Freiburg, October 11th, 2017

Dear editor,

we are happy to provide the revised version of our manuscript

On the value of water quality data and informative flow states in karst modelling.

According to the remarks of the two referees, Dr Naomi Mazzilli and Dr Arnauld Malard, we added more information on the solute transport approach of the model, the Kling-Gupta efficiency and the problem of the varying temporal resolution of the discharge end hydrochemical measurements.

Please find below a detailed point-by-point response to the reviewer comments and a changes-tracked version of the manuscript.

We want to thank the associate editor and the two referees for their valuable remarks that led to a significant improvement of our original submission and we hope that the manuscript can now be considered for publication in Hydrology and Earth System Sciences.

Sincerely,

_____

Andreas Hartmann
Juan Antonio Barberá
Bartolomé Andreo

Reviewer #1 (Naomi Mazzilli)

This study examines the information content of water quality data for a karst simulation model. The Varkarst model is applied to a spanish watershed. An initial 500.000 random parameters sets is confined using discharge and water quality data which are either taken separately or together, using the whole time series of focusing on specific flow stages. The reduction in the 25th to 75th percentiles range is used to estimate the information content of the data.

The methodology is clear, the results are commented with adequate references to related works, the illustrations are sufficient and informative. In my opinion, the manuscript deserves to be published after minor revisions.

Our response: We thank Dr Mazzilli for her positive and valuable review.

General comments

• The model equations are detailed in annex but for more clarity some information about the solute model should appear in Section 3. The lack of NO3-specific parameter is surprising and should be commented.

Our response: The model description and an elaboration of the NO_3^- dynamics was be improved in the revised version of the manuscript (subsection 3.2):

> "Solute transport simulations within the model follow the assumption of instantaneous and complete mixing within each storage (soil, epikarst, groundwater) and each of the N model compartments (Figure 3). In the particular case of NO_3^-, this implies neglecting plant uptake and release processes, which were found to be important in more humid regions (Hartmann et al., 2016) but it was found a valid assumption at Mediterranean regions such as our study site (Hartmann et al., 2013b, 2014b)."

• Parameter kE seems unsensitive. Is this related to some specificity of the flow processes on the test site ?

Our response: The parameter KE often interacts with the parameter VE. A discussion of the consequences of this interaction was added to the revised manuscript (Subsection 5.2):

> "Only a_{GW} and KE remain with low identifiability, which may be due to structural limitations of the model structure (Clark et al., 2008) or due to parameter interactions that are not explicitly considered in our approach. In fact, a lower identifiability of K_E in favour of a high identifiability of VE was found in a previous study with a similar version of the model (Hartmann et al., 2015).

• The KGE is nicely defined as a combination of linear correlation and the ratios of the mean and standard deviations of the simulations and observations. Have you had a look at whether the parameters have more influence on r, alpha or beta ?

Our response: This is a very interesting remark. An analysis or the influence of r, alpha and beta was partially done in Hartmann *et al.* (2013). In the case of this study, we omitted the analysis of the individual components of the KGE as we already considered three different variables (discharge, NO_3^- and SO_4^{2-}) and we did not want to confuse the reader by adding more dimensions to this evaluation (also we believe that the results we get is sufficient to make our point). However, we agree that now having quantified the information content of the different data types, a follow up study should

analyse in more detail how the model performs for r, alpha and beta individually. We added a respective statement to the conclusions:

> "Also, a further disaggregation of the Kling Gupta efficiency in its components, correlation, bias and variability, contains high promise for further advance of our approach."

Technical comments

• *p 5 l 14 "wtaer"*

• *p 7 l 16 "the fore" instead of "therefore"*

• *p 10 l 14 "the unsaturated state (VE Kc)" shoud be corrected in ""the saturated*

state (VE Kc)"

• *p12 l 10 "is provide"*

• *p 12 caption of Figure 6: "the he 25th"*

• *p 15 l 26 something is missing in the sentence "discharge thresholds from wich different compartments (...) of the behaviour"*

Our response: Corrections were applied to address all technical comments.

Reviewer #2 (Arnauld Malard)

This paper intends to provide an approach for reducing uncertainties in the Varkarst simulation model (= lumped model divided in compartments). 500'000 parameters sets have been confined using discharge, NO3- and SO42- measurements, (i) together or in a separate way and (i) applied on the whole time series or in sub-series corresponding to expected flow processes (floods, recession, mid-stages). Besides, datasets have been resampled in the range of the 25th to 75th percentiles using soft rules in order to assess how the observations contribute to describe the parameter. Finally, repeated simulations using the reduced 250'000 sets of parameters make it possible for the authors to identify that: - "NO3- provides most information to identify the model parameters controlling soil and epikarst dynamics for unsaturated -flow state (i.e. flood events" - "SO42- and discharge data provides most information to identify the model parameters for saturated-flow state (i.e. recession periods).

The approach sounds coherent but authors might provide more information on the model timestep and the timestep used for applying the Kling-Gupta coefficient. Indeed, measurements are of lower resolution and it is not mentioned how the authors managed that.

A few other comments - and minor corrections in the attached .pdf

Few more words on the soft rules would also be appreciated

Our response: We thank Dr Malard for his valuable recommendations. In the revised manuscript, we provided a more detailed model description (as also recommended by Dr Mazzilli in her review) including more information about the temporal resolution of the model and the observations and how they were linked within the parameter estimation (subsection 3.3):

> "For the calculation of KGE, only time steps, at which observations are available, are considered. Hence, the KGE values will only express the model performance to reflect the discharge, NO_3^- and SO_4^{2-} observations that were sampled in a 7-8 days temporal resolution (Table 1) even though the model runs on a daily time step."

Also, some more elaboration on the soft rules was provided in the methods section (subsection 3.3):

> "The threshold value of 0.2 was found by preliminary analysis. Its rather low value is meant to take into account that the simulation is exposed to various sources of uncertainty including uncertainties of the model input (observation of climate variables and their application to the entire recharge area), model structure uncertainty (representation of karst processes by conceptual mathematical formulations in a semi-distributed way), and the uncertainty of observations (discharge measurement and hydrochemical analysis, as well as their low temporal resolution)."

Both issues, the impact of lower resolution of the measurements, as well as the impact of variations in the soft rules is now discussed in more detail in the discussion section of the revised manuscript (subsection 5.3):

> "Another limitation of our research is the low resolution of the discharge and hydrochemical observations (7-8 days). Although our approach took into account this weakness by the soft rules allowing for remaining uncertainty after the reduction of our 500,000 parameter sets, we believe that a higher resolution of the observations (preferably 1 day) would have resulted in a more pronounced reduction of the initial sample and consequently to a lower remaining uncertainty."

Specific and technical comments from commented pdf

P5L3: For some event spring's peaks discharge seem to be comcomitant with EC depletion...

The resolution of the flow measurements (1 measure/week) reveals insufficient to ensure the supposed concomittancy...

Our response: True, we rephrased this statement.

P5L17: From where? Epikarst, Unsaturated zones or drainage of the phreatic zone?

Our response: The phreatic zone. We clarified this in the revised manuscript.

P5L18: "Seepage" from the epikarst should not be disregarded...

Our response: We agree that seepage from the epikarst will still be abundant during this stage. We clarified this and added some elaboration why we believe that our distinction of flow states still makes sense (subsection 2.2):

> "Even though, there still might be some seepage from the soil and epikarst during this stage, the hydrochemical signature of the spring, which is dominated by the signal of the phreatic zone (Barberá and Andreo, 2015), shows that these fractions are not very important."

P6L10: daily timescale?

Our response: Yes, daily time scale. We provided this information in the revised version of the paper.

P8L15: How did you manage the differences in time step between model (daily) vs. measurements (biweekli)?

Our response: Simulations and observations are only compared by KGE at times when observations are available. If the resolution of observations were higher, more parameter sets could have been discarded by our soft rules and the precision of the simulation with the remaining parameter sets

would have been higher. We added this important information to the methods and discussion section (please see our response to the general comment of this review).

P9L26: The "combined" state should be explicitly mentioned in Figure 2.

Our response: We updated Figure 2 and its caption accordingly.

P10L14: replace: "saturated"

Our response: The word was be replaced.

P11: marked areas in Fig 5

Our response: Unfortunately, there is no comment explaining the marked areas in Fig 5. For the revisions, we will assume that they were only included to facilitate the review but do not require any modification to the manuscript.

P11L10: Use "SO42-"

Our response: Corrected.

P12L10: typo

Our response: Corrected.

P12L10: typo

Our response: Corrected.

P13L8: clarify

Our response: We refer to the entire time period that also includes the periods of river influence. The statement was clarified accordingly.

P13: Please Make this figure bigger...(Fig 7)

Our response: The figure was enlarged in the new version of the manuscript.

P13L25: typo

Our response: Corrected.

P16L13: ...; one being a known period...

Our response: The sentence was changed accordingly.

On the value of water quality data and informative flow states in karst modelling

Andreas Hartmann[1, 2], Juan Antonio Barberá[3], Bartolomé Andreo[3]

[1]Faculty of Environment and Natural Resources, University of Freiburg, Germany
5 [2]Department of Civil Engineering, University of Bristol, UK
[3]Department of Geology and Centre of Hydrogeology of the University of Malaga (CEHIUMA), Malaga 29071, Spain

Correspondence to: Andreas Hartmann (andreas.hartmann@hydrology.uni-freiburg.de)

10 **Abstract.** If properly applied, karst hydrological models are a valuable tool for karst water resources management. If they are able to reproduce the relevant flow and storage processes of a karst system, they can be used for prediction of water resources availability when climate or land use are expected to change. A common challenge to apply karst simulation models is the limited availability of observations to identify their model parameters. In this study, we quantify the value of information when water quality data (NO_3^- and SO_4^{-2}) is used in addition to discharge observations to estimate the

15 parameters of a process-based karst simulation model at a test site in Southern Spain. We use a three-step procedure to (1) confine an initial sample of 500,000 model parameter sets by discharge and water quality observations, (2) identify alterations of model parameter distributions through the confinement, and (3) quantify the strength of the confinement for the model parameters. We repeat this procedure for flow states, at which the system discharge is controlled by the unsaturated zone, the saturated zone, and the entire time period including times when the spring is influenced by a nearby river. Our

20 results indicate that NO_3^- provides most information to identify the model parameters controlling soil and epikarst dynamics during the unsaturated flow state. During the saturated flow state, SO_4^{-2} and discharge observations provide the best information to identify the model parameters related to groundwater processes. We found reduced parameter identifiability when the entire time period is used as the river influence disturbs parameter estimation. We finally show that most reliable simulations are obtained when a combination of discharge and water quality date is used for the combined unsaturated and

25 saturated flow states.

