to avoid contamination of the extraction wells.

The approach applied in this study is based on multivariate time-series analysis and a reduction of dimensions, which brings about a reduction of complexity so that the approach can be incorporated into a water-supply decision support system as a CCP. The CCP approach is based on a combination of the self-organizing map (SOM) and Sammon's mapping (SM) for the analysis of time-series of proxy indicator parameters (groundwater head, temperature and electrical conductivity). SOMs, also known as Kohonen or feature maps, are self-organizing neural network algorithms that are used for pattern recognition and the visualization of high-dimensional data sets in a low-dimensional display (Kohonen, 2001). Due to their proficiency in anomaly detection, self-organizing maps have a wide range of applications from fault detection in industrial processes (Kohonen et al., 1996; Fuertes et al., 2010), to monitoring water quality in a drinking-water distribution network (Mustonen et al., 2008). Sammon's projections are based on a non-

linear mapping algorithm that preserves inherent structures when projecting high-dimensional data sets to a low-dimensional space (Sammon, 1969). In comparison to SOMs, Sammon's maps are able to visualize cluster structures (including temporal aspects) in the data with greater precision (Mustonen et al., 2008). To decrease the computational load, the weight vectors of the SOMs can be used as a starting point for the Sammon's mapping, the combination is referred to as SOM-SM (Lischeid, 2009).

The overall aim of this study is to evaluate the multivariate SOM-SM approach as to applicability to define critical control limits for groundwater extraction in an alluvial aquifer dominated by heterogenous river-groundwater interaction. The objectives are to test the multivariate statistical analysis on frequently monitored parameters (groundwater head, electrical conductivity and temperature) and to identify critical system states that represent imminent contamination of riverine-groundwater extraction wells.

8.3 Experimental Field Site, Monitoring System and Data Management

The experimental field site, approximately 0.9 km^2, is located in the lower Birs Valley in North-west Switzerland (Fig. 24). It is an alluvial system with a shallow, coarse, gravelly aquifer. The River Birs was canalized at the end of the 19th century and since then, the river bed has incised several metres into the former floodplain, from which it was subsequently disconnected. The Birs runs 75 km through the Swiss Jura and joins the River Rhine in Basel, creating a catchment area of 866 km^2. The mean annual flow near the confluence of the tributary and the Rhine is 15.4 m^3 s^{-1} and storm flows can reach up to 383 m^3 s^{-1}.

The aquifer material mainly consists of carbonate gravel, often well-rounded, with variable sediment sorting, and few intercalations of clay or silt layers, resulting in several orders of magnitude variance in hydraulic properties. The carbonate gravel components of the aquifer are of Triassic to Jurassic origin with hydraulic conductivities between 3 x 10^{-3} and 10^{-2} ms^{-1}. The aquifer bottom is defined by an aquitard consisting of Tertiary
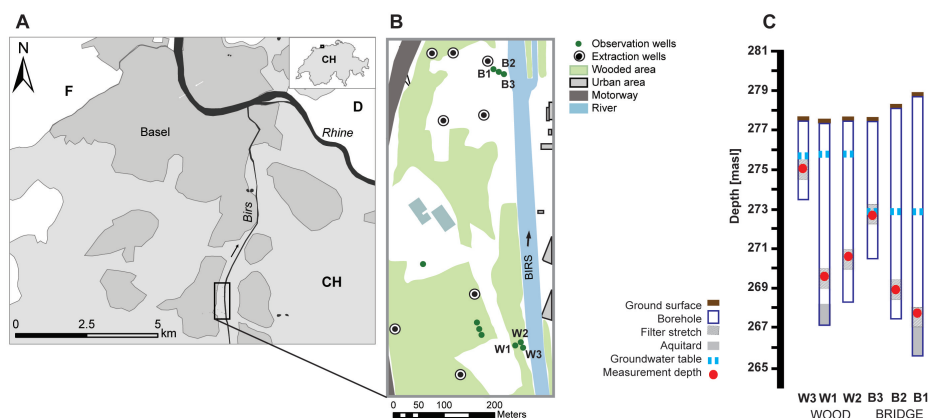
Figure 24: Panels A and B are schematic maps showing the locations of the six observation wells relative to the river and the groundwater extraction wells. Panel C shows the depth of the filter stretches and instrument installation in the observation wells relative to the surface topology and aquifer bottom (where reached).

deposits of the *Elsässer Molasse*, generally with very low hydraulic conductivities (kf 10^{-8} to 10^{-5} ms^{-1}). The maximal depth to groundwater is 29 m and the thickness of the saturated zone varies between 0.6 m and 10 m. Most of the water is supplied to the aquifer by river water infiltration, lateral inflow from the local catchment area, and intermittent artificial recharge. Eight active extraction wells in the study area supply water to six communities, approximately 51'000 people, amounting to 5 x 10^{6} m^{3} drinking water per year. The experimental field site for river-groundwater interaction was set up in the gravel aquifer in the vicinity of the river bank (Fig. 24).

The time-series for the proxy analysis were recorded in six groundwater observation wells located within 100 m of groundwater extraction wells susceptible to contamination due to river water infiltration (Fig. 24). The location of the observation wells was based on regional and local groundwater flow modelling results (Affolter et al., 2010). Groundwater head distribution was calculated from pressure measurements; temperature and electrical

conductivity were measured directly. Time-series of spectral absorption co-efficient (SAC, measured at 254 nm) and particle density (2 - 10 μm) were available only for three observation wells (B1, B2 and B3). The time-series were checked for outliers and errors due to instrument maintenance. The measurement resolution used in the analysis was 0.5 h.

Three time periods were used to study the behaviour during high discharge events as a basis for system understanding and pattern recognition. The criteria for selecting the events were based on the integrity of the data set (no missing values) and the magnitude of the event (river stage). The dimension of the resulting data matrix was 2388 x 17.

8.4   Data Analysis

The time-series were normalized separately for each time period used in the analysis (Jan. 09, Dec. 09 / Jan. 10 and May 10). The normalized time-series were combined to form one data set, which was used for the SOM analysis and Sammon's mapping. The SOMs were calculated using the SOM Toolbox, developed at the Laboratory of Information and Computer Science in the Helsinki University of Technology. Both steps of the analysis (SOM and SM) were calculated with MATLAB (Mathworks).

Self-Organizing Maps are based on vector quantization (VQ). This methodology allows an alternative representation of a complex data set (x, original data) by means of reference vectors (mi), also referred to as codebook, model or weight vectors. The reference vectors share the probability distribution function with the original data set, permitting a decrease in dimensionality and complexity without losing essential detail concerning distribution of the original measurements (x). The reference vectors can be linearly initiated, based on the two eigenvectors with the largest eigenvalues, or randomly initiated. Linear initiation was used in this study. They are then associated with nodes (units) in a regular grid, where the dimension is based on the ratio between the two largest eigenvalues (Kohonen, 2001). The grid in this study had a sheet form constructed of 18 x 13 nodes.

The best-matching unit (c, BMU) is found by calculating the Euclidean

distance between the reference vectors associated with each node and the input vector. The BMU describes the unit, where the distance between the input data vector and the reference vector is minimal (Eq. 2). The implementation of Eq. 2 can vary (Vesanto and Alhoniemi, 2000).

$$c = arg \min_{i}\{\|x - mi\|\} \tag{2}$$

The BMU is used to pinpoint positions with minimal distance between the input data set and the reference vectors associated with the grid. The following step is called a relaxation process of the grid, whereby the nodes in the proximity of the BMU are modified based on a neighbourhood function. The neighbourhood function (h_{ci}) is a smoothing kernel, e.g. a Gaussian function (Eq. 3).

$$h_{ci}(t) = \alpha(t) \cdot \exp\left(-\frac{d_{ci}^2}{2\sigma^2(t)}\right) \tag{3}$$

Thus, based on a Gaussian function, the influence of a node identified as a BMU decreases with increasing distance between c (coordinates of the BMU) and node i (d_{ci}). $\sigma(t)$ in Eq. 3 defines the width of the kernel (radius of the topological neighbourhood) and usually decreases monotonically in time (Kohonen, 2001). $\alpha(t)$ is a scalar-valued learning-rate factor and also decreases with time ($0 < \alpha < 1$). The updating and modification of the reference vectors ($m_i(t+1)$) in the grid is a function of the previous state ($m_i(t)$, the neighbourhood function ($h_{ci}(t)$) and the input data vector ($x(t)$) (Eq. 4).

$$m_i(t + 1) = m_i(t) + h_{ci}(t)[x(t) - m_i(t)] \tag{4}$$

where t=0,1,2,... is a discrete-time coordinate (steps). The result is a set of codebook vectors, an attractor, which can be approximated to a 2-D display of the probability density function of the input data vector.

To emphasize the temporal resolution and visualization of variation in the original data set, especially over time, the output of the SOMs were further analysed using a Sammon's mapping algorithm. The best-matching

units derived from the SOM analysis were used to calculate coordinates that formed the basis for the projection space in the Sammon's mapping (SM) algorithm.

Sammon's Mapping is a nonlinear mapping algorithm. It is based on multidimensional scaling (MDS), whereby the relative distances in the data are preserved to the greatest extent possible (Sammon, 1969). The error (E, Eq. 5) expresses how well the present configuration of n points in the projection space fits the N points in the input space. The objective of the optimization problem (Eq. 5) is to minimize the Sammon's Stress (E) by using a gradient descent approach based on interpoint distances in the input space ($d_{ij}*$) and in the projection space (d_{ij}).

$$E = \frac{1}{\sum\limits_{i<j} d_{ij}^*} \sum\limits_{i<j}^{N} \frac{[d_{ij}^* - dij]^2}{d_{ij}^*} \tag{5}$$

To prevent the problematic of local minima, *a priori* knowledge can be used to initialize the map. As a starting point, the codebook vectors derived from the SOMs can be used (Mustonen et al., 2008; Lischeid, 2009). The CPU time, which can be significant when using Euclidean distances to minimize stress (Lerner et al., 1998; Kolehmainen et al., 2003), can also be decreased by using the coordinates of the best matching nodes (c) calculated from the SOM. Henceforth, the combination of using the SOM codebook vectors to initiate the Sammon's projection is referred to as SOM-SM. The squared Pearson correlation coefficient (*R2*, coefficient of determination) gives an indication of the truthfulness of the SOM-SM projection by measuring the correlation between the Euclidean distances in the input matrix and those of the projection matrix. The correlation coefficient thus becomes a measure of representative strength of the projection in portraying the input data.

8.5 Results

8.5.1 Time-series

Three time periods were used for the analysis: 17. January - 03. February 2009 (time period Jan. '09), 23. December 2009 - 08. January 2010 (time period Dec. '09/Jan. '10) and 03. - 20. May 2010 (time period May '10). All three time periods included at least one high discharge event. Time-series from two observation wells (B1 and W3) are shown in Fig. 25. They are located at opposite ends of the study area (B1: north; W3: south) and represent the measurements closest to the river (W3) and those furthest away (B1), as well as the top of the aquifer (W3) and the lowest level of the aquifer (B1). The two observation wells B1 and W3 are representative of the range of values and behaviour displayed by all six observation wells used in the analysis.

The high discharge events are recognizable as increases in groundwater head. The smaller, daily fluctuating signal observed in groundwater head can be attributed to the pumping regime in the groundwater extraction wells, which was overriden or stopped during the high discharge events. The difference in absolute head level measured in B1 and W3, about 4 m, is due to the location of the observation wells: they are approximately 600 m apart from each other, giving rise to a hydraulic gradient of 0.0067 over the study area. The middle line of graphs in Fig. 25 shows the temperature time-series recorded in B1 and W3. While the temperature in B1 remained relatively constant within and between the three time periods, the temperature recorded in W3 showed greater variation between time periods than within them. Nevertheless, small fluctuations (0.75 - 1.5 $°$C) were observed during high discharge events. Electrical conductivity in B1 showed little variation, both between and within time periods Jan. '09 and Dec. '09/Jan. '10 (mean 500 μS cm^{-1}). Although electrical conductivity measurements in B1 also showed little variation in May '10, the mean was slightly higher (520 μS cm^{-1}). The electrical conductivity measurements in W3 showed strong decreases during high discharge events, sometimes accompanied by initial increases prior to dropping by up to 120 μS cm^{-1}. Each increase

in groundwater head not induced by groundwater extraction was associated with a decrease in electrical conductivity.
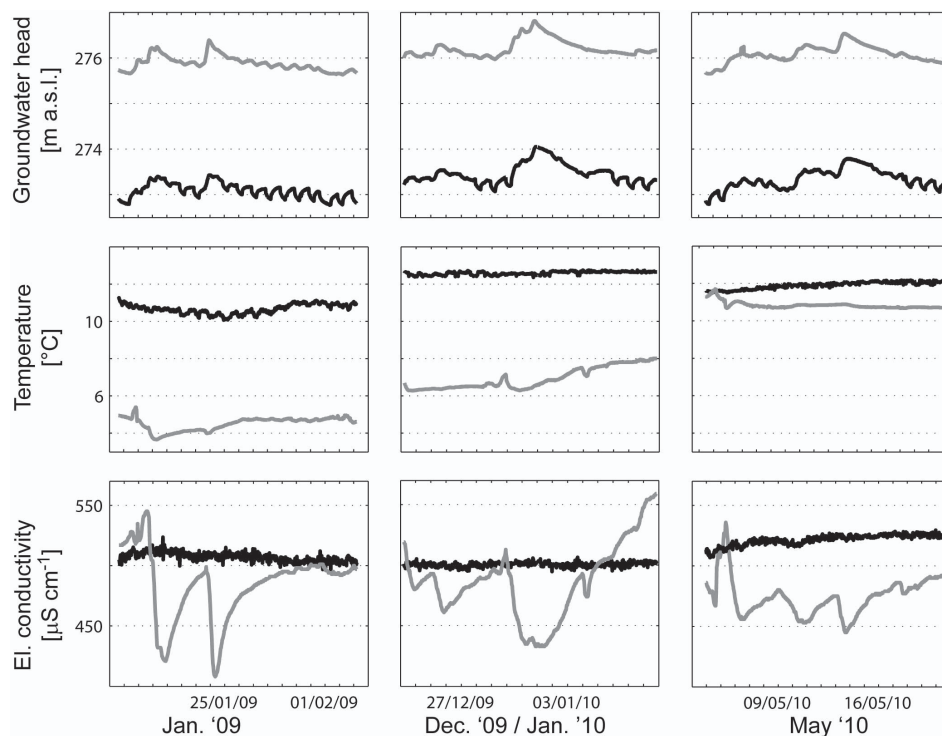
Figure 25: Groundwater head, temperature and electrical conductivity time-series (selection from input data set). The grey lines show the data from the observation well closest to the river (W3, 4m) and the black lines the data from the observation well furthest away from the river (B1, 17m).

8.5.2 SOM-SM Analysis

The SOM-SM analysis was carried out with 17 time-series: groundwater head, temperature and electrical conductivity measurements from six observation wells. One electrical conductivity time-series could not be used due to measurement error (W1). The three time periods (Jan. '09, Dec. '09/Jan. '10 and May '10), each containing 796 measurement time points, were analyzed as one data set. The dimension of the resulting matrix was

2388 x 17. The topographic error (topology preservation) and quantization error (map resolution) of the self-organizing map were 0.03 and 1.21, respectively. The correlation coefficient calculated for the normalized input data and the Sammon's projection using the SOM-SM approach was 0.92, indicating a good representation of the internal structure of the data set (information content) by the SOM-SM. The visualization of the combined SOM-SM analysis result is shown in Fig. 26. Each circle (point) represents one measurement time (ti) and all 17 variables, thus indicating the system state at time ti. The temporal distance between two points equals the measurement resolution of 0.5 h. The distance between two points in Fig. 26 in the plot is determined by the similarity of the two points with respect to system state. The shading of the points shows to which of the three time period the points belong. The development of system state, represented by the location of points in the Sammon's projection, is similar between the three events: each time period starts in the upper half of the plot, moves towards the lower half and back up to the upper half, creating a u-, or v-shaped trajectory. The events during Jan. '09 and Dec. '09/Jan. '10 reach further into the lower half of the plot than the development of system state during time period May '10.

