

Simulation of the Water Table Elevation in Shallow Unconfined Aquifers by means of the ERA5 Soil Moisture Dataset: The Umbria Region Case Study

PAOLINA BONGIOANNINI CERLINI,^a LORENZO SILVESTRI,^a SILVIA MENICONI,^b AND BRUNO BRUNONE^b

^a CIRIAF/Centre for Climate and Climate Change, University of Perugia, Perugia, Italy

^b Department of Civil and Environmental Engineering/Centre for Climate and Climate Change, University of Perugia, Perugia, Italy

(Manuscript received 24 June 2020, in final form 9 February 2021)

ABSTRACT: This paper concerns the simulation of the water table elevation in shallow unconfined aquifers where infiltration is assumed as the main mechanism of recharge. The main aim is to provide a reliable tool for groundwater management that satisfies water supply managers. Such a tool is a candidate as a physically based alternative to the use of empirical methods or general circulation models. It is based on the use of two widely available sets of data: the water table elevation measurements and soil moisture time series. In fact, the former are usually provided by government agencies on public websites whereas the latter are included in the atmospheric global datasets (reanalysis). It is notable that data from reanalysis are accessible to any citizen and organization around the world on an open-access basis (e.g., Copernicus). In the proposed method, the measured water table elevations are correlated quantitatively with the water fluxes toward the aquifer evaluated using the soil moisture data from ERA5 reanalysis (provided by ECMWF) within a Richards equation–based approach. The analysis is executed using data from the Umbria region (Italy) on both a daily and monthly scale. In fact, these are the time intervals of interest for a proper management of groundwater resources. The proposed relationships include both a logarithmic and linear term and point out the possible different regimes of the shallow aquifers with regard to the recharge due to infiltration. These different mechanisms reflect in the different role played by the water fluxes toward the aquifer in terms of water table elevation changes according to the considered time scale.

KEYWORDS: Reanalysis data; Atmosphere-land interaction; Planning; Soil moisture; Moisture/moisture budget

1. Introduction

Since groundwater is the largest source of water supply for agricultural, industrial, and domestic use, the more and more frequent falling of groundwater tables and polluted aquifers—as clear warnings of a crisis of the water body—seriously worry water authorities. Traditionally, such an issue was addressed to by regulating the pattern of human use from both the quantitative and qualitative point of view. Nowadays, to prevent irreversible serious consequences, the effects of climate change must also be taken into account (e.g., Bardsley et al. 2013; Karamouz et al. 2013; Sekhar et al. 2013). It is worth noting that, with respect to surface water resources, the groundwater behavior and its connections with the climate are more complex and difficult to model.

Without ignoring the need to control the evolving patterns of human use, the main route for developing management strategies for a sustainable use of groundwater resources is refining a physically based numerical model. In fact, such a tool makes it possible to simulate the behavior of the aquifer of interest for given scenarios, that is, for a given degree of exploitation of the groundwater resource and recharge. On the basis of the results of the water balance, water company managers can figure out the sustainable volume of water that can be withdrawn to match the

users' demand according to the water table behavior. In fact, such a feature is a clear indicator of the condition of the aquifer on the whole and dictates the management rules. Moreover, it plays an important role also from the economical point of view since the energy cost of the pumping from the aquifer depends strongly on the water table elevation. That said, a proper evaluation of the aquifer recharge, as the natural supply of the groundwater reservoir, is extremely important.

Within the numerical modeling of the aquifer behavior, recharge is a given boundary condition. According to de Vries and Simmers (2002), recharge is categorized as “diffuse” (or “direct”) and nondiffuse (or “localized” or “focused”). In the first case, it originates mainly from precipitation that infiltrates vertically from the surface directly to the water table whereas in the second case it collects in streams or topographic depressions before it infiltrates. In other words, the mechanism of the direct recharge is the infiltration through the vadose zone whereas within the indirect recharge the percolation to the water table happens through the beds of surface water courses (Lerner and Simmers 1990). Moreover, recharge may be affected by macropore flow through root channels and desiccation cracks as well as in the vadose zone preferential flow may occur due to unstable wetting fronts and differentiated soil physical characteristics (Lerner 1997).

According to the literature, beyond soil properties, main factors affecting recharge are meteorology, vegetation, and topography (e.g., Huet et al. 2016; Yin et al. 2011; Touhami et al. 2014). Unavoidable uncertainties about such control mechanisms—as an example, those connected to meteorological variability and land-use change (Jinno et al. 2009)—make recharge prediction quite difficult. To solve this problem, there

Corresponding author: Bruno Brunone, bruno.brunone@unipg.it

Earth Interactions is published jointly by the American Meteorological Society, the American Geophysical Union, and the Association of American Geographers.

DOI: 10.1175/EI-D-20-0011.1

© 2021 American Meteorological Society. For information regarding reuse of this content and general copyright information, consult the [AMS Copyright Policy](#) (www.ametsoc.org/PUBSReuseLicenses).

are two options for the approach to follow. The first approach is based on the use of simplified models whereas the second one prescribes the one of the general circulation models (GCMs).

Within the first approach, many contributions—below some examples are reported—are available addressing the problem from different points of view. In [Gogolev \(2002\)](#), the water-balance and Richards equation–based methods are compared to assess groundwater recharge for a hypothetical homogeneous profile and three real profiles for a deep aquifer. [Bonta and Müller \(1999\)](#) proposed a model providing the long-term groundwater recharge, based on the Glugla method and tested by considering historic lysimeter records, by using the average annual precipitation, runoff, potential evaporation, and crop-yield information. [Omorinbola \(1986\)](#) evaluated the groundwater accretion by means of empirical equations as correlated to the magnitude of the saturated zone thickness, since such a parameter fluctuates with the rate of groundwater recharge. Soil moisture balance models as well as a regional runoff/storm duration relationship for assessing the effective precipitation have been used in [Leach \(1982\)](#). In [Chinnasamy et al. \(2013\)](#), the prediction of the groundwater resource availability in India is based on the use of satellite-derived remote sensing data and the obtained results are compared with data from wells. In [Bjerklie et al. \(2011\)](#), future trends of groundwater recharge in Long Island Sound have been evaluated from GCM forecasts by assuming different scenarios in terms of carbon emissions. Within the water table fluctuation (WTF) method, valid for unconfined aquifers, the recharge is assumed as proportional to the measured rise of the water table, with the specific yield being the coefficient of proportionality (e.g., [Varni et al. 2013](#); [Healy and Cook 2002](#)). The premise of the WTF models is a very simplified water budget, in which the rise of the water table elevation is due only to the recharge.

As mentioned, the alternative approach is to use a GCM. In this case, the main problem for the water supply managers is the inhouse expertise. In fact, usually water companies have not a suitable staff who may deal with GCMs. A possible, less radical option could be to integrate the results of a GCM, provided by a research center or a government agency, into the case of interest (i.e., the aquifer to manage). However, even the statistical downscaling techniques, which are needed to obtain reliable results, could be out of reach for a water supply company.

In the above scenario, in principle global atmospheric datasets, which are the results of the combination of models with observations, could be of interest. In fact, they are based on the climate reanalysis of archived observations, concerning the recent history of the atmosphere, land surface, and oceans. Reanalysis, providing estimates of the atmospheric parameters (e.g., air temperature, pressure and wind at different altitudes, rainfall, and surface parameters), include millions of observations into a stable data assimilation system. Depending on the type of the reanalysis and version (see below), such estimates, extending back several decades, are available for all locations on earth with given spatial and temporal resolution. Therefore, if their viability is proven, global atmospheric datasets from reanalysis could be a good compromise between the use of simplified models and the very complex GCMs.

