



# Prediction, time variance, and classification of hydraulic response to recharge in two karst aquifers

A. J. Long<sup>1</sup> and B. J. Mahler<sup>2</sup>

<sup>1</sup>US Geological Survey, 1608 Mountain View Rd, Rapid City, South Dakota, 57702, USA

<sup>2</sup>US Geological Survey, 1505 Ferguson Lane, Austin, Texas, 78754, USA

Correspondence to: A. J. Long (ajlong@usgs.gov)

Received: 7 August 2012 – Published in Hydrol. Earth Syst. Sci. Discuss.: 16 August 2012

Revised: 1 December 2012 – Accepted: 18 December 2012 – Published: 24 January 2013

**Abstract.** Many karst aquifers are rapidly filled and depleted and therefore are likely to be susceptible to changes in short-term climate variability. Here we explore methods that could be applied to model site-specific hydraulic responses, with the intent of simulating these responses to different climate scenarios from high-resolution climate models. We compare hydraulic responses (spring flow, groundwater level, stream base flow, and cave drip) at several sites in two karst aquifers: the Edwards aquifer (Texas, USA) and the Madison aquifer (South Dakota, USA). A lumped-parameter model simulates nonlinear soil moisture changes for estimation of recharge, and a time-variant convolution model simulates the aquifer response to this recharge. Model fit to data is 2.4 % better for calibration periods than for validation periods according to the Nash–Sutcliffe coefficient of efficiency, which ranges from 0.53 to 0.94 for validation periods. We use metrics that describe the shapes of the impulse-response functions (IRFs) obtained from convolution modeling to make comparisons in the distribution of response times among sites and between aquifers. Time-variant IRFs were applied to 62 % of the sites. Principal component analysis (PCA) of metrics describing the shapes of the IRFs indicates three principal components that together account for 84 % of the variability in IRF shape: the first is related to IRF skewness and temporal spread and accounts for 51 % of the variability; the second and third largely are related to time-variant properties and together account for 33 % of the variability. Sites with IRFs that dominantly comprise exponential curves are separated geographically from those dominantly comprising lognormal curves in both aquifers as a result of spatial heterogeneity. The use of multiple IRF metrics in PCA is a novel method to characterize, compare, and classify the way in which different sites

and aquifers respond to recharge. As convolution models are developed for additional aquifers, they could contribute to an IRF database and a general classification system for karst aquifers.

## 1 Introduction

An understanding of how key hydrologic variables, such as spring flow and groundwater levels, are likely to respond to potential future climate scenarios is critical for water management. Karst aquifers are likely to be particularly susceptible to changes in short-term climate variability, because the cavernous porosity of these aquifers allows rapid replenishment by focused recharge through streambeds and sinkholes (White, 1988), the amount and timing of which is tightly linked to precipitation and antecedent moisture conditions (e.g., Long, 2009; Jukić and Denić-Jukić, 2011). The karstic Edwards aquifer has been identified as being particularly vulnerable to climate-change effects because of high use, strong links to climatic inputs, and large variability in precipitation and multi-year droughts (Loáiciga et al., 1996).

High-resolution weather-forecast models have been adapted to simulate regional climate change on the basis of boundary conditions taken from coarser-resolution general circulation models (e.g., Mearns et al., 2009; Hostetler et al., 2011). With regional climate models continually improving, convolution modeling is a promising approach to estimate how hydrologic systems will respond to local-scale climate scenarios. Convolution is a time-series method that has been widely used in rainfall-runoff models to simulate streamflow in response to infiltration on a watershed (Dooge, 1973;

Jakeman and Hornberger, 1993; Pinault et al., 2001; Simoni et al., 2011) and also has been applied to groundwater response (Hall and Moench, 1972; Long and Derickson, 1999; Von Asmuth et al., 2002; Jukić and Denić-Jukić, 2006).

Convolution is particularly useful for simulating karst hydrologic systems, which respond rapidly to changes in precipitation but for which site-specific response is difficult if not impossible to simulate with physically based, distributed-parameter models. Unsteady and nonuniform flow in variably saturated conduits and pressurized flow in filled conduits, as described by Reimann et al. (2011), are spatially and temporally variable conditions that can be simulated by convolution, even if the conduit network is not defined explicitly. There is interest in classifying different types of karst aquifers (e.g., Smart and Worthington, 2004), but only a few of the proposed approaches are quantitative (Covington et al., 2009, 2012; Labat et al., 2011; Luhmann et al., 2011). Convolution modeling estimates the impulse-response function (IRF), which characterizes the functioning of the groundwater system independently from system inputs. Other time-series methods such as Fourier spectral analysis or wavelet analysis are useful for characterizing hydraulic or hydrochemical responses in karst aquifers (e.g., Massei et al., 2006), but these are dependent on system inputs.

Here we explore methods that could be applied to model site-specific hydraulic responses with the intent of future application to projected climate simulations from improved high-resolution, dynamical climate models. We compare hydraulic-response characteristics at several sites in two karst aquifers: the Edwards aquifer (Texas, USA) and the Madison aquifer (South Dakota, USA). We describe the application of convolution models on a site-specific basis and quantification of the predictive accuracy of model output. We examine and quantify time-variant properties, a condition common in karst systems especially when epiphreatic conduits exist (Jeannin, 2001). The model simulates nonlinear soil-moisture changes for estimation of recharge and uses a time-variant convolution process to simulate the aquifer response to this recharge. We use metrics that describe the shapes of the IRFs obtained from convolution modeling to make comparisons among sites and between two aquifers in a novel approach that could be useful for karst aquifer classification.

## 2 Methods

The model consists of simulating two processes in series. The first is the process of precipitation becoming recharge; the second is recharge transitioning into a system response (i.e., functioning of the groundwater system) such as spring flow or groundwater level. The IRFs estimated by modeling were then used in aquifer classification.

### 2.1 Estimating recharge

Effective and surplus precipitation are terms used in watershed modeling to describe the variable fraction of daily precipitation that results in streamflow response. On karst land surfaces, we assume that direct runoff is small or negligible (e.g., Miller and Driscoll, 1998; Carter and Driscoll, 2006), and groundwater recharge can be estimated similarly to effective-precipitation methods, as described by Jakeman and Hornberger (1993; Eqs. 1–4). First, a daily soil-moisture index  $s$  is estimated, which is weighted by a backward-in-time exponential decay function that operates on the past daily rainfall record, as described by

$$s_i = cr_i + (1 - \kappa_i^{-1})s_{i-1} \quad (1)$$

$$= c \left[ r_i + (1 - \kappa_i^{-1})r_{i-1} + (1 - \kappa_i^{-1})^2 r_{i-2} + \dots \right] \\ i = 0, 1, \dots, N \quad 0 \leq s \leq 1, \quad (2)$$

where  $c$  is a normalizing parameter that limits  $s$  to values between 0 and 1 (dimensionless),  $\kappa$  adjusts the effect of antecedent conditions and is related to evapotranspiration (dimensionless),  $r$  is total daily rainfall (cm), and  $i$  is the time step (days).

In watershed modeling,  $c$  is calculated to satisfy an assumption that effective precipitation within the watershed is equal to outflow. However, the groundwater recharge area that affects a well or spring is not precisely defined, and so  $c$  cannot be determined empirically for groundwater applications, which also may be true for karst watersheds. Therefore, we optimize  $c$  through model calibration. Air temperature, which affects evapotranspiration and soil drying rates, is accounted for by adjusting  $\kappa$  by daily air temperature:

$$\kappa_i = \alpha \exp[(20 - T_i)f] \quad f > 0, \quad (3)$$

where  $\alpha$  is a scaling coefficient (dimensionless),  $T$  is daily mean air temperature ( $^{\circ}\text{C}$ ), and  $f$  is a temperature modulation factor ( $^{\circ}\text{C}^{-1}$ ). Equation (3) has the primary effect of increasing the value of  $s$  during cool periods ( $0 < T < 20^{\circ}\text{C}$ ) when evapotranspiration is low. Daily recharge  $u_i$  (cm) is then calculated as a fraction of daily precipitation by

$$u_i = r_i s_i. \quad (4)$$

Equations (1)–(4) describe a lumped-parameter model of recharge, because several physical processes are lumped into a few parameters. Only precipitation that occurs either as rain or melting snow was included in the calculation of  $s$ . A method was established to estimate the occurrence of snow precipitation and melting for future periods on the basis of simulated air temperature. To determine the form of precipitation for each day, an air temperature threshold value ( $T_s = 0^{\circ}\text{C}$ ) was set, below which precipitation was assumed to occur as snow. To determine days when melting occurs,

a melting threshold value  $T_m$  was estimated. If daily snow-depth data are available, this melting threshold can be determined empirically as the average air temperature for days when snow depth decreased to zero from a previous day with a snow depth greater than zero. Sublimation was accounted for by estimating the fraction of snow moisture remaining after sublimation ( $S_f$ ; Long, 2009). Snow precipitation was multiplied by  $S_f$  and summed for the continuous series of snow-precipitation days prior to each snowmelt day. This sum was added to each snowmelt day in the daily rainfall record, because snowmelt was assumed to have the same effect as rainfall on the value of  $s$ .