In order to identify hazardous situations, defined as critical system states, SAC and particle density were used as independent measurements to create the shading in Fig. 27 panels A and B, respectively. The normalized SAC and particle density measurement values recorded in B3 were used, as B3 was the closest of the three observation wells equipped with additional instrumentation to the river. Neither SAC, nor particle densities were used in the SOM-SM calculations, they are merely used for display purposes and as independent indicators for recent river water. The values used for the representation were normalized to avoid between-time period differences and to highlight variation within each time period and thus the behaviour of the variables during high discharge events. SAC and particle density both increased during the high discharge events as river water infiltrated. As shown in panel A, Fig. 27, system states with high SAC values are located in the lower area of the plot and system states with low values in the upper

92
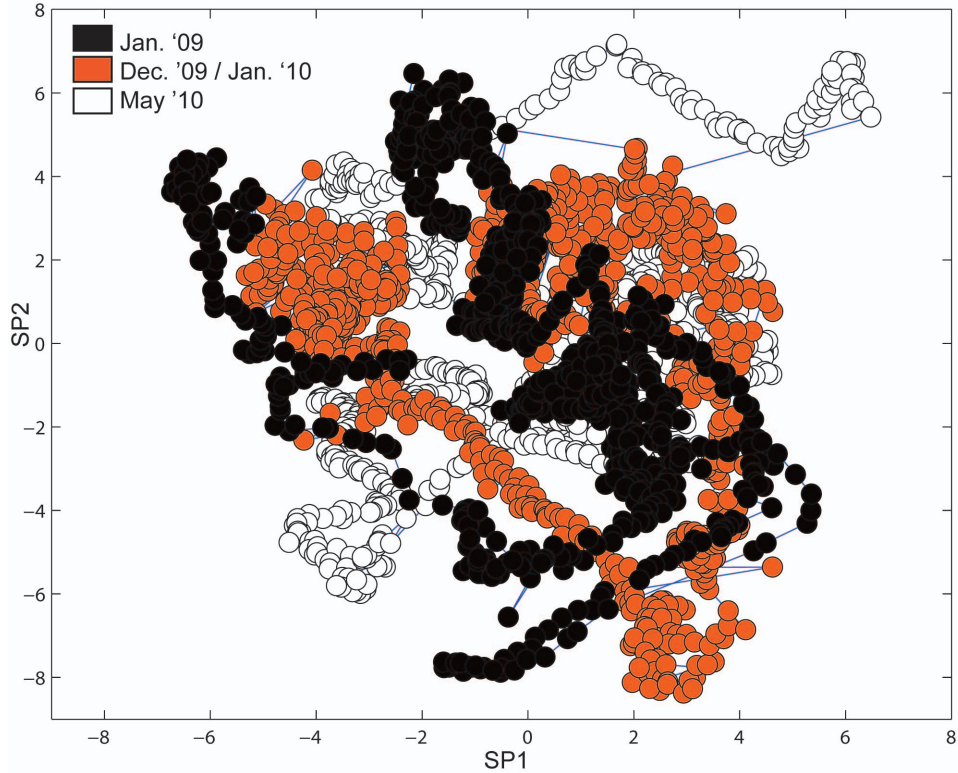
Figure 26: Visualization of the results from the SOM-SM analysis using groundwater head, temperature and electrical conductivity measurements from six groundwater observation wells during three periods with high discharge events. The axes shown are the Sammon projection axes 1 and 2 (SP1 and SP2). The shading of the points shows to which period the point belongs. Each point represents one measurement set consisting of 17 variables (one measurement time point).

area. Similarly, system states with high particle densities are located in the lower areas of panel B. Only time periods Dec. '09/Jan. '10 and May '10 were used to display the particle density results, as the particle density measurements in the Jan. '09 period could not be used due problems with the instrumentation. Three brief periods of high particle densities were observed (two during the Dec. '09/Jan. '10 period, and one during May '10).
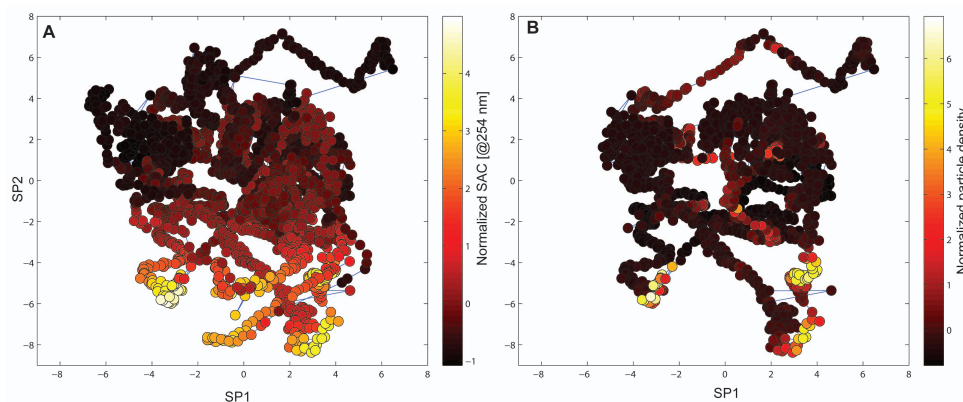
Figure 27: SOM-SM visualization based on groundwater head, temperature and el. conductivity. Panel A shows the results from the three time periods and normalized SAC measurements (from B3) to create the shading of the points. The shading in panel B is derived from normalized particle densities measured in B3. Panel B shows the results from periods Dec. '09/Jan. '10 and May '10 only.

The distance between two points can be used to quantify the change in system state as a function of time. Fig. 28 (panels A - D) shows the distances representing changes occuring over 0.5 h (measurement resolution) together with groundwater head recorded in W3. The distances between two points in the SOM-SM projection reflects the magnitude of the effect of the high discharge events on the groundwater-quality time-series used in the analysis. Strong fluctuations in measured parameter time-series and associated large changes in system state will lead to large distances, suggesting rapid changes in any of the parameters used for the analysis. In the case of Jan. '09, el. conductivity decreased and groundwater head increased rapidly, causing

94

large changes in system state over 0.5 h. The changes occuring during May '10 did not have the same magnitude and thus resulted in smaller distances. Panels D and E in Fig. 28 show two points that could potentially be used in a CCP approach: a) hexagon when the distance indicating system state change is larger than previous distances and the direction of the change (panel E) is towards system states that were previously identified as critical (e.g. Fig. 27), and b) square when a large change in system state has occured (large distance) and the system has moved further towards critical levels.
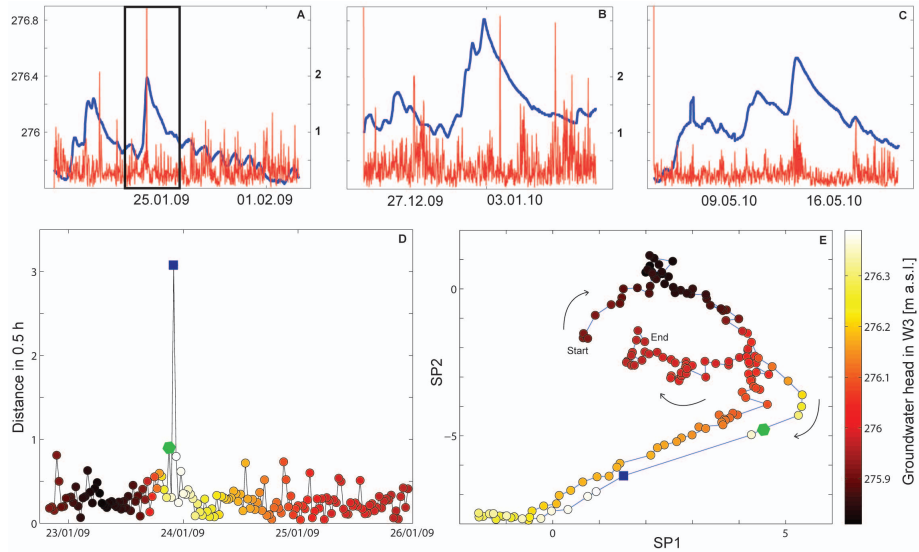
Figure 28: Panels A - C show the distances between two points in the SOM-SM projection together with groundwater head recorded in W3. The distances represent the changes in system state over 0.5 h (input data resolution). Greater distances (1 and above) can be observed during the high discharge events during all three observation periods. Panels D and E show the detail during the second high discharge event in Jan. '09. The shading shows groundwater head level recorded in W3. The hexagon indicates an initial warning and the square the point in time when a warning would be issued to the water supplier to alter the pumping regime.

8.6   Discussion

In this study, time-series of three groundwater parameters were analysed to identify critical system states, when contamination of extraction wells near rivers was imminent. The three high discharge events studied differed from each other in terms of magnitude and velocity in which the measured groundwater parameters (groundwater head, electrical conductivity and temperature) responded. All groundwater head time-series showed a strong response to all three events. Groundwater head appeared strongly influenced by pressure wave propagation through the aquifer, whereby multiple influences, such as groundwater extraction and river-stage fluctuations, may interfere with each other (Lewandowski et al., 2009). The observed decrease in electrical conductivity was probably a consequence of dilution effects caused by infiltration of river water with lower ion concentrations than the groundwater. The magnitude of the decrease in electrical conductivity depended on the distance between the observation well and the river, with the largest decrease in W3, the observation well closest to the river. Some events also showed an initial increase in el. conductivity, e.g. first event in Jan. '09 or May '10, this increase was, however, only observed during events not preceeded by other high discharge situations. The increase in electrical conductivity was assigned to a *first-flush* effect due to higher water levels reaching areas with high concentrations of soluable ions not normally inundated, leading to higher ionisation levels of the water (*washing out*). This effect was rapidly lost due to dilution effects as more river water infiltrated. Both effects, in- and decrease of electrical conductivity, were lost in the the observations wells further away from the river (B1, Fig. 24). Although, electrical conductivity showed a strong response to high discharge events, the non-conservative behaviour of the parameter may cause problems and lead to inaccuracy (Sheets et al., 2002; Cirpka et al., 2007). Temperature was the most conservative parameter in all the observation wells, showing only minor reactions to the high discharge events. Seasonality and river temperature strongly influenced groundwater temperature close to the river. Fig. 25 shows the difference in temperature between the events measured in W3,

indicating a strong dependency on river water temperature. The effectivity of temperature as an indicator for river water infiltration is however limited in the experimental site, when river and groundwater temperatures are similar, which usually occurs at least twice a year (in the spring and autumn). These three indicator parameters have all been used in previous studies assessing groundwater quantity and quality (Conant, 2004; Keery et al., 2007; Vogt et al., 2010). Their individual use in a HACCP approach is however limited due to their dependency on seasonality, river water composition and general hydrological situation.

In this study, a multivariate analysis method based on relative responses to high discharge events in the river, was applied to extract inherent patterns, which can be used in an adaptive approach to groundwater monitoring and management. The multivariate, nonlinear SOM-SM method was used to analyse time-series of all three indicator parameters (groundwater head, electrical conductivity and temperature) for three time periods including high discharge events. The SOM-SM approach focuses on the difference between individual situations (system states) and allows the user to categorize the situations into *critical* and *non-critical* system states (Mustonen et al., 2008; Camplani et al., 2009).

Besides the classification into critical and non-critical system states, the rate at which changes in system state occur is an important aspect when considering critical control points. In this analysis, the difference between two consecutive measurement points was assessed using their similarity. During high discharge events, changes in any or all of the measured time-series will result in a change in system state. Changes in system state are defined by the distance between two consecutive points in the SOM-SM projection. The smaller the distance between two points in the SOM-SM projection (Fig. 26), the greater the similarity between the system states (situations). In accordance, clusters represent similar system states and the location of points in the projection relative to each other is a measure of similarity. The joint analysis of all three observation periods was used to eliminate the potential for effects due to seasonality or other inter-event differences to dominate the SOM-SM pattern. As the coordinates used for the Sam-

mon's projection are based on the eigenvectors of the input data matrix, an analysis of each event individually could not be generalized. The joint approach was used to identify general patterns observed during high discharge events that represent the behaviour of the system, independent of individual events or situations. High point densities thus signify frequent situations, whereas low point densities indicate rarer situations, or system states. The low-density point clusters should indicate critical system states, as they occur less frequently in the time-series used for the analysis. Accordingly, the points in the lower section of Fig. 26 are more likely to represent critical system states, i.e. situations when contamination of the extraction wells may occur. As opposed to data sets used in previous studies, e.g. Mustonen et al. (2008) or Camplani et al. (2009), the relative difference between critical and non-critical system states derived from the data set used in this study were significantly smaller. Nevertheless, the distance between two points pinpointed changes in system state between two measurements: the greater the distance between two points, the greater the change in system state. Besides the visual analysis, a detailed distance analysis can also be used to identify situations, where the system state has reached temporary equilibrium and changes between system states decrease. As, in general, more changes occur during high discharge events than during average situations, the point in time when the distances decrease can identify *return-to-normal* situations. As shown in Fig. 28, infiltration events can either result in large distances (e.g. January 2009), or be represented by a cluster of intermediate distances (e.g. May 2010).

The observations and patterns in the projection need to be validated with an independent data set to avoid self-fulfillment. In this study, spectral absorption coefficient (SAC) and particle density (2 - 10 μm) were used as independent parameters to validate the interpretation of the SOM-SM projection. The origin of SAC in the groundwater observation wells was riverine and indicated the presence of organic matter. SAC has also been used as an indicator for faecal contamination (Stadler et al., 2010). Building on this, high SAC measurements indicated critical system states, i.e. points in time when increased riverine infiltrate is detected and contamina-

tion of the groundwater extraction wells may occur. Particle density is also a parameter that has been used as an indicator for faecal bacteria contamination (Auckenthaler et al., 2002; Pronk et al., 2007). Although the size range recorded in this experiment did not coincide with a particular type of bacteria and was larger than viral pathogens, the presence of elevated particle densities in groundwater suggests surface water infiltration. These two additional parameters are thus both strongly associated with river water infiltration and faecal pollution and can be used as indicators of critical situations, where drinking-water extraction wells may suffer from microbial contamination. Both SAC and particle density showed the highest values in the lower area of the SOM-SM projection, supporting the observation of point density clustering suggesting the location of points in this area represent critical system states. The high discharge event in May '10 was smaller in terms of parameter response to the high discharge in the river, and associated infiltration, e.g. decrease in electrical conductivity. The consistent postition of points representing critical system states in the SOM-SM projection, as identified by elevated SAC measurements and particle density, suggests that the combination of groundwater head, temperature and el. conductivity measurements can be used to identify situations, where the groundwater extraction wells may be contaminated if no measures are taken to avert the hazard, e.g. stopping extraction. The critical system states during each time period were located in the same area of the projection (lower half), indicating a similar behaviour during high discharge events. As shown in Fig. 28, once the critical system states have been identified using independent time-series, the combined analysis of location of points (clustering to classify) and distance between consecutive points can be applied to a HACCP approach. In agreement with Dominguez et al. (2007), the SOM-SM approach to process supervision is an important part of a decentralized and distributed management system and can be included in a HACCP-approach.

As the SOM-SM method is based on inter-measurement differences, the effectiveness increases with the magnitude of change in system state. Rapid infiltration and associated changes in measured parameters will evoke large

distances, which can then be used to issue a warning to the water supplier and provide sufficient time for necessary management options to be taken. However, events with only small or gradual changes could be missed if the minimal distance for detection is not adequately defined. Small or gradual events are less likely to pose a threat to the drinking-water supply, as the time between infiltration and extraction is increased. The increased time allows the intrinsic filter mechansims of the aquifer material and natural die-off of pathogens to improve the quality, thus minimising the contamination hazard. However, persistent contamination may still jeopardize water quality. Consequently, the definition of the distance between points signifying the onset of potential contamination requires validation using a range of events, including large and small, as well as rapid and gradual changes in system state.

8.7   Conclusions

The timely identification of potential contamination of groundwater extraction wells is key to an adaptive approach to groundwater management. The SOM-SM analysis provides a means by which critical control limits, defined by the inter-point distances and relative point locations in the projection, can be applied to groundwater extraction-well protection. The multivariate statistical analysis using groundwater head, electrical conductivity and temperature was able to identify changes in system state indicating the arrival of rapidly infiltrating river water and with it, potential contamination.

Although groundwater head, temperature and el. conductivity are commonly measured parameters in many groundwater observation wells, the reliability of the individual parameters as indicators of river water infiltration may be poor. While hydrogeological properties, such as hydraulic conductivity and length of flowpath through the aquifer, play a major role in the parameter response to high discharge events, the behaviour is also dependent on hydrological and seasonal situations. The resulting variability of potential system states is driven by complex behaviour of the parameters during high discharge events and the dependency on other factors including

seasonality.

Combining the SOM and Sammon's mapping methods as a multivariate proxy analysis provided the basis for identifying critical system states. SAC and particle density, although both parameters are not usually recorded continuously in groundwater observation wells, provided independent information on critical system states. The resulting maps showed patterns based on the measured parameters during three different high discharge events. After successful validation by independent time-series, the generalised behaviour during high discharge events can be used to identify situations, when the system is strongly influenced by infiltrating river water. The time of detection of infiltrating river water is based on changes occuring in the measured time-series and can be used to issue warnings to the water supplier. Besides issuing warnings concerning the detection of critical system states, information on when extraction can be resumed is an important element of the management process. The resumption of extraction requires further validation, which may include groundwater flow modelling.

The characterisation of high discharge events and their relationship to water quality can help determine optimal management strategies. Additional degrees of freedom can be granted to the management of extraction wells near rivers, which have high-quality water during most of the year, but where water quality is severely compromised during high discharge events. The complex task of identifying hazardous situations for drinking-water quality in a heterogenous environment can thus be simplified by using a SOM-SM proxy analysis as part of a HACCP procedure.

8.8 Acknowlegdements

9 Groundwater Flow Simulation

Large parts of the Birs Valley aquifer have been modelled in previous studies addressing various aspects of groundwater flow in urban and suburban environments. The existing groundwater-flow models have been used to study questions concerning well-head protection areas, the efficiency of artificial aquifer-recharge and flow around a weir (Epting et al., 2009; Affolter et al., 2010). All are finite-difference models constructed using the Groundwater Modelling Software (GMS 7.1 and previous versions) provided by Aquaveo (previously Environmental Modelling Systems Inc.). The flow simulation was solved using MODFLOW 2000 (McDonald et al., 2000). A high-resolution groundwater-flow simulation model (*Reinacherheide*) was embedded in an existing model (*Aesch-Reinach*, Affolter et al. (2010)). Many of the boundary conditions and aquifer characteristics defining the *Reinacherheide* model (for example, porosity and hydraulic conductivity) were based on the existing larger model (*Aesch-Reinach*). The resolution (spatial and temporal) was significantly increased to match the depth-oriented installation of the measurement instrumentation and the temporal resolution of the extraction-well pumping data. Special attention was paid to river-groundwater interaction, as it is one of the dominating factors influencing groundwater flow in the valley.

The objective was to create a small-scale model (limited to the size of the well field) to simulate river-groundwater interaction and the impact and consequence of extraction well management based on the combined method of self-organizing maps and Sammon's projection (SOM-SM).

9.1 Model Description

The model *Reinacherheide* has a homogenous horizontal resolution (5 x 5 m) and a variable vertical resolution distributed over four layers. The thickness of the layers is a function of surface topology, aquitard, mean groundwater head level and the filter stretches of the observation and the extraction wells (Fig. 29). The aquitard is described in Chapter 3 and was implemented as the impervious bottom of the model. The upper constraint was interpolated

102

from a high-resolution point data set kindly provided by the *University of Applied Sciences Northwestern Switzerland (School of Architecture, Civil Engineering and Geomatics)*.

