Characterization of spatial-temporal varying riverbed hydraulic conductivity and its role on the estimation of river-aquifer exchange fluxes with data assimilation

Qi Tang

Energie & Umwelt / Energy & Environment
Band / Volume 429
ISBN 978-3-95806-339-6

JÜLICH
Forschungszentrum

Schriften des Forschungszentrums Jülich
Reihe Energie & Umwelt / Energy & Environment Band / Volume 429

Forschungszentrum Jülich GmbH
Institut für Bio- und Geowissenschaften
Agrosphäre (IBG-3)

Characterization of spatial-temporal varying riverbed hydraulic conductivity and its role on the estimation of river-aquifer exchange fluxes with data assimilation

Qi Tang

Schriften des Forschungszentrums Jülich
Reihe Energie & Umwelt / Energy & Environment   Band / Volume 429

ISSN 1866-1793   ISBN 978-3-95806-339-6

Bibliografische Information der Deutschen Nationalbibliothek.
Die Deutsche Nationalbibliothek verzeichnet diese Publikation in der
Deutschen Nationalbibliografie; detaillierte Bibliografische Daten
sind im Internet über http://dnb.d-nb.de abrufbar.

Herausgeber und Vertrieb:	Forschungszentrum Jülich GmbH Zentralbibliothek, Verlag 52425 Jülich Tel.: +49 2461 61-5368 Fax: +49 2461 61-6103 zb-publikation@fz-juelich.de www.fz-juelich.de/zb

Umschlaggestaltung:	Grafische Medien, Forschungszentrum Jülich GmbH
Druck:	Grafische Medien, Forschungszentrum Jülich GmbH

Schriften des Forschungszentrums Jülich
Reihe Energie & Umwelt / Energy & Environment, Band / Volume 429

D 82 (Diss., RWTH Aachen, Univ., 2018)

ISSN 1866-1793
ISBN 978-3-95806-339-6

Vollständig frei verfügbar über das Publikationsportal des Forschungszentrums Jülich (JuSER)
unter www.fz-juelich.de/zb/openaccess.

Content

Abstract

Interactions between surface water and groundwater play an essential role in hydrology, hydrogeology, ecology, and water resources management. When modelling the river-aquifer interactions, a proper characterization of riverbed properties such as the riverbed topography and the riverbed hydraulic conductivity (K_{rb}) can be important for the prediction of exchange fluxes between a river and an aquifer. These riverbed properties are changing in time and space. Specifically, flood events may change the riverbed elevation as well as the riverbed texture and structure, which in turn can influence the K_{rb} and river-aquifer exchange fluxes. One main objective of this PhD-work was to investigate the role of different K_{rb} patterns on prediction of hydrologic states and river-aquifer exchange fluxes and to evaluate methodologies for improving the characterization of the spatial and temporal variability of K_{rb}, in combination with different conceptualizations of the heterogeneity of the riverbed. In particular, it was evaluated whether variants of the Ensemble Kalman Filter (EnKF), an ensemble based data assimilation technique, can reproduce such interactions. EnKF is commonly used in subsurface flow and transport modelling for estimating states and parameters. However, EnKF only performs optimally for multi-Gaussian distributed parameter fields, but the spatial distribution of K_{rb} often shows complex non-multi-Gaussian patterns, which are related to flow velocity dependent sedimentation and erosion processes. In this work, multiple types of heterogeneous K_{rb} patterns, based on different geostatistical models, were evaluated and compared.

A first synthetic study considered a 3-D river-aquifer model including a river in hydraulic connection to a homogeneous aquifer using a conductance based groundwater model. A riverbed with non-multi-Gaussian distributed hydraulic parameters with channelized structures was assumed as a virtual reference. In a series of data assimilation experiments three different geostatistical models for the spatial distribution of the riverbed hydraulic parameters were studied: (i) heterogeneous with non-multi-Gaussian distribution (channelized structures), (ii) heterogeneous with non-multi-Gaussian distribution (elliptic structures) and (iii) heterogeneous with multi-Gaussian distribution. For the non-multi-Gaussian scenarios, stochastic realizations of non-multi-Gaussian distributed riverbeds were inversely conditioned to state measurements, taken from the virtual reality, using EnKF and the normal score ensemble Kalman filter (NS-EnKF). For the multi-Gaussian distribution, the stochastic realizations of riverbed properties have multi-Gaussian distributed hydraulic parameters and are also conditioned to the same state measurements with EnKF. It was concluded that both EnKF and NS-EnKF improve the characterization of non-multi-Gaussian riverbed properties, hydraulic heads and exchange fluxes by piezometric head assimilation, and only NS-EnKF could preserve the initial distribution of K_{rb}. In addition, it was found that the best results were achieved for channel-distributed non-multi-Gaussian stochastic realizations and with parameter updating. However, differences between the simulations were small and non-multi-Gaussian riverbed properties seem to be of less importance for subsurface flow than non-multi-Gaussian aquifer properties.

In a second study, the role of these heterogeneity patterns of K_{rb} was again explored but using a fully integrated hydrological model which can simulate complex, variably saturated subsurface flow. A similar synthetic 3-D river-aquifer model was set up. The reference model was constructed with a heterogeneous riverbed using the same non-multi-Gaussian patterns in the form of meandering channels as in the first study. Data assimilation was used again to test the ability of different K_{rb} patterns to reproduce hydraulic heads, K_{rb} and river-aquifer exchange fluxes. Both fully saturated as well as variably saturated conditions underneath the riverbed were tested. The data assimilation experiments with EnKF were carried out for four types of geostatistical models of K_{rb} fields: (i) spatially homogeneous; (ii-iv) three different geostatistical models similar to the ones used in the first study. For all data assimilation experiments, state variables and K_{rb} were updated by assimilating hydraulic heads. For saturated conditions, heterogeneous geostatistical models allowed a better characterization of net exchange fluxes than a homogeneous approximation. Among the three heterogeneous models, the performance of non-multi-Gaussian models was superior to the performance of the multi-Gaussian model, but the two tested non-multi-Gaussian models showed only small differences in performance from one another. For the variably saturated conditions both the multi-Gaussian model and the homogeneous model performed clearly worse than the two non-multi-Gaussian models, while the two non-multi-Gaussian models did not show much difference in performance. This clearly shows that characterizing heterogeneity of K_{rb} is important. Moreover, particularly under variably saturated flow conditions the mean and the variance of K_{rb} do not provide enough information for exchange flux characterization and additional histogram information of K_{rb} provides crucial information for the reproduction of exchange fluxes.

In the third study, a model was set up for the Upper Emme catchment in Switzerland with a fully integrated hydrological model. Due to a 300-year flood event that happened on 24.07.2014 in the Emme River in Switzerland, the riverbed topography and probably the K_{rb} were greatly changed. Multiple data assimilation experiments were carried out with EnKF, including the periods before and after the flood for the year 2014, to detect the spatial and temporal variation of these riverbed properties and to characterize the river-aquifer interaction. The following scenarios were simulated: (i) with/without consideration of changes in riverbed topography as observed by drone measurements; (ii) with/without update of K_{rb} and aquifer hydraulic conductivity (K_{aq}) before the flood; (iii) with/without update of K_{rb} and/or K_{aq} after the flood, in 2014. The performance of the data assimilation was evaluated by evaluating the reproduction of the hydraulic states for the following year 2015. The scenario with pre-flood hydraulic parameters and pre-flood riverbed topography, not updated after the flood, resulted in the largest root mean square error (RMSE) of heads (RMSE (*h*) = 76.6 cm). Using the post-flood riverbed topography instead of the pre-flood riverbed topography reduced the RMSE (*h*) by 24% to 57.9 cm. If in addition to using the post-flood riverbed topography also K_{rb} and K_{aq} were updated through data assimilation after the flood, the smallest RMSE (*h*) was obtained (34.8 cm). This implies a reduction of RMSE (*h*) of 55% compared to using pre-flood riverbed topography and hydraulic conductivity (*K*). On the other hand, the prediction of surface water discharge was not affected much by these changes. In summary, it could be shown that updating K_{rb} and K_{aq} combined with the post-flood riverbed topography after a major flood event is important for groundwater flow modelling in the period after the flood, because changes induced by such floods have a significant effect on piezometric heads.

Zusammenfassung

Wechselwirkungen zwischen Oberflächenwasser und Grundwasser spielen eine maßgebliche Rolle in Hydrologie, Hydrogeologie, Ökologie und Wasserressourcen-Management. Bei der Modellierung von Wechselwirkungen zwischen Flüssen und Grundwasserleitern kann eine korrekte Charakterisierung von Flussbett-Eigenschaften, wie etwa der Flussbett-Topografie oder der hydraulische Flussbett-Leitfähigkeit (*K*), für die Vorhersage von Austauschprozessen zwischen Flüssen und Grundwasserleitern von großer Bedeutung sein. Diese Flussbett-Eigenschaften verändern sich über Zeit und Raum. Speziell Überflutungsereignisse können die Lage des Flussbettes sowie dessen Textur und Struktur verändern, was wiederum Einfluss auf die Flussbett-*K* sowie auf Austauschflüsse zwischen Fluss und Grundwasserleiter haben kann. Ein Hauptziel dieser Arbeit war die Untersuchung der Rolle von unterschiedlichen Flussbett-*K*-Mustern auf die Vorhersage von hydrologische Zustandsgrößen und Austauschflüssen zwischen Fluss und Grundwasserleiter sowie die Evaluation von Methoden für eine verbesserte Charakterisierung der räumlichen und zeitlichen Variabilität der Flussbett-*K*, in Kombination mit unterschiedliche Konzeptualisierungen der Flussbett-Heterogenität. Insbesondere wurde evaluiert, ob Varianten des „Ensemble-Kalman-Filters“ (EnKF), einer Ensemble-basierten Datenassimilationstechnik, solche Wechselwirkungen reproduzieren können. Der „Ensemble-Kalman-Filter“ (EnKF) wird häufig bei der Modellierung von Strömungs- und Transportprozessen im Untergrund für die Abschätzung von Zustandsgrößen und Parametern verwendet. Jedoch funktioniert EnKF nur bei einer Multi-Gauß´schen Verteilung der Parameterfelder optimal, wohingegen die räumliche Verteilung der Flussbett-*K* oft komplexe nicht-Multi-Gauß´sche Muster aufweist, die mit fließgeschwindigkeitsabhängigen Sedimentations- und Erosionsprozessen zusammenhängen. In dieser Arbeit wurden mehrere Typen heterogener Flussbett-*K*-Muster, basierend auf verschiedene geostatistische Modelle, evaluiert und verglichen.

Eine erste synthetische Studie wurde mithilfe eines dreidimensionalen Fluss-Grundwasser-Modells durchgeführt, bei dem der Fluss über das sogenannte „Leakage“-Konzept in hydraulischer Verbindung mit einem homogenen Grundwasserleiter steht. Ein Flussbett mit nicht-Multi-Gauß-verteilten hydraulischen Parametern mit Kanalstrukturen wurde als virtuelle Referenz angenommen. In einer Serie von Datenassimilations-Experimenten wurden drei verschiedene geostatistische Modelle für die räumliche Verteilung von Flussbett-*K* Parametern untersucht: (i) heterogen mit nicht-Multi-Gauß´scher Verteilung (Kanalstrukturen), (ii) heterogen mit nicht-Multi-Gauß´scher Verteilung (elliptische Strukturen), (iii) heterogen mit Multi-Gauß´scher Verteilung. Für die nicht-Multi-Gauß-Szenarien wurden stochastische Realisierungen von nicht-Multi-Gauß-verteilten Flussbetten unter Verwendung von EnKF und dem „normal-score“ Ensemble-Kalman-Filter (NS-EnKF) invers auf Zustandsinformationen aus der virtuellen Referenz konditioniert. Für das Multi-Gauß-Szenario

wiesen die stochastischen Realisierungen der Flussbetteigenschaften Multi-Gauß-verteilte hydraulische Parameter auf und wurden auch mithilfe von EnKF auf die gleichen Zustandsinformationen konditioniert. Es stellte sich heraus, dass sowohl EnKF als auch NS-EnKF die Charakterisierung von nicht-Multi-Gauß'schen Flussbett-Eigenschaften, Piezometerhöhen und Austauschflüssen durch die Assimilation von Piezometerhöhen verbessern konnte, wobei nur durch den NS-EnKF die anfängliche Verteilung der Flussbett-*K* erhalten wurde. Zudem stellte sich heraus, dass die besten Ergebnisse für nicht-Multi-Gauß'sche stochastische Realisierungen mit Kanalstrukturen unter Verwendung von Parameteranpassung erreicht wurden. Jedoch waren die Unterschiede zwischen den Simulationen gering und nicht-Multi-Gauß'sche Flussbett-Eigenschaften scheinen für Strömungsprozesse im Untergrund weniger wichtig zu sein als eine nicht-Multi-Gauß´sche Verteilung von Eigenschaften des Grundwasserleiters.

In der zweiten Studie wurde ebenfalls die Rolle dieser Heterogenitäts-Muster von Flussbett-*K* erforscht, aber diesmal unter Nutzung eines voll integrierten hydrologischen Modells, welches die Simulation komplexer, variabel gesättigter Strömungsprozesse im Untergrund ermöglicht. Ein ähnliches synthetisches dreidimensionales Fluss-Grundwasser-Modell wurde aufgebaut. Das Referenzmodell wurde mit einem heterogenen Flussbett konstruiert, unter Verwendung derselben nicht-Multi-Gauß´schen Muster in Form von mäandernden Kanälen wie in der ersten Studie. Datenassimilation wurde wieder verwendet, um die Fähigkeit verschiedener Flussbett-*K*-Muster zur Reproduktion von Piezometerhöhen, Flussbett-*K* und Fluss-Grundwasser-Austauschflüssen zu überprüfen. Sowohl völlig gesättigte als auch variabel gesättigte Zustände unterhalb des Flussbetts wurden getestet. Die Datenassimilations-Experimente mit EnKF wurden für vier Typen geostatistischer Modelle von Flussbett-*K*-Feldern durchgeführt: (i) räumlich homogen; (ii)-(iv) drei unterschiedliche geostatistische Modelle, die denjenigen aus der ersten Studie gleichen. Für alle Datenassimilations-Experimente wurden Zustandsvariablen und Flussbett-*K* durch die Assimilation von Piezometerhöhen angepasst. Bei gesättigten Bedingungen erlaubten die heterogenen geostatistischen Modelle eine bessere Charakterisierung der Netto-Austausch-Flüsse als eine homogene Annäherung. Unter den drei heterogenen Modellen war die Leistungsfähigkeit der nicht-Multi-Gauß-Modelle höher als die des Multi-Gauß-Modells, allerdings wiesen die beiden getesteten nicht-Multi-Gauß-Modelle nur kleine Leistungsunterschiede untereinander auf. Bei den variabel gesättigten Bedingungen schnitten sowohl das Multi-Gauß-Modell als auch das homogene Modell ganz klar schlechter ab, als die beiden nicht-Multi-Gauß-Modelle, deren Leistungsfähigkeit nicht sehr verschieden war. Dies macht deutlich, dass die Charakterisierung der Heterogenität der Flussbett-*K* wichtig ist. Zudem stellen der Mittelwert und die Standardabweichung der Flussbett-*K* vor allem bei variabel gesättigten Strömungsbedingungen nicht genügend Informationen für die Charakterisierung der Austausch-Flüsse zur Verfügung, wohingegen zusätzliche Histogramm-Informationen zur Flussbett-*K* wesentliche Informationen zur Reproduktion von Austausch-Flüssen liefern.

In der dritten Studie wurde ein voll integriertes hydrologisches Modell des Einzugsgebietes der oberen Emme in der Schweiz aufgestellt. Während eines 300-Jahre-Überflutungsereignisses der Emme am 24.07.2014 hat sich die Flussbetttopografie und möglicherweise auch die Flussbett-*K* stark verändert. Mehrere Datenassimilations-Experimente wurden mit EnKF durchgeführt, um die räumliche und zeitliche Variation dieser Flussbett-Eigenschaften zu ermitteln und die Fluss-Grundwasserleiter-Wechselwirkung zu charakterisieren. Diese beinhalteten die Zeitperioden vor und nach der Überflutung im Jahre 2014. Die folgenden Szenarien wurden simuliert: (i) mit/ohne Berücksichtigung der Veränderungen in der Flussbetttopografie, die über Drohnen-Messungen beobachtet wurden; (ii) mit/ohne Aktualisierung der Flussbett- und Grundwasserleiter-*K* vor der Überflutung; (iii) mit/ohne Aktualisierung der Flussbett- und/oder Grundwasserleiter-*K* nach der Überflutung, im Jahr 2014. Die Leistungsfähigkeit der Datenassimilation wurde durch einen Vergleich der Reproduktion der hydraulischen Zustandsgrößen für das Folgejahr 2015 evaluiert. Das Szenario mit den hydraulischen Parametern sowie der Flussbetttopografie vor der Überflutung, die nach der Überflutung nicht aktualisiert wurden, ergaben den größten Wert der Wurzel der mittleren quadratischen Abweichungen der Piezometerhöhen (RMSE (h) = 76,6 cm). Die Nutzung der Flussbetttopografie nach der Überflutung anstelle derjenigen vor der Überflutung reduzierte den RMSE (h)-Wert um 24 % auf 57,9 cm. Der kleinste RMSE (h)-Wert von 34,8 cm wurde erreicht, indem zusätzlich zur Nutzung der Flussbetttopografie nach der Überflutung auch Flussbett- und Grundwasserleiter-*K* durch Datenassimilation nach der Überflutung aktualisiert wurden. Dies bedeutet eine Reduzierung des RMSE (h)-Wertes um 55 % im Vergleich zur Nutzung der Flussbetttopografie und *K* vor der Überflutung. Allerdings wurde die Vorhersage des Oberflächenwasserabflusses durch diese Veränderungen nicht sehr stark beeinflusst. Zusammenfassend kann festgestellt werden, dass die Anpassung von Flussbett- und Grundwasserleiter-*K* in Kombination mit der Flussbetttopografie nach dem großen Überflutungsereignis wichtig ist für die Modellierung des Grundwasserflusses in der Zeit nach der Überflutung, da die Veränderungen, die von solchen Überflutungsereignissen hervorgerufen werden, einen wesentlichen Effekt auf Piezometerhöhen haben.

List of Figures

List of Tables

List of Acronyms

EnKF	ensemble Kalman filter
GW	groundwater
HGS	HydroGeoSphere
MP	multiple-point
NS-EnKF	Normal Score Ensemble Kalman Filter
RMSE	root mean square error
SW	surface water
K	hydraulic conductivity
K_{rb}	riverbed hydraulic conductivity
K_{aq}	aquifer hydraulic conductivity

Chapter 1 Introduction

1.1 Importance of surface water - groundwater interaction

Water resources management is challenged by an increasing demand for water by agriculture, industry and human activities. Although surface water (SW) as well as groundwater (GW) serve as important components of global freshwater resources, the fundamental management unit for water resources was normally either the SW or the GW, as SW and GW were considered as separated bodies (Winter et al., 1998; Vaessen and Brentführer, 2014). However, these two bodies interact with each other and changes in one single body affect the quantity and quality of the other one (Winter et al., 1998). For example, the maintenance of a minimum ecological flow in a river system is strongly controlled by the groundwater flow contribution to the river discharge, especially in the dry seasons (Sophocleous, 2002). Contaminants in the river can infiltrate into the aquifer and thus pollute drinking water that is pumped from the aquifer (Hendricks Franssen et al., 2011). Groundwater storage can serve as a buffer for flood and drought during extreme events (Jones, 2011; Scanlon et al., 2012). The SW-GW transition zone is ecologically significant as complex physiochemical environmental conditions are provided for species due to the exchange process (Brunke and Gonser, 1997). Groundwater discharge areas linked to the river are essential for winter fish survival as groundwater temperature is higher than the surface water (Hayashi and Rosenberry, 2002). Therefore an effective, conjunctive management of SW and GW was proposed within an integrated water resources management framework, requiring a clear and better understanding of SW-GW interaction (Winter et al., 1998; Vaessen and Brentführer, 2014).

1.2 Challenges in modeling river-aquifer system

Studies on SW-GW interactions began in the 1960s (e.g. Rorabaugh, 1964) and the main interest for studying these interactions was the eutrophication of lakes and later acid rain (Sophocleous, 2002). In recent years, exchange between near-channel water and in-channel water like the river-aquifer system attracts more and more attention. The groundwater flow contribution to the river and the corresponding feedback from the river are varying in space and time (Ellis et al., 2007; Krause et al., 2007). Quantification of such dynamic river-aquifer exchange processes is challenging. One key point is the selection of hydrological models which can consider both components and the interaction between them. Until now several of such models have been developed, for a review see Furman (2008). Conductance based groundwater models are the most commonly used models for simulating river-aquifer interaction, such as MODFLOW (McDonald and Harbaugh, 1988) and SPRING (Delta h Ingenieugesellschaft mbH, 2006). The advantage is that these models are relatively easy to set up, and need relatively little computation time (Hogan et al., 2004). However, in these models the surface water heads are pre-defined, unsaturated zones are neglected and exchange fluxes between the river and the aquifer are calculated according to a simplified, linear relationship. An alternative are more complex, integrated hydrological models which are designed to simulate the surface water and groundwater in one framework, such as InHM (VanderKwaak, 1999), MODHMS (Panday and Huyakorn, 2004), Parflow (Kollet and Maxwell, 2006), CATHY (Camporese et al., 2010), and HydroGeoSphere (HGS) (Therrien et al., 2010; Aquanty Inc, 2016). For all of these models, the three dimensional Richards' equation is used to describe the water flow in the subsurface domain. The surface water flow is described by either the two dimensional diffusion wave (e.g. InHM, MODHMS, CATHY and HGS) or the kinematic wave (e.g. Parflow) approximation to the Saint Venant equations. Coupling between the surface and the subsurface flow is done by either the first-order exchange coefficient approach (e.g. InHM, HGS and MODHMS) or the continuity of pressure approach (e.g. Parflow, HGS, CATHY). This type of models can simultaneously solve the surface water and groundwater equations and the interaction between the river and the aquifer is considered in a physically based, consistent manner. A detailed comparison between these two types of hydrological models is described by Furman (2008) and Brunner et al. (2010).

Another challenging point is to properly quantify model uncertainties induced by initial conditions, model forcings and model parameters, as these uncertainties can have a large impact on the modelling of river-aquifer interaction (Saha et al., 2017). Major uncertainty for simulating water flow in the river-aquifer system is related to the unknown geological parameters, i.e. the riverbed properties such as the riverbed bed topography and riverbed hydraulic conductivity (K_{rb}) (Li et al., 2017). These two riverbed properties can have a large impact on simulating the river-aquifer exchange fluxes (Woessner, 2000; Hester and Doyle, 2008). On the one hand, riverbed topography determines the exchange flow magnitude and pattern (Tonina and Buffington, 2007). The exchange flux between the river and the aquifer is enhanced by convexities and concavities (Harvey and Bencala, 1993). In one study, removing pool-riffle sequences reduced the hyporheic flow by a maximum of 48% and increased the longer residence time exchange flow, and removing channel sinuosity also increased the proportion of long residence time flow (Kasahara and Wondzell, 2003). Riverbed topography together with the stream water slope also control the solute transport (Harvey and Bencala, 1993). On the other hand, the K_{rb}, especially the heterogeneity of K_{rb} significantly influences the estimation of exchange fluxes (Cardenas et al., 2004; Fleckenstein et al., 2006) and flow residence time (Tonina et al., 2016). A separate riverbed layer with heterogeneous hydraulic conductivity values was needed for estimating peak flows of groundwater discharge (Kalbus et al., 2009). These hydraulic conductivity values can vary over one or two orders of magnitude (Käser et al., 2009; Xi et al., 2015; Song et al., 2016) in space and a homogeneous equivalent value can lead to large errors when estimating groundwater level (Lackey et al., 2015) and exchange fluxes under certain conditions (Irvine et al., 2012). Heterogeneity of hydraulic conductivity is also important for solute transport (Ryan and Boufadel, 2006), i.e. the assessment of contaminant movement and remediation (Anderson et al., 1999). Moreover, these two riverbed properties are varying in time, which makes the prediction of exchange fluxes complicated and highly dynamic, especially during extreme events like droughts and floods (Schubert, 2006). A scour up to 0.06 m and a total fluctuation of 0.17 m was detected in the riverbed height during the flood in January 2015 along the Great Miami River at Charles M. Bolton Water Plant in southwest Ohio, USA (Birck, 2006). Similar results were found by Simpson and Meixner (2012). Besides, the flood induced erosion can increase K_{rb} as well as the infiltration rate from the river to the aquifer (Doppler et al., 2007; Mutiti and Levy, 2010; Grischek and Bartak, 2016). On the other hand, after deposition starts, K_{rb} can be reduced by up to 75% due to the accumulation of fine materials transported by the river (Nowinski et al., 2011; Simpson and Meixner, 2012; Ulrich et al., 2015).

The ensemble based data assimilation technique, e.g. the ensemble Kalman filter (EnKF) (Evensen, 1994), is a powerful method for accounting for model as well as measurement uncertainties and can thus significantly improve the predictions of uncertain models by sequentially assimilating measurement data. Particularly, EnKF can be used together with hydrological modeling for improving the characterization of river-aquifer interaction. Stream flow can be better estimated by assimilating both stream flow and pressure head data with EnKF (Camporese et al., 2009). Specifically, the hydraulic parameters, i.e. the K_{aq} or K_{rb} can be successfully estimated together with model states by EnKF (Chen and Zhang, 2006; Hendricks Franssen and Kinzelbach, 2008; Hendricks Franssen et al., 2011; Xie and Zhang, 2013). Also, temporally and spatially variable K_{rb} can be characterized by sequential data assimilation with EnKF (Kurtz et al., 2012; Kurtz et al., 2013; Kurtz et al., 2014).

1.3 Motivation, research objectives and structure of the thesis

When modeling the heterogeneous K_{rb}, most of the above studies adopted a geostatistical model using the multi-Gaussian assumption. This is because limited information is available on the spatial K_{rb} variation and Gaussian statistics are commonly used in stochastic hydrogeology. However, already existing studies indicated that in reality more complex spatial patterns of K_{rb} with non-multi-Gaussian features are encountered (Anderson et al., 1999) which can be related to flow velocity dependent sedimentation and erosion patterns (Min et al., 2013; Sebok et al., 2015). Field measurement found that K_{rb} can be neither normally nor log-normally distributed (Genereux et al., 2008; Leek et al., 2009; Cheng et al., 2011), but can show a bimodal distribution (Springer et al., 1999). However, until now, none of these studies considered the non-multi-Gaussian patterns of K_{rb} when modelling river-aquifer interactions and whether such complex patterns influence the magnitude and the spatial patterns of river-aquifer exchange fluxes is still unknown. Moreover, no SW-GW interaction studies inversely estimated these non-multi-Gaussian distributed heterogeneous K_{rb} by sequential data assimilation. Therefore, the objective of the work in this PhD-dissertation is to improve characterizing the spatiotemporal varying K_{rb} by assimilating measurement data and to explore the role of complex patterns of K_{rb} on predicting river-aquifer exchange fluxes. This is tested for a simplified synthetic 3-D river-aquifer case using first a conductance based groundwater model and later a physically based, integrated hydrological model considering also complex saturation conditions under the riverbed. When non-multi-Gaussian distribution is taken into account, compared to the standard EnKF, the Normal Score Ensemble Kalman Filter (NS-EnKF) may be a better choice since it outperforms the standard EnKF in terms of characterizing hydraulic conductivity and updating piezometer heads (Zhou et al., 2011; Li et al., 2012; Schöniger et al., 2012; Crestani et al., 2013). Therefore we also explore whether NS-EnKF leads to a better characterization of non-multi-Gaussian K_{rb} (and river-aquifer exchange fluxes) than the standard EnKF.

Moreover, as described in section 1.1, a precise characterization of river-aquifer exchange fluxes is highly relevant for real world water management around streams in the context of drinking water supply, sustainability of groundwater use and ecological aspects. One example is the Upper Emme catchment in the Northern pre-alps of Switzerland which supplies around 40 % of drinking water for the Swiss capital Bern (Käser and Hunkeler, 2016). The pumping behavior strongly influences the SW-GW interactions around the drinking water stations which can influence the drinking water quality and the riparian ecosystem (Blau and Muchenberger, 1997). It is also shown that the riverbed permeability is the key point controlling the SW-GW interaction in this area (Schilling et al., 2017). Specifically, a 300-year flood occurred on July 24th of 2014 which changed the riverbed topography and probably also the K_{rb}. Therefore, a real world catchment model is set up for this catchment to further explore the role of the spatiotemporal variation of flood induced K_{rb} on the river-aquifer interaction.

The thesis is structured as follows. Chapter2 briefly describes the two hydrological models used for simulating the river-aquifer interactions and the EnKF approach. Chapter 3 presents the impact of three different heterogeneous K_{rb} patterns on the characterization of hydraulic heads, K_{rb} and river-aquifer exchange fluxes. A synthetic three dimensional river-aquifer model was set up using a simplified, conductance based model. Three different geostatistical models were used to generate the K_{rb} fields: one non-multi-Gaussian model with channelized structures, one non-multi-Gaussian model with elliptical structures and one multi-Gaussian model. A riverbed generated using the non-multi-Gaussian model with channelized structures was taken as a virtual reference case. Normal score transformation was implemented into the standard EnKF scheme and was later used for the two non-multi-Gaussian models. In Chapter 4, we explored further the importance of complex heterogeneous patterns of K_{rb} but with a physically based, integrated hydrological model for a synthetic study. Besides the three heterogeneous K_{rb} patterns, a homogeneous equivalent was included in the comparison. Moreover, both saturated and variably saturated conditions beneath the riverbed were considered for this study. In Chapter 5, the spatial and temporal variation of the K_{rb} induced by a 300-year flood was characterized and its impact on the prediction of hydraulic heads, surface water discharge and river-aquifer exchange fluxes was further explored for a real world catchment in the Emme catchment in Switzerland. The temporal variation of riverbed properties was characterized by not only time series of piezometric head data near the river, but also drone pictures, taken before and after the flood in the catchment. These drone pictures provide information about the riverbed elevation changes throughout the flood period. In this study again an integrated hydrological model was used combined with data assimilation. Finally, Chapter 6 briefly summarizes the outcomes of these numerical experiments and gives an outlook on possible future work.

Chapter 2 Theory and methods

2.1 Flow equations

2.1.1 Groundwater flow in porous media

Transient variably saturated groundwater flow in porous media is described by the three dimensional Richards' equation:

$$\frac{\partial \theta}{\partial t} = -\nabla \cdot \mathbf{q} \pm Q \qquad (2.1)$$

where θ is the water content [-], t is time [T], Q is external sinks and sources [L^3 L^{-3} T^{-1}] and **q** is the fluid flux [L T^{-1}]:

$$\mathbf{q} = -\mathbf{K} \cdot k_r \nabla(\psi + z) \qquad (2.2)$$

where k_r is relative permeability of the porous medium [-], ψ is the pressure head [L], z is the elevation above sea level [L] and **K** is hydraulic conductivity tensor [L T^{-1}] given by

$$\mathbf{K} = \frac{\rho g}{\mu} \mathbf{k} \qquad (2.3)$$

where ρ is the water density [M L^{-3}], g is gravitational acceleration [L T^{-2}], μ is the water viscosity [M L^{-1} T^{-1}], and **k** is the permeability tensor of the porous medium [L^2].

In both SPRING and HGS, the control volume finite element method is used to solve the groundwater flow equation. The mesh grid can be discretized as regular/irregular rectangular/triangular cells.

The van Genuchten equations are used to describe the relationship between ψ, θ and k_r for the unsaturated flow (Van Genuchten, 1980):

$$\theta(\psi) = \theta_r + \frac{\theta_s - \theta_r}{\left[1 + (\alpha|\psi|)^n\right]^m} \tag{2.4}$$

$$k_r = \theta^{0.5}\left[1 - \left(1 - \theta^{\frac{1}{m}}\right)^m\right]^2, \; m = 1 - \frac{1}{n} \tag{2.5}$$

where θ_s is the saturated soil water content [-], θ_r is the residual water content [-], α is the inverse of the air entry pressure head [L^{-1}], and *n* is the pore size distribution index [-].

2.1.2 Surface water flow

Surface water flow is based on the shallow water wave theory described by the one- or two-dimensional Saint Venant equations using the kinematic wave approximation (Lighthill and Whitham, 1955), the diffusion wave approximation, or the dynamic wave approximation. The selection of the approximation method is determined by various criteria like the Froudé numbers, boundary conditions, the kinematic wave numbers (Vieira, 1983). A detailed comparison among the different approximation methods is given by Liggett and Woolhiser (1967). In this study, the diffusive wave approximation is used:

$$\frac{\partial \phi_0 h_{sw}}{\partial t} = -\nabla \cdot d_0 \mathbf{q_0} - d_0 \Gamma \pm Q_0 \tag{2.6}$$

where ϕ_0 is the surface flow domain porosity [-], d_0 is the depth of the flow [L] , Γ is the fluid exchange rate [L^3 L^{-3} T^{-1}] between the surface domain and the subsurface domain, Q_0 is external sources and sinks [L T^{-1}] and $\mathbf{q_0}$ is the fluid flux [L T^{-1}] given by

$$\mathbf{q_0} = -\mathbf{K_0} \cdot k_{r0} \nabla(d_0 + z) \tag{2.7}$$

where $\mathbf{K_0}$ is surface conductance [L T^{-1}] dependent on the Manning´s equation, and k_{r0} is a dimensionless factor varying between 0 and 1, accounting for the additional resistance in horizontal conductance caused by obstruction storage. If a critical depth (the depth of flow where the specific energy is at a minimum) boundary condition is applied at the downstream of the surface flow nodes, stream discharge Q_r is calculated based on the upstream water level and the channel geometry:

$$Q_r = \sqrt{gd_0} \tag{2.8}$$

HGS solves the surface flow and subsurface flow equations simultaneously for each time step in a joint equation system which allows two-way coupling when accounting for the river-aquifer interaction. Instead, SPRING is a conductance based groundwater flow model, in which river stages are prescribed, for example from a pre-calibrated hydraulic model (Knapton, 2009), and the river stages do not receive the feedback from the groundwater flow simulation.