## 2.2 Convolution

If a system input, after being transported through a medium, results in a system output that is dispersed in time according to a characteristic waveform, either time variant or invariant, then this system can be simulated by convolution. When applied to hydrologic modeling, this is another type of lumped-parameter model that also has been described as a transfer-function model and a linear-reservoir model (e.g., Nash, 1959; Von Asmuth et al., 2002). Convolution is described by

$$y(t) = \int_0^{\infty} h(t - \tau)u(\tau)d\tau, \quad (5)$$

where  $y(t)$  is the system output, or response;  $u(\tau)$  is the system input, or forcing function;  $h(t - \tau)$  is an impulse-response function (IRF);  $\tau$  and  $t$  are time variables corresponding to system input and output, respectively, where  $t - \tau$  represents the delay time from system input to output (Dooge, 1973; Olsthoorn, 2008). For uniform time steps, the discrete form of Eq. (5) is

$$y_i = \sum_{j=0}^i h_{i-j}u_j \quad i = 0, 1, \dots, N, \quad (6)$$

where  $N$  is the number of time steps in the output record. The IRF describes the system output  $y$  that results from an instantaneous unit input  $u_j$ . The IRF has been described by many different terms, including instantaneous unit hydrograph, transfer function, and kernel (e.g., Nash, 1959; Berendrecht et al., 2003; Jukić and Denić-Jukić, 2006).

Exponential, lognormal, and Pearson type III (gamma) functions, all of which are left skewed, have been used as approximations of IRFs (Nash, 1959; Von Asmuth et al., 2002; Berendrecht et al., 2003). The exponential curve has a peak response, or mode, at zero, whereas lognormal and Pearson type III curves have peak responses  $> 0$  and are similarly shaped. We use exponential curves, lognormal curves, or a combination of the two to approximate IRFs. The exponential curve is defined as

$$h(t) = a\lambda \exp(-\lambda t), \quad (7)$$

where  $a$  is a scaling coefficient, and  $\lambda$  determines the mean and variance of the system output time  $t$  as

$$\mu = \lambda^{-1} \quad (8)$$

and

$$\sigma^2 = \lambda^{-4}, \quad (9)$$

respectively. The lognormal curve is defined as

$$h(t) = \frac{b}{t\sqrt{2\pi\varepsilon}} \exp\left[-\frac{1}{2\varepsilon}(\ln t - \omega)^2\right], \quad (10)$$

where  $b$  is a scaling coefficient, and  $\omega$  and  $\varepsilon$  determine the mean and variance of  $t$  as

$$\mu = \exp(\omega + \varepsilon/2) \quad (11)$$

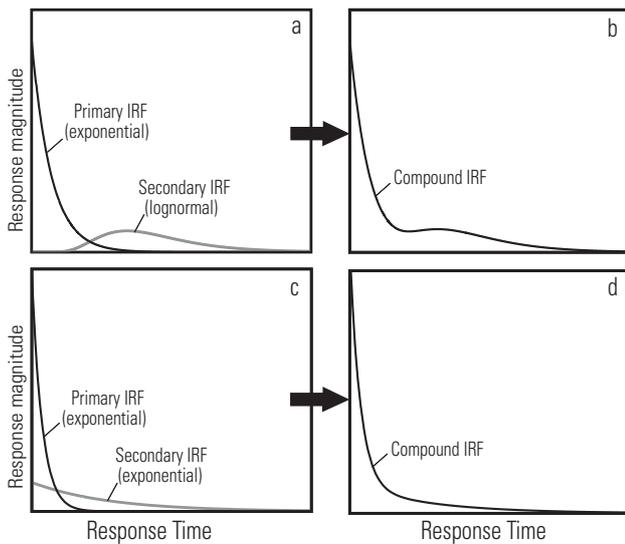
and

$$\sigma^2 = \mu^2[\exp(\varepsilon) - 1], \quad (12)$$

respectively. Parameters of Eqs. (1)–(12) are listed in Table 1. The length of the IRF quantifies the system memory, or time that the response to the impulse effectively persists. Because exponential and lognormal curves are asymptotic and thus have infinite length with infinitesimal magnitude after some point in time, system memory is defined herein as time  $t_m$  on the IRF time scale at which 95 % of the curve area is between time  $t = 0$  and  $t_m$ .

Numerical simulation of flow in karst settings is complicated by the existence of quick-flow and slow-flow components, e.g., flow through large conduits and through small fractures (Pinault et al., 2001). In some cases, a single exponential or lognormal IRF can adequately represent the quick-flow and slow-flow components of karst groundwater flow, where the first part of the curve, including the initial peak, represents quick flow, and the tail of the curve represents slow flow. In other cases, a secondary IRF with a long tail that represents all or part of the slow-flow component may be useful (e.g., Long, 2009). We combine the primary and secondary IRFs by superposition in a compound IRF. Pinault et al. (2001) used a similar approach to represent the quick-flow and slow-flow components in karst aquifers. Denić-Jukić and Jukić (2003) used one function for the IRF peak and another function for the tail in a composite IRF. In the approach herein, the compound IRF can consist of any combination of exponential or lognormal curves. A compound IRF consisting of at least one lognormal curve can result in a bimodal, or double-peaked, distribution of response times, which could be a result of two-domain flow (Long and Putman, 2006). A compound IRF consisting of two exponential curves is useful when quick flow and slow flow are separated into a sharp initial peak and long tail, respectively (Fig. 1).

If the scaling coefficients  $a$  and  $b$  are set to unity, the areas under the IRFs also are equal to unity. Adjusting the



**Fig. 1.** Examples of compound impulse-response functions (IRFs) consisting of the superposition of an exponential and a lognormal curve (**a** and **b**) and two exponential curves (**c** and **d**).

values of these coefficients allows their use as conversion factors to account for the different dimensions between the system input and output. This also allows for unequal curve areas between the primary and secondary IRF and allows the curve area for time-variant IRFs to change. For simulation of groundwater levels, a datum  $h_0$  at which hydraulic head equals zero must be established. Conceptually, this is the level to which hydraulic head would converge if the local system-input recharge was eliminated. This system input is assumed to be the only source of recharge close enough to cause hydraulic-head fluctuation and the only source that causes hydraulic head to rise above  $h_0$ .

### 2.3 Nonlinearity and time variance

Nonlinearity and time variance have distinct meanings in time-series analysis as described by Jenkins and Watts (1968). By this definition, the recharge estimation method described by Eqs. (1)–(4) represents a nonlinear process. Antecedent soil-moisture effects result in nonlinear watershed flow processes (Jakeman and Hornberger, 1993; Simoni et al., 2011). In contrast, convolution, which represents groundwater flow in response to recharge, is defined as a linear system that can be either time invariant or time variant, depending on whether the IRF is static or changing in time (Jenkins and Watts, 1968). Most commonly, time invariance is assumed in convolution models (e.g., Von Asmuth et al., 2002; Denić-Jukić and Jukić, 2003). Time-variant IRFs, however, might be critical for simulating karst aquifers because of horizontal and vertical heterogeneity, where fluctuating water levels saturate or desaturate different parts of the aquifer having different conduit or fracture networks. For application to karst watersheds, Pinault et al. (2001) used

**Table 1.** Model parameters [–, not applicable].

Parameter	Description	Equation	Estimation method
$c$	Soil moisture parameter	1	Optimized
$\kappa$	Soil moisture parameter	3	Optimized
$f$	Soil moisture parameter	3	Optimized
$T_s$	Snow precipitation threshold	–	Assumed 0 °C
$T_m$	Snowmelt threshold	–	Empirical
$S_f$	Sublimation fraction	–	Optimized
$\lambda$	Exponential IRF shape parameter	7	Optimized
$a$	Exponential IRF curve area	7	Optimized
$\omega$	Lognormal IRF shape parameter	10	Optimized
$\varepsilon$	Lognormal IRF shape parameter	10	Optimized
$b$	Lognormal IRF curve area	10	Optimized
$h_0$	Hydraulic-head datum	–	Optimized

time-variant IRFs that increased and decreased continuously with hydraulic head, which allowed a change in IRF size; a disadvantage of this method for our application is that the IRF shape does not change. For karst aquifers, Jukić and Denić-Jukić (2006) used IRFs that varied continuously in size and also shape according to an index of antecedent recharge; a disadvantage of this method is the large number of parameters necessary to define the IRF.

We present a method to represent time-variant IRFs that change in size and shape, but with a minimal number of parameters, which is important for any lumped-parameter model. First, the precipitation record was separated into wet and dry periods, which were defined as years in which the annual precipitation was either above or below the long-term mean, respectively, determined for the period of record at each site. Other methods for defining wet and dry periods also might be useful, depending on the study area (e.g., use of a drought index). Second, the shape parameters and scaling coefficients of the IRFs were estimated separately for the wet and dry periods, with the assumption of time invariance within each of these two periods. This is a method not previously used and includes a simplifying assumption that the IRF changes abruptly from wet to dry periods; model validation indicates that this is not a detrimental assumption. Further, the IRF change from wet to dry periods might occur quickly in some cases because of large heterogeneities that exist in karst aquifers. The main advantages of this method are that it requires minimal parameters and is well suited to aquifer classification, because wet-period and dry-period IRFs are distinct and clearly defined.

