Joint-inversion of spring flow and transport signatures: a multi-purpose approach for characterization and forecast of a karst system

Alireza Kavousi[1], Thomas Reimann[2], Thomas Wöhling[3], Steffen Birk[4], Andrew J Luhmann[5], Jannes Kordilla[6], Torsten Noffz[6], Martin Sauter[6], and Rudolf Liedl[3]

[1]IGW
[2]TU Dresden
[3]Technische Universität Dresden
[4]University of Graz
[5]Wheaton College
[6]University of Göttingen

November 22, 2022

Abstract

Characterization of karst systems, especially the assessment of structure and geometry of conduits along with forecast of state-variables, are essential for groundwater quality/quantity management and implementation/rehabilitation of large-scale engineering projects in karst regions. These objectives can be fully met by utilizing process-based discrete-continuum models, such as MODFLOW-2005 CFPv2, as employed here. However, such tools should be used with the caveat of the potential non-uniqueness of results. This research focuses on the joint-inversion of discharge, water temperature, and solute concentration signatures of Freiheit Spring in Minnesota, USA, in response to a spatiotemporally small-scale hydraulic and transport experiment. Adopting the multi-model concept to address conceptual uncertainty, seven distinctive model variants were considered. Spring hydro-chemo-thermo-graphs for all variants were simultaneously simulated, employing joint-inversion by PEST. Subsequently, calibrated models were compared in terms of calibration performance, parameter uncertainties and reasonableness, as well as forecast capability. Overall, results reveal the reliability of the discrete-continuum flow and transport modeling, even at a spatiotemporally small-scale, on the order of meters and seconds. All conceptualized variants suggest almost identical conduit tracer passage sizes which are close to the flood-pulse method estimates. In addition, the significance of immobile conduit-associated-drainable storages in karst hydrodynamic modeling, which is uniquely provided in our model code, was highlighted. Moreover, it was demonstrated that the spring thermograph and hydrograph carry more information about the aquifer characteristics than the chemograph. However, this last result can be site-specific and depends on the scale of the experiment and the conceptualized variants of the respective hydrological state.

Joint-Inversion of Spring Flow and Transport Signatures: A Multi-Purpose Approach for Characterization and Forecast of a Karst System

Alireza Kavousi[1], Thomas Reimann[1], Thomas Wöhling[2], Steffen Birk[3], Andrew J. Luhmann[4], Jannes Kordilla[5], Torsten Noffz[5], Martin Sauter[5], Rudolf Liedl[1],

[1] Institute for Groundwater Management, Technische Universität Dresden, 01069 Dresden, Germany

[2] Institute for Hydrology and Meteorology, Technische Universität Dresden, 01069 Dresden, Germany

[3] Institute of Earth Sciences, NAWI Graz Geocenter, Karl-Franzens-Universität Graz, 8010 Graz, Austria

[4] Department of Geology and Environmental Science, Wheaton College, Wheaton, 60187 IL, USA

[5] Dept. Applied Geology, Georg-August-Universität Göttingen, 37077 Göttingen, Germany

Corresponding author: Alireza Kavousi (alireza.kavousi@tu-dresden.de)

Key points:

- First application of discrete-continuum models for joint-inversion of measured hydro-chemo-thermographs at the scale of meters and seconds
- Spring thermograph and hydrograph carry more information about the aquifer characteristics than the chemograph
- In contrast to the presumption of conduit volume overestimation by the flood-pulse analysis, our models give comparable estimates

Abstract

Characterization of karst systems, especially the assessment of structure and geometry of conduits along with forecast of state-variables, are essential for groundwater quality/quantity management and implementation/rehabilitation of large-scale engineering projects in karst regions. These objectives can be fully met by utilizing process-based discrete-continuum models, such as MODFLOW-2005 CFPv2, as employed here. However, such tools should be used with the caveat of the potential non-uniqueness of results. This research focuses on the joint-inversion of discharge, water temperature, and solute concentration signatures of Freiheit Spring in Minnesota, USA, in response to a spatiotemporally small-scale hydraulic and transport experiment. Adopting the multi-model concept to address conceptual uncertainty, seven distinctive model variants were considered. Spring hydro-chemo-thermo-graphs for all variants were simultaneously simulated, employing joint-inversion by PEST. Subsequently, calibrated models were compared in terms of calibration performance, parameter uncertainties and reasonableness, as well as forecast capability. Overall, results reveal the reliability of the discrete-continuum flow and transport modeling, even at a spatiotemporally small-scale, on the order of meters and seconds. All conceptualized variants suggest almost identical conduit tracer passage sizes which are close to the flood-pulse method estimates. In addition, the significance of immobile conduit-associated-drainable storages in karst hydrodynamic modeling, which is uniquely provided in our model code, was highlighted. Moreover, it was demonstrated that the spring thermograph and hydrograph carry more information about the aquifer characteristics than the chemograph. However, this last result can be site-specific and depends on the scale of the experiment and the conceptualized variants of the respective hydrological state.

1 Introduction

Karst aquifers contribute to the global groundwater abstraction with ca. 13%, supplying potable water for almost one-tenth of the world's population (Stevanović, 2019). These types of groundwater resources are characterized by a solutionally-developed hierarchical discrete conduit network of generally unknown structure and geometry, imbedded into a porous rock-matrix. The resultant multimodality of flow and storage, widely conceptualized as bimodality, or even sometimes as trimodality, of karst (Atkinson, 1985; Ford & Williams, 2007; Sauter, 1992; Worthington, 1999) leads to adverse engineering consequences for water quality/quantity management (Bakalowicz, 2011), and implementation of large hydro-structures (Milanovic, 2004), in karst regions. In other words, the spatial distribution of storage and permeability fields within a karst system is not just dramatically changing, but also generally unknown. Consequently, system state-variables of hydrological (i.e., discharge and hydraulic heads) and water physicochemical (i.e., temperature and solute concentration) signatures, in-short "aquifer signatures" or so-called "hydro-chemo-thermo-graphs," can vary extremely within the spatiotemporal domain in an unpredictable and undesirable fashion.

To meet the abovementioned challenges, a great deal of the karst literature has been devoted to direct and indirect characterization of karst systems, especially the conduits. Speleological survey is the common direct characterization method of karst conduits based on field observations. Although these kind of investigations can be very informative, in practice they are also very limited because traversable conduits (i.e., caves) are neither always present, nor always accessible, nor always representative of the active flow system (Jeannin et al., 2007). Consequently, indirect karst characterization based on the recorded time-series of aquifer signatures has been of major interest.

From a methodological perspective, the indirect characterization methods are either based on hydrogeological principles through application of statistical measures/methods (e.g., Mangin, 1975; Shuster & White, 1971) or sophisticated numerical models through adaptation of inverse theory (e.g., Borghi et al., 2016). Springs, the representative monitoring site for the global behavior of karst systems (Jeannin & Sauter, 1998) comprise the only groundwater data collection point in many cases and are commonly utilized and preferred in this context.

Indirect karst characterization methods based on the spring signatures can be further categorized into two classes as naturally-driven recharge-events and artificially-driven hydraulic/tracer-tests (modified from Geyer et al. (2013)). Flood-pulse analysis (Ashton, 1966), known also as pulse train analysis (Ford & Williams, 2007; Wilcock, 1968), is a routine recharge-event-based method for estimation of phreatic conduit volume (e.g., Ryan & Meiman, 1996), which has also been applied to artificial hydraulic tests (e.g., Luhmann et al., 2012). The original method calculates the phreatic conduit volume as the bulk water discharged between the commencement of hydrograph and chemograph changes, induced by a recharge event. The time period has slightly been modified by some researchers, accounting for water displacement from epikarst and unsaturated zone (e.g., Williams, 1983), or justifying by the system-state (e.g., Luhmann et al., 2012). Although the flood-pulse method provides an intuitive measure for karst conduit sizes, it is generally criticized that it tends to overestimate the phreatic conduit volume, because of ignoring the drained water from the matrix flow system during the period of calculation (Birk et al., 2006; Williams, 1983).

Parameter estimation by inverse application of distributed numerical models has been applied as a sophisticated indirect method for karst aquifer characterization for over two decades (e.g., Larocque et al., 2000, 1999). The method is applied to the spring signatures of either naturally- (e.g., Kavousi et al., 2020) or artificially-driven (e.g., Luhmann et al., 2012) basis. However, there are generally few inverse model applications which can characterize karst conduits. This is due to the technical limitations imposed by many models to represent karst systems as well as the data scarcity and the resulting ambiguity of the model calibration.

Several distributed numerical modeling approaches have been developed for the simulation of karst systems, as reviewed by Ghasemizadeh et al. (2012), Hartmann et al. (2014), and Kovács and Sauter (2007). Among the approaches, only the discrete-continuum one can directly employ measured system-state variables, boundary conditions, parameters and processes taking place within conduit and matrix compartments of real-world karst systems in a process-based way (Kovács & Sauter, 2007). Therefore, inverse modeling utilizing such hybrid models cannot only simulate the observed system-state variables (e.g., hydro-chemo-thermographs), but also support the system characterization.

Discrete-continuum models are composed of two compartments, i.e., discrete features of conduits embedded into a rock-matrix continuum (Király, 1998). Over the last three decades, several discrete-continuum codes (e.g., de Rooij et al., 2013; Király et al., 1995; Liedl et al., 2003; Malenica et al., 2018; Reimann et al., 2018; Shoemaker et al., 2008; Tinet et al., 2019) and discrete-continuum enabled general-purpose codes (e.g., Cornaton, 2007; Diersch, 2014; Panday et al., 2013; Therrien et al., 2010; Zimmerman, 2006) have been developed, highlighting the interest of the karst modeler community in the approach.

It has been demonstrated that the application of joint-inversion techniques for the simultaneous simulation of different groundwater flow and transport signatures can improve the inverse modeling, in terms of reducing uncertainty and non-uniqueness of estimated parameters (Bravo et al., 2002; Gailey et al., 1991; Harvey & Gorelick, 1995; Xu & Gómez-Hernández, 2016). Borghi et al. (2016) showed that this hypothesis is also supported for discrete-continuum modeling. They utilized the GROUNDWATER code for synthetic sets of karst aquifers with different conduit configurations. Considering three adjustable parameters for an idealized model, Borghi et al. (2016) revealed that the joint-inversion by means of gradient-based optimization techniques, such as those offered by PEST (model-independent Parameter ESTimation code; Doherty, 2019), are efficient and promising for karst characterization. However, few joint-inversion applications of discrete-continuum models to real-world flow and transport signatures have been reported thus far. As a matter of fact, combined utilization of hydro-chemo-thermo-graphs as a calibration target can reduce the ambiguity of the inversion, but requires more efforts for simultaneous simulation of flow and transport.

Kavousi (2015) and Kavousi et al. (2020) considered model variants of conduit network configurations and simultaneously simulated the measured long-term hydro-chemo-thermo-graphs of a large-scale karst system in Iran, employing the MODFLOW-2005 CFPv2, in-short "CFPv2." Their simulations with over ten adjustable parameters were still statistically robust, and hence serve as a proof of concept for the applicability of the process-based discrete-continuum modeling approach by CFPv2. Mohammadi et al. (2018) used the CFPv2 for joint-inversion of the measured hydrograph and tracer breakthrough curve of a karst spring in Iran. Chang et al. (2019) coupled CFPv2 with ERCH, a lumped reservoir model for epikarst, and successfully simulated spring discharge and the tracer breakthrough curve from a karst aquifer in southwest China.

Results of any inverse model application can be highly affected by the considered conceptual model. Two major approaches have been used to develop hydrogeological conceptual models (Enemark et al., 2019): (1) the consensus and (2) the multi-model approach. In the consensus approach, the current state of knowledge on the site (acquired from the available data and information) would be integrated into a single conceptual model, such that the model would be sequentially updated to a new state of knowledge in the future (see Brassington & Younger (2010) and Enemark et al. (2019)). However, in the multi-model approach, alternative plausible conceptual models are being developed and tested in parallel, at any state of knowledge of the site (see Enemark et al. (2019) and Neuman and Wierenga (2003)). The multi-model approach is not aimed to find a single best model, but rather to find an ensemble of alternative conceptual models, accounting for the fact that "the hydrogeological functioning of a system can be interpreted in different ways" (Enemark et al., 2019). This approach is especially superior when the data/information on the system is scarce (Enemark et al., 2019; Neuman & Wierenga, 2003), and hence, is adopted in this study.

This work demonstrates the joint-inversion/forecast of two artificially-driven hydraulic and tracer injection experiments with the following objectives: (1) Verify the applicability of the discrete-continuum modeling approach for simultaneous simulation of flow, water temperature, and solute concentration at a small spatiotemporal scale, focusing on potential model code improvement at that scale as a future outlook plan; (2) Reveal conceptual model uncertainty by developing and testing a series of distinctive conceptual models through the multi-model approach; (3) Discuss the quantity and significance of estimated parameters; (4) Inspect the importance of different spring flow and transport signatures in karst system characterization.

2 Material and Methods

2.1 Case Study

Freiheit Spring (MN23:A00041) is located in the karst region of Fillmore County, SE Minnesota, United States. The hydrogeology of this region has been extensively investigated by the Minnesota Geological Survey (MGS) and Department of Natural Resources (DNR) (e.g., Mossler, 2008; Runkel et al., 2003; Steenberg & Runkel, 2018; Steenberg, 2014). Moreover, Luhmann (2011) investigated the variation of physicochemical characteristics of Freiheit Spring water over a three-year period and conducted two short-term controlled recharge experiments

(Luhmann et al., 2015; Luhmann et al., 2012). Here, we focus on different aspects of joint-inversion of recorded spring signatures for the first experiment, while the second experiment will be considered to test the forecast capability of our model. In the following, we provide relevant information about the geological and hydrogeological settings of the Freiheit karst system, as well as the controlled recharge experiments.

Freiheit Spring is emerging from the Stewartville Formation (~15-meter below the formation top), at an elevation of 359.66 masl (Steenberg & Runkel, 2018). The Stewartville Formation consists of limestone and dolostone sub-horizontally overlain and underlain by the Ordovician Dubuque-Maquoketa and Prosser-Cummingsville Formations, respectively. These formations are also partially or entirely comprised of limestone and dolostone, evidently associated with different karst features, e.g., sinkholes, sinking-streams, caves, and springs (see Mossler (2008), Runkel et al. (2003), and Steenberg (2014) for the details on geological formations).

Analyzing the results of several qualitative and quantitative tracer tests, DNR (2020) delineated the groundwater springshed of many karst springs in the state of Minnesota, including the Freiheit Springshed (**Figure 1**). The areas of the Freiheit surface watershed and groundwater springshed are 0.914 and 6.507 km^2, respectively. The combined area of the watershed and springshed, which overlap in the upgradient parts of the groundwater springshed, compares favorably with that obtained based on the normalized base-flow discharge method of Quinlan and Ray (1995), as mentioned by Luhmann et al. (2012).

Flow and transport in the Freiheit karst system are evidently influenced by the effect of preferential conduit flow, based on the following evidence:

1. Development of secondary solutional porosity in the bedrock formations of the Galena Group and overlaying formations were frequently reported (see Runkel et al. (2003)).
2. Inspecting the Minnesota Karst Features Database (Green et al., 2018), five sinking streams and 28 sinkholes are located within the surface watershed and Freiheit springshed.
3. Estimated flow velocities for some sink to spring tracer tests, within the groundwater springshed, are in the range of conduit flow (i.e., on the order of km/day).