2.1.3 Surface water – groundwater flow coupling

The flow coupling between the subsurface and the surface domain can be done by either the common node approach or the dual node approach. The common node approach allows a fully coupled way to consider the GW-SW interaction because consistent heads are assumed for the surface and the uppermost subsurface nodes. The dual node approach assumes a thin layer between the surface layer and the uppermost subsurface layer and a first-order exchange coefficient is used to account for the permeability of this thin layer (sediment). This allows a higher numerical stability compared to the common node approach. Liggett et al. (2012) pointed out that if the coupling length (the thickness of the thin layer) in the dual node approach is set to a very small value (approaches zero), the effects of the dual node approach on simulation results are minimal and results are comparable to the common node approach. The exchange fluxes between the surface and the subsurface domain are then defined by the sediment K, the coupling length and the head gradient according to Darcy's law:

$$q_{exe} = \frac{k_{re}\mathbf{K}_{rb}}{l_{exch}}\left(h_{sw} - h_{gw}\right) \tag{2.9}$$

Where q_{exe} is the exchange flux ($d_0\Gamma$ in equation (2.6)) [L T^{-1}], k_{re} is the relative permeability term for the exchange flux, $\mathbf{K}_{rb}$ is the riverbed hydraulic conductivity, l_{exch} is the coupling length [L], h_{gw} is the groundwater level [L] calculated from equation (2.1) and h_{sw} is the surface water level [L] pre-defined as river stages in SPRING or calculated from equation (2.6) in HGS.

In SPRING, the leakage coefficient [T^{-1}] is defined as

$$\alpha = \frac{\mathbf{K}_{rb}}{l_{exch}} \tag{2.10}$$

and k_{re} is set to one. In HGS, k_{re} is the same as k_r in Equation (2.5) during the exfiltration process (water flows from subsurface to surface), and if infiltration (flow from surface to subsurface) occurs, k_{re} is defined as below:

$$k_{re} = \begin{cases} S_{exch}^{2(1-S_{exch})} & when\ d_0 < H_s \\ 1 & when\ d_0 > H_s \end{cases} \tag{2.11}$$

where $S_{exch} = \dfrac{d_0}{H_s}$ and H_s is the obstruction height [L].

2.2 The Ensemble Kalman filter

Data assimilation allows constraining states and parameters of numerical flow models by incorporating observations to correct the model evolution in state space. It is a powerful tool for uncertainty quantification, which has first been developed for the estimation of system states and later extended for the estimation of parameters (Chen and Zhang, 2006; Hendricks Franssen and Kinzelbach, 2008). In this work, the ensemble Kalman filter (EnKF) (Evensen, 1994) is used as the data assimilation technique for simultaneously updating state variables and parameters for both the conductance based groundwater model SPRING and the integrated hydrological model HGS. EnKF consists of three basic equations: the forecast equation, the measurement equation and the analysis equation. The forecast equation is solved by EnKF several times with different input parameters, following a Monte Carlo approach, to take uncertainties into account. Model uncertainties are taken into account by drawing random samples from the multivariate pdf of the initial conditions, forcing functions and parameters. The corresponding predicted state vectors, called ensemble members or stochastic realizations, are used to represent the statistical distribution of the model states. For each realization, the augmented state vector $\mathbf{x}_i$ consists of the state variables as well as model parameters:

$$\mathbf{x}_i = \begin{pmatrix} \mathbf{x}_s \\ \mathbf{x}_p \end{pmatrix}_i \tag{2.12}$$

where $\mathbf{x}_i$ is the augmented state vector containing the model states and parameters, $\mathbf{x}_s$ is the vector with model states, $\mathbf{x}_p$ the vector with model parameters and *i* the realization counter. In our study, $\mathbf{x}_s$ is hydraulic head, and $\mathbf{x}_p$ is log-transformed parameter. Therefore the $\mathbf{x}_i$ can be rewritten as:

$$\mathbf{x}_i = \begin{pmatrix} \mathbf{h} \\ \mathbf{Y} \end{pmatrix}_i \tag{2.13}$$

where **h** is hydraulic head and $\mathbf{Y} = \log_{10}(\mathbf{K})$. Here **K** can refer to the leakage coefficient α, K_{rb}, K_{aq}, or both K_{rb} and K_{aq}, depending on the simulation model and assimilation scenarios.

For each time step, the model states at the current time step are predicted from the previous time step by the model forecast equation:

$$\mathbf{x}_{t,i} = M(\mathbf{x}_{t-1,i}) \tag{2.14}$$

where *t* is the time counter, and *M* the simulation model (SPRING or HGS in this dissertation).

To account for the measurement uncertainty, the original measurements at time step *t* are perturbed according the measurement equation

$$\mathbf{y}_{t,i} = \mathbf{y}_t + \boldsymbol{\varepsilon}_{t,i} \tag{2.15}$$

where $\mathbf{y}_{t,i}$ is the vector of perturbed measurements, $\mathbf{y}_t$ is the vector with original measurements at time step *t*, and $\boldsymbol{\varepsilon}_{t,i}$ is the vector with observation errors usually generated from a normal distribution with zero mean and standard deviation equal to the measurement error.

The augmented state vector is then updated according the analysis equation by comparing the model simulation with the perturbed measurements:

$$\mathbf{x}_{t,i}^{a} = \mathbf{x}_{t,i} + \mu \mathbf{G}\left(\mathbf{y}_{t,i} - \mathbf{H}\mathbf{x}_{t,i}\right) \tag{2.16}$$

where $\mathbf{x}_{t,i}^{a}$ is the augmented state vector containing the updated model states and parameters, $\mathbf{x}_{t,i}$ is the vector with the forecasted states obtained from the dynamic model in Equation 2.14, μ is a damping factor varying between 0 and 1, **H** is the measurement operator matrix mapping the simulated states to the observation locations, and **G** is the Kalman gain which weights the relative importance of the model forecast and the observations. The Kalman gain is calculated by:

$$\mathbf{G} = \mathbf{C}\mathbf{H}^{\mathrm{T}}\left(\mathbf{H}\mathbf{C}\mathbf{H}^{\mathrm{T}} + \mathbf{R}\right)^{-1} \tag{2.17}$$

where **R** is the measurement error covariance matrix, and **C** is the model error covariance matrix containing covariances of the model states and parameters and given by

$$\mathbf{C} = \begin{pmatrix} \mathbf{C_{hh}} & \mathbf{C_{hY}} \\ \mathbf{C_{Yh}} & \mathbf{C_{YY}} \end{pmatrix} \tag{2.18}$$

Once the updating step is finished, the updated model states and parameters will be used as input for the model forecasts (the prediction function) for the next computation time step. Each time when measurements are available, equations 2.15-2.18 are applied, until the end of the simulation period.

As EnKF works optimally only when parameters or states follow a Gaussian distribution, Zhou et al. (2011) suggested the Normal score transform EnKF (NS-EnKF) scheme for updating states and parameters, which could outperform the standard EnKF if states and parameters have a non-Gaussian distribution. In this work, the NS-EnKF is only implemented for the model SPRING for joint updating of hydraulic heads and leakage coefficients. In the NS-EnKF, simulated piezometric heads and/or leakage coefficients are transformed using anamorphosis functions according to Johnson and Wichern (2002):

$$y = G^{-1}\left[F(\mathbf{x})\right] \tag{2.19}$$

$$F_j = \frac{j - 0.5}{N} \tag{2.20}$$

where **x** is the original vector, F the empirical cumulative distribution function (CDF), G the theoretical standard normal CDF, j the rank of **x** and N the ensemble size. For each variable, the anamorphosis function is created independently at each grid cell location and time step. In this work, the transformed hydraulic heads and parameter can be written as:

$$h_{t,i}^{trans} = G^{-1}\left[\frac{j_{h_{t,i}} - 0.5}{N}\right] \tag{2.21}$$

$$Y_{t,i}^{trans} = G^{-1}\left[\frac{j_{Y_{t,i}} - 0.5}{N}\right] \tag{2.22}$$

where the superscript *trans* indicates the variable after transformation. The corresponding model covariance matrix can be written as:

$$\mathbf{C} = \begin{pmatrix} \mathbf{C}_{\mathbf{h}^{\mathbf{trans}}\mathbf{h}^{\mathbf{trans}}} & \mathbf{C}_{\mathbf{h}^{\mathbf{trans}}\mathbf{Y}^{\mathbf{trans}}} \\ \mathbf{C}_{\mathbf{Y}^{\mathbf{trans}}\mathbf{h}^{\mathbf{trans}}} & \mathbf{C}_{\mathbf{Y}^{\mathbf{trans}}\mathbf{Y}^{\mathbf{trans}}} \end{pmatrix} \tag{2.23}$$

The anamorphosis function used to transform the perturbed measurements is the same as the one applied on the simulated heads at that particular location:

$$\mathbf{y}_{t,i}^{trans} = \phi\left(\mathbf{y}_{t,i}\right) \tag{2.24}$$

where $\phi(\cdot)$ is the anamorphosis function created based on the simulated heads. Since most measurements will not correspond exactly with one of the points that define the anamorphosis function and some may be outside the range, it is necessary to interpolate and extrapolate the transformed heads along the anamorphosis function. The linear spline is selected for interpolation, as it has been proven to be a stable and reasonable choice when the ensemble size is large enough (*N*>200) (Schöniger et al., 2012). The mean slope of the anamorphosis function was estimated and used for extrapolation, which was found in test experiments to be a stable solution. The mean slope *m* at the current time step is calculated by

$$m_t = \frac{x_{t,\max}^{trans} - x_{t,\min}^{trans}}{x_{t,\max} - x_{t,\min}} \tag{2.25}$$

where the subscript *max* refers to the maximum value of $\mathbf{x}_t$ among *i* realizations and *min* the minimum value.

When normal score transformation for the simulation and measurements is done, the transformed state vector is updated by

$$\mathbf{x}_{t,i}^{a,trans} = \mathbf{x}_{t,i}^{trans} + \mu\mathbf{G}\left(\mathbf{y}_{t,i}^{trans} - \mathbf{H}\mathbf{x}_{t,i}^{trans}\right) \tag{2.26}$$

where $\mathbf{x}_i^{trans} = \begin{pmatrix} \mathbf{h}^{trans} \\ \mathbf{Y}^{trans} \end{pmatrix}_i$

(2.27)

It should be noticed that it is possible to transform only the heads or only **Y**. As a last step, the updated heads and/or **Y** are back-transformed using the inverse anamorphosis functions according to equation (2.24). In this step interpolation/extrapolation is needed, still using the linear spline for interpolation and mean slope for extrapolation. The back transformed variables are used as inputs for the simulation of the next time step, and the procedure is repeated until the end of the simulation period is reached.

Chapter 3 The role of spatial patterns of riverbed hydraulic conductivities on characterization of river-aquifer exchange fluxes using a conductance based groundwater model*

3.1 Introduction

Exchange processes between surface water and groundwater play an essential role for hydrology, hydrogeology, ecology, and water resources management (Brunke and Gonser, 1997; Hayashi and Rosenberry, 2002; Sophocleous, 2002). The main uncertain factors for predicting river-aquifer water exchange fluxes are riverbed and aquifer properties (Storey et al., 2003; Saenger et al., 2005). A better characterization of riverbed structures representing more realistic properties may lead to an improved estimation of river-aquifer exchange fluxes (Kurtz et al., 2012). Traditionally, these media are considered homogeneous (Fox and Durnford, 2003) and the models for quantifying the exchange fluxes are simplified.

Field measurements and inverse modelling show that in the real world riverbed hydraulic conductivities may vary over several orders of magnitude (Calver, 2001). Several field surveys also indicate that the spatial distribution of riverbed hydraulic conductivities exhibit non-Gaussian features (Springer et al., 1999; Genereux et al., 2008; Leek et al., 2009; Cheng et al., 2011; Sebok et al., 2015). Springer et al. (1999) found a bimodal distribution of K for five reattachment bars in the Colorado River (Grand Canyon National Park, USA). Genereux et al. (2008) conducted a detailed field experiment in a 250 m long river reach of West Bear Creek (North Carolina, USA) and found that measured riverbed hydraulic conductivities are neither normally nor log-normally distributed. Cheng et al. (2011) measured vertical streambed hydraulic conductivities at 18 sites along a 300 km reach of Platte River (Nebraska, USA) and evaluated whether the measured values were normally distributed.

* adapted from: Tang, Q., Kurtz, W., Brunner, P., Vereecken, H., and Franssen, H.-J. H., 2015, Characterisation of river–aquifer exchange fluxes: The role of spatial patterns of riverbed hydraulic conductivities: Journal of hydrology, v. 531, p. 111-123.

For nine sites a normal distribution could be confirmed by several statistical tests. However, for the other sites the statistical tests were not significant, which was attributed to the presence of river tributaries with varying sediment loads. Several studies also suggest that there can be a distinct spatial pattern of cross-sectional river bed hydraulic conductivities (Genereux et al., 2008; Min et al., 2013; Sebok et al., 2015), which is thought to be related to flow velocity dependent spatially distinct sedimentation and erosion patterns. Some papers (Genereux et al., 2008; Leek et al., 2009; Sebok et al., 2015) additionally provide maps of the spatial distribution of measured riverbed conductivities showing spatial patterns that can hardly be described by a purely Gaussian distribution.

Flow and transport modelling indicates that heterogeneity of riverbed properties has a large impact on river-aquifer exchange fluxes (Wroblicky et al., 1998; Woessner, 2000; Salehin et al., 2004; Kalbus et al., 2009; Irvine et al., 2012; McCallum et al., 2014). In earlier work, we analyzed temporal changes in riverbed hydraulic conductivities, which could be generated by floods and sedimentation processes (Kurtz et al., 2012). It was found that sequential data assimilation can detect the changes in the riverbed with some delay. Kurtz et al. (2014) analyzed the value of temperature measurements to characterize heterogeneous riverbeds. In other works it was analyzed whether heterogeneous riverbeds (with Gaussian distributed heterogeneous riverbed conductivities) can be replaced with a few zones with spatially homogeneous riverbed conductivities (Kurtz et al., 2013). In practice not enough detailed knowledge is available on the spatial variation of riverbed hydraulic conductivities and Gaussian statistics are used for modelling, if heterogeneity is taken into account at all. However, non-Gaussian patterns probably have a significant influence on the magnitude and the spatial patterns of river-aquifer exchange fluxes, which can be of great importance for the prediction of transport processes of heat and contaminants in river-aquifer systems. Non-multi-Gaussian patterns of riverbed hydraulic conductivities could result in very different net exchange fluxes between streams and aquifers compared to multi-Gaussian distributions with the same geostatistical parameters. It was demonstrated that non-multi-Gaussian patterns in aquifers result in a flow and transport behavior which is very different from multi-Gaussian patterns with the same global statistics (e.g., Gómez-Hernández and Wen, 1998; Zinn and Harvey, 2003). Fleckenstein et al. (2006) and Frei et al. (2009) represented facies distribution of aquifer heterogeneities and investigated the dynamics of river–aquifer exchange fluxes. However, in their studies, only aquifer heterogeneities were treated as non-multi-Gaussian and riverbed hydraulic conductivities were the same as the underlying aquifer hydraulic conductivities. Consequently, until now, such non-multi-Gaussian patterns have not been taken into account for the generation of riverbed hydraulic conductivities; neither were non-multi-Gaussian distributed conductivities updated using inverse methods or data assimilation. This study therefore focuses on investigating the impact of the non-multi-Gaussian distribution of riverbed hydraulic conductivities on model states and river-aquifer exchange fluxes.

A number of already established simulation techniques developed to characterize the spatial variability of aquifer heterogeneities (Zinn and Harvey, 2003; Khodabakhshi and Jafarpour, 2013) can also be applied for the characterization of spatially variable riverbed structures. Geostatistical simulation techniques can model spatial heterogeneity by generating equally likely stochastic realizations of the spatially variable geological medium. One typical approach is the sequential simulation algorithm (Gómez-Hernández and Journel, 1993) based on a variogram to generate a conditional realization from a multi-Gaussian random function. Elfeki and Dekking (2001) proposed a Markov chain model to characterize geological heterogeneities constrained on well data. Another approach is the multiple-point (MP) geostatistical technique (Guardiano and Srivastava, 1993) which expanded the traditional sequential simulation by avoiding the definition of a random function based on two-points geostatistics (Hu and Chugunova, 2008). A comparison between simulations generated by the multiple point geostatistical method and variogram-based geostatistics showed that the reproduction of the hydraulic conductivity field generated by MP methods can better represent certain geological media (Mariethoz et al., 2010). We assume that the multiple point geostatistical method can also be used to generate more realistic parameter distributions of riverbeds. A next step is the inverse conditioning of the non-multi-Gaussian parameter distribution to hydraulic head data.

Inverse modelling techniques are also called indirect methods which encompass model identification and parameter estimation. Carrera et al. (2005) reviewed the recent progress of inverse modelling for aquifer characterization and tried to find similarities between well-established methods, including the pilot point method, zonation method and sequential self-calibration. Carrera and Neuman (1986) used a maximum likelihood method called the zonation method to estimate hydraulic conductivities and possibly other parameters for a limited number of zones in which the aquifer is divided. The division of the aquifer in a limited number of zones reduces the number of parameters to be estimated and allows a unique, stable solution of the inverse problem. Carrera and Neuman (1986) proposed the solution of the inverse problem by an iterative approach solving the groundwater flow problem, which results in a hydraulic head solution which is consistent with the parameters. RamaRao et al. (1995) proposed the pilot point method for solving the inverse problem in groundwater flow systems, locating pilot points where there are no measurements. The sequential self-calibration method was proposed by Gómez-Hernández et al. (1997) and generates equally likely realizations of transmissivity fields conditioned to both transmissivities and heads. The main step forward of this approach is that a non-unique solution is sought to the inverse problem and multiple equally likely solutions are calculated. A comparison of seven inverse modelling methods for groundwater flow was presented by Hendricks Franssen et al. (2009). They showed that Monte Carlo based inverse modelling methods, which calculate multiple equally solutions to the inverse problem, generally outperform other inverse methods.

The Ensemble Kalman Filter (EnKF) (Evensen, 1994) is a Monte Carlo based inverse method. Instead of calculating one solution with a dynamical simulation model (in this paper a hydrological model) multiple solutions are calculated. The multiple solutions are calculated for different model inputs, like for example different spatial distributions of input parameters. Also other model input can be made uncertain. The different model inputs characterize the model input uncertainties and are sampled from multivariate probability density functions. The multiple solutions are used to calculate the model covariance matrix, containing the covariances between all model states. EnKF is a purely stochastic method because the observations are treated as random variables by adding perturbations to the measurements (Burgers et al., 1998). EnKF can be extended to estimate parameters together with states and was applied for estimating hydraulic conductivities for a transient groundwater flow problem by Chen and Zhang (2006). As it is suited to condition to observations and performs well for non-linear models, it becomes a robust tool to deal with flow and transport problems in complex geological media. Hendricks Franssen and Kinzelbach (2008) used EnKF combined with the 2-D saturated transient groundwater flow equation to update both model states and parameters. A damping factor was introduced to avoid the filter inbreeding problem, which is an underestimation of the model variance related to a limited number of ensemble members used to approximate the model covariance matrix. Camporese et al. (2009) estimated stream flow using the CATHY model and incorporated the assimilation of both stream flow and pressure head data for two synthetic cases: a soil column experiment and a V-catchment scale experiment. Results showed that EnKF increased the accuracy of the prediction. Also Bailey and Bau (2010; 2012), and Pasetto et al. (2012), amongst others, analyzed joint assimilation of pressure and discharge data for a coupled surface-subsurface problem. These papers however did not focus specifically on the riverbed and did not update riverbed hydraulic properties. Hendricks Franssen et al. (2011) jointly updated K_{aq} and leakage coefficients in a real-time application of EnKF for the Limmat Valley aquifer in Zurich, Switzerland. Kurtz et al. (2012) characterized time-dependent leakage coefficients by updating time-dependent model parameters with EnKF using covariance inflation to improve the estimation performance. Kurtz et al. (2013) updated spatially variable leakage coefficients.

EnKF only works optimally when parameters or states follow a Gaussian distribution. Kerrou et al. (2008) investigated and compared the ability of direct methods and inverse modelling in characterizing non-multi-Gaussian parameter fields under the wrong assumption of a multi-Gaussian random function model. They observed that two points multi-Gaussian techniques were not able to detect the non-multi-Gaussian structures and may lead to inaccurate groundwater flow and mass transport predictions. However, it is unclear whether river-aquifer exchange fluxes are affected in a similar way to flow and transport in aquifers. One approach to deal with non-Gaussian patterns of states and parameters is provided by Sun et al. (2009). They combine a grid-based EnKF with a Gaussian mixture modelling approach for handling non-Gaussian distributions. Another solution is the implementation of a normal score transformation which renders states and parameters Gaussian (Zhou et al., 2011) using anamorphosis functions for log-conductivities and piezometer heads. Zhou et al. (2011) showed that the Normal Score Ensemble Kalman Filter (NS-EnKF) outperforms the standard EnKF both in terms of characterizing hydraulic conductivity and updating piezometer heads, and in a later work (Zhou et al., 2012) it was found that superior results are achieved even when no hydraulic conductivity data are available. Li et al. (2012) showed that NS-EnKF still outperforms EnKF in case of a wrong assumption about the prior geostatistical model, but a reasonable estimate of the histogram of the conductivity data is needed for the model to be successful. Schöniger et al. (2012) indicated that for variables showing complex dependency structures in space like solute concentration, NS-EnKF might not outperform classical EnKF. Crestani et al. (2013) corroborated these results and found that a modified NS-EnKF, with one normal score transformation for all grid nodes (not grid node specific) gave better results than the original NS-EnKF.

In this work, as a synthetic reality a non-multi-Gaussian distribution of riverbed hydraulic conductivities is assumed. Note that previous studies considered riverbed heterogeneities as Gaussian distributed patterns. However, as motivated before, in the real world non-Gaussian riverbed properties are expected to be common. Given this non-Gaussian distribution, we want to analyze the impact of the adopted geostatistical model (either the multi-Gaussian one or a non-multi-Gaussian distribution) on the identification of model states, parameters and fluxes that characterize the river-aquifer interaction. How significant are the inaccuracies introduced if we impose a multi-Gaussian distribution of riverbed properties when in reality properties are non-multi-Gaussian This work aims at analyzing how well data assimilation methods like the EnKF or the NS-EnKF can inversely condition these riverbed conductivities, and whether the NS-EnKF outperforms the more classical EnKF for non-multi-Gaussian distributions. The non-multi-Gaussian fields of riverbed hydraulic conductivities are modelled with multiple point geostatistical methods. In summary, this paper investigates: (i) how important it is to represent non-multi-Gaussian distributions of riverbed hydraulic conductivities in the model (comparison with multi-Gaussian assumption and an alternative erroneous non-multi-Gaussian assumption), and (ii) whether NS-EnKF can outperform classical EnKF for the characterization of non-multi-Gaussian riverbeds under the non-multi-Gaussian assumption. For the first objective, experiments with three different geostatistical models for riverbed characterization were made: a non-multi-Gaussian model exhibiting channels, a non-multi-Gaussian model with the same bimodal histogram but without channels and a multi-Gaussian model with the same mean and variance as the non-multi-Gaussian models. Next, EnKF experiments were performed for the three geostatistical models by assimilating hydraulic heads with update of both states and parameters. Additionally, NS-EnKF was implemented for the two non-multi-Gaussian models to evaluate the performance of normal score transformation. A series of numerical experiments are carried out with a three-dimensional synthetic river-aquifer model for different scenarios. The performance of EnKF/NS-EnKF is evaluated in terms of its ability to predict the hydraulic heads, reproduce the K_{rb} fields and estimate the exchange fluxes between river and aquifer.

3.2 Methods and Materials

The governing continuity hydraulic equation for a 3D unsaturated-saturated groundwater flow problem in a river-aquifer system is described by the three dimensional Richards' equation 2.1. The software SPRING (Delta h Ingenieugesellschaft mbH, 2006) is used to solve the 3D unsaturated-saturated groundwater flow problem in a river-aquifer system with help of the finite element method, using regular rectangles in the simulation studies presented in this study. We present here the NS-EnKF in terms of the joint updating of piezometric heads and riverbed hydraulic conductivities (expressed as leakage coefficients (α)) with help of assimilation of piezometric heads to simulate river-aquifer interactions. The main steps and equations for NS-EnKF are summarized as follows:

1. Generation of initial ensemble. In our work, non-multi-Gaussian distributed parameter fields are generated using the direct sampling algorithm (Mariethoz et al., 2010), a multiple-point geostatistical simulation technique (Caers and Zhang, 2004; Zhang, 2008). Both non-multi-Gaussian fields with channels and non-multi-Gaussian fields without channels are generated by the direct sampling algorithm. multi-Gaussian distributed parameter fields are generated using a sequential multi-Gaussian simulation technique (Gómez-Hernández and Journel, 1993).

2. Forecast of state vectors. In this step, the states at the current time step are estimated from the previous time step by the transient flow model. The stochastic realizations of $\log_{10}(\alpha)$, together with other input information, are used here as input to solve the 3-D river-aquifer flow problem with the software SPRING. The calculated piezometric heads are denominated "forecasted heads" in order to distinguish them from the updated heads after data assimilation. For each of the realizations, the prediction equation is given by:

$$\mathbf{x}_{t,i} = M\left(\mathbf{x}_{t-1,i}\right) \tag{3.1}$$

where $\mathbf{x}$ is the augmented state vector containing the model states (here the simulated piezometric heads) and parameters (here the log-transformed leakage coefficient):

$$\mathbf{x}_i = \begin{pmatrix} \mathbf{h} \\ \mathbf{Y} \end{pmatrix}_i \tag{3.2}$$

where **h** is the simulated heads and $\mathbf{Y} = \log_{10}(\boldsymbol{\alpha})$. *M* is the simulation model (SPRING in this work) , *t* is the time step counter and *i* is the realization number.

3. Normal-Score transformation. This is only done for certain simulation scenarios (see section 3.3.4). The ensemble of state vector forecasts and leakage coefficient fields provides the basis for the normal-score transform of heads and leakage coefficients. At each location and time step a probability density function of hydraulic heads and leakage coefficients can be constructed. In addition, both simulated piezometric heads and leakage coefficients are transformed by creating

anamorphosis functions for each variable independently, at all grid cell locations and all time steps according equations 2.19-2.22.

4. Updating. In this step simulated heads, possibly transformed, from the model are compared with measurements, which are also transformed in case the simulated heads are transformed. The original measurements are perturbed with a series of normally distributed measurement errors. These measurement errors are defined a priori on the basis of expert knowledge:

$$\mathbf{y}_{t,i} = \mathbf{y}_t + \boldsymbol{\varepsilon}_{t,i} \tag{3.3}$$

where $\mathbf{y}_{t,i}$ is the perturbed measurement vector, $\mathbf{y}_t$ is the vector with the measurements at the current time step and $\boldsymbol{\varepsilon}_{t,i}$ is the vector containing random errors. If the perturbed measurements are transformed, this is done using the same anamorphosis function as for the simulated heads at that location. The transformed perturbed measurements can be written as:

$$\mathbf{y}_{t,i}^{trans} = \phi\left(\mathbf{y}_{t,i}\right) \tag{3.4}$$

where $\phi(\cdot)$ is the anamorphosis function created based on the simulated heads and the interpolation/extrapolation. Next, the Kalman gain **G** is calculated as:

$$\mathbf{G} = \mathbf{C}\mathbf{H}^{\mathrm{T}}\left(\mathbf{H}\mathbf{C}\mathbf{H}^{\mathrm{T}} + \mathbf{R}\right)^{-1} \tag{3.5}$$

where **C** is the covariance matrix of the, possibly transformed, states and parameters and **R** is the covariance matrix of, possibly transformed, measurement errors. The structure of **C** for the example of transformed states and variables is as follows:

$$\mathbf{C} = \begin{pmatrix} \mathbf{C}_{\mathbf{h}^{\mathbf{trans}}\mathbf{h}^{\mathbf{trans}}} & \mathbf{C}_{\mathbf{h}^{\mathbf{trans}}\mathbf{Y}^{\mathbf{trans}}} \\ \mathbf{C}_{\mathbf{Y}^{\mathbf{trans}}\mathbf{h}^{\mathbf{trans}}} & \mathbf{C}_{\mathbf{Y}^{\mathbf{trans}}\mathbf{Y}^{\mathbf{trans}}} \end{pmatrix} \tag{3.6}$$

The analysis equation updates the transformed, simulated heads (and transformed $\log_{10}$ (α)) accounting for the observations:

$$\mathbf{x}_{t,i}^{a,trans} = \mathbf{x}_{t,i}^{trans} + \mu\mathbf{G}\left(\mathbf{y}_{t,i}^{trans} - \mathbf{H}\mathbf{x}_{t,i}^{trans}\right) \tag{3.7}$$

where $\mathbf{x}_i^{trans} = \begin{pmatrix} \mathbf{h}^{trans} \\ \mathbf{Y}^{trans} \end{pmatrix}_i$. It should be noticed that it is also possible to transform only the heads or only $\log_{10}(\alpha)$ or none of the two. In the latter case the updating equation reduces to the standard EnKF-method for joint updating of states and parameters.

5. Back transformation if transformation was done at step 3. The updated heads (and $\log_{10}(\alpha)$) are back transformed after data assimilation using the inverse anamorphosis functions. For this step, the original anamorphosis functions are used but now in general interpolation/extrapolation is needed. After finishing this step, the algorithm returns to step 2.

3.3 Synthetic experiments

3.3.1 Model setup

We carried out numerical experiments for a simplified synthetic three-dimensional river-aquifer model with a domain size of 500m × 250m × 10m, see Figure 3.1. The model is discretized into 125,000 grid cells at a spatial resolution of 10m × 10m × 0.1m. For two reasons a high vertical model resolution was chosen. The first reason is the improved representation of vertical variations in saturation. The second reason is the improvement of numerical stability. The simulation period of one year is discretized into 365 time steps. A river is conceptualized into eight lines of river nodes (408 nodes totally), which are situated in the top layer of the domain and are in fact nodes which represent the riverbed-aquifer interface, connected to the river. The river nodes are situated in the middle of the simulation domain. We assume that the aquifer is homogeneous and only the river bed is spatially heterogeneous. The aquifer has a log-transformed hydraulic conductivity of -3 $\log_{10}$ m/s. The initial river stage at the western boundary is 410m and decreases to the eastern boundary with a slope of 0.01. The model forcing data for this synthetic model are transient river stages calibrated from real world discharge data of river Sihl which are taken from Kurtz et al. (2014). See Figure 3.2 (a) for further details. Otherwise, both western and eastern boundaries are impermeable. Northern and southern boundaries are assigned prescribed heads with a yearly cycle whose values are given by a phase shifted sine function with an amplitude of 1m (see Figure 3.2 (b)). Along the northern and southern boundaries, time dependent fixed heads are spatially homogeneous. A model spin-up of 50 days, using average hydrologic conditions, is made for both the reference runs and all other ensemble runs so that model states and parameters are in dynamic equilibrium and the values for the initial states are meaningful. The final simulated heads from the spin-up period were used as initial heads for the forward model run. This model set-up results in infiltration of river water into the aquifer in the western part of the model domain during most of the simulation period, whereas the opposite occurs most of the time in the eastern part of the simulation domain. Partially saturated conditions below the riverbed prevail in the eastern part of the domain. At 30 observation points piezometric head is monitored (ten measurements directly below the riverbed, for the uppermost model layer and others north and south of the river, also for the uppermost model layer) and no observations are available for constraining parameters (K and α). See also Figure 3.1 for the position of the measurement locations. Figure 3.2 provides a summary of the model dynamics and boundary conditions for both the reference runs and data assimilation runs. The temporal pattern of river stage variation is the same for each river node, but the absolute values of the river stage vary across the river nodes. Hydraulic parameters according to van Genuchten were employed for characterizing flow under unsaturated conditions, using three different parameter combinations as provided by the software SPRING (for further details of the parameter values see Kurtz et al., 2013).

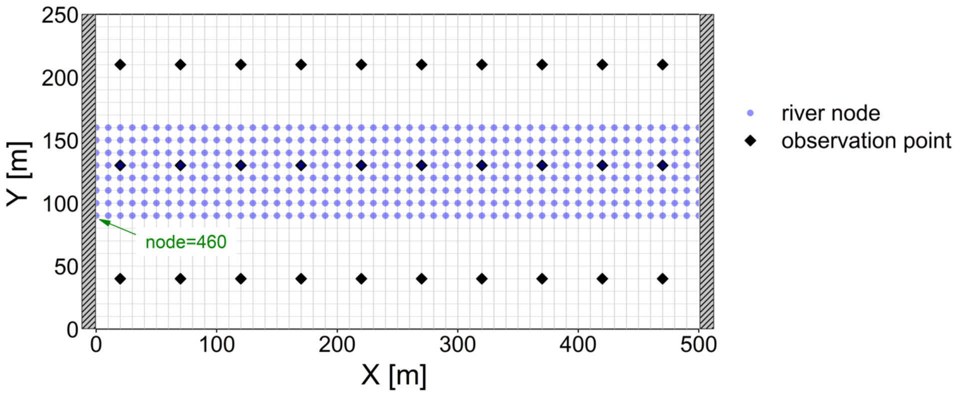

Figure 3.1: Setup of the synthetic 3D model including the observation points and the nodes which are located below the river (in blue). Node number 460 is also indicated as a reference for Figure 3.2.

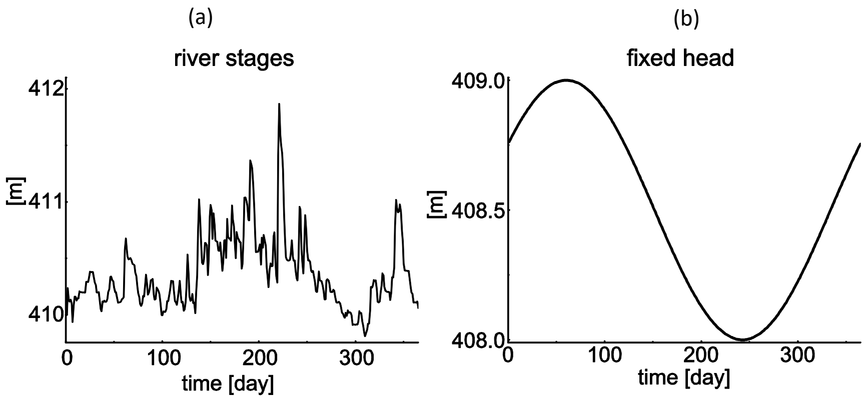

Figure 3.2: Model dynamics and boundary conditions: (a) temporal evolution of river stages at node=460; (b) temporal evolution of fixed heads along northern and southern boundaries.

3.3.2 Multiple reference runs

For the synthetic experiment in total ten different reference fields of riverbed hydraulic conductivities were defined. The performance of different prior geostatistical models (see section 3.3.3) was evaluated for each of the ten reference fields. Ten reference fields were generated instead of one reference field because results can be quite influenced by the specific features generated in the reference field (e.g., Schöniger et al., 2012).

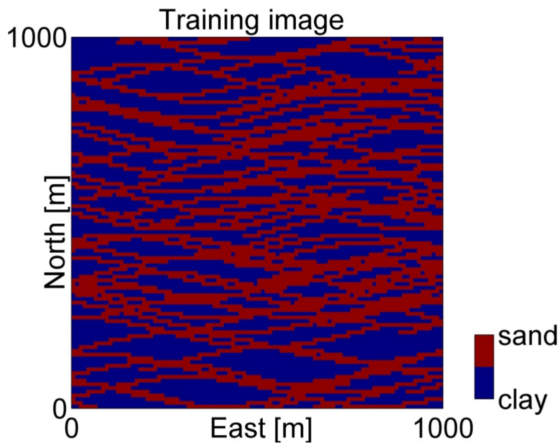

Figure 3.3: The training image used for the generation of the ten non-multi-Gaussian reference fields with channels.