Four hydraulic conductivity zones were defined per model layer (the same in each layer). The boundaries of the zones followed the terrace structure of the surface topology. The initial values used had already been manually calibrated for the larger model (*Aesch-Reinach*, Affolter et al. (2010)). The hydraulic conductivity for all 16 zones was transiently calibrated using PEST (Doherty, 2005) and adapted manually where necessary. The calibration period used was January 2009 with a 1-hour resolution and 12 observation points. The final hydraulic conductivity ranged between 3 x 10^{-3} and 2 x 10^{-2} m s^{-1}.

Lateral boundary conditions

The lateral constraints of the model are defined by three types of boundary conditions: Dirichlet, Neumann and Cauchy (Fig. 30). The southern and western boundaries are implemented with a Neumann condition by defining a transient flow. The flow was based on hyraulic modelling carried out within a previous study on nitrate transport and based on catchment areas and seasonal precipitation-driven groundwater recharge (Huggenberger et al., 2006). The locations of the southern and western boundaries were based on the saturated zone: the aquitard rises on the edges of the ancient river bed, significantly reducing the thickness of the saturated zone and thus providing a boundary for the model. The southwestern and northern boundaries were included as *constant heads* defined by a Dirichlet condition, for which groundwater head time-series from three observation wells provided a transient potential. The River Birs forms the eastern model boundary and was included as semi-permeable boundary (Cauchy condition). The permeability is determined by the hydraulic gradient between the river stage and adjoining groundwater head and the conductance of the river bed.

River-bed conductance

A significant factor in determining groundwater flow and quality in the study area is the conductivity of the river bed. While river-groundwater interaction is a function of the difference in head between the river and the
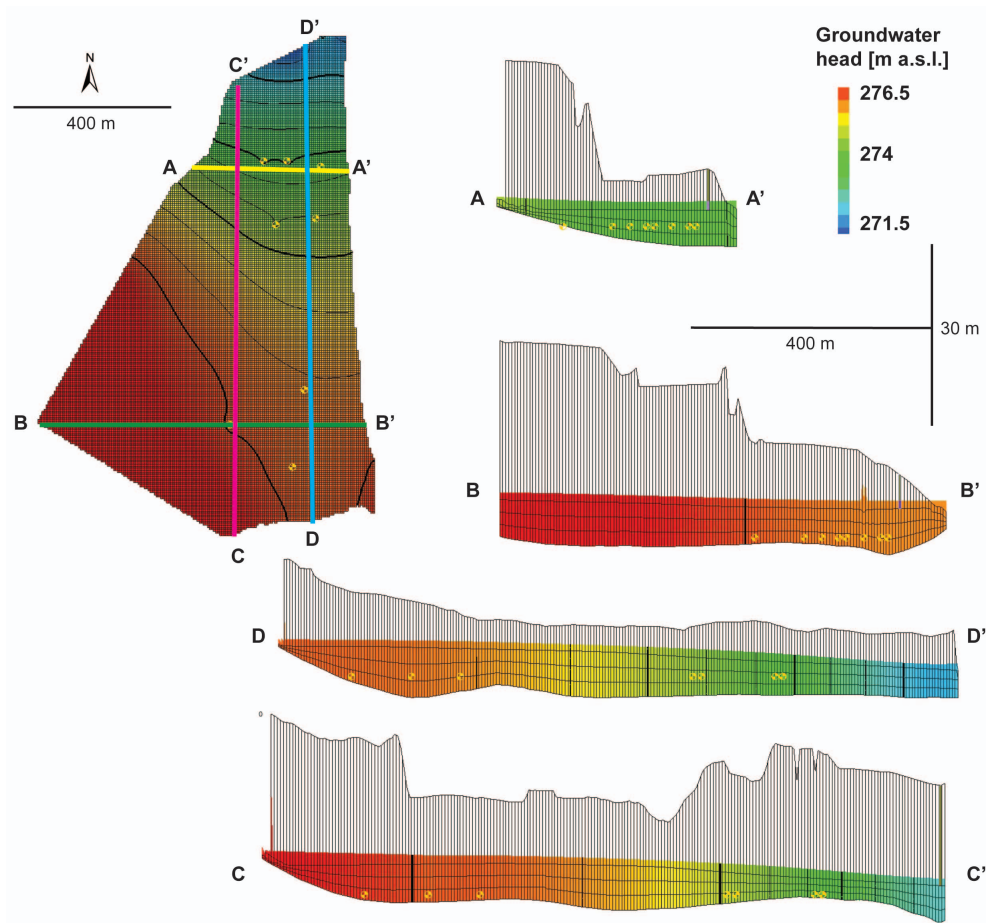
Figure 29: *Reinacherheide* model geometry. The vertical scale of the four sections (AA', BB', CC', DD') is 10x the horizontal scale. The terraces are visible in profiles AA' and BB'. The top layer is generally the largest, but mostly encompasses the unsaturated zone. The shading corresponds to groundwater head indicating the thickness of the saturated zone.
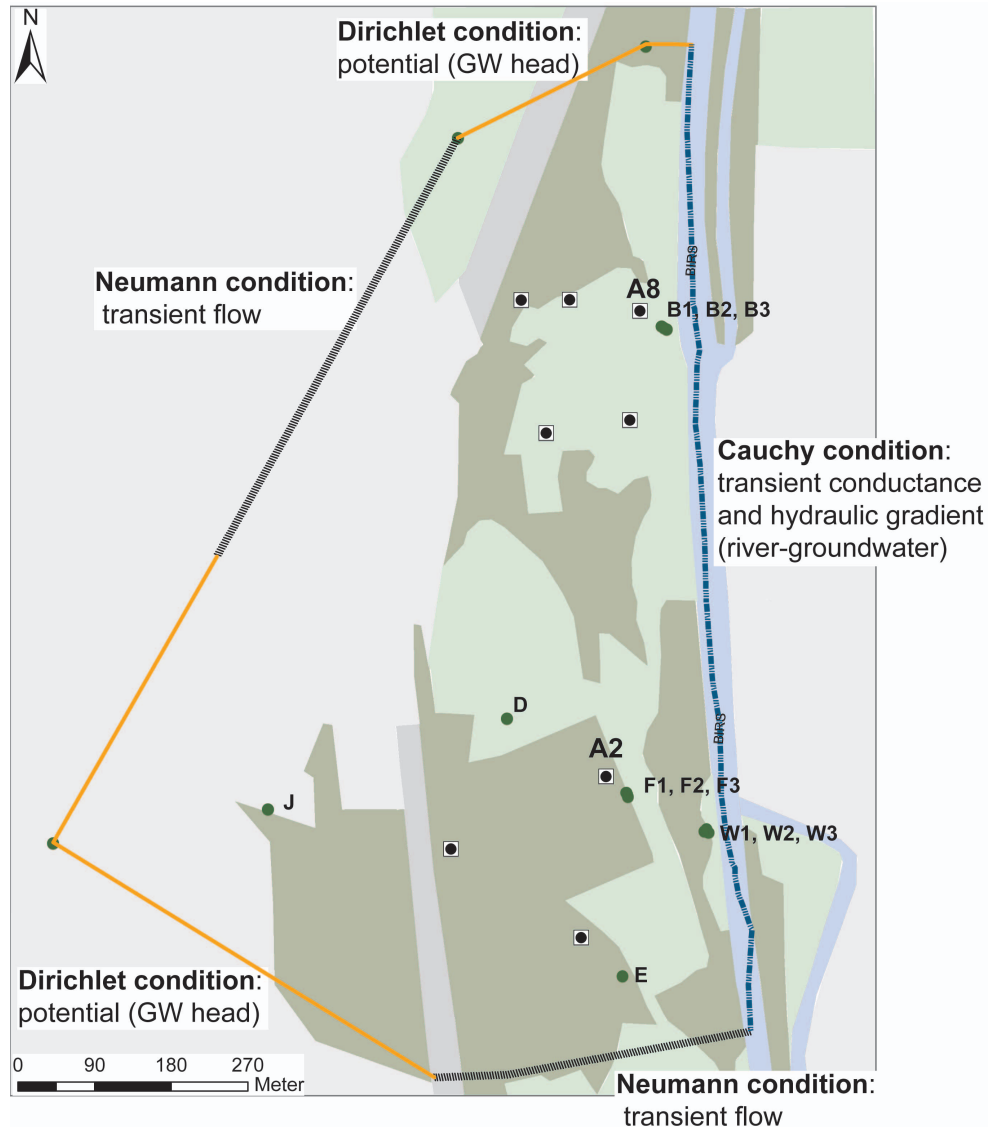
Figure 30: Boundary conditions in the *Reinacherheide* groundwater-flow simulation model. The Neumann boundaries are defined by a transient flow based on seasonal fluxes. The Dirichlet boundary conditions are based on groundwater head time-series in three observation wells. River-groundwater interaction is implemented as a Cauchy boundary using a conductance time-series and hydraulic gradients between the river and the groundwater levels.

groundwater (determined by model environment), the conductivity can be determined by the user. Three conductance values were implemented based on river-stage fluctuations. The minimal conductance value, determining river-groundwater interaction during average hydrological situations, was taken from previous studies (Affolter et al., 2010). The function determining river-groundwater interaction applied in this model is based on river-stage fluctuations, rather than discharge, as it was in the previous studies. The use of river-stage fluctuations was chosen, as river-stage measurements were available for the study area, while direct discharge measurements were not available. In addition, the profile of the river bed was not even, so, calculating discharge would give rise to additional errors and uncertainty. Two further conductance values were used to differentiate between small and larger high discharge events (Fig. 31). The gradient between two river-stage measurements, i.e. the change in river stage over an hour, was used as a proxy indicator for forces resulting in an opening of the river bed by dislodging coarse components and/or removal of fine sediment from the river-bed matrix. Thus, large changes in river stage, usually brought about by larger high discharge events, give rise to larger conductivity values (Fig. 31).

Measured vs modelled

Two of the components determining the error between measured and modelled groundwater-head time-series are the absolute value of groundwater head (too much or too little) and the shape of the curve (time-series similarity). While the error in absolute value of groundwater head is strongly variable (between zero and 80 cm, in extreme), the shape of the curve is very similar in measured and modelled time-series (Fig. 32). The correlation between the measured and modelled values ranges between 0.64 and 0.99 (January 2009). In general, the closer the observation wells are to the river, the weaker the correlation between the measured and the modelled values (for example, J, 530 m from the river, has a correlation coefficient of 0.99, whereas W3, closest to the river (4 m), and has a correlation coefficient of 0.64). Two other factors influencing the difference between modelled and measured groundwater head are suggested by the results of the PCA and have been discussed in Chapter 6. Sand and silt lenses in the area around
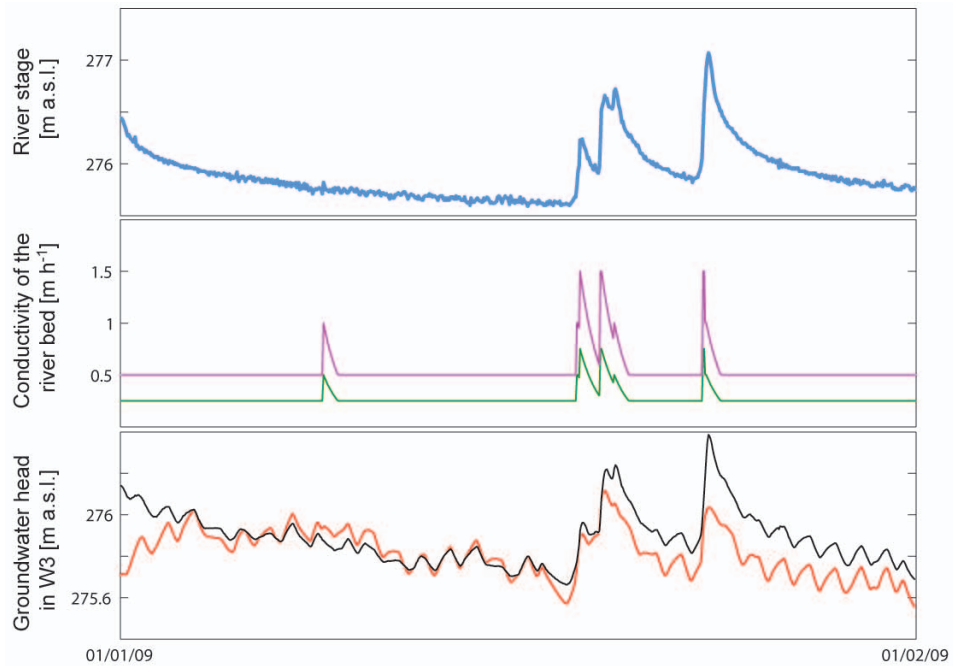
Figure 31: River stage, two river-bed conductivities and groundwater head (measured and modelled). The purple conductivity values are higher than the green ones because the river bed to the north of the *Heidebruggli* is more permeable than in the southern stretch of the study area, where the river bed has a lower gradient and flow velocity. The largest difference between the measured (black) and modelled (red) groundwater heads was observed during the high discharge events.

the W-cluster lead to small-scale effects that could not be included in the model, partly due to the geometric resolution of the model and partly because the exact composition of the aquifer is not known. W2 and F2 both have lower correlation values than would be expected based on neighbouring observation wells. The model is not able to capture the heterogeneity of the aquifer in these areas. Based on the results of the geoelectrical study carried out by Wüest (2010) and genetic structure of the aquifer, the presence of a paleochannel could lead to such differences between the measured and the modelled time-series. The inclusion of such heterogeneities requires a different approach to modelling and may include geostatistical analysis of aquifer composition. Nevertheless, the model is able to summarize the major influences on groundwater head in the well-field and can thus be used to simulate the effect of different management options (pumping regime) on the groundwater head distribution and flow as a function of groundwater extraction and river-groundwater interaction.

9.2   Scenario Analysis

The groundwater-flow simulation model was used to test different management options, i.e. pump settings, during two of the observation periods described in Chapter 8 (January 2009 and May 2010). The extraction wells are managed according to their susceptibility to riverine contamination. Guidelines based on river stage recorded approximately 4 km downstream of the well-field determine the point in time, when extraction has to be stopped. Accordingly, extraction was stopped during peak flows and recommenced a day after the river stage reached a threshhold level determined by operational water-supply manager. Fig. 33 shows the extraction regime of the two extraction wells most susceptible to contamination due to infiltrating river water during the two observation periods January 2009 and May 2010 (A2 and A8). The susceptibility of these two wells was determined by finding elevated bacterial counts in samples taken after high discharge events. By Swiss law, no *E. coli* or *Enterococcus* sp. should be found in 100 ml of water. The presence of these indicator bacteria leads to the assumption that further
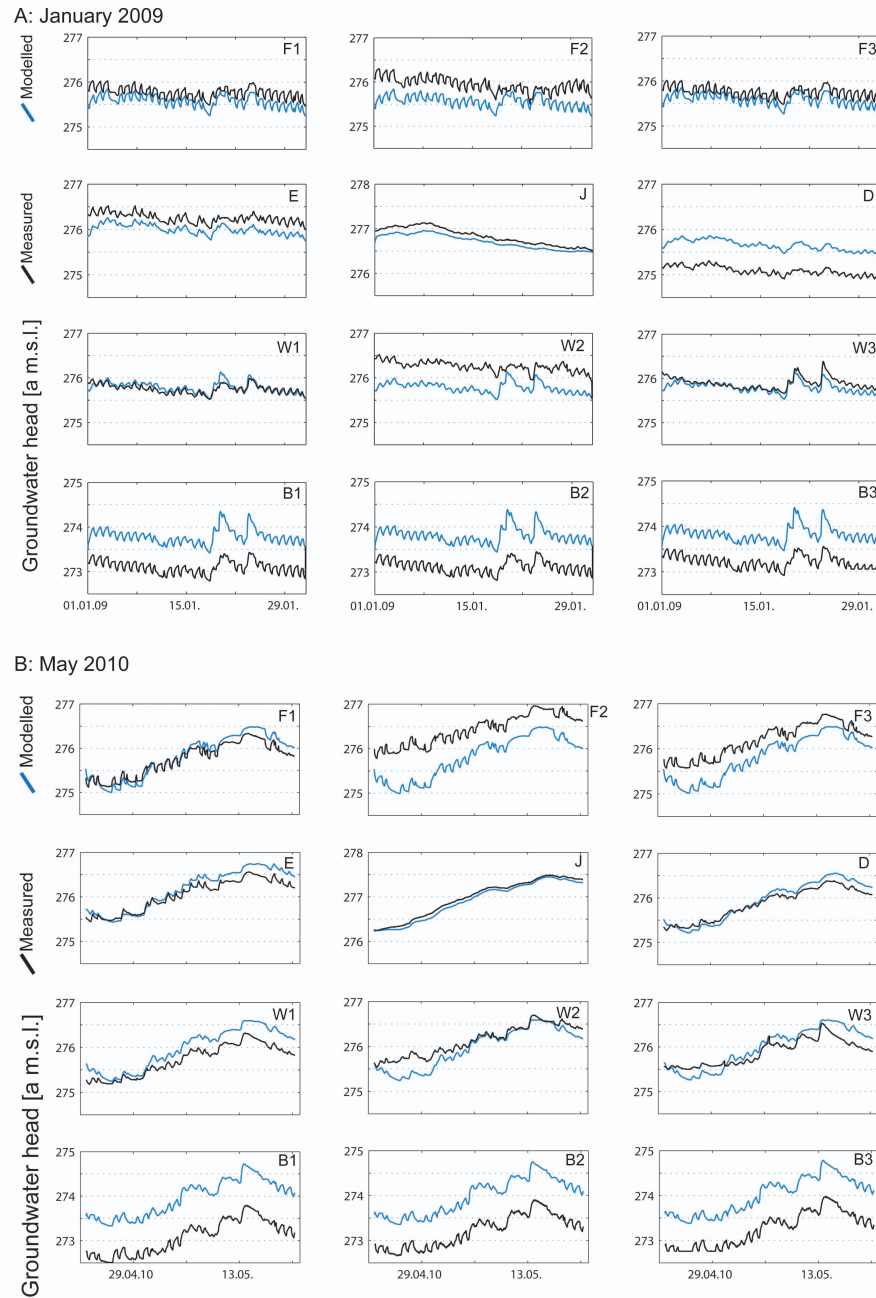
Figure 32: Measured (black) and modelled (blue) groundwater head time-series during the two observation periods January 2009 (group A) and May 2010 (group B). Modelled groundwater head was consistently greater than measured head in B1, B2 and B3 during both observation periods.