As many as four IRFs were used to simulate the system output: primary IRFs ( $IRF_{w1}$  and  $IRF_{d1}$ ) and secondary IRFs ( $IRF_{w2}$  and  $IRF_{d2}$ ), where the subscripts, w and d, refer to wet and dry periods, respectively. The IRF curve areas can be different for wet and dry periods and primarily are meaningful in their relative magnitudes for comparison of the responses of these two periods, which may provide insight into system functioning. For example, if a wet-period IRF has twice the area of a dry-period IRF, then the wet-period response is twice that of the dry period for the same amount of recharge.

## 2.4 Classification of sites

Hydraulic-response characteristics can differ among sites, which can represent different flow systems within an aquifer, i.e., different networks of conduits, fractures, and other pore spaces. Because the functioning of a groundwater system can be characterized by its IRF (Von Asmuth and Knotters, 2004), a large number of physically based parameters that describe an aquifer can largely be summarized by the IRF. This lumping of a large and complex parameter set into a few simple parameters is ideal for aquifer classification. The IRF is suitable for comparison of sites and aquifers in climatically different locations, because this method characterizes the aquifer independently from system inputs; e.g., this comparison does not depend on differences in rainfall frequency, variability, or intensity between different locations. Von Asmuth and Knotters (2004) combined moments of the IRF with characteristics of the system input to determine four metrics that can be used to classify the combined spatial and temporal aspects of the groundwater system. Because of the large variety and complexity of karst groundwater-flow systems and the challenges associated with their classification, our focus is to isolate the spatial differences of these systems, and thus we use the IRF only for classification. Because of the large variety of IRF shapes that might apply to karst aquifers, 16 metrics were used to fully describe this shape. The use of principal component analysis (PCA) provided quantifiable relations among sites and helped describe the meaning of these 16 metrics with only three principal components that could be related to system functioning.

We used metrics, which can be determined for any parametric or nonparametric IRF, to quantify several characteristics of the IRF shapes (Table 2). To define these metrics, the IRF was assumed to be a frequency distribution of the transit times of the response quantity, either hydraulic head or flow, and moments of the distribution (i.e., mean, variance, skewness, and kurtosis) and other metrics describing the shape of the distribution were computed. Metrics were selected that quantify the IRF shape independently from scale so that comparisons were not weighted by the overall IRF magnitude, which might vary for climatically different locations or different distances from the recharge area. Ratios were used for many of the metrics, because these are scale independent. Metrics include two scale-independent moments, skewness (skw) and kurtosis (krt), and five ratios: standard-deviation : mean (SDMn), standard-deviation : memory (SDMm), mean : memory (MnMm), mode : memory (MdMm), and peak-height : area (PHA; the highest peak was used for bimodal distributions). These seven metrics were quantified for wet and dry periods separately (Table 2), resulting in 14 metrics. For time-invariant systems, wet and dry metrics are equal. In addition, the wet : dry area ratio (WDA) was included, which is the ratio of the wet-period to dry-period IRF area ( $WDA = 1.0$  for time-invariant systems). Finally, the metric WDD is the area that

**Table 2.** Impulse-response function (IRF) metrics. Metrics quantified separately for wet and dry periods are designated by the subscript “w” or “d”, respectively, at the end of the abbreviations.

IRF Metric	Abbreviation
Skewness	skw
Kurtosis	krt
SD : mean ratio	SDMn
SD : memory ratio	SDMm
Mean : memory ratio	MnMm
Mode : memory ratio	MdMm
Peak-height : area ratio	PHA
Wet : dry area ratio	WDA*
Wet-dry shape difference	WDD*

\* Not defined separately for wet and dry periods.

the wet- and dry-period IRFs do not overlap divided by the total area of both IRFs. This metric quantifies the difference in IRF shape between wet and dry periods and results in a total of 16 metrics (Table 2).

PCA was used to assess similarities, differences, and groupings of sites on the basis of IRF shape, as described by the 16 metrics. PCA is a linear transformation of data in multidimensional space, where the transformed axes, or principal components, align with the greatest variances in the multivariate dataset (Davis, 2002). Each principal component is a new variable that is a linear combination of all the original variables. PCA is helpful for elucidating patterns that would otherwise be obscured in attempting to assess a large number of metrics. The software MATLAB (<http://www.mathworks.com>) was used for PCA.

## 2.5 Modeling considerations

Convolution models provide a convenient way to assess system memory, which is an important consideration in any hydrologic model. At a minimum, input corresponding to a time period equal to the system memory is required prior to the start of a calibration period. This initial time period is referred to as model spin-up, the output from which is of questionable validity, because antecedent effects of the system are not fully accounted for.

The length of the observed record also must be considered in light of the system memory. There is less confidence in the predictive strength of a model if the observed record is shorter than the system memory than if it is longer, because, in the former case, the IRF tail is not fully tested against observation. Ideally, the validation period alone should be longer than the system memory, and if it is several times longer, then the full range of the IRF is tested several times over. The *memory ratio* is defined as the ratio of the system memory to the length of the observed record, where a value of less than unity is desirable.

The use of secondary IRFs, the choice of curve type, and time variance were evaluated and selected for each site on the basis of model fit for the validation period. Inclusion of secondary IRFs and simulating a system as time variant increase model parameterization and complexity over the simpler options. In some cases the simplest model resulted in a better fit for the validation period than one with added complexity. Although model fit might improve for the calibration period as parameters are added, if the validation period indicates that this added complexity is not helpful, the model is overparameterized (Von Asmuth et al., 2002). If the overall model fit is poor, this model might not be appropriate, or the input data might not represent the recharge area well.

### 3 Model application and results

Hydraulic responses at several sites with data for spring flow, groundwater level, stream base flow, or cave drip (Table 3) were simulated for two karst aquifers: the Edwards aquifer in south-central Texas and the Madison aquifer in western South Dakota (Fig. 2). Weather station data for precipitation and air temperature were obtained from the National Climatic Data Center (2012; Table S1 in the Supplement). Generally, the recharge areas for the sites simulated are small, and data from a single weather station were used as model input. The specific recharge area for each site was assumed to be directly upgradient from the site (Fig. 2). The length of each recharge area is not precisely defined, but we assumed that the primary effect on the response at each site does not extend a long distance along the length of the formation outcrop. The weather station either within or nearest the recharge area for each site was used, and the next nearest station was used to fill periods of missing data if necessary (Table S2 in the Supplement). Equations (1), (3), (4), (6), (7), and (10) were programmed in MATLAB and executed on a daily time step. The models were calibrated to hydrologic data for the sites (Table 3; US Geological Survey, 2012). Model parameters (Table 1) were optimized by using the “lsqcurvefit” function in MATLAB to minimize the differences (residuals) between simulated and observed values. This is a subspace trust-region method and is based on the interior-reflective Newton method (Coleman and Li, 1994, 1996).

#### 3.1 Study areas

The Edwards aquifer in south-central Texas is a well-developed karst aquifer contained within the Edwards Group. Surface recharge to the Edwards aquifer occurs from direct precipitation and sinking streams that cross onto the outcrop area of the Edwards Group from the west and northwest (Fig. 2). The aquifer dips to the south and southeast below the land surface. Groundwater flow generally is to the east and northeast, and discharge from the aquifer occurs

**Table 3.** Sites simulated.

Site label (Fig. 2)	Site name	USGS site number <sup>a</sup>	Site type
Edwards aquifer sites			
HCV	Hill Country Village	293522098291201	Well
LVL	Lovelady	8159000	Well
FM1796	Medina FM1796	292618099165901	Well
Bxr	Bexar Co. Index well	292845098255401	Well
Dow	Dowell well	300835097483401	Well
Bud	Buda well	300510097504001	Well
BARsp	Barton Springs	8155500	Spring
COMsp	Comal Springs	8168710	Spring
Madison aquifer sites			
LA88C	Spearfish GC well	442854103505602	Well
Tilf	Tilford well	441759103261202	Well
RG	Reptile Gardens well <sup>c</sup>	435916103161801	Well
RFsp	Rhoads Fork Spring	06408700	Spring
FALr	Fall River	06402000	Spring complex
WCL	Windy City Lake	433302103281501 <sup>b</sup>	Cave water body
LScr	Little Spearfish Creek	06430850	Base flow
SPFcr	Spearfish Creek	06430900	Base flow
Cave drip			
CTD	Caving Tour drip	433302103281508 <sup>d</sup>	Cave drip
RmDr	Room Draculum drip	433302103281509 <sup>d</sup>	Cave drip
Fractured-rock watershed			
BEVcr	Beaver Creek	06402430	Watershed

<sup>a</sup> Flow and water-level data from US Geological Survey (2012) unless otherwise indicated. <sup>b</sup> Partial water-level record was estimated from water levels in well 7-11 MDSN (Fig. 2) with data from the South Dakota Department of the Environment and Natural Resources in Pierre, South Dakota ( $R^2 = 0.99$  for correlation between the two sites; Helsel and Hirsch, 2002). <sup>c</sup> Streamflow at USGS streamgage 06407500 (US Geological Survey, 2012) was used to estimate recharge. <sup>d</sup> Data were provided by M. Ohms of Wind Cave National Park in 2012.

at several large springs. The hydrogeology of the Edwards aquifer is described in detail in Maclay and Small (1983), Small et al. (1996), and Lindgren et al. (2004).

The Madison aquifer in western South Dakota is a well-developed karst aquifer composed of limestone and dolostone (Greene and Rahn, 1995). It is contained within the regionally extensive Madison limestone of Mississippian age, referred to locally as the Pahasapa limestone. This formation is exposed at the land surface on all flanks of the Black Hills and dips radially outward in all directions below the land surface (Fig. 2); the outcrop of the Madison limestone is the recharge area for the Madison aquifer. The hydrogeology of the Madison aquifer in the Black Hills area is described in detail in Driscoll and Carter (2001).