For the generation of the ten non-multi-Gaussian reference fields with channels, a training image is generated by SGeMS (Remy et al., 2009), which reflects the channelized structures, see Figure 3.3. The training image is composed of two different facies, and the fraction of highly permeable facies (in red) is set to 0.4. In the training image, the channels correspond to sand and the background represents low permeable materials like clay. The direct sampling method (Mariethoz et al., 2010) is used as a multiple-point geostatistical simulation technique to generate the patterns for the non-multi-Gaussian references with channels. In order to complete the generation of the non-multi-Gaussian reference fields, the two facies were independently populated with log transformed leakage coefficient values generated by sequential Gaussian simulation using the GCOSIM3D code (Gómez-Hernández and Journel, 1993). Table 3.1 gives the geostatistical parameters for generating the multi-Gaussian patterns within each of the facies of the non-multi-Gaussian models. A spherical variogram model was adopted in all cases. The mean $\log_{10}(\alpha)$ of the riverbed channels was three orders of magnitude larger than the little permeable parts of the riverbed.

Table 3.1: Variogram parameters used to generate stochastic realizations of $\log_{10}(\alpha)$ for each of the two facies in the non-multi-Gaussian reference fields with channels. Background represents the material between the highly permeable channels/ellipsoidal structures. The same parameters were used to generate stochastic realizations for the non-multi-Gaussian model with channels and the non-multi-Gaussian model with ellipsoidal structures. The geostatistical parameters for creating stochastic realizations for the multi-Gaussian model are also given.

Facies	Variogram type	Mean ($\log_{10}$(m/s))	Range (m)*	Sill($\log_{10}$(m/s))
Channel /ellipsoidal	Spherical	-2.0	100	0.5
Background	Spherical	-5.0	100	0.5
multi-Gaussian	Spherical	-3.8	100	2.7

*: in all directions (x, y and z).

The resulting reference $\log_{10}(\alpha)$ fields are shown in Figure 3.4. Each reference field shows a pronounced bimodality with distributed connected channels (one example of the histograms for these reference fields is given in Figure 3.5). Forward model runs are made for the ten reference fields for a simulation period of 365 days and the calculated heads serve to collect synthetic hydraulic head observations at 30 observation points. These synthetic data are assimilated on a daily basis in the data assimilation experiments. Piezometric head measurements are unbiased with standard deviation of 0.05m. No (riverbed) hydraulic conductivity data are available as direct data in the experiments.

3.3.3 Three geostatistical models for generation of riverbed heterogeneities

Three different a priori geostatistical models (used in data assimilation with EnKF) were used to investigate the role of patterns of riverbed heterogeneities:

(i) Non-multi-Gaussian distributed riverbed properties with connected channels (same geostatistical model as reference).

(ii) Non-multi-Gaussian distributed riverbed properties with the same bimodal distribution of $\log_{10}(\alpha)$ as in (i), but with ellipsoidal instead of channelized structures.

(iii) multi-Gaussian distributed riverbed properties.

Notice that the only difference between (i) and (ii) is the spatial organization of the riverbed hydraulic conductivities, the main difference being the spatial continuity, which is much higher for the channelized structures than for the ellipsoidal structures.

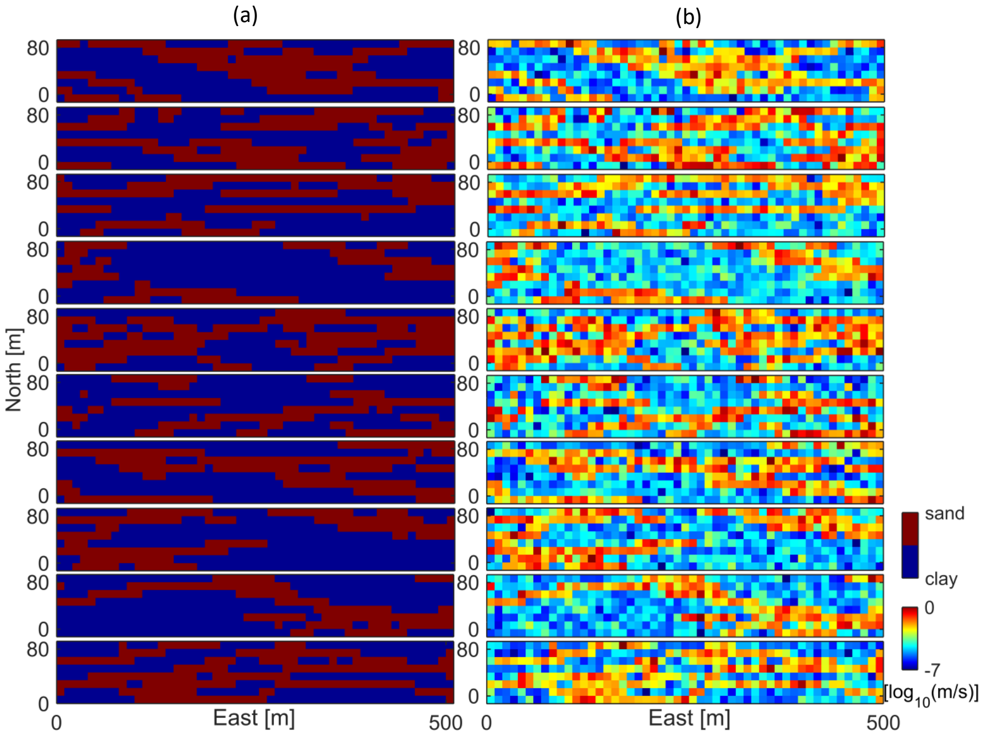

Figure 3.4: (a) The ten reference facies distributions and (b) the ten associated reference distributions of $\log_{10}$ (α).

Table 3.1 provides also the basic statistics used to generate the stochastic realizations for each of the geostatistical models. Figure 3.5 shows the log-conductivity histograms for the three geostatistical models.

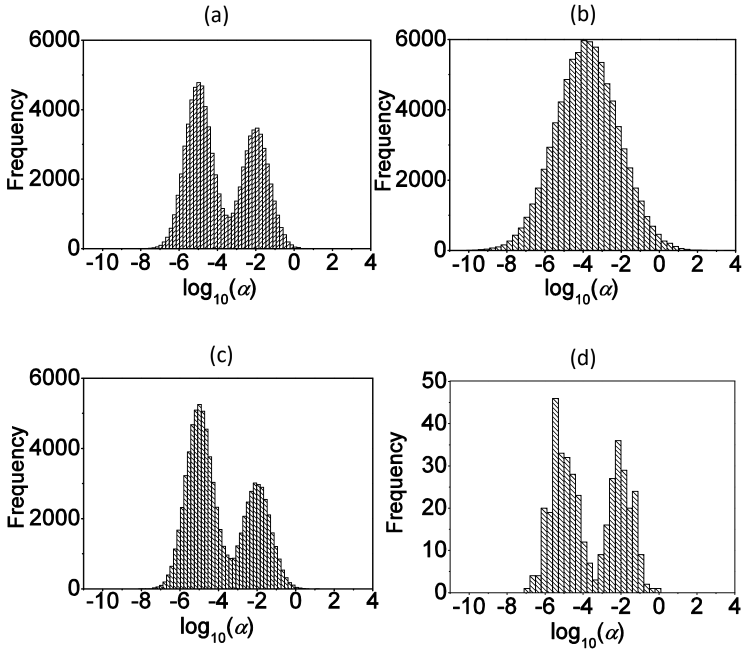

Figure 3.5: Histograms of $\log_{10}(\alpha)$ calculated over all river nodes for the non-multi-Gaussian model with channels (a), the multi-Gaussian model (b), and the non-multi-Gaussian model without channels (c). An example of a histogram for a true reference field (Reference field number 1) is shown as (d).

Stochastic realizations for the two non-multi-Gaussian models are generated with help of two training images created by SGeMS (Remy et al., 2009). The training image for the non-multi-Gaussian model with channels is the same as detailed in section 3.3.2, and the training image for the non-multi-Gaussian model with ellipsoidal structures is shown in Figure 3.6. This last training image is composed of two different facies, again with a proportion of 0.4 for the highly permeable facies to make the two training images comparable for geostatistical simulation. The direct sampling method (Mariethoz et al., 2010) is also used to generate facies distributed patterns for the two non-multi-Gaussian models.

Stochastic realizations for the two non-multi-Gaussian models are completed by populating the two facies independently with log transformed leakage coefficient values. These leakage values are generated by sequential Gaussian simulation using the GCOSIM3D code (Gómez-Hernández and

Journel, 1993), see Figure 3.6. Table 3.1 gives the geostatistical parameters for generating the multi-Gaussian patterns within each of the facies of the non-multi-Gaussian models.

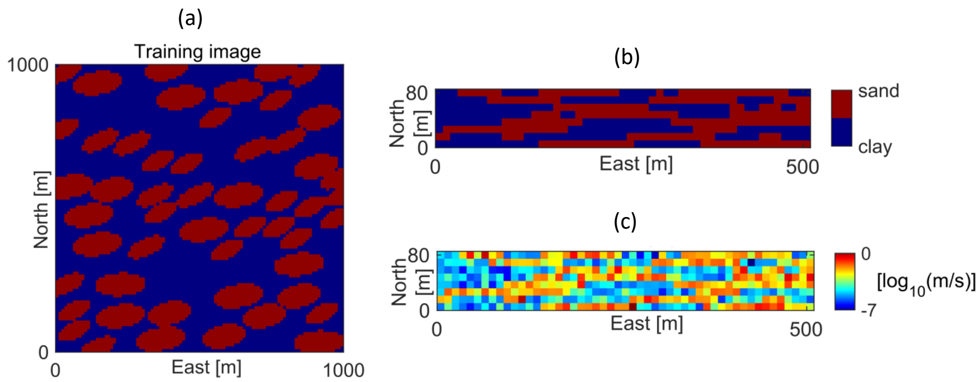

Figure 3.6: Training images used to generate an ensemble of $\log_{10}(\alpha)$ realizations for non-multi-Gaussian models with ellipsoidal structures (a). The training image is squared and deviates from the rectangular riverbed as the training image serves as a database for geological structures, from which samples are taken. An example is provided of a stochastic realization generated from the training image with channelized structures. First, a binary pattern of facies is generated with the direct sampling method (b). Second, within each facies multi-Gaussian distributed values are generated with GCOSIM3d (c).

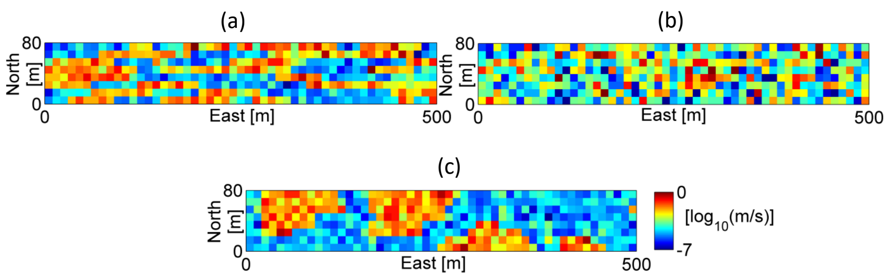

Figure 3.7: Examples of stochastic realizations of $\log_{10}(\alpha)$ drawn from the non-multi-Gaussian model with channels (a), the multi-Gaussian model (b) and the non-multi-Gaussian model without channels (c).

The multi-Gaussian stochastic realizations of $\log_{10}(\alpha)$ are generated using GCOSIM3D with the same arithmetic mean and variance as for the non-multi-Gaussian models. Figure 3.7 provides examples of generated stochastic realizations of K_{rb} fields, which are later used as initial ensembles of $\log_{10}(\alpha)$. Table 3.1 gives the geostatistical parameters used for defining the multi-Gaussian random model.

The variance of the riverbed $\log_{10}(\alpha)$ for the multi-Gaussian case was selected in order to get a variance which is similar to the overall variance of the non-multi-Gaussian case.

3.3.4 Data assimilation experiments

Simulations were performed with 200 stochastic realizations of leakage coefficients for eight scenarios and for each of the ten reference fields. The number of stochastic realizations was limited to 200 given restrictions on CPU-time. The stochastic realizations generated according to the three geostatistical models are the starting point for conditioning to hydraulic head data with help of data assimilation methods. Piezometric heads and leakage coefficients are updated with assimilation of piezometric heads. A damping factor α of 0.1 (Section 3.2.2) is used in the EnKF/NS-EnKF scheme. For comparison purpose, unconditional (open loop) simulations are performed with exactly the same model settings.

Table 3.2: Definition of the eight simulation scenarios (open loop and five different data assimilation scenarios).

Assumption	Scenarios	Update h	Update α	Transform h	Transform α
non-multi-Gaussiar (channel)	1: open loop	×	×	×	×
	2: hl	√	√	×	×
	3: hl_hl	√	√	√	√
non-multi-Gaussian (ellipsoidal)	4: ellip_open loop	×	×	×	×
	5: ellip_hl	√	√	×	×
	6: ellip_hl_hl	√	√	√	√
multi-Gaussian	7:multi_open loop	×	×	×	×
	8: multi_hl	√	√	×	×

√: yes. ×: no.

A detailed description of different scenarios is given in Table 3.2. As indicated before, three different geostatistical models were evaluated and for each of these models open loop simulations were performed (scenarios 1, 4 and 7). For each of the geostatistical models also joint updating of hydraulic heads and leakage coefficients with EnKF was evaluated (scenarios 2, 5 and 8). In addition, for the non-multi-Gaussian models this updating was also evaluated with NS-EnKF (instead of EnKF), these are the scenarios 3 and 6.

3.3.5 Performance assessment

Performance evaluation measure is the root mean square error (RMSE) of model states (simulated heads), model parameters ($\log_{10}(\alpha)$) and leakage fluxes (Q) to quantify the difference between model predictions and reference values:

$$\mathrm{RMSE}(X)=\sqrt{\frac{1}{N_r N_t N_{nodes}}\sum_{r=1}^{N_r}\sum_{t=1}^{N_t}\sum_{i=1}^{N_{nodes}}\left(X_{i,t,r}^{sim}-X_{i,t,r}^{ref}\right)^2} \tag{3.8}$$

where N_r is the number of realizations, N_t is the number of time steps, N_{nodes} is the total number of model nodes over the complete domain, the superscripts *sim* and *ref* refer to the model predictions at a certain time step before data assimilation and the reference values at the same time step, respectively. We calculated *RMSE* also for individual time steps:

$$\mathrm{RMSE}(X,t)=\sqrt{\frac{1}{N_r N_{nodes}}\sum_{r=1}^{N_r}\sum_{i=1}^{N_{nodes}}\left(X_{i,t,r}^{sim}-X_{i,t,r}^{ref}\right)^2} \tag{3.9}$$

RMSE measures how close model predictions are to the true value over the complete model domain. We also calculated the relative error for evaluation, as it expresses the improvement compared to the open loop simulation.

$$R(X)=\frac{RMSE_{scenario}(X)}{RMSE_{openloop}(X)} \tag{3.10}$$

3.4 Results and discussion

The different scenarios described previously were analyzed. The errors for simulated heads, updated $\log_{10}(\alpha)$ and estimated leakage fluxes calculated according to these different scenarios (Table 3.2) averaged over ten references are summarized in Table 3.3.

Table 3.3: Performance measures (standardized, relative errors) for the different scenarios averaged over ten references.

simulation scenarios	$RMSE_{allnodes}(h)$	$RMSE(\alpha)$	$RMSE(Q)$
1: open loop	1.00	1.00	1.00
2: hl	0.73	0.91	0.79
3: hl_hl	0.72	0.93	0.84
4: ellip_openloop	1.00	1.00	1.00
5: ellip_hl	0.62	0.90	0.82
6: ellip_hl_hl	0.60	0.92	0.85
7: multi_openloop	1.00	1.00	1.00
8: multi_hl	0.71	0.95	0.91

3.4.1 Influence of river bed patterns

This subsection centers on the impact of adopting a wrong geostatistical model for the spatial distribution of riverbed hydraulic conductivities and results are compared with the case that the correct model assumption was used. In particular, it is investigated whether the specific spatial pattern of $\log_{10}$ (α) matters; non-multi-Gaussian simulations with connected channels and disconnected ellipsoidal clusters are compared.

First, results are compared for a non-multi-Gaussian model exhibiting channels, a non-multi-Gaussian model with the same bimodal histogram but without channels and a multi-Gaussian model. The comparison is centered here on the case hydraulic head data are used to update model states and $\log_{10}$ (α), with EnKF (scenarios 2, 5 and 8 in Table 3.3). Errors for these three scenarios (in terms of RMSE) are smaller than for open loop simulations (scenario 1). For example, the relative improvement for RMSE (h) evaluated over all nodes was 27% and 29% for the multi-Gaussian and the channelized scenarios, respectively and even 38% for the non-multi-Gaussian ellipsoidal scenario. Updated $\log_{10}$ (α) and leakage fluxes show less improvement, especially for the multi-Gaussian assumption. The characterization of $\log_{10}(\alpha)$ (results for leakage fluxes in brackets) in terms of RMSE improved 5% (9%) for multi-Gaussian fields, 10% (18%) for non-multi-Gaussian fields without channels and 9% (21%) for non-multi-Gaussian with channels. In general, from Table 3.3 it can be observed that for non-multi-Gaussian fields with channels the characterization of leakage coefficients and fluxes improves more with data assimilation than for the other geostatistical models which used the wrong model assumption. However, the differences between the two non-multi-Gaussian models are only small and probably not significant.

The temporal evolution of RMSE for simulated heads for different scenarios is shown in Figure 3.8 (a). At the beginning of the simulation period, we cannot observe a pronounced benefit from data assimilation. However, after 100 days EnKF improves the simulation results compared to the open loop run and the non-multi-Gaussian model with channels performs better than the multi-Gaussian model and the non-multi-Gaussian model with ellipsoidal structures, for this specific reference. In general, data assimilation results in an improved characterization of hydraulic heads compared to the open loop run, but the differences are not very large. To analyze the ability of characterization of riverbed hydraulic conductivities, the temporal evolution of RMSE for updated $\log_{10}(\alpha)$ is shown in Figure 3.8 (b). Figure 3.8 (b) illustrates that a non-multi-Gaussian model with channels outperforms the other geostatistical models.

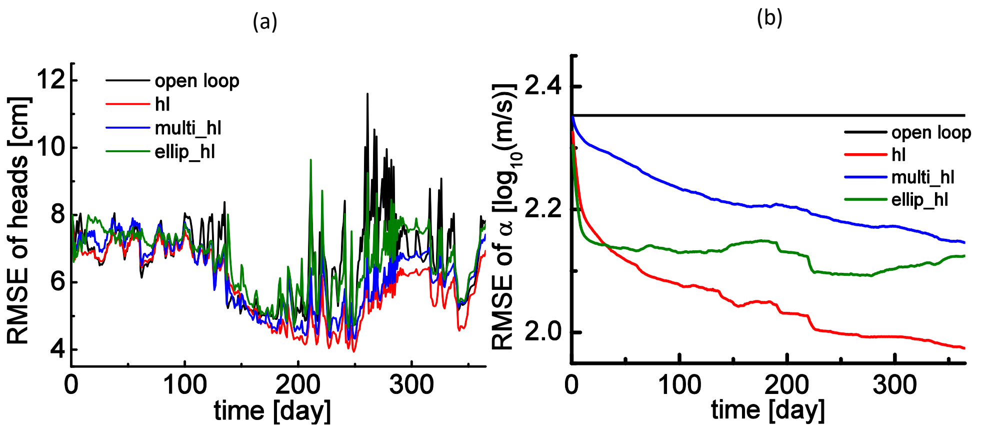

Figure 3.8: Temporal evolution of RMSE for (a) simulated heads (b) and updated $\log_{10}(\alpha)$ evaluated over all model/river nodes and all realizations after each time step (for reference 1). Displayed are the RMSE for the open loop run and scenarios with different patterns for riverbed hydraulic conductivities, conditioned on hydraulic head data with EnKF.

Figure 3.9 displays the final updated ensemble mean $\log_{10}(\alpha)$ fields for the scenarios 2, 5 and 8. After one year of assimilation, the updated $\log_{10}(\alpha)$ field for the case with channels captures the basic structures of the reference field, although the absolute magnitude for the leakage values differs from the reference. For the multi-Gaussian case the updated riverbed patterns are similar to the initial distribution, and hardly recover the basic patterns of the reference field from the measurement data.

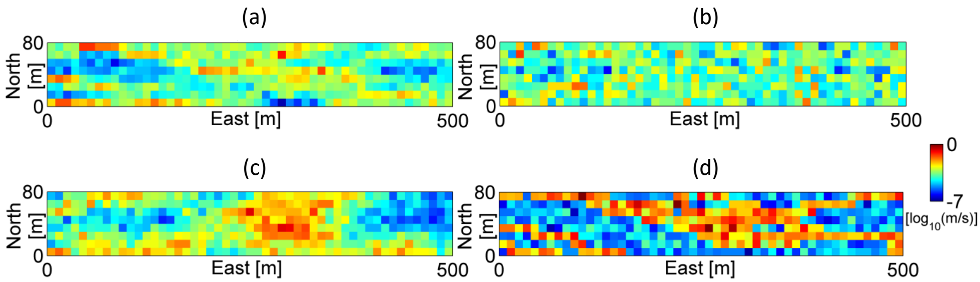

Figure 3.9: The reference $\log_{10}(\alpha)$ field together with final updated $\log_{10}(\alpha)$ fields (ensemble mean) after 365 time steps with EnKF for the three simulation scenarios: (a) hl: non-multi-Gaussian model with channels; (b) multi_hl: multi-Gaussian model; (c) ellip_hl: non-multi-Gaussian model without channels and (d): Reference field number 1.

Although EnKF improves the characterization of K_{rb} for the non-multi-Gaussian conditions, none of them could preserve the bimodal histogram.

Figure 3.10 shows the boxplot of RMSE of piezometric heads, updated $\log_{10}(\alpha)$ and leakage fluxes, calculated over ten reference fields and for different scenarios. Root mean square errors are lower after data assimilation (compared to open loop simulations).

The comparison of the performance of the three geostatistical models reveals that best results are obtained with a non-multi-Gaussian model with channels and the non-multi-Gaussian model with ellipsoidal structures. It is not surprising that the non-multi-Gaussian model with channels performs best, as all ten reference fields were also drawn from this geostatistical model. Again, it is more remarkable that the other non-multi-Gaussian model performs similar, pointing to the fact that the statistical distribution of leakage coefficients matters, but not so much the spatial pattern of it. If –

erroneously – a multi-Gaussian model is assumed for $\log_{10}(\alpha)$, results are worse, especially for characterization of leakage coefficients and to a lesser extent also for flux characterization. However, the characterization of fluxes is not so much affected as typically is found in groundwater studies (e.g., Gómez-Hernández and Wen, 1998; Zinn and Harvey, 2003). For example, synthetic experiments on groundwater flow and transport of solutes in groundwater revealed that breakthrough of contaminants can be much faster in time when channels are present, compared to the multi-Gaussian case (e.g., Wen and Gómez-Hernández, 1998). In addition, it was found that inverse modelling with an erroneous multi-Gaussian assumption for a non-multi-Gaussian aquifer is not able to delineate channels or improve transport predictions significantly, stressing the importance of the correct adoption of the geostatistical model for the aquifer (Kerrou et al., 2008). This might indicate that non-multi-Gaussian riverbed patterns matter less for flow characterization than non-multi-Gaussian aquifer patterns. This is most probably related to the vertical water fluxes through the riverbed, so that horizontally oriented channels have less influence on the flow, whereas in aquifers flow is predominantly horizontal and controlled by channels.

It was assumed in this study that geostatistical parameters which describe spatial heterogeneity of riverbed hydraulic conductivities within the facies are perfectly known. In reality these parameters are (very) uncertain as well. It is expected that this assumption has not a major impact on the outcomes of this study. For the case that a multi-Gaussian assumption was adopted, assuming correlation structures which deviate strongly from the non-multi-Gaussian reference, the performance was not much worse than for the case with the correct, non-multi-Gaussian assumption. It is expected that a minor mistake in the correlation structure within facies would have a smaller impact. For the same reasons, it is expected that a mistake in the training image would also have a more limited impact in this study. Uncertainty of the training image was not studied in this paper, but first studies have been published where uncertainty of the training image was considered (e.g., Khodabakhshi and Jafarpour, 2013).

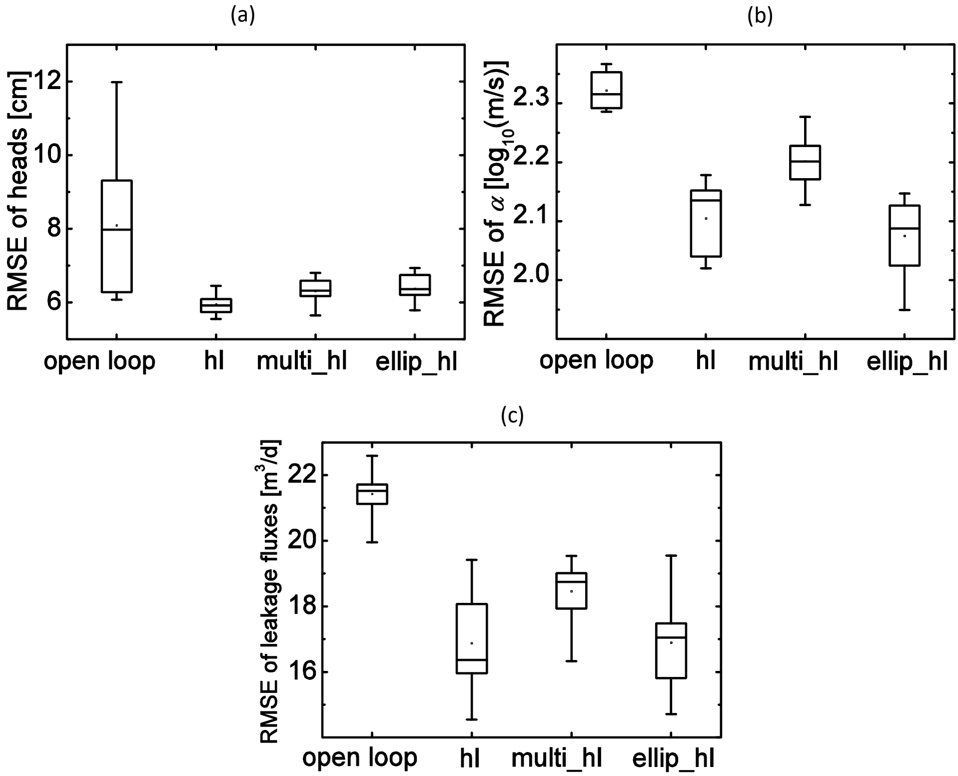

Figure 3.10: Boxplot over ten references for (a) simulated heads, (b) updated $\log_{10}(\alpha)$ and (c) estimated leakage fluxes over all model (river) nodes, all realizations and all time steps for scenarios starting with different K_{rb} patterns.

3.4.2 Performance of NS-EnKF (EnKF versus NS-EnKF)

Results presented in section 3.4.1 were generated by data assimilation using the classical Ensemble Kalman Filter. EnKF is suboptimal for non-multi-Gaussian distributions and therefore NS-EnKF was tested as an alternative data assimilation procedure. The results obtained with EnKF (section 3.4.1) showed that the bimodal distribution of $\log_{10}(\alpha)$ could not be captured even although some of the simulations with EnKF started from a bimodal distribution.

Normal score transformation of both hydraulic heads and $\log_{10}(\alpha)$ was done in the simulation scenarios 3 and 6, which will be compared with the simulation scenarios 2 and 5. For hydraulic head characterization, it is indeed found that NS-EnKF gives better results than classical EnKF, with a 1% - 2%

additional reduction of RMSE (see Table 3.3). However, the characterization of leakage coefficients and leakage fluxes slightly worsens with NS-EnKF (compared to classical EnKF).

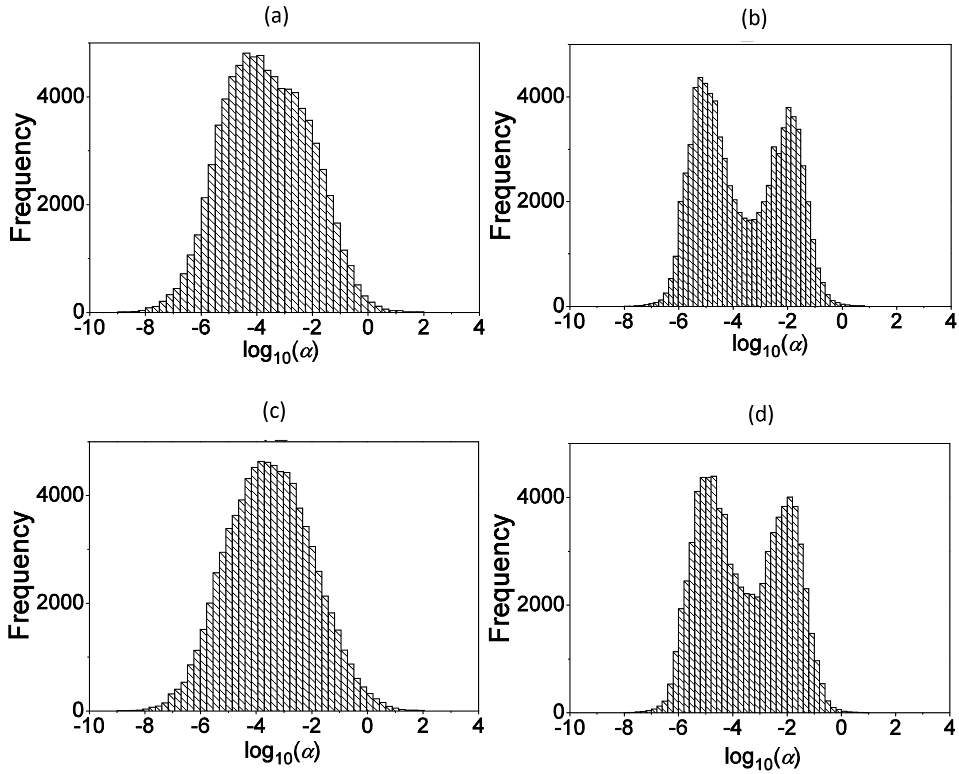

Figure 3.11: Histograms of final updated $\log_{10}(\alpha)$ for different scenarios, displayed are values for all river nodes and all stochastic realizations: (a) scenario hl: non-multi-Gaussian model with channels, updated with EnKF; (b) scenario hl_hl: non-multi-Gaussian model with channels, updated with NS-EnKF; (c) scenario ellip_hl: non-multi-Gaussian model without channels, updated with EnKF; (d) scenario ellip_hl_hl: non-multi-Gaussian model without channels, updated with NS-EnKF.

Figures 3.11 and 3.12 show the histograms and updated leakage coefficient fields, respectively. These figures illustrate that the transformation of $\log_{10}(\alpha)$ allows preserving the shape of the original histogram. Additionally, the updated $\log_{10}(\alpha)$ fields without channels (but assuming non-multi-Gaussian statistics) represent more realistic riverbed patterns after calibration with NS-EnKF (compared to EnKF). These results indicate that NS-EnKF is able to preserve the initial K_{rb} patterns. However, these patterns apparently do not yield an improved estimation of the leakage fluxes and leakage coefficients.

The boxplots in Figure 3.13 show the statistics over ten references which illustrate that both for the non-multi-Gaussian model with channels and the non-multi-Gaussian model without channels, NS-EnKF is not able to outperform EnKF. There might be multiple reasons for these results. First, leakage coefficients and leakage fluxes can only be improved more than for classical EnKF if the piezometric head data are able to provide more information about $\log_{10}(\alpha)$ than in the standard case. However, the sensitivity of head data for identifying channels is not too high. The results for section 3.4.1 already illustrated that the difference between a multi-Gaussian and non-multi-Gaussian prior model in terms of reproduction of leakage coefficients and leakage fluxes was small. This is illustrative for a limited sensitivity to identify channels. A limitation might be that calculations are very CPU-intensive and therefore limited to 200 stochastic realizations. Especially for NS-EnKF it is important to use a high number of stochastic realizations to determine the anamorphosis function and therefore results for NS-EnKF could be affected by sampling errors induced by the relatively small number of stochastic realizations. However, a test for simulations with 500 realizations gave similar results as for 200 stochastic realizations.

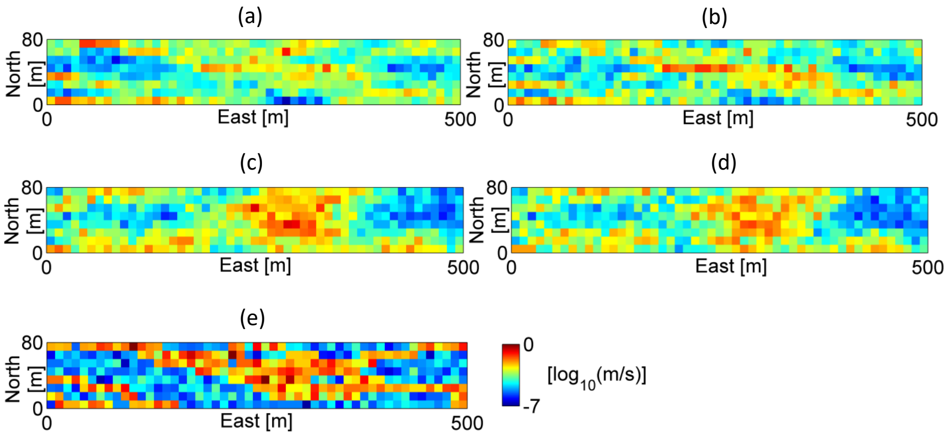

Figure 3.12: The reference $\log_{10}(\alpha)$ field together with final updated $\log_{10}(\alpha)$ fields (ensemble mean) for different scenarios: (a) scenario hl: non-multi-Gaussian model with channels, updated with EnKF; (b) scenario hl_hl: non-multi-Gaussian model with channels, updated with NS-EnKF; (c) scenario ellip_hl: non-multi-Gaussian model without channels, updated with EnKF; (d) scenario ellip_hl_hl: non-multi-Gaussian model without channels, updated with NS-EnKF and (e) reference field number 1.