Currently, several research centers and agencies provide reanalyses, such as the National Centers for Environmental Prediction–National Center for Atmospheric Research (NCEP–NCAR), the Japanese Meteorological Agency (JMA), the European Centre for Medium-Range Weather Forecasts (ECMWF), and the National Aeronautics and Space Administration (NASA).

The first reanalysis product released by ECMWF is ERA-15 covering approximately 15 years, from December 1978 to February 1994. The second reanalysis archive, ERA-40, refers to about 40 years (from 1957 to 2002). As a precursor to a revised extended reanalysis product to replace ERA-40, ECMWF released ERA-Interim ([Dee et al. 2011](#)), which covers the period from 1979 to the present. The most recent reanalysis product is ERA5 ([C3S 2017](#)), extending from 1950 to 2020 (in progress), as a part of the Copernicus Climate Change Services, instituted by the European Commission.

In a previous paper ([Cerlini et al. 2017](#)), a solely qualitative preliminary check pointed out a link between the trend of the local water table measurements in the Umbria region (Italy) and the one of the soil moisture data from ERA-Interim dataset. Such promising results encouraged an in-depth analysis that concerned two main aspects: (i) to refine a physically based model for simulating the water flux toward the aquifer by using the soil moisture datasets, and (ii) to use the meanwhile realized ERA5 reanalysis, which offers important changes with respect to ERA-Interim ([Albergel et al. 2018](#); [Hersbach et al. 2020](#)). Precisely, ERA5 is characterized by an improved temporal (from 6 hourly to hourly) and spatial (from 79 km in the horizontal dimension and 60 levels in the vertical to 31 km and 137 levels) resolution. Moreover, ERA5 leads to significant improvements in the representation of the land surface variables and precipitation. As will be shown below with regard to the soil moisture behavior, this reflects in the much better quality of the information included in the ERA5 dataset with respect to ERA-Interim.

In this paper, attention is focused on the simulation of the water table elevation in shallow unconfined aquifers where the main recharge mechanism originates from infiltration. The proposed method for capturing the behavior of the water table elevation is based on the use of the soil moisture dataset from ERA5, within a Richards equation–based approach, and water table elevation measurements executed by the Umbria region. With respect to [Cerlini et al. \(2017\)](#), the executed quantitative analysis provides two relationships that enable simulating the water table elevation as a function of the water flux toward the aquifer on a daily and monthly scale, respectively.

This paper is organized as follows. The second section describes data source and criteria for their selection as well as the model for evaluating the water flux toward the aquifer. The third section introduces the proposed method, the chosen dimensionless quantities, and statistical parameters. In the fourth section, the examined case study is discussed and two relationships for simulating the water table elevation on a daily and monthly scale are proposed. The last section offers remarks about the practical interest and the suitability of the method. The [appendix](#) is dedicated to a detailed check of the performance in terms of water budget in the vadose zone

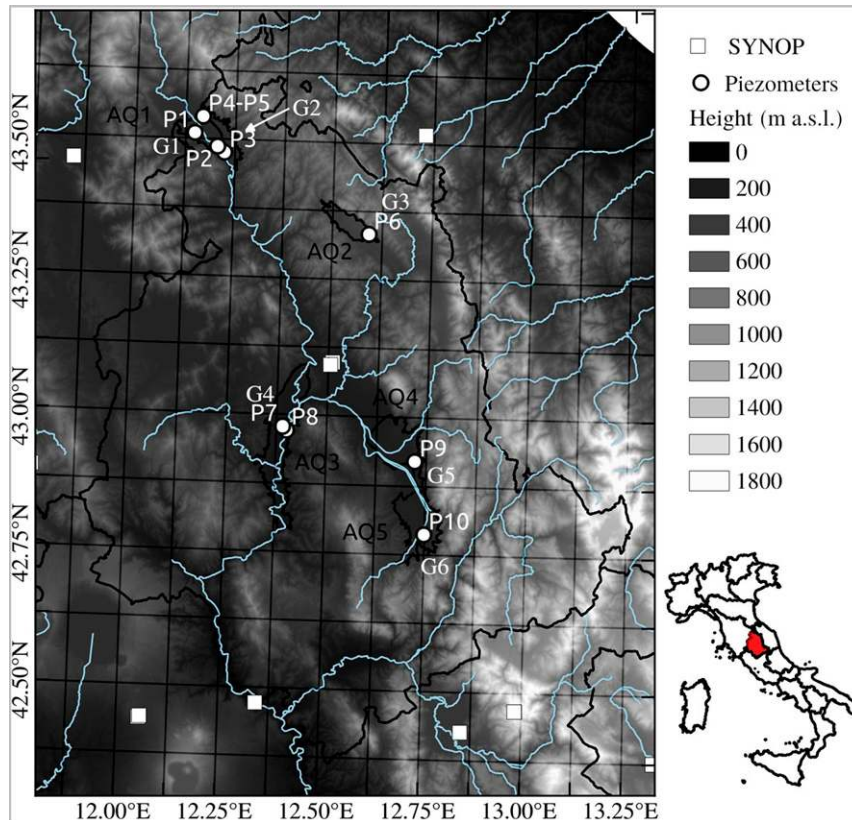


FIG. 1. Orographic map of the Umbria region; white dots indicate piezometers, AQ1–AQ5 indicate the selected aquifers, white squares indicate the synoptic stations (SYNOP) assimilated in the ERA5 reanalysis, and G1–G6 indicate the selected grid points (Table 2); grid lines mark the ERA5 grid points G_i , with a resolution of 0.125° (≈ 15 km).

of the infiltration model based on the soil moisture data from ERA5.

2. Data

A physically based alternative to the use of GCMs for groundwater management requires the preliminary identification of the significant quantities describing the behavior of the aquifer. Beyond the obvious requirement of a clear link with the investigated phenomenon, there are two essential conditions that such quantities must fulfill to be the basis of a robust and “quite” easy to use management tool: (i) to be available easily and (ii) to ensue from independent sources. The first requirement indicates open-access datasets as a strong option. This is the case, as examples, of the measurements collected by public authorities and published on their official website for the citizens’ community as well as the reanalysis provided by research centers and agencies. The second requirement prevents inappropriate relationships (spurious correlations). According to the above statements, two significant independent quantities describing the behavior of the aquifer are the time series of the water table elevation measurements and soil moisture at different depths. In fact, the former dataset

reflects the overall situation of the water body whereas the latter one is linked to its recharge mechanisms. Note that both the above datasets are widely available since water table monitoring by means of piezometers is the most popular method for groundwater control as well as data from reanalysis cover the whole world. In the below analysis, as a representative example of real cases, the attention is focused on Umbria, a hilly region in central Italy (Fig. 1), where a quite dense piezometric monitoring network is available.

a. Water table elevation measurements

The piezometers used in the below analysis are part of the piezometric regional network managed by the Regional Environmental Protection Agency [Agenzia Regionale per la Protezione Ambientale (ARPA)] of the Umbria region of Italy (ARPA 2008). Some piezometers have been active since 2001, when the monitoring network was established, whereas some others were added later to increase the number of the observations.

