



This discussion paper is/has been under review for the journal Hydrology and Earth System Sciences (HESS). Please refer to the corresponding final paper in HESS if available.

Inverse modelling of in situ soil water dynamics: accounting for heteroscedastic, autocorrelated, and non-Gaussian distributed residuals

B. Scharnagl¹, S. C. Iden², W. Durner², H. Vereecken³, and M. Herbst³

¹Department of Soil Physics, Helmholtz-Zentrum für Umweltforschung – UFZ, 06102 Halle, Germany

²Institute of Geoecology, Technische Universität Braunschweig, 38106 Braunschweig, Germany

³Agrosphere Institute (IBG-3), Forschungszentrum Jülich, 52425 Jülich, Germany

Received: 26 January 2015 – Accepted: 3 February 2015 – Published: 19 February 2015

Correspondence to: B. Scharnagl (benedikt.scharnagl@ufz.de)

Published by Copernicus Publications on behalf of the European Geosciences Union.

HESSD

12, 2155–2199, 2015

Inverse modelling of in situ soil water dynamics

B. Scharnagl et al.

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures



Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



Abstract

Inverse modelling of in situ soil water dynamics is a powerful tool to test process understanding and determine soil hydraulic properties at the scale of interest. The observations of soil water state variables are typically evaluated using the ordinary least squares approach. However, the underlying assumptions of this classical statistical approach of independent, homoscedastic, and Gaussian distributed residuals are rarely tested in practice. In this case study, we estimated the soil hydraulic properties of a homogeneous, bare soil profile from field observations of soil water contents. We used a formal Bayesian approach to estimate the posterior distribution of the parameters in the van Genuchten–Mualem (VGM) model of the soil hydraulic properties. Three likelihood models that differ with respect to assumptions about the statistical features of the time series of residuals were used. Our results show that the assumptions of the ordinary least squares did not hold, because the residuals were strongly autocorrelated, heteroscedastic and non-Gaussian distributed. From a statistical point of view, the parameter estimates obtained with this classical statistical approach are therefore invalid. Since a test of the classic first-order autoregressive (AR(1)) model led to strongly biased model predictions, we introduced an modified AR(1) model which eliminates this critical deficit of the classic AR(1) scheme. The resulting improved likelihood model, which additionally accounts for heteroscedasticity and nonnormality, lead to a correct statistical characterization of the residuals and thus outperformed the other two likelihood models. We consider the corresponding parameter estimates as statistically correct and showed that they differ systematically from those obtained under ordinary least squares assumptions. Moreover, the uncertainty in the parameter estimates was increased by accounting for autocorrelation in the observations. Our results suggest that formal Bayesian inference using a likelihood model that correctly formalizes the statistical properties of the residuals may also prove useful in other inverse modelling applications in soil hydrology.

Inverse modelling of in situ soil water dynamics

B. Scharnagl et al.

[Title Page](#)

[Abstract](#)

[Introduction](#)

[Conclusions](#)

[References](#)

[Tables](#)

[Figures](#)



[Back](#)

[Close](#)

[Full Screen / Esc](#)

[Printer-friendly Version](#)

[Interactive Discussion](#)



1 Introduction

Modelling of in situ soil water dynamics requires knowledge of the soil hydraulic properties that characterize the soil at the site of interest. Various methods have been developed to determine the soil hydraulic properties on soil samples taken to the laboratory.

5 Among these laboratory methods, the multi-step outflow method (van Dam et al., 1994; Durner and Iden, 2011) and the evaporation method (Šimůnek et al., 1998; Peters and Durner, 2008) have found widespread application because they allow for the determination of both the soil water retention function and the soil hydraulic conductivity function from a single experiment.

10 However, there is ever increasing evidence in the soil hydrological literature that the hydraulic properties determined on soil cores in the laboratory are of limited use for predictive modelling of soil water dynamics under field conditions (e.g. Mertens et al., 2005; Baroni et al., 2010). The possible reasons for this are manifold. For example, Mallants et al. (1997) suggested that the sample volume analysed in the laboratory
15 may not be a representative elementary volume. Another likely reason was discussed by Basile et al. (2003), who stressed that laboratory experiments typically aim at the determination of the soil hydraulic functions that belong to the primary drainage curve, that is, starting from an initial condition that is as closely as possible to complete saturation. This condition, however, basically never occurs in situ, and the so derived
20 properties may therefore not be representative for in situ conditions. Closely related to the argument of Basile et al. (2003) is the fact that hysteresis is typically neglected in modelling of in situ soil water dynamics, even though it is widely acknowledged to be an intrinsic feature of soils which affects soil water dynamics (e.g. Dane and Wierenga, 1975; Lenhard et al., 1991). Other effects, like dynamic non-equilibrium (Diamantopoulos and Durner, 2012), which may be significant in some situations but are neglected in
25 standard modelling approaches may further limit the usefulness of the hydraulic properties derived in the laboratory to predict water dynamics in field soils.

HESSD

12, 2155–2199, 2015

Inverse modelling of in situ soil water dynamics

B. Scharnagl et al.

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures



Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



Inverse modelling of in situ soil water dynamics

B. Scharnagl et al.

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures



Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



Given these limitations it is promising to make use of in situ observations of soil water state variables to obtain more useful estimates of the soil hydraulic properties by inverse modelling (Vereecken et al., 2008). The soil hydraulic properties derived in this way give the best possible description of the observational data. Since they are conditional on the model formulation used for parameter inference, they are also deemed particular useful for predictive modelling. However, as the soil hydraulic properties are part of the flow model, they may compensate for deficits in the model formulation and for errors in the forcing variables.

From a statistical point of view, the estimation of soil hydraulic properties by inverse modelling of in situ observations of soil water state variables is afflicted with two complicating features of this type of observational data. The first complicating feature is that in situ observations generally do not contain sufficient information to warrant accurate and precise estimation of the soil hydraulic properties (Scharnagl et al., 2011). The reason for this is that in situ data cover only a relative narrow range of soil water states. This is true even for observations taken close to the surface where soil water dynamics are highest.

To account for the limited amount of information of in situ observations, it is particularly beneficial to make explicit use of prior information about the soil hydraulic parameters in the inference process. This can be done using Bayesian statistics, which offers a formal way to combine the information on the parameters before assimilating field data and the information contained in the observational data. This formal Bayesian approach was applied by Hou and Rubin (2005), who incorporated prior information about the soil hydraulic properties in the calibration of a one-layer soil hydrological model using neutron probe and time domain reflectory (TDR) measurements. They used the principle of minimum relative entropy (Woodbury and Ulrych, 1993) to define a prior probability density function of the soil hydraulic parameters from information about the expected values, variances, as well as the lower and upper bounds of the parameters derived from the ROSETTA pedotransfer function (Schaap et al., 2001). Scholer et al. (2011) investigated the information content of ground-penetrating radar measure-

HESSD

12, 2155–2199, 2015

Inverse modelling of in situ soil water dynamics

B. Scharnagl et al.

[Title Page](#)

[Abstract](#)

[Introduction](#)

[Conclusions](#)

[References](#)

[Tables](#)

[Figures](#)



[Back](#)

[Close](#)

[Full Screen / Esc](#)

[Printer-friendly Version](#)

[Interactive Discussion](#)



ments for estimating the soil hydraulic properties of a four-layer profile. They used prior information about the soil hydraulic parameters either taken from Carsel and Parrish (1988) or derived from independent measurements on core samples taken at the field site. Scholer et al. (2011) concluded that prior information helped to substantially constrain the parameter space, especially when a prior density was used that includes information about the correlation structure of the soil hydraulic parameters. Concurring results were presented by Scharnagl et al. (2011) who used spatially distributed TDR measurements to estimate the effective soil hydraulic parameters of a single-layer profile. They implemented a Monte Carlo approach that allows to estimate the correlation structure of the soil hydraulic parameters predicted by ROSETTA, which is not available from the standard output of this pedotransfer function. Scharnagl et al. (2011) found that only if the full covariance matrix was used to define the prior density of the soil hydraulic parameters, the Bayesian approach was effective and robust in constraining the parameters to a small region of the parameter space.

The second complicating feature of in situ observations is that the corresponding residuals, that is, the deviations between observations and model predictions, are often strongly autocorrelated. The obvious reason for this are model errors, that is, errors in the model structure as well as errors in the forcing variables that are used to run the model, which all together result in a systematic misfit between observations and model predictions. Doherty and Welter (2010) provide an overview of the various formal and informal statistical approaches that are currently used to explicitly treat autocorrelation in the residuals.

Here, we adopted a formal statistical approach, which requires that the autocorrelation structure of the residuals is considered in the inference process. This can be done using the classical first-order autoregressive (AR(1)) model. This model has found widespread application in many fields of science, such as econometrics (e.g. Cochrane and Orcutt, 1949; Beach and MacKinnon, 1978) or catchment hydrology (e.g. Sorooshian and Dracup, 1980; Kuczera, 1983). Quite surprisingly, up to now it has received very little attention in the soil hydrological literature, where autocorrelation

Inverse modelling of in situ soil water dynamics

B. Scharnagl et al.

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures



Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



of the residuals as well as possible deviations from the assumptions of homogeneity and normality are typically neglected in the inference process. We are currently aware of only one study that considered autocorrelation of the residuals in inverse modelling of soil hydrological processes. Wöhling and Vrugt (2011) implemented the classical AR(1) scheme to estimate the soil hydraulic properties of a five-layer profile from multiple, parallel time series of soil water content and pressure head recorded at five depths. They found that the calibrated model did not provide a simultaneous fit of both pressure head and soil water content observations, with the model predicted soil water contents deviating substantially and consistently from the corresponding observations. Wöhling and Vrugt (2011) suspected that the reason for this failure are structural errors in the soil hydrological model that cannot be captured by the AR(1) scheme.

In this study, we considered a simple example of inverse modelling of in situ soil water dynamics: a bare, homogeneous soil profile for which a time series of soil water content observations near the surface was available. The inverse problem was posed in a formal Bayesian framework and solved using Markov chain Monte Carlo (MCMC) simulation. The general problem of limited information content of in situ data was taken into account by using an informative prior distribution of the soil hydraulic parameters, as was previously done by Scharnagl et al. (2011) at the same field site. We applied three different likelihood models that differ in their underlying assumptions about the statistical features of the error. The first likelihood model corresponds exactly to the ordinary least-squares approach, which is the de facto standard in inverse modelling of soil hydrological processes. The second likelihood model implements the classical AR(1) model to account for autocorrelation, which was previously used by Wöhling and Vrugt (2011) in a soil hydrological application. This second likelihood model additionally allows for heteroscedastic and non-Gaussian distributed residuals. In the third likelihood model, we introduced a modified AR(1) scheme that overcomes a critical deficiency of the classical scheme.

where θ ($\text{cm}^3 \text{cm}^{-3}$) is the soil water content, K (cm h^{-1}) is the soil hydraulic conductivity, h (cm) denotes the pressure head, t is time (h), and z defines the vertical coordinate (cm, positive upward). To solve the Richards equation for given initial and boundary conditions, we used the HYDRUS-1-D code of Šimůnek et al. (2008).