#### 3.2 Sites simulated

For the Edwards aquifer, groundwater levels in six wells and flow from two large springs were simulated (Fig. 2, Table 3). For the Madison aquifer, three wells, one spring, one spring complex, one cave water body, and stream base flow for two Madison limestone watersheds were simulated (Fig. 2, Table 3). In addition, cave drip at two sites and streamflow in one fractured-rock watershed located in the Madison study area were simulated for comparison to the aquifer sites (Fig. 2, Table 3). The observed and simulated flows and groundwater levels are shown for two of the sites in Fig. 3, as examples, and for all of the sites in Fig. S1 in the Supplement.

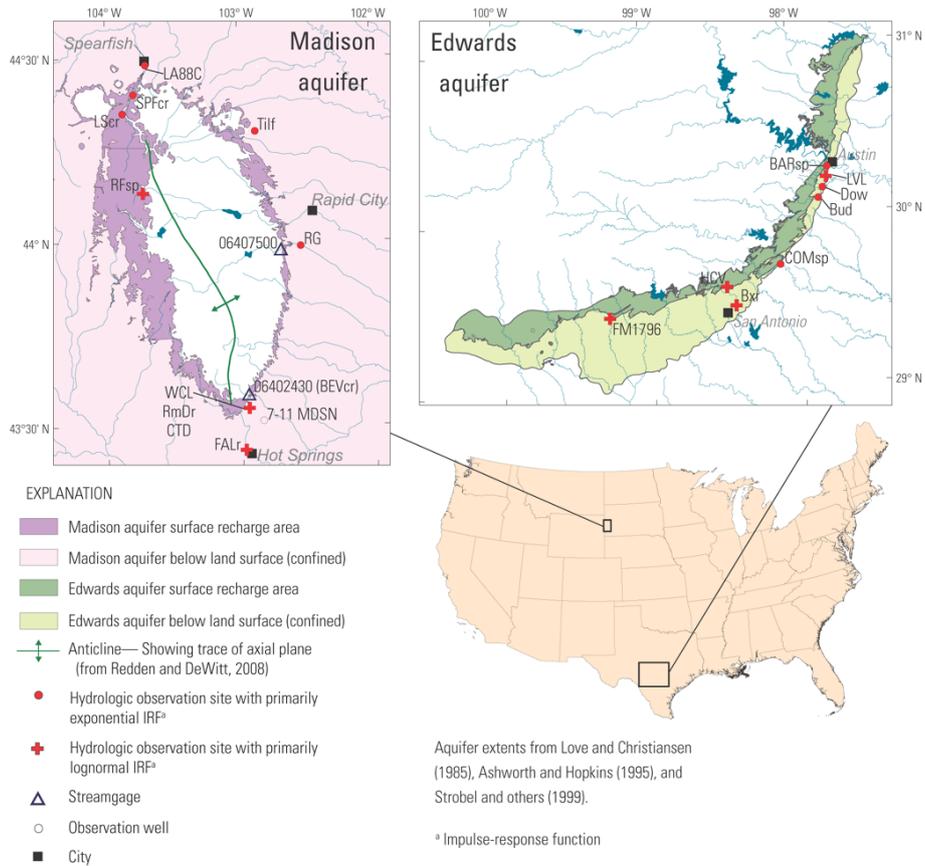


Fig. 2. Study areas showing the Madison and Edwards aquifers.

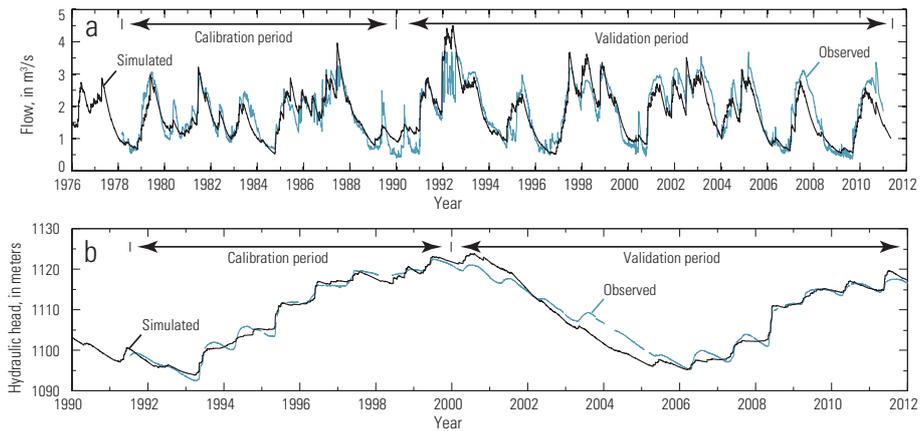


Fig. 3. Modeling results for (a) Barton Springs and (b) the Tilford well showing observed and simulated values.

Daily precipitation and air temperature were used as model input for all sites, except for the Reptile Gardens well (site RG; Fig. 2, Table 3). This site is located in an area where the primary recharge source is sinking streamflow on the Madison limestone outcrop, which was estimated as described in Hortness and Driscoll (1998) on a daily time step for streamflow at streamgage 06407500 (Fig. 2).

These estimates of sinking-stream recharge were used as direct input for convolution for site RG, without the use of Eqs. (1)–(4).

Hydrologic-response data for all of the Edwards aquifer sites consisted of direct water-level or spring flow measurements, but additional description or data manipulation was required for some of the Madison aquifer sites, e.g., because

of indirect spring flow measurements. Observed streamflow at Little Spearfish Creek and Spearfish Creek (sites LScr and SPFCr) was used as an estimate of groundwater discharge, which is 97 % and 86 %, respectively, of total streamflow (Driscoll and Carter, 2001). Overland runoff in these watersheds rarely occurs because of the highly porous karst terrain (Miller and Driscoll, 1998). Groundwater discharge to the Fall River (site FALr) is about 96 % of streamflow, primarily flowing from a complex of artesian springs (Rahn and Gries, 1973; Back et al., 1983; Driscoll and Carter, 2001). A 4-month moving average of streamflow for site FALr was used as a surrogate for total observed spring flow to remove anthropogenic variability resulting from municipal water use and wastewater discharge affecting this site. Sinking-stream recharge from Beaver Creek (site BEVcr) was used as a source of recharge in addition to precipitation for simulation of Windy City Lake (site WCL), and the model for this site therefore included an additional IRF for sinking-stream recharge. Simulated streamflow for site BEVcr at a daily time step was used as model input for the period prior to 1991, because data were not available for that period.

### 3.3 Calibration and validation of models

The models were validated by (1) calibrating each model to part of the record for the system output, (2) executing the model under those conditions for the remaining observation period (validation period), and (3) examining the similarity between the simulated and observed system outputs for the validation period. Calibration periods ranged from 0.9 to 70 yr, and validation periods ranged from 0 to 23.8 yr (Table 4). Before parameter optimization could be executed effectively, initial values needed to be estimated so that the true minimum of residuals could be achieved. Therefore, model parameters were adjusted by trial and error until the simulated and observed hydrographs were roughly similar. These parameter values then were used as initial values in parameter optimization for the calibration periods. Calibration and validation periods were evaluated on the basis of the Nash–Sutcliffe coefficient of efficiency (Nash and Sutcliffe, 1970; Legates and McCabe, 1999), which is a measure of the similarity between simulated and observed time-series records, hereafter referred to as *model fit*. The coefficient of efficiency  $E$  is defined as

$$E = 1 - \frac{\sum (y_{\text{obs}} - y_{\text{sim}})^2}{\sum (y_{\text{obs}} - y_{\text{mean}})^2}, \quad (13)$$

where  $y_{\text{obs}}$  and  $y_{\text{sim}}$  are daily time series of the observed and simulated system outputs, respectively, and  $y_{\text{mean}}$  is the mean of  $y_{\text{obs}}$ . Theoretically,  $E$  could vary from  $-\infty$  (poorest fit) to unity (perfect fit) and is the ratio of the magnitude of residuals (numerator) to the overall variability in the observed record (denominator) subtracted from unity. An  $E$  value of zero indicates that the observed mean ( $y_{\text{mean}}$ ) is an equally

good predictor as is the simulation ( $y_{\text{sim}}$ ; Legates and McCabe, 1999).

Values of  $E$  were calculated for the calibration and validation periods separately ( $E_c$  and  $E_v$ , respectively; Table 4). For comparison of residuals for different periods, the denominator of the second term in Eq. (13) must be consistent across all cases. Therefore, the denominator was calculated on the basis of the total period, and the numerator was calculated on the basis of the period of interest only. Because the total period is longer than the partial period of interest, the denominator was scaled down to be consistent with the time period of the numerator by

$$E = 1 - \frac{\sum (y_{\text{obs}} - y_{\text{sim}})^2}{\sum (y_{\text{obs}} - y_{\text{mean}})^2 \left( l_p / l_T \right)}, \quad (14)$$

where  $l_p$  and  $l_T$  are the lengths of the partial and total periods, respectively. This method provides a direct comparison of residuals for different periods, even if these periods have different fluctuation amplitudes, and thus is robust for comparing short periods, where fluctuation amplitudes might be small in comparison to the overall record. Because parameters were optimized for the calibration periods only,  $E_c$  values were 2.4 % higher on average than  $E_v$  values.