Table 2: Scenarios for extraction wells A2 and A8 during the observation periods January 2009 and May 2010.

Scenario	Name	Regime	A2 (m^3 h^{-1})	A8 (m^3 h^{-1})
1	Nil	constant	0	0
2	Full	constant	141	80
3	Original	intermittent	0 / 141	0 / 80
4	Day / Night	intermittent	0 / 141	0 / 80
5	SOM-SM	intermittent	0 / 141	0 / 80

faecal pathogens are present in the water and that it is therefore not safe for consumption and extraction has to be stopped. Particle tracking was used to delineate the 10-day capture zones, from where potential contamination is likely to reach the extraction well. The 10-day capture zone is defined by Swiss law as a protection zone for each groundwater extraction well.

The scenarios simulated were based on different management options for the two groundwater extraction wells most susceptible to riverine contamination: A2 and A8 (Fig. 30). The scenarios used to assess the pumping regimes for these two wells is summarized in Tab. 2. The travel path of twenty particles was calculated for each well in each layer: both A2 and A8 were implemented in two layers, giving rise to 40 particles each. The scenario analysis aims (a) to outline the circumstances when fresh river water is extracted and may pose a hazard to drinking-water quality (critical situations), and (b) how contamination can be avoided by applying a multivariate analysis as described in Chapter 8.

9.2.1 January 2009

The high discharge events in January 2009 have been described in Chapters 6 and 8. The time period is characterized by two high discharge events in succession. The flow velocity and direction in the model area represented by flow vectors (arrows) in Fig. 34. During low discharge situations, the flow direction was parallel to the river, while during high discharge situations,
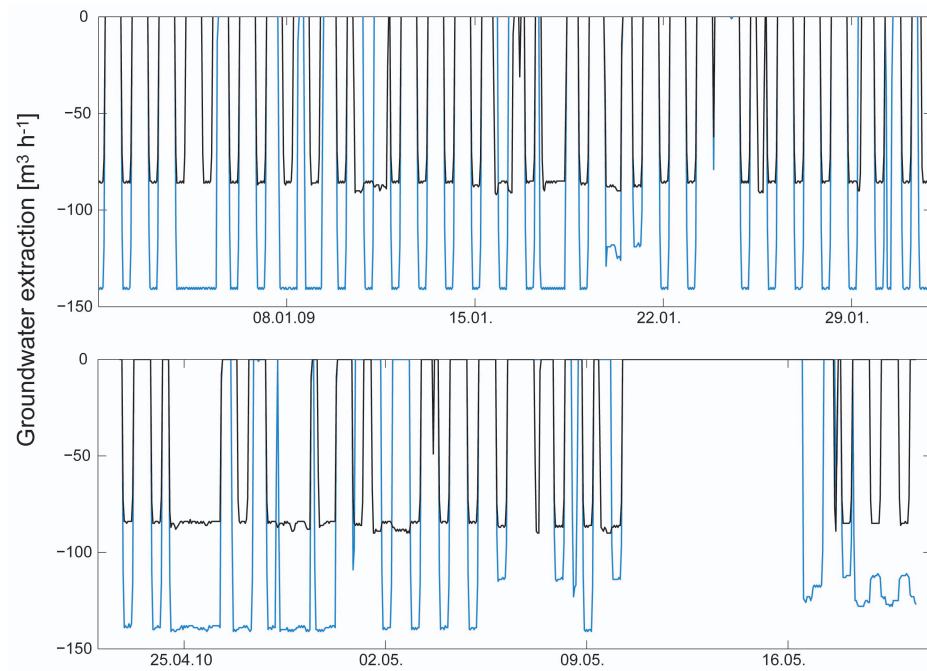
Figure 33: Groundwater extraction in January 2009 and May 2010 observation periods for A2 (blue line, max. 141 m^3 h^{-1}) and A8 (black line, max. 80 m^3 h^{-1}) (cf. Fig. 30). Drinking-water extraction ceased during both observation periods, but for considerably longer during the May 2010 event as multiple smaller events occurred after each other and the river stage did not decrease as rapidly as during January 2009.

the vector field changed to indicate stronger infiltration. The flow velocity vectors in Fig. 34 are matched to their magnitude (0 - 25 m h $^{-1}$). The largest velocities (larger arrows) were simulated around the *Heidebruggli* cluster (inset, Fig. 34) and in the central part of the model. The largest difference between discharge situations was also noted in the vicinity of the *Heidebruggli* cluster. Based on the flow velocity and dependency on river stage, this area shows the greatest fluctuations in terms of river-water infiltration. Flow velocities were smallest near the southern boundary of the model area. No distinguishable influence on velocity distribution (spatial or magnitude) was attributable to the different extraction regimes.
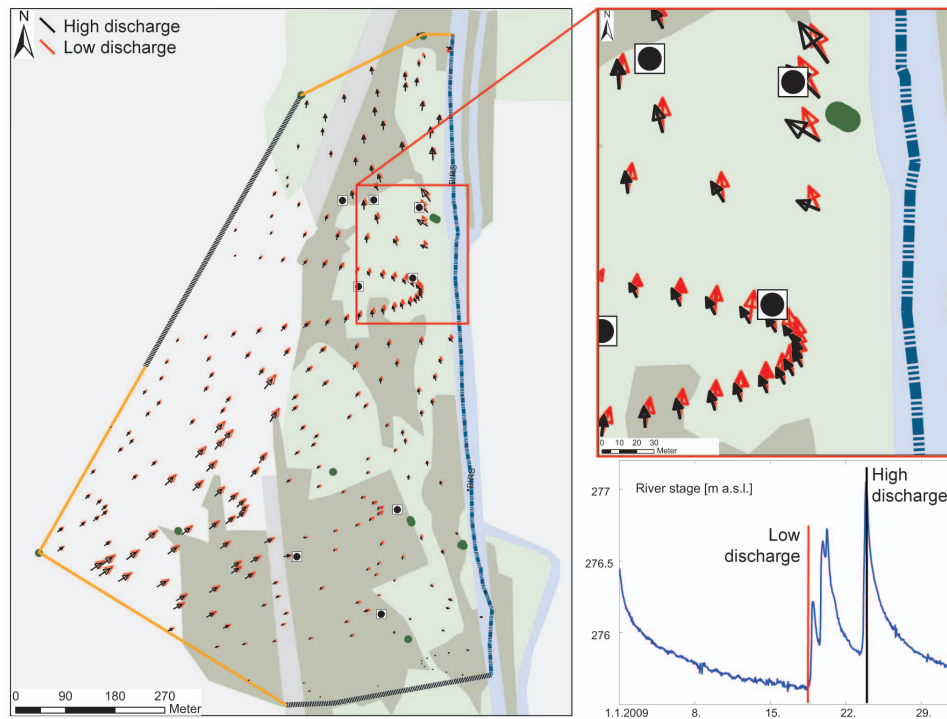
Figure 34: The flow field in the model area was dependent on river stage in January 2009. During high discharge events (black arrows), the vectors are oriented at a steeper angle away from the river than during low discharge situations (red arrows), indicating increased infiltration. Flow velocities range between 0 (smallest arrows) and 25 m h $^{-1}$ (largest arrows).

Scenarios 1 (nil) and 2 (full) are clearly distinguishable from the other scenarios (3-5), as shown in Fig. 35. These two extreme situations, i.e. no extraction and full extraction, show that infiltration of river water cannot be avoided during the high discharge events modelled, however, the amount of infiltration and the length of the passage through the aquifer can be managed. The 10-day particle tracks for scenarios 3, 4 and 5 show a similar distribution and cannot be visually separated from each other (panel C, Fig. 35). A more detailed analysis is presented in Fig. 36, where the amount of river-water infiltration was calculated based on the amount attributable to the extraction regime applied. The SOM-SM approach substantially reduced the amount of water that infiltrated during the time period following high discharge events (blue line). Apart from reducing the amount of water infiltrating, the flow path through the aquifer was increased by using the SOM-SM approach (Fig. 35, panel D). The length of the particle path is longer for the SOM-SM approach than for the scenario using full extraction (s.2). The brown lines in Fig. 35, panel D, show that the extraction path of the particles reached the river closer to the extraction well, than when using the SOM-SM approach (blue lines), thereby reducing the length of the time in the aquifer (velocities are near constant between scenarios) and the path length. The filter effect of the aquifer diminishes with these two factors, increasing the potential for contamination of the extraction wells.

The results from the flow simulation model for January 2009 show that the SOM-SM approach is able to identify situations when extraction should be stopped. The approach showed a more conservative management regime, reducing the amount of extraction for the two extraction wells A2 and A8, than originally adopted by the water supplier. Further attention should be paid to identifying the time when extraction can be resumed, so as not to give rise to an unnecessary conservative management regime.

9.2.2 May 2010

The high discharge events in May 2010 have been described in Chapters 7 and 8. The observation period is in the Spring and encompasses aquifer
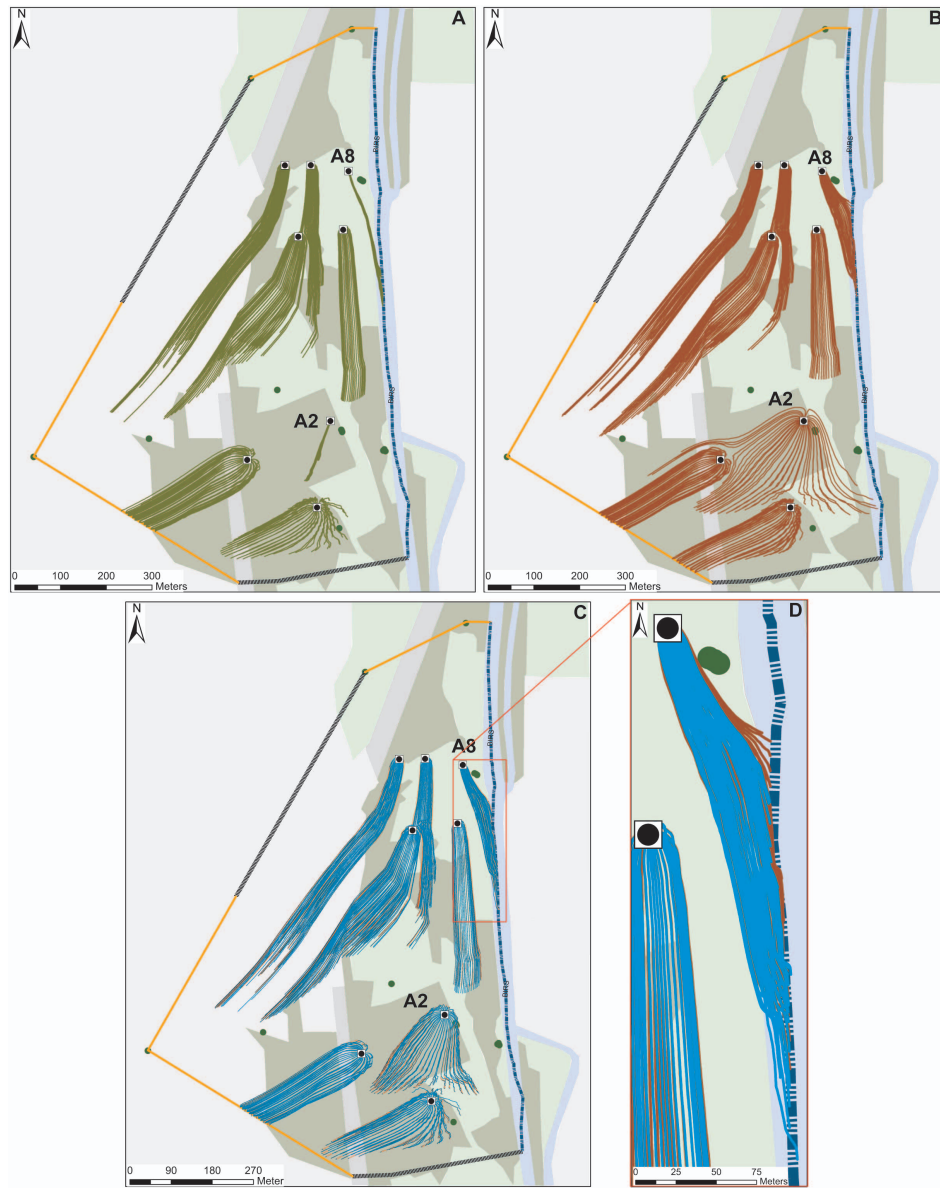
Figure 35: 10-day particle paths for all of the extraction wells in the well-field in January 2009. Panel A: scenario 1, where no groundwater was extracted from A2 or A8. Panel B: scenario 2, where the maximal groundwater volume was extracted from A2 and A8. Panel C: scenarios 3 (original), 4 (SOM-SM) and 5 (Day / Night). Panel D: scenarios 2 (full, brown) and 4 (SOM-SM, blue), where the difference in particle path length between the two scenarios is clearly visible.
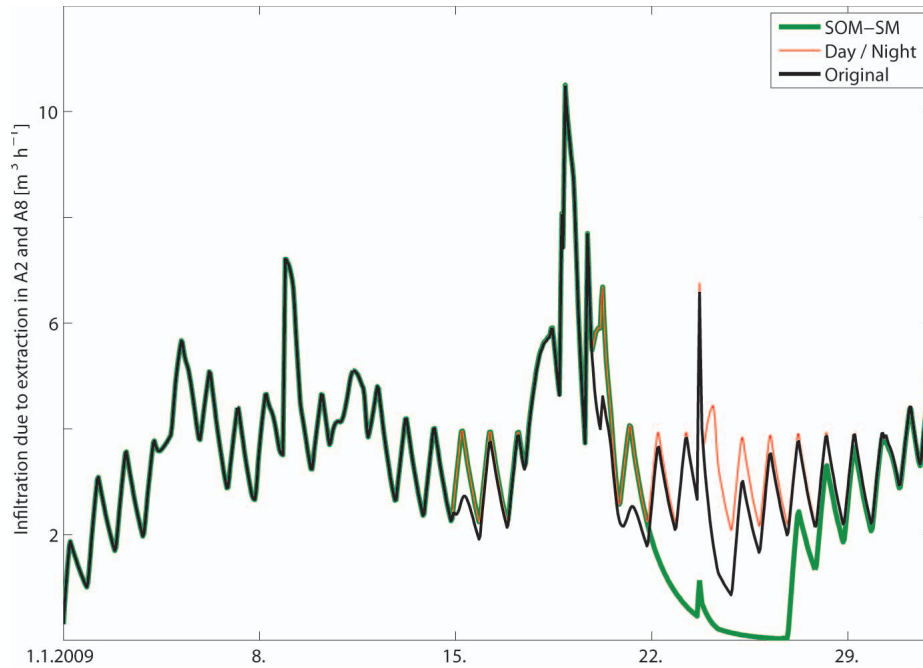
Figure 36: The time-series show river-water infiltration attributable to extraction in A2 and A8 in January 2009. The original scenario (actually run by water supplier) is shown in black. Day / Night (thin red line) is based on regular extraction during the night (normal regime). By applying the SOM-SM approach to managing the groundwater extraction wells A2 and A8, the amount of water infiltrating during and after the the high discharge events could be reduced (thick green line).

recharge. Different from the decreasing trend in January 2009, groundwater head in May 2010 showed an increasing trend, as indicated by measurements in observation well J (Fig. 32). The changes occurring (system state, cf. Chapter 8) are smaller and more gradual than during January 2009. Extraction wells A2 and A8 were out of use for a longer period in May 2010 than in January 2009 (Fig. 33).

Flow velocity and vector orientation were not substantially influenced by the extraction scenario applied and, unlike January 2009, the difference between high and low discharge situations was also negligible (Fig. 37). The orientation of the flow vectors was similar to the low discharge situations in January 2009, suggesting a weaker riverine influence.

As shown in panels A and B in Fig. 38, the scenario *no extraction* (s.1) leads A8 to draw no water immediately from the river, while *full extraction* (s.2) leads to extraction from the river. As for January 2009, the scenarios 3, 4 and 5 show similar particle tracks that are visually hardly differentiable (panel C Fig. 38). By applying the SOM-SM approach in May 2010, direct extraction of water from the river could be avoided. Fig. 35, panel D, shows the particle tracks (10-days) for the SOM-SM approach (blue) and for the full extraction scenario (brown). Besides avoiding extraction from the river, the length of flow through the aquifer was also increased.