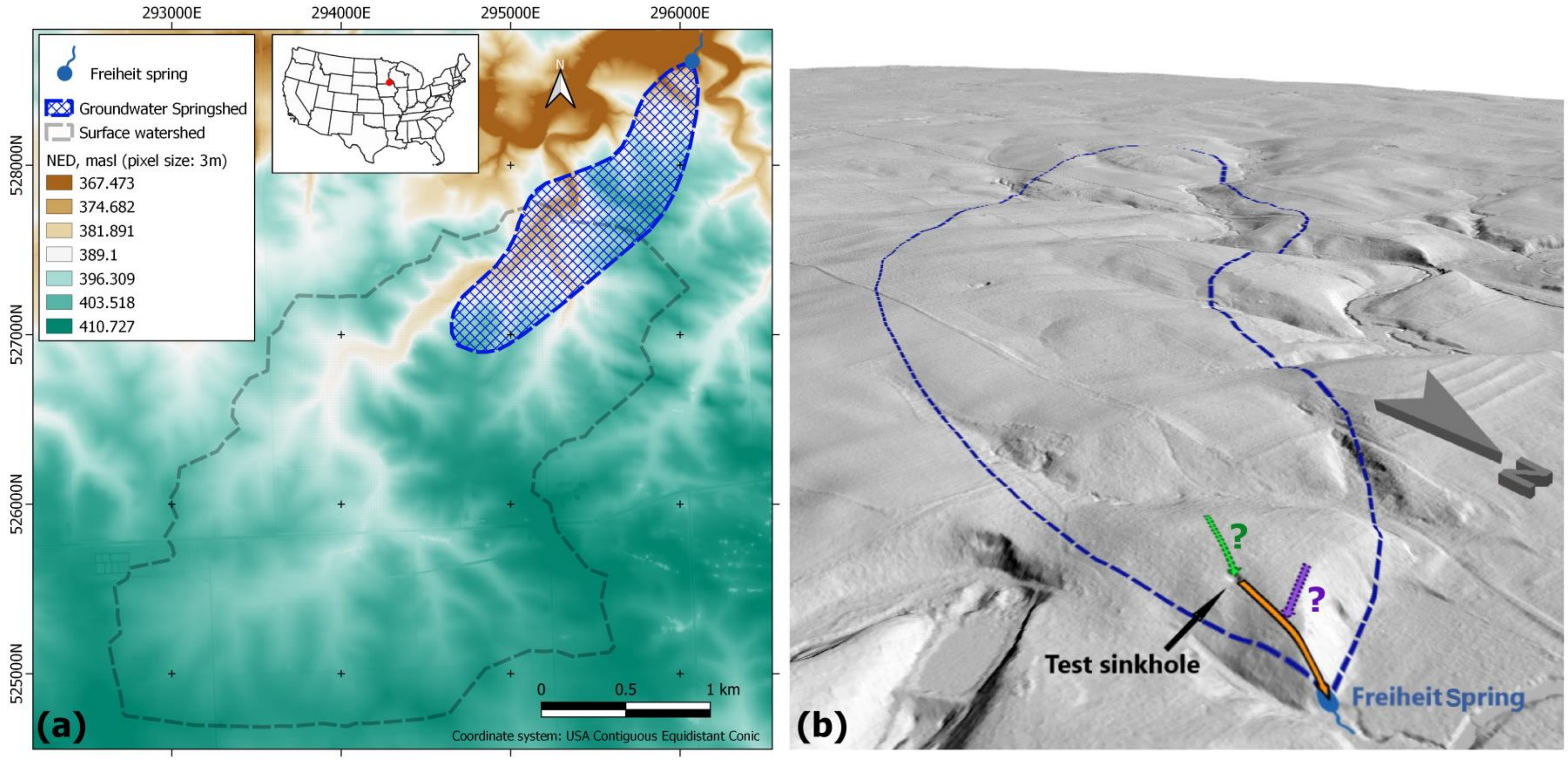

Figure 1. (a) Groundwater springshed and surface watershed of Freiheit Spring on a LiDAR-based digital terrain model of a National Elevation Dataset, NED (b) 3D hill-shade view of the aquifer terminal part, presenting the test sinkhole, spring, groundwater springshed, and considered conduits (vertical exaggeration: 1.5; camera field of view: 45˚). The orange arrow from the test sinkhole to the spring indicates the only definite conduit path, while the green and purple dotted arrows represent potential conduits located upstream and downstream of the test sinkhole, respectively.

4. Variation of the water quality and quantity of Freiheit Spring is significant and flashy, such that discharge, temperature, and specific conductivity, during the measurement period of 2008 to 2011, ranged from 10 to 385 $l \cdot s^{-1}$, 5.6 to 11.6 ºC, and 0.3 to 0.7 $mS \cdot cm^{-1}$, respectively.
5. The spring water tends to be turbid during high flows. This phenomenon could not be observed unless the opening sizes and water velocity in preferential flow paths be suitable for the suspension and transport of solid particles.

2.1.1 Experiments

Luhmann (2011) conducted two multi-tracer controlled recharge experiments at the downgradient part of the Freiheit Springshed, nearby the spring, on August 30th and September 2nd, 2010 (**Figure 1**). Water with known elevated temperature and solute concentration (including salt, uranine, and deuterated water for the first experiment and salt for the second experiment) was injected into a sinkhole, with straight horizontal and vertical distances of 95 and 19 m from the spring, respectively (**Figure 1**). The Freiheit Spring responses (including discharge, temperature, specific conductivity, and concentrations of uranine, deuterium, and suspended sediment signatures) were continuously recorded by data-loggers at a high temporal resolution (seconds) or grab samples.

In the first experiment, 13.065 ±2% m^3 of water was injected (Luhmann et al., 2012). The background spring water temperature and chloride concentration were 9.08 °C and 11.8 ppm, respectively, while the values for injected water were 24.1 °C and 1529 ppm. The main relevant outcomes of the experiment are summarized below (see Luhmann et al. (2012) for the details):

- Spring discharge started to increase shortly after injection of water into the test sinkhole and well before the other signatures, suggesting submerged full-pipe flow conditions.
- Turbidity was the next signature to rise and peak. While minor amounts of sediment were derived from the conduit flow path, most of the sediment was likely derived from the sinkhole and peaked before the other tracers at the spring due to decreasing velocity during the hydrograph recession or preferential flow along higher velocity pathways.

- Arrival of the injected water at the spring coincided with the initial rise in the conservative solutes (i.e., uranine, chloride, and deuterium), which all peaked and recovered almost identically.
- The rise and peak times of the dampened temperature signal occurred later than the same features for the conservative solute breakthrough curves.

A similar field-scale experiment was conducted three days later (Luhmann et al., 2015). This time, ~12.6 m^3 of water, at 21.5 °C temperature and with 33.02 kg of dissolved NaCl salt, was injected. However, the water was released with two almost identical pulses, separated by 25 minutes. Some rainfall occurred between the two experiments that caused more background variability in spring signatures before the second experiment, but the hydrodynamic conditions were very similar, such that the background discharges before the first and second experiments were 26.7 and 26.8 $l \cdot s^{-1}$, respectively. In general, the spring responses for the double-pulse tracer experiment were very similar to the single pulse tracer experiment three days earlier (see Luhmann et al. (2012) and Luhmann et al. (2015) for the detailed explanations).

The experiments documented the unique combination of hydraulic pressure, advection, dispersion, thermal conduction, and flow exchange processes in a real-world karst system at a spatiotemporally small-scale size that have not yet been fully simulated by a process-based model. It should be mentioned that Luhmann et al. (2012) simulated the spring chemo-thermo-graph for the first experiment, solving advective-dispersive solute and heat transport equations, as well as thermal exchange for heat transport. However, flow exchange between the conduit and surrounding matrix was neglected as the simulation was based on a sole pipe transport model, using the COMSOL Multiphysics commercial code. Moreover, flow was not simulated by Luhmann et al. (2012), but the flow velocity was required for the transport simulations and assumed as an instantaneous constant quantity throughout the whole pipe at each simulation time-step (see Luhmann et al. (2012) for the details).

2.2 Conceptual Models

Distinct conceptual models of the terminal downgradient part of the Freiheit karst system were developed, considering the multi-model concept. All the conceptual models are identical in terms of considered compartments, processes, and boundary conditions. However, they are distinct with respect to the conduit configuration (i.e., structure) and/or zonation (i.e., number of conduit sections).

2.2.1 System Compartments

We conceptualize the karst aquifer as a system with three compartments: (1) conduit, (2) matrix, and (3) conduit associated drainable storage (CADS). The matrix and conduit represent the main reserve and flow dynamic compartments, in accordance to the conceptual models of karst systems (e.g., Ford & Williams, 2007; Worthington, 1999). The CADS compartment can be assumed as a part of the conduit system. However, the water storage in CADS is of higher importance than the flow. This compartment is comparable to the “annex systems to drain, ASD” in Mangin’s conceptual model (Mangin, 1975), evidently reported for some karst aquifers (e.g., Maréchal et al., 2008; Raeisi et al., 1999). CADS reservoirs, which can be formed by solutional enlargement of fractures and cavities, are filled by almost stagnant water, but have direct association with the conduit compartment (for more detailed description of the CADS conceptualization, see Kavousi et al. (2020) and Reimann et al. (2014)).

2.2.2 Processes

Flow and transport processes within the aforementioned compartments were conceptualized based on the best functionality of the adopted numerical modeling approach and the current state of knowledge on the site. **Table 1** presents the assumed processes for different aquifer compartments. The matrix and conduits were conceptualized as the laminar and turbulent flow domains, respectively. However, the CADS is considered as an immediately responding immobile reservoir with direct association to the conduit compartment.

Considering the test sinkhole connected to a conduit path, solute and heat transport are considered only within the conduit compartment, such that the CADS and matrix surrounding the conduits can have diluting effects through exchange.

Chemical reactions are ignored for solute transport because of the more or less conservative behavior of the solute tracer employed, i.e., chloride. Therefore, it can be assumed that the tracer moves conservatively under pure advective-dispersive processes within the conduits. Temperature is considered with comparable convective-dispersive processes; however, its signal could be further affected by damping due to the heat conduction process within the matrix environment surrounding the conduit. CADS is considered as a mixing reactor with respect to both solute and heat transport processes.

Table 1. Considered flow and transport processes in different aquifer compartments.

Aquifer compartment	Processes		
	Flow dynamic	**Solute transport**	**Heat transport**
Matrix	- Laminar flow - Interaction with conduits	- Non-reactive and non-retarded	- 2D heat conduction within the conduit walls
Conduit	- Turbulent flow - Interaction with matrix - Interaction with CADS	- 1D advection - 1D dispersion	- 1D advection - 1D dispersion - 2D radial diffusion across boundary layer
CADS	- Immobile immediate-response reservoirs - Interaction with conduits	- Immobile mixing reservoir	- Immobile mixing reservoir

2.2.2.1 Ignorance of Partially-Saturated Conduit Slow and Transport

Time scales of preferential flow and transport processes in vertical shafts might be fairly ignorable in karst flow and transport modeling at regional large-scale (e.g., Kavousi et al., 2020), especially with daily or even hourly measurement frequencies. Experiments on the Freiheit karst system conducted over a short sink to spring distance of almost one-hundred meters and measured on the frequency of seconds. Therefore, the recorded delay between the pool-water release and the hydrograph rise (which was 3.25 minutes for the first experiment), as well as the associated transport signal delay can, at-least partially, be attributed to the required flow-through-time for vertical water passage to join the phreatic part of the system. Considering the elevation difference between the test sinkhole and the spring, Luhmann et al. (2012) assumed a vertical conduit infeeder, with almost nineteen-meter length, for the connection between test sinkhole and submerged conduit. Nevertheless, this study focusses on the saturated flow and transport, neglecting the partially-saturated processes in the vertical infeeder, as the water was released into a sinkhole that was assumed to have direct connection with the karst preferential flow paths, and the time period required for the recharge water to reach the phreatic conduits was estimated to be only a few minutes (Luhmann et al., 2012).

Therefore, despite the existence of a vertical infeeder beneath the test sinkhole (see Luhmann et al. (2012)), passage of water and associated solute and heat transport toward the spring were assumed to be in a phreatic condition during the whole experiment period. Thus, the conduits were supposed to be located at a sub-horizontal level, parallel to the contact surface of the Stewartville-Prosser Formations and below the spring level, such that they are always functioning at a phreatic full-pipe condition.

2.2.3 Boundary Conditions

Freiheit Spring and its predefined groundwater springshed are regarded as a fixed-head and Neumann no-flow boundary conditions, respectively (see section 0 and Figure 1). The aquifer boundary condition and spatial distribution of recharge are not known. However, when the experiments were conducted, the aquifer was draining during its recession. Therefore, recharge contribution from the surface watershed can be neglected, compared to that of the groundwater springshed. Consequently, the latter, which was recognized as the main aquifer zone (DNR, 2020), is assumed as the model domain and surrounded by no-flow boundaries (**Figure 1**). The spatiotemporally small-scale of the experiments, which are limited to the aquifer downgradient part for durations of only two hours, underpin the simplifying assumption of no-flow boundaries for the groundwater springshed. Recharge to the aquifer consists of two components: the antecedent recharge accounted for the pre-experiment background spring discharge and the injected water into the test sinkhole.

2.2.3.1 Antecedent Recharge Component

Prior to the implementation of the first and second experiments, the spring was in recession at 26.7 and 26.8 $l \cdot s^{-1}$. Spring discharge two hours after the first injection was reduced to ~25.5 $l \cdot s^{-1}$ upon the transmission of the hydraulic pulse. This slight one-liter discharge reduction was taken into account for accurate introduction of the antecedent pre-experiment recharge component for the first test. However, the spring discharge for the second experiment was almost identical, before and after the experiment, and hence, a constant antecedent recharge component is considered there.

The antecedent recharge component is further apportioned between two sub-components of distributed recharge and conduit inflows. The distributed recharge ensures initial aquifer matrix head distribution towards the conduits, draining the karst aquifer during the recession. This sub-component is defined by the estimated recharge of the region which was ~0.31 m during the year of interest (estimated from raster data provided by Smith and Westenbroek (2015)). The conduit inflows account for the recharge drained by the conduits of distant aquifer parts, which are of unknown structure and not simulated in order to reduce the overburden transport simulation time. Accordingly, the long-term distributed recharge over the groundwater springshed, i.e., the model domain, accounted for ~9.1 $l \cdot s^{-1}$ of spring discharge, while the rest was defined as the conduit inflows, i.e., a known flux boundary condition at the non-spring conduit ends.

It should be pointed out that the initial rock-matrix/water temperature and chloride concentration for the pre-experiment recharge component are presumed based on the spring background data. Accordingly, constant values of 9.08 and 9.31 ºC are considered for the initial rock-matrix/water temperature of the antecedent recharge component of the first and second experiments, respectively. Similarly, the chloride concentration of the first and second pre-experiment recharge components as well as of the corresponding rock-matrix water are set to the background values of 11.7 and 5.46 ppm, respectively (see Luhmann et al. (2012 and 2015) for more detailed description of the pre-experiment spring water characteristics).

2.2.3.2 Recharge Component of the Experiments

13.065 m^3 of water with a chloride concentration of 1529 ppm and temperature of 24.1 ºC is considered as the known flux boundary condition at the test sinkhole for the first experiment. Regarding the second experiment, 12.6 m^3 of pool-water, with an estimated chloride concentration of 1361 ppm and temperature of 21.5 °C, is assumed as two identical separate pulses, with a 25-minute delay in between.