Primary and secondary IRFs were included in the initial parameter optimization for all sites. Following this initial optimization, secondary IRFs were eliminated for some sites if one or both of these IRFs were minimized to a small curve area as a result of optimization. In these cases, the minimized IRFs were omitted if this did not decrease the  $E$  for the validation period ( $E_v$ ) by more than 0.02. In some cases, omitting secondary IRFs resulted in increased  $E_v$  values, which indicated that secondary IRFs resulted in overparameterization. For example, using secondary IRFs for the Medina FM1796 well resulted in an  $E_v$  value of 0.81, but using only primary IRFs resulted in a higher  $E_v$  value (0.91), even though the two cases fit the calibration period equally well ( $E_c = 0.97$ ; Table 4). Using a similar approach, the particular combination of exponential and lognormal curves to use as primary and secondary IRFs was tested to determine the best fit for each site (Table 4; Table S3 in the Supplement).

All models were assumed to be time variant for initial parameter optimization. Similar shapes for the wet-period and dry-period IRFs indicated that the system might be time invariant. Time invariance for these cases was tested by using the same IRFs for wet and dry periods. If  $E_v$  did not decrease by more than 0.02 as a result, then parsimony was preferred, and the system was simulated as time invariant, requiring fewer parameters than a time-variant model. For some sites,  $E_v$  was highest for the time-invariant case; for example,  $E_v$  for the Lovelady (LVL) well was 0.69 for the time-variant case and 0.85 (Table 4) for the time-invariant case, which indicated that the time-variant model was overparameterized. For other sites, time variance was critical; for

**Table 4.** Summary of modeling results ( $E_c$ , coefficient of efficiency; IRF, impulse-response function; w1, wet-period primary; w2, wet-period secondary; d1, dry-period primary; d2, dry-period secondary; Exp, exponential curve; Logn, lognormal curve; –, not applicable).

Site label (Fig. 2)	Calibration period, in years	Validation period, in years	$E_c$ (calibration period) <sup>a</sup>	$E_v$ (validation period) <sup>a</sup>	IRF <sub>w1</sub>	IRF <sub>w2</sub>	IRF <sub>d1</sub>	IRF <sub>d2</sub>	Time invariant	Memory ratio <sup>b</sup>	Predictive strength
Edwards aquifer sites											
HCV	32.1	15.7	0.85	0.70	Logn	–	Logn	–	Yes	0.25	Medium
LVL	4.2	10.4	0.90	0.85	Logn	–	Logn	–	Yes	0.12	High
FM1796	12.0	23.8 <sup>c</sup>	0.97	0.91	Logn	–	Exp	–	No	2.44	Low
Bxr	20.0	21.0	0.69	0.53	Exp	Logn	Exp	Logn	Yes	0.73	Low
Dow	3.2	10.7	0.82	0.71	Exp	–	Exp	–	No	0.43	Medium
Bud	4.1	10.2	0.90	0.76	Exp	–	Exp	–	No	0.17	Medium
BARsp	11.8	21.0	0.87	0.74	Exp	–	Exp	–	No	0.10	Medium
COMsp	70.0	13.0	0.64	0.62	Exp	–	Exp	–	Yes	0.05	High
Madison aquifer sites											
LA88C	14.4	7.8	0.90	0.78	Exp	Logn	Exp	Logn	No	0.32	Medium
Tilf	8.5	12.0	0.99	0.94	Exp	Logn	Exp	–	No	0.50	High
RG	2.0	22.6	1.00	0.92	Exp	–	Exp	–	Yes	0.13	High
RFsp	20.2	9.0	0.87	0.63	Exp	Logn	Exp	Logn	No	0.38	Low
FALr	61.9	10.9	0.85	0.77	Exp	Logn	Exp	Logn	No	0.66	Medium
WCL	9.3	16.0	0.99	0.91	Logn	–	Logn	–	No	5.01	Low
LScr	23.2	0.0	0.76	–	Exp	Exp	Exp	Exp	No	0.27	Low
SPFcr	10.2	11.8	0.58	0.73	Exp	Exp	Exp	Exp	Yes	0.40	High
Cave drip											
CTD	0.9	1.0	0.10	0.80	Exp	Logn	–	–	– <sup>d</sup>	0.05	High
RmDr	1.1	0.0	0.97	–	Logn	–	–	–	– <sup>d</sup>	20.83	Low
Fractured-rock watershed											
BEVcr	6.2	14.8	0.50	0.71	Exp	Exp	Exp	–	No	0.09	High

<sup>a</sup> Dimensionless; computed on a daily time step. <sup>b</sup> Ratio of system memory to full system-response period, dimensionless. <sup>c</sup> Also includes a backward validation period of 12.5 yr. <sup>d</sup> Period was too short to determine time variance. Simulation had one wet period and no dry period.

example,  $E_v$  for the Tilford (Tilf) well was 0.39 for the time-invariant case and 0.94 (Table 4) for the time-variant case.

### 3.4 Assessing model fit and predictive strength

The length of the validation period is an indicator of the length of time for which a model can be projected into the future with confidence on the basis of climate-model predictions. A goal here was to have the longest validation periods possible, while maintaining a minimum, or target,  $E_v$ . In general,  $E_v$  increases with an increasing calibration period, which decreases the length of the validation period. This inverse relation between  $E_v$  and validation period length results in a compromise that must be considered when setting a target  $E_v$ . Also, for time-variant systems, the calibration periods necessarily included both wet and dry periods. Therefore, the length of the validation period was limited in some cases by the occurrence of wet and dry periods for a particular system.

The value of  $E_v$  is an indicator of the expected model error for the future period, but does not account for error in the output from a climate model that is used as input for the hydraulic-response model. If simulated future precipitation

and temperature are not within the range of observation, then the prediction accuracy of aquifer response is less certain than if these inputs stay within the range of observation. The only way to determine  $E$  for a future period is to test the model when these data become available.

A target  $E_v$  was set at 0.70. Two of the sites (LScr and RmDr) had very low  $E_v$  values ( $< 0.5$ ) for any validation-period length tested, and thus this period was set to zero with no  $E_v$  value listed (Table 4). Three other sites (Bxr, COMsp, and RFsp) had  $E_v$  values that were less than the target (Table 4). Two of these sites had  $E_c$  (calibration period) values that also were less than 0.70, and achieving the target  $E_v$  was not possible for any validation period. The  $E_v$  value for the remaining sites ranged from 0.70 to 0.94 (Table 4).

An additional measure of predictive strength is the memory ratio, where a value of less than unity is desirable. Two sites, Windy City Lake (WCL) and the Medina FM1796 well, had memory ratios greater than unity (Table 4); a memory ratio greater than unity is common for systems such as these with long memories resulting from long-term groundwater storage properties associated with these sites.

Most of the sites (62%) were simulated as time variant (Table 4), and model spin-up periods were at least as long as

system memories for all sites. For Windy City Lake, Medina FM1796 well, and Fall River, data were not available for the first part of the spin-up period. Mean values of precipitation and temperature, calculated from the periods of record, were used as surrogates for the missing data for these three sites. The lengths of these periods in years were 56 (1856–1911), 61 (1840–1900), and 24 (1887–1910) for these three sites, respectively. The effects of using mean values were tested by replacing the mean values with the subsequent period of data, which changed the  $E_V$  values by less than 0.02 for these three sites. This indicates model insensitivity to precipitation and temperature variability in the early parts of the spin-up periods.

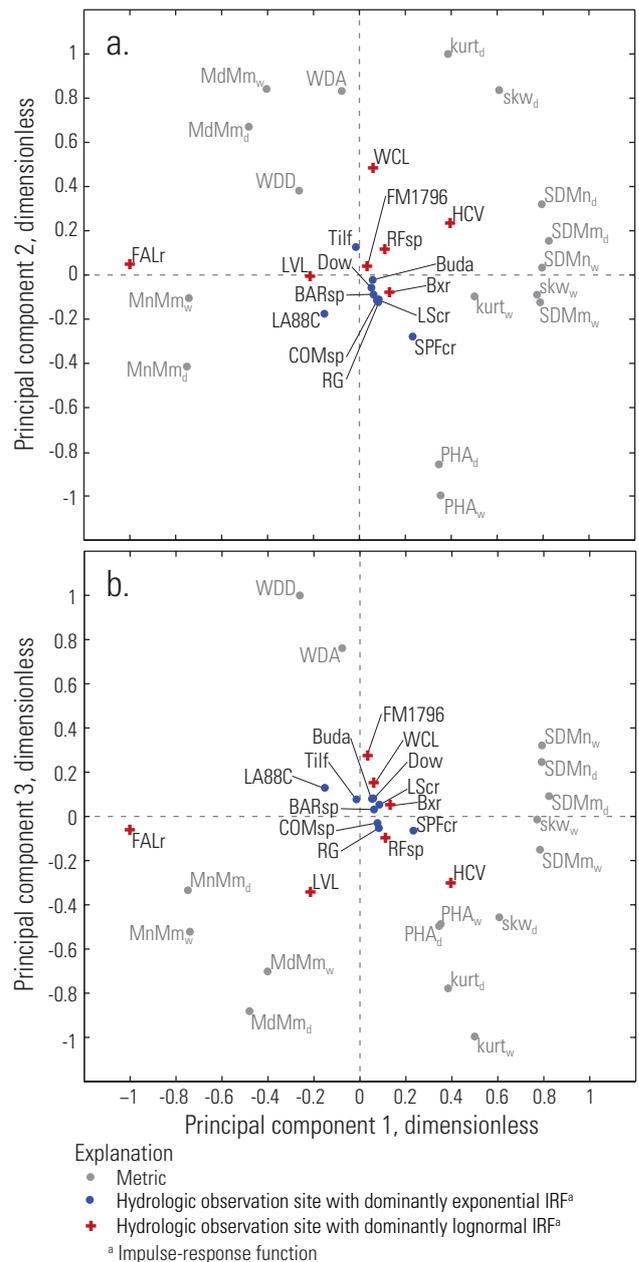
The predictive strength of each site was rated as high, medium, or low on the basis of the memory ratio and the  $E$  ratio, which is defined as  $(E_C - E_V) / E_C$ , or the fractional difference in  $E$  between the calibration and validation periods (Table 4). A small  $E$  ratio indicates that the model fit degraded little for the validation period. Sites with a memory ratio  $\leq 0.5$  and  $E$  ratio  $\leq 0.1$  were rated as high; sites with memory ratio  $\leq 1.0$  and  $E$  ratio  $\leq 0.2$  were rated as medium; and all other sites not meeting these criteria were rated as low.