The amount of infiltration due to extraction for A2 and A8 was also calculated for May 2010 (Fig. 39). In contrast to the conservative January 2009 scenario calculations, the SOM-SM scenario led to a greater extraction volume than the original extraction regime. As shown in Fig. 39, infiltration was reduced during five days in the original scenario before ceasing on the 10. / 11. May. The SOM-SM scenario allowed continued extraction for additional eight days, before ceasing on the 13. / 14. May. At this stage the SOM-SM approach only indicates critical situations when extraction should cease, however it is possible to create a *traffic-light* system to incorporate multiple pump settings, for example, 50%, or a differentiated approach for each extraction well. These steps would require a optimization for each well in combination with high-resolution multi-objective modelling.
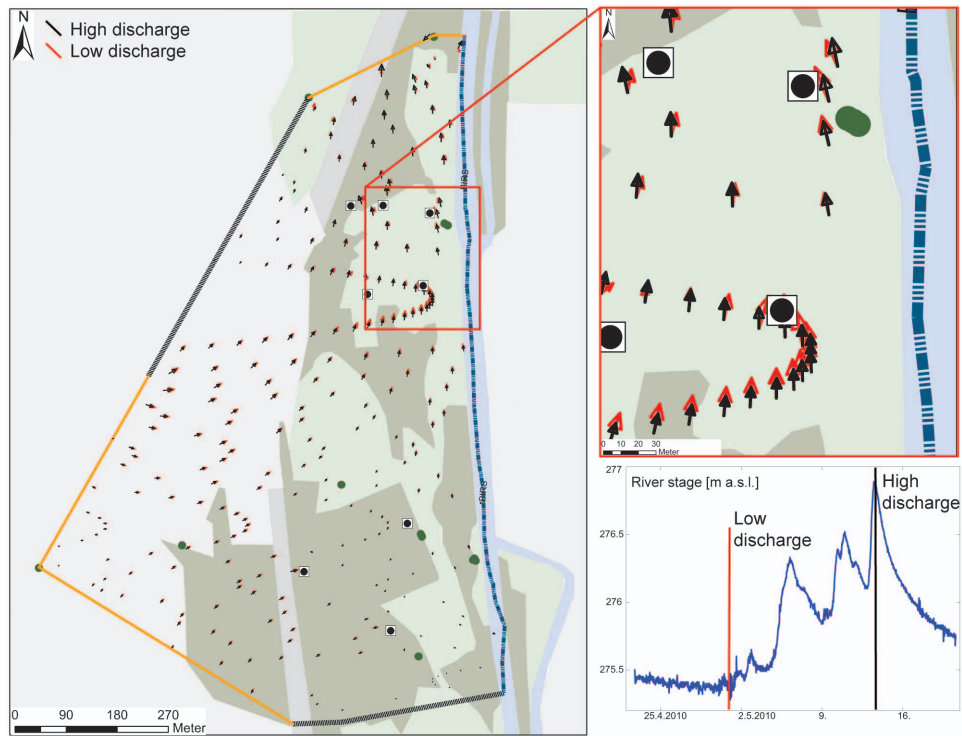
Figure 37: The flow vectors during high and low discharge situations in May 2010 show only small differences. Black arrows show the situation during high discharge events, red arrows low discharge situations. Even flow in the areas most prone to changing flow directions, *Heidebruggli*, appears parallel to the river flow direction.
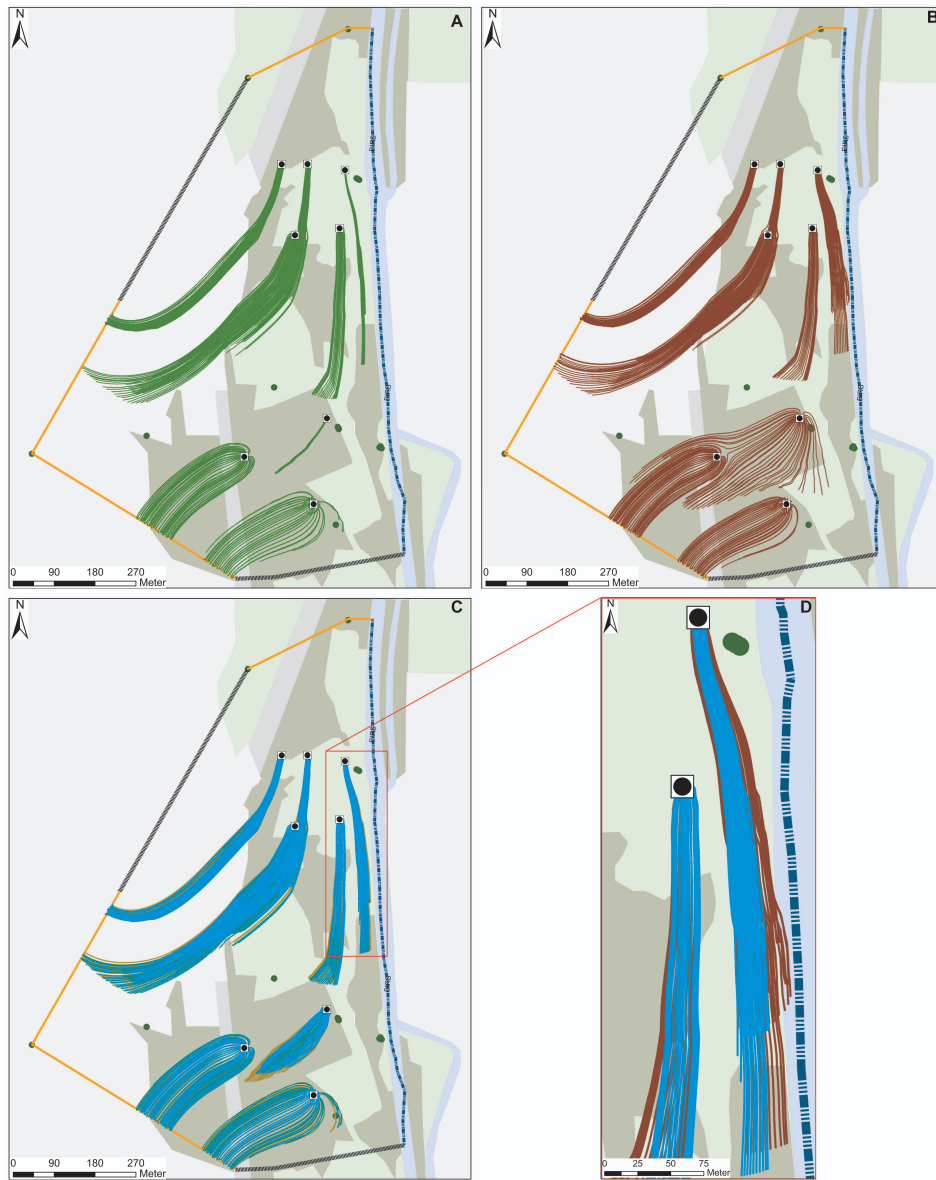
Figure 38: 10-day particle paths for all of the extraction wells in the well-field in May 2010. Panel A: scenario 1, where no groundwater was extracted from A2 or A8. Panel B: scenario 2, where the maximal groundwater volume was extracted from A2 and A8. Panel C: scenarios 3 (original), 4 (SOM-SM) and 5 (Day / Night). Panel D: scenarios 2 (full, brown) and 4 (SOM-SM, blue), where the immediate extraction of river water is avoided using the SOM-SM approach.

118
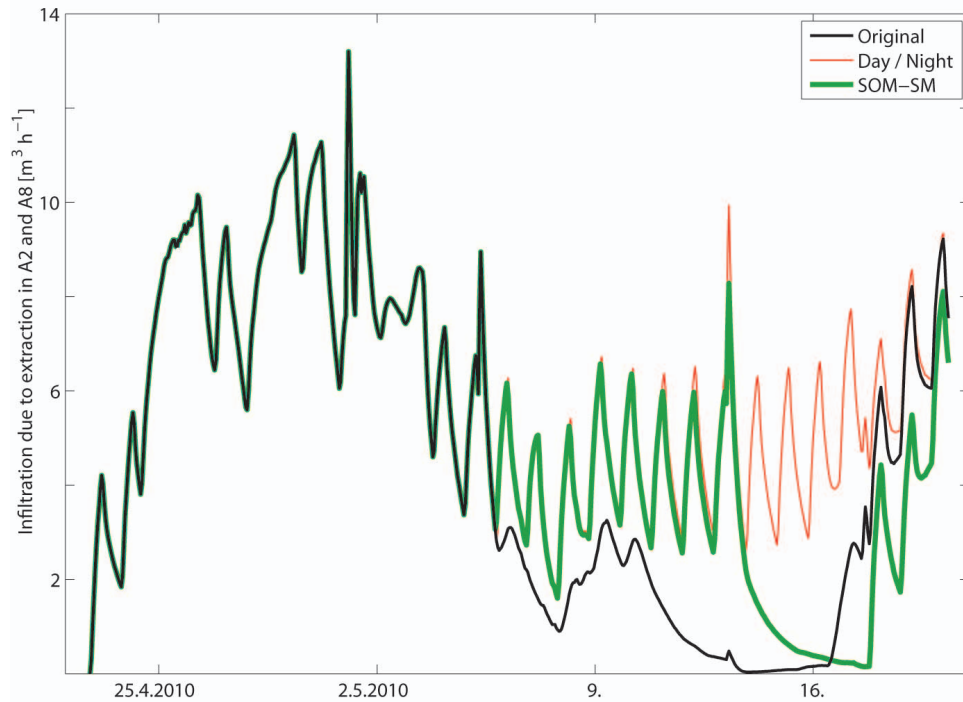
Figure 39: The time-series show river-water infiltration attributable to extraction in A2 and A8 in May 2010. The original scenario (actually run by water supplier) is shown in black. Day / Night (thin red line) is based on regular extraction during the night (normal regime). The SOM-SM approach (thick green line) allowed continued extraction until the 13. / 14. May, after which extraction is stopped for four days.

9.2.3 Conclusion Scenario Analysis

The simulation of the five pump-management scenarios for January 2009 and May 2010 illustrates the application of the SOM-SM analysis as a differentiated approach to identifying potential contamination due to river-water infiltration. As shown in Fig. 40, the two periods differed in terms of the strength of the riverine influence: high discharge events in the river showed a greater influence on groundwater head in January 2009 than they did in May 2010. In both situations the SOM-SM analysis recommended a change to the pumping regime to protect the two groundwater extraction wells susceptible to riverine contamination. The SOM-SM approach identified the point in time at which extraction should be stopped, and when it could be resumed. Rather than being based on a fixed value, i.e. river stage, the multivariate statistical method used the variation in the measured time-series to identify situations that were classed as *critical* (cf. Chapter 8). The recommended extraction based on the SOM-SM analysis led to a more conservative pumping regime in January 2009 and a briefer period with no extraction in May 2010. The particle tracks in Fig. 40 show that the extraction of fresh riverine water (in A8) in January 2009 could only be avoided by a complete cessation of extraction in A8. However, during the May 2010 situation, a more differentiated approach including a reduction of extraction, rather than complete cessation would also be possible and should be considered in the future. The current pump installation in A8 however, does not allow a graduated pump management, only *ON/OFF*.

The relative amount of river-water infiltration was used as a proxy for the extraction of river water during the two observation periods. As river-groundwater interaction in the study area is a dominant factor in determining groundwater quality and quantity, the proxy can be applied, however it cannot be used to assess potential contamination, which is also strongly dependent on water age and travel distance through the aquifer. To address these questions and design an individual warning system for each groundwater extraction well, a combined flow and transport model should be applied.

The results of the groundwater-flow modelling show that the SOM-SM

analysis and interpretation can be used to identify critical situations, however, further attention should be paid to the time when extraction can be resumed. At the onset of an event, the changes in parameter values, for example, rise in groundwater head or decrease in electrical conductivity, are mostly rapid, whereas, the return to non-critical situations takes longer and is not marked by clear or rapid changes. Under these circumstances, a groundwater-flow model and particle tracking is able to show the origin of the water extracted at a specific time. Groundwater-flow models can thus be used to support the validation of the SOM-SM analysis, especially in defining *retun-to-normal* situations.
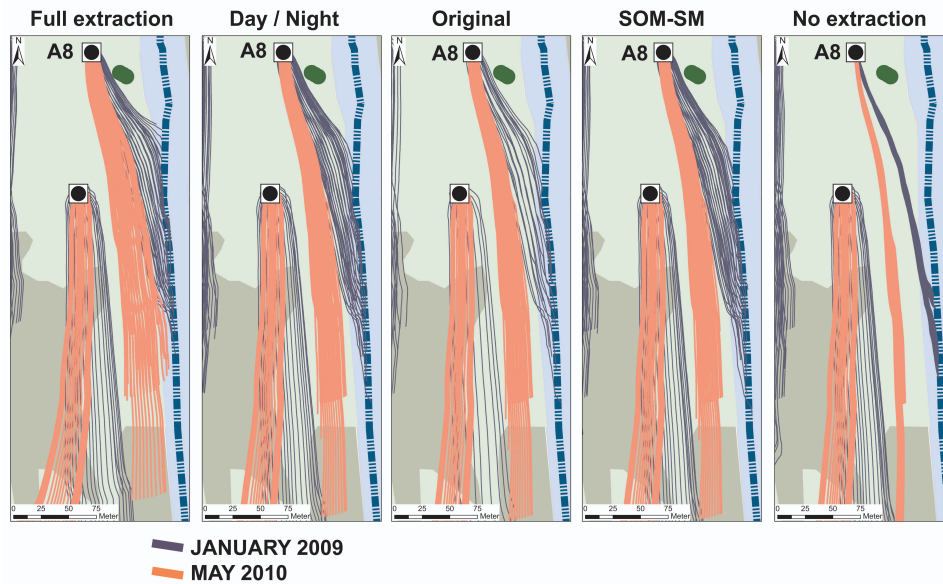
Figure 40: 10-day particle paths for all of the extraction wells in the wellfield for January 2009 (blue) and May 2010 (orange). A8 is the observation well most susceptible to contamination from riverine sources and can draw water directly from the river. The two situations show that infiltration of river water cannot be avoided, even with no extraction, however, the amount of infiltration and the 10-day capture zones can be adapted using a differentiated (SOM-SM) approach to pump management.

10 Synopsis

The aim of this project was to improve the protection of groundwater extraction wells near rivers and to increase the degrees of freedom available to the managment of fluvial planes with multiple stakeholders, including nature reserves, energy and drinking-water production and urban development. The individual chapters of this thesis describe the steps taken to develop an approach to identifying and assessing endangerment to drinking-water quality. This chapter summarizes the steps and describes how they can lead to a product, or methodology that can be used for similar projects in other well fields. A flowchart summarizing the steps is presented in Fig. 41. This chapter also outlines further research that can lead to multi-criteria decision support systems for drinking-water suppliers with extraction wells near rivers.

10.1 Systems Analysis

Knowledge of the principal processes governing the geological, hydrogeological, biological and environmental characteristics of the well field and surroundings, needs to be acquired prior to the installation of a monitoring network to track changes in system state and identify critical system states. The collation of all information available in a central database can help provide an overview and simplify data management. The simultaneous consideration of different sources of information, for example, event-oriented bacterial studies and geoelectrical measurements, can give essential insight into the functioning of the system under observation. The information can be summarized to give rise to system profiles for individual groundwater bodies, which can be used in multi-objective management analysis (Epting et al., 2008).

Besides the substantial knowledge available on the behaviour of viral and bacterial pathogens in groundwater, high-resolution, event-oriented sampling studies are important for the understanding of local behaviour and potential endangerment of groundwater extraction wells due to riverine contamination.

Process Dynamics
  - Sedimentological Studies
  - Geoelectrical / -physical Experiments
  - Bacterial Sampling

Monitoring Network
  - Observation Well Instrumentation
  - Telemetry
  - Database Options

Preliminary Analysis
  - Principal Component Analysis
  - Tracer Studies
  - Groundwater Flow Modelling

Define Indicator Parameters
and Observation Wells

Measurement

SOM-SM
Analysis

NO   Endangerment

YES

Issue Warning

Figure 41: Flowchart showing the individual steps taken to set up a ground-water quality monitoring system.

Groundwater-flow models can be used to further understand the processes occuring in the system, to test hypotheses and can provide valuable support when designing, interpreting and validating management scenarios for groundwater extraction wells. They are, however, also strongly dependent on the quantity and quality of the information available, as insufficient or incorrect knowledge and information can lead to erroneous results and therefore a poor basis for a management tool. Complete knowledge on processes or the heterogeneity of well-fields is not possible to obtain and uncertainties must be treated as such. Multiple objectives and individual pumping regimes for wells in a multi-well field require high-resolution optimization and transport models.

10.2 Measurement Network

The installation of new boreholes is costly and should only be carried out with knowledge of the processes affecting the system under consideration. While the proximity to mains power may be useful, it is not essential if the installation is intended to record and transmit for longer periods of time. Two limiting factors for transmission are the power supply and the placement of a transmitter antenna. In this case, the location of the boreholes permitted placing a car battery in a water-tight box and burying it in the soft ground. However, in urban environments, or areas where vandalism may be a problem, this solution is not applicable. The antenna requires a satellite connection and should not be obvious to potential vandals. There are solar-panel solutions available, these also require sufficient sunlight and a vandal-free zone. These factors also require consideration when identifying suitable positions for measurement installations. The use of existing boreholes is desirable, as long time-series encompassing different hydrological situations may be available. The number of boreholes actually used to identify critical situations can later be reduced so that only the wells most sensitive to rapid river-water infiltration are used in the analysis and pumping-well management.

The installation of measurement instruments at different depth levels,

as was done in this project, can provide useful information on groundwater quality at different depths of the aquifer. The use of three separate boreholes with filter stretches at different depths, e.g. W-cluster, can reflect the small-scale heterogeneity of the aquifer structure. The data collected at W2 showed a significantly different pattern in January 2009 than the time-series recorded in the other two wells only a few metres away (W1 and W3). The interpretation of the measurement data collected from this cluster relies on additional information, for example, from geoelectrical studies carried out by Wüest (2010).

Accordingly, not all observation wells are equally representative of the source of potential contamination, in this case river-water infiltration. The determination of the observation wells best suited for the hazard analysis can start once a data set is available. A principal component analysis provided information on the suitability of each observation well for the detection of infiltrating river water. The variance-based method helps to identify observation wells that respond rapidly and strongly to changes in groundwater quality and links the observed changes to their source (i.e. rapid river-water infiltration). By these means, the radius of influence of river-water infiltration can be determined and indicator wells identified. Besides the applicability to identifying indicator observation wells, the method can also be used to identify indicator parameters for detecting rapid river-water infiltration.