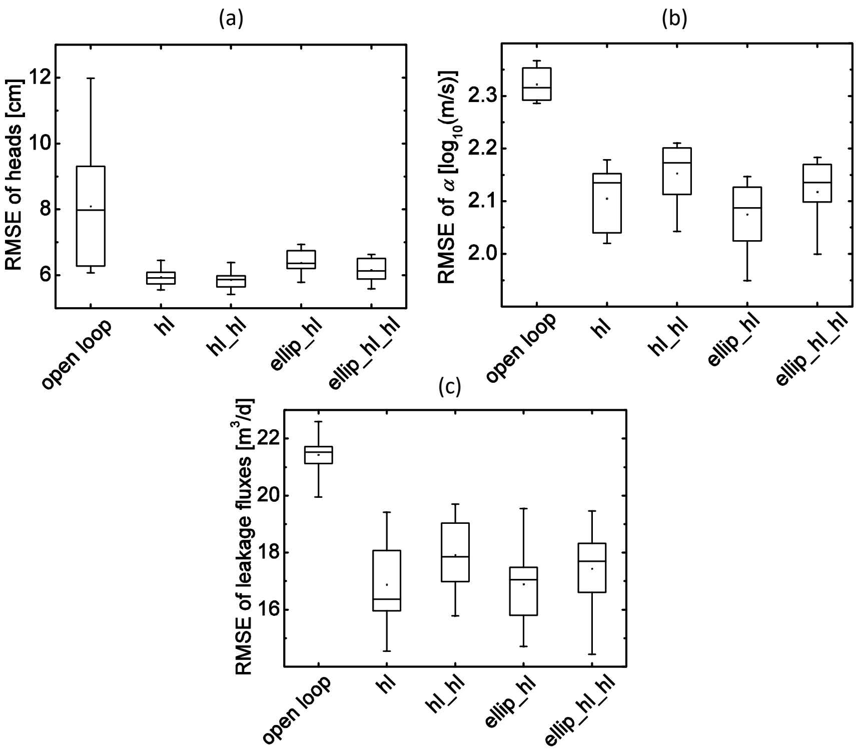

Figure 3.13: Boxplot over ten references for (a) simulated heads, (b) updated $\log_{10}(\alpha)$ and (c) estimated leakage fluxes over all model (river) nodes, all realizations and all time steps for different scenarios updated with EnKF (hl and ellip_hl) or NS-EnKF (hl_hl and ellip_hl_hl).

3.5 Summary and conclusions

This synthetic study investigated for a riverbed displaying non-multi-Gaussian structures like channels the impact of the adopted geostatistical model for the inverse characterization (with the ensemble Kalman Filter (EnKF)) of model states, riverbed properties and river-aquifer exchange fluxes. For aquifers the correct characterization of channels has an important impact on flow and transport predictions. In this study it is found that the reproduction of channels by multiple point geostatistical methods results in a better characterization of states, parameters and fluxes than a multi-Gaussian model, but differences are very small compared to a non-multi-Gaussian model without channels. This can most probably be explained by the predominant vertical flow through the riverbeds, and the expected horizontal orientation of layers and channels. The conclusion is that non-multi-Gaussian riverbed properties have less influence on flow behavior than non-multi-Gaussian hydraulic conductivity distributions in aquifers. We have reached this conclusion using a model that only simulates vertical exchange fluxes between the river and the aquifer. The majority of numerical models (e.g. MODFLOW) are based on a conductance approach, and our analysis suggests that non-multi-Gaussian structures do not improve the predictive value of a conductance type model. However, the question has again to be analyzed in a 3D, fully coupled model for surface water-groundwater interaction. Such a complimentary study will provide further insight to what extent complexity in the geological structure can improve simulation results for a given model type. It is also expected that heat and solute transport simulations are more sensitive to non-multi-Gaussian patterns of riverbed hydraulic conductivities than flow simulations alone. This study tested also whether the Normal Score ensemble Kalman filter (NS-EnKF) was able to give better inversion results than the classical EnKF. This could be expected, as hydraulic heads and leakage coefficients did not follow a Gaussian distribution in this study, and the EnKF performs optimally for Gaussian distributions only, whereas NS-EnKF can handle better bimodal distributions as present in this case. However, the performance of EnKF and NS-EnKF was very similar in this study. This can be related to a too limited sensitivity of piezometric heads to identify channels in the riverbed.

Chapter 4 The influence of riverbed heterogeneity patterns on river-aquifer exchange fluxes under different connection regimes using an integrated hydrological model*

4.1 Introduction

4.1.1 Background

Quantifying river-aquifer interactions is important for understanding flow and transport mechanisms between rivers and aquifers. It is also important for assessing the impact of climate change on water resources (Goderniaux et al., 2009) and balancing human water demands and ecosystem base flow maintenance in arid and semi-arid regions (Zhou et al., 2013). An important requirement for the estimation of river-aquifer exchange fluxes is a proper representation of riverbed properties, e.g. K_{rb} (Brunner et al., 2017). In early studies, riverbeds were often neglected, and if considered they were typically strongly simplified and modeled as homogeneous layers (e.g. Fox and Durnford, 2003). However, field measurements and laboratory analysis showed that riverbed properties can be different from the properties of the underlying aquifer, and that K_{rb} can vary over several orders of magnitude within a single reach (Calver, 2001; Leek et al., 2009). Studies also indicated a large impact of riverbed heterogeneity on the prediction of river-aquifer exchange fluxes. It is therefore important to assess the impact of riverbed heterogeneities in numerical flow models (Kalbus et al., 2009; Irvine et al., 2012; Lackey et al., 2015; Schilling et al., 2017). Irvine et al. (2012) found that a heterogeneous riverbed could be replaced by a homogeneous equivalent while maintaining the accuracy of the prediction of infiltration fluxes in a losing stream system as long as the hydraulic connectivity between river and aquifer is not different between the calibration period and the prediction period.

* adapted from: Tang, Q., Kurtz, W., Schilling, O. S., Brunner, P., Vereecken, H., and Hendricks Franssen, H.-J., 2017, The influence of riverbed heterogeneity patterns on river-aquifer exchange fluxes under different connection regimes: Journal of Hydrology. Under review for Journal of Hydrology.

Based on the investigations of Irvine et al. (2012), Schilling et al. (2017) developed an efficient method to rapidly assess the potential for unsaturated flow conditions to appear in heterogeneous riverbed-heterogeneous aquifer systems using key statistical variables. Lackey et al. (2015) compared the estimation of stream depletion using both homogeneous and heterogeneous K_{rb} fields, which were varying over two orders of magnitude along the streambed. They found that a homogeneous conceptualization of the riverbed properties always led to errors in the estimate.

In several studies the impact of riverbed heterogeneity was assessed by data assimilation (Hendricks Franssen et al., 2011; Kurtz et al., 2012; Kurtz et al., 2013; Kurtz et al., 2014). Hendricks Franssen et al. (2011) successfully estimated spatially variable leakage coefficients (together with spatially variable K_{aq}) with data assimilation for the river - groundwater flow system in the upper Limmat valley in Switzerland. Kurtz et al. (2013) explored whether high resolution characterization of spatially heterogeneous riverbeds was required in a variably saturated groundwater system with strong river-aquifer interactions and found that it is important to represent at least the basic zones of heterogeneity. Kurtz et al. (2014) also implemented a data assimilation scheme for joint assimilation of groundwater temperature data and piezometric head data in the Upper Limmat Valley near Zurich (Switzerland) and tested cases with heterogeneous riverbeds and found that both model states and parameters can be better predicted by updating aquifer and riverbed hydraulic parameters. None of those earlier studies accounted for more complicated heterogeneous patterns of K_{rb}, such as multimodal non-Gaussian distributions, which are expected to occur in practice (Cheng et al., 2011). To fill this gap, we investigated in a previous study the impact of different simple and complex patterns of heterogeneous K_{rb} on the characterization of hydraulic heads, riverbed properties, and river-aquifer exchange fluxes (Tang et al., 2015).

4.1.2 Limitation of previous studies

In order to estimate river-aquifer exchange fluxes correctly, it is important that the selected hydrological model calculates these in a physically based manner. The majority of numerical groundwater models (e.g. MODFLOW (McDonald and Harbaugh, 1988), SPRING (Delta h Ingenieugesellschaft mbH, 2006)) are based on a conductance approach. The disadvantage of conductance approaches is that the non-linear unsaturated flow behavior underneath the river is usually not taken into account, and lateral flow within the riverbed is also neglected (Brunner et al., 2010). Our previous study (Tang et al., 2015) used the model SPRING and was based on this kind of approach. Fully-integrated physically-based surface-subsurface models are increasingly applied to provide a physically more consistent description of surface water-groundwater exchange processes. Examples are provided by VanderKwaak and Sudicky (2000) (InHM), Panday and Huyakorn (2004) (MODHMS), Kollet and Maxwell (2006) (Parflow), Camporese et al. (2010) (CATHY), Therrien et al. (2010) and Brunner and Simmons (2012) (HGS). Such integrated hydrological models can better capture the dynamics of surface water-groundwater interactions but require additional input parameters (e.g. parameters for simulation of surface water flow) compared to more simplified approaches, which simulate only the groundwater flow component.

Only few data assimilation studies were carried out with fully coupled surface water-groundwater models. Examples are provided by Camporese et al. (2009) and Rasmussen et al. (2015), amongst others. Both groundwater head and stream discharge observations were used for assimilation in those works and Rasmussen et al. (2015) also considered the estimation of spatially homogeneous leakage coefficients. Kurtz et al. (2016) developed a data assimilation scheme for the terrestrial system modelling platform TerrSysMP (Shrestha et al., 2014). It allows joint updating of both model states and parameters of the CLM (land surface) and Parflow (subsurface including overland flow) modules. However, until now no studies updated heterogeneous riverbeds in fully coupled surface water-groundwater models.

4.1.3 Contribution of this paper

Tang et al. (2015) used a conductance based model to conclude that complex non-multi-Gaussian riverbed patterns did not result in significantly different river-aquifer exchange fluxes compared to multi-Gaussian riverbed patterns (with the same overall mean K_{rb} and variance as the non-multi-Gaussian riverbed). In this work, we combine now the use of fully coupled surface water-groundwater models with the evaluation of the impact of complex heterogeneous patterns of K_{rb} on river-aquifer exchange fluxes. The heterogeneous patterns of K_{rb} are estimated by the ensemble Kalman filter (Evensen, 1994), a data assimilation approach. This will provide further insights to what extent representing the geological complexity of a riverbed is required for the characterization of states and fluxes under both fully saturated and variably saturated conditions underneath the riverbed in the river-aquifer system.

The experiments consist of two different cases. In one case, saturated conditions for the riverbed and the underlying aquifer prevail and full hydraulic connection between the river and aquifer is thus maintained. In the second case, variably saturated conditions occur within and underneath the riverbed and as outlined above, due to the unsaturated conditions the estimation problem becomes non-linear and it is expected that it will be more challenging to infer river-aquifer exchanges fluxes and model states and parameters. The role of different K_{rb} patterns for this configuration is therefore much unclearer and results are expected to show more variation. The analysis was made for four different geostatistical models of K_{rb} including non-multi-Gaussian fields with and without channelized structures, multi-Gaussian fields and homogeneous fields.

4.2 Theory and methods

4.2.1 Coupled surface-subsurface flow simulations with HydroGeoSphere

If river-aquifer interactions are modelled by the conductance approach, river stages are needed as a predefined input and are conceptualized as fixed head dependent boundary conditions in the river-aquifer model (Knapton, 2009). The river-aquifer exchange fluxes are then calculated as a linear function of the gradient of hydraulic heads between groundwater and surface water (Brunner et al., 2010). However, if the groundwater level is lowered, an unsaturated zone can develop and once the river and the aquifer are hydraulically disconnected exchange fluxes are no longer changing linearly as a function of head differences between the surface water body and the aquifer (Brunner et al., 2009). Compared to conductance-based models, a fully-coupled surface water-subsurface flow model allows a dynamic, two-way coupling when accounting for the river-aquifer interaction. Exchange fluxes are then calculated in a fully coupled manner. HydroGeoSphere (HGS) (Therrien et al., 2010; Brunner and Simmons, 2012; Aquanty Inc, 2016) is such a fully coupled, physically-based model for groundwater and surface water flow simulations and used for the work presented in this paper.

In HGS, variably saturated subsurface flow is simulated with the three-dimensional Richards equation and surface water flow is solved using a one- or two-dimensional approximation of the Saint Venant equations. The surface flow and subsurface flow equations are solved simultaneously for each time step at the interface nodes in a joint equation system, and complete water balance and solute budgets are calculated. HGS has successfully been used to simulate the interactions between groundwater and surface water (Goderniaux et al., 2009; Partington et al., 2013) , the interactions between surface water-, groundwater and vegetation (e.g. Banks et al., 2011; Schilling et al., 2014), micro-topographic wetland runoff (Frei et al., 2010) and also large scale solute transport (Blessent et al., 2011).

In HGS, two approaches can be used for flow coupling between subsurface and surface domains. In the first approach, the top layer of nodes represents both surface and subsurface domains, assuming that the same heads are assigned for the surface and the uppermost subsurface nodes. The surface and subsurface flow equations are solved simultaneously for those surface-subsurface interface nodes. The second approach uses Darcy’s law for flow relations between the surface nodes and the first layer of subsurface nodes, with an assumption that they are separated by a thin layer where the leakage occurs. A first-order exchange coefficient is introduced and exchange fluxes are calculated linearly dependent on the pressure difference between the surface and subsurface domain. A detailed comparison between the two approaches is given in Liggett et al. (2012). In both of these approaches subsurface and surface flow equations are fully coupled. In this work, the second approach (called the dual node approach) is selected to represent the relation between surface domain and subsurface domain.

While the dual node approach follows a similar conceptualization as the leakage principle in conductance based models, the representation of exchange fluxes between river and aquifer in HGS in the unsaturated zone is more realistic because it is limited by the relative permeability. Moreover, if the coupling length l_{exch} of the dual node approach is set to a very small value, the effects of the dual node approach on simulation results are minimal and results are comparable to the common node approach, while providing a higher numerical stability compared to the common node approach (Liggett et al., 2012).

4.2.2 Data assimilation with ensemble Kalman filter

The data assimilation procedure of this study can be described by the following steps:

(1) For each realization of K_{rb} generated according to section 4.3.3, the model states at the current time step are forecasted using the solution from the previous time step as initial condition with HGS:

$$\mathbf{x}_{t,i} = M\left(\mathbf{x}_{t-1,i}\right) \tag{4.1}$$

where $\mathbf{x}$ is the augmented state vector containing the model states (here the simulated piezometric heads) and parameters (here the log-transformed K_{rb}):

$$\mathbf{x}_i = \begin{pmatrix} \mathbf{h} \\ \mathbf{Y} \end{pmatrix}_i \tag{4.2}$$

where **h** is the simulated heads and $\mathbf{Y} = \log_{10}\left(\mathbf{K_{rb}}\right)$. *M* the numerical flow model (HGS in this study), *t* is the time step counter and *i* the realization number.

(2) After every time step of the simulation, EnKF is applied to update both model states and parameters according to:

$$\mathbf{x}_{t,i}^{a} = \mathbf{x}_{t,i} + \mu\mathbf{G}\left(\mathbf{y}_{t,i} - \mathbf{H}\mathbf{x}_{t,i}\right) \tag{4.3}$$

where $\mathbf{x}_{t,i}^{a}$ is the updated augmented state vector, $\mathbf{x}_{t,i}$ is the forecasted state vector, μ is a damping factor varying between 0 and 1, **H** is the measurement operator matrix mapping the simulated states to the observation locations, and **G** is the Kalman gain which weights the relative importance of the model forecast and the observations.

According to Burgers et al. (1998), measurements need to be perturbed to account for the measurement uncertainty in order to guarantee the correct estimation of the posterior variance of the updated ensemble with EnKF, which can be written as :

$$\mathbf{y}_{t,i} = \mathbf{y}_{t} + \boldsymbol{\varepsilon}_{t,i} \tag{4.4}$$

where $\mathbf{y}_{t,i}$ is the vector of perturbed measurements, $\mathbf{y}_{t}$ is a vector with the original measurements at time step *t*, and $\boldsymbol{\varepsilon}_{t,i}$ is a vector with normally distributed measurement errors with zero mean and specified variance accounting for the measurement uncertainty.

The Kalman gain **G**, which determines the weight assigned to the model simulation on the one hand and the measurements on the other hand, is given by

$$\mathbf{G} = \mathbf{C}\mathbf{H}^{\mathrm{T}}\left(\mathbf{H}\mathbf{C}\mathbf{H}^{\mathrm{T}} + \mathbf{R}\right)^{-1} \tag{4.5}$$

where **R** is the measurement error covariance matrix, and **C** is the covariance matrix with covariances between states, between states and parameters, and between parameters according to:

$$\mathbf{C} = \begin{pmatrix} \mathbf{C}_{\mathbf{hh}} & \mathbf{C}_{\mathbf{hY}} \\ \mathbf{C}_{\mathbf{Yh}} & \mathbf{C}_{\mathbf{YY}} \end{pmatrix} \tag{4.6}$$

Once the updating step (step (2)) is finished, the updated states and parameters are used as inputs for the simulation of the next time step, the algorithm returns to step (1), and the procedure is repeated until the end of the simulation. The coupling of EnKF and HGS is described in Kurtz et al. (2017).

4.3 Synthetic flow modeling experiments

4.3.1 Overview

To investigate the importance of K_{rb} patterns for river-aquifer interactions, simulations with complex, non-multi-Gaussian reference K_{rb} fields are compared to simulations with multiple geostatistical models for K_{rb}. The model simulation period is two years in total. Data assimilation experiments were performed in the first year, and the second year served as a verification period. Before the data assimilation runs, a spin up run of one year with the same transient model forcings as for the data assimilation experiments was carried out. The modeling experiments consist of the following steps:

1) Ten reference K_{rb} fields are generated, and the corresponding river-aquifer interactions are simulated during two years with HGS. The first year is the assimilation period and the second year the verification period. The K_{rb} fields, the simulated hydraulic heads and the simulated (net) exchange fluxes serve as the 'truth'.

2) A large number of initial K_{rb} fields are generated for each of the different geostatistical models, using geostatistical simulation algorithms.

3) Data assimilation experiments were carried out for a one year period for each of the four geostatistical models, using the initial K_{rb} fields as input. The updated K_{rb} fields and the simulated heads and fluxes are subsequently evaluated against the reference fields.

4) The estimated fields of K_{rb} parameters were used as input for verification experiments for the following one year period with different hydrological conditions and without data assimilation, for

each of the four geostatistical models. The simulated heads and fluxes are then evaluated against the reference fields.

4.3.2 Three dimensional river-aquifer model setup

4.3.2.1 Model domain

In this study, a three-dimensional synthetic river-aquifer model is simulated. The subsurface domain (i.e., the aquifer) has a spatial extent of 500 m × 250 m × 10.5 m, while the surface domain (i.e., the river) is defined as a channel centered on top of the aquifer, with spatial dimensions of 500 m × 80 m × 0.5 m. Both the surface and subsurface domain are discretized by 3D-blocks of size 10 m x 10 m x 0.1 m resulting in a grid of 131,250 cells in total. The numerical model is outlined in Figure 4.1. The model is inclined along the x-direction with a slope of 0.01 m/m. The riverbed (including the riverbanks) is conceptualized by a rectangular cross-section with eight rows of elements. This constrains the surface water to flow within the channel.

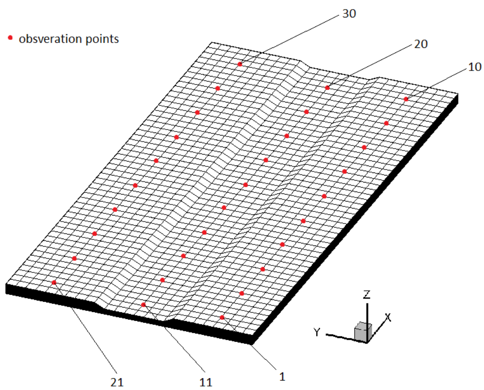

Figure 4.1: The synthetic 3-D river-aquifer model domain with locations of synthetic hydraulic head observations (red solid points).

4.3.2.2 Twin cases set up: saturated vs. variably saturated

The importance of K_{rb} patterns for river-aquifer interactions is evaluated for two cases using the same model domain but different states of connection between the river and the aquifer. The 'saturated' case indicates that in the simulated model, riverbed and the subsurface below the riverbed are fully saturated and the river and the aquifer are hydraulically connected. Alternatively, the 'variably saturated' case refers to the model which is partially unsaturated under the river and the system is either fully hydraulically disconnected (i.e., an unsaturated zone is present under the entire streambed) or in a transitional state with partly saturated and unsaturated conditions. This variably saturated case is obtained by adjusting the ratio of K in the aquifer and riverbed according to Brunner et al. (2009), where, in order for an unsaturated zone to occur, the following equation needs to hold true:

$$\frac{K_{rb}}{K_{aq}} \leq \frac{h_{rb}}{h_{rb} + d} \tag{4.7}$$

where h_c is the thickness of the riverbed and d is the ponded water depth.

As an example, the saturation map of the variably saturated case for the reference K_{rb} field no. 6 is displayed in Figure 4.2.

4.3.2.3 Model parameterization

The aquifer and riverbed are conceptualized as gravel-sand material; van Genuchten parameters for aquifer and riverbed are α=3.48 m^{-1} and n=1.75, according to Li et al. (2008). Specific storage is set to 10^{-4} m^{-1} and porosity is 0.25. The coupling length is set to a very small value of 10^{-3} m for calculating the exchange fluxes between the surface and the subsurface domain. Like in the study of Tang et al. (2015), the aquifer is homogeneous with a hydraulic conductivity value of 10^{-3} m/s and only riverbeds are heterogeneous. For the saturated flow case the geometrical mean value of K_{rb} is equal to -6 $\log_{10}$ m/s. Under this condition, the aquifer underneath the riverbed is saturated and the relationship between exchange fluxes and head difference from surface water to groundwater is approximately linear. For the variably saturated case, the geometrical mean value of K_{rb} was lowered by two magnitudes.

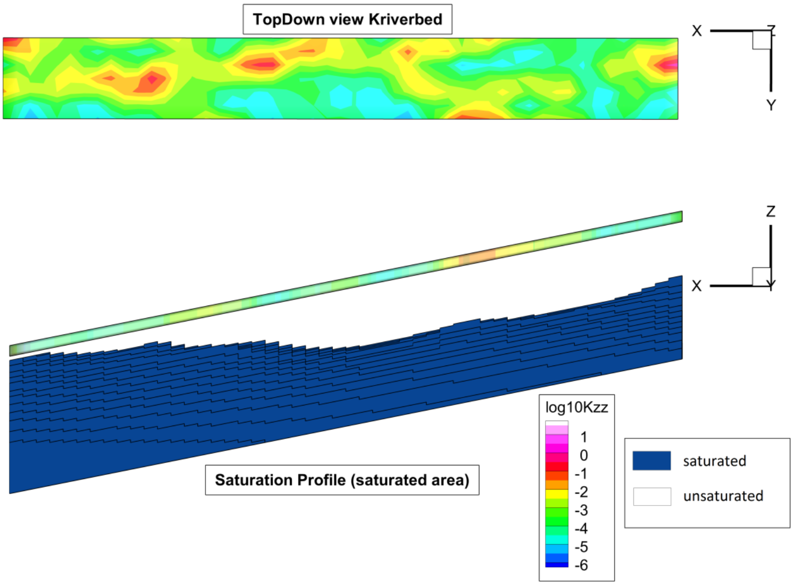

Figure 4.2: 2-D view of the $\log_{10}(K_{rb})$ map (top row) and connectivity conditions under the river (bottom row) for the reference K_{rb} field No. 6 of the variably saturated case. Displayed is the initial condition of the assimilation period. The $\log_{10}(K_{rb})$ field is colored according to the magnitude of K_{rb}, and the saturated regions underneath the river are shown in blue.

4.3.2.4 Boundary conditions

Surface water flow is generated through transient fixed head boundary conditions defined for the upstream nodes along the river cross-section (x = 0, y = 80 to 160). A critical depth boundary condition is defined at the downstream part (x = 500, y= 80 to 160). In the subsurface domain, transient prescribed head boundary conditions are defined for the two planes y = 0 and y = 250. No flow boundaries are applied to the aquifer nodes along x = 0 and x = 500. Figure 4.3 summarizes the boundary conditions used for the saturated and the variably saturated cases over the complete simulation period. For each of the two cases, the same transient boundary conditions are used for the spin-up runs and the assimilation experiments. For the saturated case, in the assimilation period the transient fixed heads defining the inflow at the river inlet are taken from Kurtz et al. (2014). Transient prescribed heads for the subsurface domain are set between z = 404.5 m and 405.5 m. In the verification period, lower prescribed head boundary values are used for the subsurface domain, but otherwise the setup is similar to the assimilation period. For the variably saturated case, in the assimilation period, river stages are the same as for the saturated case and prescribed heads in the subsurface domain are 5 m lower than for the saturated case. The verification period for the variably saturated case has the same subsurface prescribed heads as the assimilation period but river stages have much larger fluctuations than in the assimilation period.

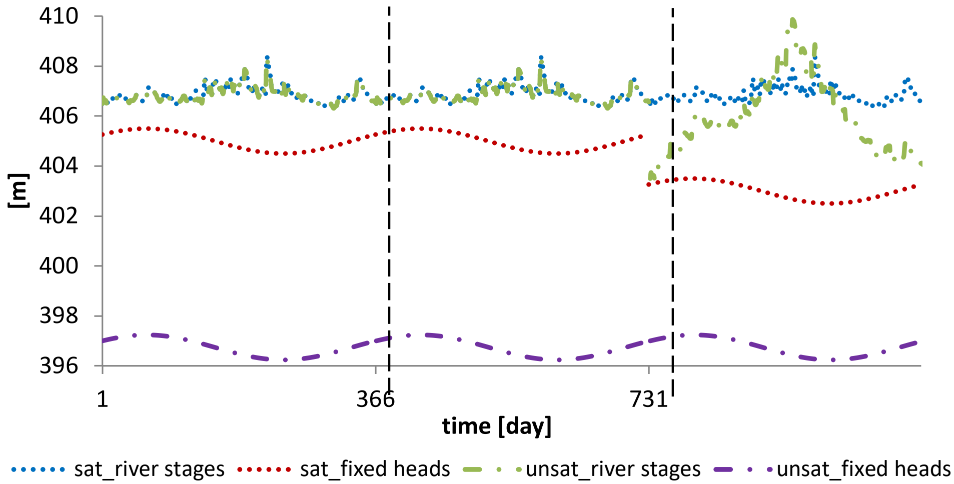

Figure 4.3: Fixed head boundary conditions for the subsurface domain for the saturated case (red round dots; abbr. sat_fixed heads) and for the variably saturated case (purple long dash dots; abbr. unsat_fixed heads). Transient river stages for the upstream river nodes are shown for the saturated case with blue round dots (abbr. sat_river stages), and for the variably saturated case with green long dash dots (abbr. unsat_river stages).

4.3.3 Hydraulic conductivity fields

4.3.3.1 Reference riverbed hydraulic conductivity

A non-multi-Gaussian K_{rb} field with channelized structures was used as the reference K_{rb}. For this purpose, a training image reflecting channelized structures is generated using the software SGeMS (Remy et al., 2009), see Figure 4.4 (a). The training image consists of two facies: channels occupy a proportion of 0.4 of the whole image, and the rest symbolizes background material. The two facies represent two materials with high and low permeability (e.g. sand and clay). In order to avoid random effects in the results of statistical analysis, which might be induced if only one particular reference field is used, ten different K reference fields are generated. The ten realizations of facies distributions are generated using the direct sampling method (Mariethoz et al., 2010). Within each facies, $\log_{10}(K_{rb})$ values are generated independently by sequential multi-Gaussian simulation using GCOSIM3D (Gómez-Hernández and Journel, 1993). The corresponding geostatistical parameters are provided in Table 4.1.

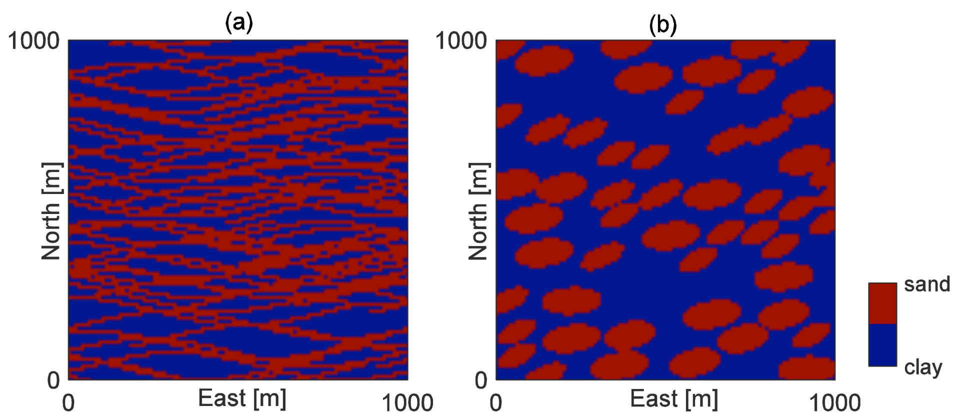

Figure 4.4: Two training images for the generation of stochastic realizations for non-multi-Gaussian K fields: (a) with channelized structures and (b) with elliptic structures. The left one is also used for generating ten reference K_{rb} fields.

Table 4.1: Geostatistical parameters for the two facies in the non-multi-Gaussian distributed K fields and geostatistical parameters for the multi-Gaussian fields.

Facies	Variogram type	Mean ($\log_{10}$(m/s))	Range (m)	Sill ($\log_{10}$(m/s))
channel/ellipse	Spherical	-4.3	100	0.5
background	Spherical	-7.3	100	0.5
multi-Gaussian	Spherical	-6.1	100	2.7

4.3.3.2 Initial riverbed hydraulic conductivity used for the data assimilation experiments

As outlined in section 4.3.1, four different geostatistical models were used to generate initial K_{rb} fields for the assimilation experiments: (1) non-multi-Gaussian distributed fields with channelized structures, (2) non-multi-Gaussian fields with elliptic structures, (3) multi-Gaussian fields and (4) homogeneous fields. As for the reference K_{rb} fields, non-multi-Gaussian fields are simulated with the direct sampling method. The multi-Gaussian fields are generated by sequential multi-Gaussian simulation.

The two types of non-multi-Gaussian heterogeneous K_{rb} fields are generated as follows:

1) Two training images (for each of the two types of non-multi-Gaussian fields independently, see Figure 4.4) are generated by SGeMS. The training image for the non-multi-Gaussian model with channelized structures is the same as used for the generation of the ten reference fields. Each of the two training images is composed of two different facies with a proportion of 0.4 for channels/ellipses.

2) In total 200 stochastic realizations of the spatial distribution of facies are generated with help of the direct sampling method, using the training images from step 1.

3) Each facies is independently populated with multi-Gaussian distributed $\log_{10}(K)$ values, using the sequential Gaussian simulation technique. The geostatistical parameters used for defining the variograms for each facies are given in Table 4.1.

For the multi-Gaussian heterogeneous K_{rb} fields, a similar mean and variance of riverbed $\log_{10}(K_{rb})$ are used as for the non-multi-Gaussian random fields by sequential multi-Gaussian simulation. The mean and variance for the multi-Gaussian random fields is calculated from the mean and variance of each of the two facies of the non-multi-Gaussian fields, using the geometric mean. A more detailed description of the generation technique of non-multi-Gaussian and multi-Gaussian riverbed fields is given in Tang et al. (2015).

The stochastic realizations of $\log_{10}(K_{rb})$ values for the equivalent homogeneous case are generated by taking the geometric mean of each of the 200 random fields for the non-multi-Gaussian case with channelized structures. Figure 4.5 shows examples of stochastic realizations of K_{rb} for each of the four geostatistical models. Figure 4.6 displays the histograms of K_{rb}, calculated over all 200 stochastic realizations for each of the three geostatistical models with heterogeneous *K*. The figure illustrates that the histogram is bimodal for the two non-multi-Gaussian cases, but Gaussian for the multi-Gaussian case.

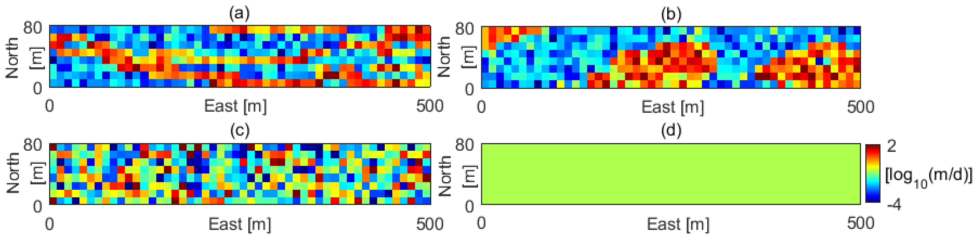

Figure 4.5: Examples of stochastic realizations of initial K_{rb} fields: (a) non-multi-Gaussian field with channelized structures; (b) non-multi-Gaussian field with elliptic structures; (c) multi-Gaussian field; (d) homogeneous field.

4.3.4 Modelling strategy

4.3.4.1 Spin-up run

The spin-up process includes two parts: first a quasi-steady state simulation that runs for 10,000 days with constant boundary conditions corresponding to the forcings of the first time step of the assimilation period, followed by an additional one year exit spin-up run with the 200 stochastic realizations of K_{rb} fields, using the same transient boundary conditions as for the assimilation experiments. The one year quasi-steady state simulation departs from dry initial conditions. The following one year exit spin-up runs are made for each of the four geostatistical models and two cases (fully saturated and variably saturated) and result in different initial heads for each of the different realizations of K_{rb}. These differing initial conditions are necessary to reflect the different interactions resulting from the different *K* fields, and allow generating an adequate ensemble spread. In summary, the spin-up runs provide the initial hydraulic heads needed for the data assimilation experiments.