### 3.5 Classification of sites

Sixteen metrics that characterize the shapes of the IRFs for the eight Edwards aquifer sites and eight Madison aquifer sites were used in PCA (Table 2; Table S4 in the Supplement). The values were log transformed and standardized to a mean of zero and standard deviation of unity before applying PCA. Principal components 1, 2, and 3 (PC1, PC2, and PC3) explain 51, 19, and 14%, respectively of the variance in the dataset. The relations between the sites and the metrics were plotted in principal component space (Fig. 4).

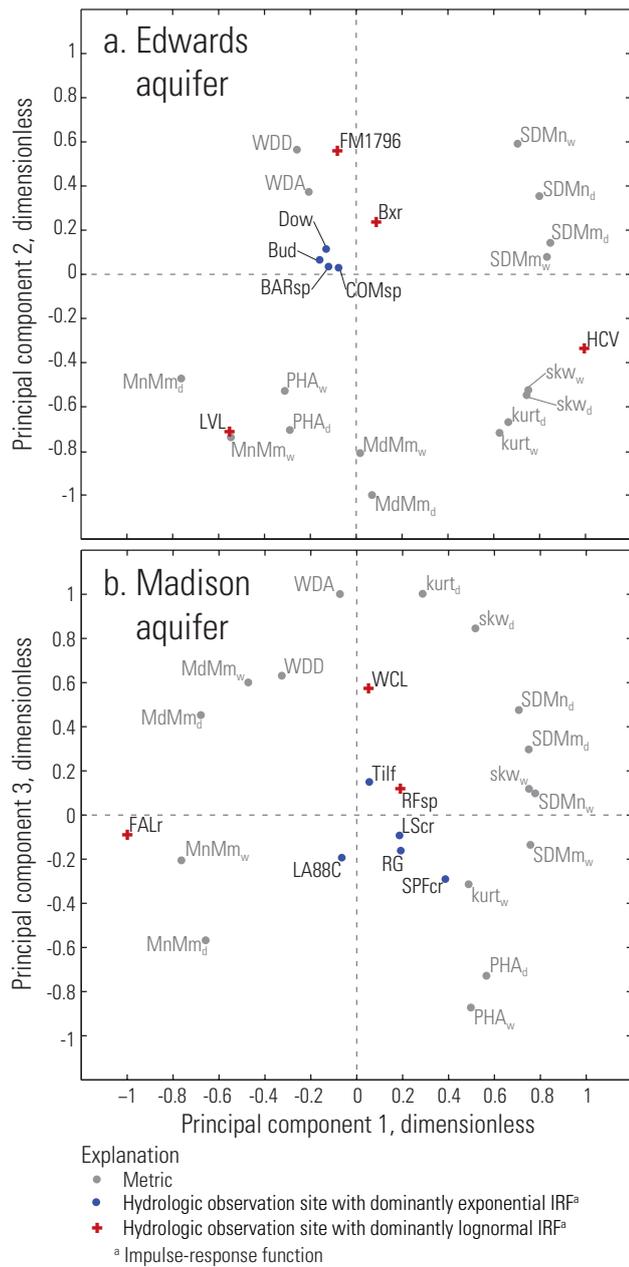
In general, the metrics are well distributed in principal component space, indicating that each contains unique information related to the shape of the IRF and verifying their inclusion in the analysis (Fig. 4). PC1, PC2, and PC3 describe three generalized variables that account for 84 % of the variability in the IRF shapes. PC1 represents the IRF skewness and temporal spread, i.e., the width of the main body of the IRF. PC1 separates those metrics that include standard deviation, skewness, and kurtosis from those that include the mean or mode in relation to the memory (Fig. 4). PC2 represents differences in the wet- and dry-period counterparts of each metric; for metrics on the positive side of PC1, the dry-period counterparts plot more positively on PC2 than the wet-period counterparts, and for those on the negative side of PC1, the opposite is true (Fig. 4a). PC3 represents differences in wet and dry periods overall, with the metrics WDD and WDA separated from other metrics on this axis (Fig. 4b).

To assess differences between the two aquifers, PCA was done separately for each aquifer. Time-variant IRFs, which have metrics that differ between wet and dry periods, were



**Fig. 4.** Principal component biplots showing simulated sites and metrics for the Edwards and Madison aquifers for (a) the first two principal components and (b) the first and third principal components. Plotting positions were scaled to range from  $-1$  to  $1$  to show sites and metrics on a similar scale. Site and metric labels are shown in Tables 2 and 3. See Table S6 in the Supplement for plotting positions.

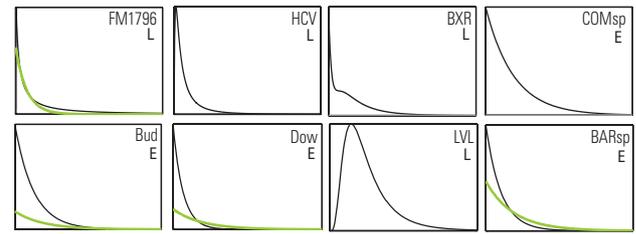
applied to 50 % of the Edwards aquifer sites and 75 % of the Madison aquifer sites (Table 4). The wet- and dry-period counterparts of four metrics (krt, skw, SDMm, SDMn) plotted farther apart for the Madison aquifer than for the Edwards aquifer (Fig. 5), indicating that the largest differences between the wet- and dry-period IRFs are associated with the



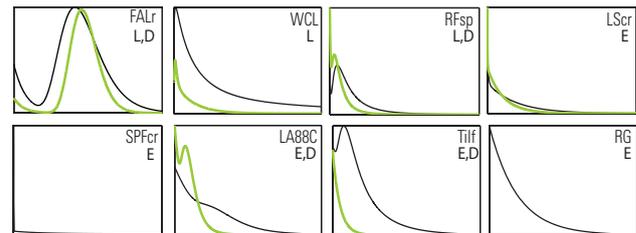
**Fig. 5.** Principal component biplots showing simulated sites and metrics for the first two principal components for the (a) Edwards aquifer and (b) Madison aquifer. Plotting positions were scaled to range from  $-1$  to  $1$  to show sites and metrics on a similar scale. Site and metric labels are shown in Tables 2 and 3.

Madison aquifer. These metrics dominate the positive side of PC1 for both aquifers (Fig. 5). The wet and dry counterparts for the remaining metrics (primarily negative on PC1) plotted about the same distance apart for the two aquifers. Differences in the IRF shapes for wet and dry periods (Fig. 6) are likely the result of different parts of a heterogeneous aquifer being saturated during these different climatic periods; large

Edwards Aquifer



Madison Aquifer



— IRF, wet period  
 — IRF, dry period  
 E = Dominantly exponential  
 L = Dominantly lognormal  
 D = Double peak

**Fig. 6.** Impulse-response functions (IRFs) plotted for scale-independent comparison. The vertical axes extend from zero to the maximum peak height, and the horizontal axes extend from zero to the system memory.

differences in IRF shape for the Madison aquifer suggest that this aquifer has the larger heterogeneity of the two aquifers.

Similar IRF shapes plot in similar locations in principal component space; large separation indicates large differences in IRF shape. The separation of each site to every other site was quantified by the Euclidean distance in principal component space for the analysis that included both aquifers (Fig. 4). The mean separation distance for Madison aquifer sites was 2.4 times greater than for the Edwards. This difference was statistically significant at the 95 % confidence level according to the rank sum test (Helsel and Hirsch, 2002). This indicates that the IRFs are more similar for the Edwards aquifer than for the Madison aquifer, which is further evidence that the Madison aquifer has the larger heterogeneity of the two.