10.3 Critical Situations

After the indicator observation wells and indicator parameters have been selected, the time-series can be processed and critical situations, or system states, identified. Critical situations occur when rapid river water infiltration threatens to contaminate the groundwater extraction wells. The definition of their manifestation in the analysis (SOM-SM projection) is essential to the robustness and representativeness of this approach. This was done using a multivariate data set (groundwater head, temperature and electrical conductivity) from six observation wells. The combination of neural-network based self-organizing maps and the nonlinear Sammon's projection (SOM-

SM) used the variance between measurement points in time to create a representation of the similarity between system states. The information from each set of measurements at one point in time is summarized into one point in the projection and respresents the system state at that point in time. The complexity and the dimensionality of the data set is simplified and can be used to continuously assess actual groundwater quality influenced by infiltrating river water. The SOM-SM analysis, however, requires validation and further parameters, not used in the analysis, to identify the critical system states. The critical situations, or system states, were defined as situations when rapid river-water infiltration was detected and that, therefore, the potential for contamination of the drinking-water extraction wells existed.

Not all parameters used in the SOM-SM analysis in Chapter 8 were available online. Based on the results from Chapter 6, only W1 and W3 were equipped with an option for automatic data-transfer. As shown in Fig. 42, the use of only two observation wells significantly reduced the complexity of the SOM-SM projection. As less data was used in constructing the projection, less information on system state could be derived from the analysis and the projection became simpler. Based on only five variables, the trajectory through time is straighter, as it does not contain as much variation as the analysis based on 17 variables. While the scaling of the SOM-SM projection and the position of the points in the projection differ between panel A and panel B in Fig. 42, the development over time is similar in relation to the beginning and end points. In this case, the essential information concerning changes in system state is adequately captured by five variables as: a) the shape of trajectory in panel A (W1 and W3) is similar to that in panel B (all six observation wells), b) the trajectory develops in the same direction in both panels in Fig. 42, and c) the distances between the points in panel A are greater during the declining part of the trajectory (beginning marked with d in panel A).

The essential information necessary to issue a warning when a critical situation is detected can therefore be based on the two groundwater observation wells W1 and W3. However, a full set of variables, in this case the 17 time-series (three parameters and six observation wells), provides suffi-
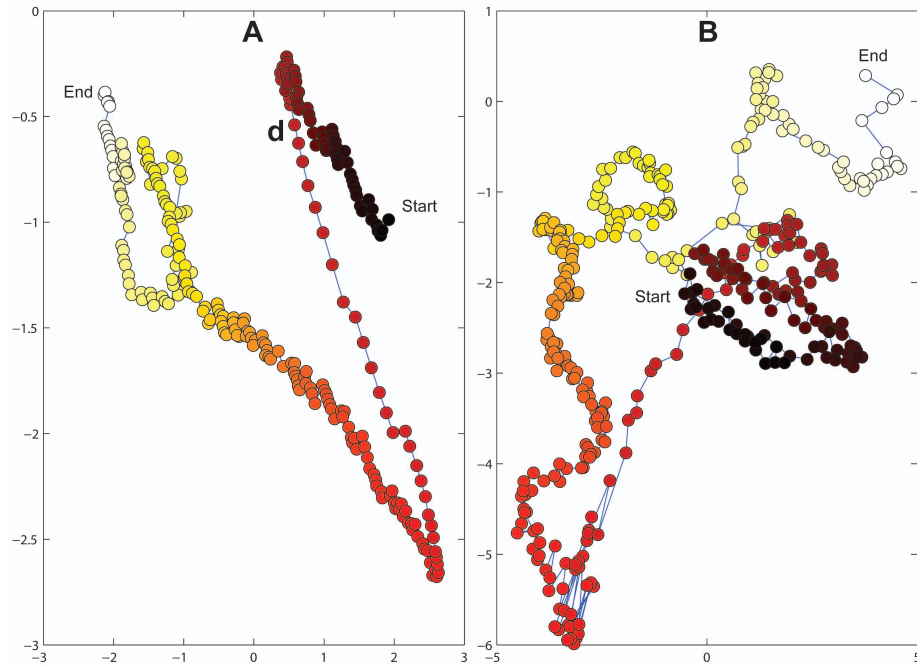
Figure 42: Both panels show the same time period, but with a different number of variables used for the analysis. Progressing time is visualized by the shading of the points (dark-light). Panel A shows a SOM-SM projection based on groundwater head, temperature and electrical conductivity from W1 and W3 (five variables). Panel B shows a SOM-SM projection based on groundwater head, temperature and electrical conductivity from all six observation wells (W1, W2, W3, B1, B2 and B3, 17 variables). Electrical conductivity measurements for W1 could not be used in the analysis.

127

cient information to assess the process dynamics and to verify the suitability and ability of the selected observation wells to represent the actual system state in terms of river-water infiltration. A preliminary assessment therefore encompasses an analysis of all parameters available (measurements and observation wells), in combination, for example, with a principal component analysis. Following the preliminary analysis and feasibility study for the telemetry installation, the time-series from selected indicator observation wells, W1 and W3, can be used to assess actual groundwater quality.

10.4 Alerting and Feedback Water Supply

The next step was to create a database to manage the data from the different sources. Information on actual pumping has to be available in order to be able to issue warnings for individual pumps and assess the effect of a management operation (alteration of pumping regime). The system will take some time to respond after a warning has been issued. The response time is a function of operational aspects, i.e. time lag between issue of warning and pump regime change, and aquifer response. The second aspect, aquifer response, depends on the hydrological situation, including hydraulic gradients between the river and the aquifer. Issuing a further warning should be avoided until the effects of the initial warning can be observed.

Provided a graduated extraction pattern is possible (both multiple pumps and graduated volumes for each individual pump), a multi-objective optimization model can be used to assess the different options available to the pump operator. The water supply pump operator can thus be supported in their decision to a) continue extraction, b) stop extraction or c) throttle extraction for each, all, or a combination of the extraction wells in the well-field.

11 Conclusion

Multivariate analysis methods, for example, principal component analysis, show a strong ability to identify patterns and relationships between different parameters (Dominguez et al., 2007; Mustonen et al., 2008; Corona et al., 2010; Ündey et al., 2010). Multivariate analysis methods have frequently been used in groundwater quality assessment (Helena et al., 2000; Lewandowski et al., 2009; Lischeid, 2009), but, process control requires information derived from real-time monitoring data. The application of multivariate alert, or warning systems have been in use in industrial process monitoring for some years (Dominguez et al., 2007). Recently, such applications have been extended to drinking-water distribution networks and used in connection with disease outbreaks (Mustonen et al., 2008; Kroll and King, 2010; Burkom et al., 2011; Koch and McKenna, 2011). The US Environmental Protection Agency has recently issued a software to alert water suppliers following the detection of contamination in distribution networks based on multivariate time-series (CANARY, Murray et al. (2010)). The proposed concept in this thesis extends the area of protection, i.e. the number of critical control points in an HACCP approach for drinking-water management, to include the resource groundwater. By applying such an adaptive groundwater-extraction management scheme, source protection is included and a multibarrier concept can be implemented. The complexity arising due to heterogenous river-groundwater interaction, is substantially reduced, so that the proposed methodology can either be used by the water supplier directly, or integrated into an automated control system.

The methodology presented in this thesis aims to identify and assess potential contamination of groundwater extraction wells due to infiltrating river water. The approach is based on multivariate analysis of groundwater time-series from several observation wells representing different zones and depths of the aquifer. The complexity of multiple time-series, each showing different behaviour during and between high discharge events, is reduced by using the combined approach of self-organizing maps and Sammon's projections (SOM-SM). While reducing the dimensionality, the inherent infor-

mation is preserved so that the approach enables a real-time assessment of groundwater quality in a model-free environment. Although the actual assessment of groundwater quality based on system state analysis does not require iterative groundwater flow or transport simulations, the validation of the method should be based on other methods used in hydrogeological investigations and independent time-series. Two of these methods were applied in this project: 1) microbial, event-oriented sampling, and 2) groundwater-flow modelling. The microbial sampling showed that individual parameters are often poor indicators of groundwater quality. The bi-weekly sampling campaign was subject to a strong *hit-and-miss* characteristic identified during the high-resolution sampling campaign. The choice of observation wells sampled can be optimized by using principal component analysis (PCA) to indentify the dominant processes at each observation well and determine the wells most sensitive to river-water infiltration. Further microbial sampling campaigns should include a range of high discharge events, while focussing on the observation wells most representative of river-water infiltration. Ideally, a real-time SOM-SM analysis is carried out simultaneously and help determine the microbial sampling frequency. The groundwater flow simulations were used to test management options derived from the SOM-SM analysis. The application of the flow model was effective at showing the consequences of different management options and highlighting additional degrees of freedom open to the water supplier during high discharge events. The shifts in system state identified in the SOM-SM analysis, i.e. the time when extraction should be stopped, was presented to the water supplier as a catalogue of scenarios simulated using the groundwater-flow model, thus providing a better understanding of the SOM-SM approach to adaptive groundwater management. The measurement setup, including resolution, and choice of observation wells, provided the SOM-SM analysis with information on actual groundwater quality in the aquifer. As soon as the warning is issued, the water supplier is able to act to prevent contamination of the extraction wells. This precautionary approach provides the water supplier with more degrees of freedom, including more time, than other measures based on measurements in the extraction wells, for example, turbidity thresholds.

Further attention should be paid to the real-time updating and validation of the SOM-SM projection. While the combined analysis of the three situations (cf. Chapter 8) indicates consistency in the critical system state location in the projection, additional high discharge events including high-resolution microbial sampling will help to fine-tune the warning and its recall (*return-to-normal*) issued to the water supplier. The advancement of telecommunication systems and data-management tools enables rapid development of real-time, multiparameter monitoring systems. Additional fault detection steps may need to be incorporated in the real-time analysis to prevent measurement or transmission errors falsifying the SOM-SM projection (Murray et al., 2010).

The proposed methodology can be included as a first step (source protection) in a multibarrier approach of an adaptive water management, or decision support system. The adaptability is derived from the flexibility of the method in terms of parameters used, field of employment and the independence of measurement thresholds. The parameters used in this project (groundwater head, temperature and electrical conductivity) may be equally significant in other areas, however they can be easily replaced by other parameters to suit the prevalent conditions. The application of the methodology can be extended to include other measured time-series, for example, pipe pressure recorded in the distribution network. As the SOM-SM approach is not dependent on individual parameters, or pre-defined thresholds, it can be transferred to other well-fields and experimental field sites to address similar questions.

12 References

Affolter, A., P. Huggenberger, S. Scheidler, and J. Epting, 2010: Adaptive groundwater management in urban areas: Effect of surface water-groundwater interaction using the example of artificial groundwater recharge and in- and exfiltration of the river Birs (Switzerland). *GRUNDWASSER*, **15 (3)**, 147–161.

Allen, D. J., W. G. Darling, D. Goody, D. Lapworth, A. Newell, A. Williams, D. Allen, and C. Abesser, 2010: Interaction between groundwater, the hyporheic zone and a Chalk stream: a case study from the River Lambourn, UK. *HYDROGEOLOGY JOURNAL*, **18 (5)**, 1125–1141.

Arntzen, E. V., D. R. Geist, and P. E. Dresel, 2006: Effects of fluctuating river flow on groundwater/surface water mixing in the hyporheic zone of a regulated, large cobble bed river. *RIVER RESEARCH AND APPLICATIONS*, **22 (8)**, 937–946.

Auckenthaler, A., G. Raso, and P. Huggenberger, 2002: Particle transport in a karst aquifer: natural and artificial tracer experiments with bacteria, bacteriophages and microspheres. *WATER SCIENCE AND TECHNOLOGY*, **46 (3)**, 131–138.

BAG, 2004: *Schweizerisches Lebensmittelbuch*. Bundesamt für Gesundheit.

Bauser, G., H.-J. H. Franssen, H.-P. Kaiser, U. Kuhlmann, F. Stauffer, and W. Kinzelbach, 2010: Real-Time Management of an Urban Groundwater Well Field Threatened by Pollution. *ENVIRONMENTAL SCIENCE & TECHNOLOGY*, **44 (17)**, 6802–6807.

BAZ, 2011: 15 Millionen für neue Trinkwasseranlage. http://bazonline.ch/, Article 16232997.

BBC, 2011: Cardiff University-led team finds New Delhi water bug. http://www.bbc.co.uk/news/uk-wales-south-east-wales-12989782.

Bernataviciene, J., G. Dzemyda, O. Kurasova, and V. Marcinkevicius, 2006: Optimal decisions in combining the SOM with nonlinear projection methods. *EUROPEAN JOURNAL OF OPERATIONAL RESEARCH*, **173 (3)**, 729–745.

Bitterli-Brunner, P., H. Fisher, and H. P., 1984: Geologischer Atlas der Schweiz (1:25.000). Blatt 1067 Arlesheim.

Borchardt, M. ., N. Haas, and R. Hunt, 2004: Vulnerability of drinking-water wells in La Crosse, Wisconsin, to enteric-virus contamination from surface water contributions. *APPLIED AND ENVIRONMENTAL MICROBIOLOGY*, **70 (10)**, 5937–5946.

Brookes, J., M. Hipsey, M. Burch, R. Regel, L. Linden, C. Ferguson, and J. Antenucci, 2005: Relative value of surrogate indicators for detecting pathogens in lakes and reservoirs. *ENVIRONMENTAL SCIENCE & TECHNOLOGY*, **39 (22)**, 8614–8621.

Burkom, H. S., L. Ramac-Thomas, S. Babin, R. Holtry, Z. Mnatsakanyan, and C. Yund, 2011: An integrated approach for fusion of environmental and human health data for disease surveillance. *STATISTICS IN MEDICINE*, **30 (5, Sp. Iss. SI)**, 470–479.

Camplani, M., B. Cannas, A. Fanni, G. Pautasso, G. Sias, P. Sonato, and Asdex Upgrade Team, 2009: Tracking of the Plasma States in a Nuclear Fusion Device Using SOMs. *ENGINEERING APPLICATIONS OF NEURAL NETWORKS, PROCEEDINGS*, Brown, DP and Draganova, C and Pimenidis, E and Mouratidis, H, Ed., SPRINGER-VERLAG BERLIN, Communications in Computer and Information Science, Vol. 43, 430–437, 11th International Conference on Engineering Applications of Neural Networks, London, ENGLAND, AUG 27-29, 2009.

Cattell, R. B., 1966: The scree test for the number of factors. *MULTIVARIATE BEHAVIOURAL RESEARCH*, **1(2)**, 245–276.

Cirpka, O. A., M. N. Fienen, M. Hofer, E. Hoehn, A. Tessarini, R. Kipfer, and P. K. Kitanidis, 2007: Analyzing bank filtration by deconvoluting time series of electric conductivity. *GROUND WATER*, **45 (3)**, 318–328.

Cliff, N., 1988: The Eigenvalues-greater-than-one rule and the reliability of components. *PSYCHOLOGICAL BULLETIN*, **103 (2)**, 276–279.

Codex, 1997: *Codex Alimentarius*. Codex Alimentarius Commission, cac/rcp 1-1969, rev.3 (1997), amended 1999, second ed.

Conant, B., 2004: Delineating and quantifying ground water discharge zones using streambed temperatures. *GROUND WATER*, **42 (2)**, 243–257.

Corona, F., M. Mulas, R. Baratti, and J. A. Romagnoli, 2010: On the topological modeling and analysis of industrial process data using the SOM. *COMPUTERS & CHEMICAL ENGINEERING*, **34 (12, Sp. Iss. SI)**, 2022–2032.

Dahm, C., N. Grimm, P. Marmonier, H. Valett, and P. Vervier, 1998: Nutrient dynamics at the interface between surface waters and groundwaters. *FRESHWATER BIOLOGY*, **40 (3)**, 427–451, Workshop on River in the Landscape: Riparian and Groundwater Ecology, KASTANIENBAUM, SWITZERLAND, FEB 05-11, 1996.

Dash, R. R., E. V. P. B. Prakash, P. Kumar, I. Mehrotra, C. Sandhu, and T. Grischek, 2010: River bank filtration in Haridwar, India: removal of turbidity, organics and bacteria. *HYDROGEOLOGY JOURNAL*, **18 (4)**, 973–983.

Dizer, H., G. Grutzmacher, H. Bartel, H. Wiese, R. Szewzyk, and J. Lopez-Pila, 2004: Contribution of the colmation layer to the elimination of coliphages by slow sand filtration. *WATER SCIENCE AND TECHNOLOGY*, **50 (2)**, 211–214.

Doherty, J., 2005: *PEST - Model-Independent Parameter Estimation*. Watermark Numerical Computing, 5th ed.

Dominguez, M., J. J. Fuertes, P. Reguera, I. Diaz, and A. A. Cuadrado, 2007: Internet-based remote supervision of industrial processes using self-organizing maps. *ENGINEERING APPLICATIONS OF ARTIFICIAL INTELLIGENCE*, **20 (6)**, 757–765.

Duhigg, C., 2009: That tap water is legal but may be unhealthy. http://www.nytimes.com/2009/12/17/us/17water.html.

Epting, J., P. Huggenberger, and L. Glur, 2009: Integrated investigations of karst phenomena in urban environments. *ENGINEERING GEOLOGY*, **109 (3-4)**, 273–289, doi:10.1016/j.enggeo.2009.08.013.

Epting, J., C. Regli, and P. Huggenberger, 2008: Groundwater protection in urban areas incorporating adaptive groundwater monitoring and management - Reconciliation of water engineering measures along rivers. *Adaptive and Integrated Water Management: Coping with Complexity and Uncertainty*, Wostl, CP and Kabat, P and Moltgen, J, Ed., SPRINGER-VERLAG BERLIN, HEIDELBERGER PLATZ 3, D-14197 BERLIN, GERMANY, 97–123, International Conference on Adaptive and Integrative Water Management, Basel, SWITZERLAND, NOV 12-15, 2007.