The available observation frequency of the water table elevation is daily with the value, h_w^d , obtained as a median of the hourly observations previously checked within a quality control procedure (hereinafter, the subscript d indicates daily

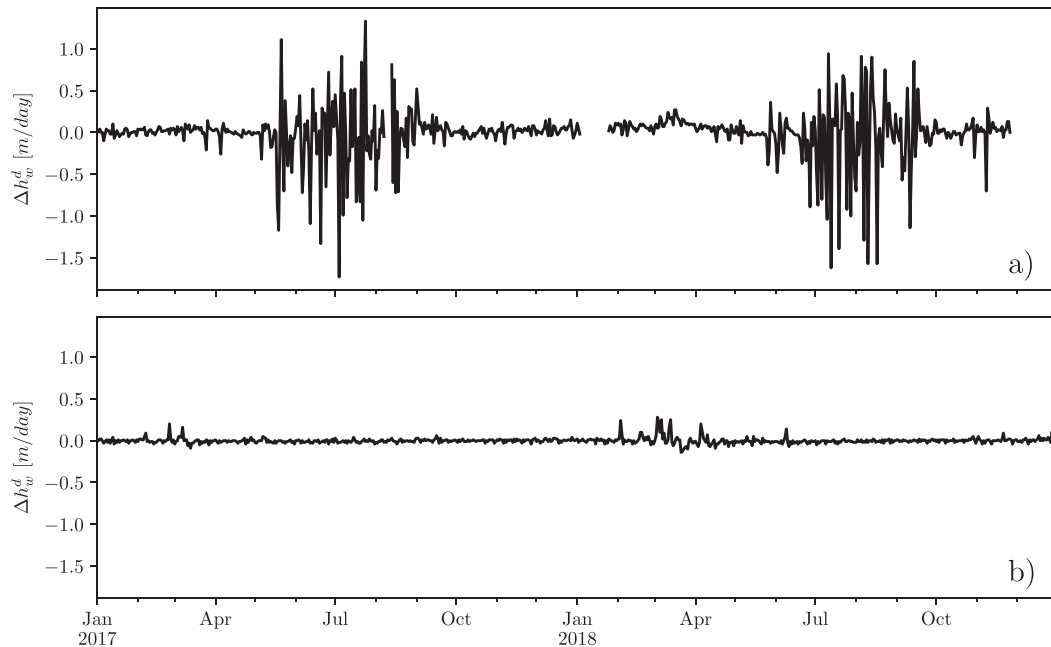


FIG. 2. Daily water table increment Δh_w^d for two neighboring piezometers: (a) the under-review piezometer (Cerbara), with pumping influence, and (b) the reference piezometer (Pistrino—P1), with no pumping influence.

quantities). Precisely, a minimum of three hourly observations has been considered sufficient for evaluating h_w^d (ARPA 2008).

The piezometers considered in the below analysis have been chosen according to the following criteria: (i) the type of the monitored aquifer, (ii) whether pumping occurs in the proximities, and (iii) the value of the mean depth of the water table D with respect to ground level.

With regard to the first criterion, attention has been focused on unconfined and alluvial aquifers; the selected ones are shown in Fig. 1 and indicated as AQ1, AQ2, AQ3, AQ4, and AQ5, respectively. Then the relevant time series of the water table oscillation have been examined and the datasets with a percentage of missing data larger than 60% have been excluded.

The existence of an orderly pumping for drinkable water supply or irrigation purposes—as a reason for exclusion—has been verified not only by consulting local water companies but also by pointing out large water table oscillations occurring in given time intervals (e.g., those when the irrigation is active). An example of this checking procedure is shown in Fig. 2, where the water table daily change Δh_w^d , measured in the under-review piezometer (Fig. 2a) is compared with the one in a neighboring one known as not affected by pumping, assumed as a reference (Fig. 2b). Figure 2 curves clearly show that the oscillations taking place in the reference piezometer are one order of magnitude smaller than those in the examined one, which has been thereby excluded.

Piezometers where the mean water table depth D ranges between 4 and 10 m have been included in the analysis. It is worth noting that the lower limit of such a range of D values is larger enough than the depth investigated by the ERA5 hydrology model (i.e., 2.89 m; see below) whereas the upper limit

can be assumed as a rational maximum depth of shallow aquifers where the infiltration plays the role of main recharge mechanism. According to Seibert et al. (2003), the selected range of D (i.e., several meters) authorizes to assume that the connection between the vadose zone and the aquifer is unidirectional and the vadose zone–groundwater interaction is negligible.

As a final result, 10 piezometers, fulfilling the abovementioned criteria, have been considered in the following analysis; the main characteristics of these piezometers as well as the relevant grid points are reported in Table 1.

b. ERA5 reanalysis

The ERA5 reanalysis (Hersbach et al. 2020) includes two different realizations: one deterministic with a high resolution (HRES) and a 10-member ensemble reduced resolution (EDA). In this paper, the high-resolution version of ERA5 has been used for evaluating the soil moisture time series.

Since the horizontal resolution of HRES is about 0.28° (≈ 31 km), in Umbria there are about 20 grid points. To obtain the model grid points closer to the selected piezometers, ERA5 data have been bilinearly interpolated to the higher-resolution grid of about 0.125° shown in Fig. 1. Then, the soil moisture variable at the location of the considered piezometers (Fig. 1), have been extracted at grid points G1, G2, G3, G4, G5, and G6 at an hourly resolution from 2001 to 2018.

Reanalysis are produced by using a constant model framework in order to ensure data consistency through several decades. In particular, ERA5 is produced by using the Integrated Forecasting System (IFS) (version Cy41r2; ECMWF 2016b). This system includes several components that interact together: the atmospheric

TABLE 1. Piezometers, from the ARPA Umbria network, selected for the analysis. Columns from left to right: identification number P and name identifying the piezometer; latitude and longitude (in decimal degrees) and height (in meters above sea level as extracted from the digital elevation model); the corresponding ERA5 grid point G and aquifer (Fig. 1); starting year of measurements; mean depth of the water table (m) with respect to the ground level; standard deviation of the water table measurements (m).

| P | Name | Lat | Lon | Height | G | Aquifer | Start | D | $\sigma_{h_w^d}$ |
|-----|-------------------------------------|---------|---------|--------|----|---------|-------|-----|------------------|
| P1 | Pistrino | 43.5099 | 12.1475 | 294 | G1 | AQ1 | 2001 | 4.5 | 0.6 |
| P2 | Piosina | 43.4862 | 12.2055 | 281 | G2 | AQ1 | 2006 | 4.9 | 0.5 |
| P3 | Riosecco | 43.4756 | 12.2232 | 286 | G2 | AQ1 | 2001 | 4.5 | 0.5 |
| P4 | San Giustino superficiale | 43.5403 | 12.1671 | 303 | G1 | AQ1 | 2006 | 8.8 | 2.3 |
| P5 | San Giustino profondo | 43.5403 | 12.1671 | 303 | G1 | AQ1 | 2006 | 9.9 | 2.5 |
| P6 | Gubbio | 43.3333 | 12.5916 | 461 | G3 | AQ2 | 2006 | 8.4 | 2.1 |
| P7 | Barche | 42.9779 | 12.3902 | 162 | G4 | AQ3 | 2001 | 5.9 | 0.8 |
| P8 | Pescheto | 42.9713 | 12.4000 | 160 | G4 | AQ3 | 2001 | 7.4 | 0.5 |
| P9 | San Eraclio | 42.9203 | 12.7213 | 216 | G5 | AQ4 | 2001 | 6.4 | 0.7 |
| P10 | San Giacomo di Spoleto superficiale | 42.7870 | 12.7486 | 240 | G6 | AQ5 | 2006 | 5.4 | 3.2 |