The recharge components are approximated by uniform rectangular functions. The time period of recharge is 214 s for the first experiment, based on the field observation of the test sinkhole flooding period. However, the periods of test sinkhole flooding were not reported for the first and second pulses of the second experiment. The recharge period for these pulses is estimated based on the ratio of time-difference between hydrograph and chemograph rise-times (i.e., 625 s) to the observed test sinkhole flooding period (i.e., 214 s) for the first experiment. Accordingly, the duration of the first and second pulses of the second experiment are calculated as 172 s and 159 s, considering the corresponding time-differences between hydrograph and chemograph rise-times of 502 s and 464 s.

2.2.4 Model Variants

In order to investigate the conceptual uncertainty, seven variants of feasible conduit configuration and zonation for the test sinkhole-Freiheit Spring connection are conceptualized, considering a multi-model approach (**Figure 2**). The variants are limited to the test sinkhole and spring location, assuming inflow at the non-spring conduit ends, as discussed in section 0. The variants that can warrant all potential cases of conduit configuration of the Freiheit karst system in its terminal section are described in the following:

- Model variant I is the simplest, and has only one conduit section as the tracer passage, connecting the test sinkhole and spring (**Figure 2**). This single conduit is comparable to that of Luhmann et al. (2012). However, as it was formerly discussed, potential flow and transport interactions (i.e., between conduits, CADS, and matrix) can make a big difference between the results of our models with that of Luhmann et al. (2012).
- Model variant II is similar to model variant I, though the conduit section of the tracer passage is further split into two sections in the former (**Figure 2**). This conduit zonation would test whether likely constrictions of conduit passage or changes in aquifer characteristics (i.e., conduit parameters) can help to achieve a better result. It is worth mentioning that the addition of more conduit sections could likely result in a better model fit, which is not the main objective of this research. The conduit zonation for the tracer passage in model variant II is comparable to the zonation considered for some other model variants, as described below (**Figure 2**).
- Model variant III is comparable to model variant I, but has an additional upstream conduit section linked to the test sinkhole (**Figure 2**). This upstream conduit would allow to check for potential improvement in simulation by considering the back-flooding effect in the upstream conduits.
- Model variant IV is a combination of model variants II and III and comprises of three conduit sections (**Figure 2**).
- Model variant V considers the test sinkhole as a sole input tributary for the experiment, which is connected to a downstream lateral tributary conduit in the middle of the tracer passage (**Figure 2**). The test sinkhole is here assumed to be connected to a phreatic conduit which joins a lateral conduit in its path towards the spring, such that the upstream inflow is set to zero and all conduit inflow is attributed to the lateral mixing conduit.
- Model variants VI and VII, which are the combination of model variants II and V, and IV and V, respectively, test the contribution of inflow from both upstream and lateral tributary mixing conduits (**Figure 2**).

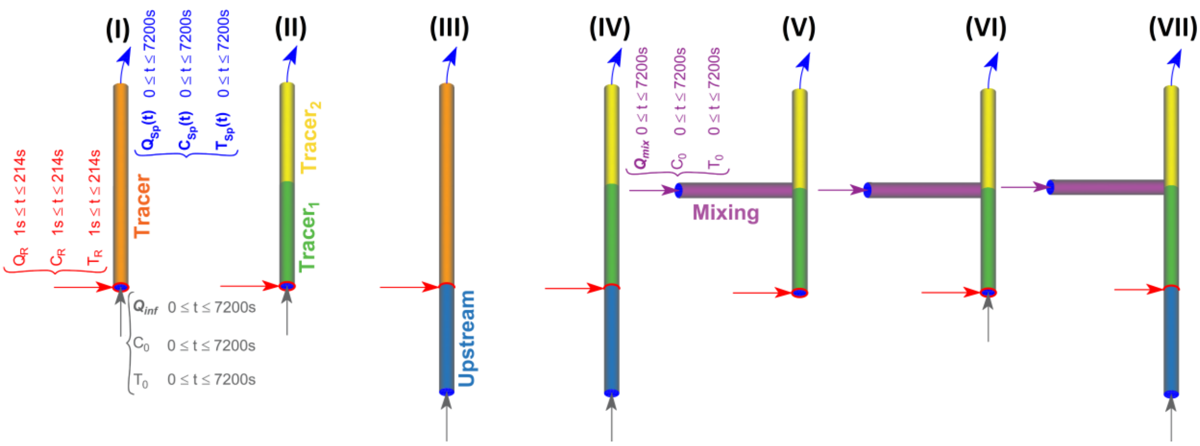

Figure 2. Conceptual models for conduit configuration and zonation. Conduit sections are indicated in colors, beside the associated names. “Tracer” is the conduit section bridging across the test sinkhole and spring, which is further divided into “Tracer$_1$” and “Tracer$_2$” in some model variants. “Mixing” and “Upstream” are the conduit sections for the lateral tributary mixing and upstream of the test sinkhole, respectively. Red, grey, purple, and blue arrows symbolize the injected water, upstream conduit inflow, mixing conduit inflow and spring outflow, respectively. Q, C and T stand for the discharge, solute concentration and water temperature, respectively. The subscript Sp, R, inf, mix, and 0 correspond to the spring, recharge for the experiment, inflow, downstream mixing, and background values, respectively. The mentioned time periods correspond to the first experiment.

Since the configuration of conduits between the test sinkhole and spring is unknown, numerous conduit configurations can be assumed and tested. However, considering a holistic overview, all possible conduit configurations can fall within the proposed seven distinctive conduit configurations. In fact, model variants I to IV would assess if the test sinkhole is located over a main conduit, assuming no further conduit junctions in the tracer passage, while model variant V assumes that the test sinkhole is associated with a minor conduit tributary, connected to the main conduit downstream. Model variants VI and VII are further testing for lateral conduit inflow contribution in the center of the tracer passage (**Figure 2**).

It might be assumed that all the defined model variants are subsets of model variant VII. For example, if the diameter of the mixing conduit section is set to zero, the model variants IV and VII are identical. However, it is emphasized that, since the conduit diameter and wall roughness are being separately estimated, the conduit diameter cannot be reduced to a diameter of zero in the course of inversion. For example, a high value of wall roughness coincides with a low value of conduit diameter (i.e., a ratio of > ~3:1) causes model failure due to numerical instabilities. Nevertheless, only model variants II and IV which are chosen to investigate potential conduit constriction can be assumed as the subsets of variants “I” and “III,” respectively.

2.3 Numerical Model

2.3.1 Simulation Code: CFPv2

MODFLOW-2005 CFPv2, in-short "CFPv2," is employed for the simulations. Here we give a brief explanation of the code. Further details can be found in Birk (2002), Harbaugh (2005), Liedl et al. (2003), Reimann et al. (2014, 2018), and Shoemaker et al. (2008).

CFPv2 is the research version of MODFLOW-2005 CFP-M1, enhanced by addition of solute and heat transport subroutines originating from the Carbonate Aquifer Void Evolution, CAVE code. The code considers a modified representation of the karst system by associating a new compartment for immobile conduit reservoirs, as well as further improvements in computational schemes (Reimann et al., 2018). CFPv2 has been successfully tested with analytical solutions for some idealized cases (Reimann et al., 2018) and adopted for modeling of flow and transport in real-world karst systems (Chang et al., 2019; Hu & Xu, 2016; Karay & Hajnal, 2015; Kavousi et al., 2020; Mohammadi et al., 2018; Sullivan et al., 2019; Xu, Hu, Davis, & Cao, 2015; Xu, Hu, Davis, & Kish, 2015).

2.3.1.1 Flow Modeling

CFPv2 utilizes MODFLOW-2005 for three-dimensional simulation of laminar flow in the matrix continuum (Harbaugh, 2005). Neighboring matrix nodes can be associated with conduit nodes, modeled as cylindrical pipes. Flow in the conduits can occur under laminar or turbulent flow conditions (Shoemaker et al., 2008), whereas in this work it is always in a turbulent state due to the high flow velocities observed in the experiments.

Computation of conduit flows and exchanges with associated matrix and CADS depends on the matrix and conduit heads, which is solved by Newton-Raphson iterations satisfying Kirchhoff's law for discharge at conduit nodes (Reimann et al., 2014; Shoemaker et al., 2008). The coupling of matrix continuum/CADS to the embedded conduit network is crucial for CFPv2 and allows rigorous numerical realizations of karst groundwater systems. The coupling is achieved by a head-dependent exchange flow term between matrix and associated conduits, while the CADS is immediately exchanging flow with the associated conduits, such that the heads of CADS and associated conduits are always the same.

2.3.1.2 Solute Transport Modeling

Considering the solute conservative behavior and flow velocities, the solute transport module, STM, Package of CFPv2, is designated to model one-dimensional advective-dispersive transport inside conduits under turbulent flow condition. CADS at the associated conduit nodes is considered as a simple mixing reactor, where solute exchange is coupled with the flow (see Reimann et al. (2018) for the details of processes implemented in the CFPv2).

2.3.1.3 Heat Transport Modeling

The heat transport module, HTM, Package of CFPv2, accounts for convective-dispersive heat transport in conduits, as well as conductive heat transfer between the conduit and the surrounding rock-matrix through a thermal boundary layer, as well as within the rock-matrix (Birk, 2002; Birk et al., 2006). Likewise in STM, CADS is considered as a simple mixing reactor for heat exchange in HTM (Reimann et al., 2018). Details of heat (as well as solute) transport numerical schemes can be found in Birk (2002) and Reimann et al. (2018).

2.3.2 Numerical Model Representation

The modeling area is discretized by 100 m^2 (i.e., 10 m × 10 m) cells in two layers, based on the 3D data on geological formations (i.e., Steenberg, 2014). The calibration period is temporally discretized by three stress-periods, where the first one is considered as steady-state. This stress-period is required to reproduce a matrix head for the remaining modeling period, justified by the given reasons in section 0. The steady-state stress-period is followed by two transient stress-periods of 214 seconds for the injection and 6986 seconds for the pulse transmission. Gradual increase of time-step length is considered to increase the accuracy of calculations, while reducing the computation time. In total, 287 time-steps are considered for the whole calibration period.

Table 2 gives the list of adjustable parameters, beside their relevant aquifer compartments and relevant input-files (see Reimann et al. (2018) and Harbaugh (2005) for detailed description of the input-files). Overall, there are 7 conduit, 2 CADS, and 3 matrix adjustable parameters.

All model input parameters are considered to be homogeneously distributed across the domain, except for the first five parameters in Table 2, which were imposed by zonation based on the conduit sections for the model variants II to VII. CADS is assumed to be present in all conduit sections across all nodes, such that their width can be optimized. Moreover, it should be noted that the conduit inflows are only estimated for the model variants VI and VII, where the total inflow can be shared between the mixing and upstream inflows. In other words, the allocation ratio between conduit inflows in the variants VI and VII is not fixed but computed by the model. The conduit upstream or mixing inflows are not estimated and treated as fixed for the other model variants, because only one such parameter was introduced there which can deliver all conduit inflow. Consequently, the number of adjustable parameters for the model variants I to VII are successively 10, 15, 15, 20, 20, 22, and 27.

Table 2. List of model adjustable parameters, including their notations, units, relevant aquifer components and relevant input-files.

No.	**Parameter**	**symbol†**	**Unit**	**Aquifer component**	**MODFLOW-2005 / CFPv2 input-file**
1	Conduit diameter	D_c	m	Conduit	CFP
2	Conduit tortuosity	τ_c	-	Conduit	CFP
3	Conduit wall rough.	k_c	m	Conduit	CFP
4	Exchange coef.	K_{ex}	$m \cdot s^{-1}$	Conduit	CFP
5	CADS width	W_{CADS}	m	CADS	CFP
6	Conduit upstr. inflow	Q_{inf}	$m^3 \cdot s^{-1}$	Conduit	CRCH
7	Conduit mixing inflow	Q_{mix}	$m^3 \cdot s^{-1}$	Conduit	CRCH
8	Conduit recharge	CRCH	-	Conduit	CRCH
9	CADS recharge	CADS-RCH	-	CADS	CRCH
10	Horizontal hydr. cond.	K_h	$m \cdot s^{-1}$	Matrix	LPF
11	Rock specific heat	$c_{p,r}$	$J \cdot kg^{-1} \cdot K^{-1}$	Matrix	HTM
12	Rock thermal cond.	k_r	$W \cdot m^{-1} \cdot K^{-1}$	Matrix	HTM

† Parameters with further zonation would have slightly modified symbols, which would be indicated in the text.

All matrix layers are modeled with similar parameters and comprise a homogeneous but anisotropic matrix domain with a K_x:K_z anisotropy ratio of 10:1 (Runkel, 2020, personal communication). Concentrated recharge to the test sinkhole node can be either diverted to the model as the CRCH (conduit recharge) or CADS-RCH (recharge to the CADS), such that experiment recharge would be directed into these two components. Since concentrated recharge is defined through the RCH MODFLOW package, prior information was taken into account to accomplish this. Therefore, the model can alternatively transmit some portion of the experiment recharge to the matrix system too, if this matrix recharge is supported by the inversion.

The amount, temperature, and concentration of recharge components are assigned to the relevant CFPv2 input-files according to the values in section 0.

2.3.3 Calibration Procedure

Inverse problems in hydrogeology are typically solved by history matching, i.e., fitting model-simulated to field-observed signatures, in both steady- and transient-state simulations, which results in a set of parameter estimates that produce a satisfactory level of match (Anderson et al., 2015).

Field observations are scarce for history matching of many environmental systems, and it is often necessary to integrate all measured and simulated data into a single objective function for simultaneous inversion (Hill & Tiedeman, 2007). For this purpose, weighting of different signatures within a single objective function of least-squares is very efficient. The weighted least-squares objective function, *Φ*, is the sum of squared-weighted-residuals, i.e., differences between measured and simulated signatures (Neuman, 1973):

$$\Phi = \sum_{i=1}^{n}\left(\sum_{j=1}^{m}\left[w_{i,j}\left(M_{i,j} - S_{i,j}\right)\right]^2\right) = \sum_{i=1}^{n}\left(\sum_{j=1}^{m}\left(w_{i,j}r_{i,j}\right)^2\right) \qquad (1)$$

where *n* and *m* refer to the number of signature groups and individual signatures, respectively; $w_{i,j}$ is the weight assigned to the *j*-th signature of *i*-th signature group; $M_{i,j}$ is the *j*-th measured signature of *i*-th signature group; $S_{i,j}$ is the simulated signature, equivalent for $M_{i,j}$; and $r_{i,j}$ is the residual of difference between $M_{i,j}$ and $S_{i,j}$. There are three time-series of measured signatures for Freiheit Spring, i.e., spring discharge (Q), solute concentration (C), and water temperature (T), which are included in the weighted least-squares objective function.

Weights in an objective function are critical for a simultaneous calibration (i.e., joint-inversion), as they represent the degree of certainty that is assigned to different measurements. In an ideal statistical context, any individual weight should directly express the measurement error of the associated observation, though the ideal rarely holds when models are applied in practice (Anderson et al., 2015). Here the weights are assigned as inversely proportional to the standard deviation of measured signatures, such that equal weights are considered within each group (according to the guidelines by Doherty (2019)). Accordingly, w_Q, w_C, and w_T are 371.116 s^1m^{-3}, 6.393 m^3kg^{-1}, and 1.52 $^{\circ}C^{-1}$, respectively. Minimization of the weighted least-squares objective function is implemented using PEST (Doherty, 2019), in the parameter estimation mode.

Matrix and conduit heads are not recorded within the Freiheit karst system. However, considering the spring type, the outlet elevation (i.e., 359.66 masl) is assumed as the matrix head for the corresponding spring node at the first steady-state stress period. This assumption would constrain the inverse problem to stick with realistic heads. This head value beside the prior information designated for realistic concentrated recharge estimation and inflow allocation (for model variants VI and VII) are normally engaged as a part of the weighted least-squares objective function (see Doherty (2019)).