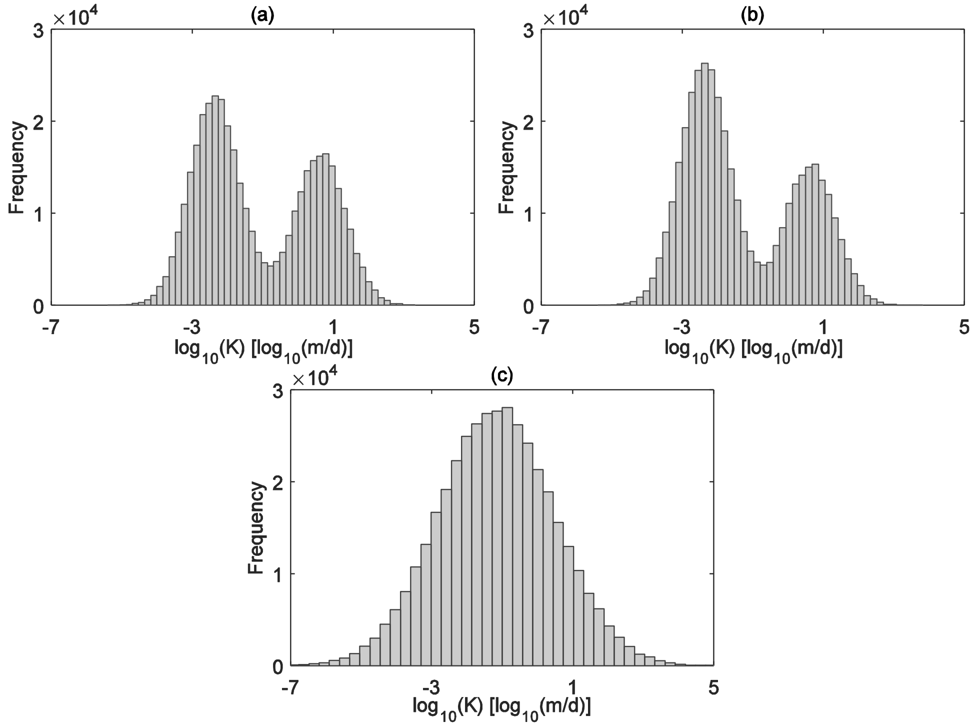

Figure 4.6: Histograms of K_{rb} for three geostatistical models, calculated over all the 200 realizations: (a) non-multi-Gaussian fields with channelized structures; (b) non-multi-Gaussian fields with elliptic structures; (c) multi-Gaussian fields.

4.3.4.2 Observations obtained from the reference runs

For the data assimilation experiments, 30 piezometric head observations are taken from the sixth layer (saturated case) or bottom layer (variably saturated case) and used as virtual observations. 10 of the 30 measurements are beneath the riverbed and the other 20 in the aquifer north and south of the river. These virtual observations are taken from each of the ten reference runs. A measurement error of 5 cm is imposed.

4.3.4.3 Data assimilation experiments

Data assimilation experiments are performed for each of the four geostatistical models, using in total 200 stochastic realizations per model. The damping factor for the parameter update is set to 0.1. In the data assimilation experiments, hydraulic heads and K_{rb} are updated daily by assimilating the 30 virtual piezometric head data. Open loop simulations, without data assimilation, are also performed for comparison purposes. Table 4.2 summarizes the assimilation scenarios.

Table 4.2: Simulation scenarios for data assimilation experiments. The symbol '√' represents the update of h/$\log_{10}(K)$ is done, while '×' represents not.

Riverbed patterns	Scenarios	Update h	Update $\log_{10}(K_{rb})$
non-multi-Gaussian (channel)	channel_open loop	×	×
	channel_hK	√	√
non-multi-Gaussian (elliptic)	ellip_open loop	×	×
	ellip_hK	√	√
multi-Gaussian	multi_open loop	×	×
	multi_hK	√	√
homogeneous	homo_open loop	×	×
	homo_hK	√	√

4.3.4.4 Verification experiments

The estimated hydraulic head and K_{rb} fields at the end of the assimilation period are used as input fields for the verification experiments. This is done for both the data assimilation runs with update of heads and $\log_{10}(K_{rb})$ and for the open loop simulations, and for each of the four geostatistical models of K_{rb}. In the one year verification period the parameter fields are not modified and no data assimilation is done.

4.3.4.5 Performance measures

The Root Mean Square Error (RMSE) is used to evaluate the characterization of model states, riverbed hydraulic conductivities and river-aquifer exchange fluxes. RMSE is evaluated separately both for the data assimilation as well as the verification periods. The RMSE for hydraulic heads is calculated for all model nodes (including both surface nodes and subsurface nodes) over all simulation time steps:

$$\mathrm{RMSE}(h) = \sqrt{\frac{1}{n_t n_{nodes\_all}} \sum_{i=1}^{n_t} \sum_{j=1}^{n_{nodes\_all}} \left(\overline{h}_{i,j}^{f} - h_{i,j}^{ref}\right)^2} \tag{4.8}$$

where n_t is the number of simulation time steps, n_{nodes_all} the total number of model nodes, the overbar indicates ensemble average, *h* hydraulic heads, the superscript *f* indicates model simulations and the superscript *ref* the reference field.

The RMSE for the updated K_{rb} is calculated over all riverbed elements at the final simulation time step:

$$\mathrm{RMSE}(K_{rb}) = \sqrt{\frac{1}{n_{river\_ele}} \sum_{j=1}^{n_{river\_ele}} \left(\overline{\log_{10}(K_{rb\,j}^{\,a})} - \log_{10}(K_{rb\,j}^{\,ref})\right)^2} \tag{4.9}$$

where $\overline{\log_{10}(K_{rb\,j,n_t}^{\,a})}$ is the updated ensemble mean value of $\log_{10}(K_{rb})$ for the j^{th} element of the riverbed at the final time step of the assimilation period and n_{river_ele} is the number of river elements.

The RMSE for exchange fluxes is calculated over all the river nodes for all time steps:

$$\mathrm{RMSE}(Q) = \sqrt{\frac{1}{n_t} \sum_{i=1}^{n_t} \left(\overline{Q_i} - Q_i^{ref}\right)^2} \tag{4.10}$$

$$Q_i = \sum_{k=1}^{n_{nodes\_river}} Q_k \tag{4.11}$$

where n_{nodes_river} is the number of the river nodes.

For an individual time step, the RMSE for hydraulic heads and net exchange fluxes is calculated according to:

$$\mathrm{RMSE}(h,t) = \sqrt{\frac{1}{n_{nodes\_all}} \sum_{j=1}^{n_{nodes\_all}} \left(\overline{h}_{i,j,t}^{f} - h_{i,j,t}^{ref} \right)^2} \tag{4.12}$$

$$\mathrm{RMSE}(Q,t) = \left| \overline{Q_{i,t}} - Q_{i,t}^{ref} \right| \tag{4.13}$$

4.4 Results

4.4.1 Saturated case

4.4.1.1 Assimilation period

Compared to the open loop simulations, data assimilation improved the characterization of model states and parameters, especially when the models included heterogeneous riverbed structures. Since the simulation of heads and exchange fluxes is influenced by model dynamics, average RMSE (*h*,*t*) and RMSE (*Q*,*t*) calculated over ten references are shown in Figure 4.7 (a, c). Displayed are results for the four geostatistical models. From Figure 4.7 (a) we see that only at the beginning of the data assimilation period RMSE (*h,t*) is clearly higher for ellip_hK and homo_hK compared to the other two models, but later RMSE shows a substantial reduction. After a simulation period of 50 days, channel_hK and ellip_hK have very similar errors, with lower RMSE than for multi_hK and homo_hK. Within the assimilation period, the maximum difference in RMSE (*h*,*t*) among the four geostatistical models is 7 cm.

The boxplot of RMSE (K_{rb}) calculated over ten reference models is displayed in Figure 4.8 (a). The final updated ensemble mean riverbed $\log_{10}(K_{rb})$ field for the four different geostatistical models is shown for one of the reference fields in Figure 4.9 (a - d). The non-multi-Gaussian models with channelized structures resemble the true K_{rb} field best, while, non-surprisingly, the homogeneous model deviates most. Nevertheless, the differences between the two non-multi-Gaussian models are small.

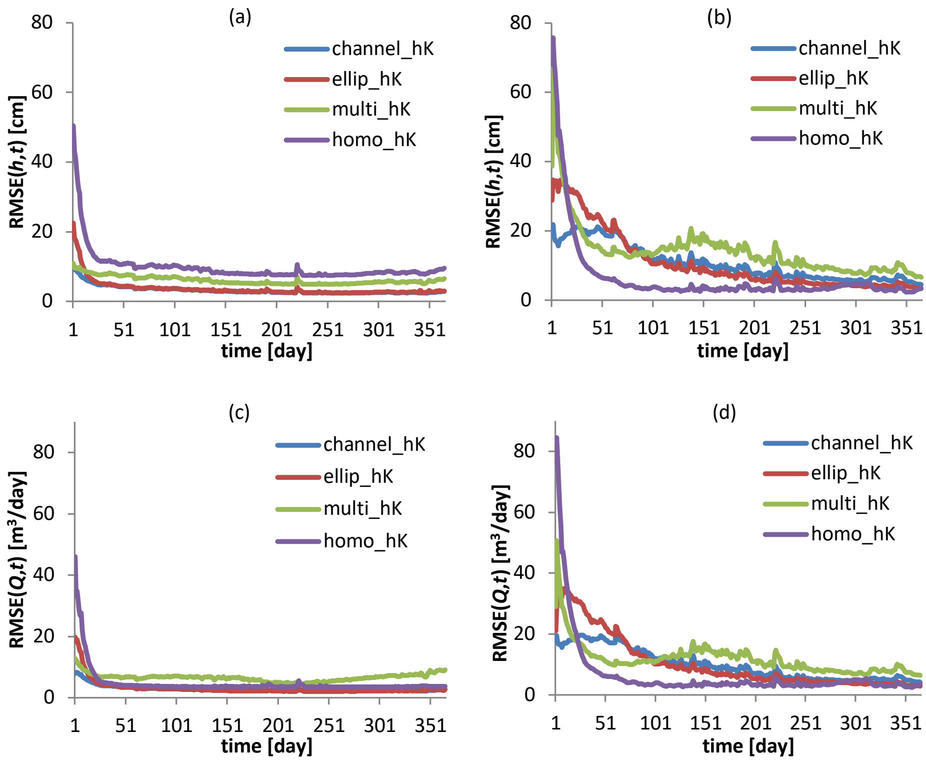

Figure 4.7: Average RMSE (*h*,*t*) for (a) the saturated case and (b) the variably saturated case; average RMSE (*Q*,*t*) for (c) the saturated case and (d) the variably saturated case. These RMSEs were calculated over ten references and for the four geostatistical models, the saturated case and assimilation period.

The large differences in the characterization of K_{rb} patterns between the non-multi-Gaussian models and the other models also manifest themselves in terms of exchange fluxes: Compared to the non-multi-Gaussian model with channelized structures, the average RMSE (*Q*) over ten references is 61.3% higher for the non-multi-Gaussian model with elliptic structures and 32.6% higher for the multi-Gaussian model. The homogeneous model performs considerably worse than the others, which resulted in the highest errors: the RMSE (*Q*) is 271.3% higher than for the non-multi-Gaussian model with channelized structures.

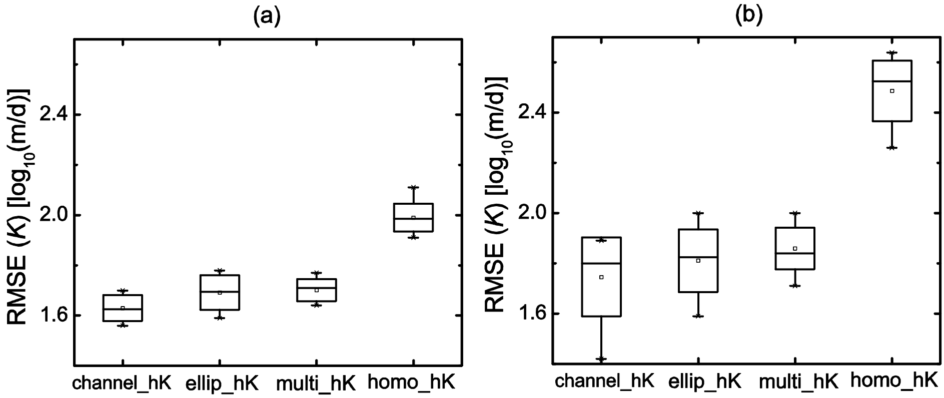

Figure 4.8: Boxplots of RMSE (K_{rb}) calculated over ten references (K_{rb} is evaluated at the end of the assimilation period) for the four geostatistical models for (a) the saturated case and (b) the variably saturated case.

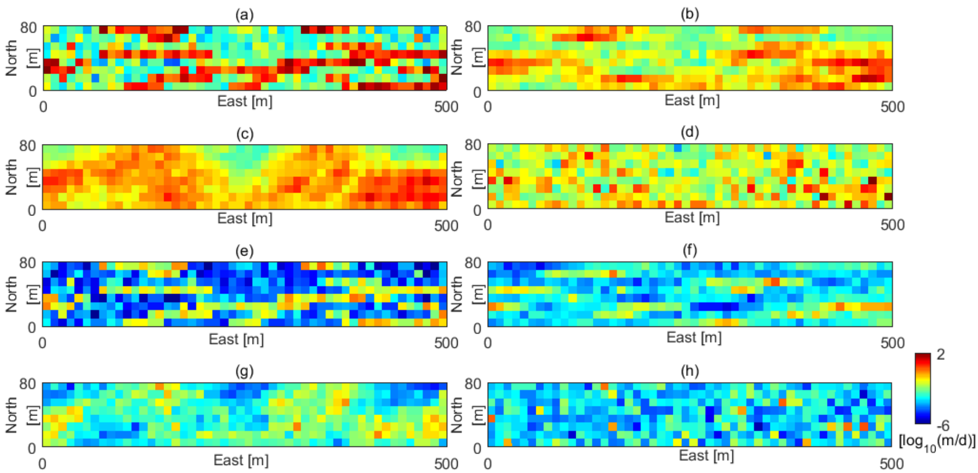

Figure 4.9: Ensemble mean riverbed $\log_{10}$ (K_{rb}) fields at the end of the assimilation period for the saturated case (a - d) and the variably saturated case (e - h). Shown are the reference (a, e) and different geostatistical models: (b, f) channel_hK; (c, g) ellip_hK; (d, h) multi_hK.

4.4.1.2 Verification experiment

The results of the verification experiments are evaluated in terms of RMSE (h) and RMSE (Q), shown in Table 4.3. Values for the ten individual reference fields and their average are provided. The non-multi-Gaussian model with channelized structures results in the smallest RMSE (h) in four out of ten cases, the non-multi-Gaussian model with elliptic structures in five out of ten, and the multi-Gaussian model in one. The non-multi-Gaussian models thus clearly outperform the other models. The average RMSE (h) is also lower for the heterogeneous models (2.9 - 3.2 cm) than for the homogeneous model (6 cm). Among the three heterogeneous models, the differences are minor. Slightly smaller RMSE (h) are found for two non-multi-Gaussian models (average RMSE (h) 2.9 - 3.0 cm) than the multi-Gaussian model (average RMSE (h) 3.2 cm).

Table 4.3: RMSE (h) (cm) and RMSE (Q) (m^3/d) for the verification period for the saturated case. Shown are results for the ten reference fields and the four geostatistical models: the non-multi-Gaussian model with channelized structures (channel), the non-multi-Gaussian model with elliptic structures (ellip), the multi-Gaussian model (multi) and the homogeneous model (homo).

	channel		ellip		multi		homo	
	RMSE (h)	RMSE (Q)	RMSE (h)	RMSE (Q)	RMSE (h)	RMSE (Q)	RMSE (h)	RMSE (Q)
Ref_1	3.16	2.07	2.72	2.00	3.24	1.86	4.75	3.19
Ref_2	2.78	1.60	3.06	1.97	3.05	2.22	3.97	0.95
Ref_3	2.65	2.55	2.91	3.11	2.97	3.19	7.97	5.87
Ref_4	3.38	2.55	3.10	2.63	4.71	1.69	7.88	4.53
Ref_5	2.94	3.61	2.68	3.86	2.65	3.77	5.02	3.57
Ref_6	3.20	0.95	3.26	0.69	3.32	0.55	7.80	2.72
Ref_7	2.79	0.32	2.73	0.25	2.73	0.39	4.65	3.04
Ref_8	3.46	1.04	3.19	1.34	3.37	0.65	4.60	0.60
Ref_9	2.67	1.69	2.96	1.53	3.43	0.70	8.83	5.33
Ref_10	2.93	0.60	2.60	0.47	2.87	8.45	4.47	1.88
Average	**3.00**	**1.70**	**2.92**	**1.79**	**3.23**	**2.35**	**6.00**	**3.17**

Also in terms of characterization of river-aquifer exchange fluxes the heterogeneous models outperform the homogeneous models. The average RMSE (Q) is 3.17 m^3/d for the homogeneous models compared to 1.7 – 2.3 m^3/d for the heterogeneous models; the RMSE (Q) is about 86.5% larger for the homogeneous model than for the non-multi-Gaussian model with channelized structures. Among the three heterogeneous models, like for the RMSE (h), similar performance is observed for the two non-multi-Gaussian models (average RMSE (Q) between 1.7 - 1.8 m^3/d), while the RMSE (Q) for the multi-Gaussian model is 41.2% larger than the non-multi-Gaussian model with channelized structures.

4.4.2 Variably saturated case

4.4.2.1 Assimilation period

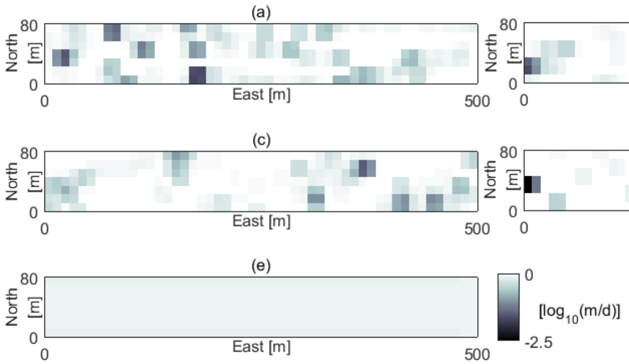

Figure 4.10: Maps with exchange fluxes for the variably saturated case at the end of the assimilation period. Shown are the reference (a) and results for the different geostatistical models: (b) channel_hK; (c) ellip_hK; (d) multi_hK; (e) homo_hK. Negative fluxes indicate infiltration from the river into the aquifer.

The average RMSE (*h*,*t*) and RMSE (*Q*,*t*) calculated over all ten references are shown in Figure 4.7 (b, d). The errors are larger in the beginning of the assimilation period and vary strongly between the different models due to the different initial conditions. However, towards the end of the assimilation period, the RMSE (*h*,*t*) and RMSE (*Q*,*t*) for the four models become smaller and more similar. Within the first 100 days, data assimilation allows reducing the RMSE (*h*,*t*) by more than 80% for all four geostatistical models of K_{rb}, especially for the homogeneous model which produces the smallest RMSE (*h*) at the end of the assimilation period. Figure 4.8 (b) shows the boxplot of RMSE (K_{rb}) for the updated K_{rb} fields. In six out of ten cases the non-multi-Gaussian case with channelized structures shows the best characterization of K_{rb} in terms of RMSE (K_{rb}). Figure 4.9 (e - h) indicates that only the non-multi-Gaussian model with channelized structures could capture some of the channels present in the reference K_{rb} fields; the multi-Gaussian model deviates strongly from that reference pattern. Overall, heterogeneous fields allow better capturing the spatial distribution of exchange fluxes, although their exact spatial position is often not correct. This is illustrated in Figure 4.10. The relatively good performance of the homogeneous model in reproducing exchange fluxes compared to the heterogeneous models is related to very large flux errors for a few grid cells in the heterogeneous cases, which have an important influence on the performance statistics, even though we focus on the net fluxes.

4.4.2.2 Verification experiment

The verification experiment with variably saturated conditions was carried out with stronger variations in river stage than during the assimilation phase. The RMSE (*h*) and RMSE (*Q*) scores of the verification period are provided in Table 4.4. As in the assimilation period, the RMSE (*h*) for the three heterogeneous models are smaller than for the homogeneous model. In total, the non-multi-Gaussian model with channelized structures results in the smallest RMSE (*h*) for three out of ten reference cases, the non-multi-Gaussian model with elliptic structures six out of ten, and the multi-Gaussian model in one out of ten cases. The two non-multi-Gaussian models result in the lowest mean RMSE (*h*). The average RMSE (*h*) for the non-multi-Gaussian model with channelized structures is 18.2 cm and for the non-multi-Gaussian model with elliptic structures it is 17.1 cm. The multi-Gaussian model shows only a slightly larger mean RMSE (*h*) (19.6 cm). The homogeneous model leads to a RMSE (*h*) of 40.4 cm, which is 121% higher than for the non-multi-Gaussian model with channelized structures.

The two non-multi-Gaussian models result in the best characterization of net exchange fluxes: in two out of ten cases the non-multi-Gaussian model with channelized structures outperforms all other geostatistical models, and in eight out of ten cases the non-multi-Gaussian model with elliptic structures. The average RMSE (*Q*) of the non-multi-Gaussian model with channelized structures and the non-multi-Gaussian model with elliptic structures are 3.45 m^3/d and 2.49 m^3/d, respectively. Although the homogeneous riverbed has clearly the largest RMSE (*h*) and RMSE (*K*), for the characterization of net exchange fluxes the homogeneous model performs only slightly worse (RMSE (*Q*) = 5.72 m^3/d) compared to the multi-Gaussian model (5.22 m^3/d). The mean RMSE (*Q*) for the homogeneous model is 65.8% larger than for the non-multi-Gaussian model with channelized structures.

Table 4.4: RMSE (*h*) (cm) and RMSE (*Q*) (m^3/d) for the verification period for the variably saturated case. Shown are results for the ten reference fields and the four geostatistical models.

	channel		Ellip		multi		homo	
	RMSE (*h*)	RMSE (*Q*)	RMSE (*h*)	RMSE (*Q*)	RMSE (*h*)	RMSE (*Q*)	RMSE (*h*)	RMSE (*Q*)
Ref_1	14.48	2.45	13.67	0.75	19.54	6.08	42.92	7.16
Ref_2	18.69	3.31	20.87	2.02	21.73	5.40	44.85	2.86
Ref_3	18.12	2.69	19.05	1.52	18.39	3.40	59.61	14.01
Ref_4	18.90	4.28	17.72	2.11	17.62	3.78	37.96	2.49
Ref_5	21.08	4.87	17.64	2.86	21.44	6.92	32.21	4.05
Ref_6	21.53	3.55	18.66	4.46	22.84	6.38	41.22	4.04
Ref_7	14.37	5.26	15.50	6.38	17.14	7.14	41.21	8.36
Ref_8	18.58	2.72	14.28	2.23	19.24	3.76	50.48	4.36
Ref_9	15.83	3.17	15.65	0.79	17.67	4.43	23.55	3.94
Ref_10	20.81	2.19	17.45	1.79	20.28	4.86	29.94	5.91
Average	**18.24**	**3.45**	**17.05**	**2.49**	**19.59**	**5.22**	**40.40**	**5.72**

4.4.3 Correlation analysis

In a final step, linear correlations were calculated between different observations and simulated heads for all riverbed nodes, all references and all simulation scenarios, with the aim to find out whether head observations in the assimilation process provided necessary and sufficient information or not. The percentage of grid cells for which the absolute value of the linear correlation coefficient was >0.5 and >0.2 was calculated, for three different time steps (after 50, 180 and 365 days), all four different geostatistical models and the two different saturation cases. The percentage of grid cells represents an average value calculated over all reference fields. Table 4.5 summarizes these scores. The correlations for the homogeneous model are much stronger than for the heterogeneous models for all the displayed time steps and for both saturated and variably saturated cases. For the homogeneous model 99.6% - 100% of the grid cells show a correlation coefficient >0.2, while for the heterogeneous model this percentage is between 8.5% and 32.5%.

Table 4.5: The percentage of grid cells for which the absolute value of the linear correlation coefficient was >0.5 and >0.2, for three different time steps, the four different geostatistical models and the two different considered cases. The percentage of grid cells is an average value calculated over all reference fields.

model	K_{rb} type	t=50 days		t=180 days		t=365 days	
		>0.5	>0.2	>0.5	>0.2	>0.5	>0.2
Saturated case	channel	2.60%	21.66%	2.44%	21.39%	1.60%	15.04%
	ellip	2.40%	18.43%	2.45%	19.88%	1.42%	10.31%
	multi	1.32%	8.47%	1.15%	10.43%	1.13%	8.16%
	homo	99.01%	99.56%	99.78%	100.00%	98.69%	99.78%
Variably saturated case	channel	0.87%	15.91%	1.06%	16.51%	1.16%	16.43%
	ellip	0.85%	11.53%	1.09%	19.06%	2.77%	32.53%
	multi	0.59%	9.64%	0.81%	9.91%	2.11%	24.90%
	homo	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%

The high correlation between the observations and the simulated hydraulic heads for the homogeneous model likely relates to the fact that the different observations all spatially coincide with the one homogeneous K_{rb}, while for heterogeneous riverbeds observations provide more local information to the surrounding grid elements, and not directly relate to the entire riverbed. This highlights that for homogeneous riverbeds measurement data can be more informative to update the single (and unknown) K_{rb}. As in the homogeneous case the representation of the riverbed is strongly simplified, the measurement data allow the identification of this unknown K_{rb} value.

4.5 Discussion

The results of this study reveal that for all four geostatistical models and both the saturated and variably saturated cases, data assimilation improved the characterization of hydraulic heads. In terms of characterizing K_{rb}, data assimilation provided an improvement only when heterogeneous models of K_{rb} were used. Finally, in most simulation scenarios the estimation of net river-aquifer exchange fluxes was improved through data assimilation. At the end of the assimilation period of the saturated case, the two non-multi-Gaussian models performed clearly better than the other two geostatistical models, while in the variably saturated case the four different geostatistical models showed similar performance.

One important question is whether an erroneous heterogeneous/homogeneous approximation of a heterogeneous, non-multi-Gaussian riverbed with channelized structures allows characterizing the river-aquifer exchange fluxes adequately. Performance is evaluated based on the results in the verification period. Comparison is first made between an erroneous heterogeneous non-multi-Gaussian model (in this work the non-multi-Gaussian model with elliptic structures) and a correct heterogeneous non-multi-Gaussian model (here the non-multi-Gaussian model with channelized structures): Results provided in sections 4.4.1.2 and 4.4.2.2 illustrate that both for the saturated and the variably saturated case differences between the two non-multi-Gaussian models were minor, especially for the variably saturated case, as the erroneous heterogeneous non-multi-Gaussian model outperformed even the correct heterogeneous non-multi-Gaussian model. Next, the performance of the multi-Gaussian model and the (correct) non-multi-Gaussian model with channelized structures is compared: For the saturated case the average RMSE (Q) of the multi-Gaussian model is slightly larger, although five out of ten times it outperforms the correct non-multi-Gaussian model. For the variably saturated case the multi-Gaussian model performs clearly worse than the (correct) non-multi-Gaussian model with channelized structures.

These comparisons highlight that for the saturated, hydraulically connected river-aquifer systems, the error in the characterization of net river-aquifer exchange fluxes does not significantly differ between a non-multi-Gaussian model with channelized structures, a non-multi-Gaussian model with elliptic structures and a multi-Gaussian model. This is a similar conclusion as for the river-aquifer model analyzed by Tang et al. (2015), where the pattern of the heterogeneous K_{rb} had little influence on the characterization of river-aquifer exchange fluxes. However, for variably saturated river-aquifer systems which are not hydraulically connected everywhere, the error in reproducing net river-aquifer exchange fluxes is smaller for the non-multi-Gaussian models than for a multi-Gaussian model. This implies that knowing only the mean and the variance of the K_{rb} contains not sufficient information for estimating the net river-aquifer exchange fluxes, and knowing the histogram of the K_{rb} provides valuable additional information. On the other hand, the small differences between the two non-multi-Gaussian models imply that information on exact patterns and connectivity is not important for improving the estimation of river-aquifer exchange fluxes. It is possible that the estimation of solute fluxes between river and aquifer would be more affected by the K_{rb} pattern.

The homogeneous model results in larger errors than the correct non-multi-Gaussian model in terms of average RMSE (*Q*), although the homogeneous model outperforms the correct model three out of ten times in both the saturated and the variably saturated cases. It is somewhat surprising that for the variably saturated case the characterization of net exchange fluxes with the homogeneous model was not much worse than for the heterogeneous models, given that both hydraulic heads and K_{rb} were significantly better characterized using the heterogeneous models. It was also expected that the heterogeneous models would outperform the homogeneous model stronger for the variably saturated case than for the saturated case (Irvine et al., 2012). Different reasons can be postulated for the relatively good performance of the homogeneous model (albeit with larger RMSE values than for the heterogeneous non-multi-Gaussian models):

1) Although a homogeneous approximation of a heterogeneous riverbed induces errors, in this study the heterogeneous approximation also induced errors, which is related to a lack of information from hydraulic head observations (both in terms of the limited information content of hydraulic heads towards reproducing fluxes, as well as in terms of the limited number of available head observations). This is one limitation of this study. In the study of Irvine et al. (2012) the homogeneous approximation of the heterogeneous riverbed was made with help of the measured net river-aquifer exchange flux over the studied stream reach. This is the best information possible over the studied stream reach. In this work, both for the homogeneous and the heterogeneous approximations of the true heterogeneous riverbed, only a limited amount of hydraulic head data is available.

This is a key difference between the problem considered and the upscaling problem analyzed by Irvine et al. (2012).

2) The absolute correlations between measured hydraulic heads and modelled riverbed hydraulic heads were larger for the homogeneous case than for the heterogeneous cases, illustrating that for the homogeneous case measurements were more informative. Although for both the homogeneous approximation of the heterogeneous riverbed and the characterization of the heterogeneous riverbed only a limited number of point measurements was used, those point measurements were more informative to estimate a single unknown homogeneous value than to estimate spatially heterogeneous fields.

3) For the variably saturated case, for which the homogeneous approximation should be more problematic according to the findings of Irvine et al. (2012), for a large part of the riverbed the saturation condition (i.e., saturated or unsaturated) did not change over time. This implies that the sketched saturation condition in Figure 4.2 was not very dynamic over time, and under such conditions the homogeneous approximation is less problematic than for highly dynamic variably saturated conditions. However, this is controlled by both the dynamic model forcings and the ratio of K_{rb} and K_{aq}. This is another limitation of this study.

Altogether points (1)-(3) show that for approximating a heterogeneous riverbed with a homogeneous value it is important to distinguish between the upscaling problem as outlined by Irvine et al. (2012) and the inverse problem, where typically a limited number of measurement data is available (as was the case in this study). Moreover, in this study the aquifer is simplified and assumed to be known and homogeneous over the whole simulation period. The homogeneity assumption affects the river-aquifer exchange fluxes and simplifies the parameter estimation problem. In reality, a more complex, heterogeneous aquifer is common. Therefore, further study should focus on the role of a spatially heterogeneous aquifer on the estimation of riverbed properties with different geostatistical models for those riverbed properties. In addition, further evaluation is planned for a real-world case study to inversely estimate riverbed properties by data assimilation. This will provide further insight into understanding the role of complex heterogeneous patterns within a more dynamic and uncertain river-aquifer system.

4.6 Conclusions

We compared four geostatistical models of riverbed hydraulic conductivity (K_{rb}) for simulating river-aquifer interactions with the integrated surface water-groundwater model HydroGeoSphere. HydroGeoSphere calculates a fully coupled feedback between surface water and groundwater, and river-aquifer exchange fluxes are better approximated when an unsaturated zone is present between surface water and aquifer. In this work, the reference (“true”) K_{rb} field was drawn from a non-multi-Gaussian distribution with channelized structures. Four different geostatistical models were compared: a homogeneous model, a multi-Gaussian model, a non-multi-Gaussian model with elliptic structures and a non-multi-Gaussian model with channelized structures. 200 stochastic realizations of K_{rb} fields were generated for each of the four geostatistical models, and served as input parameter files for model simulations and data assimilation experiments. Model simulations were done for (1) a setup with saturated conditions in and below the riverbed and (2) another setup with variably saturated conditions. Piezometric head measurement data were used for updating model states (heads) and parameters (K_{rb}) with the ensemble Kalman filter. The performance was evaluated in terms of its ability to characterize hydraulic heads, K_{rb} and net river-aquifer exchange fluxes. From the analysis above, we conclude:

1) Assimilation of hydraulic head data improved the characterization of hydraulic heads, K_{rb} and river-aquifer exchange fluxes for all four geostatistical models, even though in three out of four cases the prior geostatistical models did not coincide with the reference geostatistical model.

2) In the case of a fully saturated river-aquifer system, both the non-multi-Gaussian and the multi-Gaussian geostatistical models outperform the homogeneous model in terms of characterization of model states, parameters and exchange fluxes. However, the differences in performance between the three heterogeneous models are minor, indicating that the spatial pattern of K_{rb} has, in most cases, only a limited influence on river-aquifer exchange fluxes, similar to what has been shown for a one-way coupled model by Tang et al. (2015).

3) In the case of variably saturated conditions in and under the riverbed, conclusions are similar for the characterization of model states and parameters. However, concerning the characterization of net river-aquifer exchange fluxes results are slightly different. The performance of a multi-Gaussian K_{rb} field and a homogeneous equivalent K_{rb} are clearly worse than of the two non-multi-Gaussian models. The two different non-multi-Gaussian models, however, performed similarly. It can thus be concluded that for a variably saturated river-aquifer system the exact spatial pattern of the heterogeneous riverbed hydraulic conductivities is not important for characterizing river-aquifer exchange fluxes, but besides mean and variance also histogram information of the K_{rb} is valuable for selecting an appropriate geostatistical model. However, this might not be true for simulating solute transport can be stronger affected by connectivity and complex spatial patterns of K_{rb}.

Chapter 5 Simulating flood induced riverbed transience with physically-based modelling and ensemble Kalman filter*

5.1 Introduction

The riverbed plays a key role for modeling hydrological processes and changes, especially concerning stream-aquifer interactions. When a riverbed is considered within a hydrological simulation model, both riverbed topography and the K_{rb} are critical variables to be quantified. The riverbed topography is not only a main control for surface water - groundwater exchange fluxes (Harvey and Bencala, 1993; Woessner, 2000; Kasahara and Wondzell, 2003; Tonina and Buffington, 2007; Hester and Doyle, 2008; Cardenas, 2009; Shope et al., 2012; Zhang et al., 2017) but also for numerous biogeochemical processes in the hyporheic zone (Boano et al., 2010a; Harvey et al., 2012; Hester et al., 2013; Marzadri et al., 2014; Wildhaber et al., 2014). The K_{rb} also plays an important role in the dynamics of exchange fluxes and solute transport processes between the river and the surrounding groundwater system (e.g., Woessner, 2000; Cardenas et al., 2004; Ryan and Boufadel, 2006; Song et al., 2007; Kalbus et al., 2009; Käser et al., 2009; Irvine et al., 2012; Lackey et al., 2015; Miller et al., 2016; Song et al., 2016; Schilling et al., 2017b).