The sites compose two groups whose IRFs dominantly comprise either lognormal or exponential curves, on the basis of curve area. Seven sites have dominantly lognormal IRFs; nine sites have dominantly exponential IRFs; and the differences are evident in their plotting positions in principal component space (Figs. 4 and 5). Exponential IRFs describe hydraulic-response types that have an immediate peak response, whereas lognormal IRFs describe a lagged peak response. Therefore, exponential IRFs describe a rapid transfer of pressure from the recharge area to the well or spring, which is delayed for lognormal IRFs. Minimal displacement

of water in confined aquifers can result in a rapid transfer of pressure, but the larger displacement of water in unconfined aquifers can result in a delayed transfer of pressure. In karst aquifers, however, confined and unconfined conditions can exist in the same area, because partially and fully saturated conduits and fractures can exist at different levels in the aquifer. For some of the sites, double-peaked IRFs resulted from a combination of an exponential and a lognormal curve (Fig. 6), which might be the result of this dual nature of karst aquifers but also might result from another dual aspect of karst aquifers, i.e., quick conduit flow and slow flow in surrounding fracture networks. The Madison aquifer has four double-peaked sites, whereas the Edwards aquifer has none (Fig. 6), which is a third line of evidence that the Madison aquifer has the larger heterogeneity of the two aquifers.

Although it is difficult to determine the specific aquifer characteristics that will result in a particular IRF shape, the spatial distribution in the aquifer of the two major categories of IRFs (exponential and lognormal) suggests that these aquifer characteristics are not randomly distributed. For the Edwards aquifer, sites with dominantly exponential IRFs are in the northeastern part of the aquifer, and sites with dominantly lognormal IRFs are in the southwestern part, with the exception of site LVL (Fig. 2). For the Madison aquifer, sites with dominantly exponential IRFs are in the northeastern part of the Black Hills, and sites with dominantly lognormal IRFs are in the southwestern part (Fig. 2).

Hydrogeologic features can be identified that might be related to differences in IRF shape. Therefore, we describe these differences and attempt to relate them to IRF characteristics where possible. Sites with dominantly lognormal IRFs are located where the Edwards aquifer is wide (Fig. 2); in this area Eberts (2011) reported that groundwater flows southwesterly, then sweeps around to the northeast, with a strong regional flow component. Sites with dominantly exponential IRFs are located where the Edwards aquifer is narrow, and flow in this area is to the north-northeast along a series of conduit-controlled flow paths (Hunt et al., 2006). A hydrologic divide is between the two areas. Although not conclusive, the circuitous flow path near the lognormal IRF sites might partly account for the delayed peak response, whereas for the exponential IRF sites, the straight flow path might be related to the quick peak response. For the Madison aquifer, the lognormal IRF sites are located near the axis of the main anticline that defines the structural uplift of the Black Hills (Fig. 2). Fracturing and subsequent conduit development resulting from this anticline might be related to IRF characteristics.

Differences between springs and wells also were considered. Springs commonly discharge water from a large surrounding area, whereas none of the wells used in this study are pumped, so the zone of influence is small by comparison. Despite these differences, no distinction between springs and wells is evident from PCA (Figs. 4 and 5), which indicates

that the difference between springs and wells is small compared with regional heterogeneity in the aquifers.

Metric correlation is an indication of redundancy in the dataset (Table S5 in the Supplement). Positively correlated metric pairs generally are identified by proximity of plotting positions in Fig. 4a and b. For example, the group of five metrics that plot farthest to the right resulted in five metric pairs that have Pearson's  $r$  correlation coefficient  $\geq 0.9$  (Helsel and Hirsch, 2002). Negatively correlated metric pairs plot on opposite sides of the origin. The number of highly correlated pairs with  $r \geq |0.95|$  is only 2.3% of the total number of metric pairs. These pairs are  $MnMm_w/SDMn_w$ ,  $MnMm_d/SDMm_d$  (negative  $r$ ), and  $SDMn_d/SDMm_d$  (positive  $r$ ). The small number of highly correlated metric pairs indicates a low level of redundancy. These correlations do not correspond exactly to the separation distances in Fig. 4, because principal components 1–3 do not explain the total variance in the dataset.

#### 4 Conclusions

The convolution models are well suited to link climate scenarios to hydrologic responses critical to human water supply and ecosystems. As climate models improve in simulating storm processes, projected precipitation and air-temperature records can be used as input for hydrologic models to investigate how water levels and spring discharge might change in the future. The output from the hydrologic models can, in turn, serve as input for models evaluating vulnerability of karst ecosystems to potential changes in groundwater level and spring flow. Unlike most physically based, distributed-parameter models, convolution models are calibrated to data at short time steps that are well suited to short-term variability characteristic of karst groundwater responses. The impulse-response function (IRF) is the core of the convolution model and characterizes the groundwater-system functioning.

The use of multiple IRF metrics in principal component analysis (PCA) is a novel method to characterize, compare, and classify the way in which multiple sites and aquifers respond to recharge. When applied to several sites in two karst aquifers in the United States (the Edwards and Madison aquifers), this classification indicates a wide range of IRF shapes within a single karst aquifer that vary spatially and temporally. The IRF shape is a result of the aquifer's pore geometry and connectivity, which is horizontally and vertically heterogeneous in karst aquifers. The first three principle components (PC1, PC2, and PC3) account for 84% of the variability in the IRF shapes for the Edwards and Madison aquifers; these components represent differences in (1) IRF skewness and temporal spread, (2) wet- and dry-period counterparts of each metric, and (3) wet- and dry-period IRFs overall, respectively. PC1 accounts for 51% of IRF variability for the two aquifers, and PC2 and PC3 combined, which

largely represent time-variant properties, account for 33 % of IRF variability.

Three lines of evidence indicate that the Madison aquifer has larger spatial heterogeneity than the Edwards aquifer. First, the Madison aquifer has larger spatial differences in IRF shapes on the study-area scale than does the Edwards aquifer. Second, the Madison aquifer has the largest temporal differences in IRF shapes at individual sites as a result of different parts of the heterogeneous aquifer being saturated during different climatic periods. Third, the Madison aquifer has four sites with double-peaked IRFs, possibly because of large differences in quick and slow flow, whereas the Edwards aquifer has none.

Sites with IRFs that dominantly comprise exponential curves are separated geographically from those dominantly comprising lognormal curves in both aquifers. Differences in groundwater flow paths (i.e., direct or circuitous), groundwater divides, and structural geologic features might be related to these different response types. No distinction between springs and wells is evident from PCA, which indicates that differences between springs and wells are small compared with regional heterogeneity in the aquifers.

Currently, IRFs have been developed for only two karst aquifers, but as convolution models are developed for additional sites in additional aquifers, they could contribute to an IRF database and, moreover, a general classification system for karst aquifers. Of particular interest will be comparison of telogenetic and eogenetic systems, humid and arid systems, and diffuse- and conduit-controlled systems. Characterization and classification of time-variant properties will be useful for assessing model uncertainty for simulation of future precipitation scenarios that are outside of the range of observation. The availability of long-term records for precipitation, air temperature, water level, and spring flow in many areas would facilitate this effort.

**Supplementary material related to this article is available online at: <http://www.hydrol-earth-syst-sci.net/17/281/2013/hess-17-281-2013-supplement.zip>.**

*Acknowledgements.* The authors would like to thank Steffen Birk, Jos Von Asmuth, and an anonymous referee for their review comments, which led to substantial improvements in the manuscript. Also, Johnathan Bumgarner provided helpful comments on an early draft of the manuscript. We thank Jonathan McKaskey, Thomas Sample, and Jennifer Bednar for assistance with obtaining data and creating figures and tables. Any use of trade, firm, or product names is for descriptive purposes only and does not imply endorsement by the US Government.

Edited by: M. Bakker

## References

- Ashworth, J. B. and Hopkins, J.: Major and minor aquifer of Texas, Texas Water Development Board Report 345, 60 pp., 1995.
- Back, W., Hanshaw, B. B., Plummer, L. N., Rahn, P. H., Rightmire, C. T., and Rubin, M.: Process and rate of dedolomitization: Mass transfer and  $^{14}\text{C}$  dating in a regional carbonate aquifer, *Geol. Soc. Am. Bull.*, 94, 1415–1429, 1983.
- Berendrecht, W. L., Heemink, A. W., Van Geer, F. C., and Gehrels, J. C.: Decoupling of modeling and measuring interval in groundwater time series analysis based on response characteristics, *J. Hydrol.*, 278, 1–16, 2003.
- Carter, J. M. and Driscoll, D. G.: Estimating recharge using relations between precipitation and yield in a mountainous area with large variability in precipitation, *J. Hydrol.*, 316, 71–83, 2006.
- Coleman, T. F. and Li, Y.: On the convergence of reflective newton methods for large-scale nonlinear minimization subject to bounds, *Math. Prog.*, 67, 189–224, 1994.
- Coleman, T. F. and Li, Y.: An interior, trust region approach for nonlinear minimization subject to bounds, *SIAM J. Optimiz.*, 6, 418–445, 1996.
- Covington, M. D., Wicks, C. M., and Saar, M. O.: A dimensionless number describing the effects of recharge and geometry on discharge from simple karstic aquifers, *Water Resour. Res.*, 45, W11410 1–16, 2009.
- Covington, M. D., Luhmann, A. J., Wicks, C. M., and Saar, M. O.: Process length scales and longitudinal damping in karst conduits, *J. Geophys. Res.-Earth*, 117, F01025, doi:10.1029/2011JF002212, 2012.
- Davis, J. C.: *Statistics and Data Analysis in Geology*, John Wiley & Sons, Inc., Hoboken, New Jersey, 2002.
- Denić-Jukić, V. and Jukić, D.: Composite transfer functions for karst aquifers, *J. Hydrol.*, 274, 80–94, 2003.
- Dooge, J. C. I.: *Linear theory of hydrologic systems*, US Dep. of Agriculture and Tech. Bull., Washington, DC, 1468 pp., 1973.
- Driscoll, D. G. and Carter, J. M.: Hydrologic conditions and budgets for the Black Hills of South Dakota, through water year 1998, US Geol. Surv., Reston, Virginia, WRIR 2001–4226, 143 pp., 2001.
- Eberts, S. M. (Ed.): *Hydrogeologic settings and groundwater-flow simulations for regional investigations of the transport of anthropogenic and natural contaminants to public-supply wells – Investigations begun in 2004*, US Geol. Surv., Reston, Virginia, PP 1737–B, 127 pp., 2011.
- Greene, E. A. and Rahn, P. H.: Localized anisotropic transmissivity in a karst aquifer, *Ground Water*, 33, 806–816, 1995.
- Hall, F. R. and Moench, A. F.: Application of the convolution equation to stream-aquifer relationships, *Water Resour. Res.*, 8, 487–493, 1972.
- Helsel, D. R. and Hirsch, R. M.: *Statistical methods in water resources*, US Geol. Surv., TWRI, Book 4, Ch. A3, 522 pp., 2002.
- Hortness, J. E. and Driscoll, D. G.: Streamflow losses in the Black Hills of western South Dakota, US Geol. Surv., Reston, Virginia, WRIR 98–4116, 99 pp., 1998.
- Hostetler, S. W., Alder, J. R. and Allan, A. M.: Dynamically downscaled climate simulations over North America: Methods, evaluation and supporting documentation for users, US Geol. Surv., Reston, Virginia, OFR 2011–1238, 64 pp., 2011.