The soil hydraulic properties, $\theta(h)$ and $K(h)$, were described with the van Genuchten–Mualem (VGM) model (van Genuchten, 1980). The water retention function $\theta(h)$, which is expressed here in terms of effective saturation S (dimensionless), is given by:

$$S(h) = \frac{\theta(h) - \theta_r}{\theta_s - \theta_r} = \begin{cases} \frac{1}{(1+|\alpha h|^n)^m} & \text{for } h \leq 0 \\ 1 & \text{for } h > 0 \end{cases} \quad (2)$$

where θ_r and θ_s ($\text{cm}^3 \text{cm}^{-3}$) represent the residual and saturated water content, respectively, and α (cm^{-1}), n , and $m = 1 - 1/n$ (both dimensionless) are shape parameters. The hydraulic conductivity function $K(h)$ is given by:

$$K(h) = K_s S(h)^L \left(1 - \left(1 - S(h)^{\frac{1}{m}} \right)^m \right)^2 \quad (3)$$

where K_s (cm h^{-1}) is the saturated hydraulic conductivity, and L (dimensionless) is an additional shape parameter that increases the flexibility of the conductivity model.

The upper boundary condition was defined using hourly averaged measurements of precipitation, air temperature, relative humidity, wind speed, incoming shortwave radiation, and atmospheric pressure taken at a meteorological station located about 100 m west of the measurement site. As input to the HYDRUS-1-D model, we used hourly precipitation P , and potential evaporation E_{pot} , where E_{pot} was calculated from observed meteorological variables using the method of Allen et al. (1998). A mixed boundary condition was used at the upper boundary that switches between a flux condition and a pressure head condition when the pressure head at the upper boundary reaches the lower limit of $h_{\text{UB}}^{\text{min}} = -100\,000$ cm.

**Inverse modelling of
in situ soil water
dynamics**

B. Scharnagl et al.

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures



Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



Title Page	
Abstract	Introduction
Conclusions	References
Tables	Figures
⏪	⏩
◀	▶
Back	Close
Full Screen / Esc	
Printer-friendly Version	
Interactive Discussion	



By assuming that the standardised residuals follow a Gaussian distribution, the likelihood of an individual standardised residual $\underline{\varepsilon}_i$, and hence, of the corresponding observation \tilde{y}_i is:

$$p(\tilde{y}_i | \mathbf{x}_1, \mathbf{x}_2, \sigma_{\text{const}}) = \frac{1}{\sqrt{2\pi}\sigma_{\text{const}}} \exp\left(-\frac{\varepsilon_i^2}{2}\right) \quad i = 1, \dots, N. \quad (9)$$

Based on the additional assumption that the residuals are independent, the joint likelihood of all observation $\tilde{y}_1, \dots, \tilde{y}_N$ is:

$$p(\tilde{y}_1, \dots, \tilde{y}_N | \mathbf{x}_1, \mathbf{x}_2, \sigma_{\text{const}}) = \prod_{i=1}^N p(\tilde{y}_i | \mathbf{x}_1, \mathbf{x}_2, \sigma_{\text{const}}). \quad (10)$$

The ordinary least-squares approach is the de facto standard in inverse modelling of soil hydrological processes (e.g. Jacques et al., 2002; Sonnleitner et al., 2003; Wollschläger et al., 2009; Schelle et al., 2012). However, its underlying assumptions are rarely tested in practice. In this study, the assumptions of homoscedastic, independent, and Gaussian distributed residuals did not hold in the case of inverse modelling of in situ soil water dynamics and good statistical practice then requires that the likelihood model is revised accordingly.

Likelihood 2: The second likelihood model allows for heteroscedastic and non-Gaussian distributed residuals and implements the classical AR(1) scheme to account for serial dependence of the residuals. To parametrise heteroscedasticity, we modelled the SD of the i th residual as a function of the simulated water content, $\sigma_i(y_i)$, using a piecewise cubic Hermite interpolating polynomial (PCHIP, Fritsch and Carlson, 1980). The PCHIP approach offers great flexibility in modelling the underlying, but a priori unknown variation of the SD of the residuals as

Inverse modelling of
in situ soil water
dynamics

B. Scharnagl et al.

Title Page	
Abstract	Introduction
Conclusions	References
Tables	Figures
◀	▶
◀	▶
Back	Close
Full Screen / Esc	
Printer-friendly Version	
Interactive Discussion	



Discussion Paper | Discussion Paper | Discussion Paper | Discussion Paper | Discussion Paper

a function of the simulated water content. It was inspired by the “free-form” approach (Bitterlich et al., 2004; Iden and Durner, 2007) to model the soil hydraulic properties.

The PCHIP is calculated on four points, defined by matrix Ω (size 4×2). The relationship $\sigma_i(y_i)$ can then formally be written as:

$$\text{PCHIP} : \Omega = \begin{bmatrix} \theta_a & \sigma_a \\ \theta_b & \sigma_b \\ \theta_c & \sigma_c \\ \theta_d & \sigma_d \end{bmatrix} \rightarrow \sigma_i(y_i) \quad i = 1, \dots, N \quad (11)$$

where $\theta_{a,\dots,d}$ and $\sigma_{a,\dots,d}$ represent the x and y ordinates of the four points, respectively. The x ordinates of the first and last of these points were defined by $\theta_a = 0.9 \tilde{y}_{\min}$ and $\theta_d = 1.1 \tilde{y}_{\max}$, respectively, where \tilde{y}_{\min} denotes the minimum and \tilde{y}_{\max} the maximum of the soil water content observations. This definition of θ_a and θ_d was chosen to avoid extrapolation of the PCHIP. The remaining two x ordinates, θ_b and θ_c , were set equidistantly between the minimum and maximum values. The corresponding y ordinates, $\sigma_{a,\dots,d}$, are unknown parameters of the likelihood model.

In a next step, we consider the autocorrelation of the residuals using the classical AR(1) scheme. Its application presumes homoscedasticity of the residuals (e.g. Box et al., 2008; Evin et al., 2013) and has therefore to be applied to the standardised residuals. The classical AR(1) scheme is defined by:

$$\eta_i(\mathbf{x}_1, \mathbf{x}_2, \Omega, \phi) = \varepsilon_i - \phi \varepsilon_{i-1} \quad i = 2, \dots, N \quad (12)$$

where η_i represents the decorrelated standardised residuals and ϕ is the autocorrelation coefficient. The decorrelated residuals η_i can be considered as the remaining, unexplained deviations between observations and corresponding model predictions after removal of the systematic deviations that can be explained by

autocorrelation. The autocorrelation coefficient ϕ is an additional parameter of the likelihood model. In general, for a AR(1) process to be stationary, the autocorrelation coefficient must satisfy $|\phi| < 1$ (e.g. Box et al., 2008).

The classical AR(1) scheme in Eq. (12) was successfully applied to inverse modelling in catchment hydrology (e.g. Sorooshian and Dracup, 1980; Kuczera, 1983; Bates and Campbell, 2001; Vrugt et al., 2009b). It failed, however, in the soil hydrological application reported by Wöhling and Vrugt (2011). To understand the reason for this failure, it is instructive to look at the expectation of the decorrelated and restandardised residuals $\underline{\eta}_i$. Taking expectations on both sides of Eq. (12) gives:

$$\begin{aligned} E(\eta_i) &= E(\underline{\varepsilon}_i) - E(\phi)E(\underline{\varepsilon}_{i-1}) \\ &= (1 - E(\phi)) E(\underline{\varepsilon}_i) \quad i = 2, \dots, N \end{aligned} \quad (13)$$

where we used $E(\underline{\varepsilon}_i) = E(\underline{\varepsilon}_{i-1})$, which follows directly from the basic assumption given in Eq. (7). The expectation of the decorrelated residuals shows that the classical AR(1) scheme may introduce bias in the decorrelated residuals because the expectation of η_i will always be closer to zero than the expectation of $\underline{\varepsilon}_i$. This disadvantage is most pronounced if $E(\phi) \rightarrow 1$, in which case $E(\eta_i)$ approaches zero even if $E(\underline{\varepsilon}_i)$ differs markedly from zero. In fact, as we illustrated in this study, this bias may distort the inference process and may lead to meaningless results, very similar to those obtained by Wöhling and Vrugt (2011).

The transformation of the residuals implied by the AR(1) scheme in Eq. (12) needs to be accounted for in the formulation of the likelihood model. To do so, we need to restandardise the decorrelated residuals as:

$$\underline{\eta}_i(\mathbf{x}_1, \mathbf{x}_2, \mathbf{\Omega}, \phi) = \frac{\eta_i}{\sqrt{1 - \phi^2}} \quad i = 2, \dots, N \quad (14)$$

where $\sqrt{1 - \phi^2}$ is just the SD of the decorrelated residuals (e.g. Box et al., 2008).

Inverse modelling of in situ soil water dynamics

B. Scharnagl et al.

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures

⏪

⏩

◀

▶

Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



If we further assume that the restandardised decorrelated residuals $\underline{\eta}_i$ follow a Student distribution with unit variance, the likelihood of a single observation is given by:

$$p(\tilde{y}_i | \mathbf{x}_1, \mathbf{x}_2, \mathbf{\Omega}, \phi, \nu) = \frac{\Gamma\left(\frac{\nu+1}{2}\right) \sqrt{\frac{\nu}{\nu-2}}}{\Gamma\left(\frac{\nu}{2}\right) \sqrt{\pi\nu} \sqrt{1-\phi^2} \sigma_i} \left(1 + \frac{\underline{\eta}_i^2}{\nu-2}\right)^{-\frac{\nu+1}{2}} \quad i = 2, \dots, N \quad (15)$$

where Γ denotes the gamma function and $\nu > 2$ is a kurtosis parameter that controls the tailing behaviour of the distribution. For $\nu \rightarrow 2$, the tailing of the Student distribution becomes increasingly heavy, while for $\nu \rightarrow \infty$ the Student distribution converges to the Gaussian distribution. This is illustrated in Fig. 1a. The Student distribution has been proposed for the use in parameter inference by West (1984) and Lange et al. (1989), among others, because it has the ability to adequately fit heavy-tailed distributions, which also makes the inference process more robust against possible outliers in the observations.