1   Introduction

It is estimated that around 10-15% of emerged Earth surface is covered by soluble rocks that are susceptible to be karstified (Ford and Williams, 2013). Today, aquifers developed in such type of rocks roughly supply with groundwater to a quarter of

30 world´s population. The importance of groundwater resources from karst aquifers is not only limited to satisfy the fresh water demand of large regions with some millions of inhabitants (e.g. Austria or Slovenia), but also it guarantees the water supply in small settlements where karst waters are the only source of drinking water.

The intrinsic characteristics of karst aquifers such as the development of a secondary porosity through enlarged conduits and fractures, as well as the duality in the recharge (diffuse *vs* concentrated), result in a high permeability media (Bakalowicz, 2005; White and White, 2003). The shallower parts of the aquifers, including soil and epikarst (i.e. unsaturated zone), play a key role for the understanding of system functioning. Epikarst is characterized by slow percolation of air and water into narrow fissures, inducing water storage, and by a rapid drainage through connected conduit system promoting flow concentration (Aquilina et al., 2006; Ford and Williams, 2013; Labat et al., 2000). Thus, in the first top meters of aquifer rock, biogeochemical processes occur in a multiphase environment (gas, liquid and solid) and recharge waters rapidly acquire their chemical composition, keeping practically unaltered until reach the emergence points. Rapid drainage impedes that such physical-chemical processes may attenuate naturally a potential contaminant entering into the system. Therefore, karst aquifers are especially vulnerable to the contamination despite that the unsaturated zone, jointly with soil and epikarst, acts chemically as reaction layer able to modify the groundwater quality in a substantial way.

Simulation models are a common tool to address water management questions such as the impacts of climate and land use changes on karst water resources (Hartmann et al., 2014a). In order to provide reliable predictions those models need to include the most relevant processes of karst systems and various approaches have been developed to include karst processes in distributed and lumped karst simulation models (Ghasemizadeh et al., 2012; Hartmann et al., 2014a; Hartmann and Baker, 2017; Kovacs and Sauter, 2007; Sauter et al., 2006). The choice of the model approach is usually due to the required purpose. A key challenge in all of these karst modelling approaches is the identification of the model parameters. Methods to explore and analyse karst systems can provide prior knowledge on karst system properties (Goldscheider and Drew, 2007) that can be used to gain prior information of karst model parameters such as hydraulic conductivities or catchment boundaries. However, capturing the entire heterogeneity of karst systems with those methods is commonly impossible (Hartmann et al., 2013a) and inverse parameter estimation schemes, for instance automatic calibration by observed discharge, have to be applied.

Work with automatic calibration approaches early showed that using only discharge observations for model calibration allows to identify up six model parameters (Jakeman and Hornberger, 1993; Wheater et al., 1986; Ye et al., 1997). More recent work also revealed that including disinformative periods in the calibration, i.e. periods when errors in the observation can be expected, may significantly bias the results of model calibration and evaluation of hydrological models (Beven et al., 2011; Beven and Westerberg, 2011; Kauffeldt et al., 2013). Due to the complexity of karst processes, karst models usually require more than 6 model parameters to reflect the most important hydrological processes. Some studies tried to compensate for this apparent lack of information by using auxiliary data such as gravimetric information (Mazzilli et al., 2012), artificial tracer experiments (Hartmann et al., 2012; Oehlmann et al., 2015), or hydrochemical information (Charlier et al., 2012; Hartmann et al., 2013b, 2016). However, to our knowledge the problem of disinformative observations, either discharge observations or auxiliary information, has not been addressed explicitly in karst modelling studies.

This study proposes a new approach to quantitatively assess the information content of discharge and hydrochemical information for karst model calibration including periods with disinformative observations. A process-based model is used to

simulate the hydrodynamic and hydrochemical (NO_3^- and SO_4^{-2}) behaviour of a karst system, at which the unsaturated zone dynamics dominates under recharge conditions, controlling groundwater flow and solute transport processes. During specific periods, the discharge and chemistry of the system is influenced by the surface flow of a nearby river, which constitutes disinformative periods for model parameter estimation. A new parameter estimation approach is employed to estimate the information content of the different types of calibration data during pre-defined flow states that focus on time periods dominated by unsaturated zone discharge, saturated zone discharge, and periods that include the disinformative observations. Even though applied to one particular study site this approach can easily be transferred to any hydrological system where different observation types are available for calibration.
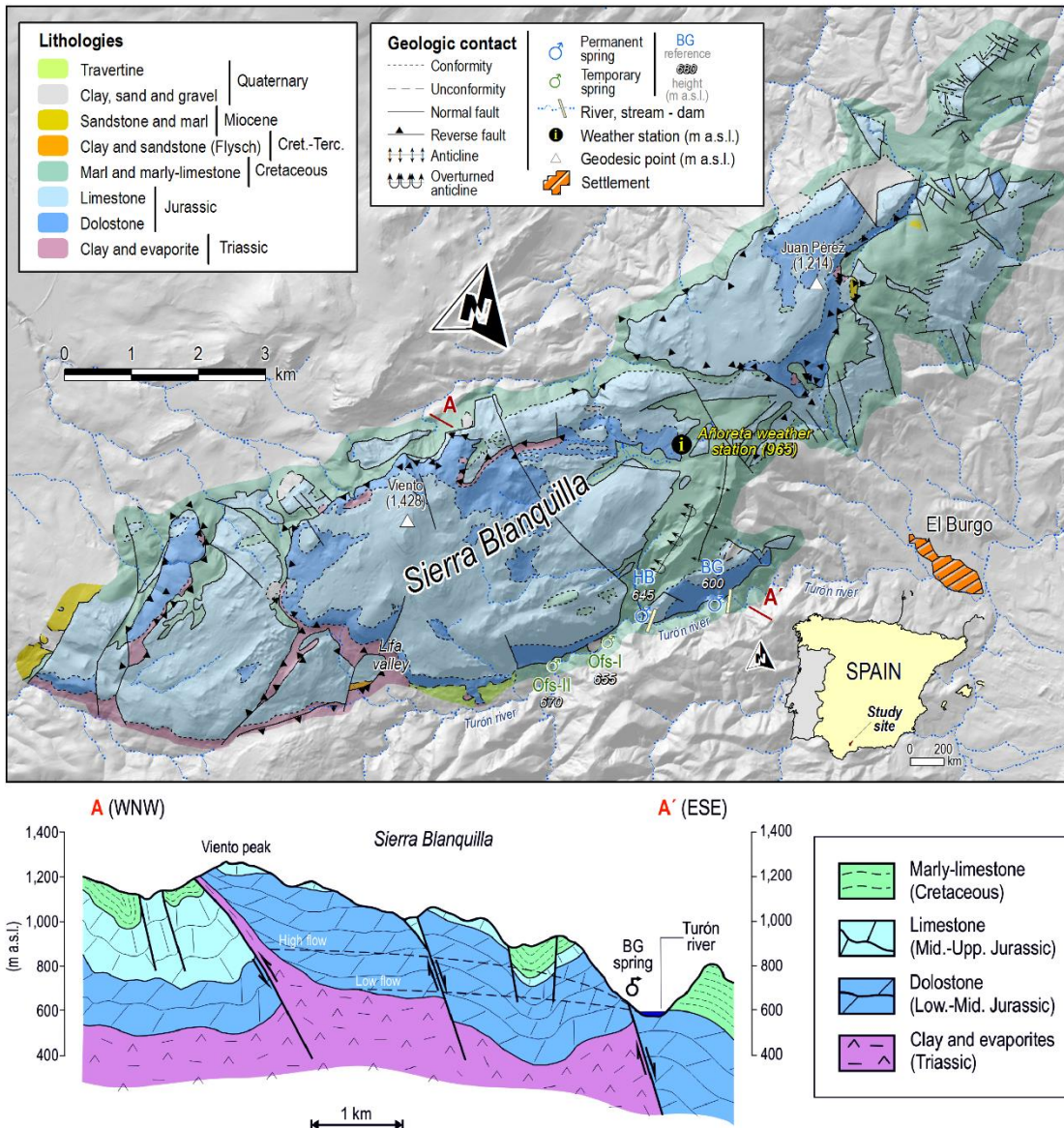
2   Study site description

The experimental area is located in the Eastern Ronda Mountains, at the NW of Málaga province (S Spain). It consists of steep and rugged NE-SW oriented reliefs (e.g. Sierra Blanquilla), reaching a maximum height of 1,428 m a.s.l. (Viento peak; Figure 1~~Figure 1~~). Geologically, three main stratigraphic groups can be differentiated (Cruz-Sanjulián, 1974; Martín-Algarra, 1987, Figure 1~~Figure 1~~): (i) clays and evaporites of upper Triassic age (the older formation); (ii) a thick (up to 500 m) carbonate sequence of Jurassic dolostones and limestones forming the main aquifer (i.e. Sierra Blanquilla); and (iii) Cretaceous-Paleogene marls and marly limestones as the uppermost materials. The geological structure of Sierra Blanquilla is constituted by a NE-SW oriented box-shaped anticline, plunging towards NE (Martín-Algarra, 1987), with a flat and wide hinge, as well as subvertical flanks. The folded structure is also fractured by two set of faults N50-70E and N150E oriented (Fernández, 1980). From the point of view of the karst landscape development, in plateau areas the horizontal bedding planes of carbonate exposures jointly to the high precipitation rate have favoured the formation of exokarstic features including karrenfields, dolines, uvalas, shafts and swallets, as result of intense karstification processes.

2.1   Karst hydrogeology

Sierra Blanquilla carbonate aquifer is permeable by fracturation and karstification. Recharge is mostly produced by rainwater infiltration through the carbonate exposures, although seepage from a losing river and streams also account for groundwater input (Barberá and Andreo, 2015, 2012). Natural groundwater discharge is preferentially conducted toward the SE border of the aquifer (Figure 1~~Figure 1~~), through several springs that constitute the discharge area towards the Turón river valley (Barberá, 2014). Among them, El Burgo (BG, 600 m a.s.l.) and Hierbabuena (HB, 645 m a.s.l.) perennial springs drain most of the groundwater of the hydrogeological system (Figure 1~~Figure 1~~). During high flow periods, when the total flow of the BG and HB springs exceeds 1.1 $m^3 \cdot s^{-1}$, two overflow springs (OfsI, 655 m a.s.l.; and OfsII, 670 m a.s.l.), located upstream of the permanent ones, activate after heavy rainfall events (Barberá and Andreo, 2015). Low flow is established when the permanent groundwater flow (from BG and HB springs) is below to 0.2 $m^3 \cdot s^{-1}$.