Riverbed topography as well as its texture and materials may change in time and space due to erosion and clogging processes at the river-aquifer interface. This will also cause changes in K_{rb} and the corresponding surface water - groundwater exchange fluxes (Anderson et al., 1999; Calver, 2001; Doppler et al., 2007; Genereux et al., 2008; Leek et al., 2009; Rosenberry and Pitlick, 2009; Boano et al., 2010b; Cuthbert et al., 2010; Hatch et al., 2010; Nowinski et al., 2011; Rosenberry, 2011; Sebok et al., 2015; Ulrich et al., 2015; Xi et al., 2015; Burnette et al., 2016; Grischek and Bartak, 2016; Wang et al., 2017). Both measurements from field and laboratory experiments as well as results from

* adapted from: Tang, Q., Schilling, O. S., Kurtz, W., Brunner, P., Vereecken, H., and Hendricks Franssen, H.-J., 2018, Simulating flood induced riverbed transience using unmanned aerial vehicles, physically-based hydrological modelling and the ensemble Kalman filter. Under review of Water Resources Research.

numerical models provide evidence of this: Nowinski et al. (2011) observed temporally varying K_{rb} on a riverbed point bar at a meander scale throughout the period of one year; originally high K_{rb} decreased due to the accumulation of fine materials transported by the river. Grischek and Bartak (2016) reviewed and documented the temporal variation of the leakage coefficient from 1971 to 2015 for the Elbe River in Germany. The leakage coefficient is defined as the ratio of K_{rb} to the thickness of the clogging layer. The results from the field tests showed that the leakage coefficient increased by a factor of 1.3 to 3.3 during the observed period. Ulrich et al. (2015) investigated the spatiotemporal variability of riverbed permeability and the mechanisms behind it by monitoring the clogging process with different methods at the Russian River near Forestville, California, from May to November in 2012. A decrease of permeability was observed from the onset of September due to the formation and accumulation of biomass on the riverbed.

During and after a flood event, scouring and deposition can occur which changes the riverbed topography and the K_{rb} by disturbing the sediment and re-sorting the grain size of the materials (e.g., Kim et al., 1999; Freer et al., 2002; Birck, 2006; Schubert, 2006; Doppler et al., 2007; Mutiti and Levy, 2010; Levy et al., 2011; Tu, 2011; Harvey et al., 2012; Simpson and Meixner, 2012). Birck (2006) explored the impact of high river stages on scouring of the riverbed and the corresponding changes of K_{rb} along the Great Miami River at Charles M. Bolton Water Plant in southwest Ohio, USA. A scouring of up to 0.06 m and a total fluctuation of 0.17 m was detected during the flood in January 2015. However, the contribution of scouring to the variation of overall K_{rb} was minimal. Doppler et al. (2007) provided field evidence of temporal varying leakage coefficients playing an important role when simulating groundwater flow for the upper Limmat valley in Zürich, Switzerland. The groundwater flow model needed to be recalibrated after a flood event, as the infiltration rate from the river into the aquifer was increased due to the flood induced erosion and the corresponding increase in K_{rb}. Schubert (2006) observed a highly dynamic clogging process and a significant temporal variation in the permeability of the clogged layer with rising and lowering water level in the lower Rhine region, Germany. Mutiti and Levy (2010) investigated the temporal evolution of K_{rb}

during a storm event beginning on June 12, 2005 at Charles M. Bolton Water Plant in southwest Ohio, USA. Groundwater flow and heat transport were simulated with the model VS2DHI (Hsieh et al., 2000). An increase of K_{rb} by almost one magnitude was observed due to the loss of fine materials during the scouring process. Simpson and Meixner (2012) used a series of surface and groundwater flow models to simulate a synthetic flood event and its effect on K_{rb} and surface - groundwater interactions. The total amount of scouring and filling was 12mm and 16mm, respectively; relatively small compared to the flood flow rate. The vertical K_{rb} was cumulatively increased by a factor of 14 in the beginning 15 hours of the flood event before the deposition started, and reduced afterwards 75% due to the generation of thin and low-conductivity layers on top of the riverbed.

These findings make it evident that it is important to characterize the spatiotemporal variation of K_{rb} induced by a flood, as well as the corresponding dynamic influence on the hydraulic states and fluxes of the river-aquifer system. Most of the applications mentioned above directly measured the K_{rb} by either field or laboratory experiments. However, indirect methods such as inverse modeling can also be an efficient approach for estimating K_{rb}. Data assimilation is one possible inverse modeling method for this purpose (Chen and Zhang, 2006; Hendricks Franssen and Kinzelbach, 2008). Until now it has already been successfully applied for estimating K_{rb} in several cases (e.g., Hendricks Franssen et al., 2011; Kurtz et al., 2012; Kurtz et al., 2013; Kurtz et al., 2014; Tang et al., 2015). Hendricks Franssen et al. (2011) estimated leakage coefficients for a limited number of zones, together with K_{aq} using a 3-D variably saturated groundwater flow model of the Limmat valley, Switzerland. The prediction of hydraulic heads was improved if both parameters were updated using the ensemble Kalman filter (EnKF). Kurtz et al. (2012) successfully captured the spatial and temporal variation of K_{rb} for a synthetic 3-D river-aquifer model with joint updating of piezometer heads and K_{rb} by EnKF. Later, the importance of K_{rb} heterogeneity in the Limmat valley river-aquifer system was studied by Kurtz et al. (2013). The heterogeneous K_{rb} as well as the piezometric heads can be even better estimated when both groundwater temperature data and piezometric heads were assimilated (Kurtz et al., 2014). Tang et al. (2015) investigated the impact of different K_{rb} patterns on the estimation of hydraulic heads, K_{rb}, and river-aquifer exchange fluxes for a synthetic experiment with a 3-D conductance based groundwater model. The surface water heads are pre-defined as fixed head boundary and exchange fluxes between the river and the aquifer are calculated according to Darcy's law. Using this kind of flow model approach, we found that complex K_{rb} patterns did not have a significant influence on the characterization of river-aquifer exchange fluxes. This was further explored by Tang et al. (2017) using a physically - based, fully integrated surface water-groundwater model able to simulate variably saturated conditions. Under fully saturated conditions, results were consistent with those shown in Tang et al. (2015); under variably saturated conditions, however, the histogram of K_{rb} provided useful additional information for the characterization of net river-aquifer exchange fluxes, suggesting that the pattern of K_{rb} is nevertheless important.

To the authors' knowledge, however, until now no studies inversely estimated the spatiotemporal variation of a flood induced K_{rb} with the help of hydraulic head measurements combined with transient riverbed topography data in an integrated hydrological model. Whether such K_{rb} variations during and after a flood event can be reproduced by hydraulic head data is still unknown, and the importance of K_{rb} variations for the prediction of hydraulic heads and fluxes between surface water and groundwater is also still unknown. Therefore, in this paper we investigate the spatially and temporally varying K_{rb} induced by a 300-year flood event and analyze its effect on the prediction of hydraulic heads, surface water discharge and river-aquifer exchange fluxes.

For this purpose, a 3-D river-aquifer model was set up for the site where the flooding occurred (Emme valley, Switzerland) using the physically based, fully integrated surface water-groundwater model HydroGeoSphere (HGS), and the hydraulic response to the flood event was simulated. Two high resolution datasets of riverbed topography, one before and one after the flood event, were used to define the riverbed topography in the model. An ensemble of different K_{aq} and K_{rb} fields was generated using sequential Gaussian simulations, and served as input files for the HGS model. Data assimilation experiments were carried out using the EnKF. After data assimilation, one year verification experiments were carried out to evaluate the performance of EnKF to reproduce the hydraulic heads and surface water discharge.

The data assimilation experiments were designed to systematically answer the following questions:

1) Assuming that only information on the riverbed topography is available for the pre- and post-flood period, but no hydraulic information: Does a change in riverbed topography have an influence on hydraulic heads, surface water discharge and river-aquifer exchange fluxes and is it important to consider this change

2) Assuming that measurements of hydraulic heads are available for the pre- and post-flood period, but no information on the post-flood riverbed topography: Could assimilation of the hydraulic head measurements of the pre-flood period allow calibrating the K_{aq} and K_{rb} sufficiently well so that the hydraulic heads and SW discharge of the post flood period can be reproduced Moreover, if K_{aq} and K_{rb} are continuously updated through data assimilation, but the riverbed topography remains unchanged, can the updated parameters compensate for the missing riverbed topography and produce as good results as when the correct (post-flood) riverbed topography would be used

3) Assuming that both hydraulic head measurements and riverbed topography information are available for the pre- and post-flood period**:** How fast can data assimilation based on EnKF adapt the K_{aq} and K_{rb} using information of hydraulic heads and riverbed topography

4) Does the correct estimation of K_{aq} or the correct estimation of K_{rb} play a more important role for the simulation of exchange fluxes How does this change if information on the flood-induced changes to the riverbed topography is available

5.2 Materials and methods

5.2.1 Integrated hydrological model HydroGeoSphere

HydroGeoSphere (HGS) (Therrien et al., 2010; Brunner and Simmons, 2012; Aquanty Inc, 2016) is a physically-based, integrated hydrological model designed for accounting for all components in the water cycle. The surface water and groundwater flow equations are solved simultaneously. Surface water flow is described by the two-dimensional Saint Venant equation. Groundwater flow is implemented by the three-dimensional Richards' equation. The dual node approach is used for the flow coupling between the surface and the subsurface domain. Van Genuchten functions (Van Genuchten, 1980) are used to describe the relationship between the hydraulic conductivity of the porous medium, the soil water content and pressure.

HGS has been used for the simulation of many different hydrological systems, e.g. for the interactions between groundwater, surface water and vegetation (e.g. Banks et al., 2011; Schilling et al., 2014; Ala-aho et al., 2017), for the detection of the sensitivity of catchment scale dynamics to the different parameters (Cornelissen et al., 2016), to track catchment scale surface water levels and overland flow routing (Ameli and Creed, 2017), to explore the hydrological dynamics of wetlands (Liu et al., 2016) and micro-topographic wetland runoff (Frei et al., 2010), as well as for large scale solute transport (Blessent et al., 2011). In many instances HGS was calibrated against different types of data (e.g. Karan et al., 2014; Schilling et al., 2014; Schilling, 2017) using the automatic inverse code PEST (Doherty, 2015). Recently, HGS has been coupled to a sequential data assimilation routine (EnKF-HGS, (Kurtz et al., 2017)).

5.2.2 Ensemble Kalman Filter

The data assimilation approach used in this study is the ensemble Kalman filter (EnKF) described by Evensen (1994). It has first been developed for the estimation of system states and later extended for the estimation of parameters (Chen and Zhang, 2006; Hendricks Franssen and Kinzelbach, 2008). The following description contains the EnKF algorithm as it is coupled to HGS.

EnKF consists of three basic equations: the forecast equation, the observation equation and the updating equation. Being a Monto-Carlo based inverse modelling method, EnKF uses a (large) number of stochastic realizations of model states and parameters to represent the model uncertainties. The augmented state vector includes both the model states and model parameters. For each stochastic realization, the augmented state vector can be written as:

$$\mathbf{x}_i = \begin{pmatrix} \mathbf{x}_s \\ \mathbf{x}_p \end{pmatrix}_i \tag{5.1}$$

where $\mathbf{x}_s$ is the vector with model states and $\mathbf{x}_p$ the vector of model parameters. The subscript i is the realization number. In our case, the augmented state vector contains one type of state variable (hydraulic head) and one type of model parameter (log-transformed hydraulic conductivity), therefore the state vector can be rewritten as:

$$\mathbf{x}_i = \begin{pmatrix} \mathbf{h} \\ \mathbf{Y} \end{pmatrix}_i \tag{5.2}$$

where **h** is hydraulic head and $\mathbf{Y} = \log_{10}(\mathbf{K})$. Here **K** can include K_{rb}, K_{aq}, or both K_{rb} and K_{aq}, depending on the assimilation scenario.

The model states at the current time step are predicted from the previous time step using the integrated surface-subsurface flow model HGS:

$$\mathbf{x}_{t,i} = M\left(\mathbf{x}_{t-1,i}\right) \tag{5.3}$$

where *M* is the forward model and the subscript *t* represent the time step.

The observations available at time step *t* are perturbed according the observation equation:

$$\mathbf{y}_{t,i} = \mathbf{y}_t + \boldsymbol{\varepsilon}_{t,i} \tag{5.4}$$

where $\mathbf{y}_{t,i}$ is the vector of perturbed measurements, $\mathbf{y}_t$ is the vector with original measurements at time step *t*, and $\boldsymbol{\varepsilon}_{t,i}$ is the vector with observation errors usually generated from a normally distribution with zero mean and standard deviation equal to the measurement error.

Combining the model prediction with the observations, the state vector is updated according the following analysis equation:

$$\mathbf{x}_{t,i}^{a} = \mathbf{x}_{t,i} + a\mathbf{G}\left(\mathbf{y}_{t,i} - \mathbf{H}\mathbf{x}_{t,i}\right) \tag{5.5}$$

where $\mathbf{x}_{t,i}^{a}$ is the augmented state vector containing the updated model states and parameters, $\mathbf{x}_{t,i}$ is the vector with the forecasted states obtained from the dynamic model, *a* is a damping factor varying between 0 and 1, **H** is the measurement operator matrix mapping the simulated states to the observation locations, and **G** is the Kalman gain which weights the relative importance of the model forecast and the observations. The Kalman gain is calculated by:

$$\mathbf{G} = \mathbf{C}\mathbf{H}^{\mathrm{T}}\left(\mathbf{H}\mathbf{C}\mathbf{H}^{\mathrm{T}} + \mathbf{R}\right)^{-1} \tag{5.6}$$

where **R** is a diagonal covariance matrix representing the measurement errors at individual observation locations, and **C** is the covariance matrix of the model states and parameters, given by

$$\mathbf{C} = \begin{pmatrix} \mathbf{C}_{\mathbf{hh}} & \mathbf{C}_{\mathbf{hY}} \\ \mathbf{C}_{\mathbf{Yh}} & \mathbf{C}_{\mathbf{YY}} \end{pmatrix} \tag{5.7}$$

Once the updating step is done, the updated model states and parameters will be used as input for the forward model for the next computation time step. Each time when observations are available, equations 5.4-5.7 are applied. The coupling of HGS and EnKF is described in detail by Kurtz et al. (2017).

5.2.3 Description of the study site

The Upper Emmental catchment in the Northern pre-alps of Switzerland is home to one of the largest drinking water stations of the Swiss capital Bern (Blau and Muchenberger, 1997; Käser and Hunkeler, 2016; Schilling et al., 2017a). The Emme River, which is the primary source of recharge for the alluvial aquifer of the Upper Emmental, has an average annual discharge of 4.4 m^3/s and is characterized by highly-dynamic discharge behavior (Käser and Hunkeler, 2016). The drinking water station is situated on the valley bottom towards the outlet of the catchment at an elevation of 690 m ASL (Figure 5.1). At that location, the topographic gradient is approximately 0.9 % (Käser and Hunkeler, 2016). A total rate of 0.4 m^3/s of groundwater (GW) is pumped from 8 wells located on a wellfield in the ultimate vicinity of the Emme River. The unconfined alluvial aquifer, out of which the GW is pumped, is composed of highly permeable quaternary alluvial gravel (~80%) and sand (~20%) with an average hydraulic conductivity of around 4×10^{-3} m/s, and is limited below by an impermeable layer of freshwater molasses (Würsten, 1991). The average thickness of the aquifer around the drinking water station is 25 m (Würsten, 1991). Per well, GW is abstracted from a single depth, which in the three upstream wells (wells 1-3) is 10 m and in the five downstream wells (wells 4-8) is 15 m below ground (see Figure 5.1). An extensive monitoring network (see Figure 5.1) has been put in place in order to manage the drinking water abstraction (Blau and Muchenberger, 1997; Kropf et al., 2014; Lapin et al., 2014). Multiple studies have been carried out in order to better characterize and quantify the river-aquifer interactions: Figura et al. (2013, 2015) characterized and predicted the evolution of GW temperature at the drinking water wellfield based on future climate scenarios, indicating an increasing trend. Käser and Hunkeler (2016) used temperature and electrical conductivity measurements in the Emme River to show that directly upstream of the wellfield river water is mostly infiltrating into the subsurface, whereas due to a narrowing of the aquifer downstream of the wellfield GW is exfiltrating back into the stream. Two weirs located in the Emme River at the height of the wellfield are also influencing SW-GW interactions, producing locally losing conditions upstream and locally gaining conditions downstream of the weirs. Schilling et al. (2017a) carried out a multi-tracer study and found that the GW pumped at the drinking water wellfield contains approximately 50% older GW and 50% SW that infiltrated directly upstream of the wellfield. They showed that the key parameter controlling the SW-GW interactions of this system, and therefore mixing between recently infiltrated SW and older GW in the subsurface, is the permeability of the riverbed. Unfortunately, the K_{rb} around the wellfield is very variable and poorly known, but the mean is estimated to be approximately two orders of magnitude lower than the K_{aq} (Schilling et al., 2017a). Multiple studies show that the pumping rate strongly influences the SW-GW interactions around the drinking water wellfield, and this can pose a major threat both to drinking water quality as well as to the riparian ecosystem health (Blau and Muchenberger, 1997; Schilling et al., 2017a).

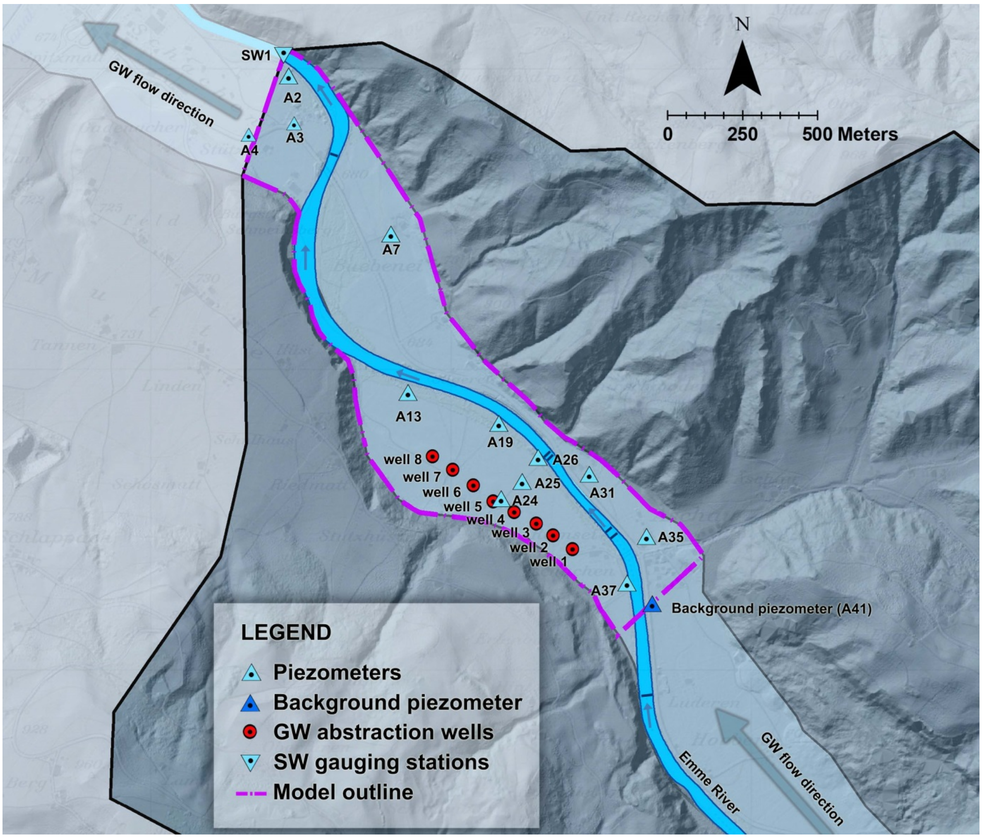

Figure 5.1: Location of the measurement stations of piezometer heads and surface water discharge within the study site.

5.2.4 Riverbed topography before and after the flood event

A violent flood with a return period of only 300 years and a peak discharge of 350 m^3/s occurred on July 24th, 2014, and substantially altered the riverbed topography and probably also changed riverbed properties. The riverbed topography was recorded in high resolution both before the flood (on February 12th 2014) and after the flood (on March 20th 2015). The riverbed topography was obtained with through-water photogrammetry of remotely sensed images taken from different angles. The method was reviewed in detail by Feurer et al. (2008). A high-resolution digital elevation model (DEM) of the floodplain was available from Swisstopo (2010). Based on the two riverbed DEMs and the floodplain DEM, two DEMs of the study area around the drinking water wellfield, representing the topography before and after the flood, were generated (Figure 5.2).

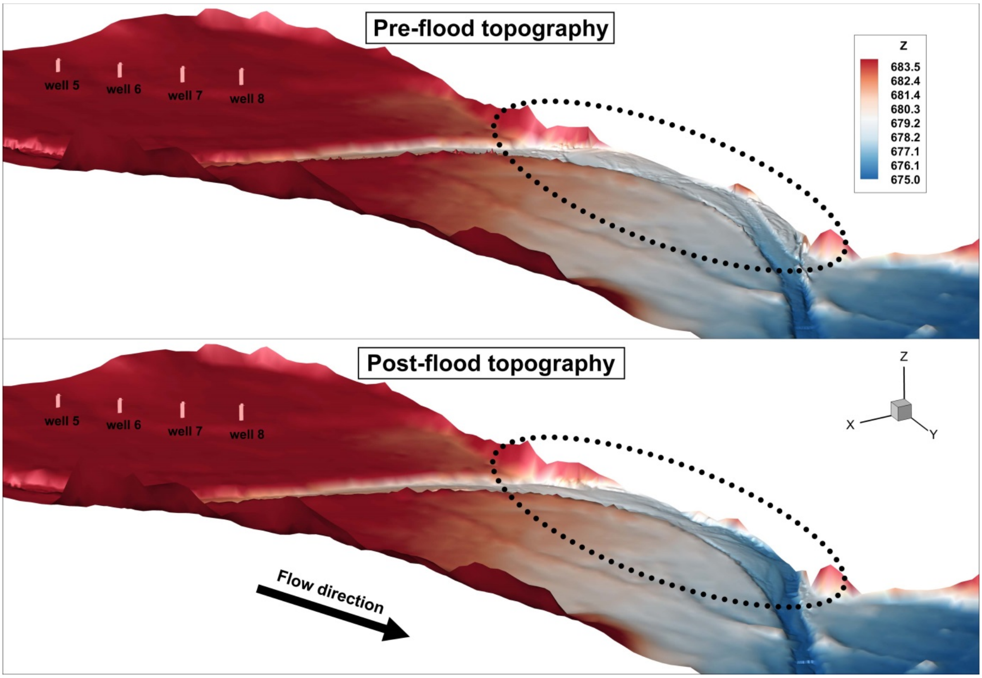

Figure 5.2: Two riverbed topography profiles before (a) and after the flood (b) (Schilling, 2017). The section where a lot of change happened is highlighted within the ellipse region.

5.3 The 3-D HGS model setup

5.3.1 Model conceptualization

5.3.1.1 Model discretization

The conceptual and numerical models used for this study have been described in detail by Schilling et al. (2017a): Vertically, the model is divided into 15 layers, and proportional sub-layering is used: the top 5 layers cover each 0.61 %, the following 4 layers each 6.1 %, and the bottom 5 layers each 12 % of the total aquifer depth, which results in 0.28 m, 2.8 m and 5.5 m thick layers at the location where the aquifer has the largest vertical extent (46 m). Horizontally, the model is discretized by equilateral triangles of 17.5 m side length on the floodplain and 8.5 m side length within the Emme River, resulting in 10983 elements per layer, i.e. 5645 nodes per slice. The model outline is illustrated in Figure 5.1. The automatic time-stepping scheme implemented in HGS was used.

5.3.1.2 Model parametrization

The riverbed was conceptualized as a slightly clogged layer within the Emme River boundaries, spanning across the 4 topmost model sublayers. An average K value for the riverbed of 2.8 x 10^{-5} m/s was found to best reproduce the observed mixing ratios between older GW and freshly infiltrated SW in the wellfield (Schilling et al., 2017a). For the aquifer, a homogeneous value of 2.9 x 10^{-3} m/s, which is about two magnitudes higher than the riverbed, is used to represent the K_{aq}. These homogeneous K values were used as starting points for the generation of spatially distributed K fields (described below in the section 5.3.2). The porosity of the aquifer layers was set to 0.15, representing a typical value of gravel-sand aquifers (Fetter, 2000; Anderson et al., 2015), and 0.41 for the riverbed layers. The unsaturated van Genuchten parameters α and β were set to 3.48 m^{-1} and 1.75, respectively, representing typical values for gravel-sand aquifers according to (Li et al., 2008). The residual saturation (S_{wr}) was fixed at 0.05. The coupling length (l_{exch}) between the SW and the GW domain was set to a very small value of 0.001 m, representing an optimal compromise between head continuity and numerical stability (see de Rooij (2017)).

5.3.1.3 Boundary conditions

The daily changing transient boundary conditions were based on measurement time series aggregated to daily values. Precipitation was conceptualized as a second type, specified flux boundary condition, and input time series of precipitation were obtained from the nearby weather station in Langnau i. E. operated by MeteoSuisse ([lat/long]: 07°48.33' / 46°56.38', elevation [mASL]: 745). Evapotranspiration was addressed by correcting the precipitation input time series with the actual evapotranspiration rate, calculated after Spreafico and Auer (2005), and using solar radiation and air temperature measurements of the same weather station as input. The Emme River inflow was also conceptualized as a specified flux boundary condition, and discharge input time series were obtained from two measurement stations located 5 km upstream of the upstream model boundary. The SW outflow boundary was conceptualized as a critical depth boundary condition. On the upstream boundary of the model, GW inflow was implemented as a first type, constant head boundary condition using measured GW levels from the piezometer located directly on the upstream model boundary (A41). To account for mounding due to losing conditions underneath the river in the upstream, the hydraulic head underneath the riverbed was fixed 1 m above the hydraulic head measured in that piezometer. For the downstream, constant head boundary condition, GW levels measured in a piezometer on the downstream model boundary (A3) were used. The lateral model boundaries and the bottom of the model were set impermeable. The GW pumps of the drinking water station were implemented as single node specified flux boundary conditions at the corresponding depth of each well. GW abstraction time series were provided by the water works association of the region of Bern (WVRB).

5.3.2 Hydraulic conductivity fields

In this study, spatially heterogeneous K_{aq} and K_{rb} fields were adopted. As Tang et al. (2017) demonstrated, multi-Gaussian distributed *K* fields provide good results for fully-connected river-aquifer systems when an integrated hydrological model is used, and non-multi Gaussian *K* fields are not required. We therefore used a multi-Gaussian geostatistical model for the generation of both random K_{rb} and K_{aq} fields. The generated random *K* fields served as the initial parameter fields for the data assimilation experiments. Figure 5.3 shows the initial ensemble mean as well as one stochastic realization of the random *K* fields for both riverbed and aquifer.

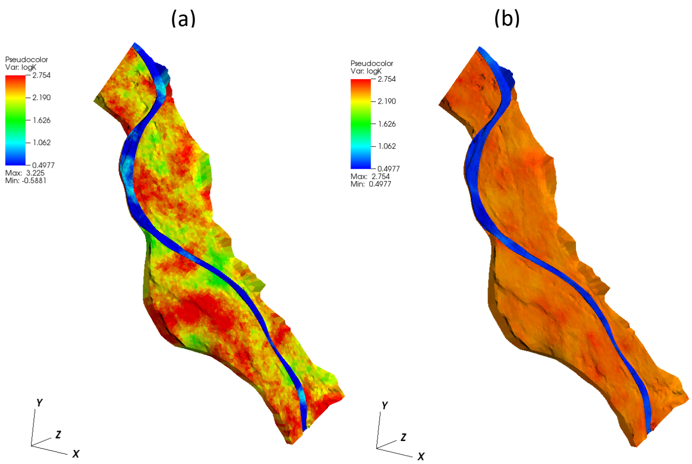

Figure 5.3: (a) One stochastic realization of the initial K_{aq} and K_{rb} field and (b) the ensemble mean of the initial *K* fields, calculated over 128 realizations.

For the aquifer, 128 stochastic realizations of multi-Gaussian distributed random K_{aq} fields were generated with the sequential Gaussian simulation algorithm in SGeMS (Remy et al., 2009). The geostatistical parameters used for generating the variogram are listed in Table 5.1. A spherical variogram was selected, with a mean value of -2.53 $\log_{10}$ (m/s) and a sill value of 0.1 $\log_{10}$ (m^2/s^2). The semi-variogram range was set to 150 m which is approximately 10% of the horizontal model domain size and which was used in absence of better information on the range parameter. The sequential Gaussian simulations were conditioned to two hydraulic conductivity values obtained from pumping tests (Würsten, 1991).

Table 5.1: Geostatistical parameters of the variogram used for generating the K_{aq} and K_{rb} fields.

	Type	Mean ($\log_{10}$ m/s)	Nugget	Range (m)	Sill ($\log_{10}$ m/s)
K_{aq}	Spherical	-2.53	0	150	0.1
K_{rb}	Spherical	-4.55	0	150	0.3

For the riverbed, random K_{rb} fields were generated in a similar way. Table 5.1 provides the corresponding geostatistical parameters for the variogram. A mean value of -4.55 $\log_{10}$ (m/s) was used to generate the random K_{rb} fields, which is approximately 2 magnitudes lower than the K_{aq} values. The semi-variogram range was set to 150 m, and a sill value of 0.3 $\log_{10}$ (m^2/s^2) was used. The generation of the random K_{rb} fields was unconditional as no measurement data were available for the riverbed.

5.4 Numerical experiments with EnKF-HGS

5.4.1 Overview

Simulation experiments were carried out using the numerical model described in section 5.3.1. Data assimilation was performed using measurement data from the year 2014, and measurements from the year 2015 were used as a verification dataset. Due to the riverbed changing flood on July 24th 2014, starting on July 25th 2014 two models with different streambed topography were used: one model with the pre-flood riverbed topography, and one with the post-flood topography. Multiple assimilation scenarios were tested: In order to systematically investigate the influence of riverbed topography on the inverse estimation of riverbed and K_{aq} as well as the prediction of hydraulic heads and exchange fluxes, the data assimilation experiments were organized according to ten different scenarios, divided into two groups. In the first group, changes in riverbed topography induced through the flood event were considered, whereas in the second group these changes were not considered. Within each of the two groups, five scenarios were implemented, including one ‘open loop’ scenario without any update of model states (hydraulic heads) or parameters (K_{aq} and K_{rb}), and four scenarios with data assimilation. For all of these four scenarios, data assimilation was carried out before the flood event (updating hydraulic heads, K_{rb} and K_{aq}); however, the four scenarios differed in the parameters which were updated after the flood event: i) no update; ii) only K_{rb}; iii) only the K_{aq}; iv) both the K_{aq} and K_{rb}. Performance of different simulation scenarios was evaluated by the root mean square error (RMSE) of hydraulic heads and surface water discharge measured during the verification period.

5.4.2 Observation dataset

Hydraulic heads within the model domain were measured at seven observation points (see Figure 5.1): A4, A7, A19, A24, A25, A26 and A35. The SW measurement station SW1 is located on top of a weir directly on the downstream model boundary and measures the SW discharge leaving the study site. Figure 5.4 displays the measured hydraulic heads and the surface water discharge for the years 2014 and 2015. Among these stations, measurements from A24, A25 and A26 are available for the entire two-year simulation period, while for the locations A4, A7, A19 and A35 some gaps exist. The hydraulic head observations of the year 2014 are used for the data assimilation experiments, and the hydraulic head observations of the year 2015 are used as a validation dataset. The SW discharge measurements of the year 2015 are also used to evaluate the model performance.

5.4.3 Model spin-up

Prior to the data assimilation, a series of spin-up simulations was carried out in order to obtain suitable initial conditions: First, a steady state simulation with one randomly chosen realization of the stochastic K fields was run using boundary conditions corresponding to the annual average values of 2013. In a second step, individual exit-spin-up simulations for all 128 stochastic realizations of K fields were run for 20 days. The initial conditions calculated by the steady state spin up runs were used for these exit spin-up runs, and transient boundary conditions of the first 20 days of January 2014 were imposed. The exit-spin-up runs produced an adequate initial ensemble spread adapted to the multiple stochastic *K* fields for the subsequent data assimilation experiments.

5.4.4 Data assimilation and verification experiments

Data assimilation experiments were carried out, updating both the hydraulic heads and the hydraulic conductivities, by assimilation of the hydraulic head observations. Due to the riverbed-changing flood event, the assimilation year 2014 was divided into two periods: A 'pre-flood period' from 01.01.2014 until 24.07.2014, the day the flood occurred, and a 'post-flood period' from 25.07.2014 until 31.12.2014. As outlined in Section 5.4.1, the data assimilation scenarios were divided into two groups: In Group 1 the pre-flood riverbed was used, and the riverbed topography remained unchanged before and after the flood event (denoted by '_old'), while in Group 2 a new riverbed corresponding to the measured post-flood topography was used for the post-flood period (denoted by '_new'). As outlined in Section 5.4.1, each of the two groups consisted of five different simulation scenarios:

1) An 'open loop' scenario without assimilation and update, denoted by 'OL'.
2) A scenario where hydraulic heads, K_{aq} and K_{rb} were updated during the pre-flood period, but without update during the post-flood period, denoted by 'DA_OL'.
3) A scenario where hydraulic heads, K_{aq} and K_{rb} were updated during the pre-flood period, but during the post-flood period only hydraulic heads and K_{rb} were updated, denoted by 'DA_hKr'.
4) A scenario with the same pre-flood updates as for scenarios 1 and 2, but where during the post-flood period only hydraulic heads and K_{aq} were updated, denoted by 'DA_hKa'.
5) A scenario where hydraulic heads, K_{aq} and K_{rb} were updated both during the pre-flood period and the post-flood period of the assimilation year 2014, denoted by 'DA_hKrKa'.

A detailed description of all the simulation scenarios is provided in Table 5.2.

Table 5.2: Simulation scenarios with different riverbed topography (RBT), updating variables, and for different time periods.

	Before flood (2014.1.1-2014.7.24)				After flood (2014.7.25-2014-12-31)				Verification year 2015
	RBT	Update *h*	Update *Kr*	Update *Ka*	RBT	Update *h*	Update *Kr*	Update *Ka*	RBT
OL_old	old	N	N	N	old	N	N	N	old
OL_new	old	N	N	N	new	N	N	N	new
DA_OL_old	old	Y	Y	Y	old	N	N	N	old
DA_OL_new	old	Y	Y	Y	new	N	N	N	new
DA_hKr_old	old	Y	Y	Y	old	Y	Y	N	old
DA_hKr_new	old	Y	Y	Y	new	Y	Y	N	new
DA_hKa_old	old	Y	Y	Y	old	Y	N	Y	old
DA_hKa_new	old	Y	Y	Y	new	Y	N	Y	new
DA_hKrKa_old	old	Y	Y	Y	old	Y	Y	Y	old
DA_hKrKa_new	old	Y	Y	Y	new	Y	Y	Y	new

In order to eliminate any potential effects of initial conditions during the performance evaluation, pure verification experiments were carried out using measurements of the year 2015. For each of the different simulation scenarios, the same riverbed topography as used during post-flood period in 2014 was also used for the verification period. The simulated hydraulic heads and the calibrated *K* fields of the last time step of the assimilation period (i.e., end of December 31st 2014) were taken as initial conditions and parameter fields for the verification period. During the verification period, no data assimilation or update of states or parameters was carried out.