model, the land model [the soil hydrology scheme of the Tiled ECMWF Scheme for Surface Exchanges over Land (H-TESSEL)], the Nucleus for European Modelling of the Ocean (NEMO), the ECMWF Ocean Wave Model (ECWAM), and the data assimilation system (4D-Var). As a consequence, each quantity extracted from ERA5 reanalysis, for example, the soil moisture, is the result of the interaction between all the mentioned components. Precisely, the atmospheric model interacts through surface fluxes, as well as through all the meteoric species, with the H-TESSEL land surface model (Balsamo et al. 2009; ECMWF 2016b). Moreover, the observations assimilated within the atmospheric model (for a complete list, see Hersbach et al. 2020) indirectly interact with the surface model component. The last model component, 4D-Var, is responsible of combining observations and short range forecasts (background or first guess) to produce the best estimate of the initial conditions for each model component. This system is very complex, and a detailed explanation is beyond the scope of this paper. An exhaustive description of the ERA5 data assimilation system, together with all the observations used by the 4D-Var, is reported in Hersbach et al. (2020), Dee et al. (2011), and ECMWF (2016a). For a specific description of the Land Data Assimilation System (LDAS), which is responsible of providing the initial conditions for the H-TESSEL model, the reader may refer also to de Rosnay et al. (2014, 2013). It is worth noting that two main sources of observations directly influence the soil moisture analysis (Albergel et al. 2012): the surface observations of temperature and relative humidity from synoptic stations (SYNOP), measured at 2 m above the ground level (the so-called screen level), and *MetOp-A* and *MetOp-B* Advanced Scatterometer (ASCAT) soil moisture data from satellites. The synoptic stations present in the analyzed domain are shown in Fig. 1. Screen-level parameters are indirectly related to soil moisture, while satellites provide a more direct measurement of the surface soil moisture. Since the latter source is capable of describing only the top few centimeters of the soil (Albergel et al. 2012), the root-zone soil moisture is estimated by propagating downward this information by means of the H-TESSEL model. The assimilation of satellites and synoptic stations data in the land surface model has improved the ECMWF forecast performances in the atmospheric boundary

layer (Drusch and Viterbo 2007; de Rosnay et al. 2013, 2014; Fairbairn et al. 2019). Since ERA5 uses this type of assimilation, all the improvements reflect also in the data used in the analysis.

The H-TESSEL model computes the infiltration and evaporation rate at the surface by solving a moisture balance and then the water flux through the unsaturated zone by means of the Richards equation:

$$\frac{\partial \theta}{\partial t} = -\frac{\partial F_w}{\partial z} + S_\theta, \tag{1}$$

where θ = volumetric soil water content (or soil moisture) expressed in meters cubed per meters cubed, t = time, z = soil depth, F_w = soil moisture flux per unit area (positive downward) expressed in meters cubed per meters squared per second, and S_θ = root extraction sink term in meters cubed per meters cubed per second.

Equation (1) is solved over a soil column, assumed as homogeneous and discretized into four layers, reaching a depth of 2.89 m (the size of the soil layers Δz_k is indicated in Fig. 3); θ is given at the layer center, whereas F_w is calculated at the interface (see below).

The gridpoint features are listed in Table 2; in Fig. 4, as an example, the soil moisture time series for the grid point G4 is shown. In this figure, data from ERA-Interim are also reported as an example of the clear differences in terms of the soil moisture series between the two global datasets. It is worth pointing out that, according to Hersbach et al. (2020), the local discontinuity in ERA5 dataset between 2009 and 2010 has been eliminated.

A measure of the difference in terms of temporal scales between the soil layers is given by the coefficient of variation (CV) of the soil moisture:

$$CV_k = \frac{\sigma_{\theta_k}}{\bar{\theta}_k}, \tag{2}$$

where σ_{θ_k} and $\bar{\theta}_k$ are the standard deviation and the mean value of soil moisture over the k th soil layer evaluated over the considered temporal range (2001–18), respectively. According to Iqbal et al. (2005) and Cameron (1978), as a response to the atmospheric forcing, the deeper the soil layer, the smaller CV (Table 2). Moreover, the entity of such a difference between

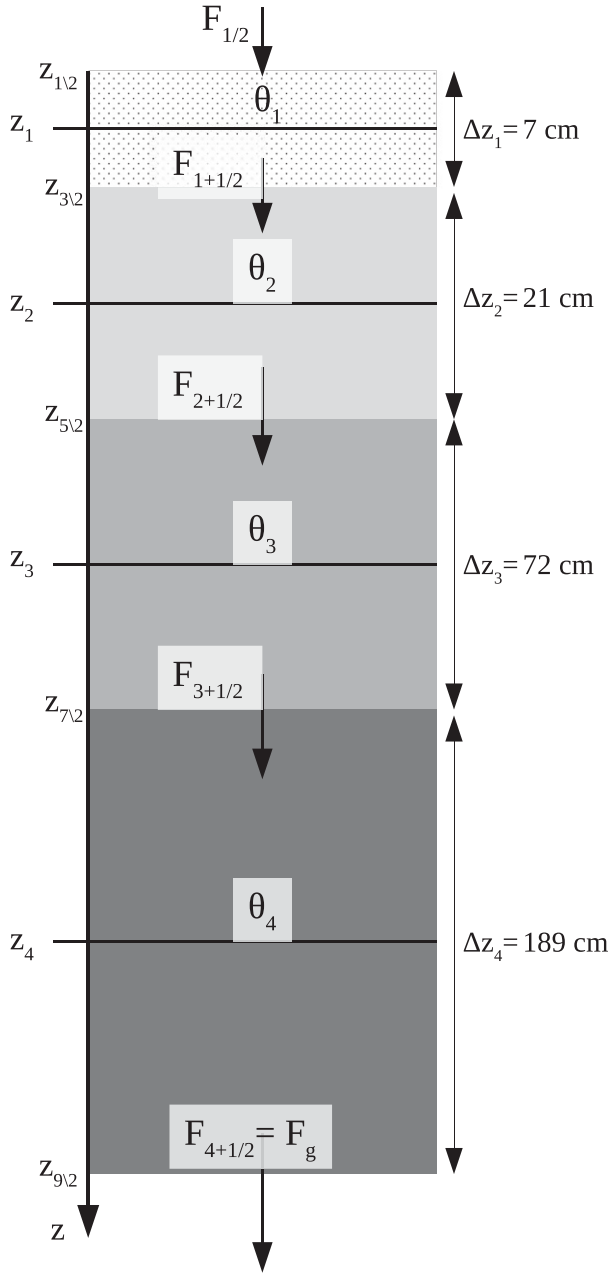


FIG. 3. Discretization of the soil column used for solving the Richards equation with the soil moisture contents given at the layer center z_k and the water fluxes evaluated at the interface $z_{k+1/2}$.

the fourth and the other soil layers can be ascribed to the large difference in terms of layer depth. In fact, the depth of the fourth layer is almost twice the total depth of the first three ones (Fig. 3).

c. Water flux toward the aquifer

According to Eq. (1), in the executed simulations, the water flux through the vadose zone has been obtained by means of the following relationship:

$$F_w = \lambda(\theta) \frac{\partial \theta}{\partial z} - \gamma(\theta), \quad (3)$$

where $\lambda(\theta)$ ($\text{m}^2 \text{s}^{-1}$) and $\gamma(\theta)$ (m s^{-1}) are the hydraulic diffusivity and conductivity, respectively. In the below simulations, the vegetation transpiration component S_θ , in the fourth soil layer has been neglected with respect to the drainage term $\partial F_w / \partial z$.

Hydraulic conductivity and diffusivity have been evaluated by means of the van Genuchten formulation (van Genuchten 1980):

$$\gamma = \gamma_{\text{sat}} \frac{\{[1 + (\alpha p)^n]^{1-1/n} - (\alpha p)^{n-1}\}^2}{[1 + (\alpha p)^n]^{(1-1/n)(l+2)}} \quad (4)$$

and

$$\lambda = \gamma \left| \frac{dp}{d\theta} \right|, \quad (5)$$

where γ_{sat} = saturated hydraulic conductivity, and the pressure head p is obtained on the basis of the value of the soil moisture given by ERA5 by means of the following relationship:

$$\theta(h) = \theta_{\text{res}} + \frac{\theta_{\text{sat}} - \theta_{\text{res}}}{[1 + (\alpha p)^n]^{(1-1/n)}}, \quad (6)$$

where θ_{sat} and θ_{res} are saturated and residual soil moisture, respectively, and n and α are constants that depend on the soil type. The soil texture type has been extracted from the ERA5 grid points closer to the selected piezometers, by following the same procedure used for the soil moisture variable. In the analyzed area, the loamy soil is the only relevant soil texture type, as given by the FAO dataset (FAO 2003). The correspondent values of the constant parameters θ_{sat} , θ_{res} , n , and α used in Eqs. (4)–(6) are reported in Table 3. The reader may refer to (ECMWF 2016b, chapter 8.6.2) for a detailed description of how such parameters have been obtained.