3

The main hydrological feature in the test site, Turón River, crosses intermittently the carbonate exposures in the southern border of Blanquilla aquifer (Figure 1Figure 1). The surface flow has been demonstrated to alter the hydrodynamic functioning of both perennial springs (Barberá and Andreo, 2015), which are partly affected by the existence of two regulation dams (20-25 m high) built over the Turón riverbed, just several tens of meters downstream from the springs (Figure 1Figure 1). In high flow periods, both headwaters and groundwater discharge from Sierra Blanquilla aquifer maintain the river flow, while during low flow conditions, the Turón river is exclusively fed by karst groundwater.
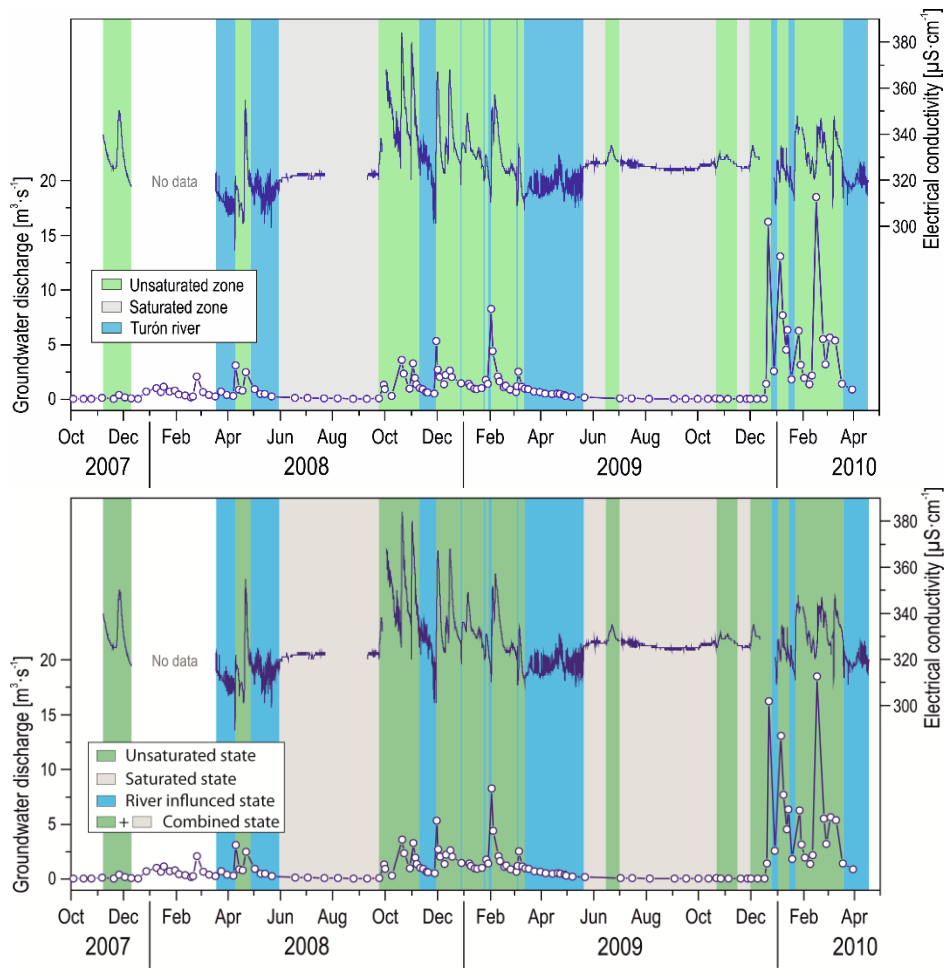
Figure 1: Geographic, geological and hydrogeological features of Sierra Blanquilla carbonate aquifer.

2.2 Dominant hydrogeological processes

Electrical conductivity (EC) has been used as global physical-chemical marker for distinguishing the hydrochemical states

5 that characterize El Burgo spring discharge. Generally, EC peaks ~~are~~ seem to be concomitant with maximum spring

discharge at event scale, which evidence that more mineralized groundwater is drained immediately after each rainfall

episode (green shaded areas in Figure 2~~Figure 2~~). Barberá and Andreo (2015) stated that this high EC groundwater is also

characterized by higher Alkalinity and $logPCO_2$ values and higher Ca^{+2} and TOC contents, suggesting predominant

limestone dissolution in the shallower parts of the aquifer. This spring behaviour reflect a functioning based on a "piston effect", by which groundwater stored in the epikarst reservoir is pulled out to the unsaturated and saturated zone until the discharge zone by a subsequent recharge pulse. Therefore, unsaturated flow dominates under high water conditions in El Burgo spring (state 1 - unsaturated zone, in Figure 2~~Figure 2~~).

5   Under low flow conditions (no rainfall, grey shaded areas in Figure 2~~Figure 2~~), EC levels in groundwater remain quite stable in the range of 320-330 µS/cm. This provides the chemical baseline of the system (state 2 - saturated zone, in Figure 2~~Figure 2~~), which is dependent on the accumulated rainfall on each hydrological year. The lower and less variable EC values of groundwater compared with those obtained under high ~~wtaer~~water conditions can be explained by the loss of aggressiveness of groundwater (degassed waters respect to CO_2) flowing through the system as consequence of the lack of aquifer recharge

10   (Barberá and Andreo, 2015). Therefore, groundwater drainage under low water conditions consists of a system of slower flows coming from capacitive compartments of the aquifer (~~matrix~~the phreatic zone). In these circumstances, the functioning of the hydrogeological system is mainly dominated by the saturated zone (state 2 - saturated zone, in Figure 2~~Figure 2~~). Even though, there still might be some seepage from the soil and epikarst during this stage, the hydrochemical signature of the spring, which is dominated by the signal of the phreatic zone (Barberá and Andreo, 2015), shows that these fractions are not

15   very important.

Figure 2: Decomposition of El Burgo spring flow in selected hydrochemical states from EC and discharge time series: (1) unsaturated zone dominates discharge; (2) saturated zone dominates discharge; and (3) discharge (and EC) influenced by Turón river flow; the combination of unsaturated and saturated states represents the combined flow state.

Marked dilutions in groundwater mineralization (below the chemical baseline of the system), which very often occur during the spring recession after flood events, are also observed in the chemograph of El Burgo spring (preferentially from March to June, in Figure 2~~Figure 2~~). Since Turón river waters are less mineralized than groundwater and that the temporary storage of surface water in the nearby river dam favours water mixing, surface water dilutes groundwater from the spring (state 3 – Turón river, in Figure 2~~Figure 2~~). This occurs when the river stage is higher than groundwater level in the discharge zone, promoting water flow towards the aquifer (Barberá and Andreo, 2015).

3 Methodology

3.1 Available data

Continuous daily measurements of precipitation and air temperature were recorded at Añoreta weather station (Figure 1Figure 1) and discrete sampling campaigns for meteoric water chemistry (NO_3^- and SO_4^{-2}, among other) were performed in a rain collector installed to the north of Viento peak (Figure 1Figure 1), from August 2007 to April 2010. From meteorological data, potential evapotranspiration was calculated on a daily time scale using Thornthwaite's approach (Thornthwaite, 1948). Discontinuous measurements of Turón river flow in two selected sections (T_{up} and T_{dn} ; Figure 1Figure 1), upstream and downstream of the permanent and temporary springs, were conducted during the same study period to quantify the net groundwater discharge from Sierra Blanquilla aquifer. Simultaneously, a representative sampling of the chemical composition (NO_3^- and SO_4^{-2}) of karst groundwater was performed (daily to biweekly) at El Burgo spring. Additionally, hourly data of EC was recorded at this outlet. Detailed methodological procedures can be found in Barberá and Andreo (2015). The environmental tracers NO_3^- and SO_4^{-2} were chosen as complementary time series for the model development because they are expected to provide distinctive chemical signatures that characterize flow and transport processes in the soil and epikarst (nitrogen cycling) and saturated zone (dissolution of evaporites at the aquifer basement) of Sierra Blanquilla aquifer. Table 1Table 1 provides a summary of all available data. In addition, the information about the three differentiated flow states of the system (subsection 2.2) was used to provide an independent consideration of observations that can be attributed to time periods of state 1 (unsaturated zone), state 2 (saturated zone), and all states including the period influenced by Turón River dynamics (state 3).

Table 1: Main characteristics of the time series of hydrodynamic and hydrochemical data used in this study.

Sampling site	Parameter	Unit	n	Max	Min	Mean	CV (%)	Average sampling frequency	Period
Añoreta weather st.	Rainfall (accumulated)	mm·day^{-1}	959	71	0	3.3	-	1 day	16/08/2007 - 31/03/2010
	Air temperature (daily mean)	ºC	959	14.9	2.6	8	-	1 day	16/08/2007 - 31/03/2010
Viento rain collector	NO_3^-	mg·l^{-1}	38	23	0	3	2	15 days *	04/10/2007 - 16/02/2010
	SO_4^{-2}	mg·l^{-1}	38	4.9	0.3	1.2	1	15 days *	04/10/2007 - 16/02/2010
Turón river	Discharge (GW component)	m^3·s^{-1}	132	18,5	0.06	1,63	169	7 days	16/08/2007 - 30/03/2010
El Burgo spring	Electrical conductivity (EC)	µS·cm^{-1}	17,296	384	288	326	3	1 hour	07/11/2007 - 15/04/2010
	NO_3^-	mg·l^{-1}	130	21.2	0.8	5.1	56	8 days	01/08/2007 - 30/03/2010
	SO_4^{-2}	mg·l^{-1}	130	24.4	4.2	11.4	49	8 days	01/08/2007 - 30/03/2010

(*) Sampling frequency was dependent on the occurrence of rainfall episodes

3.2 The model

VarKarst model was previously developed at a neighbour karst system in Southern Spain (Hartmann et al., 2013b) and it was successfully applied at different karst systems around Europe (Brenner et al., 2016; Hartmann et al., 2013a, 2014b, 2016). It

includes the variability of karst system properties by statistical distribution functions (Figure 3~~Figure 3~~). Explicitly, it considers the spatial variability of (i) soil and epikarst depths, (ii) fractions of concentrated and diffuse recharge to the groundwater, (iii) epikarst hydrodynamics, and (iv) groundwater hydrodynamics by distribution functions that are applied to a set of *N* model compartments. This allows the simulation of variably dynamic pathways of water and solutes through the