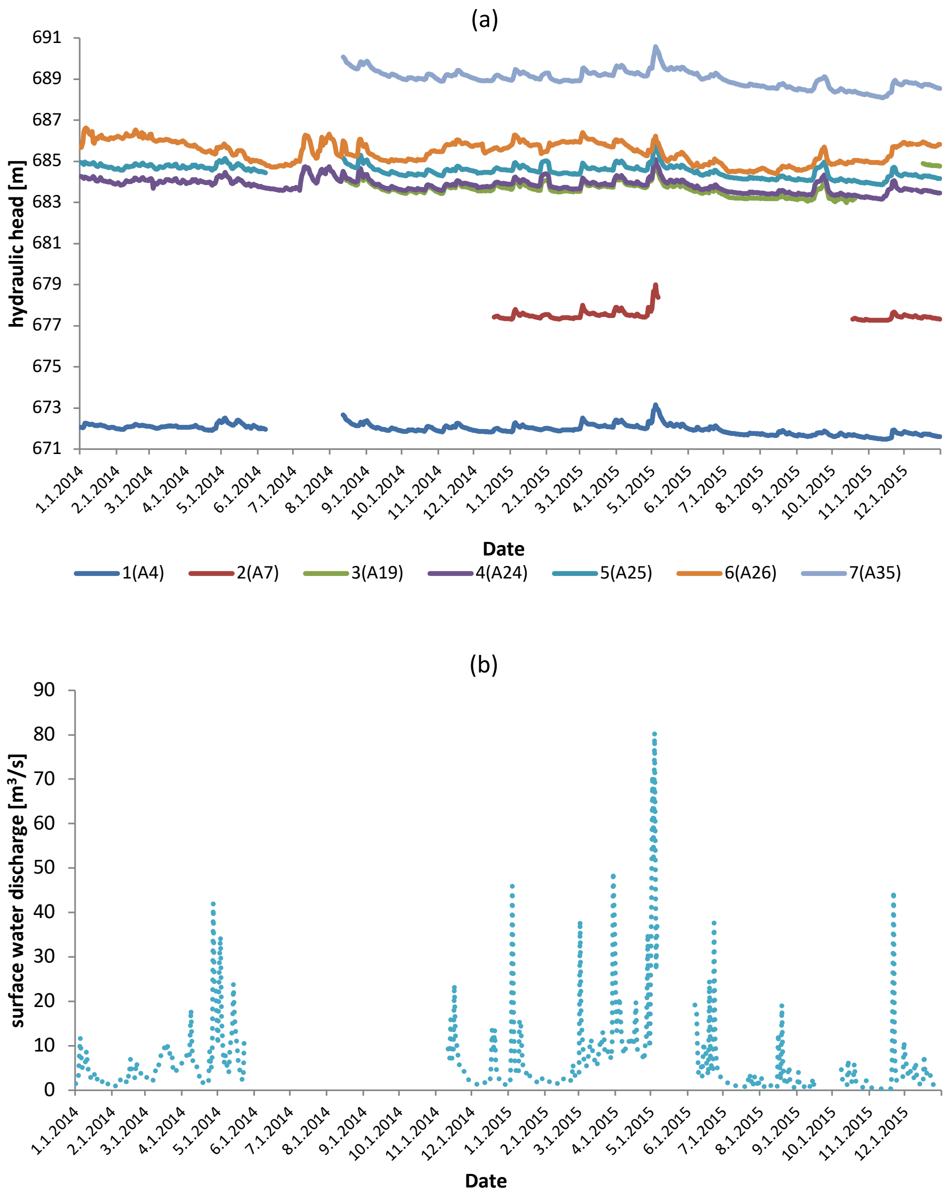

Figure 5.4: Measurements of (a) hydraulic head and (b) surface water discharge for the two-year simulation period 2014 and 2015.

5.4.5 Performance measures

The Root Mean Square Error (RMSE) was used to evaluate the model performance by comparing simulated and measured hydraulic heads and surface water discharge for the verification period. The

RMSE for the simulated hydraulic heads was calculated by comparing the ensemble mean value of the simulated heads to the measured hydraulic heads, averaged over all time steps and all observation locations:

$$\mathrm{RMSE}(h)=\sqrt{\frac{1}{n_t n_{\mathrm{nodes\_obs}}}\sum_{i=1}^{n_t}\sum_{j=1}^{n_{\mathrm{nodes\_obs}}}\left(\overline{h}_{i,j}^{f}-h_{i,j}^{\mathrm{obs}}\right)^2} \tag{5.8}$$

where h is the hydraulic head, n_t the number of simulation time steps, $n_{\mathrm{nodes\_obs}}$ the total number of the observation locations which differs for different simulation periods, the overbar indicates ensemble average, the superscript f indicates simulations and the superscript obs the observations. In our case, n_t equals to 365 and $n_{\mathrm{nodes\_obs}}$ is 7.

The RMSE for the surface water discharge is calculated by comparing the ensemble mean value of the simulated flow rate at the outlet, defined as the 'critical depth', with the measurement value, averaged over all time steps:

$$\mathrm{RMSE}(Q)=\sqrt{\frac{1}{n_t}\sum_{i=1}^{n_t}\left(\overline{Q}_i^{f}-Q_i^{\mathrm{obs}}\right)^2} \tag{5.9}$$

where Q is the surface water discharge.

Temporal evolution of RMSE for the simulated hydraulic heads and surface water discharge is also calculated by

$$\mathrm{RMSE}(h,t)=\sqrt{\frac{1}{n_{\mathrm{nodes\_obs}}}\sum_{j=1}^{n_{\mathrm{nodes\_obs}}}\left(\overline{h}_{j,t}^{f}-h_{j,t}^{\mathrm{obs}}\right)^2} \tag{5.10}$$

$$\mathrm{RMSE}(Q,t)=\left|\overline{Q}_t^{f}-Q_t^{obs}\right| \tag{5.11}$$

5.5 Results and discussion

5.5.1 Reproduction of heads

The scores of the RMSE (h) for different simulation scenarios are provided in Table 5.3. In the verification period, the RMSE (h) for the open loop run with the old riverbed topography (scenario OL_old) is 0.766m and with the new riverbed topography (scenario OL_new) 0.579m. The RMSE (h) was therefore reduced by 24% through the incorporation of the post-flood riverbed topography. This indicates that changes in riverbed topography have an influence on the estimation of hydraulic heads, and that the new and correct riverbed topography leads to a better reproduction of hydraulic heads if no calibration of the parameter fields is made.

If in the verification period the old riverbed topography is used, and if hydraulic heads, and K_{aq} and K_{rb} were only updated in the pre-flood period (scenario DA_OL_old) the RMSE (h) is 0.659m, which is 13.4% lower than its open loop counterpart (scenario OL_old). This shows that data assimilation can improve the simulation of hydraulic heads. However, the improvement is less compared to the scenario which used the new riverbed topography but where otherwise no update was carried out (scenario OL_new). This clearly illustrates that it is crucial to consider the actual riverbed topography of river-aquifer systems. If, for the model with the old riverbed topography, states and parameters are also updated in the post-flood period (scenario DA_hKrKa_old), the RMSE (h) is reduced by 51.7% (RMSE (h) = 0.370m) compared to the open loop with the old riverbed topography. Interestingly, this is only slightly larger than the scenario with continuous update of states and parameters throughout both the pre- and the post-flood periods and the new riverbed topography (scenario DA_hKrKa_new) (RMSE (h) = 0.348m). This last comparison illustrates that if a strong flood event changes the riverbed topography, but no information on these changes is available, data assimilation of hydraulic heads, in which both states as well as K_{rb} and K_{aq} are updated, can almost make up for the missing information on riverbed topography, at least in terms of the reproduction of hydraulic head.

Table 5.3: RMSE of simulated heads and surface water discharge for different simulation scenarios.

	RMSE(*h*) (m)	RMSE(*Q*) (m^3/s)
OL_old	0.766	4.17
OL_new	0.579	4.15
DA_OL_old	0.659	4.14
DA_OL_new	0.581	4.11
DA_hKr_old	0.449	4.06
DA_hKr_new	0.630	4.13
DA_hKa_old	0.513	4.09
DA_hKa_new	0.357	4.23
DA_hKrKa_old	0.370	4.04
DA_hKrKa_new	0.348	4.18

If the post-flood riverbed topography is used, but hydraulic heads, K_{aq} and K_{rb} are only updated in the pre-flood period (scenario DA_OL_new) the RMSE (*h*) is 0.581m. This performance is similar to the open loop counterpart (scenario OL_new) (RMSE (*h*) = 0.579m). This poor performance is probably due to the fact that the K_{rb} fields are estimated based on the pre-flood riverbed topography, which is not in correspondence with the post-flood riverbed topography. Nevertheless, the RMSE (*h*) is still 11.6% lower than the scenario where the riverbed topography was not updated after the flood (scenario DA_OL_old). This again highlights the importance of updating riverbed topography for estimating hydraulic heads. If with the post-flood riverbed topography parameters are also updated during the post-flood (scenario DA_hKrKa_new), the RMSE (*h*) is reduced to 0.348m, which is the best among all the scenarios. It is 39.9% lower than the corresponding open loop run OL_new and 40.1% lower than the scenario with only update in the pre-flood period. The reduction is less if compared to the scenario with the pre-flood riverbed topography (scenario DA_hKrKa_old, reduction 51.7 %), because the post-flood riverbed topography already provides information on the riverbed and estimation of the K_{rb} fields is less important than for the scenario where the old riverbed topography is used.

That information on riverbed topography is important can be further seen by comparing scenarios where both K_{aq} and K_{rb} are updated to scenarios where only one of them is updated. The former one refers to scenarios DA_hKrKa_old and DA_hKrKa_new, while the latter one refers to scenarios DA_hKr_old, DA_hKr_new, DA_hKa_old and DA_hKa_new. The RMSE (*h*) shows that both for the pre- or post-flood riverbed topography, updating only K_{aq} or only K_{rb} results in a poorer performance than updating both. With the pre-flood topography, the RMSE (*h*) for the scenario with update of only K_{rb} is 0.449m, and with update of only K_{aq} 0.513m. These values are 21.3% and 38.6% higher than the scenario with update of both K_{aq} and K_{rb} (RMSE (*h*) = 0.37m), respectively. Nevertheless, the performance comparison demonstrates that updating K_{rb} is more important than updating K_{aq}, most likely because the K_{aq} could already be sufficiently estimated during the pre-flood period, and because the flood event induced changes in K_{rb}, and/or because the wrong riverbed topography is compensated. If the post-flood riverbed topography is used after the flood event, the RMSE (*h*) for the scenario where only K_{rb} is updated is 0.63m and if only K_{aq} is updated it is 0.357m. This shows that if the new riverbed topography is incorporated in the model, updating K_{aq} is important and re-estimation of K_{rb} is not necessary anymore.

5.5.2 Evolution of aquifer and riverbed hydraulic conductivity

Figure 5.5 shows the ensemble mean *K* fields at the initial (t = 01.01.2014) and the final time step (t = 24.07.2014) of the pre-flood period. The $\log_{10}$ (K_{aq}) values are between -5.29 and -1.30 $\log_{10}$ (m/s) for the final updated field and between -4.44 and -2.19 $\log_{10}$ (m/s) for the initial field. Compared to the initial *K* field, the final updated ensemble mean *K* field shows a larger fluctuation by nearly two orders of magnitude. As no measurement data is available for K_{rb}, and the only two K_{aq} measurements are used for generating the initial K_{aq} fields, it is not possible to directly evaluate if the updated *K* fields are closer to the true values. However, these updated *K* fields improved the characterization of heads by 29% in the pre-flood period.

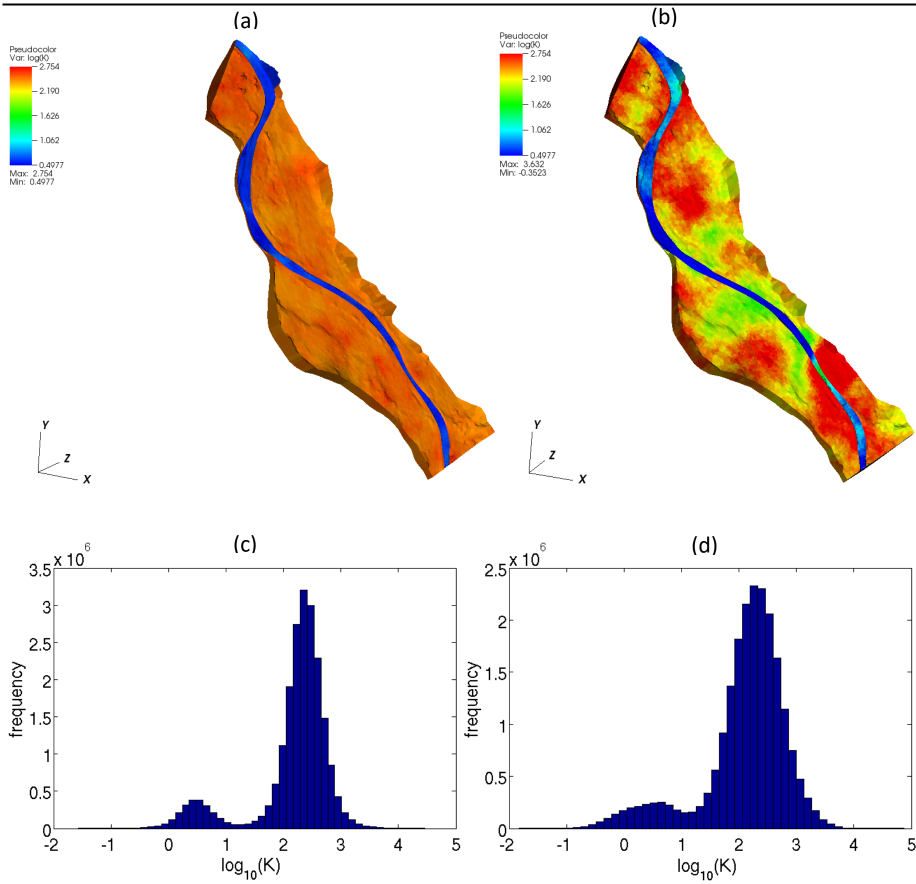

Figure 5.5: The ensemble mean of K fields (a) at the initial (t = 01.01.2014) and (b) the final time step of the pre-flood period (t = 24.07.2014) after data assimilation together with the corresponding histograms (c and d).

The temporal evolution of K_{rb} averaged over all the riverbed elements is displayed for different simulation scenarios in Figure 5.6. In the pre-flood period these four scenarios are exactly the same. The following results focus on the post-flood period. The estimated K_{rb} in the post-flood period is lower for scenarios with the pre-flood riverbed topography than for the scenarios with the post-flood one, which highlights the fact that the K_{rb} does not need to be altered that much for a good reproduction of hydraulic heads if the riverbed topography is also updated. Besides, either with the old or the new riverbed topography, the reduction of K_{rb} is larger for scenarios where both K_{aq} and K_{rb} are updated than for scenarios where only K_{rb} is updated.

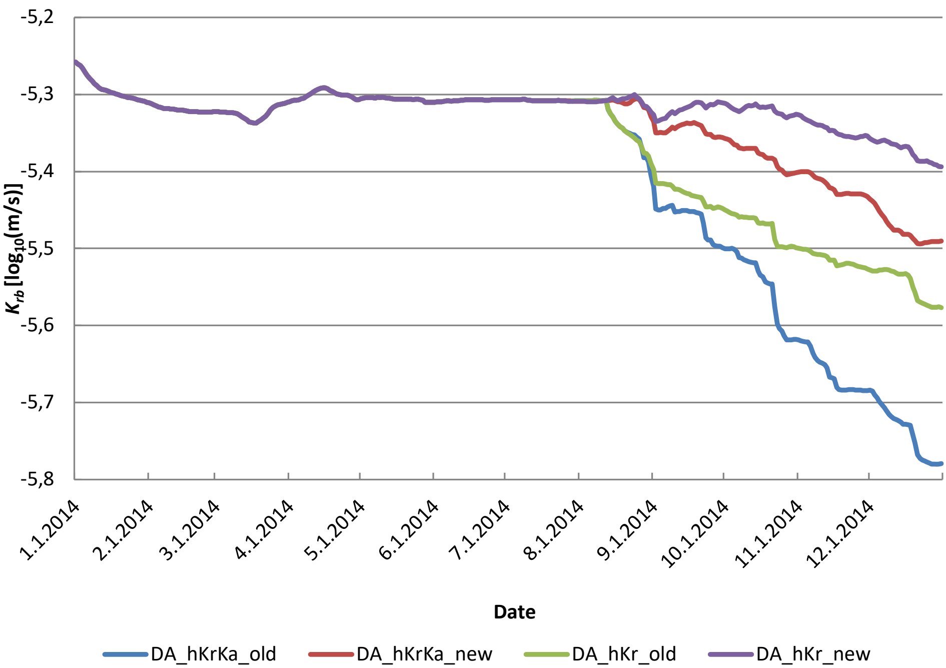

Figure 5.6: Temporal evolution of average K_{rb} for different simulation scenarios for the assimilation year 2014. Shown are four scenarios where different variables are updated in the post-flood period: scenario with update of heads, K_{rb} and K_{aq} using the old riverbed topography (blue line, DA_hKrKa_old), scenario with update of heads, K_{rb} and K_{aq} using the new riverbed topography (red line, DA_hKrKa_new), scenario with update of heads and K_{rb} using the old riverbed topography (green line, DA_hKr_old) and scenario with update of heads and K_{rb} using the new riverbed topography (purple line, DA_hKr_new).

Figure 5.7 shows the ensemble mean K fields at the end of the post-flood period (t = 31.12.2014) for scenarios with data assimilation also during the post-flood period. Scenarios with the post-flood riverbed topography show slightly higher and smoother K values for both the aquifer and the riverbed than scenarios with the pre-flood riverbed topography, which is also consistent with the analysis above and the head results in section 5.5.1.

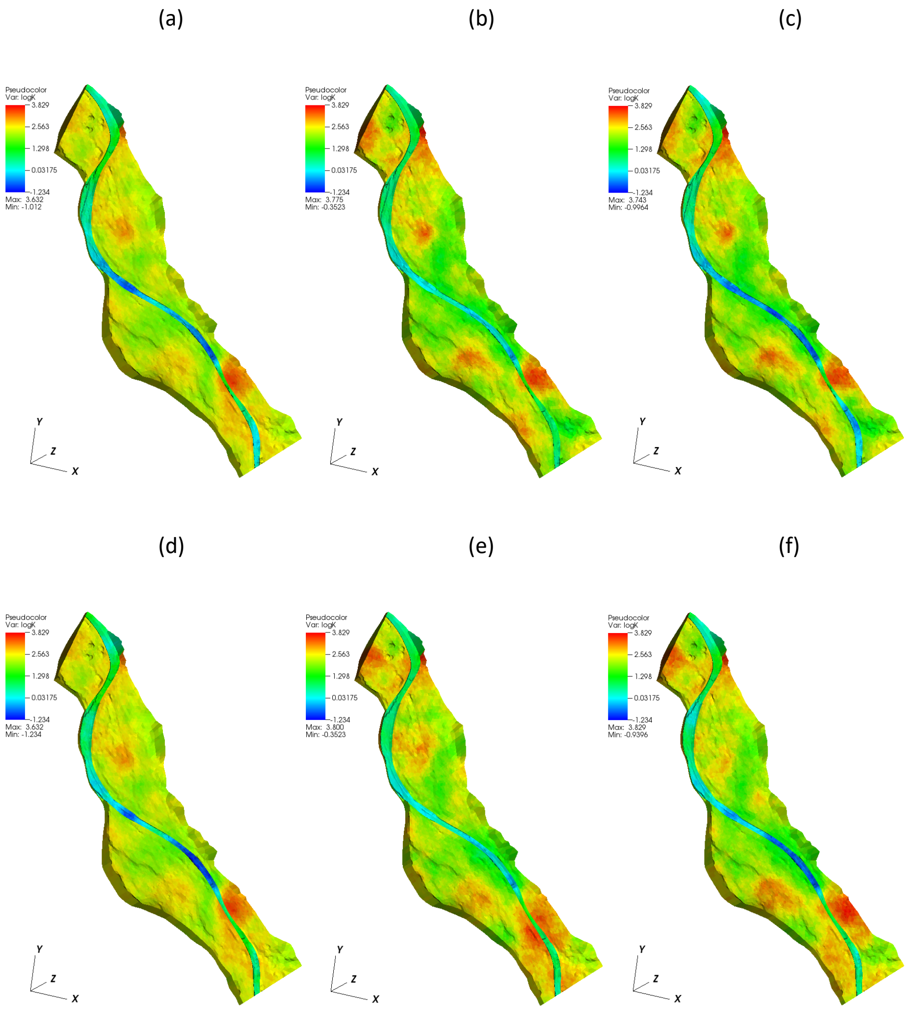

Figure 5.7: The final updated *K* fields at the end of the post-flood period (t = 2014.12.31) for different simulation scenarios: (a) scenario DA_hKr_old; (b) scenario DA_hKa_old; (c) scenario DA_hKrKa_old; (d) scenario DA_hKr_new; (e) scenario DA_hKa_new; (f) scenario DA_hKrKa_new. Scenarios using the old topography are displayed in the upper row, while scenarios using the new topography are displayed in the bottom row.

5.5.3 Simulation of exchange fluxes and surface water (SW) discharge (*Q*)

Figure 5.8 displays the temporal evolution of exchange fluxes between surface and subsurface water. In the pre-flood period all plotted scenarios give the same results as there are no differences between the runs. Differences among the scenarios are minor in the post-flood period, but become larger in the second half of 2015. Differences are particularly large between scenarios with the pre-flood and scenarios with the post-flood riverbed topography, which is consistent with the reduction of K_{rb} (see Figure 5.6). The lower the K_{rb} is, the less infiltration is observed. This shows that although data assimilation was able to compensate for a wrong riverbed topography in the post-flood period, the exchange fluxes between river and aquifer are clearly different for the scenarios with the correct and the wrong riverbed topography. Thus, it has to be concluded that the use of a wrong riverbed topography can only be partly corrected by data assimilation as states can be reproduced well, but exchange fluxes differ clearly between the scenarios with the correct and wrong riverbed topography.

RMSE (Q) for different simulation scenarios and evaluation period are provided in Table 5.3. The difference of RMSE (Q) between the two open loop runs with different riverbed topographies is 0.02 m^3/d. This indicates that an update of the riverbed topography also slightly improves the representation of SW discharge, but the impact is very small. If the pre-flood riverbed topography is used and hydraulic heads, K_{aq} and K_{rb} are updated only in the pre-flood period (scenario DA_OL_old), the RMSE (Q) is 4.14 m^3/d, which is only 0.03 m^3/d lower than for the open loop counterpart OL_old. However, if all of these variables are continuously updated by EnKF in the post-flood period (scenario DA_hKrKa_old), the RMSE (Q) is decreased to 4.04 m^3/d. This shows that without information on the post-flood riverbed topography, data assimilation based on EnKF can nonetheless improve the characterization of SW discharge via updating of parameters and hydraulic heads. Altogether, scenarios where both K_{aq} and K_{rb} are updated (scenario DA_hKrKa_old) resulted in the smallest RMSE (Q). This is consistent with the results for hydraulic heads discussed in section 5.5.1.

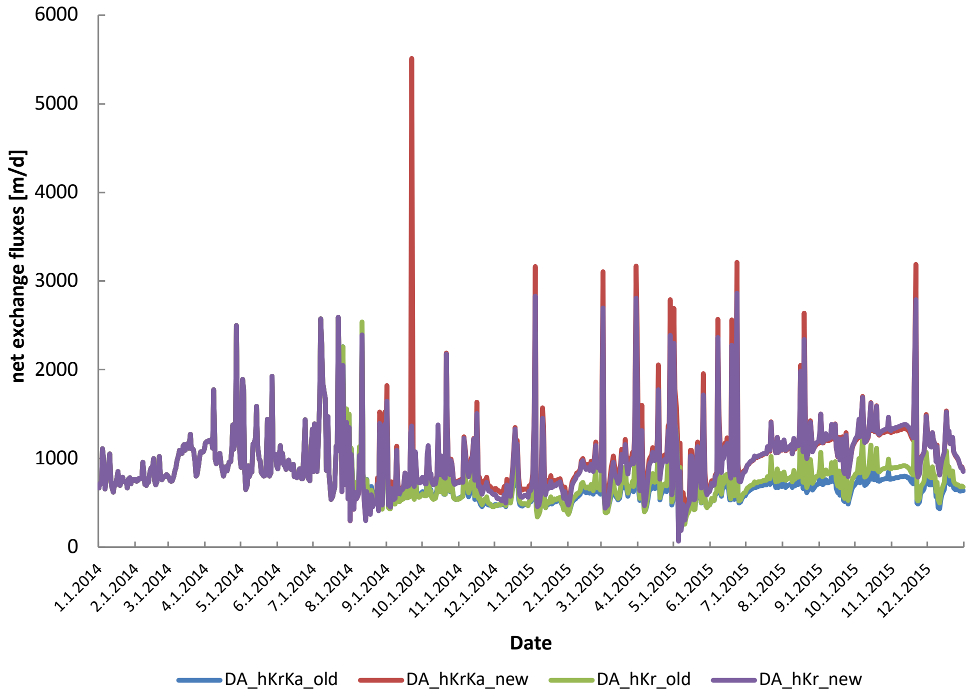

Figure 5.8: Temporal evolution of net exchange fluxes between the surface and subsurface domain. Shown is the mean value calculated over 128 realizations. Positive values indicate infiltration from the surface water into the subsurface.

If the post-flood riverbed topography is included in the simulations, the RMSE (*Q*) for the scenario where hydraulic heads, K_{aq} and K_{rb} are updated in the pre-flood period (scenario DA_OL_new) is 0.04 m^3/d lower than for the open loop counterpart (scenario OL_new). However, data assimilation in the post-flood period did not improve SW discharge characterization significantly. Only the scenario with update of K_{rb} (scenario DA_hKr_new) provided minimally better results than the corresponding open loop run (scenario OL_new). The temporal evolution of simulated SW discharge against the measurement (not displayed) also indicates that differences between the different simulation scenarios are small, and differences with the measurements are mostly small except for periods with large discharge rates, for example in May 2015. This illustrates that the updating the riverbed topography and K_{rb} has a small impact on the SW discharge, although the impact on the simulation of hydraulic heads and exchange fluxes is strong. This is mainly because the total amount of exchange fluxes is much smaller than the total amount of the SW discharge.

5.6 Conclusion

In this study, we investigated the spatial and temporal variation of the riverbed hydraulic conductivity (K_{rb}) and topography induced through a 300-year flood event and investigated its influence on the simulation of hydraulic heads, surface water discharge and river-aquifer exchange fluxes using the physically-based, fully integrated hydrological model HydroGeoSphere. Both riverbed topography data before and after the flood event were available as input for the simulation model. Data assimilation experiments were performed with the EnKF for pre- and post-flood assimilation periods, either under consideration of the changes of riverbed topography, or ignoring them. Heterogeneous multi-Gaussian distributed K_{aq} and K_{rb} fields were used as the initial parameter fields and later, together with hydraulic heads, updated by EnKF-HGS. Verification experiments of the post-flood year (2015) were used for evaluating the performance of the different parameter/states/topography updating schemes. The following conclusions can be drawn from the simulation experiments:

1) Changes in riverbed topography have a significant influence on the prediction of hydraulic heads and river-aquifer exchange fluxes, and a minor influence on the prediction of surface water discharge. Incorporation of information on the modified riverbed topography obtained via remotely sensed through-water photogrammetry improves the simulation of hydraulic heads and exchange fluxes significantly.

2) The estimation of K_{aq} and K_{rb} with help of data assimilation allows a better estimation of hydraulic heads and SW discharge, even if assimilation is only carried out before the flood event. The parameter estimation is especially important if changes in riverbed topography are not taken into account in the post-flood simulations, allowing compensating for this problem. However, the exchange fluxes between river and aquifer and clearly different for the scenarios with the correct

riverbed topography and the wrong riverbed topography, indicating that data assimilation only partly can correct for wrong information on the post-flood riverbed topography.

3) Both observations of the changes in riverbed topography and of hydraulic heads provide useful information for the simulation of river-aquifer systems. Without the riverbed topography information, estimation of both K_{rb} and K_{aq} is important. However, if information on the transient changes of riverbed topography is available, a re-estimation of K_{rb} seems to be less important.

Chapter 6 Summary and outlook

Integrated water resources management requires full consideration of surface water – groundwater interaction, especially when such interactions can strongly influence the water quantity and quality of surrounding drinking water stations (Winter et al., 1998). Our study site, the Upper Emme catchment in Switzerland, is such an area with strong river-aquifer interaction. Simulating these river-aquifer interactions relies strongly on the proper representation of riverbed properties like the riverbed topography and riverbed hydraulic conductivity (K_{rb}) (Schilling et al., 2017). Riverbed properties show a strong spatial variability and are also highly dynamic in time due to erosion and deposition processes especially in relation with extreme events like floods. Spatial and temporal variability of riverbed properties make the simulation of river-aquifer interaction very uncertain and dynamic. Specifically, previous studies usually adopted the multi-Gaussian assumption to model the spatial variability of K_{rb}. However, in reality, the K_{rb} shows more complex non-multi-Gaussian spatial patterns (Springer et al., 1999), which might have significant influence on the estimation of river-aquifer exchange fluxes. In order to investigate the role of spatial patterns of r K_{rb} as well as temporal variability, the sequential data assimilation technique ensemble Kalman filter (EnKF) was used, which is a powerful tool to inversely estimate model parameters by accounting for model uncertainties as well as measurement uncertainties. In this PhD-dissertation, the aim was to improve characterization of the spatiotemporal variation of K_{rb} by EnKF, and to investigate the role of riverbed topography and complex heterogeneous K_{rb} patterns on simulating river-aquifer exchange fluxes. These complex heterogeneous K_{rb} patterns are generated according to different geostatistical models under either the multi-Gaussian assumption or the non-multi-Gaussian assumption. Numerical experiments are carried out first for a simplified synthetic 3-D river-aquifer case using a conductance based groundwater model, and later for a similar synthetic case using a physically based, integrated hydrological model under both fully saturated and variably saturated conditions beneath the riverbed. These two synthetic experiments provide a controlled environment for process-studies and for detecting model sensitivities, as all model states and fluxes of the virtual 'truth' are exactly known at all times, which allows to precisely determine the efficiency of data assimilation and parameter estimation. Afterwards, a real world 3-D river-aquifer case study was tested for the Upper Emme catchment using the integrated hydrological model to explore the role of transient riverbed properties on the river-aquifer interaction. In the Emme catchment the riverbed properties greatly changed because of a 300-year flood which occurred on July 24th of 2014. For the real-world case two riverbed topography profiles obtained from through-water photogrammetry of remotely sensed images before and after the flood event were used, together with more traditional hydrogeological data types.

For the two synthetic studies, ten reference K_{rb} fields, which served as synthetic truths, were generated using a non-multi-Gaussian model with channelized structures, as several experimental studies show that this could be a more realistic model for the spatial variability of K_{rb} and since until now no study considered such type of K_{rb} pattern for modeling river-aquifer interaction. Three different geostatistical models were used to generate the initial heterogeneous K_{rb} fields, including one non-multi-Gaussian model with channelized structures, one non-multi-Gaussian model with elliptical structures and one multi-Gaussian model. For the synthetic study with the integrated hydrological model, a homogeneous geostatistical model was also used to generate the initial K_{rb} fields.

The role of different K_{rb} patterns was explored with these two synthetic experiments. Both of the two synthetic studies revealed that EnKF can improve the characterization of hydraulic heads and river-aquifer exchange fluxes. The K_{rb} characterization could also be improved, even if the prior geostatistical models differed from the reference (true) geostatistical model. Under fully saturated conditions beneath the riverbed, and both for a conductance based groundwater model and an integrated hydrological model, the differences in performance among the correct non-multi-Gaussian model, the erroneous non-multi-Gaussian model and the erroneous multi-Gaussian model were minor. This indicates that K_{rb} patterns have only a minor influence on river-aquifer interactions.

Moreover, NS-EnKF, a variant of EnKF which can handle better non-Gaussian distributed states and parameters due to a normal score transformation, did not show a better performance than standard EnKF, although parameters showed a non-multi-Gaussian distribution. The only main advantage was that with NS-EnKF, the histogram shape of K_{rb} could be preserved. Under variably saturated conditions, for the characterization of net river-aquifer exchange fluxes, an erroneous multi-Gaussian model performed clearly worse than the erroneous non-multi-Gaussian model and the correct non-multi-Gaussian model, while the two non-multi-Gaussian models gave similar performance. This indicates that for variably saturated conditions, again complex heterogeneous K patterns such as connectivity do not matter so much for characterizing net river-aquifer exchange fluxes, but now, and in contrast to fully saturated conditions, the histogram provides valuable information besides the mean and the variance value.

For the real world study, the characterization of both hydraulic heads and river-aquifer exchange fluxes was also improved by EnKF. The maximum improvement by EnKF for estimating the exchange fluxes was 3%, which is relatively small due to the relative small amount of exchange fluxes compared to the surface water discharge in this catchment. On the other hand, for the prediction of hydraulic head, the maximum RMSE-reduction with data assimilation, averaged over all observation locations, was 55% for the scenario where riverbed topography, hydraulic heads, K_{rb} and K_{aq} were updated.

Besides the K_{rb}, riverbed topography also plays an important role in simulating a river-aquifer system. Our real world case study shows that considering the changes in both the riverbed topography and K_{rb} leads to the best simulation results for hydraulic head and stream-aquifer exchange fluxes. However, if the riverbed topography information is missing, the K_{rb} can be estimated by EnKF and account for the missing information on topography, resulting in only slightly worse results than the simulation taking into account the changes of riverbed topography. If the riverbed topography is available, the re-estimation of K_{rb} is not that important.