2.3.4 Performance Assessment

The extent to which model outputs (i.e., simulated signatures) are in agreement with their pertinent field records (i.e., measured signatures) is detectible from the value of the objective function (Doherty, 2019). However, some other statistical measures of goodness of fit have usually been adopted for this reason (see Anderson et al. (2015); Knoben et al. (2019); Pushpalatha et al. (2012); Wöhling et al. (2013); Zheng and Bennett (2002)).

We consider three criteria to assess the fitting statistics: (1) root mean square error, RMSE which is in the scale of measured values; (2) Nash-Sutcliffe efficiency, NSE (Nash & Sutcliffe, 1970), which is widely used as a normalized metric for model performance assessment in hydrological simulations; and (3) Kling-Gupta efficiency, KGE (Gupta et al., 2009), which is another commonly used normalized metric for model performance assessment. As a rule of thumb, the smaller the RMSE and the higher the KGE or NSE values (i.e., the more close to the unity), the better the model performance. Unlike the NSE, which ranges from zero to one, the KGE can reach negative values, such that -0.41 was established as the mathematical threshold between the "good" and "bad" model performances (Knoben et al., 2019).

2.3.5 Parameter Uncertainty Reduction

Observations considered in an inverse-problem would contribute to the reduction of prior uncertainties at which the parameters are being estimated. Relative parameter uncertainty variance reduction, RUV, is a powerful measure to compare the degree of parameter uncertainty reduction, given a set of observations (Doherty, 2015):

$$RUV_I = 1 - \frac{\sigma_{uI}^2}{\sigma_I^2} \qquad (2)$$

where σ_{uI}^2 and σ_I^2 are the prior and posterior uncertainty variances for the *I*-th parameter, respectively. RUV_I ranges from zero to one, such that the higher the value, the more the reduction of parameter uncertainty (Doherty, 2015). RUV_I is considered as a more robust method in comparison to the other similar statistics for parameter comparison, such as parameter sensitivity, identifiability, and relative error reduction (see Doherty and Hunt (2009) and Doherty (2015)).

3 Results and Discussions

3.1 Model Calibration

All model variants of Freiheit Spring are calibrated to the observations from the first experiment using PEST (Doherty, 2019). **Figure 3** presents the measured and simulated hydro-chemo-thermo-graphs for the model variant II as a sample fit. The graph is normalized to the range, such that each measured and simulated signature is subtracted by the corresponding measured pre-experiment background, then divided by the relevant measured peak value (**Figure 3**).

It should be noticed that normalization to the range tends to exaggerate the discrepancies between measured and simulated signatures, e.g., the maximum discrepancy between measured and simulated discharge (as the weakest modeled signature) is smaller than 4.6%, which is still less than the half of discharge measurement error of ±10%, as reported by Luhmann et al. (2012). Therefore, **Figure 3** reveals that the CFPv2 could fairly reproduce the signatures observed at Freiheit Spring through joint-inversion. **Table 3** provides the model calibration performance criteria for all calibrated model variants, including RMSE, NSE, and KGE, beside the weighted least-squares objective function from PEST (i.e., Φ), which combines all observations and prior information into a single performance criterion. Model variants VII, IV, VI, and II achieve better Φ performance in comparison to the rest. Among these variants, variant II is simpler as it has only 15 adjustable parameters. However, all model variants are still acceptable in terms of the fitting statistics. Results indicate that the model performance can be improved by further zonation of the tracer conduit section (compare model variants I and III with the other variants, cf. **Table 3** and **Figure 2**). Moreover, the fitting statistics for the model variant V, which has conduit inflow solely as downstream mixing, is the worst among the model variants with downstream mixing inflow (i.e., V, VI, and VII).

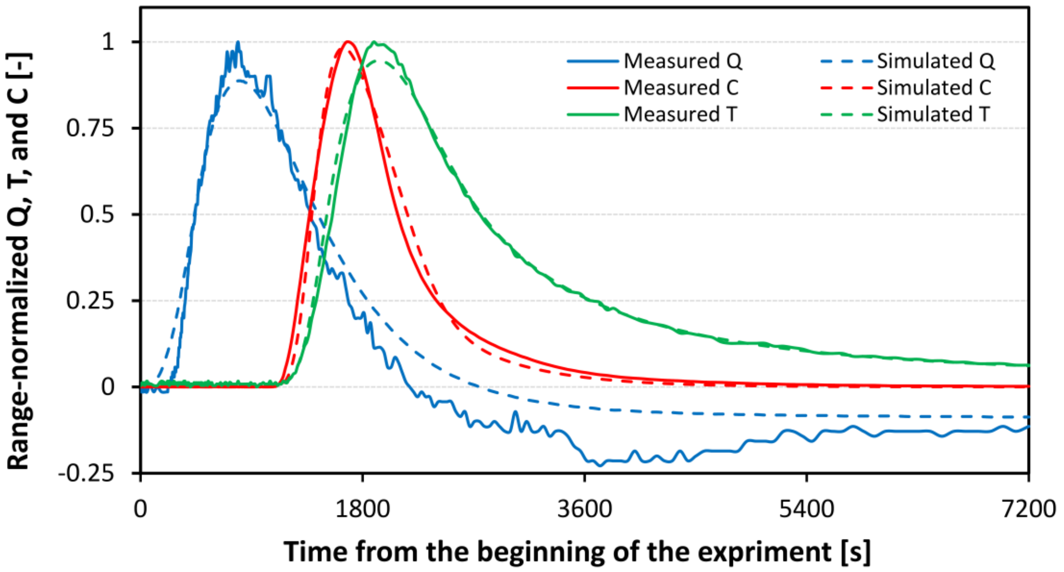

Figure 3. Range-normalized measured and calibrated hydro-chemo-thermo-graphs (i.e., Q, C, and T signatures) of the first experiment, for the model variant II. Note that the normalization is performed based on the ranges of measured values, such that the pre-experiment background and peak values reach the values of zero and one, respectively.

Table 3. Weighted least-squares objective function of PEST (Φ), and calibration performance criteria (i.e., RMSE, NSE, and KGE) of different signatures across all model variants (ranked from the best to the worst according to the Φ value), beside the number of adjustable parameters. Note that the red to green color-scale for the NSE and KGE criteria have the same scale, such that the darker the shade of green in the cell, the closer the value is to unity, i.e., the better the fitting statistic. However, the red to green color-scale for the RMSE and Φ can only be compared vertically (i.e., within each signature group), such that in each of the columns of RMSE or Φ criteria, greener cells correspond to better fitting statistics.

Rank	Model variant	No. of adj. parameters	Φ [-]	$RMSE_Q$ [l·s^{-1}]	$RMSE_C$ [ppm]	$RMSE_T$ [°C]	NSE_Q [-]	NSE_C [-]	NSE_T [-]	KGE_Q [-]	KGE_C [-]	KGE_T [-]
1	VII	27	14.49	0.497	0.015	0.039	0.966	0.991	0.996	0.893	0.963	0.995
2	IV	20	15.67	0.553	0.013	0.048	0.958	0.993	0.995	0.874	0.979	0.996
3	VI	22	16.52	0.528	0.016	0.039	0.962	0.989	0.996	0.876	0.974	0.997
4	II	15	18.20	0.584	0.016	0.045	0.953	0.989	0.995	0.868	0.971	0.997
5	V	20	32.51	0.600	0.032	0.096	0.950	0.958	0.979	0.853	0.877	0.968
6	I	10	43.89	0.908	0.025	0.059	0.886	0.974	0.992	0.759	0.895	0.995
7	III	15	43.95	0.965	0.022	0.032	0.871	0.980	0.998	0.733	0.956	0.995

According to all criteria, fitting statistics for the discharge signature (i.e., hydrograph) are somewhat weaker than the transport ones (cf. **Figure 3** and the NSE and KGE values in **Table 3**). However, the discrepancy between measured and simulated discharges is still smaller than the measurement error and can be attributed to the following reasons:

1. The beginning of measured hydrograph rise had a 3.25 minute delay relative to the start of the experiment, which is attributed to the required time for recharge water to

reach the phreatic conduits. Then the hydraulic pulse was assumed to propagate along the flow path at the speed of sound, under submerged conditions (Luhmann et al., 2012). Since the partially-saturated vertical conduit flow is ignored in our simulations, the delay for the simulated discharge is inevitably shorter (see section 0).

2. Small oscillations in the measured hydrograph likely resulted from the measurement method, which was comprised of a 120° v-notch weir and pressure transducer data-logger (see Luhmann (2011) for the details).
3. An unusual discharge drop from the pre-experiment level (at ~2,200 s) and recovery (since ~4,500 s to the end of simulation) can be observed. Luhmann et al. (2012) attributed this odd flow behavior to syphoning, flow inertia, or hysteresis effects associated with the transition of some portion of conduits from full-pipe to open-channel flow. These flow processes are neither distinguishable at the current state of knowledge of the site nor covered by the available model tools.

It is again emphasized that the reported ±10% error in the discharge measurements (Luhmann et al., 2012) is still greater than the maximum discrepancy between measured and simulated discharge in all model variants. Therefore, one may consider that the relatively weaker fit for the hydrograph is mainly stemmed from the associated measurement error.

Moreover, one may be able to achieve better fitting statistics by considering additional zonation of conduits adopting a highly regularized inversion, e.g., by Tikhonov regularization (Doherty, 2019), if a perfect fit is the desired modeling outcome. However, we only consider two sections, as this yields a sufficiently fine joint-inversion of all signatures at the cost of fewer adjustable parameters, i.e., lower degree of uncertainty.

3.2 Statistical Assessment of Estimated Parameters

Table 4 presents the average, maximum, and minimum of the estimated parameter values across different model variants. All estimated parameter values are within their reported feasible ranges, except for the tortuosity and rock thermal parameters, which will be discussed in the relevant following sections.

In order to have a favorable comparison of calibrated parameters in all model variants, each estimated value and their 95% confidence intervals are plotted in a separate plot on a normalized scale (**Figure 4**). The normalized estimated parameter value of each variant is calculated as one minus the ratio of the actual estimated parameter value to the corresponding average value across all model variants. Likewise, normalizations of 95% upper and lower confidence limits are performed by adding the relative deviation of the respective confidence limits to the normalized parameter values (**Figure 4**). Adopting this normalization scheme, the variation of normalized parameter estimates highlights the deviation of estimated values across the model variants, while the size of confidence intervals are mainly applicable to compare the certainty at which the parameters were estimated by PEST in each model variant (**Figure 4**). It's worth mentioning that the PEST calculated 95% confidence limits are based on the linear approximation of the posterior covariance matrix (see Doherty (2019)).

Comparing the normalized values for segmented parameters, except for the conduit tortuosity, the estimated values exhibit a narrower uncertainty interval for the parameters corresponding to the tracer conduit sections, in comparison to those of the upstream and mixing ones (**Figure 4**). Specifically, diameters of conduit sections for the tracer passage (i.e., D_c, D_{c1}, and D_{c2}) have the most constant and certain estimates across all model variants.

Estimated parameters for the model variants I and II, which have neither upstream nor mixing conduit sections, are associated with the highest certainties across all model variants, on average. This observation suggests that the conducted experiment was more appropriate to estimate parameters for the tracer conduit passage (especially the conduit diameter) than those of the upstream or mixing ones.

Table 4. Average, minimum, and maximum of estimated parameter values across different model variants.

No.	Parameter	Conduit section	Parameter symbol†	Unit	Average estimate	Min. estimate	Max. estimate	Model Variant
1	Conduit diameter	Tracer	D_c	m	0.365	0.363	0.367	I, III
2		Tracer1	D_{c1}	m	0.332	0.300	0.392	II, IV, V, VI, VII
3		Tracer2	D_{c2}	m	0.416	0.325	0.502	II, IV, V, VI, VII
4		Mixing	$D_{c,\ Mixing}$	m	0.922	0.561	1.392	V, VI, VII
5		Upstr.	$D_{c,\ Upstr.}$	m	1.823	1.546	2.000	III, IV, VII
6	Conduit tortuosity	Tracer	τ_c	-	5.336	5.266	5.406	I, III
7		Tracer1	τ_{c1}	-	5.879	1.164	8.199	II, IV, V, VI, VII
8		Tracer2	τ_{c2}	-	4.347	2.108	8.408	II, IV, V, VI, VII
9		Mixing	$\tau_{c,\ Mixing}$	-	1.064	1.000	1.191	V, VI, VII
10		Upstr.	$\tau_{c,\ Upstr.}$	-	1.118	1.000	1.208	III, IV, VII
11	Conduit wall rough.	Tracer	k_c	m	0.684	0.645	0.724	I, III
12		Tracer1	k_{c1}	m	0.717	0.375	1.000	II, IV, V, VI, VII
13		Tracer2	k_{c2}	m	0.605	0.185	1.000	II, IV, V, VI, VII
14		Mixing	$k_{c,\ Mixing}$	m	0.629	0.559	0.768	V, VI, VII
15		Upstr.	$k_{c,\ Upstr.}$	m	0.805	0.616	0.998	III, IV, VII
16	Exchange coef.	Tracer	K_{ex}	$m \cdot s^{-1}$	1.31E-04	1.27E-04	1.35E-04	I, III
17		Tracer1	K_{ex1}	$m \cdot s^{-1}$	9.98E-05	4.39E-05	1.60E-04	II, IV, V, VI, VII
18		Tracer2	K_{ex2}	$m \cdot s^{-1}$	2.09E-04	1.83E-04	2.42E-04	II, IV, V, VI, VII
19		Mixing	$K_{ex,\ Mixing}$	$m \cdot s^{-1}$	1.28E-05	8.34E-06	1.72E-05	V, VI, VII
20		Upstr.	$K_{ex,\ Upstr.}$	$m \cdot s^{-1}$	5.16E-05	1.45E-05	7.74E-05	III, IV, VII
21	CADS width	Tracer	W_{CADS}	m	7.39E-03	6.37E-03	8.40E-03	I, III
22		Tracer1	W_{CADS1}	m	4.45E-03	3.05E-03	5.38E-03	II, IV, V, VI, VII
23		Tracer2	W_{CADS2}	m	6.51E-02	1.58E-02	1.02E-01	II, IV, V, VI, VII
24		Mixing	$W_{CADS,\ Mixing}$	m	5.01E-03	1.48E-03	1.02E-02	V, VI, VII
25		Upstr.	$W_{CADS,\ Upstr.}$	m	1.05E-03	4.84E-05	2.20E-03	III, IV, VII
26	Conduit inflow	Mixing	$Inflow_{Mixing}$	$m^3 \cdot s^{-1}$	1.38E-02	1.35E-02	1.40E-02	VI, VII
27		Upstr.	$Inflow_{Upstr.}$	$m^3 \cdot s^{-1}$	3.80E-03	3.57E-03	4.03E-03	VI, VII
28	Conduit recharge	-	CRCH	-	0.689	0.568	0.819	All
29	CADS recharge	-	CADS-RCH	-	0.311	0.181	0.432	All
30	Horizon. hydr. cond.	-	K_h	$m \cdot s^{-1}$	1.87E-06	1.01E-06	3.14E-06	All
31	Rock specific heat	-	$c_{p,r}$	$J \cdot kg^{-1} \cdot K^{-1}$	5.00E+03	5.00E+03	5.00E+03	All
32	Rock thermal cond.	-	k_r	$W \cdot m^{-1} \cdot K^{-1}$	3.297	2.920	3.632	All