Esty, D., M. Levy, T. Srebotnjak, A. deSherbinin, and C. Kim, 2006: *Pilot 2006 Environmental Performance Index*. New Haven: Yale Center for Environmental Law & Policy.

Figueras, M. J. and J. Borrego, 2010: New Perspectives in Monitoring Drinking Water Microbial Quality. *INTERNATIONAL JOURNAL OF ENVIRONMENTAL RESEARCH AND PUBLIC HEALTH*, **7 (12)**, 4179–4202.

Fisher, K. and C. Phillips, 2009: The ecology, epidemiology and virulence of Enterococcus. *MICROBIOLOGY*, **155**, 1749–1757.

Foster, S. and P. Chilton, 2003: Groundwater: the processes and global significance of aquifer degradation. *PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY OF LONDON SERIES B-BIOLOGICAL SCIENCES*, **358 (1440)**, 1957–1972.

Frind, E. O., J. W. Molson, and D. L. Rudolph, 2006: Well vulnerability: A quantitative approach for source water protection. *GROUND WATER*, **44 (5)**, 732–742.

Fuertes, J. J., M. Dominguez, P. Reguera, M. A. Prada, I. Diaz, and A. A. Cuadrado, 2010: Visual dynamic model based on self-organizing maps for supervision and fault detection in industrial processes. *ENGINEERING APPLICATIONS OF ARTIFICIAL INTELLIGENCE*, **23 (1)**, 8–17.

Furlow, B., 2005: Clean drinking water: a double-edged sword? *NEW SCIENTIST*, **2528**, 47–49.

Gerbersdorf, S., T. Jancke, B. Westrich, and D. M. Paterson, 2008: Microbial stabilization of riverine sediments by extracellular polymeric substances. *GEOBIOLOGY*, **6 (1)**, 57–69.

Gerbrands, J., 1981: On the Relationships between SVD, KLT and PCA. *PATTERN RECOGNITION*, **14 (1-6)**, 375–381.

Glur, L., 2008: Geophysikalische Methoden zur Erkundung Fluvialer Ablagerungen und von Karstphaenomenen. M.S. thesis, Applied and Environmental Geology, University of Basel.

Goldscheider, N., M. Pronk, and J. Zopfi, 2010: New insights into the transport of sediments and microorganisms in karst groundwater by continuous monitoring of particle-size distribution. *GEOLOGIA CROATICA*, **63 (2)**, 137–142.

Grisey, E., E. Belle, J. Dat, J. Mudry, and L. Aleya, 2010: Survival of pathogenic and indicator organisms in groundwater and landfill leachate

through coupling bacterial enumeration with tracer tests. *DESALINA-TION*, **261 (1-2)**, 162–168.

Gronewold, A. D., L. Myers, J. L. Swall, and R. T. Noble, 2011: Addressing uncertainty in fecal indicator bacteria dark inactivation rates. *WATER RESEARCH*, **45 (2)**, 652–664.

Gruden, C., S. Skerlos, and P. Adriaens, 2004: Flow cytometry for microbial sensing in environmental sustainability applications: current status and future prospects. *FEMS MICROBIOLOGY ECOLOGY*, **49 (1)**, 37–49, 5th International Symposium on Subsurface Microbiology, Copenhagen, DENMARK, SEP 08-13, 2002.

Gupta, V., W. P. Johnson, P. Shafieian, H. Ryu, A. Alum, M. A. adn S. A. Hubbs, and T. Rauch-Williams, 2009: Riverbank Filtration: Comparison of Pilot Scale Transport with Theory. *ENVIRONMENTAL SCIENCE & TECHNOLOGY*, **43**, 669–676.

Hamilton, P. D., P. Gale, and S. J. T. Pollard, 2006: A commentary on recent water safety initiatives in the context of water utility risk management. *ENVIRONMENT INTERNATIONAL*, **32 (8)**, 958–966.

Havelaar, A., 1994: Application of a HACCP to Drinking-Water Supply. *FOOD CONTROL*, **5 (3)**, 145–152.

Helena, B., R. Pardo, M. Vega, E. Barrado, J. Fernandez, and L. Fernandez, 2000: Temporal evolution of groundwater composition in an alluvial aquifer (Pisuerga River, Spain) by principal component analysis. *WATER RESEARCH*, **34 (3)**, 807–816.

Hoehn, E. and O. A. Cirpka, 2006: Assessing residence times of hyporheic ground water in two alluvial flood plains of the Southern Alps using water temperature and tracers. *HYDROL. EARTH SYST. SCI.*, **10**, 553–563.

Hooper, R., 2009: Top 11 compounds in us drinking water. http://www.newscientist.com, Article dn16397.

Hrudey, S. E., E. J. Hrudey, and S. J. T. Pollard, 2006: Risk management for assuring safe drinking water. *ENVIRONMENT INTERNATIONAL*, **32**, 948–957.

Huggenberger, P., J. Epting, I. Spottke, C. Regli, and E. Zechner, 2006: INTERREG III A-Projekt MoNit: Modellierung der Grundwasserbelas-

tung durch Nitrat im Oberrheingraben. Landesanstalt für Umwelt, Messungen und Naturschutz Baden-Würtemberg. Teilprojekte Nitratherkunft (Nitrat-Transportmodellierung) und Fluss-Grundwasser-Interaktion (regionale hydrologische Grundlagedaten), Landesanstalt für Umwelt, Messungen und Naturschutz Baden-Würtemberg. Teilprojekte Nitratherkunft (Nitrat-Transportmodellierung) und Fluss-Grundwasser-Interaktion (regionale hydrologische Grundlagedaten).

Huggenberger, P., E. Hoehn, R. Beschta, and W. Woessner, 1998: Abiotic aspects of channels and floodplains in riparian ecology. *FRESHWATER BIOLOGY*, **40**, 407–425.

Jackson, D. A., 1993: Stopping Rules in Principal Component Analysis: A Comparison of Heuristical and Statistical Approaches. *ECOLOGY*, **74 (8)**, 2204–2214.

Joerin, F., G. Cool, M. J. Rodriguez, M. Gignac, and C. Bouchard, 2010: Using multi-criteria decision analysis to assess the vulnerability of drinking water utilities. *ENVIRONMENTAL MONITORING AND ASSESSMENT*, **166 (1-4)**, 313–330.

John, D. and J. Rose, 2005: Review of factors affecting microbial survival in groundwater. *ENVIRONMENTAL SCIENCE & TECHNOLOGY*, **39 (19)**, 7345–7356.

Kaiser, H., 1960: The Application of Electronic-computers to Factor-analysis. *EDUCATIONAL AND PSYCHOLOGICAL MEASUREMENT*, **20 (1)**, 141–151.

Keery, J., A. Binley, N. Crook, and J. W. N. Smith, 2007: Temporal and spatial variability of groundwater-surface water fluxes: Development and application of an analytical method using temperature time series. *JOURNAL OF HYDROLOGY*, **336 (1-2)**, 1–16.

Koch, M. W. and S. A. McKenna, 2011: Distributed Sensor Fusion in Water Quality Event Detection. *JOURNAL OF WATER RESOURCES PLANNING AND MANAGEMENT-ASCE*, **137 (1)**, 10–19.

Kohonen, T., 2001: *Self-Organizing Maps*. Springer.

Kohonen, T., E. Oja, O. Simula, A. Visa, and J. Kangas, 1996: Engineering applications of the self-organizing map. *PROCEEDINGS OF THE IEEE*, **84 (10)**, 1358–1384.

Kolehmainen, M., P. Ronkko, and A. Raatikainen, 2003: Monitoring of yeast fermentation by ion mobility spectrometry measurement and data visualisation with Self-Organizing Maps. *ANALYTICA CHIMICA ACTA*, **484 (1)**, 93–100.

Krause, S., A. Bronstert, and E. Zehe, 2007: Groundwater-surface water interactions in a North German lowland floodplain - Implications for the river discharge dynamics and riparian water balance. *JOURNAL OF HYDROLOGY*, **347**, 404–417.

Kroll, D. and K. King, 2010: Methods for evaluating water distribution network early warning systems. *JOURNAL AMERICAN WATER WORKS ASSOCIATION*, **102 (1)**, 79+.

Kukkula, M., P. Arstila, M. Klossner, L. Maunula, C. vonBonsdorff, and P. Jaatinen, 1997: Waterborne outbreak of viral gastroenteritis. *SCANDINAVIAN JOURNAL OF INFECTIOUS DISEASES*, **29 (4)**, 415–418.

Lautenbach, S., J. Berlekamp, N. Graf, R. Seppelt, and M. Matthies, 2009: Scenario analysis and management options for sustainable river basin management: Application of the Elbe DSS. *ENVIRONMENTAL MODELLING & SOFTWARE*, **24**, 26–43.

Lerner, B., H. Guterman, M. Aladjem, I. Dinstein, and Y. Romem, 1998: On pattern classification with Sammon's nonlinear mapping - An experimental study. *PATTERN RECOGNITION*, **31 (4)**, 371–381.

Lewandowski, J., G. Lischeid, and G. Nützmann, 2009: Drivers of water level fluctuations and hydrological exchange between groundwater and surface water at the lowland River Spree (Germany): field study and statistical analyses. *HYDROLOGICAL PROCESSES*, **23 (15, Sp. Iss. SI)**, 2117–2128, General Assembly of the European-Geosciences-Union, Vienna, AUSTRIA, APR 14-18, 2008.

Lin, B., M. Syed, and R. A. Falconer, 2008: Predicting faecal indicator levels in estuarine receiving waters - An integrated hydrodynamic and ANN modelling approach. *ENVIRONMENTAL MODELLING & SOFTWARE*, **23 (6)**, 729–740.

Lischeid, G., 2009: Non-linear visualization and analysis of large water quality data sets: a model-free basis for efficient monitoring and risk assessment. *STOCHASTIC ENVIRONMENTAL RESEARCH AND RISK ASSESSMENT*, **23 (7, Sp. Iss. SI)**, 977–990.

Lischeid, G. and J. Bittersohl, 2008: Tracing biogeochemical processes in stream water and groundwater using non-linear statistics. *JOURNAL OF HYDROLOGY*, **357 (1-2)**, 11–28.

Longuevergne, L., N. Florsch, and P. Elsass, 2007: Extracting coherent regional information from local measurements with Karhunen-Loeve transform: Case study of an alluvial aquifer (Rhine valley, France and Germany). *WATER RESOURCES RESEARCH*, **43 (4)**, W04 430.

Marrone, B. L., 2009: Flow Cytometry: A Multipurpose Technology for a Wide Spectrum of Global Biosecurity Applications. *JALA*, **14 (3, Sp. Iss. SI)**, 148–156.

Martinez, S., F. Xavier Munoz, and E. Baldrich, 2010: Inductive microcoils for the fast and simple detection of bacterial presence. *SENSORS AND ACTUATORS B-CHEMICAL*, **147 (1)**, 304–309.

McDonald, M., A. Harbaugh, E. Banta, and M. Hill, 2000: *Modflow-2000, The U.S. geological survey modular ground-water flow user guide to modularization concepts and the ground-water flow process.* US Geological Survey, Open File Report 00-92.

McDowell-Boyer, L. M., J. R. Hunt, and N. Sitar, 1986: Particle Transport through Porous Media. *WATER RESOURCES RESEARCH*, **22(13)**, 1901–1921.

McFeters, G., G. Bissione, J. Jezeski, C. Thomson, and D. Stuart, 1974: Comparative Survival of Indicator Bacteria and Enteric Pathogens in Well Water. *APPLIED MICROBIOLOGY*, **27 (5)**, 823–829.

McKergow, L. A. and R. J. Davies-Colley, 2010: Stormflow dynamics and loads of Escherichia coli in a large mixed land use catchment. *HYDROLOGICAL PROCESSES*, **24 (3)**, 276–289.

Murray, R., et al., 2010: Water Quality Event Detection System for Drinking Water Contamination Warning Systems: Development, Testing and Application of CANARY. National Homeland Security Research Centre, Office of Research and Development, U.S. Environmental Protection Agency.

Mustonen, S. M., S. Tissari, L. Huikko, M. Kolehmainen, M. J. Lehtola, and A. Hirvonen, 2008: Evaluating online data of water quality changes in a pilot drinking water distribution system with multivariate data exploration methods. *WATER RESEARCH*, **42 (10-11)**, 2421–2430.

Mutiti, S. and J. Levy, 2010: Using temperature modeling to investigate the temporal variability of riverbed hydraulic conductivity during storm events. *JOURNAL OF HYDROLOGY*, **388 (3-4)**, 321–334.

Nichols, G., C. Lane, N. Asgari, N. Q. Verlander, and A. Charlett, 2009: Rainfall and outbreaks of drinking water related disease and in England and Wales. *JOURNAL OF WATER AND HEALTH*, **7 (1)**, 1–8.

Nnane, D., J. Ebdon, and H. D. Taylor, 2011: Integrated analysis of water quality parameters for cost-effective faecal pollution management in river catchments. *WATER RESEARCH*, **45**, 2235–2246.

Page, R., P. Huggenberger, J. Epting, and G. Lischeid, 2011: Risk analysis for riverine groundwater extraction. *GQ10: Groundwater Quality Management in a Rapidly Changing World*, IAHS Publ., Vol. 342.

Page, R. M. and J. Simovic, 2011: *Statistical Analysis of Monitoring Data. In: Urban Geology. Eds P. Huggenberger and J. Epting*, chap. 4.5, 86–93. Springer.

Payment, P., M. Waite, and A. Dufour, 2003: *Introducing parameters for the assessment of drinking water quality*, chap. 2, 47–77. IWA Publishing.

Peters, N. E., 2009: Effects of urbanization on stream water quality in the city of Atlanta, Georgia, USA. *HYDROLOGICAL PROCESSES*, **23**, 2860–2878.

Polat, E., 2010: The transport of noroviruses and indicator microorganisms depending on the interaction between river and groundwater during rainfall events. M.S. thesis, Infection Biology and Epidemiology, University of Basel.

Pronk, M., N. Goldscheider, and J. Zopfi, 2007: Particle-size distribution as indicator for fecal bacteria contamination of drinking water from karst springs. *ENVIRONMENTAL SCIENCE & TECHNOLOGY*, **41 (24)**, 8400–8405.

Regli, C., M. Rauber, and P. Huggenberger, 2003: Analysis of aquifer heterogeneity within a well capture zone, comparison of model data with field experiments: A case study from the river Wiese, Switzerland. *AQUATIC SCIENCES*, **65 (2)**, 111–128.

Reisner, A., K. A. Krogfelt, B. M. Klein, E. L. Zechner, and S. Molin, 2006: In Vitro Biolfilm Formation of Commensal and Pathogenic Escherichia coli Strains: Impact of Environmental and Genetic Factors. *JOURNAL OF BACTERIOLOGY*, **188 (10)**, 3572–3581.

Rogers, L. L. and F. U. Dowla, 1994: Optimization of groundwater remediation using artificial neural networks with parallel solute transport modelling. *WATER RESOURCES RESEARCH*, **30 (2)**, 457–481.

Rosenberg, D. E., T. Tarawneh, R. Abdel-Khaleq, and J. R. Lund, 2007: Modeling integrated water user decisions in intermittent supply systems. *WATER RESOURCES RESEARCH*, **43**, W07 425.

Sammon, J., 1969: A Nonlinear Mapping for Data Structure Analysis. *IEEE TRANSACTIONS ON COMPUTERS*, **C-18 (5)**, 401–409.

Sanchez-Martos, F., R. Jimenez-Espinosa, and A. Pulido-Bosch, 2001: Mapping groundwater quality variables using PCA and geostatistics: a case study of Bajo Andarax, southeastern Spain. *HYDROLOGICAL SCIENCES JOURNAL*, **46(2)**, 227–242.

Savenkoff, C., D. P. Swain, J. M. Hanson, M. Castonguay, M. O. Hammill, H. Bourdages, L. Morissette, and D. Chabot, 2007: Effects of fishing and predation in a heavily exploited ecosystem: Comparing periods before and after the collapse of groundfish in the southern Gulf of St. Lawrence (Canada). *ECOLOGICAL MODELLING*, **204 (1-2)**, 115–128.

Schmassmann, H.-J., 1981: Das Naturschutzgebiet Reinacherheide (Reinach, Basel-Landschaft). Special Issue of *Tätigkeitsberichte der Naturforschenden Gesellschaft Baselland*, Volume 31, Liestal, Switzerland.

Schmidt, C., B. Conant Jr, M. Bayer-Raich, and S. M., 2007: Evaluation of field-scale application of an analytical method to quantify groundwater discharge using mapped streambed temperatures. *JOURNAL OF HYDROLOGY*, **347**, 292–307.

Schwarzenbach, R. P., B. I. Escher, K. Fenner, T. B. Hofstetter, C. A. Johnson, U. von Gunten, and B. Wehrli, 2006: The Challenge of Micropollutants in Aquatic Systems. *SCIENCE*, **313**, 1072 – 1077.

Selker, J., N. van de Giesen, M. Westhoff, W. Luxemburg, and M. B. Parlange, 2006: Fiber optics opens window on stream dynamics. *GEOPHYSICAL RESERACH LETTERS*, **33**, L24 401.

141

Serrano, S. E. and S. R. Workman, 1998: Modeling transient stream/aquifer interaction with the non-linear Boussinesq equation and its analytical solution. *JOURNAL OF HYDROLOGY*, **206**, 245–255.

Sheets, R., R. Darner, and B. Whitteberry, 2002: Lag times of bank filtration at a well field, Cincinnati, Ohio, USA. *JOURNAL OF HYDROLOGY*, **266 (3-4)**, 162–174.

Sklash, M. G., M. K. Stewart, and P. A. J., 1986: Storm runoff generation in humid headwater catchments. 2. A case study of hillslope and low-order stream response. *WATER RESOURCES RESEARCH*, **22(8)**, 1273–1282.