In a next step, we applied the method of Fernández and Steel (1998) to introduce skewness in the symmetric Student distribution shown in Eq. (15). The skewed Student distribution is given by:

$$p(\tilde{y}_i | \mathbf{x}_1, \mathbf{x}_2, \mathbf{\Omega}, \phi, \nu, \kappa) = \frac{2c_2 \Gamma\left(\frac{\nu+1}{2}\right) \sqrt{\frac{\nu}{\nu-2}}}{\left(\kappa + \frac{1}{\kappa}\right) \Gamma\left(\frac{\nu}{2}\right) \sqrt{\pi\nu} \sqrt{1-\phi^2} \sigma_i} \times \left(1 + \frac{1}{\nu-2} \left(\frac{c_1 + c_2 \underline{\eta}_i}{\kappa^{\text{sgn}(c_1 + c_2 \underline{\eta}_i)}}\right)^2\right)^{-\frac{\nu+1}{2}} \quad i = 2, \dots, N \quad (16)$$

where sgn represents the signum function, c_1 and c_2 are correction terms that account for the shift in expectation and variance of the decorrelated and restandardised residuals caused by the skewing mechanism, respectively, and $\kappa > 0$

is a skewness parameter. For $\kappa = 1$, the skewed Student distribution is symmetric, while for $\kappa < 1$ it is left-skewed and for $\kappa > 1$ right-skewed. An illustration of this is given in Fig. 1b. The correction terms c_1 and c_2 are calculated as:

$$c_1(\nu, \kappa) = \frac{\left(\kappa^2 - \frac{1}{\kappa^2}\right) 2 \Gamma\left(\frac{\nu+1}{2}\right) \sqrt{\frac{\nu}{\nu-2}} (\nu-2)}{\left(\kappa + \frac{1}{\kappa}\right) \Gamma\left(\frac{\nu}{2}\right) \sqrt{\pi\nu} (\nu-1)} \quad (17)$$

and

$$c_2(\nu, \kappa) = \sqrt{-c_1^2 + \frac{\kappa^3 + \frac{1}{\kappa^3}}{\kappa + \frac{1}{\kappa}}}. \quad (18)$$

A detailed description of the method used for its derivation of can be found in Fernández and Steel (1998). An instructive example of an application of a skewed distribution for inference purposes can be found in Schoups and Vrugt (2010), who used the same method to introduce skewness in a symmetric exponential power distribution.

Finally, by assuming that the AR(1) scheme removes the autocorrelation in the residuals, the joint likelihood of observations $\tilde{y}_2, \dots, \tilde{y}_N$ can be calculated as:

$$p(\tilde{y}_2, \dots, \tilde{y}_N | \tilde{y}_1, \mathbf{x}_1, \mathbf{x}_2, \boldsymbol{\Omega}, \phi, \nu, \kappa) = \prod_{i=2}^N p(\tilde{y}_i | \mathbf{x}_1, \mathbf{x}_2, \boldsymbol{\Omega}, \phi, \nu, \kappa). \quad (19)$$

This joint likelihood is a conditional likelihood because it gives the joint likelihood of the $\tilde{y}_2, \dots, \tilde{y}_N$ observations given the first observation \tilde{y}_1 . The first observation is used to calculate the second decorrelated residual η_2 , but is not further considered in the calculation of the joint likelihood in Eq. (19). Note that this does not adversely affect the inference for $N \gg 1$.

Likelihood 3: The third likelihood model differs from the previous likelihood model in that it implements a modified AR(1) scheme that does not suffer from the deficiency of the classical scheme used in the previous likelihood model. The modified AR(1) scheme is given by:

$$\eta_i(\mathbf{x}, \mathbf{x}_2, \mathbf{\Omega}, \phi) = \underline{\varepsilon}_i - \phi \underline{\varepsilon}_{i-1} + \frac{\phi}{N} \sum_{j=1}^N \underline{\varepsilon}_j \quad i = 2, \dots, N \quad (20)$$

where the last term on the right hand side assures that the expectation of η_i will always be the same as the expectation of $\underline{\varepsilon}_j$. This can be seen from taking expectations of both sides of Eq. (20):

$$\begin{aligned} E(\eta_i) &= E(\underline{\varepsilon}_i) - E(\phi) E(\underline{\varepsilon}_{i-1}) + E(\phi) E(\underline{\varepsilon}_j) \\ &= E(\underline{\varepsilon}_j) \quad i = 2, \dots, N \end{aligned} \quad (21)$$

where again we used $E(\underline{\varepsilon}_i) = E(\underline{\varepsilon}_{i-1})$. All other assumptions are the same as for Likelihood 2, we solely replace Eq. (13) by the modified expression given in Eq. (20). We show in this study that the use of the modified AR(1) scheme removes the bias in the model prediction caused by the classic AR(1) scheme. Furthermore, statistical inference using Likelihood 3 resulted in reliable parameter estimates that fulfilled the underlying assumptions of the likelihood model.

As mentioned in Sect. 2.1, the time series of soil water content observations has a gap of about five days. The gap does not affect the calculations in the ordinary least-squares scheme of Likelihood 1. However, it calls for an extra step when using Likelihoods 2 and 3. This is because the autoregressive schemes applied in these likelihood models requires the same distance between neighbouring observations. For these two models, we therefore calculated two joint likelihoods, one for the first part of the time series of observations, and another one for the second part after the data gap. These two likelihoods were then multiplied to compute the likelihood to observe the entire data set.

2.4 Bayes' theorem

The joint likelihood summarizes in a probabilistic way what we can learn about the estimated parameters given the observations. However, sometimes it is desirable to make explicit use of prior knowledge about the parameters in the inference process. Bayes' theorem provides a formal way of combining information from observations with prior information about the estimated parameters. In case of Likelihood 1, Bayes' theorem can be written as:

$$p(\mathbf{x}_1, \mathbf{x}_2, \sigma_{\text{const}} | \tilde{\mathbf{y}}) \propto p(\mathbf{x}_1) p(\mathbf{x}_2) p(\sigma_{\text{const}}) p(\tilde{\mathbf{y}} | \mathbf{x}_1, \mathbf{x}_2, \sigma_{\text{const}}) \quad (22)$$

where the term $p(\mathbf{x}_1, \mathbf{x}_2, \sigma_{\text{const}} | \tilde{\mathbf{y}})$ on the left hand side is the posterior density of the estimated parameters given the observations. $p(\mathbf{x}_1)$, $p(\mathbf{x}_2)$, and $p(\sigma_{\text{const}})$ represent the prior densities of the parameters, and $p(\tilde{\mathbf{y}} | \mathbf{x}_1, \mathbf{x}_2, \sigma_{\text{const}})$ is just Likelihood 1, given in Eq. (10). Similar expressions for the posterior density are obtained for Likelihood 2 and 3.

For computational reasons, it is often not the prior, likelihood, and posterior that is used in practice, but rather its natural logarithm. This is because the prior, likelihood, and posterior may take on values larger or smaller than the largest or smallest number that can be represented by a double-precision floating-point format. This is especially true if the joint likelihood of a large number of observations is calculated, as it was the case in the present study. The log-prior, log-likelihood, and log-posterior was therefore used in all calculations, and it is these numbers that are reported here as model diagnostics.

2.5 Prior densities

In order to apply Bayes' theorem, the prior densities of the parameters need to be defined. As a prior of the VGM parameters we chose a truncated multivariate normal distribution $p(\mathbf{x}_1) \sim \mathcal{TN}(\boldsymbol{\mu}_{x_1}, \boldsymbol{\Sigma}_{x_1}, a_{x_1}, b_{x_1})$, where $\boldsymbol{\mu}_x$ denotes the mean vector of the VGM parameters, $\boldsymbol{\Sigma}_x$ is the corresponding covariance matrix, and a_{x_1} and b_{x_1} represent

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures

⏪

⏩

◀

▶

Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



ter Braak and Vrugt, 2008) and decides upon their acceptance by the Metropolis rule Metropolis et al. (1953). We did two minor changes to the standard DE-MC_(Z) scheme, which aimed at improving the performance of this MCMC scheme. Without these modifications, the exploration of the parameter space by the various Markov chains run in parallel was quite poor if Likelihood 2 or 3 was used. In fact, standard DE-MC_(Z) even failed to locate the region of the parameter space with highest posterior density in some of the test runs. The two modifications improved the performance in these cases substantially. The two modifications are as follows:

1. We implemented a block updating scheme, similar to what was used in ter Braak (2006). The idea behind block updating is that the convergence to a stationary distribution can sometimes be increased substantially if only a subset of the estimated parameters is updated in a single step. This general finding motivated Vrugt et al. (2009a) to implement a self-adaptive randomized updating scheme into DE-MC, resulting in the DiffeRential Evolution Adaptive Metropolis (DREAM) algorithm. This updating strategy did not work well in our application, presumably because it ignores the correlation structure of the VGM parameters induced by the informative prior of these parameters. Instead of the randomized subspace sampling used in DREAM, we therefore used a block updating scheme, as follows. For the first 2000 steps in each of the parallel chains, each parameter block was updated with probability 0.8. The VGM parameters were considered a block, whereas the remaining parameters were considered a block each. This simple block updating scheme increased the exploration capabilities of the Markov chains and the convergence to a stationary distribution substantially. The block updating was restricted to the initial phase of the MCMC simulation because it is not particularly efficient any more once the chains have approached to the stationary distribution.
2. Roberts and Rosenthal (2001) suggested that acceptance rates between 0.15 and 0.5 can generally be considered efficient, with an optimal value being 0.23.

HESSD

12, 2155–2199, 2015

Inverse modelling of in situ soil water dynamics

B. Scharnagl et al.

[Title Page](#)[Abstract](#)[Introduction](#)[Conclusions](#)[References](#)[Tables](#)[Figures](#)[⏪](#)[⏩](#)[◀](#)[▶](#)[Back](#)[Close](#)[Full Screen / Esc](#)[Printer-friendly Version](#)[Interactive Discussion](#)

Inverse modelling of in situ soil water dynamics

B. Scharnagl et al.

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures



Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



Within standard DE-MC, the jumping distance of a proposal is scaled such that optimal acceptance rates are obtained when sampling from Gaussian and Student posteriors (Roberts and Rosenthal, 2001; ter Braak, 2006). Here, we introduce a factor r (dimensionless) that reduces this reference scaling of the jumping distance. The reasoning for this reduction is that efficient sampling from posterior distributions that differ from the multivariate Gaussian or Student distribution generally require smaller jumps than those taken by the reference scaling used in the standard DE-MC scheme. In the present study, the reduction factor was set to $r = 1.0$ in the case of Likelihood 1, that is, no reduction, and $r = 0.5$ in the cases of Likelihood 2 and 3. With these settings, acceptance rates near the optimal value of 0.23 were obtained in all three cases.