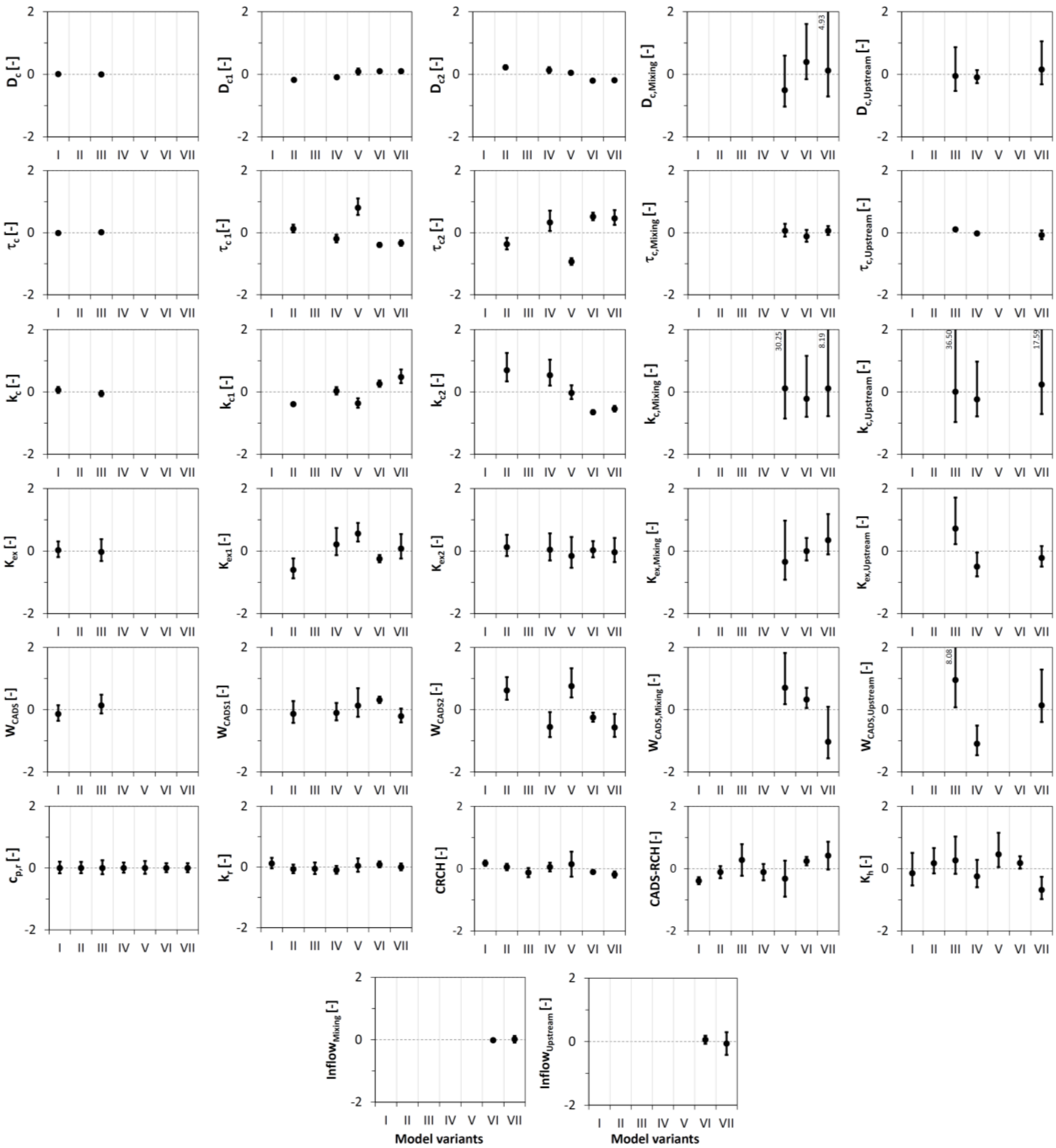

Figure 4. Normalized plots of estimated parameter values and their 95% confidence intervals for all model variants. Note that each plot corresponds to one parameter in all model variants. Deviation of the estimated parameter value from zero for each model variant highlights the deviation from the average estimated value of the parameter across all model variants, while the size of confidence intervals are mainly applicable to compare the certainty at which the parameters were estimated in each model variant.

Relative uncertainty variance reduction, of parameters RUV, is also investigated using the utilities provided for PEST post-processing (Doherty, 2019; Doherty & Hunt, 2009). For this purpose, the Jacobian matrix is recalculated for the calibrated parameter dataset excluding the prior information from the inversion. **Table 5** presents the RUV for all model variants.

Table 5. Relative parameter uncertainty variance reduction (RUV) for all model variants. Values are color-scaled based on the given legend. Rows of the table that correspond to the parameters of the tracer conduit sections are presented in boldface.

No.	Parameter	Conduit section	Parameter symbol	RUV for the model variant I	II	III	IV	V	VI	VII
1	Conduit diameter	Tracer	D_c	1.00		1.00	-	-	-	-
2		Tracer1	D_{c1}	-	0.99	-	0.99	1.00	0.99	0.99
3		Tracer2	D_{c2}	-	0.99	-	0.99	0.99	0.98	0.99
4		Mixing	$D_{c,\ Mixing}$	-	-	-	-	0.45	0.17	0.20
5		Upstream	$D_{c,\ Upstream}$	-	-	0.88	0.81	-	-	0.64
6	Conduit tortuosity	Tracer	τ_c	1.00	-	1.00	-	-	-	-
7		Tracer1	τ_{c1}	-	0.95	-	0.97	0.94	0.96	0.97
8		Tracer2	τ_{c2}	-	0.92	-	0.86	0.98	0.87	0.84
9		Mixing	$\tau_{c,\ Mixing}$	-	-	-	-	0.64	0.68	0.57
10		Upstream	$\tau_{c,\ Upstream}$	-	-	0.71	0.74	-	-	0.52
11	Conduit wall rough.	Tracer	k_c	1.00	-	1.00	-	-	-	-
12		Tracer1	k_{c1}	-	0.99	-	0.98	1.00	0.97	0.97
13		Tracer2	k_{c2}	-	0.93	-	0.93	0.96	0.98	0.98
14		Mixing	$k_{c,\ Mixing}$	-	-	-	-	0.37	0.24	0.27
15		Upstream	$k_{c,\ Upstream}$	-	-	0.38	0.32	-	-	0.31
16	Exchange coef.	Tracer	K_{ex}	0.99	-	0.98	-	-	-	-
17		Tracer1	K_{ex1}	-	0.98	-	0.94	0.97	0.97	0.91
18		Tracer2	K_{ex2}	-	0.94	-	0.89	0.97	0.90	0.91
19		Mixing	$K_{ex,\ Mixing}$	-	-	-	-	0.82	0.84	0.41
29		Upstream	$K_{ex,\ Upstream}$	-	-	0.94	0.94	-	-	0.91
21	CADS width	Tracer	W_{CADS}	0.99	-	0.99	-	-	-	-
22		Tracer1	W_{CADS1}	-	0.99	-	0.99	0.98	0.99	0.99
23		Tracer2	W_{CADS2}	-	0.98	-	0.98	0.98	0.98	0.98
24		Mixing	$W_{CADS,\ Mixing}$	-	-	-	-	0.94	0.92	0.83
25		Upstream	$W_{CADS,\ Upstream}$	-	-	0.38	0.97	-	-	0.84
26	Conduit inflow	Mixing	$Inflow_{Mixing}$	-	-	-	-	-	1.00	1.00
27		Upstream	$Inflow_{Upstream}$	-	-	-	-	-	1.00	1.00
28	Conduit rech. comp.	-	CRCH	0.96	0.93	0.94	0.93	0.84	0.93	0.93
29	CADS rech. comp.	-	CADS-RCH	0.94	0.90	0.95	0.89	0.86	0.93	0.95
30	Horizon. hydr. cond.	-	K_h	0.95	0.93	0.93	0.93	0.94	0.93	0.93
31	Rock specific heat	-	$c_{p,r}$	0.92	0.80	0.88	0.84	0.84	0.80	0.89
32	Rock thermal cond.	-	k_r	0.90	0.87	0.87	0.88	0.84	0.92	0.91

Color scale:
> 0.9
0.8 - 0.9
0.7 - 0.8
0.6 - 0.7
0.5 - 0.6
< 0.5

As it can be seen in **Table 5**, all parameters in the model variant I and II have a RUV of over 0.8, which means that all parameters are effectively sensitive for the current conceptual model and dataset. The level of uncertainty reduction is different, but still remains significant for many parameters in the other model variants. Obviously, almost all parameters from the upstream and mixing conduit sections have smaller RUV comparing to the equivalent parameters in the same model variants, and therefore, are less sensitive. However, the RUV of exchange coefficients for the upstream conduit section in the model variants IV and VII are exceptions, slightly exceeding their equivalent values for the conduit sections of the tracer passage. These results highlight that model variant II may be preferred to the others even if its calibration performance is slightly weaker.

It is worth mentioning that the matrix specific yield has also been considered as an adjustable parameter in the early model runs. However, the parameter turns out to be very insensitive in the course of calibrations of all model variants, causing occasional numerical instabilities of the inversion. Therefore, this parameter is fixed at a value of 0.05 (personal communication with Runkel, 2020) and omitted from the inversion to reduce the dimensionality of the model inversion. The insensitivity of the model to the specific yield is justifiable by the fact that the time period of the aquifer dynamic response to the experiment was only two hours, when the reserves from the conduit and CADS compartments govern the aquifer behavior more effectively than those of the matrix.

3.3 Model Testing

All calibrated model variants were considered as base models and used to reproduce the hydro-chemo-thermo-graphs of the second experiment conducted at Freiheit Spring. Comparing the NSE and KGE performance criteria values across and within model variants, the following results can be summarized (**Figure 5**):

- Model variant V has the best forecasts among the model variants.
- The best model fit to the observed signatures is for the discharges and the worst is for the solute concentrations, except for the model variant III, where the KGE of forecasted thermograph is slightly better than that of the hydrograph.
- All thermographs produced better model fits than the chemographs.

Figure 6 presents the measured and forecasted hydro-chemo-thermo-graphs for all model variants. The observed deviations between the measured and forecasted results potentially result from the uncertainties in recharge functions, because all model variants are calibrated and forecasted by rectangular recharge functions. Overall, the following deviations can be highlighted by careful inspection of the joint-forecasts for all model variants:

1. The rise and peak times are clearly better estimated for the hydrograph in comparison to the others, which obviously indicates some delayed shift of the forecasted transport spikes (**Figure 6**).
2. Forecasted peak-values are somewhat smaller than the observed ones for all signatures, in almost all model variants (**Figure 6**).
3. Like the hydrograph of the first experiment, the unaccountable effect of some flow process (e.g., due to siphoning effects) caused discrepancy between measured and forecasted discharge (esp. from the ~2500-3000 second onward; cf. **Figure 6**).

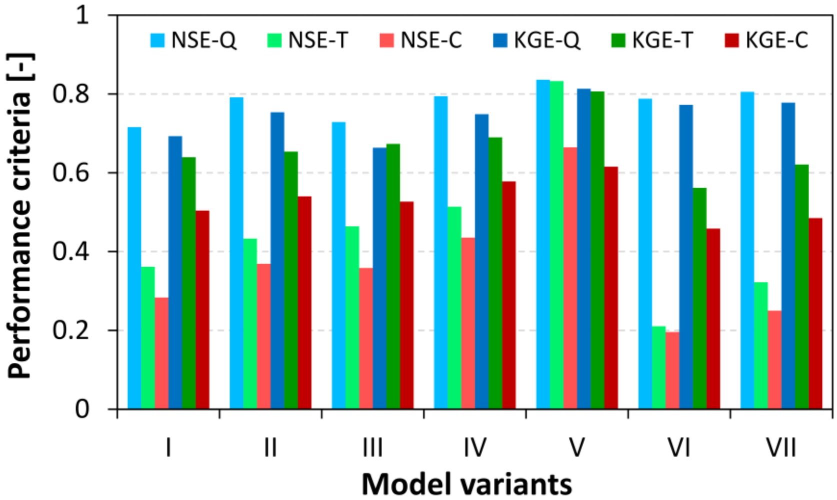

Figure 5. KGE and NSE performance criteria for the joint-forecasted hydro-chemo-thermo-graphs of all model variants.

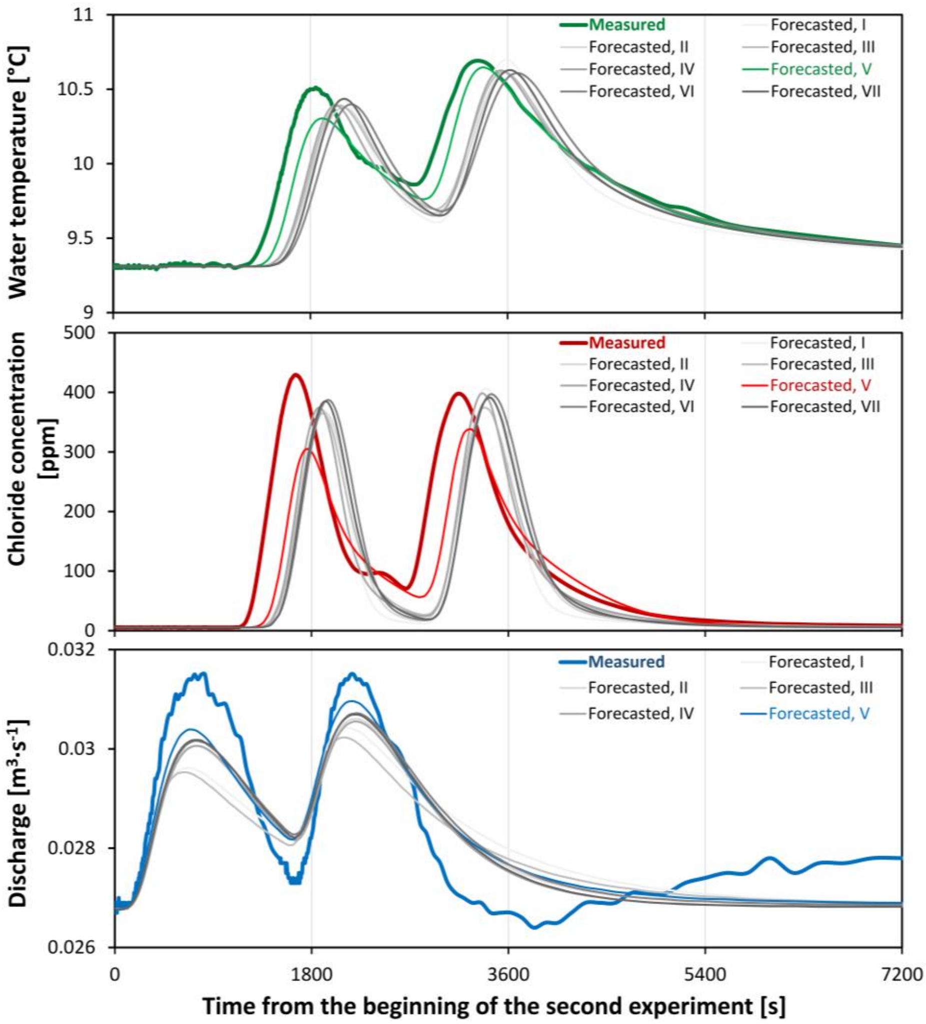

Figure 6. Measured and forecasted hydro-chemo-thermo-graphs of Freiheit Spring for all model variants. The forecasted results of model variants are presented with gray-scale colors, except for those of the model variant V, which has the best statistics (cf. **Figure 5**).