As a boundary condition at the bottom, free drainage is assumed since, as mentioned above, the mean water table depth is large compared to the soil column thickness (2.89 m for ERA5). In other words, it is postulated that the interaction between the unsaturated and the saturated zone is negligible and no upward diffusion flux from the groundwater influences the soil moisture of the above layers (Yeh and Eltahir 2005). Accordingly, the water flux toward the aquifer F_g is given by

$$F_g = \gamma_{4\text{th}}, \quad (7)$$

where $\gamma_{4\text{th}}$ is a representative value of the hydraulic conductivity that makes it possible to simulate the interaction between the vadose zone and aquifer.

3. Method

As mentioned, in the assumed groundwater recharge mechanism, the water flux toward the aquifer F_g is the main quantity responsible for the changes of the water table elevation. In other words, in a shallow groundwater F_g , which is below the root zone, reaches the water table, with a possible delay of few days, and recharges the aquifer. Within this

TABLE 2. ERA5 gridpoint features and the coefficient of variation (%) of the soil moisture for the four soil layers.

| Grid point | Lat (°) | Lon (°) | Height (m MSL) | CV ₁ (%) | CV ₂ (%) | CV ₃ (%) | CV ₄ (%) |
|------------|---------|---------|----------------|---------------------|---------------------|---------------------|---------------------|
| G1 | 43.5 | 12.125 | 529 | 24.4 | 25.4 | 26.2 | 9.7 |
| G2 | 43.5 | 12.25 | 519 | 24.4 | 25.4 | 26.3 | 9.8 |
| G3 | 43.375 | 12.625 | 459 | 24.3 | 24.6 | 25.6 | 9.4 |
| G4 | 43 | 12.375 | 390 | 30.9 | 32.3 | 32.8 | 11.5 |
| G5 | 42.875 | 12.75 | 641 | 30.6 | 33.3 | 35.4 | 12 |
| G6 | 42.75 | 12.75 | 655 | 30.3 | 30.8 | 30.4 | 10 |

approach, it is crucial to check whether, over a given period of time I , a representative value of the water table elevation h_w^I , is correlated with the corresponding integral flux F_g^I , that is, the sum over I of the fluxes toward the aquifer F_g . It is worth pointing out that, according to the needs of groundwater management, in the below analysis attention focused on two time intervals: the day ($I = d$) and month ($I = m$).

In dimensionless terms, the conjectured link between h_w^I and F_g^I is examined by considering the standardized relative water table elevation:

$$D_r^{I*} = \frac{h_w^{\max} - h_w^I}{\sigma_{h_w^I}} \tag{8}$$

and the relative flux toward the aquifer:

$$F_g^{I*} = \frac{F_g^{\max}}{F_g^I} \tag{9}$$

In Eq. (8), $\sigma_{h_w^I}$ and h_w^{\max} , the standard deviation and a reference maximum value of the water table elevation, respectively, characterize the dynamics of the aquifer in the considered period of time; in Eq. (9), F_g^{\max} is the maximum value of the flux. Note that the parameters D_r^{I*} and F_g^{I*} derive from two independent data sources: the first comes from water table observations, whereas the latter is calculated from reanalysis data as described in section 2c.

Following the usual practice, the value of appropriate statistics parameters will indicate the relevance of the dependence of F_g^{I*} on D_r^{I*} , and then the one of F_g^{I*} on h_w^I , for the considered period of time. Precisely, two parameters have been used. The

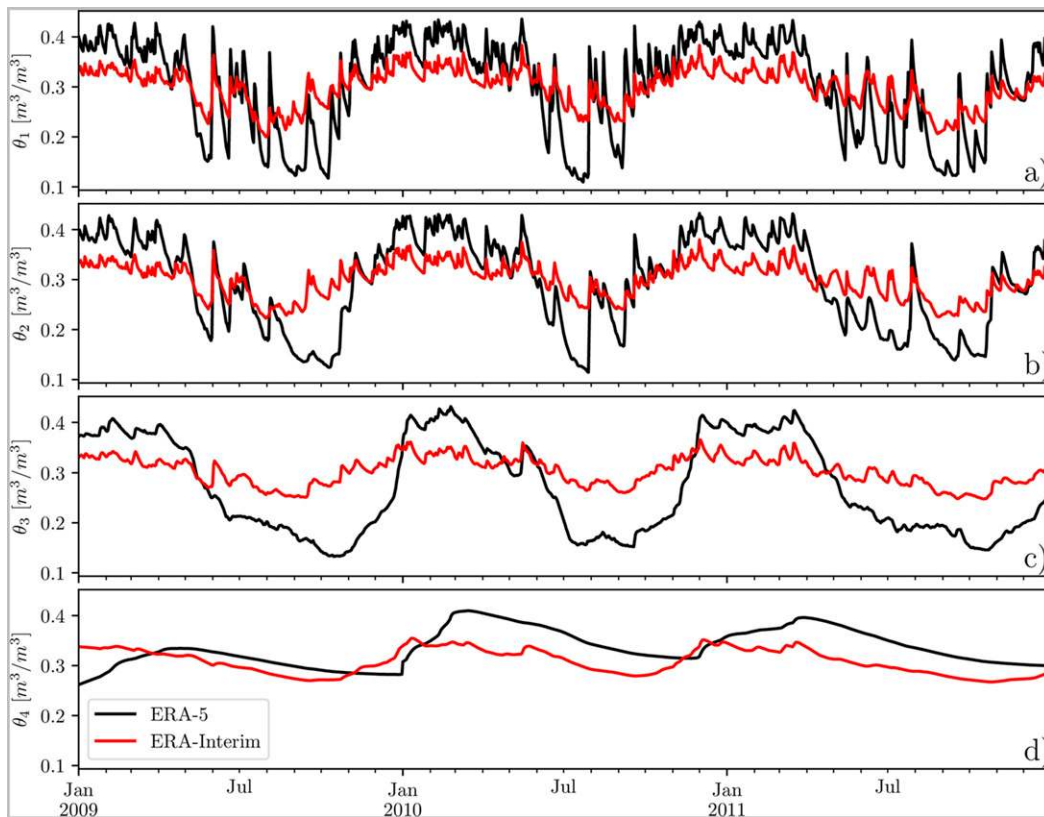


FIG. 4. Soil moisture θ_k time series from 2009 to 2011 extracted at grid point G4 for both ERA-Interim and ERA5 reanalyses: (a) soil layer 1, (b) soil layer 2, (c) soil layer 3, and (d) soil layer 4.

TABLE 3. Van Genuchten soil parameters for the land texture type considered in the analyzed grid points of Fig. 1.

| Soil texture | α | l | n | γ_{sat} | θ_{sat} | θ_{res} |
|---------------------|-------------------------|--------|------|--|-----------------------|-----------------------|
| Medium (loamy soil) | 3.14 (m ⁻¹) | -2.342 | 1.28 | 1.16 (10 ⁻⁶ m s ⁻¹) | 0.439 | 0.010 |

first is the cross-correlation function $R(\tau)$ evaluated, for the sake of generality, by means of the Spearman rank equation:

$$R(\tau) = \frac{\sum_{i=1}^{N-\tau} [(r_h^i - \bar{r}_h^+) (r_F^{i+\tau} - \bar{r}_F^-)]}{\sqrt{\sum_{i=1}^{N-\tau} (r_h^i - \bar{r}_h^+)^2} \sqrt{\sum_{i=1}^{N-\tau} (r_F^{i+\tau} - \bar{r}_F^-)^2}}, \quad (10)$$

where τ = time lag, N = number of available time steps, r_h^i and r_F^i are the rank of the water table measurements and the fluxes, respectively, at time step i , and the plus- and minus-sign subscripts on the overbar indicate the sample mean over the first and last $n - \tau$ time steps, respectively.