5   karst system. Solute transport simulations within the model follow the assumption of instantaneous and complete mixing within each storage (soil, epikarst, groundwater) and each of the N model compartments (Figure 3). In the particular case of NO_3^-, this implies neglecting plant uptake and release processes, which were found to be important in more humid regions (Hartmann et al., 2016) but it was found a valid assumption at Mediterranean regions such as our study site (Hartmann et al., 2013b, 2014b). The detailed equations of the model in the appendix and a list of all model parameters including their

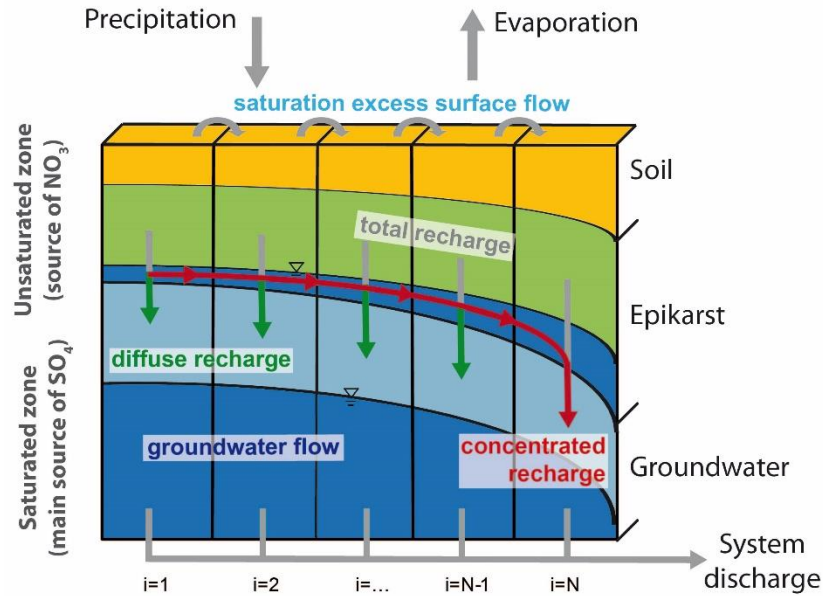10   description are provided in Table 2~~Table 2~~.

Figure 3: Schematic representation of the VarKarst model structure (modified from Hartmann et al., 2013a; modified)

3.3   Parameter estimation for the distinctive flow states and different observation types

The low resolution of observed discharge and hydrochemistry, as well as the complex karstic setting of the study site creates

15   a rather uncertain environment for modeling. For that reason, a traditional multi-objective parameter estimation was omitted as in previous studies (Hartmann et al., 2013b, 2016). Instead, a parameter estimation scheme considering "soft rules" was used to confine a large uniformly sampled set of model parameters there~~re~~-fore explicitly allowing for some uncertainty to remain but to be quantified. A similar approach was already applied successfully ~~for cases with similarly complex modelling domains: another karst system in Southern Spain (Mudarra et al., n.d.) and~~in the frame of a large-scale karst groundwater

20   recharge study (Hartmann et al., 2015, 2017).

As a measure of performance, the Kling-Gupta efficiency KGE (Gupta et al., 2009) is used. It is defined to show numbers approaching 1 for the best simulations:

$$KGE = 1 - \sqrt{(r-1)^2 + (\alpha-1)^2 + (\beta-1)^2} \qquad (1)$$

with $\alpha = \dfrac{\sigma_S}{\sigma_O}$ and $\beta = \dfrac{\mu_S}{\mu_O}$. $\qquad (2)$

r expresses the linear correlation coefficient between simulations and observations, while μ_S/μ_O and σ_S/σ_O are defined as the mean and standard deviation of simulations and observations, respectively. Consequently, α expresses the similarity of simulated and observed variability, while β quantifies the bias between them. For the calculation of *KGE*, only time steps, at which observations are available, are considered. Hence, the *KGE* values will only express the model performance to reflect the discharge, NO_3^- and SO_4^{2-} observations that were sampled in a 7-8 days temporal resolution (Table 1) even though the model runs on a daily time step.

Table 2: Description of model parameters, ranges for parameters estimation and average values found for the combined unsaturated and saturated flow states, and the entire time period including the disinformative period of river influence.

Parameter	Description	Unit	Parameter ranges		Combined unsat. and sat. states	All states
			lower	upper	mean*	mean*
A	Recharge area	[km²]	30	80	55.5	57.5
V_S	Mean soil storage capacity	[mm]	0	250	159.9	174.1
V_E	Mean epikarst storage capacity	[mm]	0	250	23.5	75.8
a_{SE}	Soil/epikarst depth variability constant	[-]	0	3	0.6	1.8
K_E	Epikarst mean storage coefficient	[d^{-1}]	15	65	49.4	43.4
a_f	Recharge separation variability constant	[-]	0	3	1.4	1.3
K_C	Conduit storage coefficient	[d^{-1}]	1	25	5.7	12.4
a_{GW}	Groundwater variability constant	[-]	0	3	1.8	1.3
c_{SO4}	Mean equilibrium concentration of SO_4^{2-}	[mg l^{-1}]	0	100	16.6	22.0
a_{SO4}	SO_4^{2-} variability constant	[-]	0	3	0.6	1.4
KGE$_Q$	performance concerning discharge	[-]	0	1	0.37	0.36
KGE$_{NO3}$	performance concerning NO_3	[-]	0	1	0.47	0.32
KGE$_{SO4}$	performance concerning SO_4^{2-}	[-]	0	1	0.58	0.40

* variability of model parameters shown in Figure 5Figure 5

For parameter estimation, an initial sample of 2500,000 parameter sets was created from predefined ranges (Table 2Table 2) that were chosen by prior knowledge and previous model experiences in the same region (Hartmann et al., 2013b, 2014b). A 4-year warm up period was set up and the model was run 2500,000 times with the initial parameter sample. Using the observed time series, the Kling Gupta Efficiency was calculated for each of the simulation runs: KGE$_Q$ (groundwater

discharge), KGE_{NO3} (NO_3^- concentrations) and KGE_{SO4} (SO_4^{-2} concentrations). Similar to Choi and Beven (2007) "soft rules" were used to reduce the initial sample of parameters in four steps:

1. All parameters sets from the initial sample with $KGE_Q < 0.2$ were discarded
2. All parameters sets from the initial sample with $KGE_{NO3} < 0.2$ were discarded
3. All parameters sets from the initial sample with $KGE_{SO4} < 0.2$ were discarded
4. All parameters sets from the initial sample with KGE_Q, KGE_{NO3}, and $KGE_{SO4} < 0.2$ at the same time were discarded

The threshold value of 0.2 was found by preliminary analysis. Its rather low value is meant to take into account that the simulation is exposed to various sources of uncertainty including uncertainties of the model input (observation of climate variables and their application to the entire recharge area), model structure uncertainty (representation of karst processes by conceptual mathematical formulations in a semi-distributed way), and the uncertainty of observations (discharge measurement and hydrochemical analysis, as well as their low temporal resolution).

The ~~same procedure~~application of the soft rules is repeated four times for observations falling into the unsaturated flow state, the saturated flow state, the combined unsaturated and saturated flow state and into the entire time period including the hydrodynamic state defined by influence of Turón river flow on groundwater discharge. For each of these time periods the four soft rules will result in a reduction of the initial sample and the prior ranges of the model parameters will experience a confinement (Hartmann et al., 2015).

3.4 Evaluation of information content and simulation uncertainty for the different flow states and different observation types

In this study, the strength of this confinement is used to assess the information content of the set of observations during the different flow states. The strength of the confinement is quantified by the reduction of the distance between the 25th and 75th percentile of each model parameter after the confinement through the soft rules. For instance, parameter c_{SO4} (Table 2~~Table 2~~) has the prior range of $0 – 100$ mg·l^{-1}. Consequently, the uniform sampling strategy for the initial sample will result in values close to 25 and 75 mg·l^{-1} for the 25th and the 75th percentile, respectively. Applying one of the soft rules may now result in values of 10 and 30 mg·l^{-1} for the 25th and the 75th percentile, respectively. Hence, the reduction of the distance between the 25th and 75th percentile is 50-20 mg·l^{-1}, i.e. a reduction of 60% took place. In this example case we would find that the observations applied through the selected soft rule provided useful information to estimate this parameter. Applying this procedure for each of the four soft rules and the four time series defined by the flow states, we can assess how (1) the different types of observations (discharge, NO_3^- and SO_4^{-2}) contribute to parameter identification, and (2) the focus on particular time periods and flow stages strengthens or weakens the confinement of the model parameters.

Particular attention is given to the comparison of the entire time period, including the times when the spring is influenced by the river, with the time periods when only the unsaturated zone and the saturated zone control the discharge of the spring. It is expected that this time period contains disinformative information for parameter estimation as the VarKarst
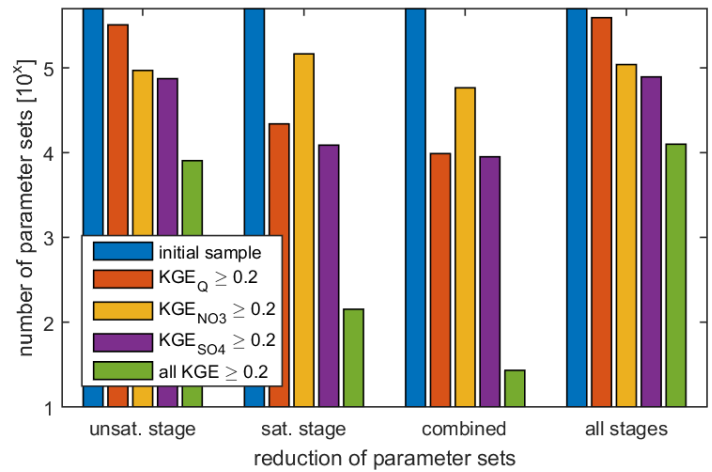
11

does not take into account the river's influence. The reduction of the 25th and 75th percentile of the model parameters is used after applying the fourth soft rule (subsection 3.3) of the combined unsaturated and saturated flow state, and the entire time period including the period that is influenced by the river to understand the impact of the disinformative information on parameter identification. In a last step, the simulation uncertainty is quantified for the two time periods by plotting the simulations of the parameters sets that remained after the fourth soft rule was applied to the two observation time series. After including the disinformative time period, a greater simulation uncertainty is expected.