- Hunt, B. B., Smith, B. A., Beery, J., Johns, D., and Hauwert, N.: Summary of 2005 groundwater dye tracing, Barton Springs segment of the Edwards aquifer, Hays and Travis Counties, Central Texas, Barton Springs/Edwards Aquifer Conservation District (BSEACD) Report of Investigations 2006–0530, 19 pp., 2006.
- Jakeman, A. J. and Hornberger, G. M.: How much complexity is warranted in a rainfall-runoff model?, *Water Resour. Res.*, 29, 2637–2649, 1993.
- Jeannin, P. Y.: Modeling flow in phreatic and epiphreatic karst conduits in the Hölloch cave (Muotatal, Switzerland), *Water Resour. Res.*, 37, 191–200, 2001.
- Jenkins, G. M. and Watts, D. G.: *Spectral Analysis and its Applications*, Holden Day, San Francisco, 525 pp., 1968.
- Jukić, D. and Denić-Jukić, V.: Nonlinear kernel functions for karst aquifers, *J. Hydrol.*, 328, 360–374, 2006.
- Jukić, D. and Denić-Jukić, V.: Partial spectral analysis of hydrological time series, *J. Hydrol.*, 400, 223–233, 2011.
- Labat, D., Masbou, J., Beaulieu, E., and Mangin, A.: Scaling behavior of the fluctuations in stream flow at the outlet of karstic watersheds, France, *J. Hydrol.*, 410, 162–168, 2011.
- Legates, D. R. and McCabe Jr., G. J.: Evaluating the use of “goodness-of-fit” measures in hydrologic and hydroclimatic model validation: *Water Resour. Res.*, 35, 233–241, 1999.
- Lindgren, R. J., Dutton, A. R., Hovorka, S. D., Worthington, S. R. H., and Painter, S.: Conceptualization and simulation of the Edwards aquifer, San Antonio region, Texas, *US Geol. Surv., Reston, Virginia, SIR 2004–5277*, 154 pp., 2004.
- Loáiciga, H. A., Maidment, D. R., Valdes, J. B., Roos, M., and Wood, E. F.: Report of the case study selection team: ASCE/EPA climate change cooperative agreement, *Am. Soc. of Civil Eng., Washington, DC*, 17 pp., 1996.
- Long, A. J.: Hydrograph separation for karst watersheds using a two-domain rainfall-discharge model, *J. Hydrol.*, 364, 249–256, 2009.
- Long, A. J. and Derickson, R. G.: Linear systems analysis in a karst aquifer, *J. Hydrol.*, 219, 206–217, 1999.
- Long, A. J. and Putnam, L. D.: Translating CFC-based piston ages into probability density functions of ground-water age in karst, *J. Hydrol.*, 330, 735–747, 2006.
- Love, J. D. and Christiansen, A. C.: *Geologic map of Wyoming*, US Geol. Surv., Reston, Virginia, Map WDB 904, 1985.
- Luhmann, A. J., Covington, M. D., Peters, A. J., Alexander, S. C., Anger, C. T., Green, J. A., Runkel, A. C., and Alexander, E. C.: Classification of thermal patterns at karst springs and cave streams, *Ground Water*, 49, 324–335, 2011.
- Massei, N., Dupont, J. P., Mahler, B. J., Laignel, B., Fournier, M., Valdes, D., and Ogier, S.: Investigating transport properties and turbidity dynamics of a karst aquifer using correlation, spectral, and wavelet analyses, *J. Hydrol.*, 329, 244–257, 2006.
- Maclay, R. W. and Small, T. A.: Hydrostratigraphic subdivisions and fault barriers of the Edwards aquifer, south-central Texas, USA, *J. Hydrol.*, 61, 127–146, 1983.
- Mearns, L. O., Gutowski, W. J., Jones, R., Leung, L. Y., McGinnis, S., Nunes, A. M. B., and Qian, Y.: A regional climate change assessment program for North America, *Eos Transactions, Am. Geophys. Un.*, 90, 311–312, 2009.
- Miller, L. D. and Driscoll, D. G.: Streamflow characteristics for the Black Hills of South Dakota, through water year 1993, *US Geol. Surv., Reston, Virginia, SIR 97–4288*, 322 pp., 1998.
- Nash, J. E.: Systematic determination of unit hydrograph parameters, *J. Geophys. Res.*, 54, 111–115, 1959.
- Nash, J. E. and Sutcliffe, J. V.: River flow forecasting through conceptual models part I – A discussion of principles, *J. Hydrol.*, 10, 282–290, 1970.
- National Climatic Data Center, National Oceanic and Atmospheric Administration, US Department of Commerce, available at: <http://www.ncdc.noaa.gov>, last access: 14 June 2012.
- Olsthoorn, T. N.: Do a bit more with convolution, *Ground Water*, 46, 13–22, 2008.
- Pinault, J. L., Plagnes, V., Aquilina, L., and Bakalowicz, M.: Inverse modeling of the hydrological and the hydrochemical behavior of hydrosystems; characterization of karst system functioning, *Water Resour. Res.*, 37, 2191–2204, 2001.
- Rahn, P. H. and Gries, J. P.: Large springs in the Black Hills, South Dakota and Wyoming, Report of Investigations – Department of Natural Resource Development, South Dakota Geological Survey, 107, p. 100, 1973.
- Redden, J. A. and DeWitt, E.: Maps showing geology, structure, and geophysics of the central Black Hills, South Dakota, 44 pp., 2 sheets, 2008.
- Reimann, T., Geyer, T., Shoemaker, W. B., Liedl, R., and Sauter, M.: Effects of dynamically variable saturation and matrix-conduit coupling of flow in karst aquifers, *Water Resour. Res.*, 47, W11503, doi:10.1029/2011WR010446, 2011.
- Simoni, S., Padoan, S., Nadeau, D. F., Diebold, M., Porporato, A., and Barrentxea, G.: Hydrologic response of an alpine watershed: application of a meteorological wireless sensor network to understand streamflow generation, *Water Resour. Res.*, 47, W10524, doi:10.1029/2011WR010730, 2011.
- Small, T. A., Hanson, J. A., and Hauwert, N. M.: Geologic framework and hydrogeologic characteristics of the Edwards aquifer outcrop (Barton Springs segment), northeastern Hays and southwestern Travis Counties, Texas, *US Geol. Surv., Denver, CO, WRIR 96–4306*, 15 pp., 1996.
- Smart, C. and Worthington, S. R. H.: Springs, in: *Encyclopedia of Caves and Karst Science*, edited by: Gunn, J., Fitzroy Dearborn, New York, 669–703, 2004.
- Strobel, M. L., Jarrell, G. J., Sawyer, J. F., Schleicher, J. R., and Fahrenbach, M. D.: Distribution of hydrogeologic units in the Black Hills Area, South Dakota, *US Geol. Surv., Reston, Virginia, HA–743*, 1999.
- US Geological Survey, National Water Information System (NWIS Web): US Geological Survey database, available at: <http://waterdata.usgs.gov/usa/nwis/nwis>, last access: 14 June 2012.
- Von Asmuth, J. R. and Knotters, M.: Characterising groundwater dynamics based on a system identification approach, *J. Hydrol.*, 296, 118–134, 2004.
- Von Asmuth, J. R., Bierkens, M. F. P., and Maas, K.: Transfer function-noise modeling in continuous time using predefined impulse response functions, *Water Resour. Res.*, 38, 1287, doi:10.1029/2001WR001136, 2002.
- White, W. B.: *Geomorphology and Hydrology of Karst Terrains*, Oxford University Press, New York, 1988.