The second parameter is the Nash-Sutcliffe efficiency (NSE) coefficient defined as

$$\text{NSE} = 1 - \frac{\sum_{i=1}^N (h_{w,i}^{\text{Obs}} - h_{w,i}^{\text{Opt}})^2}{\sum_{i=1}^N (h_{w,i}^{\text{Obs}} - \bar{h}_{w,i}^{\text{Obs}})^2}, \quad (11)$$

where $h_{w,i}^{\text{Obs}}$ and $h_{w,i}^{\text{Opt}}$ are respectively the observed and simulated water table elevation at the i th time instant, and N = total number of the available observations.

4. The Umbria region case study

As anticipated, the proposed approach, in which a correlation between h_w^l and F_g^l over the time interval l is assumed, has been checked, both on a daily and monthly scale, by considering the water table measurements executed in the Umbria region by ARPA and soil moisture data given by ERA5.

Because of the mentioned very rough spatial discretization of the soil layers used in the H-TESSSEL model, which reflects in the resolution of the available soil moisture profile, a preliminary check has concerned the ERA5 data. Precisely, the consistency in terms of water mass conservation of the soil moisture data has been verified because of the well-known crucial role played by the spatial and temporal discretization in the numerical solution of the Richards equation (e.g., Celia and Binning 1990). Moreover, as discussed below, this check indicated a proper value of the soil moisture, and then of the hydraulic conductivity, that is representative of the water flux toward the aquifer, according to Eq. (7).

a. Preliminary water budget check in the vadose zone

A detailed analysis of the water budget formulation is reported in the appendix, whereas this section presents its key factors and main results.

With the aim of assessing the sensitivity of the soil water budget to the method used for calculating interlayer properties, that plays

an important role in the numerical integration of the Richards equation (Brunone et al. 2003), two different options have been considered: (i) a constant value of the soil moisture equal to the ERA5 one as in H-TESSSEL (Fig. 5a), hereinafter referred to as the max_θ approximation; and (ii) a linear interpolation between the values of θ in two successive layers (Fig. 5b), hereinafter referred to as the lin_θ approximation.

The performance of the soil moisture profile given by ERA5 and of such approximations has been evaluated by considering the error in the daily water budget:

$$\epsilon_k^d = \frac{100}{\bar{W}_k} \left\{ \sum_{t \in d} (W_k^{t+1} - W_k^t) - \sum_{t \in d} [\beta (F_{k-1/2}^{t+1} - F_{k+1/2}^{t+1}) + (1 - \beta) (F_{k-1/2}^t - F_{k+1/2}^t)] \Delta t \right\} \quad (12)$$

with $\Delta t = 1$ h, $F_{k+1/2}$ = flux between layers k and $k + 1$, and W_k^t = soil moisture volume per unit area at layer k at time t :

$$W_k^t = \int_{z_{k-1/2}}^{z_{k+1/2}} \theta^t(z) dz, \quad (13)$$

and \bar{W}_k = daily average of $W_k(t)$. The first sum on the RHS of Eq. (12) represents the daily water integral change, while the second sum is the difference between the incoming daily integral fluxes and outgoing daily integral fluxes relative to the k th layer. The coefficient β indicates whether the interpolation is implicit ($\beta = 1$), explicit ($\beta = 0$), or semi-implicit ($\beta = 0.5$). Note that the day, as time interval for the check of the water budget, has been chosen according to both the actual needs of groundwater management and the inertia of aquifers. The water budget check concerned all regions (Fig. 1) represented by the grid points listed in Table 2. Figure 6 shows, as an example, the results for the grid point G2, representative of the aquifer AQ1 (Table 1). Precisely, for the given θ_4 time history in 2015–16 (Fig. 6a), the corresponding values of the flux toward the fourth layer, $F_{3+1/2}$ (Fig. 6b), and toward the aquifer, $F_g = F_{4+1/2}$ (Fig. 6c), within both the max_θ and lin_θ approximations for $\beta = 0.5$, are reported. With regard to $F_{3+1/2}$, the larger values obtained within the max_θ approximation are properly due to the larger values of the soil moisture and then of the hydraulic conductivity. With regard to F_g , it is important to point out that its behavior does not depend significantly on the chosen soil moisture approximation. Moreover, for both the approximations, the daily mass error in the fourth layer ϵ_4^d is very small, which confirms the consistency of the soil moisture profile given by ERA5 (Fig. 6d). Such a result is reaffirmed by the behavior of the mass error in the selected grid point in the whole examined period of time (2001–18) (Fig. 7). The same behavior characterizes the mass balance values when the explicit ($\beta = 0$) or the implicit interpolation ($\beta = 1$) is implemented (not shown).

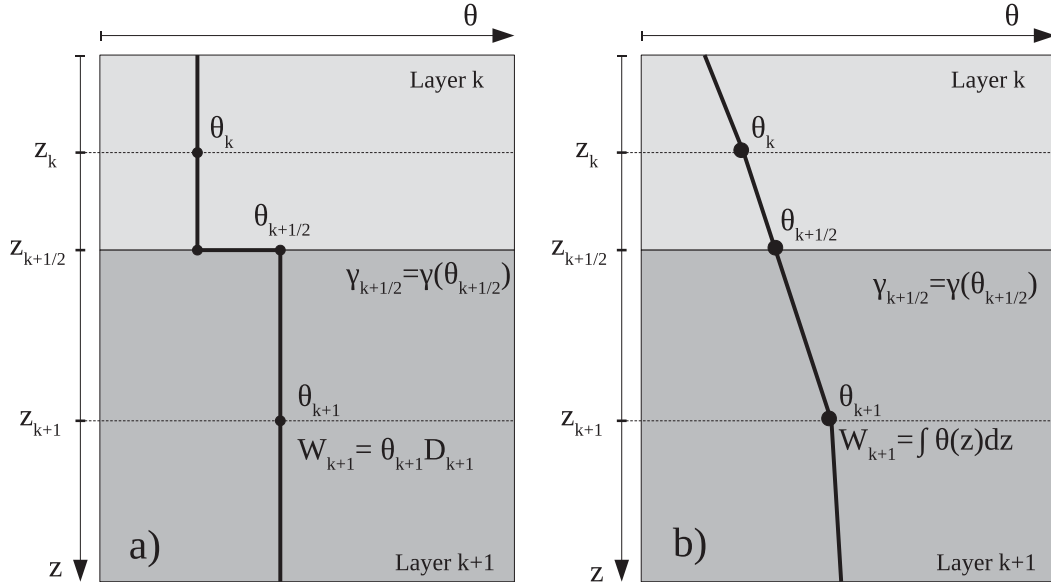


FIG. 5. Approximations used for evaluating the interlayer hydraulic properties: (a) \max_θ , and (b) lin_θ . Within the water budget check, the first approximation, used by H-TESSSEL evaluates the interlayer properties by considering the largest volumetric water content in the above layer whereas the second approximation assumes that the interlayer properties are a function of θ at the interlayer $\theta_{k+1/2}$ obtained by a linear interpolation of θ_k and θ_{k+1} between the two adjacent layers.