4 Results

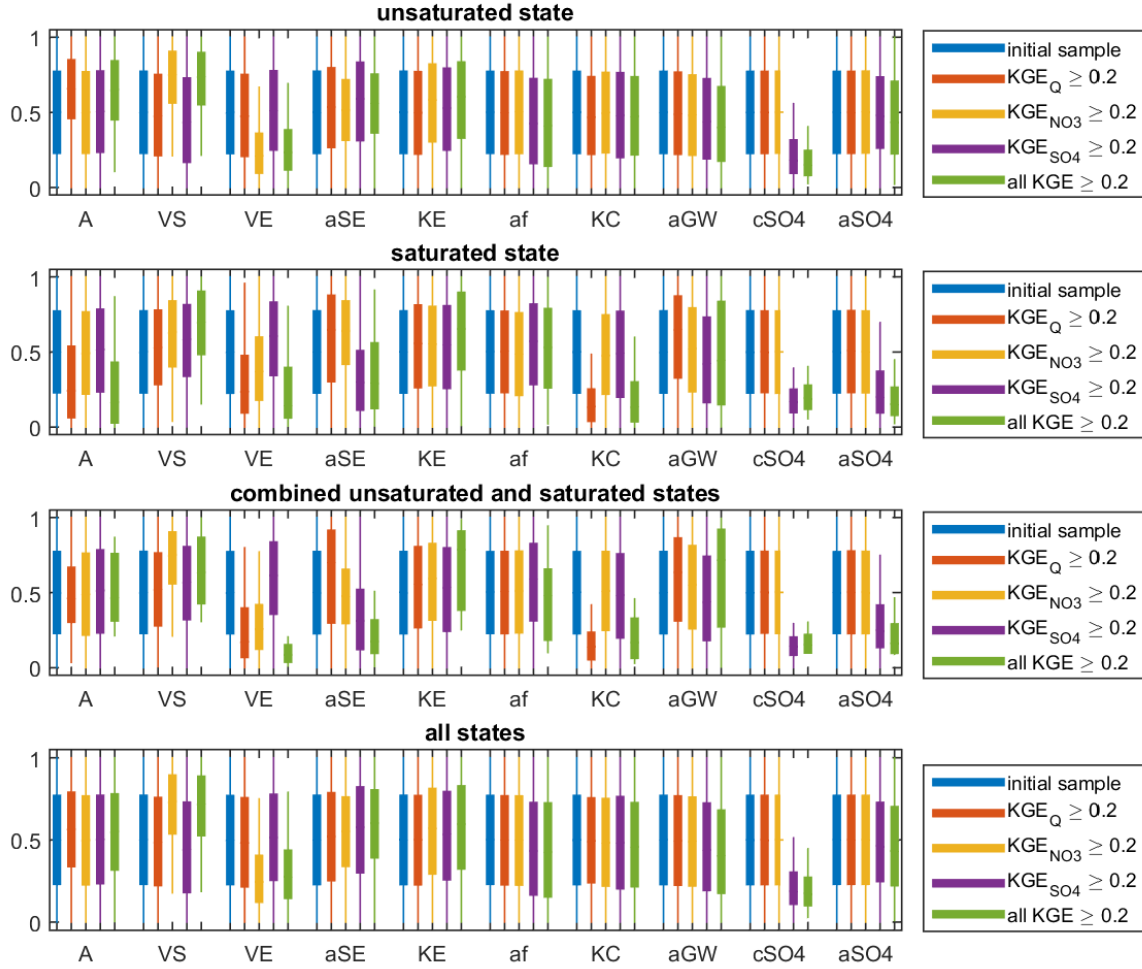### 4.1 Parameter estimation for the different flow states and different observation types

Different reductions of the initial sample are found by the different soft rules and during the different flow states (Figure 4Figure 4). The reduction by discharge ($KGE_Q \geq 0.2$) varies among the different flow states but remains rather limited. The same is seen for the individual use of the hydrochemical information ($KGE_{NO3} \geq 0.2$ or $KGE_{SO4} \geq 0.2$). However, using the combination of all soft rules (all $KGE \geq 0.2$), a significant reduction of the initial sample is obtained for all flow states. This is most evident for the combined unsaturated and saturated state. The weakest reduction of the initial sample for all soft rules is found for the consideration of all stages including the disinformative time period influenced by the river.

Figure 4: Reduction of the initial sample by the four soft rules for the unsaturated state, saturated state, combined saturated and saturated states, and all system states

The influence of the soft rules during the different flow states varies for all model parameters (Figure 5Figure 5). The reduction of the initial sample by discharge ($KGE_Q \geq 0.2$) alters the uniform distribution of the initial sample for the different flow states, mostly for the parameters A, V_E and K_C. These changes are most prominent in the unsaturated state (A), the unsaturated state (V_E and K_C) and the combined unsaturated and saturated states (A, V_E and K_C). Using NO_3^- for the reduction ($KGE_{NO3} \geq 0.2$), the parameters V_S, V_E and a_{SE} experience the strongest change of their initial distribution. This change is

12

most pronounced at the unsaturated state and the combined unsaturated and saturated states. The reduction by the observations of SO_4^{-2} concentrations ($KGE_{SO4} \geq 0.2$) mostly affects the model parameters c_{SO4} and a_{SO4}, but also find a strong impact on a_{SE}, mainly at the saturated state and the combined unsaturated and saturated state. Finally applying all information in the fourth soft rule (all $KGE \geq 0.2$), we find again an alteration of the model parameters that were affected by soft rules 1-3 (A, V_E, V_S, a_{SE}, K_C, c_{SO4} and a_{SO4}) and, additionally, a moderate alteration of K_E and a_f. This is most notable at the combined unsaturated and saturated states; using all states including the disinformative period that is influenced by the river, the alterations are generally less pronounced.
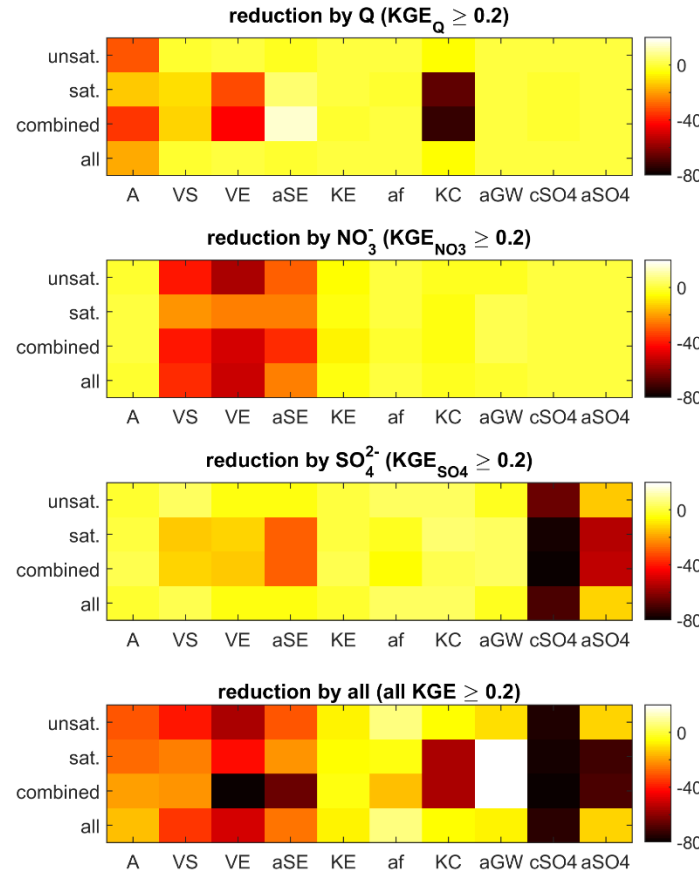

Figure 5: Distribution of model parameters (normalised by their ranges) after applying the four soft rules for the unsaturated, saturated, and combined unsaturated and saturated, and all stages.

4.2 Evaluation of information content and simulation uncertainty for the distinctive flow states and different observation types

Using the change of distance between the 25th and 75th percentile of each model parameter for the different soft rules and the different flow states we are able to quantify the information content of the available observations (Figure 6~~Figure 6~~). We find that discharge ($KGE_Q \geq 0.2$) and SO_4^{2-} ($KGE_{SO4} \geq 0.2$) provide most information during the saturated flow state, while NO_3^- reduces the distance between the two percentiles most during the unsaturated stage. The state that uses all information including the disinformative time period of river influence shows generally the weakest reduction between the 25th and 75th percentile as already indicated by Figure 4~~Figure 4~~.

Again the most evident changes of model parameter distributions are found for the combined unsaturated and saturated states: here, we see that observed discharge ($KGE_Q \geq 0.2$) provides most information on the parameter K_C, but the change of distance for the parameters A and V_E is still considerable. This is most evident in the combined unsaturated and saturated states. We find a more balanced distribution of information on the altered parameters when regarding the reduction obtained by NO_3^- ($KGE_{NO3} \geq 0.2$). Here, the change of distances is considerable (but similar) for V_S, V_E and a_{SE}. For SO_4^{-2} ($KGE_{SO4} \geq 0.2$), the alteration mostly affects c_{SO4}, followed by a considerable alteration of a_{SO4} and a moderate change of a_{SE}. Using all information to confine the initial sample (all $KGE \geq 0.2$) shows that the combined use of discharge, NO_3^- and SO_4^{-2} observations provided most information on V_E, a_{SE}, K_C, c_{SO4}, and a_{SO4}. Still considerable information is provided for A, V_S and a_f. However, no reduction of the distance between the 25th and 75th percentile is found for K_E, and even a widening takes place for a_{GW}.

Figure 6: Change of distance between the~~ he~~ 25th and 75th percentile of each model parameter when the different soft rules are applied (top to bottom) for the four flow states.

5   The proceeding analysis indicates that most information to identify the largest number of model parameters is provided by the combined unsaturated and saturated flow states using discharge, NO_3^- and SO_4^{-2} observations. It further reveals that using the entire time period, using discharge, NO_3^- and SO_4^{-2} observations and including the period that is influenced by the river, provided the fewest information; only 5 (A, V_S, V_E, a_{SE} and c_{SO4}) of the 10 model parameters show a detectable reduction of the two flow percentiles (Figure 6~~Figure 6~~, bottom).