3.4 Value of Observation Data

The model variants are calibrated using all observation data, including spring discharge (Q), solute concentration (C), and water temperature (T) signatures, i.e., hydro-chemo-thermo-graphs. This section discusses the value of different signatures on the certainty of achieved parameter values. For this reason, seven cases of observation data availability (in-short "case"), including, Q, C, T, CT, QC, QT, and QCT are assumed for all model variants. No changes were made to the estimated parameter values, which were obtained from the full observation dataset, i.e., the QCT case (as described in section 0). However, the Jacobian Matrixes are recalculated and employed by PEST to explore the value of observation data based on the extent of parameter uncertainty reduction. For this reason, the PEST-reported 95% confidence intervals of parameters for different cases are compared. In case 95% confidence limits were not reported by PEST, because of the nearly singular normal matrix, confidence limits are externally calculated by the PREDUNC1 PEST utility (Doherty, 2020). It should be mentioned that the cases without discharge observations (i.e., C, T, and CT) are definitely uncommon and undesirable for site data collection. However, they are provided to assure a comprehensive investigation of observation data value.

Figure 7 presents a bar-chart of normalized 95% confidence intervals of parameters for the model variant II, at different cases of observation data availability. Normalization of all values is performed relative to the same parameter estimates for each model variant, which were achieved by the calibration based on all observations (i.e., the QCT case). The model variant II, which was chosen as an example plot (i.e., **Figure 7**), is a variant with two conduit sections laid solely along the tracer path, from the test sinkhole to the spring. Therefore, it achieved relatively narrow 95% confidence intervals, comparing to the variants III to VII (cf. **Figure 4**). Inspecting the results for the model variant II, the following outcomes can be summarized:

1. Comparing the single signatures (i.e. Q, C, and T cases), the Q and C cases recover the narrowest and widest confidence intervals of parameters, on average.
2. Although the parameters do not behave similarly, joint use of signatures in the inversion increases the certainty at which the parameters are being estimated, such that the CT, QT, and QC cases are more capable of reducing parameter uncertainty than the relevant single cases of these combinations, based on the number of parameters with higher reduction of uncertainty across the cases (i.e., according to the average rank of parameter uncertainty reductions).

3. Q case is even more valuable than the combined CT case.
4. Although the T case is substantially more valuable than the C, the combined QC case is slightly favored over the QT one.
5. The normalized confidence intervals for the QCT case are the narrowest for all parameters (**Figure 7**). In other words, parameter estimates are most certain if the full observation dataset for model calibration are considered.
6. Comparing the QC, QT, and QTC cases, there is no significant difference in the reduction of parameter uncertainty if one uses the QT or QC case, instead of the full observation QCT dataset.

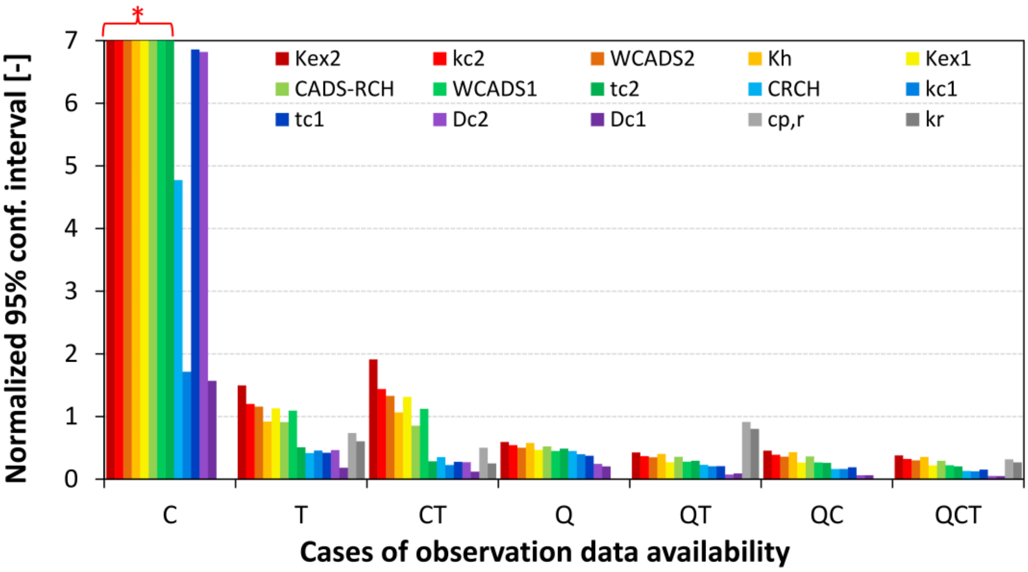

Figure 7. Normalized 95% confidence intervals of parameters for the model variant II, at different cases of observation data availability. The horizontal axis groups the values for each case sorted based on the number of incidences with reduced parameter uncertainty, such that the C and the combined QCT cases exhibit the least and the most reduction of parameter uncertainties, respectively. The values within each case are also sorted based on the average rankings across all the cases, such that the more left and warm the color, the higher the parameter uncertainty. Note that the $C_{p,r}$ and k_r parameters are absent for the C, Q, and QC cases, and therefore, they are not considered in the ranking and are simply presented by gray colors at the right-most end of the relevant cases. The bars indicated by the red star exceed the y-axis range (i.e., 7).

Figure 8 presents the cumulative probability (i.e., the exceedance probability) of normalized 95% confidence intervals of parameters for all model variants, with respect to the different cases of observation data availability. This kind of probability plot is utilized because the parameters do not behave similarly with respect to different observation data, and the normalized 95% confidence intervals could reach very large values, especially for the parameters of the non-tracer conduit sections in the model variants III to VII (cf. **Figure 8**). For the sake of simplicity in comparisons, this general rule is applicable: the more the cumulative probability line shifts to the right, the more significant the reduction of uncertainty for the corresponding case of observation data availability. The following results are achieved by comparing the uncertainty reduction of parameters in different model variants:

1. Comparing the single signatures, the C case is the least valuable in all model variants. Moreover, the Q case is the most valuable in the model variants I and II (as it was already mentioned), while the T case is most valuable in the rest of the model variants, where there are some conduit sections outside of the tracer path.
2. Although the combined use of signatures generally results in higher uncertainty reduction for most of the estimated parameters, the T case could even more effectively reduce the uncertainty of parameters in comparison to the combined use of some signatures in some model variants. Specifically, the T signature is the second most important case of observation data availability (after the QCT case) for the model variant III.
3. The QC case is the second close case to the QCT case in the model variants I and II, in terms of the average value and ranking of parameter uncertainty reduction. However, the QT case is always close to the QCT case in all model variants, and therefore, can be preferred to the QC case, considering the conceptual uncertainty.

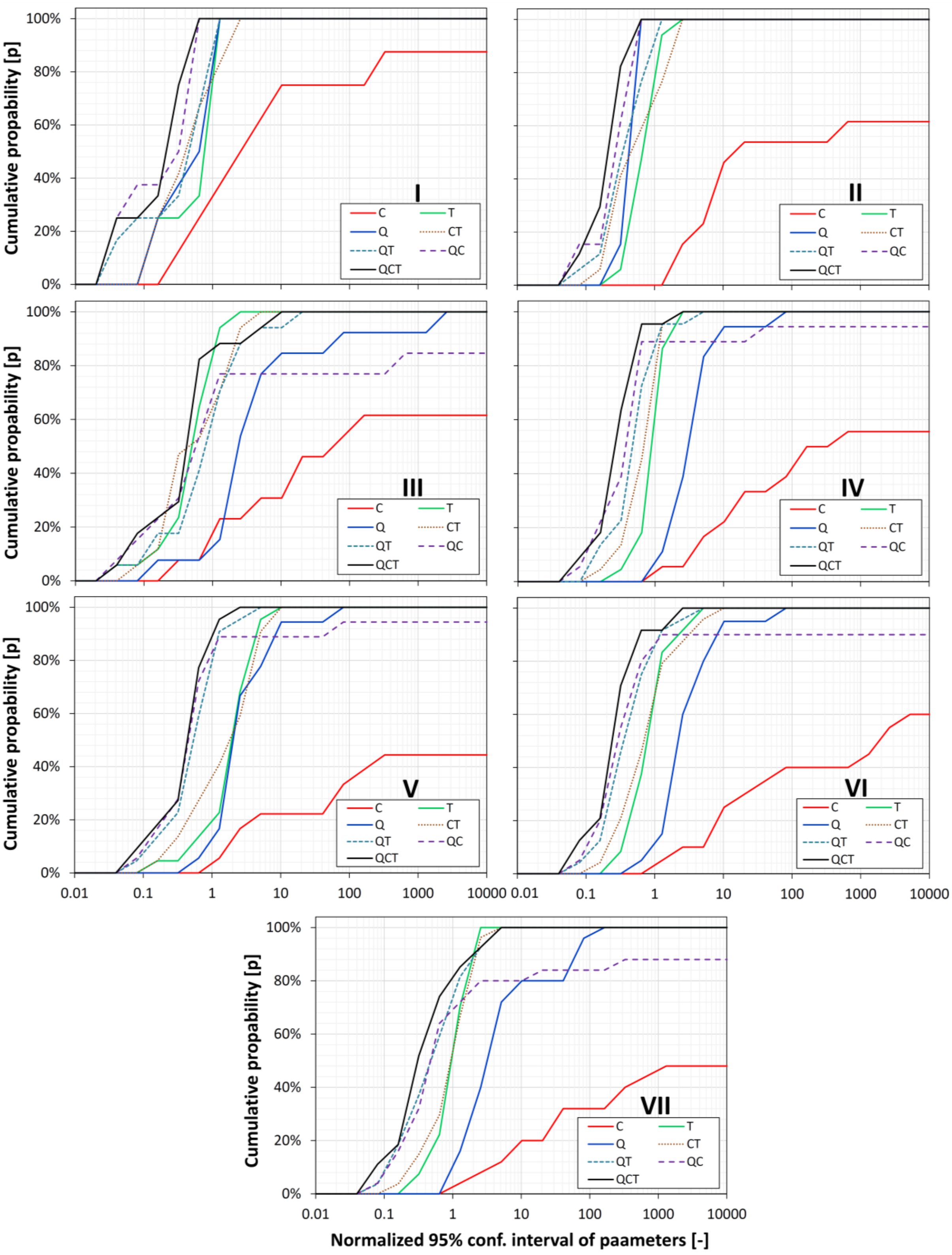

Figure 8. Cumulative probability of normalized 95% confidence interval of parameters at different cases of observation data availability across different model variants.

3.5 Modeling Insights

There are very limited observations of the size and location of the conduits inside the Freiheit karst system. Therefore, none of the candidate model variants can be disregarded. However, as the assumed model variants cover a reasonable range of feasible conduit configurations between the injection and spring point, they can provide some insightful outcomes on the aquifer characteristics and functioning at its terminal part.

3.5.1 Parameter Uniqueness and Conceptual Model Uncertainty

Inspection of parameter values across model variants demonstrate that some parameters could be almost uniquely estimated across variants, while some others are significantly diverse. Specifically, conduit associated parameters for the tracer passage conduit section are estimated at similar values and with a lower degree of uncertainty when comparing to the same parameters in the upstream and downstream mixing conduit sections.

3.5.2 Uniformity of Tracer Conduit Size

Conduit passage collapses, sediment breakdowns, and insoluble blocks may constrict karst conduits (especially at the aquifer terminal parts), and consequently regulate the observed spring signatures. Model simulations have indicated that these features may control the observed behavior of some karst aquifers (Chen & Goldscheider, 2014; Covington et al., 2009; Halihan et al., 1998; Kavousi et al., 2020).

Potential constriction in the conduit path was investigated by inspection of conduit diameters in the tracer path. Model variants II, IV, V, VI, and VII have two classes of conduit sections and therefore can recover constriction features. Although zonated conduit parameters may result in better model performance, emphasizing the importance of conduit parameters in the inversion, the estimated values for the conduit diameter of tracer passage sections are almost uniform in size, ranging from 30.0 to 50.2 cm across the variants (cf. **Table 4**). In other words, the inverse model did not suggest constriction of flow within the conduit sections. It’s also worth mentioning that Luhmann et al. (2012), using sole pipe transport simulations, suggested a conduit diameter of 7 to 8 cm for the Freiheit Aquifer. These values are obviously smaller than our results and highlight the importance of using a process-based discrete-continuum model in comparison to a sole pipe transport one, which consider uniform velocity across model, during each calculation time-step.

3.5.3 Possibility of Inflow from both Upstream and Downstream Mixing Conduits

Values achieved for the conduit inflow of mixing tributary in the model variants VI and VII are ~13.5 and ~14.0 $l \cdot s^{-1}$, respectively (cf. **Table 4**), such that ~77.1 and ~79.7 % of the conduit inflow is allocated to the lateral mixing tributary, while the rest inflow was attributed to the upstream conduit inflow there. Model variants I, II, III and IV which have no mixing tributary conduit section (i.e., 100.0 % upstream conduit inflow) and model variant V which has no upstream conduit inflow (i.e., 100.0 % mixing tributary conduit inflow) are also capable of reproducing the observed spring behavior, although the model variants with both inflows (i.e., model variants VI and VII) perform better in the model calibration. However, the model variant V, which had no upstream inflow, performs better in the forecast (cf. section 0). Overall, one may deduce that the inversion problem tends to favor contribution from both downstream mixing and upstream conduit inflows, with a higher contribution from the former. Nevertheless, this conclusion is valid as long as the aquifer is at the current hydraulic level. Other inactive vadose passages may connect with the spring, creating epi-phreatic to phreatic conditions during high-flows.

3.5.4 Significance of CADS

The estimated values for the width of CADS is on the order of one millimeter to around ten centimeters across all model variants. This may lead to the assumption that the CADS is not important in the simulated experiment (cf. **Table 4**). Neglecting the presumption, a CADS-detached version of the model variant II was constructed and calibrated. The new model has no CADS, and therefore, the concentrated recharge is also solely allocated as CRCH.

Results suggest that the CADS-free model cannot simultaneously adhere to all the observed spring signatures, such that in comparison to the CADS-bearing version of model variant II, the value of Φ is increased to 77.41, which means over a four-time increase in the objective function (cf. **Table 3**). The transport signatures can partially be simulated with the CADS-free model; however, the flow signature cannot be jointly simulated at all. The inadequacy of the CADS-free model highlights the importance of the CADS compartment in the overall aquifer functioning.

3.5.5 Necessity of Some Additional Areas for Heat Exchange

Modeling results indicate that in order to have a successful joint-inversion of all signatures for the conducted experiment, extraordinary large values for tortuosity and rock specific heat are required (**Table 4**).

No extensive analysis on cave sinuosity has been carried out in the studied karst region (Luhmann et al., 2012). However, since our models neglect the vertical passage of water in the infeeder, some large estimates of tortuosities are expected to compensate for the ignored associated loss of heat signals through some additional area for heat exchange. As a rough estimate, we consider the reported median value for tortuosity by Worthington (2014), based on 85 major cave flow paths (which was 1.44); given the elevation and horizontal differences between the test sinkhole and the spring location (which are 19 and 95 m, respectively) and the straight model length (which is 9×10 m), the model tortuosity should be around 1.82 to account for both vertical and horizontal water paths. However, the calculated tortuosity values for the model variants I and III (which had only one tracer conduit section) were 5.4 and 5.2, respectively. Therefore, adopting a large tortuosity value, the model tried to provide extra flow path length, i.e., additional area for heat exchange, which evidently exceeded the assumed loss of heat for combined horizontal and vertical water movement in circular conduits. Considering the additional conduit length due to the large tortuosity, the estimated flow velocity within the conduit should also increase, such that the flow-through-time becomes fixed and an accurate joint-inversion of the spring hydro-chemo-graph can be obtained.