Apart from these two changes to the DE-MC_(Z) scheme, we used standard settings for the algorithmic parameters reported by ter Braak (2006) and ran five chains in parallel. Convergence of the Markov chains to the stationary distribution was monitored using the interval diagnostic of Brooks and Gelman (1998). After convergence, we ran each of the five chains for another 10 000 steps. Parameter inference was based on 5000 samples selected from this final sampling stage by thinning the chains, that is, by retaining every 10th sample. The calculation of the uncertainty bounds of the model prediction is described in Schoups and Vrugt (2010), and is therefore not repeated herein.

3 Results

Figure 2 shows the results of inverse modelling of the time series of in situ soil water content observations using Likelihood 1 together with precipitation and potential evaporation fluxes that define the upper boundary condition in HYDRUS-1-D. The parameter estimates that belong to the prediction with maximum posterior density are listed in Table 1. This table also lists the corresponding mean error, root mean square error,

model efficiency (Nash and Sutcliffe, 1970), as well as the log-prior, log-likelihood, and log-posterior, which serve as model diagnostics to rate the goodness of the model prediction. Based on visual inspection of Fig. 2, the ordinary least-squares approach implemented in the first likelihood model results in a good overall description of the observational data. This impression is also corroborated by the various model diagnostics listed in Table 2.

To check the assumptions of homoscedastic, independent, and Gaussian distributed standardised residuals $\underline{\varepsilon}_i$ used in ordinary least-squares, we made use of three diagnostic plots. Figure 3a depicts the standardised residuals plotted against the simulated soil water content. This diagnostic plot indicates that the residuals are heteroscedastic. The SD of the residuals increases with increasing values of the predicted soil water content. Figure 3b shows the autocorrelation function of the residuals. This diagnostic plot shows that the residuals are strongly autocorrelated. The autocorrelation of the residuals at lag 1, representing a time interval of 3 h, is close to one. The autocorrelation decreases rather slowly with increasing lag in an approximately exponential manner, which points to an autoregressive process of first order (e.g. Box et al., 2008). At lag 60, corresponding to a temporal separation of 7.5 days, the autocorrelation is still significant, as indicated by the theoretical 95 % uncertainty bounds for a time series of uncorrelated random variables. Figure 3c depicts a quantile–quantile plot of the standardised residuals. This diagnostic plot indicates that the actual distribution of the standardised residuals differs from that of a Gaussian distribution. The actual distribution appears to be bimodal and also differs in its tailing behaviour from that of a Gaussian distribution. In summary, the three diagnostic plots show that neither of the three assumptions used in the formulation of Likelihood 1 is fulfilled. In this case, good statistical practice requires that the likelihood model is revised and that the inference process is repeated with the revised likelihood function.

The results for Likelihood 2, which allows for heteroscedastic, autocorrelated, and non-Gaussian distributed residuals, are presented in Fig. 4. The predicted soil water content deviates substantially and consistently from the observations. Clearly, the pa-

HESSD

12, 2155–2199, 2015

Inverse modelling of in situ soil water dynamics

B. Scharnagl et al.

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures



Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



log-likelihood as well as the log-posterior obtained with Likelihood 3 are substantially larger than the corresponding values obtained with Likelihood 1 (Table 2). The larger log-likelihood value already indicates that Likelihood 3 better describes the statistical features of the residuals.

Figure 6 illustrates the standardization of the residuals using the PCHIP approach and also shows the diagnostic plots used to check the underlying assumptions of the likelihood model. In Fig. 6a, the raw residuals ε_i are plotted as a function of the predicted soil water content. It becomes visible here that the variance of the raw residuals varies as function of the predicted soil water content, that is, that the ε_i are heteroscedastic. However, what becomes also clear from this figure is that the expectation of the raw residuals also varies as a function of the predicted soil water content. The soil hydrological model systematically overestimates larger observations, which is an obvious expression of bias of the process model. The results of the PCHIP approach to estimate the SD of the raw residuals as a function of the predicted soil water content is shown in Fig. 6b. The corresponding standardised residuals $\underline{\varepsilon}_i$ are depicted in Fig. 6c. The standardised residuals appear to be homoscedastic, even though the systematic deviation of the expectation of $\underline{\varepsilon}_i$ persists. We explain this by the fact that the PCHIP approach partially compensates for model bias by assigning larger SDs to the more biased residuals.

The three assumptions of homoscedasticity, independence, and skewed Student distribution of the decorrelated and restandardised residuals $\underline{\eta}_i$ that underlie Likelihood 3 are checked in the three diagnostic plots shown in Fig. 6d to f. These plots show that the three underlying assumptions are approximately fulfilled in the case of Likelihood 3. The decorrelated and restandardised residuals $\underline{\eta}_i$ used in this likelihood model appear to be approximately homoscedastic, independent, and skewed Student distributed. From a statistical point of view, the parameter estimates obtained with Likelihood 3 can therefore be considered statistically valid. It is also noteworthy that the estimated values of the kurtosis parameter ν and the skewness parameter κ indicate that the actual distribution of the residuals deviates substantially from that of a Gaussian dis-

HESSD

12, 2155–2199, 2015

Inverse modelling of in situ soil water dynamics

B. Scharnagl et al.

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures



Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



4 Discussion

The application of Likelihood 1, which makes use of the ordinary least-squares assumptions, resulted in parameter estimates that are invalid from a statistical point of view because the residuals violate all underlying assumptions of this likelihood model.

5 Neglecting autocorrelation in the residuals means to overestimate the information content of the data and therefore leads to an underestimation of parameter uncertainty. This becomes evident in the parameter sample from the posterior shown in Fig. 7: the variance of the posterior obtained with Likelihood 1 is generally smaller than that obtained with Likelihood 3. If residuals are correlated, the amount of information provided by each observation is smaller than in the case of uncorrelated observations. In fact, using a likelihood model which neglects autocorrelation essentially implies to treat the soil hydrological model as if it was perfect. This is overoptimistic because environmental models are – almost by definition – at best good approximations of real-world processes: they cannot be perfect.

15 We showed that the classical AR(1) scheme implemented in Likelihood 2 resulted in unrealistic parameter estimates and, correspondingly, to biased model predictions. This corroborates the results reported by Wöhling and Vrugt (2011), who used the classical AR(1) scheme in a soil hydrological application. We showed analytically that the AR(1) scheme introduces bias, which explains the failure of this autoregressive model in the present study as well as in the application of Wöhling and Vrugt (2011). Obviously, inverse modelling of in situ soil water dynamics seems to be much more susceptible to the bias problem induced by the classical AR(1) scheme than, for example, inverse problems in catchment hydrology, where it has been applied successfully many times (e.g. Sorooshian and Dracup, 1980; Bates and Campbell, 2001). We conjecture that this is because the systematic misfit between observations and model predictions depicted in Fig. 4 can easily be compensated for by an estimated value of the autocorrelation coefficient that is larger than actually necessary. In a recent study, Evin et al. (2014) showed that the results obtained with the classical AR(1) model applied

HESSD

12, 2155–2199, 2015

Inverse modelling of in situ soil water dynamics

B. Scharnagl et al.

[Title Page](#)

[Abstract](#)

[Introduction](#)

[Conclusions](#)

[References](#)

[Tables](#)

[Figures](#)

[⏪](#)

[⏩](#)

[◀](#)

[▶](#)

[Back](#)

[Close](#)

[Full Screen / Esc](#)

[Printer-friendly Version](#)

[Interactive Discussion](#)



to inverse modelling of stream flow can be nonrobust, indicated by strong correlations between parameters in the hydrological model and the likelihood model. We suspect that the reason for this nonrobustness might also be explained, at least partially, with the bias that is introduced by the classical AR(1) scheme.

5 The residuals obtained with Likelihood 3 clearly satisfied all assumptions of this likelihood model and furthermore demonstrate that Likelihood 3 solves the bias-problem inherent to Likelihood 2. Since the log-likelihood and log-posterior values for Likelihood 3 were substantially larger than those obtained with Likelihood 1 (Table 2), we consider this likelihood function and the parameter estimates obtained with it as clearly superior
10 and statistically valid. The model predictions obtained with Likelihood 3 slightly violate the zero-expectation assumption, which becomes apparent from the non-constant expectation of the residuals shown in Fig. 6a. This model bias can in principle be accounted for in the formulation of the likelihood model by introducing a term for bias correction, which can for example be estimated as a function of the predicted soil water content (e.g. Kennedy and O'Hagan, 2001; Erdal et al., 2012). From our point of
15 view, however, such an approach would not be meaningful in the present application. This is because the bias correction would be valid only for the depth within the soil profile where the observations were taken, but it remains unclear how this bias correction affects the predicted soil water contents at other depths within the profile and the fluxes to adjacent compartments within the water cycle. This is a fundamental difference between the Bayesian approach adopted in this study and other statistical approaches that are designed to account for model bias, such as particle filtering (e.g. Montzka et al., 2011; Vrugt et al., 2013). Within the the present statistical framework,
20 we think that the model bias can therefore only be meaningfully corrected for by a thorough revision of the process model, the assumption of a homogeneous profile, the parametrisation of the soil hydraulic properties, and the inclusion of hysteresis.

25 Likelihood 3, which allows for heteroscedastic, autocorrelated, and non-Gaussian residuals, may be equally useful in other soil hydrological applications. In general, autocorrelation of the residuals becomes increasingly apparent if observations are taken

HESSD

12, 2155–2199, 2015

Inverse modelling of in situ soil water dynamics

B. Scharnagl et al.

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures



Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



at a sufficiently high temporal frequency, which is typically the case for automated measurement setups either in the field or in the laboratory. For example, autocorrelation of the residuals is particularly pronounced in inverse modelling of multi-step outflow experiments (e.g. Diamantopoulos et al., 2012), evaporation experiments (e.g. Dettmann et al., 2014), or lysimeter data (e.g. Schelle et al., 2012). Based on the findings of the present study, we conjecture that a likelihood model that allows for heteroscedastic, autocorrelated, and non-Gaussian residuals would also prove useful in these applications. The main advantages are a more accurate quantification of the model parameters and their uncertainty. In the case of multiple time series used for parameter inference, Likelihood 3 must be extended to the multivariate case by using a first-order vector autoregressive (VAR(1)) model (e.g. Box et al., 2008). In contrast to the AR(1) model treated in this study, the VAR(1) model additionally accounts for the cross-correlation between the multiple, parallel time series of observations.