The above results indicate that, in terms of mass balance, the soil moisture profile given by ERA5 makes it possible to evaluate properly the flux toward the aquifer irrespective of the soil moisture approximation (\max_θ or lin_θ) and the interpolation chosen for evaluating the mass balance error. Accordingly, the below simulations use the \max_θ approximation with the semi-implicit interpolation and $\gamma_{4\text{th}}$ of Eq. (7) is evaluated by assuming $\theta = \theta_4$.

b. Water flux toward the aquifer versus water table elevation on a daily scale

Plots for piezometer P5 and the related cell, a representative example of the considered cases, give an idea of the outlined link between F_g^d and h_w^d (Fig. 8). Precisely (Fig. 8c), it points out the delay between F_g^d and h_w^d with the maximum value of the cross-correlation function R_{\max} and its correspondent time lag τ_{\max} pointed out. For the piezometer P5 and related cell, it results in $\tau_{\max} = 4$ days, as highlighted by the zoom in Fig. 8d; for all the other cells/piezometers, τ_{\max} is smaller than 9 days (Table 4).

The obtained small time lags and large correlation coefficients (Table 4) are confirmed by the scatterplot between daily water table levels and the daily integral fluxes for piezometer P5 of Fig. 9. This figure clearly outlines the nonlinear relation between F_g^d and h_w^d .

Such a behavior characterizes most of the piezometers (and related cells); only a few cases exhibit a very linear behavior (e.g., piezometer P3). To take into account this mixed linear–not linear behavior, the following relationship has been assumed between the dimensionless parameters D_r^{d*} and F_g^{d*} :

$$D_r^{d*} = k_{\log} \log(F_g^{d*}) + k_{\text{lin}} \left(1 - \frac{1}{F_g^{d*}} \right), \quad (14)$$

where the dimensionless coefficients k_{\log} and k_{lin} have been obtained by means of the nonlinear least squares method (Table 4).

For each piezometer, the values of the statistical parameters given by Eqs. (8) and (9) are reported in Table 4. Note that for each piezometer h_w^{max} is the 99th percentile of the water table elevation, and $\sigma_{h_w^d}$ is the mean absolute deviation (MAD) of the water table elevation, a more robust measure of the standard deviation of h_w^d .

According to Eq. (14), D_r^{d*} is the summation of a logarithmic and a linear function of F_g^{d*} , weighted by the constants k_{\log} and k_{lin} , respectively. The value of the ratio k_{lin}/k_{\log} (Table 4) indicates which component is predominant. If k_{lin}/k_{\log} tends to negative values, then the relation between D_r^{d*} and F_g^{d*} can be assumed to be almost logarithmic. As k_{lin}/k_{\log} approaches unity, the two components overlap. In this case, the linear term is more effective in correspondence of the large fluxes (i.e., small relative fluxes F_g^{d*}), while the logarithmic term plays a major role in correspondence of the small fluxes (i.e., large relative fluxes F_g^{d*}). When k_{lin}/k_{\log} is larger than unity, the relation between D_r^{d*} and F_g^{d*} becomes almost linear. This behavior is shown in Fig. 10, where the piezometers have been sorted from left to right in increasing order by the ratio k_{lin}/k_{\log} .

The linear component (represented by the dotted red curves) increases from the left to the right, while the opposite is true for the logarithmic component (dashed red lines). Moreover, the linear component approaches the constant value k_{lin} for the larger values of F_g^{d*} :

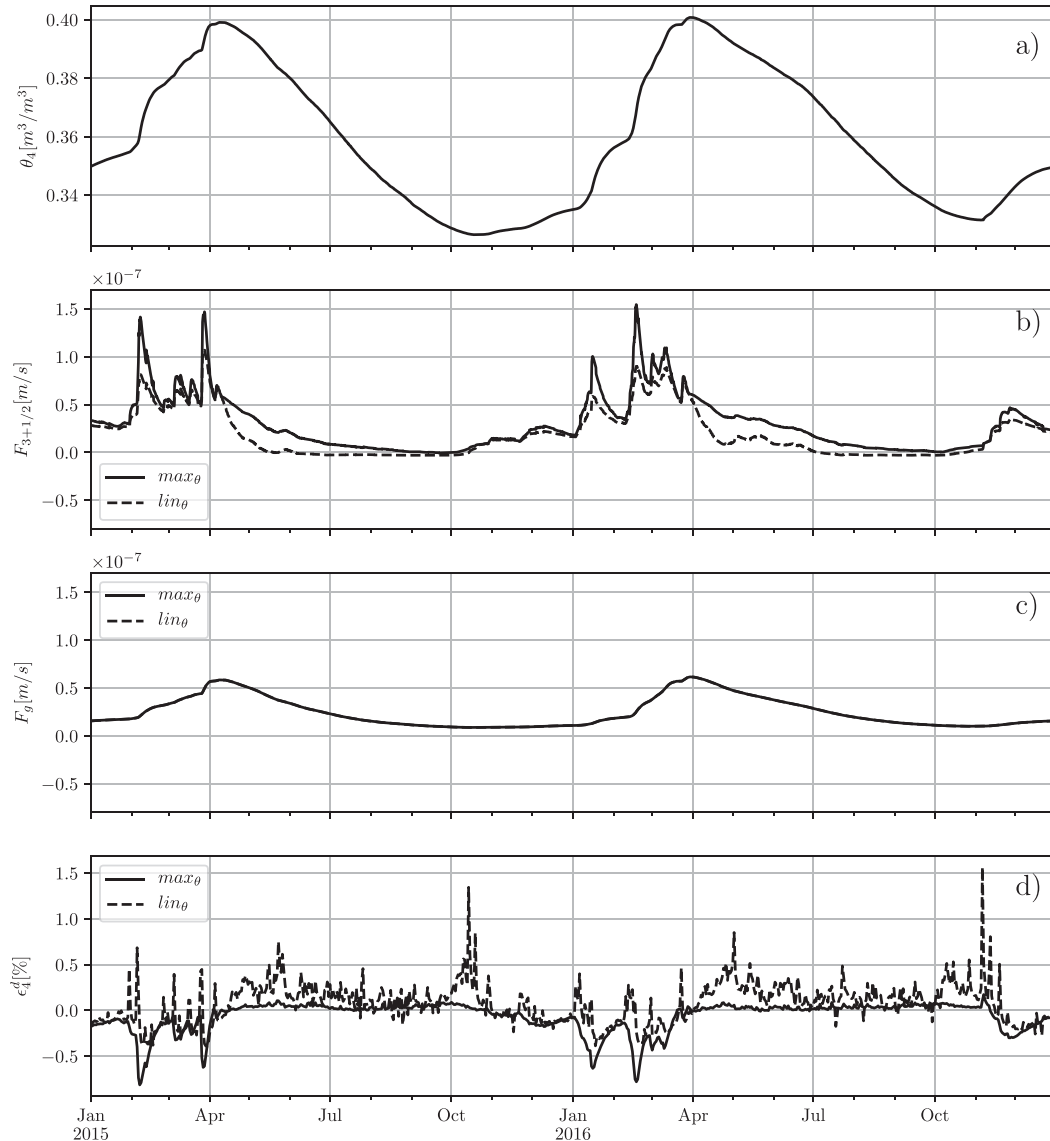


FIG. 6. Grid point G2 (43.5°N, 12.25°E) in the Umbria region (Fig. 1). Time history for both the max_θ and lin_θ approximations and $\beta = 0.5$ of (a) soil moisture at the fourth layer θ_4 , (b) flux toward the fourth layer $F_{3+1/2}$, (c) flux toward the aquifer F_g , and (d) daily mass balance error ϵ_4^d in the fourth layer.

$$\lim_{F_g^{d*} \rightarrow \infty} k_{lin} \left(1 - \frac{1}{F_g^{d*}} \right) = k_{lin}. \quad (15)$$