10 The final averages of the estimated parameters (after applying 4th soft rule; Table 2~~Table 2~~) of the combined unsaturated and saturated flow states, and the state the uses the entire set of observations are similar for the parameters A, V_S and c_{SO4}, while there is a strong difference for V_E and a_{SE}. Comparing furthermore the resulting simulation uncertainty (Figure 7~~Figure 7~~), we find that the final parameter~~s~~ set~~s~~ that were found using the entire observed time series~~used all flow states~~, including the disinformative, river influences time period, results in a larger simulation~~s~~ uncertainty than the final parameter sample that

15 used only the combined unsaturated and saturated flow states for parameter estimation.
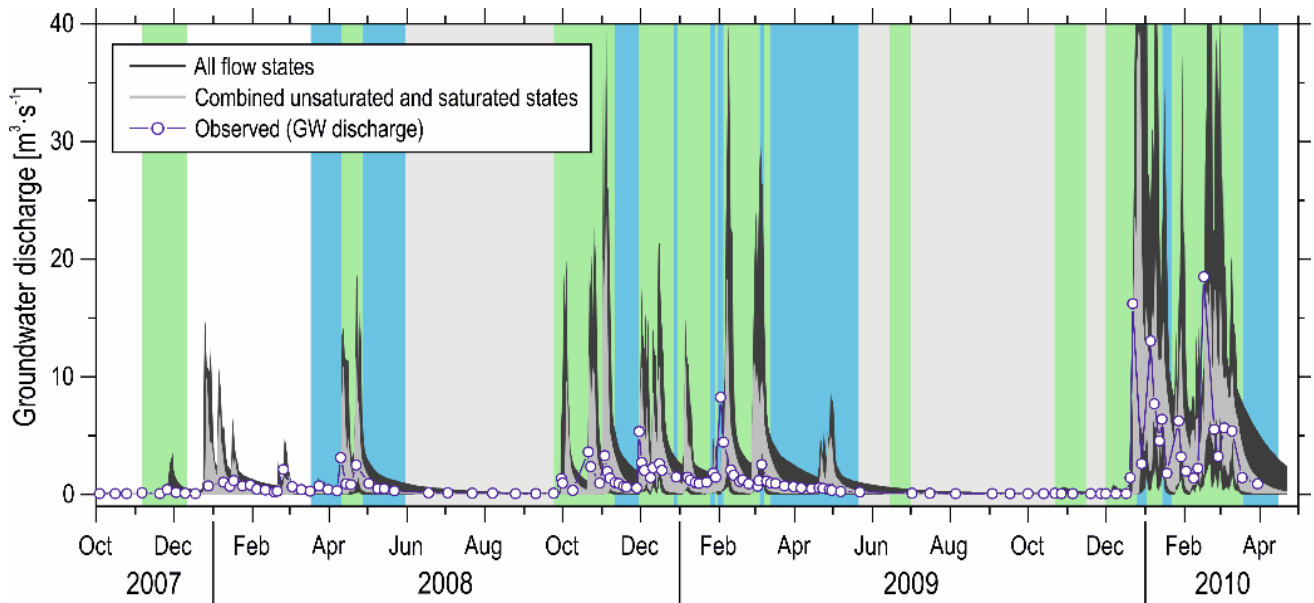
15

Figure 7: Observed discharge and the simulation uncertainty of the final parameter sample (all KGE ≥ 0.2) of the combined unsaturated and saturated flow states and the all flow states including the disinformative period of river influence. Background colours representing flow states match to that of the ~~Figure 2~~Figure 2.

5 Discussion

5.1 Application of the soft rules during the different flow states

The application of the 4 soft rules results in a general reduction of the initial sample for all flow states (Figure 4~~Figure 4~~). A weak reduction for all of the four flow states takes place when only discharge observations are applied to confine the sample. Previous research with lumped model calibration showed that the information content of discharge observations usually suffices to calibrate 5-6 parameters (Jakeman and Hornberger, 1993; Wheater et al., 1986; Ye et al., 1997); more parameters often lead to over-parametrization (Perrin et al., 2003) and equifinality (Beven, 2006). Hence, the small reduction of the VarKarst initial parameter sample may be due to the large number of model parameters (Table 2~~Table 2~~) with in the VarKarst model. The same behaviour of a weak decrease of the initial parameter sample is found when the hydrochemical observations are used individually (soft rule 2 and 3). The weakest reduction of the initial parameter sample among all four flow states is found for the entire time period that includes the periods of river influence (see discussion in subsection 5.2). When soft rule 4 (all KGE ≥ 0.2) is applied, we find the strongest reduction of the initial sample across all of the four flow states. This means, the combined information of discharge, NO_3^- and SO_4^{-2} observations provides the most information to reduce the initial sample of model parameters. Previous research already showed that hydrochemical information can reduce parameter uncertainty (Kuczera and Mroczkowski, 1998; Rimmer and Hartmann, 2014; Son and Sivapalan, 2007). In this study, a similar reduction of parameter uncertainty could be observed (Figure 5~~Figure 5~~). Depending on the applied soft rule

16

and the considered flow states the initially uniform distributions of the model parameters are altered differently. Some model parameters distributions change their mean without much change in the shape of their distribution (same distance between 25th and 75th percentile); some of the show a more confined distribution when the soft rules are applied.

5.2 Information provided by discharge and hydrochemistry during the different flow states

5   The differences of the reduction across the model parameters reveals the influence of different types of observations that were used for parameter estimation. We find that the reductions of the distance between the 25th and 75th percentile is most pronounced during the saturated stage for the discharge observations (Figure 6Figure 6). This indicates that discharge provides most information during the recession period. Information about hydrodynamic parameters A, V_E and K_C is derived directly from the discharge observations. This makes sense because hydrodynamic changes in the main discharge area of

10   Sierra Blanquilla aquifer reflect the hydraulic pressure transference from the unsaturated zone to the saturated zone of the system. Similar results were found by Wagener et al. (2003) when they applied Dynamic Identifiability Analysis to a lumped rainfall runoff model using only discharge data.

They also found that the parameters, which control the unsaturated zone and fast flow components of their model, are most identifiable during and just after the rainfall-runoff events. Our results indicates a similar behaviour by showing the strongest

15   reduction of the distance between the 25th and 75th percentile for the unsaturated zone parameters during the unsaturated flow state using the NO_3^- observations (parameters V_S, V_E and a_{SE}). This is in accordance with Reusser and Zehe (2011) who showed that model parameters that control the recession period are most sensitive during the recession period with a time dynamic resection and cluster analysis using discharge information. NO_3^- has been used almost as an ideal tracer to determine infiltration processes through the soil and epikarst in the shallower aquifer zones (Hunkeler and Mudry, 2007;

20   Mudarra et al., 2014). Thus, NO_3^- observations contribute stronger to the identification of surface and evapotranspiration processes during the unsaturated flow state. This can be explained by the relative stability of NO_3^- dynamics within the karst system under oxidizing conditions (Mudarra et al., 2014), which favours its preservation from surface to the spring.

SO_4^{-2} provided most information on the parameters c_{SO4}, a_{SO4}, and a_{SE} during the saturated state. This makes sense as SO_4^{-2} is stored within saturated zone of the system where groundwater is in touch with gypsum-bearing geological formations

25   (Triassic clays with evaporites), which are found in contact with deeper aquifer compartments. SO_4^{-2} time series provide more information about the unsaturated zone / epikarst drainage during the saturated flow stage. Such findings mean that the high chemical contrast observed in SO_4^{-2} concentrations of fresh (recently infiltrated) and old (stored) groundwater is useful to assess the relative importance of unsaturated flow and saturated flow during saturated flow stage (Barberá and Andreo, 2015; Mudarra et al., 2011).

30   The highest number of identifiable parameters is found when all information (discharge, NO_3^- and SO_4^{-2} observations) are combined and estimated during the combined unsaturated and saturated flow stages (Figure 6Figure 6). In addition to the parameters that showed an increase of identifiability at the individual stages (A, V_E, K_C, V_S, V_E , a_{SE}, c_{SO4}, and a_{SO4}), we also

17

see an increased identifiability of the parameter a_f, most probably due to parameter interactions (Pianosi et al., 2016). Only a_{GW} and KE remain with low identifiability, which may be due to structural limitations of the model structure (Clark et al., 2008) or due to parameter interactions that are not explicitly considered in our approach. In fact, a lower identifiability of K_E in favour of a high identifiability of VE was found in a previous study with a similar version of the model (Hartmann et al.,

5   2015). Compared to that, using all information during the entire time period, including the disinformative period, only five of the model parameters show a visible decrease of the distance between the 25th and 75th percentile of their distribution. Hence, the inclusion of the disinformative period led to an increase of posterior parameter uncertainty compared to using only the informative time periods represented by the unsaturated and saturated states. This was also shown by Beven and Westerberg (2011) or Beven et al. (2011), when they considered the impact disinformative discharge events.

10   The impact of the disinformative time period on the precision of the observations is clearly visible in Figure 7Figure 7. Since the model has to compensate for structural errors, i.e. the missing representation of the influence of the river on the discharge of the karst spring, it is forced to allow for a wider range of parameter combinations to account for the simulation errors. Using only the unsaturated and saturated states allows for a much better confinement of model parameters and therefore a much smaller simulation uncertainty, although showing some deviations during the periods when the fiver affects the flow

15   system of the spring (blue shaded areas in Figure 7Figure 7). Hence, similar to Kauffeldt et al. (2013), our study shows that a proper pre-analysis of the information content observations for model parameters estimation (subsection 2.2) allows for excluding disinformative information to reduce model parameter and simulation uncertainty.

5.3   Limits and transferability of the approach

The analysis of variations of the groundwater component in the Turón River flow has permitted to determine the timing,

20   duration and magnitude of the global hydrodynamic aquifer responses under influenced hydrological conditions, as well as to assess the discharge thresholds from which different compartments (i.e. flooding of relict conduit networks) of the behaviour system activate (i.e. flooding of relict conduit networks; (Barberá and Andreo, 2015). However, a more accurate decomposition of flow components from the study of spring hydrographs has not been possible due to the relatively low resolution of discharge time series (Table 1Table 1). Even though the chemical signature of groundwater that drains the

25   different aquifer zones (unsaturated zone and saturated zone), and that is affected by the Turón river, can be estimated using electric conductivity (Figure 2Figure 2), it is rather based on subjective interpretation. However, it can be argued that the previous knowledge based on the accurate interpretation of El Burgo spring chemographs has permitted a realistic flow decomposition from EC time series as our results show a clear difference of estimated parameter distributions and resulting simulation uncertainty using the unsaturated and saturated flow states and the entire time period including the disinformative

30   data. A more precise distinction between the states is only possible if specific chemical indicators are available better constrain the differentiation of flow states contributing to El Burgo spring discharge, which was not possible within the frame of this study. But even though due to subjectivity, the identification of time periods or data sets that disinformative

contributions to parameter estimation is a useful way to reduce the simulation uncertainty of hydrological models. Building on previous research on disinformative data that focussed on disinformative discharge information, our approach provides a systematic procedure that also includes hydrochemical observations to identify disinformative periods and to improve parameter estimation of models for complex hydrological systems. Another limitation of our research is the low resolution of the discharge and hydrochemical observations (7-8 days). Although our approach took into account this weakness by the soft rules allowing for remaining uncertainty after the reduction of our 500,000 parameter sets, we believe that a higher resolution of the observations (preferably 1 day) would have resulted in a more pronounced reduction of the initial sample and consequently to a lower remaining uncertainty.