5 Summary and conclusions

In this case study, we estimated the soil hydraulic properties of a homogeneous, bare soil profile from in situ observations of soil water content taken at a single location near the soil surface. This type of observational data often contain too little information to warrant accurate and precise estimates of all of the parameters in the soil hydrological model. We posed the inverse problem in a formal Bayesian approach, which allowed us to use prior information of the soil hydraulic parameters in terms of an informative prior distribution for these parameters. The prior distribution was taken from Scharnagl et al. (2011). We then applied three likelihood models, which differ in their underlying assumptions about the salient statistical features of the time series of residuals.

The first likelihood model corresponds to the ordinary least-squares approach, which is the de facto standard in inverse modelling of soil hydrological processes. However, the results obtained with Likelihood 1 showed that the underlying assumptions of this likelihood model of homoscedastic, independent, and Gaussian residuals were all vio-

HESSD

12, 2155–2199, 2015

Inverse modelling of in situ soil water dynamics

B. Scharnagl et al.

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures



Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



lated. From a statistical point of view, the parameter estimates obtained with this first likelihood model can therefore be considered invalid.

The second likelihood model makes use of the classical AR(1) scheme to account for autocorrelated residuals, allows for heteroscedasticity, and supports heavily-tailed, non-Gaussian distributions of the residuals. To describe the heteroscedastic nature of the residuals, we devised a flexible model for the relationship between model predicted soil water content and the SD of the residuals that uses cubic Hermite interpolation. Deviations from a Gaussian distribution were modelled with a skewed Student distribution. A theoretical analysis of this parameter inference scheme and the results obtained with Likelihood 2 demonstrated that the classical AR(1) model introduces bias into the inference process. Due to this bias, the results obtained with Likelihood 2 were essentially meaningless, which corroborates the findings of Wöhling and Vrugt (2011) obtained in a similar soil hydrological study and using the same classical AR(1) scheme.

In the third likelihood model, we introduced a modified AR(1) scheme, which does not suffer from the bias problem induced by the classical scheme. With this likelihood model, we obtained parameter estimates that differed significantly from those obtained with Likelihood 1, which implements the ordinary least-squares assumptions. Furthermore, parameter uncertainty was higher as compared to Likelihood 1 because the decreased information content of the calibration data caused by autocorrelation was accounted for. Since the residuals obtained with Likelihood 3 fulfilled the underlying statistical assumptions, and because Likelihood 3 lead to a much larger log-posterior than Likelihood 1, we concluded that the parameter estimates obtained with Likelihood 3 are statistically valid, and therefore, superior to the ones obtained with Likelihood 1.

The results of this study demonstrate that the ordinary least-squares approach may lead to statistically invalid estimates of the soil hydraulic parameters, because the underlying assumptions about the statistical features of the time series of residuals are not fulfilled. The current practice in inverse modelling of soil hydrological processes should therefore be revised. We introduced a new, highly flexible likelihood model that accounts for autocorrelated, heteroscedastic, and non-Gaussian residuals, which per-

HESSD

12, 2155–2199, 2015

Inverse modelling of in situ soil water dynamics

B. Scharnagl et al.

[Title Page](#)

[Abstract](#)

[Introduction](#)

[Conclusions](#)

[References](#)

[Tables](#)

[Figures](#)



[Back](#)

[Close](#)

[Full Screen / Esc](#)

[Printer-friendly Version](#)

[Interactive Discussion](#)



formed very well in the present test study. Based on our findings, we advocate the use and further extension of the new likelihood model in inverse modelling of soil hydrological processes.

Acknowledgements. We thank Marius Schmidt and Karl Schneider for providing the meteorological data used to define the upper boundary conditions. We also acknowledge the help of Nils Prolingheuer during the measurement setup and data collection. The first, fourth, and fifth author gratefully acknowledge financial support by the TERENO project and by SFB/TR 32 “Patterns in Soil–Vegetation–Atmosphere Systems: Monitoring, Modelling, and Data Assimilation” funded by the Deutsche Forschungsgemeinschaft (DFG).

The service charges for this open-access publication have been covered by a Research Centre of the Helmholtz Association.

References

- Allen, R. G., Pereira, L. S., Raes, D. S., and Smith, M.: Crop evapotranspiration (guidelines for computing crop water requirements), no. 56, in: FAO Irrigation and Drainage Paper No. 56, Food and Agricultural Organization of the United Nations, Rome, Italy, 1998. 2162
- Baroni, G., Facchi, A., Gandolfi, C., Ortuani, B., Horeschi, D., and van Dam, J. C.: Uncertainty in the determination of soil hydraulic parameters and its influence on the performance of two hydrological models of different complexity, *Hydrol. Earth Syst. Sci.*, 14, 251–270, doi:10.5194/hess-14-251-2010, 2010. 2157
- Basile, A., Ciollaro, G., and Coppola, A.: Hysteresis in soil water characteristics as a key to interpreting comparisons of laboratory and field measured hydraulic properties, *Water Resour. Res.*, 39, 1355, doi:10.1029/2003WR002432, 2003. 2157
- Bates, B. C. and Campbell, E. P.: A Markov chain Monte Carlo scheme for parameter estimation and inference in conceptual rainfall-runoff modeling, *Water Resour. Res.*, 37, 937–947, doi:10.1029/2000WR900363, 2001. 2167, 2179
- Beach, C. M. and MacKinnon, J. G.: A maximum likelihood procedure for regression with autocorrelated errors, *Econometrica*, 46, 51–58, 1978. 2159

Inverse modelling of in situ soil water dynamics

B. Scharnagl et al.

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures



Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



Inverse modelling of in situ soil water dynamics

B. Scharnagl et al.

[Title Page](#)

[Abstract](#)

[Introduction](#)

[Conclusions](#)

[References](#)

[Tables](#)

[Figures](#)

[⏪](#)

[⏩](#)

[⏴](#)

[⏵](#)

[Back](#)

[Close](#)

[Full Screen / Esc](#)

[Printer-friendly Version](#)

[Interactive Discussion](#)



- Bitterlich, S., Durner, W., Iden, S. C., and Knabner, P.: Inverse estimation of the unsaturated soil hydraulic properties from column outflow experiments using free-form parameterizations, *Vadose Zone J.*, 3, 971–981, doi:10.2136/vzj2004.0971, 2004. 2166
- Box, G. E. P., Jenkins, G. M., and Reinsel, G. C.: *Time Series Analysis: Forecasting and Control*, 3rd Edn., John Wiley and Sons, Hoboken, NJ, USA, 2008. 2166, 2167, 2175, 2181
- Brooks, S. P.: Markov chain Monte Carlo method and its application, *J. Roy. Stat. Soc. D-Stat.*, 47, 69–100, doi:10.1111/1467-9884.00117, 1998. 2172
- Brooks, S. P. and Gelman, A.: General methods for monitoring convergence of iterative simulations, *J. Comput. Graph. Stat.*, 7, 434–455, doi:10.1080/10618600.1998.10474787, 1998. 2174
- Carsel, R. F. and Parrish, R. S.: Developing joint probability distributions of soil water retention characteristics, *Water Resour. Res.*, 24, 755–769, doi:10.1029/WR024i005p00755, 1988. 2159
- Cochrane, D. and Orcutt, G. H.: Application of least squares regression to relationships containing auto-correlated error terms, *J. Am. Stat. Assoc.*, 44, 32–61, doi:10.1080/01621459.1949.10483290, 1949. 2159
- Dane, J. H. and Wierenga, P. J.: Effect of hysteresis on the prediction of infiltration, redistribution and drainage of water in a layered soil, *J. Hydrol.*, 25, 229–242, doi:10.1016/0022-1694(75)90023-2, 1975. 2157
- Dettmann, U., Bechtold, M., Frahm, E., and Tiemeyer, B.: On the applicability of unimodal and bimodal van Genuchten–Mualem based models to peat and other organic soils under evaporation conditions, *J. Hydrol.*, 515, 103–115, doi:10.1016/j.jhydrol.2014.04.047, 2014. 2181
- Diamantopoulos, E. and Durner, W.: Dynamic nonequilibrium of water flow in porous media: a review, *Vadose Zone J.*, 11, doi:10.2136/vzj2011.0197, 2012. 2157
- Diamantopoulos, E., Iden, S. C., and Durner, W.: Inverse modeling of dynamic non-equilibrium in water flow with an effective approach, *Water Resour. Res.*, 48, W03503, doi:10.1029/2011WR010717, 2012. 2181
- Doherty, J. and Welter, D.: A short exploration of structural noise, *Water Resour. Res.*, 46, W05525, doi:10.1029/2009WR008377, 2010. 2159
- Durner, W. and Iden, S. C.: Extended multistep outflow method for the accurate determination of soil hydraulic properties near water saturation, *Water Resour. Res.*, 47, W08526, doi:10.1029/2011WR010632, 2011. 2157

Inverse modelling of in situ soil water dynamics

B. Scharnagl et al.

[Title Page](#)

[Abstract](#)

[Introduction](#)

[Conclusions](#)

[References](#)

[Tables](#)

[Figures](#)

[⏪](#)

[⏩](#)

[◀](#)

[▶](#)

[Back](#)

[Close](#)

[Full Screen / Esc](#)

[Printer-friendly Version](#)

[Interactive Discussion](#)