In this thesis, we characterized the spatiotemporal varying K_{rb} and investigated the role of K_{rb} patterns on the estimation of river-aquifer exchange fluxes. All of the conclusions in this thesis are based on simulations in river-aquifer systems. As Li et al. (2012), Chen et al. (2013) and Tong et al. (2013) pointed out, ensemble based data assimilation methods such as EnKF and ensemble smoother can improve the characterization of heterogeneous K_{aq} by assimilating tracer data or solute concentration data. The impact of a heterogeneous riverbed on solute transport simulation was not investigated in this work. Therefore, in future, the sensitivity of solute transport through riverbeds with respect to non-multi-Gaussian and multi-Gaussian patterns of heterogeneous K_{rb} should also be analyzed. Moreover, further study should also investigate the role of these types of heterogeneous K_{rb} patterns on heat transport. As Kalbus et al. (2006) and Constantz (2008) pointed out, temperature can be used as a tool for the estimation of water fluxes through streambed sediments, as groundwater temperature is often relatively stable and stream temperature often shows stronger fluctuations on the daily and yearly scales. As a tracer for surface water infiltration and hydraulic conductivity, it is also inexpensive and naturally available over the complete stream reach (Anderson, 2005). Kurtz et al. (2014) found, for a study in the Upper Limmat Valley in Switzerland, that a heterogeneous riverbed can be better estimated by jointly assimilating groundwater temperature data and piezometric head data. However, Kurtz et al. (2014) adopted a multi-Gaussian K_{rb} pattern. Future work could also here focus on the role of different complex non-multi-Gaussian K_{rb} patterns on simulating heat transport and the value of additional temperature data measured in the riverbed. It is expected that both solute and heat transport are more affected by such complex spatial patterns of K_{rb} than flow only. Specifically, compared to the traditional isolated temperature data logger, temperature recorded by Distributed Temperature Sensing (DTS) can provide highly resolved data along buried cables instead of temperature data for isolated points, which allows detecting the precise location of groundwater inflows to river channels (Lowry et al., 2007; Sebok et al., 2013). DTS has also been used to detect river-aquifer exchange fluxes together with electrical imaging method (Slater et al., 2010). Vogt et al. (2010) used vertical high resolution temperature data obtained by DTS for the estimation of seepage rates to quantify river-groundwater exchange. K_{rb} and river-aquifer exchange fluxes can then inversely be estimated with the flow and heat transport model by assimilating the high resolution DTS temperature data together with the piezometer head data. An improved characterization of both the primary variables (e.g. the hydraulic heads and the K_{rb}) and the secondary variables (e.g. the surface water discharge and the river-aquifer exchange fluxes) can be expected.

References

Anderson, M. P., Aiken, J., Webb, E., and Mickelson, D., 1999, Sedimentology and hydrogeology of two braided stream deposits: Sedimentary Geology, v. 129, no. 3, p. 187-199.

Aquanty Inc, 2016, HydroGeoSphere. A Three-dimensional Numerical Model Describing Fully Integrated Subsurface and Surface Flow and Solute Transport: Waterloo, ON, Canada.

Bailey, R., and Baù, D., 2010, Ensemble smoother assimilation of hydraulic head and return flow data to estimate hydraulic conductivity distribution: Water Resources Research, v. 46, no. 12.

Bailey, R. T., and Bau, D., 2012, Estimating geostatistical parameters and spatially-variable hydraulic conductivity within a catchment system using an ensemble smoother: Hydrology and Earth System Sciences, v. 16, no. 2, p. 287-304.

Banks, E., Brunner, P., and Simmons, C. T., 2011, Vegetation controls on variably saturated processes between surface water and groundwater and their impact on the state of connection: Water Resources Research, v. 47, no. 11.

Bear, J., 1979, Hydraulics of Groundwater, 569, McGraw-Hill, New York.

Birck, M. D., 2006, Temporal variability of riverbed hydraulic conductivity at an induced infiltration site, southwest Ohio: Miami University.

Blau, R. V., and Muchenberger, F., 1997, Grundlagen für Schutz und Bewirtschaftung der Grundwasser des Kantons Bern: Nutzungs-, Schutz- und Überwachungskonzept für die Grundwasserleiter des obersten Emmentals, zwischen Emmenmatt, Langnau und Eggiwil, Synthesebericht, Wasser- und Energiewirtschaft des Kantons Bern, Bern.

Blessent, D., Therrien, R., and Gable, C. W., 2011, Large-scale numerical simulation of groundwater flow and solute transport in discretely-fractured crystalline bedrock: Advances in Water Resources, v. 34, no. 12, p. 1539-1552.

Brunke, M., and Gonser, T., 1997, The ecological significance of exchange processes between rivers and groundwater: Freshwater biology, v. 37, no. 1, p. 1-33.

Brunner, P., Cook, P. G., and Simmons, C. T., 2009, Hydrogeologic controls on disconnection between surface water and groundwater: Water Resources Research, v. 45, no. 1.

Brunner, P., and Simmons, C. T., 2012, HydroGeoSphere: a fully integrated, physically based hydrological model: Ground Water, v. 50, no. 2, p. 170-176.

Brunner, P., Simmons, C. T., Cook, P. G., and Therrien, R., 2010, Modeling surface water ‐ groundwater interaction with MODFLOW: some considerations: Groundwater, v. 48, no. 2, p. 174-180.

Brunner, P., Therrien, R., Renard, P., Simmons, C. T., and Hendricks Franssen, H.-J., 2017, Advances in understanding river - groundwater interactions: Reviews of Geophysics, p. n/a-n/a.

Burgers, G., Jan van Leeuwen, P., and Evensen, G., 1998, Analysis scheme in the ensemble Kalman filter: Monthly weather review, v. 126, no. 6, p. 1719-1724.

Caers, J., and Zhang, T., 2004, Multiple-point geostatistics: a quantitative vehicle for integrating geologic analogs into multiple reservoir models: AAPG Mem, v. 80, p. 383-394.

Calver, A., 2001, Riverbed permeabilities: Information from pooled data: Ground Water, v. 39, no. 4, p. 546-553.

Camporese, M., Paniconi, C., Putti, M., and Orlandini, S., 2010, Surface ‐ subsurface flow modeling with path ‐ based runoff routing, boundary condition ‐ based coupling, and assimilation of multisource observation data: Water Resources Research, v. 46, no. 2.

Camporese, M., Paniconi, C., Putti, M., and Salandin, P., 2009, Ensemble Kalman filter data assimilation for a process-based catchment scale model of surface and subsurface flow: Water Resources Research, v. 45, no. 10, p. W10421.

Cardenas, M. B., Wilson, J., and Zlotnik, V. A., 2004, Impact of heterogeneity, bed forms, and stream curvature on subchannel hyporheic exchange: Water Resources Research, v. 40, no. 8.

Carrera, J., Alcolea, A., Medina, A., Hidalgo, J., and Slooten, L. J., 2005, Inverse problem in hydrogeology: Hydrogeology journal, v. 13, no. 1, p. 206-222.

Carrera, J., and Neuman, S. P., 1986, Estimation of aquifer parameters under transient and steady state conditions: 1. Maximum likelihood method incorporating prior information: Water Resources Research, v. 22, no. 2, p. 199-210.

Chen, Y., and Zhang, D., 2006, Data assimilation for transient flow in geologic formations via ensemble Kalman filter: Advances in Water Resources, v. 29, no. 8, p. 1107-1122.

Cheng, C., Song, J., Chen, X., and Wang, D., 2011, Statistical distribution of streambed vertical hydraulic conductivity along the Platte River, Nebraska: Water resources management, v. 25, no. 1, p. 265-285.

Crestani, E., Camporese, M., Bau, D., and Salandin, P., 2013, Ensemble Kalman filter versus ensemble smoother for assessing hydraulic conductivity via tracer test data assimilation: Hydrology and Earth System Sciences, v. 17, no. 4, p. 1517-1531.

Delta h Ingenieugesellschaft mbH, 2006, Spring 3.3: Witten, Germany.

Doppler, T., Franssen, H.-J. H., Kaiser, H.-P., Kuhlman, U., and Stauffer, F., 2007, Field evidence of a dynamic leakage coefficient for modelling river–aquifer interactions: Journal of Hydrology, v. 347, no. 1, p. 177-187.

Elfeki, A., and Dekking, M., 2001, A Markov chain model for subsurface characterization: Theory and applications: Mathematical Geology, v. 33, no. 5, p. 569-589.

Ellis, P. A., Mackay, R., and Rivett, M. O., 2007, Quantifying urban river–aquifer fluid exchange processes: a multi-scale problem: Journal of contaminant hydrology, v. 91, no. 1, p. 58-80.

Evensen, G., 1994, Sequential data assimilation with a nonlinear quasi ‐ geostrophic model using Monte Carlo methods to forecast error statistics: Journal of Geophysical Research: Oceans, v. 99, no. C5, p. 10143-10162.

Fleckenstein, J. H., Niswonger, R. G., and Fogg, G. E., 2006, River ‐ aquifer interactions, geologic heterogeneity, and low ‐ flow management: Groundwater, v. 44, no. 6, p. 837-852.

Fox, G. A., and Durnford, D. S., 2003, Unsaturated hyporheic zone flow in stream/aquifer conjunctive systems: Advances in Water Resources, v. 26, no. 9, p. 989-1000.

Frei, S., Fleckenstein, J. H., Kollet, S. J., and Maxwell, R. M., 2009, Patterns and dynamics of river-aquifer exchange with variably-saturated flow using a fully-coupled model: Journal of Hydrology, v. 375, no. 3-4, p. 383-393.

Frei, S., Lischeid, G., and Fleckenstein, J., 2010, Effects of micro-topography on surface–subsurface exchange and runoff generation in a virtual riparian wetland—A modeling study: Advances in Water Resources, v. 33, no. 11, p. 1388-1401.

Furman, A., 2008, Modeling coupled surface–subsurface flow processes: a review: Vadose Zone Journal, v. 7, no. 2, p. 741-756.

Genereux, D. P., Leahy, S., Mitasova, H., Kennedy, C. D., and Corbett, D. R., 2008, Spatial and temporal variability of streambed hydraulic conductivity in West Bear Creek, North Carolina, USA: Journal of Hydrology, v. 358, no. 3, p. 332-353.

Goderniaux, P., Brouyère, S., Fowler, H. J., Blenkinsop, S., Therrien, R., Orban, P., and Dassargues, A., 2009, Large scale surface–subsurface hydrological model to assess climate change impacts on groundwater reserves: Journal of Hydrology, v. 373, no. 1, p. 122-138.

Gómez-Hernández, J. J., and Journel, A. G., 1993, Joint sequential simulation of multiGaussian fields, Geostatistics Troia'92, Springer, p. 85-94.

Gómez-Hernández, J. J., and Wen, X.-H., 1998, To be or not to be multi-Gaussian A reflection on stochastic hydrogeology: Advances in Water Resources, v. 21, no. 1, p. 47-61.

Gómez-Hernánez, J. J., Sahuquillo, A., and Capilla, J. E., 1997, Stochastic simulation of transmissivity fields conditional to both transmissivity and piezometric data--1. Theory: Journal of Hydrology(Amsterdam), v. 203, no. 1, p. 167-174.

Grischek, T., and Bartak, R., 2016, Riverbed Clogging and Sustainability of Riverbank Filtration: Water, v. 8, no. 12, p. 604.

Guardiano, F. B., and Srivastava, R. M., 1993, Multivariate geostatistics: beyond bivariate moments, Geostatistics Troia'92, Springer, p. 133-144.

Harvey, J. W., and Bencala, K. E., 1993, The effect of streambed topography on surface ‐ subsurface water exchange in mountain catchments: Water Resources Research, v. 29, no. 1, p. 89-98.
Hayashi, M., and Rosenberry, D. O., 2002, Effects of ground water exchange on the hydrology and ecology of surface water: Groundwater, v. 40, no. 3, p. 309-316.
Hendricks Franssen, H., Alcolea, A., Riva, M., Bakr, M., Van Der Wiel, N., Stauffer, F., and Guadagnini, A., 2009, A comparison of seven methods for the inverse modelling of groundwater flow. Application to the characterisation of well catchments: Advances in Water Resources, v. 32, no. 6, p. 851-872.
Hendricks Franssen, H., Kaiser, H., Kuhlmann, U., Bauser, G., Stauffer, F., Müller, R., and Kinzelbach, W., 2011, Operational real ‐ time modeling with ensemble Kalman filter of variably saturated subsurface flow including stream ‐ aquifer interaction and parameter updating: Water Resources Research, v. 47, no. 2.
Hendricks Franssen, H., and Kinzelbach, W., 2008, Real ‐ time groundwater flow modeling with the ensemble Kalman filter: Joint estimation of states and parameters and the filter inbreeding problem: Water Resources Research, v. 44, no. 9.
Hester, E. T., and Doyle, M. W., 2008, In ‐ stream geomorphic structures as drivers of hyporheic exchange: Water Resources Research, v. 44, no. 3.
Hogan, J., Phillips, F. M., and Scanlon, B. R., 2004, Groundwater recharge in a desert environment: the Southwestern United States, American Geophysical Union.
Hu, L., and Chugunova, T., 2008, Multiple-point geostatistics for modeling subsurface heterogeneity: A comprehensive review: Water Resources Research, v. 44, no. 11, p. W11413.
Irvine, D. J., Brunner, P., Franssen, H.-J. H., and Simmons, C. T., 2012, Heterogeneous or homogeneous   Implications of simplifying heterogeneous streambeds in models of losing streams: Journal of Hydrology, v. 424, p. 16-23.
Johnson, R. A., and Wichern, D. W., 2002, Applied multivariate statistical analysis, Prentice hall Upper Saddle River, NJ, v. 8.
Jones, J. A. A., 2011, Sustaining groundwater resources: a critical element in the global water crisis, Springer.
Kalbus, E., Schmidt, C., Molson, J., Reinstorf, F., and Schirmer, M., 2009, Influence of aquifer and streambed heterogeneity on the distribution of groundwater discharge: Hydrology and Earth System Sciences, v. 13, no. 1, p. 69-77.
Kasahara, T., and Wondzell, S. M., 2003, Geomorphic controls on hyporheic exchange flow in mountain streams: Water Resources Research, v. 39, no. 1.
Käser, D., and Hunkeler, D., 2016, Contribution of alluvial groundwater to the outflow of mountainous catchments: Water Resources Research.
Käser, D. H., Binley, A., Heathwaite, A. L., and Krause, S., 2009, Spatio ‐ temporal variations of hyporheic flow in a riffle ‐ step ‐ pool sequence: Hydrological Processes, v. 23, no. 15, p. 2138-2149.
Kerrou, J., Renard, P., Franssen, H.-J. H., and Lunati, I., 2008, Issues in characterizing heterogeneity and connectivity in non-multiGaussian media: Advances in Water Resources, v. 31, no. 1, p. 147-159.
Khodabakhshi, M., and Jafarpour, B., 2013, A Bayesian mixture-modeling approach for flow-conditioned multiple-point statistical facies simulation from uncertain training images: Water Resources Research, v. 49, no. 1, p. 328-342.
Knapton, A., 2009, Gulf water study: An integrated surface-groundwater model of the Roper River catchment, Northern Territory. Department of Natural Resources, Environment, The Arts and Sports: Water Resources Branch, NT, Australia.
Kollet, S. J., and Maxwell, R. M., 2006, Integrated surface–groundwater flow modeling: A free-surface overland flow boundary condition in a parallel groundwater flow model: Advances in Water Resources, v. 29, no. 7, p. 945-958.

Krause, S., Bronstert, A., and Zehe, E., 2007, Groundwater–surface water interactions in a North German lowland floodplain–implications for the river discharge dynamics and riparian water balance: Journal of hydrology, v. 347, no. 3, p. 404-417.

Kurtz, W., He, G., Kollet, S. J., Maxwell, R. M., Vereecken, H., and Franssen, H.-J. H., 2016, TerrSysMP-PDAF version 1.0): a modular high-performance data assimilation framework for an integrated land surface-subsurface model: Geoscientific Model Development, v. 9, no. 4, p. 1341-1360.

Kurtz, W., Hendricks Franssen, H.-J., Brunner, P., and Vereecken, H., 2013, Is high-resolution inverse characterization of heterogeneous river bed hydraulic conductivities needed and possible : Hydrology and Earth System Sciences, v. 17, no. 10, p. 3795-3813.

Kurtz, W., Hendricks Franssen, H. J., Kaiser, H. P., and Vereecken, H., 2014, Joint assimilation of piezometric heads and groundwater temperatures for improved modeling of river - aquifer interactions: Water Resources Research, v. 50, no. 2, p. 1665-1688.

Kurtz, W., Hendricks Franssen, H. J., and Vereecken, H., 2012, Identification of time - variant river bed properties with the ensemble Kalman filter: Water Resources Research, v. 48, no. 10.

Kurtz, W., Lapin, A., Schilling, O. S., Tang, Q., Schiller, E., Braun, T., Hunkeler, D., Vereecken, H., Sudicky, E., Kropf, P., Hendricks Franssen, H.-J., and Brunner, P., 2017, Integrating hydrological modelling, data assimilation and cloud computing for real-time management of water resources: Environmental Modelling & Software, v. 93, p. 418-435.

Lackey, G., Neupauer, R. M., and Pitlick, J., 2015, Effects of streambed conductance on stream depletion: Water, v. 7, no. 1, p. 271-287.

Leek, R., Wu, J. Q., Wang, L., Hanrahan, T. P., Barbet, M. E., and Qiu, H. X., 2009, Heterogeneous characteristics of streambed saturated hydraulic conductivity of the Touchet River, south eastern Washington, USA: Hydrological Processes, v. 23, no. 8, p. 1236-1246.

Li, L., Zhou, H., Hendricks Franssen, H., and Gómez-Hernández, J. J., 2012, Groundwater flow inverse modeling in non-MultiGaussian media: performance assessment of the normal-score Ensemble Kalman Filter: Hydrology and Earth System Sciences, v. 16, no. 2, p. 573-590.

Li, Q., Unger, A., Sudicky, E., Kassenaar, D., Wexler, E., and Shikaze, S., 2008, Simulating the multi-seasonal response of a large-scale watershed with a 3D physically-based hydrologic model: Journal of hydrology, v. 357, no. 3, p. 317-336.

Li, Y. J., Kokkinaki, A., Darve, E. F., and Kitanidis, P. K., 2017, Smoothing-based compressed state Kalman filter for joint state-parameter estimation: Applications in reservoir characterization and CO2 storage monitoring: Water Resources Research, p. n/a-n/a.

Liggett, J. E., Werner, A. D., and Simmons, C. T., 2012, Influence of the first-order exchange coefficient on simulation of coupled surface–subsurface flow: Journal of hydrology, v. 414, p. 503-515.

Mariethoz, G., Renard, P., and Straubhaar, J., 2010, The Direct Sampling method to perform multiple - point geostatistical simulations: Water Resources Research, v. 46, no. 11.

McCallum, A. M., Andersen, M. S., Rau, G. C., Larsen, J. R., and Acworth, R. I., 2014, River- aquifer interactions in a semiarid environment investigated using point and reach measurements: Water Resources Research, v. 50, no. 4, p. 2815-2829.

McDonald, M. G., and Harbaugh, A. W., 1988, A modular three-dimensional finite-difference ground-water flow model.

Min, L., Yu, J., Liu, C., Zhu, J., and Wang, P., 2013, The spatial variability of streambed vertical hydraulic conductivity in an intermittent river, northwestern China: Environmental earth sciences, v. 69, no. 3, p. 873-883.

Mutiti, S., and Levy, J., 2010, Using temperature modeling to investigate the temporal variability of riverbed hydraulic conductivity during storm events: Journal of Hydrology, v. 388, no. 3, p. 321-334.

Nowinski, J. D., Cardenas, M. B., and Lightbody, A. F., 2011, Evolution of hydraulic conductivity in the floodplain of a meandering river due to hyporheic transport of fine materials: Geophysical Research Letters, v. 38, no. 1.

Panday, S., and Huyakorn, P. S., 2004, A fully coupled physically-based spatially-distributed model for evaluating surface/subsurface flow: Advances in water Resources, v. 27, no. 4, p. 361-382.

Partington, D., Brunner, P., Frei, S., Simmons, C., Werner, A., Therrien, R., Maier, H., Dandy, G., and Fleckenstein, J., 2013, Interpreting streamflow generation mechanisms from integrated surface‐subsurface flow models of a riparian wetland and catchment: Water Resources Research, v. 49, no. 9, p. 5501-5519.

Pasetto, D., Camporese, M., and Putti, M., 2012, Ensemble kalman filter versus particle filter for a physically-based coupled surface-subsurface model: Advances in Water Resources.

RamaRao, B. S., LaVenue, A. M., De Marsily, G., and Marietta, M. G., 1995, Pilot point methodology for automated calibration of an ensemble of conditionally simulated transmissivity fields: 1. Theory and computational experiments: Water Resources Research, v. 31, no. 3, p. 475-493.

Rasmussen, J., Madsen, H., Jensen, K. H., and Refsgaard, J. C., 2015, Data assimilation in integrated hydrological modeling using ensemble Kalman filtering: evaluating the effect of ensemble size and localization on filter performance: Hydrology and Earth System Sciences, v. 19, no. 7, p. 2999-3013.

Remy, N., Boucher, A., and Wu, J., 2009, Applied geostatistics with SGeMS: A user's guide, Cambridge University Press.

Rorabaugh, M., 1964, Estimating changes in bank storage and ground-water contribution to streamflow: International Association of Scientific Hydrology, v. 63, no. 1, p. 432-441.

Ryan, R. J., and Boufadel, M. C., 2006, Influence of streambed hydraulic conductivity on solute exchange with the hyporheic zone: Environmental Geology, v. 51, no. 2, p. 203-210.

Saenger, N., Kitanidis, P. K., and Street, R. L., 2005, A numerical study of surface-subsurface exchange processes at a riffle-pool pair in the Lahn River, Germany: Water Resources Research, v. 41, no. 12.

Saha, G. C., Li, J., and Thring, R. W., 2017, Understanding the Effects of Parameter Uncertainty on Temporal Dynamics of Groundwater-Surface Water Interaction: Hydrology, v. 4, no. 2, p. 28.

Salehin, M., Packman, A. I., and Paradis, M., 2004, Hyporheic exchange with heterogeneous streambeds: Laboratory experiments and modeling: Water Resources Research, v. 40, no. 11, p. 18.

Scanlon, B. R., Longuevergne, L., and Long, D., 2012, Ground referencing GRACE satellite estimates of groundwater storage changes in the California Central Valley, USA: Water Resources Research, v. 48, no. 4.

Schilling, O., Doherty, J., Kinzelbach, W., Wang, H., Yang, P., and Brunner, P., 2014, Using tree ring data as a proxy for transpiration to reduce predictive uncertainty of a model simulating groundwater–surface water–vegetation interactions: Journal of hydrology, v. 519, p. 2258-2271.

Schilling, O. S., Gerber, C., Partington, D. J., Purtschert, R., Brennwald, M. S., Kipfer, R., Hunkeler, D., and Brunner, P., 2017, Advancing physically-based flow simulations of alluvial systems through atmospheric noble gases and the novel 37Ar tracer method: Water Resources Research (under review).

Schöniger, A., Nowak, W., and Hendricks Franssen, H. J., 2012, Parameter estimation by ensemble Kalman filters with transformed data: Approach and application to hydraulic tomography: Water Resources Research, v. 48, no. 4.

Schubert, J., 2006, Experience with riverbed clogging along the Rhine River, Riverbank filtration hydrology, Springer, p. 221-242.

Sebok, E., Duque, C., Engesgaard, P., and Boegh, E., 2015, Spatial variability in streambed hydraulic conductivity of contrasting stream morphologies: channel bend and straight channel: Hydrological Processes, v. 29, no. 3, p. 458-472.

Shrestha, P., Sulis, M., Masbou, M., Kollet, S., and Simmer, C., 2014, A scale-consistent terrestrial systems modeling platform based on COSMO, CLM, and ParFlow: Monthly weather review, v. 142, no. 9, p. 3466-3483.

Simpson, S. C., and Meixner, T., 2012, Modeling effects of floods on streambed hydraulic conductivity and groundwater‐surface water interactions: Water resources research, v. 48, no. 2.

Song, J., Jiang, W., Xu, S., Zhang, G., Wang, L., Wen, M., Zhang, B., Wang, Y., and Long, Y., 2016, Heterogeneity of hydraulic conductivity and Darcian flux in the submerged streambed and adjacent exposed stream bank of the Beiluo River, northwest China: Hydrogeology Journal, v. 24, no. 8, p. 2049-2062.

Sophocleous, M., 2002, Interactions between groundwater and surface water: the state of the science: Hydrogeology journal, v. 10, no. 1, p. 52-67.

Springer, A. E., Petroutson, W. D., and Semmens, B. A., 1999, Spatial and temporal variability of hydraulic conductivity in active reattachment bars of the Colorado River, Grand Canyon: Groundwater, v. 37, no. 3, p. 338-344.

Storey, R. G., Howard, K. W. F., and Williams, D. D., 2003, Factors controlling riffle-scale hyporheic exchange flows and their seasonal changes in a gaining stream: A three-dimensional groundwater flow model: Water Resources Research, v. 39, no. 2, p. 17.

Sun, A. Y., Morris, A. P., and Mohanty, S., 2009, Sequential updating of multimodal hydrogeologic parameter fields using localization and clustering techniques: Water Resources Research, v. 45, no. 7.

Tang, Q., Kurtz, W., Brunner, P., Vereecken, H., and Franssen, H.-J. H., 2015, Characterisation of river–aquifer exchange fluxes: The role of spatial patterns of riverbed hydraulic conductivities: Journal of hydrology, v. 531, p. 111-123.

Therrien, R., McLaren, R., Sudicky, E., and Panday, S., 2010, HydroGeoSphere: a three-dimensional numerical model describing fully-integrated subsurface and surface flow and solute transport: Groundwater Simulations Group, University of Waterloo, Waterloo, ON.

Tonina, D., and Buffington, J. M., 2007, Hyporheic exchange in gravel bed rivers with pool‐riffle morphology: Laboratory experiments and three‐dimensional modeling: Water Resources Research, v. 43, no. 1.

Tonina, D., de Barros, F. P., Marzadri, A., and Bellin, A., 2016, Does streambed heterogeneity matter for hyporheic residence time distribution in sand-bedded streams : Advances in Water Resources, v. 96, p. 120-126.

Ulrich, C., Hubbard, S. S., Florsheim, J., Rosenberry, D., Borglin, S., Trotta, M., and Seymour, D., 2015, Riverbed clogging associated with a California Riverbank Filtration System: An assessment of mechanisms and monitoring approaches: Journal of Hydrology, v. 529, p. 1740-1753.

Vaessen, V., and Brentführer, R., 2014, Integration of groundwater management: into transboundary basin organizations in Africa.

Van Genuchten, M. T., 1980, A closed-form equation for predicting the hydraulic conductivity of unsaturated soils: Soil science society of America journal, v. 44, no. 5, p. 892-898.

VanderKwaak, J., and Sudicky, E., 2000, Application of a physically-based numerical model of surface and subsurface water flow and solute transport: IAHS PUBLICATION, p. 515-523.

VanderKwaak, J. E., 1999, Numerical simulation of flow and chemical transport in integrated surface-subsurface hydrologic systems.

Wen, X.-H., and Gómez-Hernández, J. J., 1998, Numerical modeling of macrodispersion in heterogeneous media: a comparison of multi-Gaussian and non-multi-Gaussian models: Journal of Contaminant Hydrology, v. 30, no. 1, p. 129-156.

Winter, T. C., Harvey, J. W., Franke, L. O., and Alley, W. M., 1998, Ground water and surface water: a single resource, DIANE Publishing Inc.

Woessner, W. W., 2000, Stream and fluvial plain ground water interactions: Rescaling hydrogeologic thought: Ground Water, v. 38, no. 3, p. 423-429.

Wroblicky, G. J., Campana, M. E., Valett, H. M., and Dahm, C. N., 1998, Seasonal variation in surface-subsurface water exchange and lateral hyporheic area of two stream-aquifer systems: Water Resources Research, v. 34, no. 3, p. 317-328.

Xi, H., Zhang, L., Feng, Q., Si, J., Chang, Z., Yu, T., and Li, J., 2015, The spatial heterogeneity of riverbed saturated permeability coefficient in the lower reaches of the Heihe River Basin, Northwest China: Hydrological Processes, v. 29, no. 23, p. 4891-4907.

Xie, X., and Zhang, D., 2013, A partitioned update scheme for state - parameter estimation of distributed hydrologic models based on the ensemble Kalman filter: Water resources research, v. 49, no. 11, p. 7350-7365.

Zhang, T., 2008, Incorporating geological conceptual models and interpretations into reservoir modeling using multiple-point geostatistics: Earth Science Frontiers, v. 15, no. 1, p. 26-35.

Zhou, H., Gómez-Hernández, J. J., Hendricks Franssen, H.-J., and Li, L., 2011, An approach to handling non-Gaussianity of parameters and state variables in ensemble Kalman filtering: Advances in Water Resources, v. 34, no. 7, p. 844-864.

Zhou, H., Li, L., Franssen, H.-J. H., and Gómez-Hernández, J. J., 2012, Pattern recognition in a bimodal aquifer using the normal-score ensemble Kalman filter: Mathematical Geosciences, v. 44, no. 2, p. 169-185.

Zhou, Y., Wenninger, J., Yang, Z., Yin, L., Huang, J., Hou, L., Wang, X., Zhang, D., and Uhlenbrook, S., 2013, Groundwater–surface water interactions, vegetation dependencies and implications for water resources management in the semi-arid Hailiutu River catchment, China–a synthesis: Hydrology and Earth System Sciences, 17 (7), 2013.

Zinn, B., and Harvey, C. F., 2003, When good statistical models of aquifer heterogeneity go bad: A comparison of flow, dispersion, and mass transfer in connected and multivariate Gaussian hydraulic conductivity fields: Water Resources Research, v. 39, no. 3.

Acknowledgements

First and foremost, I would like to express my special appreciation and thanks to my supervisor Prof. Dr. Harrie-Jan Hendricks Franssen. This PhD work cannot be done without his patient and intensive supervision, guidance, support and encouragement during the past five years. Thanks for his insightful discussions and suggestions for the simulation work and editing skills for the improvement of the three papers and this PhD thesis. His background knowledge and enthusiasm in science encourage me to pursue a scientific position in future career.

I am also very grateful to Prof. Dr. Philip Brunner for the collaboration for the Emme project. He also provides many scientific advices and insightful suggestions for this PhD study. Thanks for his support for the instrument installation, field experiments and data collection in the catchment. I also appreciate the three months stay with his group in Neuchâtel and the annual project meeting there.

I would like to appreciate Prof. Dr. Harry Vereecken for his continuous support for a scientific environment for the PhD study and the kind suggestions and feedback for the biyearly PhD report. I also would like to thank Dr. Wolfgang Kurtz for his assistance and guidance through the whole PhD period. Thanks for his enthusiasm and willingness to share the data, code and modeling experience with me, and also for the helpful discussion for solving the problems during the modeling experiments. I also thank Dr. Oliver S. Schilling to provide both data and the model set-up for the real world catchment. Thanks for those skype discussion on improving the model and paper, and also for his encouragement to be positive and happy all the time in study and life.

I would also like to acknowledge the following people and groups for their support and help during my PhD study:

Prof. Dr. Philippe Renard and Dr. Julien Straubhaar for providing the direct sampling software to implement the multiple point geostatistical simulations and the three days training courses in Neuchâtel;

Dr. Daniel Käser for providing the catchment profiles and the field campaign in the research catchment;

The computing time granted by the JARA-HPC Vergabegremium and VSR commission on the supercomputer JURECA (Krause and Thörnig, 2016) at Forschungszentrum Jülich.

Guowei He for helping implement the normal score transformation algorithm into the EnKF-SPRING code and run the direct sampling executable on the Linux system;

The IBG-3 colleagues for the great time working together in the past five years.

A very special gratitude goes out to the Chinese Scholarship Council (CSC) for providing the funding for my life and study in Germany.

Last, but not least, I would like to thank my parents for all their love and encouragement. They provide me through moral and emotional support in my life. Thanks Shuo for his understanding, patience, support and encouragement especially during the final stages of the PhD study.

Band / Volume 416
Nanocrystalline Silicon Oxide in Silicon Heterojunction Solar Cells
A. Richter (2018), 166 pp
ISBN: 978-3-95806-310-5

Band / Volume 417
Gas-to-Particle Partitioning of Major Oxidation Products from Monoterpenes and Real Plant Emissions
G. Gkatzelis (2018), xii, 128 pp
ISBN: 978-3-95806-314-3

Band / Volume 418
The Dynamics of Electrons in Linear Plasma Devices and its Impact on Plasma Surface Interaction
M. Hubeny (2018), 141 pp
ISBN: 978-3-95806-316-7

Band / Volume 419
Tomographic reconstruction of gravity wave parameters from satellite-borne airglow observations
R. Song (2018), 136 pp
ISBN: 978-3-95806-317-4

Band / Volume 420
Reduktion von Edelmetallen in der Wasserstoffelektrode bei der Polymerelektrolyt-Wasserelektrolyse
P. Paciok (2018), VII, 187 pp
ISBN: 978-3-95806-320-4

Band / Volume 421
Mechanismen des Hochtemperaturrisswachstums in einem ferritischen Stahl an Luft und in Wasserdampf
T. Fischer (2018), VIII, 216 pp
ISBN: 978-3-95806-326-6

Band / Volume 422
Selbstheilende plasmagespritzte $Mn_{1,0}Co_{1,9}Fe_{0,1}O_4$-Schutzschichten in Festoxidbrennstoffzellen
N. Grünwald (2018), x, 140 pp
ISBN: 978-3-95806-327-3

Band / Volume 423
Sonochemical Synthesis of Silicon Hydride Polymers and Silicon Nanoparticles from Liquid Silanes
A. P. Cádiz Bedini (2018), viii, 132, XVIII pp
ISBN: 978-3-95806-329-7

Schriften des Forschungszentrums Jülich
Reihe Energie & Umwelt / Energy & Environment

Band / Volume 424
Synthesis and Stability Assessment of Uranium Microparticles: Providing Reference Materials for Nuclear Verification Purposes
R. Middendorp (2018), X, 145 pp
ISBN: 978-3-95806-330-3

Band / Volume 425
Herstellung und Charakterisierung von Lithiumlanthanzirkonat-Funktionsschichten für Lithium-Festkörperbatterien
T. Reppert (2018), vii, 187 pp
ISBN: 978-3-95806-331-0

Band / Volume 426
Proton Conduction and Gas Permeation through Polymer Electrolyte Membranes during Water Electrolysis
M. Schalenbach (2018), VI, 212 pp
ISBN: 978-3-95806-333-4

Band / Volume 427
Improved characterization of root zone soil moisture by assimilating groundwater level and surface soil moisture data in an integrated terrestrial system model
H. Zhang (2018), x, 125 pp
ISBN: 978-3-95806-335-8

Band / Volume 428
Manipulation of magnetism in iron oxide nanoparticle / $BaTiO_3$ composites and low-dimensional iron oxide nanoparticle arrays
L. Wang (2018), VI, 151 pp
ISBN: 978-3-95806-338-9

Band / Volume 429
Characterization of spatial-temporal varying riverbed hydraulic conductivity and its role on the estimation of river-aquifer exchange fluxes with data assimilation
Q. Tang (2018), xv, 117 pp
ISBN: 978-3-95806-339-6

Weitere ***Schriften des Verlags im Forschungszentrum Jülich*** unter
http://wwwzb1.fz-juelich.de/verlagextern1/index.asp

Energie & Umwelt / Energy & Environment
Band / Volume 429
ISBN 978-3-95806-339-6

Mitglied der Helmholtz-Gemeinschaft

JÜLICH
Forschungszentrum