6 Conclusions

In this research, a new approach to estimate the information content of water quality data and the value of identifying most informative periods for model parameter estimation has been proposed. Using soft rules to include discharge, NO_3^- and SO_4^{-2} observations into the parameter estimation procedure, we were able to reduce an initial sample of 500,000 parameter sets during pre-defined flow states; one ~~of the including~~being a known period of disinformative data. Comparing the distributions of the initial and reduced parameter sets, we were able to quantify the information contained in our observations to identify the parameters of our simulation model.

We found that the information content of the observations varies for the different states that we considered. NO_3^- provided most of its information when the unsaturated zone processes dominate the discharge behaviour of the spring. During the time when the saturated zone controls the outflow behaviour, SO_4^{-2} and discharge observations provide the best information to identify the model parameters. Including the disinformative period, the information content of all data generally decreases, as well as the uncertainty of simulations increases. We finally show that the combination of saturated and unsaturated flow states provides the most precise information about the model parameters. Due to parameter interactions, even model parameters that were not identifiable during the unsaturated or saturated flow state alone became identifiable. As a result, the simulation uncertainty is significantly reduced compared to the simulations obtained by the entire time series of observations that include the disinformative data.

Even though exemplified at a particular karst spring in Southern Spain, our approach is easily transferrable to other modelling studies that want to use water quality data for the identification of disinformative periods and for the estimation of model parameters. Our results add to previous findings on the value of removing disinformative data from model parameter estimation to reduce simulation uncertainty. Furthermore our results can help building a better communication between experimental hydrologists and modellers (Hartmann, 2016; Seibert and McDonnell, 2002) as hydrochemical data is often used for system characterization. Our study showed that NO_3^- and SO_4^{-2} that often are used for understanding the unsaturated and saturated zone processes also help to identify the corresponding process parameters in our model. Further research should therefore include the evaluation of other hydrochemical variables that can be attributed to particular hydrological processes to and their value to identify the corresponding processes in process-based simulation models. Also, a further

7 Acknowledgements

8 Appendix

The parameter $V_{mean,S}$ [mm] and the distribution coefficient a_{SE} [-] control the variability of soil depths over the N model compartments. Using them, the soil storage capacity $V_{S,i}$ [mm] for every compartment i is defined by:

$$V_{S,i} = V_{max,S} \cdot \left(\frac{i}{N}\right)^{a_{SE}} \tag{3}$$

$V_{max,S}$ [mm] represents the maximum soil storage capacity and is derived from V_S:

$$\int_0^{i_{1/2}} V_{max,S}\left(\frac{x}{N}\right)^{a_{SE}} dx = \frac{\int_0^N V_{max,S}\left(\frac{x}{N}\right)^{a_{SE}} dx}{2} ; V_S = V_{max,S}\left(\frac{i_{1/2}}{N}\right)^{a_{SE}}$$
$$\Updownarrow$$
$$V_{max,S} = V_S \cdot 2^{\left(\frac{a_{SE}}{a_{SE}+1}\right)} \tag{4}$$

Where the compartment at which the volumes on the left equal the volumes on the right is found at $i_{1/2}$. The same distribution coefficient a_{SE} is used to derive the epikarst storage distribution by the mean epikarst depth V_E [mm] (derivation of $V_{max,E}$ likewise to $V_{max,S}$ in Eq (4)):

$$V_{E,i} = V_{max,E} \cdot \left(\frac{i}{N}\right)^{a_{SE}} \tag{5}$$

Actual evapotranspiration from each soil compartment $E_{act,i}$ is calculated by:

$$E_{act,i}(t) = E_{pot}(t) \cdot \frac{\min\left[V_{Soil,i}(t) + P(t) + Q_{Surface,i}(t), V_{S,i}\right]}{V_{S,i}} \tag{6}$$

Potential evapotranspiration E_{pot} [mm] is found by the Thornthwaite equation (Thornthwaite, 1948) and $Q_{surface,i}$ [mm] is the surface inflow that originates from compartment i-1 (see Eq. (11)). $V_{Soil,i}$ [mm] is the volume of water stored in the soil at time step t. Recharge from the soil to the epikarst $R_{Epi,i}$ [mm] is found by water balance:

$$R_{Epi,i}(t) = Q_{\inf}(t) + \max\left[V_{Soil,i}(t) + P(t) + Q_{Surface,i}(t) - E_{act,i}(t) - V_{S,i}, 0\right] \tag{7}$$

with $Q_{\inf}(t)$ being the river infiltration (Eq. (5)). The epikarst storage coefficients $K_{E,i}$ [d] controls the outflow from the epikarst:

$$Q_{Epi,i}(t) = \frac{\min\left[V_{Epi,i}(t) + R_{Epi,i}(t), V_{E,i}\right]}{K_{E,i}} \cdot \Delta t \tag{8}$$

$$K_{E,i} = K_{\max,E} \cdot \left(\frac{N - i + 1}{N}\right)^{a_{SE}} \tag{9}$$

Here, $V_{Epi,i}$ [mm] is the water stored in the epikarst at time step t. $K_{\max,E}$ is found by the mean epikarst storage coefficient K_E and by applying the same distribution coefficient a_{SE}:

$$N \cdot K_E = \int_0^N K_{\max,E}\left(\frac{x}{N}\right)^{a_{SE}} dx$$

$$\Updownarrow \tag{10}$$

$$K_{\max,E} = K_E \cdot (a_{SE} + 1)$$

Surface flow to the next model compartment $Q_{Surf,i+1}$ [mm] initiates when soil and epikarst storage capacities are exceeded:

$$Q_{Surf,i+1}(t) = \max\left[V_{Epi,i}(t) + R_{Epi,i}(t) - V_{E,i}, 0\right] \tag{11}$$

The vertical percolation from the epikarst is split into diffuse ($R_{diff,i}$ [mm]) and concentrated groundwater recharge ($R_{conc,i}$ [mm]) again by a variable separation factor $f_{C,i}$ [-] and a distribution coefficient a_f [-]:

$$R_{conc,i}(t) = f_{C,i} \cdot Q_{Epi,i}(t) \tag{12}$$

$$R_{diff,i}(t) = (1 - f_{C,i}) \cdot Q_{Epi,i}(t) \tag{13}$$

$$f_{C,i} = \left(\frac{i}{N}\right)^{a_f} \tag{14}$$

The diffuse recharge reaches the groundwater compartments ($i = 1 \ldots N\text{-}1$) directly below, while concentrated recharge is routed laterally to the conduit system (compartment $i = N$, :). Similar to epikarst storage coefficients, variable groundwater storage coefficients $K_{GW,i}$ [d] are calculated. The, groundwater contributions of the matrix system $Q_{GW,i}$ [mm] in therefore found by:

$$Q_{GW,i}(t) = \frac{V_{GW,i}(t) + R_{diff,i}(t)}{K_{GW,i}}; i = 1 \ldots N - 1 \tag{15}$$

with

$$K_{GW,i} = K_C \cdot \left(\frac{i}{N}\right)^{-a_{GW}} \qquad (16)$$

The conduit system discharges from compartment N:

$$Q_{GW,i}(t) = \frac{V_{GW,N}(t) + \sum_{i=1}^{N} R_{conc,i}(t)}{K_C} ; i = N \qquad (17)$$

where conduit storage coefficient is given by K_C [d]. The discharge of the main spring Q_{main} [l s^{-1}] is comprised by the sum of the matrix and the conduit system discharge rescaled to [l s^{-1}] the recharge area A [km²]:

$$Q_{main}(t) = \frac{A}{N} \cdot \sum_{i=1}^{N} Q_{GW,i}(t) \qquad (18)$$

Solute transport within the VarKarst model follows the assumption of complete mixing for every model compartment. Hence, enrichment only takes place due to evaporation and by geogene dissolution (only SO_4^{2-}), for which varying equilibrium concentrations are defined according to:

$$c_{SO4,i} = c_{\max,SO4} \cdot \left(\frac{N-i+1}{N}\right)^{a_{SO4}} \qquad (19)$$

where a_{SO4} is a variability constant and $c_{max,SO4}$ is derived from c_{SO4} [mg l^{-1}] (similar to eq. (10)).

9   References

Aquilina, L., Ladouche, B., Dörfliger, N. and Doerfliger, N.: Water storage and transfer in the epikarst of karstic systems during high flow periods, J. Hydrol., 327(3), 472–485, 2006.

Bakalowicz, M.: Karst groundwater: a challenge for new resources, Hydrogeol. J., 13(1), 148–160, doi:10.1007/s10040-004-0402-9, 2005.

Barberá, J. A.: Hydrogeological research in the carbonate aquifers of eastern Serranía de Ronda (Málaga) in Spanish and English, PhD thesis at the Centre of Hydrogeology of the University of Málaga CEHIUMA (Spain)., 2014.

Barberá, J. A. and Andreo, B.: Hydrogeological characterisation of two karst springs in Southern Spain by hydrochemical data and intrinsic natural fluorescence, Groundw. Qual. Sustain., 281, 2012.

Barberá, J. a. A. and Andreo, B.: Hydrogeological processes in a fluviokarstic area inferred from the analysis of natural hydrogeochemical tracers. The case study of eastern Serranía de Ronda (S Spain), J. Hydrol., 523, 500–514, doi:10.1016/j.jhydrol.2015.01.080, 2015.

Beven, K. and Westerberg, I.: On red herrings and real herrings: Disinformation and information in hydrological inference,

Hydrol. Process., 25(10), 1676–1680, doi:10.1002/hyp.7963, 2011.

Beven, K., Smith, P. J. and Wood, A.: On the colour and spin of epistemic error (and what we might do about it), Hydrol. Earth Syst. Sci., 15(10), 3123–3133, doi:10.5194/hess-15-3123-2011, 2011.

Beven, K. J.: A manifesto for the equifinality thesis, J. Hydrol., 320(1–2), 18–36, 2006.