- Erdal, D., Neuweiler, I., and Huisman, J. A.: Estimating effective model parameters for heterogeneous unsaturated flow using error models for bias correction, *Water Resour. Res.*, 48, W06530, doi:10.1029/2011WR011062, 2012. 2180
- Evin, G., Kavetski, D., Thyer, M., and Kuczera, G.: Pitfalls and improvements in the joint inference of heteroscedasticity and autocorrelation in hydrological model calibration, *Water Resour. Res.*, 49, 4518–4524, doi:10.1002/wrcr.20284, 2013. 2166
- Evin, G., Thyer, M., Kavetski, D., McInerney, D., and Kuczera, G.: Comparison of joint versus postprocessor approaches for hydrological uncertainty estimation accounting for error autocorrelation and heteroscedasticity, *Water Resour. Res.*, 50, 2350–2375, doi:10.1002/2013WR014185, 2014. 2179
- Fernández, C. and Steel, M. F. J.: On Bayesian modeling of fat tails and skewness, *J. Am. Stat. Assoc.*, 93, 359–371, doi:10.1080/01621459.1998.10474117, 1998. 2168, 2169
- Fritsch, F. N. and Carlson, R. E.: Monotone piecewise cubic interpolation, *SIAM J. Numer. Anal.*, 17, 238–246, doi:10.1137/0717021, 1980. 2165
- Heimovaara, T. J. and Bouten, W.: A computer-controlled 36-channel time domain reflectometry system for monitoring soil water contents, *Water Resour. Res.*, 26, 2311–2316, doi:10.1029/WR026i010p02311, 1990. 2161
- Hou, Z. and Rubin, Y.: On minimum relative entropy concepts and prior compatibility issues in vadose zone inverse and forward modeling, *Water Resour. Res.*, 41, W12425, doi:10.1029/2005WR004082, 2005. 2158
- Iden, S. C. and Durner, W.: Free-form estimation of the unsaturated soil hydraulic properties by inverse modeling using global optimization, *Water Resour. Res.*, 43, W07451, doi:10.1029/2006WR005845, 2007. 2166
- IUSS Working Group WRB: World Reference Base for Soil Resources 2006, World Soil Resources Reports No. 103, first updated Edn., Food and Agricultural Organization of the United Nations, Rome, Italy, 2007. 2161
- Jacques, D., Šimůnek, J., Timmerman, A., and Feyen, J.: Calibration of Richards' and convection-dispersion equations to field-scale water flow and solute transport under rainfall conditions, *J. Hydrol.*, 259, 15–31, doi:10.1016/S0022-1694(01)00591-1, 2002. 2165
- Kavetski, D., Kuczera, G., and Franks, S. W.: Bayesian analysis of input uncertainty in hydrological modeling: 1. Theory, *Water Resour. Res.*, 42, W03407, doi:10.1029/2005WR004368, 2006. 2164

HESSD

12, 2155–2199, 2015

**Inverse modelling of
in situ soil water
dynamics**

B. Scharnagl et al.

[Title Page](#)[Abstract](#)[Introduction](#)[Conclusions](#)[References](#)[Tables](#)[Figures](#)[⏪](#)[⏩](#)[◀](#)[▶](#)[Back](#)[Close](#)[Full Screen / Esc](#)[Printer-friendly Version](#)[Interactive Discussion](#)

- Kennedy, M. C. and O'Hagan, A.: Bayesian calibration of computer models, *J. Roy. Stat. Soc. B*, 63, 425–464, doi:10.1111/1467-9868.00294, 2001. 2180
- Kuczera, G.: Improved parameter inference in catchment models: 1. Evaluating parameter uncertainty, *Water Resour. Res.*, 19, 1151–1162, doi:10.1029/WR019i005p01151, 1983. 2159, 2167
- Lange, K. L., Little, R. J. A., and Taylor, J. M. G.: Robust statistical modeling using the *t* distribution, *J. Am. Stat. Assoc.*, 84, 881–896, doi:10.1080/01621459.1989.10478852, 1989. 2168, 2172
- Lenhard, R. J., Parker, J. C., and Kaluarachchi, J. J.: Comparing simulated and experimental hysteretic two-phase transient fluid flow phenomena, *Water Resour. Res.*, 27, 2113–2124, doi:10.1029/91WR01272, 1991. 2157
- Mallants, D., Mohanty, B. P., Vervoort, A., and Feyen, J.: Spatial analysis of saturated hydraulic conductivity in a soil with macropores, *Soil Technol.*, 10, 115–131, doi:10.1016/S0933-3630(96)00093-1, 1997. 2157
- Mertens, J., Madsen, H., Kristensen, M., Jacques, D., and Feyen, J.: Sensitivity of soil parameters in unsaturated zone modelling and the relation between effective, laboratory and in situ estimates, *Hydrol. Process.*, 19, 1611–1633, doi:10.1002/hyp.5591, 2005. 2157
- Metropolis, N., Rosenbluth, A. W., Rosenbluth, M. N., Teller, A. H., and Teller, E.: Equation of state calculations by fast computing machines, *J. Chem. Phys.*, 21, 1087–1091, doi:10.1063/1.1699114, 1953. 2173
- Montzka, C., Moradkhani, H., Weihermüller, L., Hendricks Franssen, H.-J., Canty, M., and Vereecken, H.: Hydraulic parameter estimation by remotely-sensed top soil moisture observations with the particle filter, *J. Hydrol.*, 399, 410–421, doi:10.1016/j.jhydrol.2011.01.020, 2011. 2180
- Nash, J. E. and Sutcliffe, J. V.: River flow forecasting through conceptual models part I A discussion of principles, *J. Hydrol.*, 10, 282–290, doi:10.1016/0022-1694(70)90255-6, 1970. 2175
- Peters, A. and Durner, W.: Simplified evaporation method for determining soil hydraulic properties, *J. Hydrol.*, 356, 147–162, doi:10.1016/j.jhydrol.2008.04.016, 2008. 2157
- Roberts, G. O. and Rosenthal, J. S.: Optimal scaling for various Metropolis-Hastings algorithms, *Stat. Sci.*, 16, 351–367, doi:10.1214/ss/1015346320, 2001. 2173, 2174
- Schaap, M. G., Leij, F. J., and van Genuchten, M. T.: ROSETTA: a computer program for estimating soil hydraulic parameters with hierarchical pedotransfer functions, *J. Hydrol.*, 251, 163–176, doi:10.1016/S0022-1694(01)00466-8, 2001. 2158, 2172

Inverse modelling of in situ soil water dynamics

B. Scharnagl et al.

[Title Page](#)

[Abstract](#)

[Introduction](#)

[Conclusions](#)

[References](#)

[Tables](#)

[Figures](#)



[Back](#)

[Close](#)

[Full Screen / Esc](#)

[Printer-friendly Version](#)

[Interactive Discussion](#)



- Scharnagl, B., Vrugt, J. A., Vereecken, H., and Herbst, M.: Inverse modelling of in situ soil water dynamics: investigating the effect of different prior distributions of the soil hydraulic parameters, *Hydrol. Earth Syst. Sci.*, 15, 3043–3059, doi:10.5194/hess-15-3043-2011, 2011. 2158, 2159, 2160, 2161, 2163, 2164, 2172, 2181
- 5 Schelle, H., Iden, S. C., Fank, J., and Durner, W.: Inverse estimation of soil hydraulic and root distribution parameters from lysimeter data, *Vadose Zone J.*, 11, doi:10.2136/vzj2011.0169, 2012. 2165, 2181
- Scholer, M., Irving, J., Binley, A., and Holliger, K.: Estimating vadose zone hydraulic properties using ground penetrating radar: the impact of prior information, *Water Resour. Res.*, 47, W10512, doi:10.1029/2011WR010409, 2011. 2158, 2159
- 10 Schoups, G. and Vrugt, J. A.: A formal likelihood function for parameter and predictive inference of hydrologic models with correlated, heteroscedastic, and non-Gaussian errors, *Water Resour. Res.*, 46, W10531, doi:10.1029/2009WR008933, 2010. 2169, 2174
- Sonnleitner, M. A., Abbaspour, K. C., and Schulin, R.: Hydraulic and transport properties of the plant-soil system estimated by inverse modelling, *Eur. J. Soil Sci.*, 54, 127–138, doi:10.1046/j.1365-2389.2002.00491.x, 2003. 2165
- 15 Sorooshian, S. and Dracup, J. A.: Stochastic parameter estimation procedures for hydraulic rainfall-runoff models: correlated and heteroscedastic error cases, *Water Resour. Res.*, 16, 430–442, doi:10.1029/WR016i002p00430, 1980. 2159, 2167, 2179
- 20 ter Braak, C. J. F.: A Markov chain Monte Carlo version of the genetic algorithm Differential Evolution: easy Bayesian computing for real parameter spaces, *Stat. Comput.*, 16, 239–249, doi:10.1007/s11222-006-8769-1, 2006. 2172, 2173, 2174
- ter Braak, C. J. F. and Vrugt, J. A.: Differential evolution Markov chain with snooker updater and fewer chains, *Stat. Comput.*, 18, 435–446, doi:10.1007/s11222-008-9104-9, 2008. 2172, 2173
- 25 Topp, G. C., Davis, J. L., and Annan, A. P.: Electromagnetic determination of soil water content: Measurements in coaxial transmission lines, *Water Resour. Res.*, 16, 574–582, doi:10.1029/WR016i003p00574, 1980. 2161
- van Dam, J. C., Stricker, J. N. M., and Droogers, P.: Inverse method to determine soil hydraulic functions from multistep outflow experiments, *Soil Sci. Soc. Am. J.*, 58, 647–652, doi:10.2136/sssaj1994.03615995005800030002x, 1994. 2157
- 30