Sophocleous, M., 2002: Interactions between groundwater and surface water: the state of the science. *HYDROGEOLOGY JOURNAL*, **10**, 52–67.

Stadler, H., E. Klock, P. Skritek, R. L. Mach, W. Zerobin, and A. H. Farnleitner, 2010: The spectral absorption coefficient at 254 nm as a real-time early warning proxy for detecting faecal pollution events at alpine karst water resources. *WATER SCIENCE AND TECHNOLOGY*, **62 (8)**, 1898–1906.

Stanford, J. A. and J. V. Ward, 1993: An ecosystem perspective of alluvial rivers: connectivity adn the hyporheic corridor. *JOURNAL OF NORTH AMERICAN BENTHOLOGICAL SOCIETY*, **12 (1)**, 48–60.

Su, G. W., J. Jasperse, D. Seymour, J. Constantz, and Q. Zhou, 2007: Analysis of pumping-induced unsaturated regions beneath a perennial river. *WATER RESOURCES RESEARCH*, **43**, W08 421.

Swiss Federal Legislation, 2005: SR 817.02 Lebensmittel- und Gebrauchsgegenständeverordnung (LGV). Art. 49, Art. 49.

Taylor, R., A. Cronin, S. Pedley, J. Barker, and T. Atkinson, 2004: The implications of groundwater velocity variations on microbial transport and wellhead protection - review of field evidence. *FEMS MICROBIOLOGY ECOLOGY*, **49 (1)**, 17–26, 5th International Symposium on Subsurface Microbiology, Copenhagen, DENMARK, SEP 08-13, 2002.

ten Veldhuis, J. A. E., F. H. L. R. Clemens, G. Sterk, and B. R. Berends, 2010: Microbial risks associated with exposure to pathogens in contaminated urban flood water. *WATER RESEARCH*, **44 (9)**, 2910–2918.

Unc, A. and M. J. Goss, 2003: Movement of faecal bacteria through the vadose zone. *WATER, AIR AND SOIL POLLUTION*, **149**, 327–337.

Ündey, C., S. Ertunc, T. Mistretta, and B. Looze, 2010: Applied advanced process analytics in biopharmaceutical manufacturing: Challenges and prospects in real-time monitoring and control. *JOURNAL OF PROCESS CONTROL*, **20 (9)**, 1009–18.

UNESCO, 2009: *The United Nations World Water Development Report 3: Water in a Changing World.* Paris: UNESCO Publishing and London: Earthscan.

Vesanto, J. and E. Alhoniemi, 2000: Clustering of the self-organizing map. *IEEE TRANSACTIONS ON NEURAL NETWORKS*, **11 (3)**, 586–600.

Vesanto, J., J. Himberg, E. Alhoniemi, and J. Parhankangas, 2000: *SOM Toolbox for Matlab 5.* P.O. Box 5400, FIN-03015 HUT, Finland, SOM Toolbox Team, Helsinki University of Technology.

Vogt, T., P. Schneider, L. Hahn-Woernle, and O. A. Cirpka, 2010: Estimation of seepage rates in a losing stream by means of fiber-optic high-resolution vertical temperature profiling. *JOURNAL OF HYDROLOGY*, **380 (1-2)**, 154–164.

Wilkinson, J., A. Jenkins, M. Wyer, and D. Kay, 1995: Modeling Faecal-coliform Dynamics in Streams and Rivers. *WATER RESEARCH*, **29 (3)**, 847–855.

Wilkinson, J., D. Kay, M. Wyer, and A. Jenkins, 2006: Processes driving the episodic flux of faecal indicator organisms in streams impacting on recreational and shellfish harvesting waters. *WATER RESEARCH*, **40 (1)**, 153–161.

Wolf, S. G., M. A. Snyder, W. J. Sydeman, D. F. Doak, and D. A. Croll, 2010: Predicting population consequences of ocean climate change for an ecosystem sentinel, the seabird Cassin's auklet. *GLOBAL CHANGE BIOLOGY*, **16 (7)**, 1923–1935.

Wroblicky, G. J., M. E. Campana, H. M. Valett, and C. N. Dahm, 1998: Seasonal variation in surface-subsurface water exchange and lateral hyporheic area of two stream-aquifer systems. *WATER RESOURCES RESEARCH*, **34(3)**, 317–328.

Wüest, A., 2010: Geophysikalische untersuchungen im birstal flussgrundwasser interaktion und karstevolution. M.S. thesis, Applied and Environmental Geology, University of Basel.

Zektser, I. S. and L. G. Everett, 2004: Groundwater Resources of the World and their Use. UNESCO IHP-VI, Series on Groundwater No. 6.

13 Curriculum Vitae

Rebecca Mary Page

Date of Birth: 23 July 1980
Place of Birth: Gateshead, Tyne & Wear, United Kingdom
Citizenship: British, Swiss (*Heimatort: Basel-Stadt*)

Education

2008 - present	Ph. D., Applied and Environmental Geology, Institute of Geology and Paleontology, Department of Environmental Sciences, University of Basel, CH
2004 - 2007	BSc Environmental Engineering, ETH Zurich, CH BSc Thesis: *Indicators for the Effects of Water Exploitation on Mangrove Diversity* Supervisors: Prof. S. Hellweg, S. Pfister
2002 - 2003	MSc Applied Marine Science, University of Plymouth, UK MSc Thesis: *A Baseline Survey of the Benthic Invertebrate Fauna of the Maltese Islands* Supervisors: Prof. P. J. Schembri, Dr. D. Pilgrim
1999 - 2002	BSc Marine Biology Hons., University of Newcastle-upon-Tyne, UK BSc Thesis: *The Effect of Demersal Trawling on Benthic Communities* Supervisor: Dr. C. L. Frid
1991 - 1999	*Gymnasium Leonhard*, Basel, CH

Related Work Experience

2004	Institute for European Environmental Policy (IEEP), London, UK (Internship)
2003	MEDITS (International MEDiterranean Trawl Survey), University of Malta and Ministry for Resources and Rural Affairs, Malta (MSc field work)
2001	GeoCrust: Use of Satellite GPS Data to Map Effort and Landings of the Portuguese Crustacean Fleet, Universidad do Algarve, Portugal (Internship)

14   Appendix

Risk analysis for riverine groundwater extraction

R. M. PAGE[1], P. HUGGENBERGER[1], J. EPTING[1] & G. LISCHEID[2]

1 *Department of Environmental Sciences, Applied and Environmental Geology, University of Basel, Bernoullistr. 32, 4056 Basel, Switzerland*
rebecca.page@unibas.ch

2 *Leibniz Centre for Agricultural Landscape Research (ZALF), Institute of Landscape Hydrology, Eberswalder Str. 84, 15374 Müncheberg, Germany*

Abstract Natural filtration in aquifers can improve the quality of groundwater relative to infiltrating surface water. During high discharge events, the microbial load of the groundwater can be several orders of magnitude higher than under normal flow conditions. The risk of contamination of groundwater extraction wells through infiltrating river water has been assessed by: (1) considering hydraulic gradients between the river and the groundwater, as well as the behaviour of quality parameters, such as turbidity, electrical conductivity and spectral absorbance coefficient; (2) principal component analysis to identify areas, time periods and parameters sensitive to infiltrating river. The results illustrate the major processes influencing the groundwater body in the study area, thus enabling a real-time assessment of potential sources of contamination.

Key words river–groundwater interaction; principal component analysis; drinking water

INTRODUCTION

Groundwater extraction wells are often located near rivers, where water suppliers rely on natural groundwater recharge from the river, augmenting the amount of extractable water in the aquifer (e.g. Sheets *et al.*, 2002). Depending on the filter capacity of the aquifer material, natural filtration can improve the quality of the water in comparison to the infiltrating surface water (Taylor *et al.*, 2004). However, one of the major concerns for drinking water suppliers extracting riverine groundwater is microbial contamination (Sheets *et al.*, 2002; Regli *et al.*, 2003). During high discharge events, the concentrations of microorganisms in the groundwater can be several orders of magnitude higher then under normal flow conditions (Huggenberger, 2001; Regli *et al.*, 2003). The decline of drinking water quality during high discharge events is influenced by several factors: the amount and velocity of infiltrating river water, hydraulic conductivity of the river bed and aquifer, the distance between the extraction wells and the infiltration zone, as well as contaminant loading, e.g. bacterial concentration of the river water. The time period over which infiltration, and the concentration of contaminants, e.g. the bacterium *E. coli*, is elevated, is important for assessing management options for groundwater extraction wells near rivers (McKergow & Davies-Colley, 2010). One way to maintain the quality of the groundwater under continually changing boundary conditions (natural and operational) is through an adaptive approach to managing groundwater capture zones, considering both spatial and temporal behaviour (e.g. Seward *et al.*, 2006; Epting *et al.*, 2008). A dynamic approach to risk-oriented groundwater production systems opens well-specific time windows for the use of the groundwater resource. In addition, it allows the simultaneous consideration of multiple issues, e.g. maintaining environmental integrity while satisfying the quality and quantity aspects of groundwater supply

146

(Feyen & Gorelick, 2004). In order to optimise the extraction regime, the major processes dominating the system need to be understood and indicators identified, so that the potential risk of contamination can be continuously monitored. These include the spatial and temporal variability of hydraulic gradients as a driving force for dynamic infiltration/exfiltration processes.

This study combines continuous multi-parameter monitoring with a principal component analysis (PCA) to identify and characterize the processes potentially leading to contamination of drinking water by infiltrating river water and to identify indicator parameters and suitable observation wells. This is a first step in defining well-specific extraction management rules to avoid contamination.

STUDY AREA AND METHODS

The study area is located in the lower Birs Valley in NW Switzerland and focuses on a regional groundwater body, a management unit of groundwater, as defined in the Water Framework Directive (Rejman, 2007). The Birs Valley alluvial system is a shallow coarse gravelly aquifer. Most of the water supplied to the aquifer comes from river infiltration, natural recharge through the local catchment area, and intermittent artificial recharge. There are eight extraction wells in the study area, supplying approx. 5 million m^3 water per year.

Nine groundwater observation wells were equipped with temperature, electrical conductivity (EC), turbidity and pressure sensors with a 30-min measurement resolution. In addition, a flow-through analyser allowed pH, Spectral Absorbance Coefficient (SAC) and number of particles (2–10 μm) to be continuously recorded at three of the observation wells.

PCA can be used to extract information on dominating influences and patterns from large data sets (e.g. Gerbrands, 1981; Longuevergne *et al*., 2007; Lischeid, 2009). Each principal component (PC) is representative of an input signal responsible for a certain amount of variance observed in the original data set. The more variance that can be assigned to a signal, the more important the signal is for the observed behaviour in the measured data. All groundwater head (H, calculated from pressure), six EC and six temperature data sets were used in the PCA. In this study the resulting PCs were then compared to time-series representing the potential processes dominating the system, e.g. river stage and regional groundwater flow. The behaviour of temperature, EC, turbidity, H, pH, SAC and number of particles is discussed in connection with river stage and precipitation during a high discharge event.

RESULTS AND DISCUSSION

The first three PCs describe >80% of the variance observed in the groundwater parameters of temperature, H and EC. A detailed analysis shows that processes giving rise to PC1 and PC2 have a seasonal or regional character and strongly influence H and temperature patterns. These two PCs, by definition statistically independent, show similar temporal patterns, most likely owing to the fact that processes dominating regional groundwater recharge and seasonality are not independent. PC3 shows a strong correlation with river stage, especially during the winter months, when high discharge events were more frequent. H shows short-term fluctuations and EC is mainly determined by the process of infiltration, as described by PC3.
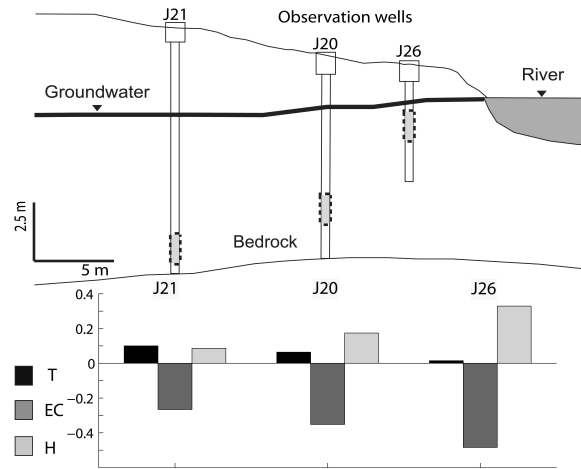
Fig. 1 Schematic representation of three observation wells, filter depths and PC3 loadings for temperature (T), EC and H.
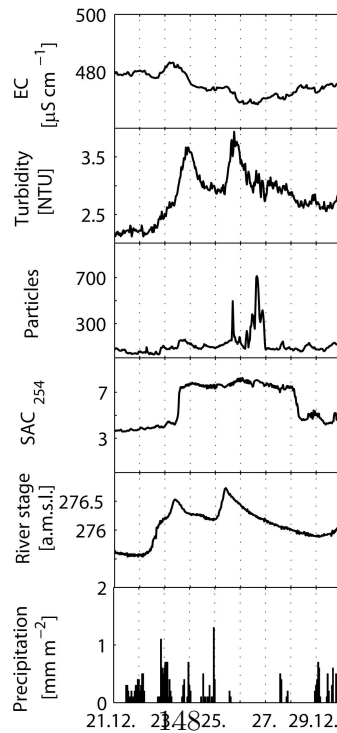

Fig. 2 Reactions in parameters following precipitation events in December 2009.

149

312   *R. M. Page et al.*

Figure 1 shows a schematic representation of three observation wells with the filter screens where the measurements were taken. In addition, the relative importance of PC3 (infiltration of river water) is shown for the three observation wells and parameters. With increasing distance from the river, the importance of PC3 for EC and H weakens. Temperature however, only shows a relatively weak relationship with PC3.

The PCA results show that the two observation wells close to the river are most sensitive to infiltration. As hydraulic gradients between the river and the groundwater are drivers for infiltration, changes in H provide an initial warning of potential infiltration. In addition, decreases in EC indicate mass transport of water from the river, which in this case has a lower EC. The time-lag between precipitation, river swelling and reactions in the groundwater parameters is important for determining an indicator parameter for infiltration. The characteristic decrease in EC is one of the first reactions after the river stage rises (Fig. 2). SAC also responds rapidly and remains elevated while other parameters, such as turbidity, tend to return to their initial levels. SAC shows the duration of elevated levels of organic matter in the groundwater. As the origin of organic matter is riverine, SAC would be a good indicator for the duration of contamination due to river water infiltration.

CONCLUSIONS

Natural variability and short-term fluctuations caused by groundwater extraction impede a threshold-approach to assessing potential contamination of groundwater extraction wells. However, the state of the system can be continually monitored by considering the combination of hydraulic fluctuations and physical quality, as recorded by H and EC or SAC, respectively. This can be especially important in areas or regions with high annual or inter-annual fluctuations.

The PCA identifies the major processes influencing the groundwater body in the study area, and provides a temporal assessment of potential sources of contamination. The three indicator parameters (H, EC and SAC) are sensitive to river water infiltration.

The method proposed is independent of absolute values, but uses the dynamics of the system to identify short-term changes that are significant for assessing potential risk of contamination of groundwater extraction wells due to infiltrating river water. The effect of variable extraction regimes on infiltration and flow paths between the river and the groundwater extraction well will be simulated using groundwater modelling tools.

REFERENCES

Epting, J., Regli, C. & Huggenberger, P. (2008) Groundwater protection in urban areas incorporating adaptive groundwater modeling and management. In: *Adaptive and Integrated Water Management, Coping with Complexity and Uncertainty* (ed. by C. Pahl-Wostl, P. Kabat & J. Möltgen), 97–123. Springer ISBN: 978-3-540-75940-9.
Feyen, L. & Gorelick, S. M. (2004) Reliable groundwater management in hydroecologically sensitive areas. *Water Resour. Res.* **40**(W0), 7408, doi:10.1029/2003WR003003.

Gerbrands, J. J. (1981) On the relationships between SVD, KLT and PCA. *Pattern Recognition* **14**, 375–381.

Huggenberger P. (2001) Wiese-Revitalisierung: Führen die Veränderungen der Sohlenstruktur zu einer Trinkwasser-gefährdung? *Regio Basiliensis* **42**(1), 63–76.

Lischeid, G. (2009) Non-linear visualization and analysis of large water quality data sets: a model-free basis for efficient monitoring and risk assessment. *Stoch. Environ. Res. Risk Assess.* **23**, 977–990.

Longuevergne L., Florsch N. & Elsass P. (2007) Extracting coherent regional information from local measurements with Karhunen-Loève transform: Case study of an alluvial aquifer (Rhine valley, France and Germany). *Water Resour. Res.* **43**(W0), 4430, doi:10.1029/2006WR005000.

McKergow, L. A. & Davies-Colley, R. J. (2010) Stormflow dynamics and loads of Escherichia coli in a large mixed land use catchment. *Hydrol. Processes* **24**, 276–289.

Regli, C., Rauber, M. & Huggenberger, P. (2003) Analysis of aquifer heterogeneity within a well capture zone, comparison of model data with field experiments; A case study from the river Wiese, Switzerland. *Aquat. Sci.* **65**, 111–128.

Rejman, W. (2007) EU Water Framework Directive *versus* real needs of groundwater management. *Water Resour. Manage.* **21**, 1363–1372.

Seward, P, Xu, Y. & Brendonck, L. (2006) Sustainable groundwater use, the capture principle, and adaptive management. *Water SA* **32**(4), 473–482.

Sheets, R. A., Darner, R. A. & Whitteberry, B. L. (2002) Lag times of bank filtration at a well field, Cincinnati, Ohio, USA. *J. Hydrol.* **266**, 162–174.

Taylor, R., Cronin, A., Pedley, S., Barker, J. & Atkinson, T. (2004) The implications of groundwater velocity variations on microbial transport and wellhead protection – review of field evidence. *FEMS Microbial Ecology* **49**, 17–26.