5 Brenner, S., Coxon, G., Howden, N. J. K., Freer, J. and Hartmann, A.: A percentile approach to evaluate simulated groundwater levels and frequencies in a Chalk catchment in Southwest England, Nat. Hazards Earth Syst. Sci. Discuss., (December), doi:doi:10.5194/nhess-2016-386, 2016.

Charlier, J.-B., Bertrand, C. and Mudry, J.: Conceptual hydrogeological model of flow and transport of dissolved organic carbon in a small Jura karst system, J. Hydrol., 460–461, 52–64, doi:10.1016/j.jhydrol.2012.06.043, 2012.

10 Choi, H. T. and Beven, K.: Multi-period and multi-criteria model conditioning to reduce prediction uncertainty in an application of TOPMODEL within the GLUE framework, J. Hydrol., 332(3–4), 316–336, doi:10.1016/j.jhydrol.2006.07.012, 2007.

Clark, M. P., Slater, A. G., Rupp, D. E., Woods, R. A., Vrugt, J. A., Gupta, H. V, Wagener, T. and Hay, L. E.: Framework for Understanding Structural Errors (FUSE): A modular framework to diagnose differences between hydrological models, 15 Water Resour. Res., 44(W00B02), doi:10.1029/2007WR006735, 2008.

Cruz-Sanjulián, J. J.: Estudio geológico del sector Cañete la Real-Teba-Osuna (Cordillera Bética, región occidental), Universidad de Granada., 1974.

Fernández, R.: Investigaciones hidrogeológicas al Norte de Ronda (Málaga), Univ. de Granada., 1980.

Ford, D. C. and Williams, P. W.: Karst Hydrogeology and Geomorphology, John Wiley & Sons., 2013.

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

Goldscheider, N. and Drew, D.: Methods in Karst Hydrogeology, edited by I. A. of Hydrogeologists, Taylor & Francis Group, Leiden, NL., 2007.

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

Hartmann, A.: Putting the cat in the box: why our models should consider subsurface heterogeneity at all scales, WIREs Water, doi:10.1002/wat2.1146, 2016.

30 Hartmann, A. and Baker, A.: Modelling karst vadose zone hydrology and its relevance for paleoclimate reconstruction, Earth-Science Rev., 1–54, doi:https://doi.org/10.1016/j.earscirev.2017.08.001, 2017.

Hartmann, A., Lange, J., Weiler, M., Arbel, Y. and Greenbaum, N.: A new approach to model the spatial and temporal variability of recharge to karst aquifers, Hydrol. Earth Syst. Sci., 16(7), 2219–2231, doi:10.5194/hess-16-2219-2012, 2012.

Hartmann, A., Weiler, M., Wagener, T., Lange, J., Kralik, M., Humer, F., Mizyed, N., Rimmer, A., Barberá, J. a., Andreo,

B., Butscher, C. and Huggenberger, P.: Process-based karst modelling to relate hydrodynamic and hydrochemical characteristics to system properties, Hydrol. Earth Syst. Sci., 17(8), 3305–3321, doi:10.5194/hess-17-3305-2013, 2013a.

Hartmann, A., Barberá, J. A., Lange, J., Andreo, B., Weiler, M., Antonio, J., Lange, J., Andreo, B., Weiler, M., Barberá, J. A., Lange, J., Andreo, B. and Weiler, M.: Progress in the hydrologic simulation of time variant recharge areas of karst systems – Exemplified at a karst spring in Southern Spain, Adv. Water Resour., 54, 149–160, doi:10.1016/j.advwatres.2013.01.010, 2013b.

Hartmann, A., Goldscheider, N., Wagener, T., Lange, J. and Weiler, M.: Karst water resources in a changing world: Review of hydrological modeling approaches, Rev. Geophys., 52(3), 218–242, doi:10.1002/2013rg000443, 2014a.

Hartmann, A., Mudarra, M., Andreo, B., Marín, A., Wagener, T. and Lange, J.: Modeling spatiotemporal impacts of hydroclimatic extremes on groundwater recharge at a Mediterranean karst aquifer, Water Resour. Res., 50(8), 6507–6521, doi:10.1002/2014WR015685, 2014b.

Hartmann, A., Gleeson, T., Rosolem, R., Pianosi, F., Wada, Y. and Wagener, T.: A large-scale simulation model to assess karstic groundwater recharge over Europe and the Mediterranean, Geosci. Model Dev., 8(6), 1729–1746, doi:10.5194/gmd-8-1729-2015, 2015.

Hartmann, A., Kobler, J., Kralik, M., Dirnböck, T., Humer, F. and Weiler, M.: Model-aided quantification of dissolved carbon and nitrogen release after windthrow disturbance in an Austrian karst system, Biogeosciences, 13(1), 159–174, doi:10.5194/bg-13-159-2016, 2016.

Hartmann, A., Gleeson, T., Wada, Y., Wagener, T., Kingdom, U., Sciences, O. and Kingdom, U.: Enhanced groundwater recharge rates and altered recharge sensitivity to climate variability through subsurface heterogeneity, Proc. Natl. Acad. Sci., 19, EGU2017-12796, doi:10.1073/pnas.1614941114, 2017.

Hunkeler, D. and Mudry, J.: Hydrochemical tracers, in Methods in karst hydrogeology, edited by N. Goldscheider and D. Drew, pp. 93–121, Taylor and Francis/Balkema, London, UK., 2007.

Jakeman, A. J. and Hornberger, G. M.: How much complexity is warranted in a rainfall-runoff model?, Water Resour. Res., 29, 2637–2649, 1993.

Kauffeldt, A., Halldin, S., Rodhe, A., Xu, C. Y. and Westerberg, I. K.: Disinformative data in large-scale hydrological modelling, Hydrol. Earth Syst. Sci., 17(7), 2845–2857, doi:10.5194/hess-17-2845-2013, 2013.

Kovacs, A. and Sauter, M.: Modelling karst hydrodynamics, in Methods in karst hydrogeology, edited by N. Goldscheider and D. Drew, pp. 65–91, Taylor and Francis/Balkema, London, UK., 2007.

Kuczera, G. and Mroczkowski, M.: Assessment of hydrologic parameter uncertainty and the worth of multiresponse data, Water Resour. Res., 34(6), 1481–1489, 1998.

Labat, D., Ababou, R. and Mangin, A.: Rainfall–runoff relations for karstic springs. Part I: convolution and spectral analyses, J. Hydrol., 238, 123–148, 2000.

Martín-Algarra, A.: Evolución geológica Alpina del contacto entre las Zonas Internas y las Zonas Externas de la Cordillera Bética (Sector Occidental), Universidad de Granada., 1987.

Mazzilli, N., Jourde, H., Jacob, T., Guinot, V., Moigne, N., Boucher, M., Chalikakis, K., Guyard, H. and Legtchenko, A.: On the inclusion of ground-based gravity measurements to the calibration process of a global rainfall-discharge reservoir model: case of the Durzon karst system (Larzac, southern France), Environ. Earth Sci., doi:10.1007/s12665-012-1856-z, 2012.

Mudarra, M., Andreo, B. and Mudry, J.: Monitoring groundwater in the discharge area of a complex karst aquifer to assess

5   the role of the saturated and unsaturated zones, Environ. Earth Sci., 65(8), 2321–2336, doi:10.1007/s12665-011-1032-x, 2011.

Mudarra, M., Andreo, B., Marín, a I., Vadillo, I. and Barberá, J. A.: Combined use of natural and artificial tracers to determine the hydrogeological functioning of a karst aquifer: the Villanueva del Rosario system (Andalusia, southern Spain), Hydrogeol. J., 22(5), 1027–1039, doi:10.1007/s10040-014-1117-1, 2014.

10   Oehlmann, S., Geyer, T., Licha, T. and Sauter, M.: Reducing the ambiguity of karst aquifer models by pattern matching of flow and transport on catchment scale, Hydrol. Earth Syst. Sci., 19(2), 893–912, doi:10.5194/hess-19-893-2015, 2015.

Perrin, C., Michel, C. and Andréassian, V.: Improvement of a parsimonious model for streamflow simulation, J. Hydrol., 279, 275–289, 2003.

Pianosi, F., Beven, K., Freer, J., Hall, J. W., Rougier, J., Stephenson, D. B. and Wagener, T.: Sensitivity analysis of

15   environmental models: A systematic review with practical workflow, Environ. Model. Softw., 79(MAY), 214–232, doi:10.1016/j.envsoft.2016.02.008, 2016.

Reusser, D. E. and Zehe, E.: Inferring model structural deficits by analyzing temporal dynamics of model performance and parameter sensitivity, Water Resour. Res., 47(7), doi:10.1029/2010wr009946, 2011.

Rimmer, A. and Hartmann, A.: Optimal hydrograph separation filter to evaluate transport routines of hydrological models, J.

20   Hydrol., 514, 249–257, doi:10.1016/j.jhydrol.2014.04.033, 2014.

Sauter, M., Kovács, A., Geyer, T. and Teutsch, G.: Modellierung der Hydraulik von Karstgrundwasserleitern – Eine Übersicht, Grundwasser, 3, 143–156, 2006.

Seibert, J. and McDonnell, J. J.: On the dialog between experimentalist and modeler in catchment hydrology: Use of soft data for multicriteria model calibration, Water Resour. Res., 38(11), 1241, doi:10.1029/2001WR000978, 2002.

25   Son, K. and Sivapalan, M.: Improving model structure and reducing parameter uncertainty in conceptual water balance models through the use of auxiliary data, Water Resour. Res., 43(1), W01415, doi:10.1029/2006wr005032, 2007.

Thornthwaite, C. W.: An Approach toward a Rational Classification of Climate, Geogr. Rev., 38(1), 55–94, doi:10.2307/210739, 1948.

Wagener, T., McIntyre, N., Lees, M. J., Wheater, H. S. and Gupta, H. V: Towards reduced uncertainty in conceptual rainfall-

30   runoff modelling: dynamic identifiability analysis, Hydrol. Process., 17(2), 455–476 [online] Available from: http://dx.doi.org/10.1002/hyp.1135, 2003.

Wheater, H. S., Bishop, K. H. and Beck, M. B.: The identification of conceptual hydrological models for surface water acidification, Hydrol. Process., 1(1), 89–109, doi:10.1002/hyp.3360010109, 1986.

White, W. B. and White, E. L.: Ground water flux distribution between matrix , fractures , and conduits : constraints on

modeling, 2003.

Ye, W., Bates, B. C., Viney, N. R., Sivapalan, M. and Jakeman, A. J.: Performance of conceptual rainfall-runoff models in low-yielding ephemeral catchments, Water Resour. Res., 33(1), 153–166, doi:10.1029/96wr02840, 1997.