Inverse modelling of in situ soil water dynamics

B. Scharnagl et al.

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures



Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



- van Genuchten, M. T.: A closed-form equation for predicting the hydraulic conductivity of unsaturated soils, *Soil Sci. Soc. Am. J.*, 44, 892–898, doi:10.2136/sssaj1980.03615995004400050002x, 1980. 2162
- Vereecken, H., Huisman, J. A., Bogaen, H., Vanderborght, J., Vrugt, J. A., and Hopmans, J. W.: On the value of soil moisture measurements in vadose zone hydrology: a review, *Water Resour. Res.*, 44, W00D06, doi:10.1029/2008WR006829, 2008. 2158
- Vrugt, J. A., Braak, C. J. F., Diks, C. G. H., Robinson, B. A., Hyman, J. M., and Higdon, D.: Accelerating Markov chain Monte Carlo simulation by differential evolution with self-adaptive randomized subspace sampling, *Int. J. Nonlin. Sci. Num.*, 10, 271–288, 2009a. 2173
- Vrugt, J. A., Braak, C. J. F., Gupta, H. V., and Robinson, B. A.: Equifinality of formal (DREAM) and informal (GLUE) Bayesian approaches in hydrologic modeling?, *Stoch. Env. Res. Risk A.*, 23, 1011–1026, doi:10.1007/s00477-008-0274-y, 2009b. 2167
- Vrugt, J. A., ter Braak, C. J. F., Diks, C. G. H., and Schoups, G.: Hydrologic data assimilation using particle Markov chain Monte Carlo simulation: theory, concepts and applications, *Adv. Water Resour.*, 51, 457–478, doi:10.1016/j.advwatres.2012.04.002, 2013. 2180
- Šimůnek, J., van Genuchten, M. T., and Wendroth, O.: Parameter estimation analysis of the evaporation method for determining soil hydraulic properties, *Soil Sci. Soc. Am. J.*, 62, 894–905, doi:10.2136/sssaj1998.03615995006200040007x, 1998. 2157
- Šimůnek, J., Šejna, M., Saito, H., Sakai, M., and van Genuchten, M. T.: The HYDRUS-1-D Software Package for Simulating the One-Dimensional Movement of Water, Heat and Multiple Solutes in Variably-Saturated Media (Version 4.0), Department of Environmental Sciences, University of California Riverside, Riverside, CA, USA, available at: http://www.pc-progress.com/Downloads/Pgm_Hydrus1D/HYDRUS1D_4.pdf (last access: 26 January 2015), 2008. 2162
- West, M.: Outlier models and prior distributions in Bayesian linear regression, *J. Roy. Stat. Soc. B*, 1, 431–439, 1984. 2168
- Wöhling, T. and Vrugt, J. A.: Multiresponse multilayer vadose zone model calibration using Markov chain Monte Carlo simulation and field water retention data, *Water Resour. Res.*, 47, W04510, doi:10.1029/2010WR009265, 2011. 2160, 2167, 2176, 2179, 2182
- Wollschläger, U., Pfaff, T., and Roth, K.: Field-scale apparent hydraulic parameterisation obtained from TDR time series and inverse modelling, *Hydrol. Earth Syst. Sci.*, 13, 1953–1966, doi:10.5194/hess-13-1953-2009, 2009. 2165

HESSD

12, 2155–2199, 2015

Inverse modelling of in situ soil water dynamics

B. Scharnagl et al.

[Title Page](#)

[Abstract](#)

[Introduction](#)

[Conclusions](#)

[References](#)

[Tables](#)

[Figures](#)



[Back](#)

[Close](#)

[Full Screen / Esc](#)

[Printer-friendly Version](#)

[Interactive Discussion](#)



Inverse modelling of
in situ soil water
dynamics

B. Scharnagl et al.

Table 1. Estimated parameters in the soil hydrological and likelihood model, respectively, as well as parameter estimates with maximum posterior density obtained using the three different likelihood models.

	parameter	unit	transformation	parameter estimates with maximum posterior density				
				lower bound	upper bound	likelihood model 1	likelihood model 2	likelihood model 3
soil hydrological model	θ_r	$\text{cm}^3 \text{cm}^{-3}$	–	0.043	0.091	0.051	0.061	0.053
	θ_s	$\text{cm}^3 \text{cm}^{-3}$	–	0.409	0.481	0.461	0.460	0.436
	α	cm^{-1}	$\log_{10}(\alpha)$	0.0028	0.0085	0.0045	0.0051	0.0037
	n	–	$\log_{10}(n)$	1.51	1.85	1.77	1.70	1.77
	K_s	cm h^{-1}	$\log_{10}(K_s)$	0.006	0.832	0.534	0.254	0.091
	L	–	–	–5.49	6.27	4.20	1.76	3.67
	h_{LB}	cm	–	–250	–50	–170	–51	–178
likelihood model	σ_{const}	$\text{cm}^3 \text{cm}^{-3}$	–	0.000	0.025	0.0121		
	σ_a	$\text{cm}^3 \text{cm}^{-3}$	–	0.000	0.050		0.0183	0.0092
	σ_b	$\text{cm}^3 \text{cm}^{-3}$	–	0.000	0.050		0.0190	0.0131
	σ_c	$\text{cm}^3 \text{cm}^{-3}$	–	0.000	0.050		0.0316	0.0218
	σ_d	$\text{cm}^3 \text{cm}^{-3}$	–	0.000	0.050		0.0498	0.0360
	ϕ	–	–	0.000	1.000		0.996	0.957
	ν	–	$1/\nu$	2.000	∞		3.32	2.89
	κ	–	–	0.500	2.000		1.27	0.85

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures

⏪

⏩

◀

▶

Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



Inverse modelling of
in situ soil water
dynamics

B. Scharnagl et al.

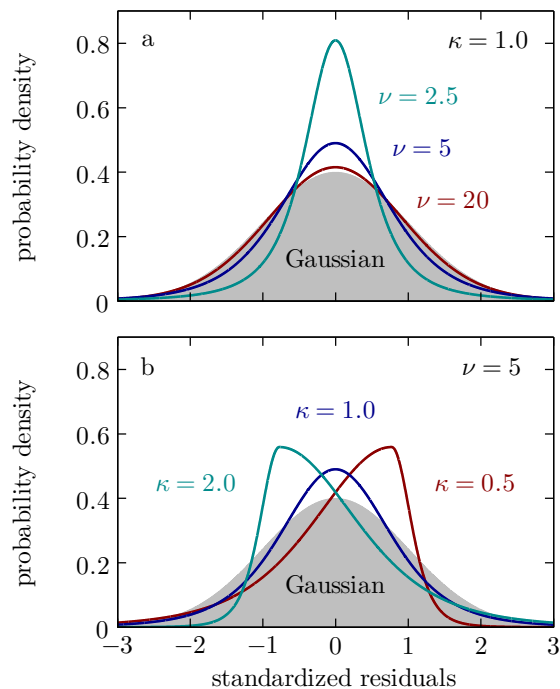


Figure 1. Illustration of the skewed Student distribution. In **(a)**, the skewness parameter is kept constant ($\kappa = 1$) while the kurtosis parameter is varied. In **(b)**, the skewness parameter is varied while the kurtosis parameter is kept constant ($\nu = 5$). To allow for a direct comparison, the corresponding Gaussian distribution is also shown. All distributions have an expectation of zero and a SD of one.

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures

◀

▶

◀

▶

Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



Inverse modelling of in situ soil water dynamics

B. Scharnagl et al.

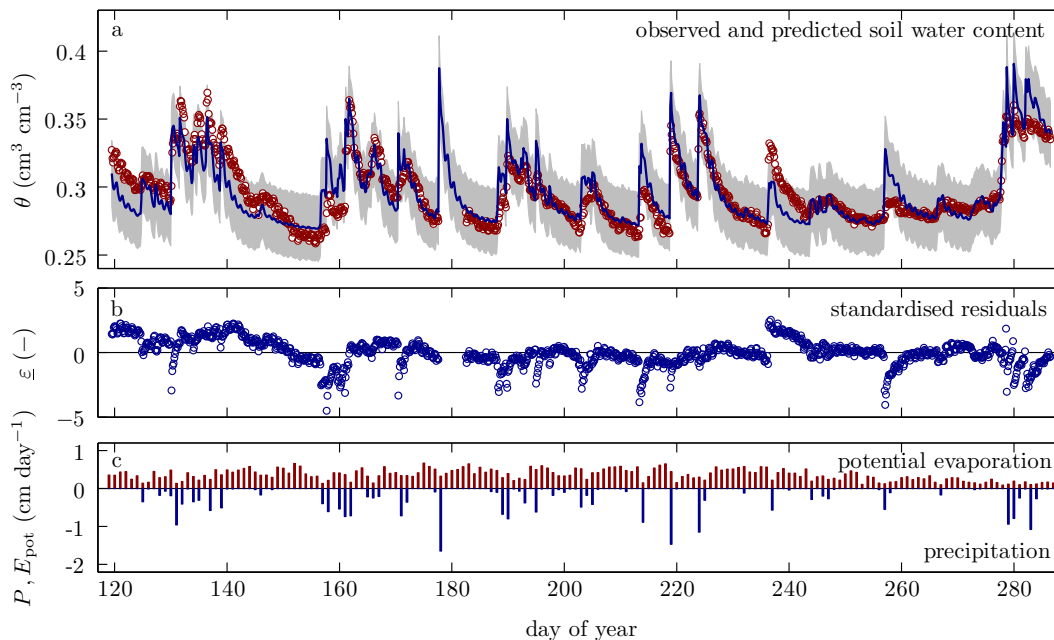


Figure 2. Model prediction of the time series of soil water content observations obtained using Likelihood 1. In **(a)**, the model prediction with maximum posterior density and the 95% uncertainty bounds of the prediction are shown together with the observations of soil water content θ . **(b)** displays the standardised residuals ε_i , corresponding to the prediction with maximum posterior density. **(c)** shows the precipitation P (negative) and potential evaporation E_{pot} (positive) fluxes used as input to the HYDRUS-1-D model. For graphical reasons, daily fluxes are shown here whereas hourly values were used to run the model.

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures



Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



Inverse modelling of
in situ soil water
dynamics

B. Scharnagl et al.

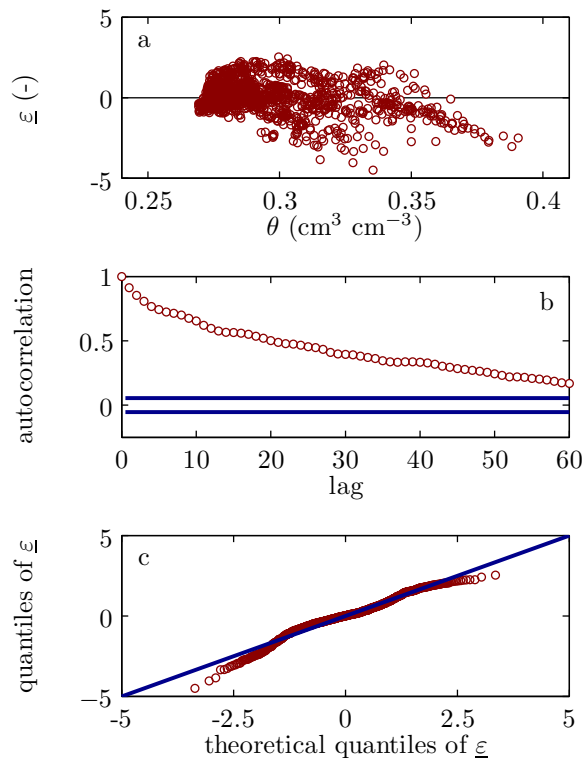


Figure 3. Diagnostic plots of the standardised residuals ε_i of the model prediction with maximum posterior density using Likelihood 1. In **(a)**, the standardised residuals ε_i are plotted over the predicted water content θ . **(b)** depicts the autocorrelation function with theoretical 95% uncertainty intervals of a time series of uncorrelated random variables, and **(c)** shows a quantile–quantile plot of the standardised residuals.