On the other hand, all model variants touched their upper bound for estimation of rock specific heat, which was set at a large value of 5000 $J\cdot kg^{-1}\cdot K^{-1}$ (which is even somewhat larger than that of water, i.e., 4195.2 $J\cdot kg^{-1}\cdot K^{-1}$). Moreover, the estimated rock thermal conductivity ranges from 2.920 to 3.632 $W\cdot m^{-1}\cdot K^{-1}$, i.e., in the range of the largest reported values for carbonate rocks (see Robertson (1988)). The large values of rock thermal parameters also suggest that the model tries to consider more heat exchange in order to simultaneously simulate all signatures.

Freiheit Spring emerges from a bedding plane parting, suggesting that a wide rectangular conduit cross-section shape may be more appropriate than a circular one (Luhmann et al., 2012). Beside this field evidence, rectangular conduits feeding Freiheit Spring is further supported by the simulations conducted by Luhmann et al. (2012). They demonstrated that the rectangular conduits can perform better than the circular one, considering sole pipe transport simulations with conduction in the surrounding rock. Water flowing through rectangular conduits would obviously have additional heat exchange area comparing to the circular conduits, at the same hydraulic radius. However, CFPv2 currently assumes only circular conduits.

3.5.6 Conduit Volume of Tracer Passage

Figure 9 presents the estimated conduit volume of the tracer passage in all model variants. Volumes, along with their pertinent 95% confidence intervals, are calculated by multiplying the length of tracer conduit sections by the relevant estimates for diameters and tortuosities. Conduit volumes range from 45.64 m^3 to 50.79 m^3, while their confidence intervals range from 13.14 m^3 to 25.86 m^3 across all model variants. Estimated conduit volume by the flood-pulse method (Ashton, 1966) was 47 ±10% m^3 (Luhmann et al., 2012). Accordingly, the estimates of conduit volume by our models tend to be very close to the estimates from the flood-pulse method.

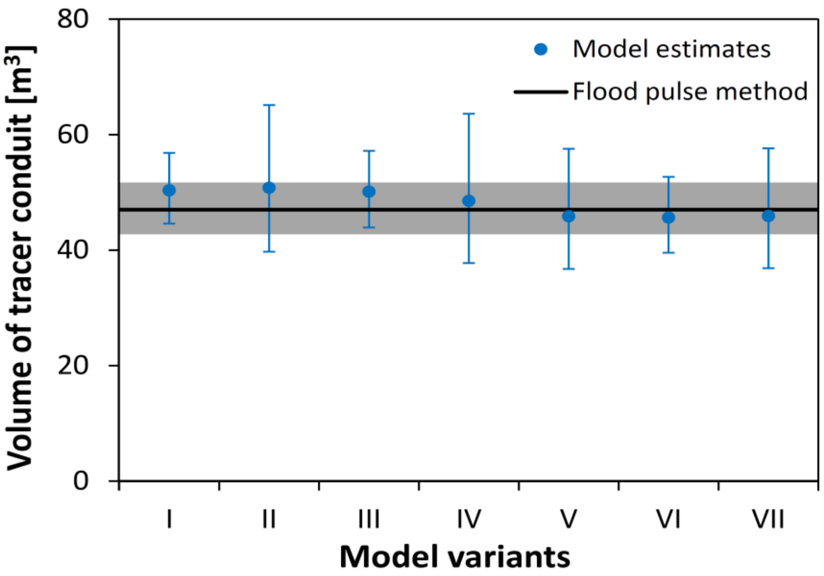

Figure 9. Comparison of estimated volume of tracer conduit passage for all model variants versus the estimate from the flood-pulse method (Ashton, 1966). The 95% confidence intervals of model estimates and error bands of the flood-pulse method are also indicated.

3.6 Model Limitations/Outlook for Spatiotemporally Small-Scale Applications

Our approach simulates most of the hydraulic and transport processes happening in a karst aquifer in a process-based way. However, its application to a spatiotemporally small-scale real-world site highlighted the effect of some processes/characteristics at this scale, which are normally overlooked:

1. Conduit cross-sectional shapes:

   Real-world karst conduits can be of different cross-sectional shape. Freiheit Spring emerges from a bedding plane parting, suggesting that a wide rectangular cross-section shape may be a more appropriate choice than a circular one (Luhmann et al., 2012). Simulation results of Luhmann et al. (2012) for Freiheit Spring also proved that the rectangular conduit shape can achieve better fitting, illustrating enhanced simulation of the thermograph. Jouves et al. (2017) investigated the geometrical shape of a cumulative length of 621 km of cave passages from France, Spain and Italy. They demonstrated that the width-height ratio for water table, looping, and maze conduit networks ranges from ~1.6 to 1.9 on average.

2. Partially-saturated vertical flow and transport processes in conduits:
   The effect of these processes can be observed at short distance and time frames, like the experiments considered for this research. However, as it was discussed in section 0, these processes were ignored in this work, due to their relative low importance and indistinguishability, as well as the current model code limitations.
3. Syphoning or inertia processes in conduit flow:
   Luhmann et al. (2012) assumed such flow processes as a likely cause of odd hydrograph fluctuations upon passage of the hydraulic pulses. Although the current version of CFPv2 is able to consider inertia in conduit flow by means of Partially Filled Pipe Storage (PFPS) functionality, the conduits are here assumed at a level below the spring, such that they are submerged, prohibiting likely model failure during inversion. Ignorance of these flow processes in our models may cause some discrepancies between measured and modelled data (as discussed in section 0).

The abovementioned model structural inadequacies give us the outlook of required discrete-continuum model improvements for spatiotemporally small-scale applications, as the next steps of CFPv2 enhancements.

4 Conclusions

A CFPv2 discrete-continuum model was applied for the joint-inversion of spring hydro-chemo-thermo-graphs in response to a sink to spring hydraulic and tracer injection experiment at a spatiotemporally small-scale size. Adopting a multi-model concept, a set of different conduit configuration variants was considered and compared in terms of calibration performance and parameter uncertainties, as well as forecast capability for a second more complex experiment at the same hydrological state.

Although the forecast capability of all model variants is surprisingly good, the models are used for interpretative aspects, i.e., as engineering calculators for site characterization and screening tools for proof of concept and recognition of potential inadequacies for a process-based discrete-continuum modeling approach at spatiotemporally small-scale applications.

The main modeling outcomes can be summarized as follows:

1) High certainty and uniqueness of estimated values for conduit diameter of tracer passages;
2) Unlikely conduit restriction in the tracer passages;
3) Potential inflow from both upstream and downstream mixing conduits with higher contribution from the latter;
4) Importance of conduit associated drainable storage, CADS, in aquifer hydrodynamic behavior;
5) Necessity of considering some additional areas for heat exchange, e.g., by considering wide rectangular conduits instead of circular ones;
6) Importance of using heat data in joint-inversion;
7) Consistency of the estimated conduit volumes for the tracer passage with those of the routine hydraulic method.

Considering the models as engineering calculators, the most notable result is that the volume of the active conduit compartment of the models is comparable to that estimated by the intuitive flood-pulse method. However, former studies have claimed that Ashton's (1966) method tends to overestimate the conduit volume, ignoring the contribution from the fissure system during the hydraulic-pulse transmission (e.g., Birk et al., 2006; Williams, 1983). While our results support the flood-pulse method results, some generic model investigations are required to thoroughly test the hypothesis under different conceptual models and parameter combinations.

As an additional part of this research, the value of flow, solute concentration, and water temperature data on the certainty degree of model adjustable parameters was investigated. Results suggest that the discharge and temperature data reduce parameter uncertainty more than the solute concentration data in all considered model variants. Moreover, the parameter estimates have the highest degree of certainty if the full observation data were used through the joint-inversion. Results further suggested that if the solute concentration data are excluded from the joint-inversion, the estimated parameters could be recovered at almost the same degree of certainty compared with the case of the full observation dataset. Therefore, considering the objective of data acquisition planning for spring behavior simulation, it is concluded that the continuous record of spring water temperature is of higher priority than the solute concentration. However, this result is case-specific and depends on the current dataset, system-state, and considered conceptual models.

5 Acknowledgement

This work was supported by the Deutsche Forschungsgemeinschaft, DFG for the iKarst project (Grant Numbers: KO 5359/3-1, LI 727/31-1, RE 4001/2-1, and SA 501/28-1). Experts from the University of Minnesota (Calvin Alexander, Anthony Runkel, and Julia Steenberg), Minnesota Department of Natural Resources (Jeff Green and John Barry), and Minnesota Department of Health (Bob Tipping), who have greatly assisted us through better understanding of the Freiheit karst system, are gratefully acknowledged. The archived joint-inversion input and output files as well as utilized model code executable are accessible through the Zenodo repository (as http://doi.org/10.5281/zenodo.4289007).

6 References

Anderson, M. P., Woessner, W. W., & Hunt, R. J. (2015). *Applied groundwater modeling: simulation of flow and advective transport*. Academic press.

Ashton, K. (1966). The analysis of flow data from karst drainage systems. *Transactions of the Cave Research Group of Great Britain*, *7*(2), 161–204.

Atkinson, T. C. (1985). Present and future directions in karst hydrogeology. *Annales - Societe Geologique de Belgique*, *108*(1977), 293–296. Retrieved from https://popups.uliege.be:443/0037-9395/index.php?id=2849

Bakalowicz, M. (2011). Management of Karst Groundwater Resources. In P. E. van Beynen (Ed.), *Karst Management* (pp. 263–282). Dordrecht: Springer Netherlands. https://doi.org/10.1007/978-94-007-1207-2_12

Birk, S. (2002). *Characterisation of karst systems by simulating aquifer genesis and spring responses: Model development and application to gypsum karst*, (Doctoral dissertation). Eberhard–Karls–Universität Tübingen, Germany.

Birk, S., Liedl, R., & Sauter, M. (2006). Karst spring responses examined by process-based modeling. *Ground Water*, *44*(6), 832–836. https://doi.org/10.1111/j.1745-6584.2006.00175.x

Borghi, A., Renard, P., & Cornaton, F. (2016). Can one identify karst conduit networks geometry and properties from hydraulic and tracer test data? *Advances in Water Resources*, *90*, 99–115. https://doi.org/10.1016/j.advwatres.2016.02.009

Brassington, F. C., & Younger, P. L. (2010). A proposed framework for hydrogeological conceptual modelling. *Water and Environment Journal*, *24*(4), 261–273. https://doi.org/10.1111/j.1747-6593.2009.00173.x

Bravo, H. R., Jiang, F., & Hunt, R. J. (2002). Using groundwater temperature data to constrain parameter estimation in a groundwater flow model of a wetland system. *Water Resources Research*, *38*(8), 28-1-28–14. https://doi.org/10.1029/2000WR000172

Chang, Y., Wu, J., Jiang, G., Liu, L., Reimann, T., & Sauter, M. (2019). Modelling spring discharge and solute transport in conduits by coupling CFPv2 to an epikarst reservoir for a karst aquifer. *Journal of Hydrology*, *569* (October 2018), 587–599. https://doi.org/10.1016/j.jhydrol.2018.11.075

Chen, Z., & Goldscheider, N. (2014). Modeling spatially and temporally varied hydraulic behavior of a folded karst system with dominant conduit drainage at catchment scale, Hochifen–Gottesacker, Alps. *Journal of Hydrology*, *514*, 41–52. https://doi.org/10.1016/j.jhydrol.2014.04.005

Cornaton, F. J. (2007). *Ground water: a 3-D ground water and surface water flow, mass transport and heat transfer finite element simulator, reference manual*. University of Neuchâtel, Neuchâtel, Switzerland.

Covington, M. D., Wicks, C. M., & Saar, M. O. (2009). A dimensionless number describing the effects of recharge and geometry on discharge from simple karstic aquifers. *Water Resources Research*, *45*(11), 1–16. https://doi.org/10.1029/2009WR008004

Diersch, H.-J. G. (2014). *FEFLOW*. Berlin, Heidelberg: Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-642-38739-5

DNR (2020). *Minnesota Groundwater Tracing Database Application*. Minnesota Dept. of Natural Resources. Retrieved June 30, 2020, from https://arcgis.dnr.state.mn.us/portal/apps/webappviewer/index.html

Doherty, J. (2015). *Calibration and Uncertainty Analysis for Complex Environmental Models, PEST: complete theory and what it means for modelling the real world*. Watermark Numerical Computing.

Doherty, J. (2019). *PEST: Model-Independent Parameter Estimation, User Manual Part I* (8th Edition). Watermark Numerical Computing.

Doherty, J. (2020). *PEST: Model-Independent Parameter Estimation, User Manual Part II: PEST Utility Support Software* (8th Edition). Watermark Numerical Computing.

Doherty, J., & Hunt, R. J. (2009). Two statistics for evaluating parameter identifiability and error reduction. *Journal of Hydrology*, *366*(1–4), 119–127. https://doi.org/10.1016/j.jhydrol.2008.12.018

Enemark, T., Peeters, L. J. M., Mallants, D., & Batelaan, O. (2019). Hydrogeological conceptual model building and testing: A review. *Journal of Hydrology*, *569*(December 2018), 310–329. https://doi.org/10.1016/j.jhydrol.2018.12.007

Ford, D., & Williams, P. (2007). *Karst Hydrogeology and Geomorphology* (2nd Edition), West Sussex, England: John Wiley & Sons Ltd. https://doi.org/10.1002/9781118684986

Gailey, R. M., Crowe, A. S., & Gorelick, S. M. (1991). Coupled process parameter estimation and prediction uncertainty using hydraulic head and concentration data. *Advances in Water Resources*, *14*(5), 301–314. https://doi.org/10.1016/0309-1708(91)90041-L

Geyer, T., Birk, S., Reimann, T., Dörfliger, N., & Sauter, M. (2013). Differentiated characterization of karst aquifers: some contributions. *Carbonates and Evaporites*, *28*(1–2), 41–46. https://doi.org/10.1007/s13146-013-0150-9

Ghasemizadeh, R., Hellweger, F., Butscher, C., Padilla, I., Vesper, D., Field, M., & Alshawabkeh, A. (2012). Review: Groundwater flow and transport modeling of karst aquifers, with particular reference to the North Coast Limestone aquifer system of Puerto Rico. *Hydrogeology Journal*, *20*(8), 1441–1461. https://doi.org/10.1007/s10040-012-0897-4

Green, J., Tipping, R., Barry, J., Brick, G., Wheeler, B., Rutelonis, J., et al. (2018). Electronic Access to Minnesota Springs, Karst Features & Groundwater Tracing Information. In *Proceedings of the 15th Multidisciplinary Conference on Sinkholes and the Engineering and Environmental Impacts of Karst and the 3rd Appalachian Karst Symposium* (pp. 167–172). National Cave and Karst Research Institute. https://doi.org/10.5038/9780991000982.1049

Gupta, H. V., Kling, H., Yilmaz, K. K., & Martinez, G. F. (2009). Decomposition of the mean squared error and NSE performance criteria: Implications for improving hydrological modelling. *Journal of Hydrology*, *377*(1–2), 80–91. https://doi.org/10.1016/j.jhydrol.2009.08.003

Halihan, T., Wicks, C. M., & Engeln, J. F. (1998). Physical response of a karst drainage basin to flood pulses: example of the Devil's Icebox cave system (Missouri, USA). *Journal of Hydrology*, *204*(1–4), 24–36. https://doi.org/10.1016/S0022-1694(97)00104-2

Harbaugh, A. W. (2005). MODFLOW-2005 , The U.S. Geological Survey Modular Ground-Water Model- the Ground-Water Flow Process. *U.S. Geological Survey Techniques and Methods 6-A16*, 253.