Inverse modelling of
in situ soil water content
dynamics

B. Scharnagl et al.

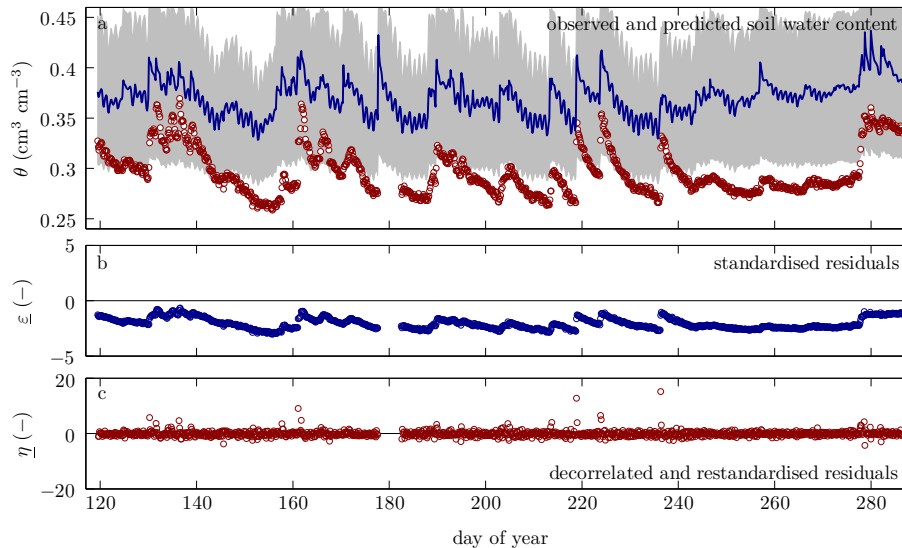


Figure 4. Model prediction of the time series of soil water content observations obtained using Likelihood 2. In **(a)**, the model prediction with maximum posterior density and the 95% uncertainty bounds of the prediction are shown together with the observations of soil water content θ . **(b)** displays the standardised residuals ε_i , corresponding to the prediction with maximum posterior density. **(c)** shows the decorrelated and restandardised residuals η_i .

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures

◀

▶

◀

▶

Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



Inverse modelling of
in situ soil water content
dynamics

B. Scharnagl et al.

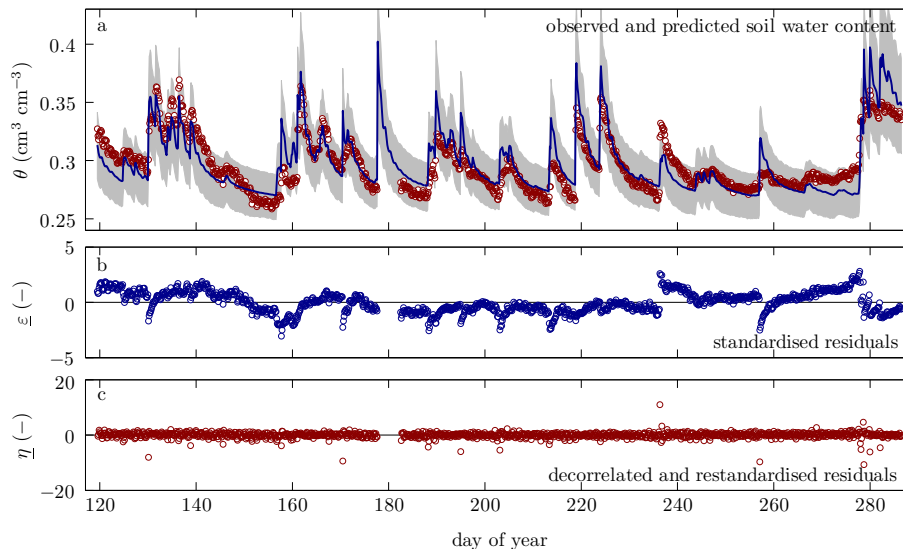


Figure 5. Model prediction of the time series of soil water content observations obtained using Likelihood 3. In **(a)**, the model prediction with maximum posterior density and the 95 % uncertainty bounds of the prediction are shown together with the observations of soil water content θ . **(b)** displays the standardised residuals ε_i , corresponding to the prediction with maximum posterior density. **(c)** shows the decorrelated and restandardised residuals η_i .

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures



Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



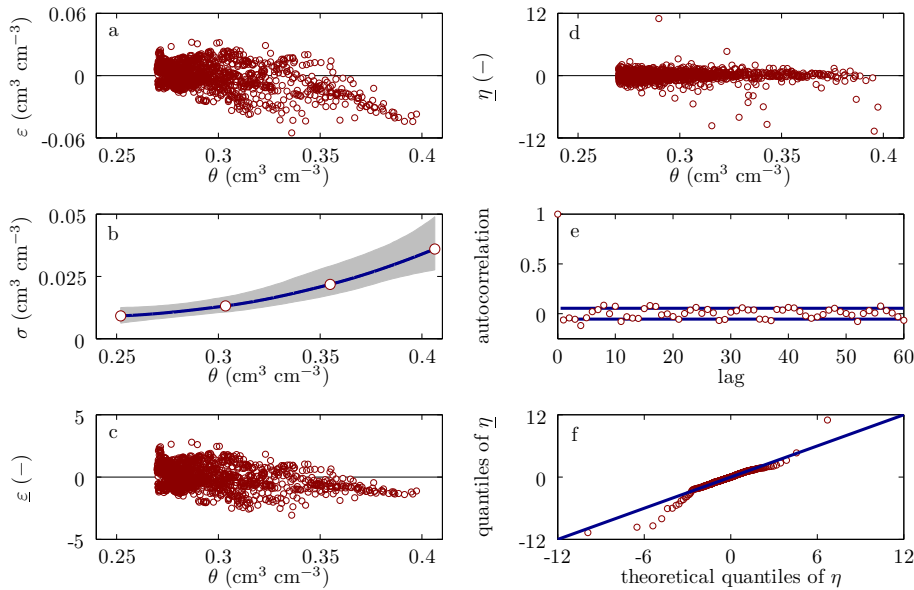


Figure 6. Performance of the PCHIP approach used for standardization of the residuals and diagnostic plots of the decorrelated and restandardised residuals corresponding to the model prediction with maximum posterior density using Likelihood 3. **(a)** shows the residuals ε_i as a function of the predicted soil water content θ . **(b)** gives the estimated SD as a function of the predicted soil water content, showing the estimated points $\sigma_{a,\dots,d}$ and the resulting function given by the PCHIP. The grey area represents the corresponding 95% uncertainty. In **(c)** the standardised residuals ε_i are plotted over the predicted water content, while **(d)** shows the decorrelated and restandardised residuals η_i . **(e)** depicts the autocorrelation function of η_i with theoretical 95% uncertainty intervals of a time series of uncorrelated random variables, and **(f)** shows a quantile–quantile plot of η_i .

Inverse modelling of in situ soil water dynamics

B. Scharnagl et al.

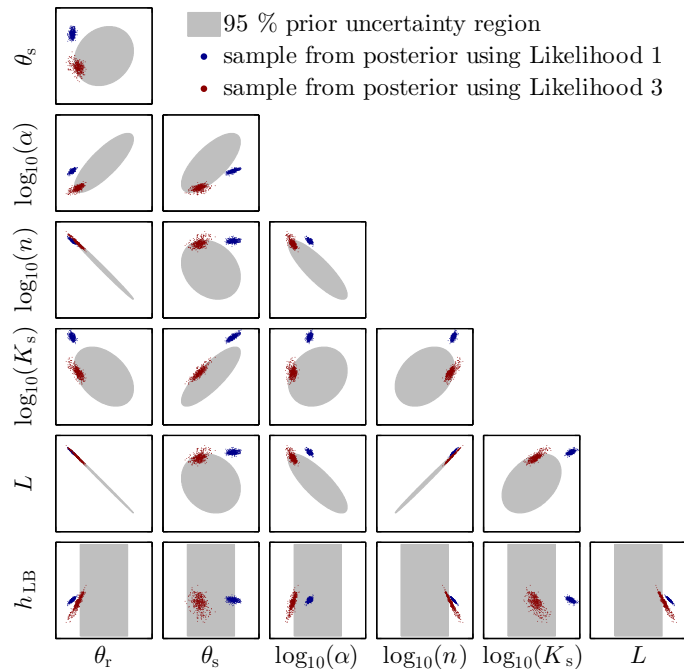


Figure 7. Pairwise scatter plots of the posterior samples of the parameters of the soil hydrological model obtained with Likelihood 1 and 3 together with the corresponding 95 % prior uncertainty regions of these parameters. The axes cover the parameter bounds listed in Table 1.

Title Page	
Abstract	Introduction
Conclusions	References
Tables	Figures
◀	▶
◀	▶
Back	Close
Full Screen / Esc	
Printer-friendly Version	
Interactive Discussion	



Inverse modelling of
in situ soil water
dynamics

B. Scharnagl et al.

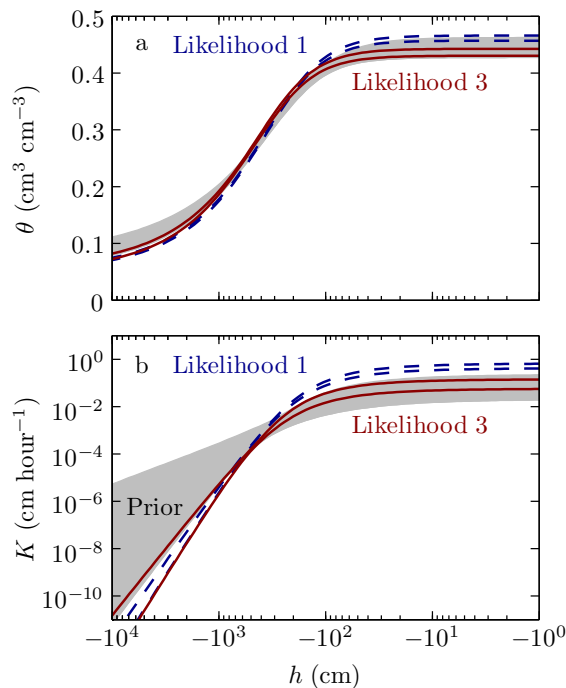


Figure 8. 95 % posterior uncertainty bounds of **(a)** the water retention function and **(b)** the hydraulic conductivity function obtained with Likelihood 1 and 3. Additionally, the corresponding 95 % prior uncertainty regions are shown as a grey area.