Hartmann, A., Goldscheider, N., Wagener, T., Lange, J., & Weiler, M. (2014). Karst water resources in a changing world: Review of hydrological modeling approaches. *Reviews of Geophysics*, *52*(3), 218–242. https://doi.org/10.1002/2013RG000443

Harvey, C. F., & Gorelick, S. M. (1995). Mapping Hydraulic Conductivity: Sequential Conditioning with Measurements of Solute Arrival Time, Hydraulic Head, and Local Conductivity. *Water Resources Research*, *31*(7), 1615–1626. https://doi.org/10.1029/95WR00547

Hill, M. C., & Tiedeman, C. R. (2007). *Effective Groundwater Model Calibration*. Hoboken, NJ, USA: John Wiley & Sons, Inc. https://doi.org/10.1002/0470041080

Hu, B. X., & Xu, Z. (2016). Numerical Simulation of Groundwater Flow and Solute Transport in a Karst Aquifer with Conduits. In M. S. Javaid (Ed.), *Groundwater - Contaminant and Resource Management*. InTechOpen. https://doi.org/10.5772/63766

Jeannin, P.-Y., & Sauter, M. (1998). Analysis of karst hydrodynamic behaviour using global approaches: a review. *Bulletin d'hydrogéologie*, *16*, 31–48.

Jeannin, Pierre-Yves, Groves, C., Philipp, H., & Häuselmann, P. (2007). Speleological investigations. In N. Goldscheider & D. Drew (Eds.), *Methods in karst hydrogeology* (pp. 25–44). London: IAH: International Contributions to Hydrogeology, 26. https://doi.org/10.1201/9781482266023-9

Jouves, J., Viseur, S., Arfib, B., Baudement, C., Camus, H., Collon, P., & Guglielmi, Y. (2017). Speleogenesis, geometry, and topology of caves: A quantitative study of 3D karst conduits. *Geomorphology*, *298*(January 2018), 86–106. https://doi.org/10.1016/j.geomorph.2017.09.019

Karay, G., & Hajnal, G. (2015). Modelling of Groundwater Flow in Fractured Rocks. *Procedia Environmental Sciences*, *25*, 142–149. https://doi.org/10.1016/j.proenv.2015.04.020

Kavousi, A. (2015). *Controlling parameters on karst aquifer hydrodynamic behavior*, (Doctoral dissertation). Shiraz University, Iran.

Kavousi, A., Reimann, Th., Liedl, R., & Raeisi, E. (2020). Karst aquifer characterization by inverse application of MODFLOW-2005 CFPv2 discrete-continuum flow and transport model. *Journal of Hydrology*, *587* (January), 124922. https://doi.org/10.1016/j.jhydrol.2020.124922

Király, L. (1998). Modelling karst aquifers by the combined discrete channel and continuum approach. *Bulletin d'Hydrogéologie*, *16*, 77–98.

Király, L., Perrochet, P., & Rossier, Y. (1995). Effect of the epikarst on the hydrograph of karst springs: a numerical approach. *Bulletin d'Hydrogéologie*, *14*, 199–220.

Knoben, W. J. M., Freer, J. E., & Woods, R. A. (2019). Technical note: Inherent benchmark or not? Comparing Nash-Sutcliffe and Kling-Gupta efficiency scores. *Hydrology and Earth System Sciences Discussions*, (July), 1–7. https://doi.org/10.5194/hess-2019-327

Kovács, A., & Sauter, M. (2007). Modelling karst hydrodynamics. In N. Goldscheider & D. Drew (Eds.), *Methods in Karst Hydrogeology* (pp. 201–222). London: IAH: International Contributions to Hydrogeology. Retrieved from https://www.taylorfrancis.com/books/e/9780429153419

Larocque, M., Banton, O., Ackerer, P., & Razack, M. (1999). Determining Karst Transmissivities with Inverse Modeling and an Equivalent Porous Media. *Ground Water*, *37*(6), 897–903. https://doi.org/10.1111/j.1745-6584.1999.tb01189.x

Larocque, M., Banton, O., & Razack, M. (2000). Transient-State History Matching of a Karst Aquifer Ground Water Flow Model. *Ground Water*, *38*(6), 939–946. https://doi.org/10.1111/j.1745-6584.2000.tb00694.x

Liedl, R., Sauter, M., Hückinghaus, D., Clemens, T., & Teutsch, G. (2003). Simulation of the development of karst aquifers using a coupled continuum pipe flow model. *Water Resources Research*, *39*(3), 1–11. https://doi.org/10.1029/2001WR001206

Luhmann, A. J., Covington, M. D., Myre, J. M., Perne, M., Jones, S. W., Alexander Jr., E. C., & Saar, M. O. (2015). Thermal damping and retardation in karst conduits. *Hydrology and Earth System Sciences*, *19*(1), 137–157. https://doi.org/10.5194/hess-19-137-2015

Luhmann, Andrew J. (2011). *Water Temperature as a Tracer in Karst Aquifers*, (Doctoral dissertation). University of Minnesota, USA.

Luhmann, Andrew J., Covington, M. D., Alexander, S. C., Chai, S. Y., Schwartz, B. F., Groten, J. T., & Alexander, E. C. (2012). Comparing conservative and nonconservative tracers in karst and using them to estimate flow path geometry. *Journal of Hydrology*, *448–449*, 201–211. https://doi.org/10.1016/j.jhydrol.2012.04.044

Malenica, L., Gotovac, H., Kamber, G., Simunovic, S., Allu, S., & Divic, V. (2018). Groundwater Flow Modeling in Karst Aquifers: Coupling 3D Matrix and 1D Conduit Flow via Control Volume Isogeometric Analysis—Experimental Verification with a 3D Physical Model. *Water*, *10*(12), 1787. https://doi.org/10.3390/w10121787

Mangin, A. (1975). *Contribution à l'étude hydrodynamique des aquifères karstiques: Première partie: Généralités sur le karst et les lois d'écoulement utilisées, Ann. Spéléol., 1974, 29, 3,* Deuxième partie: Concepts méthodologiques adoptés. Systèmes karstiques. Université de Dijon.

Maréchal, J.-C., Ladouche, B., Dörfliger, N., & Lachassagne, P. (2008). Interpretation of pumping tests in a mixed flow karst system. *Water Resources Research*, *44*(5), 1–18. https://doi.org/10.1029/2007WR006288

Milanovic, P. (2004). *Water resources engineering in karst*. CRC press.

Mohammadi, Z., Illman, W. A., & Karimi, M. (2018). Optimization of the hydrodynamic characteristics of a karst conduit with CFPv2 coupled to OSTRICH. *Journal of Hydrology*, *567*(April), 564–578. https://doi.org/10.1016/j.jhydrol.2018.10.050

Mossler, J. H. (2008). *Paleozoic stratigraphic nomenclature for Minnesota.* Minnesota Geological Survey, Report of Investigations 65.

Nash, J. E., & Sutcliffe, J. V. (1970). River flow forecasting through conceptual models part I - A discussion of principles. *Journal of Hydrology*, *10*(3), 282–290. https://doi.org/10.1016/0022-1694(70)90255-6

Neuman, S.P., & Wierenga, P. J. (2003). *A Comprehensive Strategy of Hydrogeologic Modeling and Uncertainty Analysis for Nuclear Facilities and Sites*. U.S. Nuclear Regulatory Commission Office of Nuclear Regulatory Research, 311.

Neuman, S. P. (1973). Calibration of distributed parameter groundwater flow models viewed as a multiple-objective decision process under uncertainty. *Water Resources Research*, *9*(4), 1006–1021. https://doi.org/10.1029/WR009i004p01006

Panday, S., Langevin, C. D., Niswonger, R. G., Ibaraki, M., & Hughes, J. D. (2013). *MODFLOW – USG Version 1: An Unstructured Grid Version of MODFLOW for Simulating Groundwater Flow and Tightly Coupled Processes Using a Control Volume Finite-Difference Formulation*. U.S. Geological Survey, (Techniques and Methods 6-A45), 66.

Pushpalatha, R., Perrin, C., Moine, N. L., & Andréassian, V. (2012). A review of efficiency criteria suitable for evaluating low-flow simulations. *Journal of Hydrology*, *420–421*, 171–182. https://doi.org/10.1016/j.jhydrol.2011.11.055

Quinlan, J. F., & Ray, J. A. (1995). Normalized base-flow discharge of groundwater basins: A useful parameter for estimating recharge area of springs and for recognizing drainage anomalies in karst terranes. In B. F. Beck (Ed.), *Karst Geohazards: Engineering and Environmental Problems in Karst Terrane* (pp. 149–164). Routledge, London. https://doi.org/10.1201/9780203749708-19

Raeisi, E., Zareh, M., & Eftekhari, P. (1999). Application of dye tracing for determining the characteristics of Sheshpeer karst spring , Iran. *Theoretical and Applied Karstology*, *11–12*, 109–118.

Reimann, T., Giese, M., Geyer, T., Liedl, R., Maréchal, J. C., & Shoemaker, W. B. (2014). Representation of water abstraction from a karst conduit with numerical discrete-continuum models. *Hydrology and Earth System Sciences*, *18*(1), 227–241. https://doi.org/10.5194/hess-18-227-2014

Reimann, T., Liedl, R., Birk, S., & Bauer, S. (2018). *Modifications and enhancements to CFPM1 flow subroutines and addition of transport subroutines*. Retrieved from http://tu-dresden.de/die_tu_dresden/fakultaeten/fakultaet_forst_geo_und_hydrowissenschaften/fachrichtung_wasserwesen/igw/forschung/downloads/cfpv2

Robertson, E. C. (1988). *Thermal Properties of Rocks*. US Department of the Interior: Geological Survey, Open-File Report 88-441.

de Rooij, R., Perrochet, P., & Graham, W. (2013). From rainfall to spring discharge: Coupling conduit flow, subsurface matrix flow and surface flow in karst systems using a discrete–continuum model. *Advances in Water Resources*, *61*, 29–41. https://doi.org/10.1016/j.advwatres.2013.08.009

Runkel, A. C., Tipping, R., Alexander, C. E., & Green, J. A. (2003). *Hydrogeology of the paleozoic bedrock in southeastern Minnesota*. Minnesota Geological Survey, Report of Investigations 61 (Vol. 61).

Ryan, M., & Meiman, J. (1996). An Examination of Short-Term Variations in Water Quality at a Karst Spring in Kentucky. *Ground Water, 34*(1), 23–30. https://doi.org/10.1111/j.1745-6584.1996.tb01861.x

Sauter, M. (1992). Quantification and Forecasting of Regional Groundwater Flow and Transport in a Karst Aquifer, (Doctoral dissertation). Universität Tübingen, Germany.

Shoemaker, W. B., Kuniansky, E. L., Birk, S., Bauer, S., & Swain, E. D. (2008). *Documentation of a Conduit Flow Process (CFP) for MODFLOW-2005 A product of the Ground-Water Resources Program*. U.S. Geological Survey. https://doi.org/10.3133/tm6A24

Shuster, E. T., & White, W. B. (1971). Seasonal fluctuations in the chemistry of lime-stone springs: A possible means for characterizing carbonate aquifers. *Journal of Hydrology*, *14*(2), 93–128. https://doi.org/10.1029/WR025i001p00126

Smith, E. A., & Westenbroek, S. M. (2015). *Potential Groundwater Recharge for the State of Minnesota Using the Soil-Water-Balance Model , 1996 – 2010.* Scientific Investigations Report 2015 – 5038. USGS Scientific Investigations Report. https://doi.org/10.3133/sir20155038

Steenberg, J. R., & Runkel, A. C. (2018). *Stratigraphic Positions of Springs in Southeast Minnesota*. Minnesota Geological Survey.

Steenberg, J. R. (2014). *Bedrock Geology of Houston County, Geologic Atlas of Houston County, Minnesota* (Vol. C–33). Minnesota Geological Survey.

Stevanović, Z. (2019). Karst waters in potable water supply: a global scale overview. *Environmental Earth Sciences, 78*(23), 662. https://doi.org/10.1007/s12665-019-8670-9

Sullivan, T. P., Gao, Y., & Reimann, T. (2019). Nitrate transport in a karst aquifer: Numerical model development and source evaluation. *Journal of Hydrology, 573*(March), 432–448. https://doi.org/10.1016/j.jhydrol.2019.03.078

Therrien, R., McLaren, R. G., Sudicky, E. A., & Panday, S. M. (2010). *HydroGeoSphere. A three-dimensional numerical model describing fully-integrated subsurface and surface flow and solute transport*. Groundwater Simulations Group.

Tinet, A. J., Collon, P., Philippe, C., Dewaide, L., & Hallet, V. (2019). OM-MADE: An open-source program to simulate one-dimensional solute transport in multiple exchanging conduits and storage zones. *Computers and Geosciences*, *127*, 23–35. https://doi.org/10.1016/j.cageo.2019.03.001

Wilcock, J. D. (1968). Some developments in pulse-train analysis. *Transactions of the Cave Research Group of Great Britain*, *10*(2), 73–98.

Williams, P. W. (1983). The role of the subcutaneous zone in karst hydrology. *Journal of Hydrology*, *61*(1–3), 45–67. https://doi.org/10.1016/0022-1694(83)90234-2

Wöhling, T., Samaniego, L., & Kumar, R. (2013). Evaluating multiple performance criteria to calibrate the distributed hydrological model of the upper Neckar catchment. *Environmental Earth Sciences*, *69*(2), 453–468. https://doi.org/10.1007/s12665-013-2306-2

Worthington, S. R. H. (1999). A comprehensive strategy for understanding flow in carbonate aquifers. In A. N. Palmer, M. V. Palmer, & I. D. Sasowsky (Eds.), *Karst Modeling* (Vol. 1, pp. 30–37). Karst Waters Institute, Special Publication 5.

Worthington, S. R. H., & Ford, D. C. (2009). Self-organized permeability in carbonate aquifers. *Ground Water*, *47*(3), 326–336. https://doi.org/10.1111/j.1745-6584.2009.00551.x

Worthington, S. R. H. (2014). Characteristics of channel networks in unconfined carbonate aquifers. *Bulletin of the Geological Society of America*, *127*(5–6), 759–769. https://doi.org/10.1130/B31098.1

Xu, T., & Gómez-Hernández, J. J. (2016). Characterization of non-Gaussian conductivities and porosities with hydraulic heads, solute concentrations, and water temperatures. *Water Resources Research*, *52*(8), 6111–6136. https://doi.org/10.1002/2016WR019011

Xu, Z., Hu, B. X., Davis, H., & Kish, S. (2015). Numerical study of groundwater flow cycling controlled by seawater/freshwater interaction in a coastal karst aquifer through conduit network using CFPv2. *Journal of Contaminant Hydrology*, *182*, 131–145. https://doi.org/10.1016/j.jconhyd.2015.09.003

Xu, Z., Hu, B. X., Davis, H., & Cao, J. (2015). Simulating long term nitrate-N contamination processes in the Woodville Karst Plain using CFPv2 with UMT3D. *Journal of Hydrology*, *524*(2015), 72–88. https://doi.org/10.1016/j.jhydrol.2015.02.024

Zheng, C., & Bennett, G. D. (2002). *Applied contaminant transport modeling* (2nd Edition), Wiley-Interscience, New York.

Zimmerman, W. B. J. (2006). Introduction To COMSOL Multiphysics, 1–26. In W. B. J. Zimmerman (Ed.), *Multiphysics Modeling with Finite Element Methods,* World Scientific Publishing. https://doi.org/10.1142/9789812773302_0001