As the ratio k_{lin}/k_{log} increases, k_{lin} gets closer and closer to the mean relative standardized depth D_r^{d*} (dashed gray horizontal line), defined as

$$\overline{D_r^{d*}} = \frac{h_w^{\max} - \overline{h_w^d}}{\sigma_{h_w^d}}, \quad (16)$$

where $\overline{h_w^d}$ is the mean daily water table elevation. Values larger than $\overline{D_r^{d*}}$ denote negative anomalies of the water table depth (dry periods), and values smaller than $\overline{D_r^{d*}}$ denote positive anomalies (wet periods). The direct link between $\overline{D_r^{d*}}$

and k_{lin}/k_{log} could provide a useful tool for simulating the time behavior of the water table elevation for the piezometers with a very large ratio, such as the Piosina (P2) one. Note the link between the spread of the water table elevation, measured by the MAD (Table 4), and k_{lin}/k_{log} . The larger this ratio is, the smaller is the spread of the water table elevation, and the smaller is the range of values of the relative elevation. Piezometers behaving mainly linearly exhibit smaller oscillations than those characterized by a logarithmic behavior. Such a nonlinear behavior between the flux and the water table elevation, taken into account by the logarithmic term, could be due to a weak interaction between the saturated and unsaturated zone. In other words, small-spreading shallow aquifers show a predominant linear behavior, whereas the behavior of the large-spreading aquifers is nonlinear.

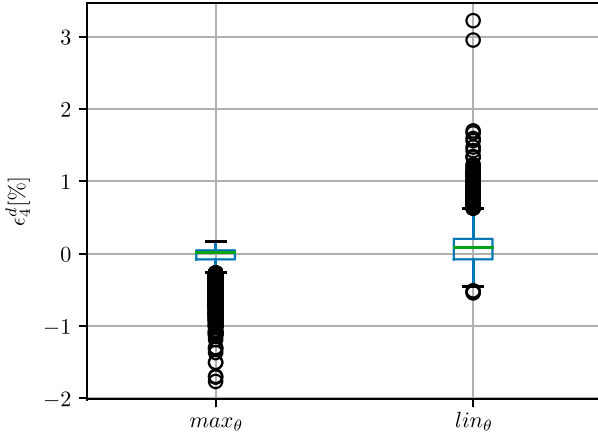


FIG. 7. Grid point G2 (43.5°N, 12.25°E) in the Umbria region (Fig. 1). Boxplots of the mass error ϵ_4^d in the fourth layer for the period (2001–18).

The water table elevation is obtained by substituting Eqs. (8) and (9) in Eq. (14):

$$h_w^d = h_w^{\max} - \sigma \left[k_{\log} \log \left(\frac{F_g^{\max}}{F_g^d} \right) + k_{\text{lin}} \left(\frac{F_g^{\max} - F_g^d}{F_g^{\max}} \right) \right]. \quad (17)$$

In Fig. 11, the values of the water table elevation given by Eq. (14) are compared with the observed ones.

To assess the reliability of the proposed method, the NSE coefficient has been evaluated between the observed and simulated value of the water table elevation (Table 4). The NSE values

indicate that the proposed model is satisfactory for all piezometers, with the only exception of P7, where the NSE is slightly smaller than 0.5. To be precise (Fig. 11), both the phase and larger scales of the water table behavior are very well captured by Eq. (14). For the sake of completeness, the last columns of Table 4 reports also the value of NSE obtained by assuming only a linear relation [i.e., by substituting $k_{\log} = 0$ in Eq. (14)]. All of these values are smaller than the values obtained by a log-linear relationship. In particular, the NSE undergoes a drastic reduction for those piezometers characterized by negative values of the ratio k_{lin}/k_{\log} as P4, P5, P9, and P10. This confirms the important role of this ratio on determining whether the behavior of the piezometer is mainly logarithmic or linear.

c. Water flux toward the aquifer versus water table elevation on a monthly scale

Having in mind that, for management purposes, a monthly forecast is more effective than the daily one, the successive step of the analysis focused on the monthly values h_w^m and F_g^m following the same approach. In this case, the monthly standardized relative depth D_r^{m*} and the monthly relative flux F_g^{m*} are obtained by substituting h_w^m and F_g^m in Eqs. (8) and (9), respectively. The reference parameters characterizing the piezometers, h_w^{\max} , MAD, and F_g^{\max} , have been kept equal to those of the daily quantities (Table 4).

Accordingly, a relation of the same type of Eq. (14) has been assumed to be valid also for the monthly quantities:

$$D_r^{m*} = k_{\log}^m \log(F_g^{m*}) + k_{\text{lin}}^m \left(1 - \frac{1}{F_g^{m*}} \right), \quad (18)$$

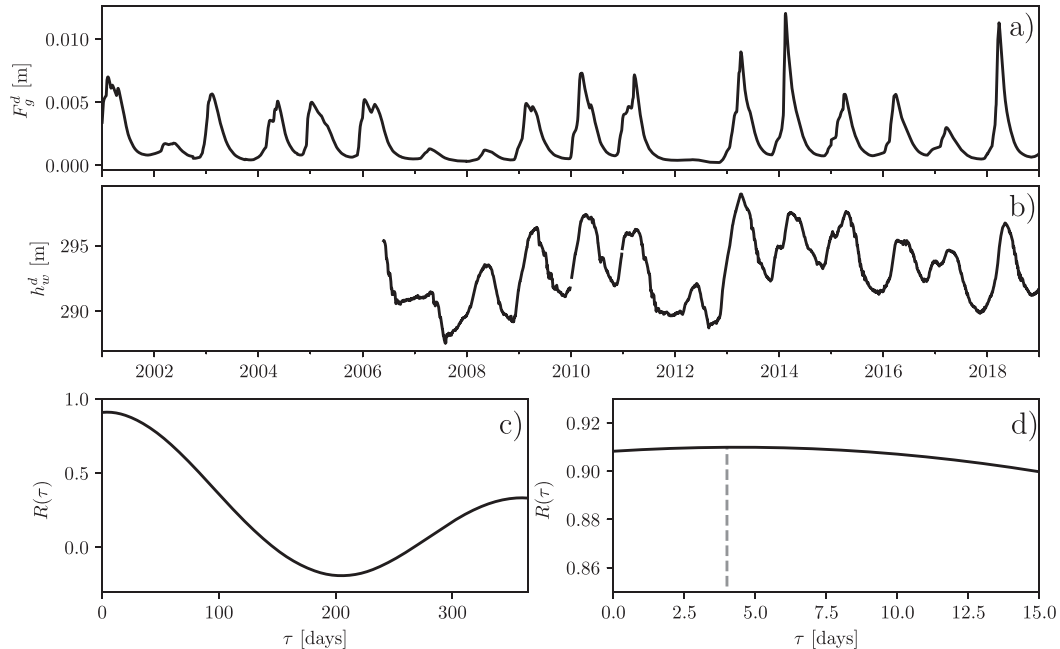


FIG. 8. Cross-correlation function for piezometer P5: (a) time series of the daily integral fluxes F_g^d ; (b) time series of the daily water table elevation h_w^d ; (c) cross-correlation function $R(\tau)$; (d) behavior of the cross-correlation function for the small values of the time lag.

TABLE 4. Daily statistical properties of water table observations and interpolation parameters. Columns from left to right: piezometer identifier; maximum Spearman coefficient R_{\max} as in Eq. (10); time lag of maximum correlation τ_{\max} ; 99th percentile of h_w , h_w^{\max} ; mean absolute deviation of h_w ; absolute maximum value of flux F_g^{\max} ; mean relative standardized depth as in Eq. (16) D_r^{d*} ; simulation parameters and their ratio k_{\log} , k_{lin} , and k_{lin}/k_{\log} ; Nash–Sutcliffe efficiency coefficient as in Eq. (11) for the logarithmic–linear relation; Nash–Sutcliffe efficiency coefficient for the linear relation.

| P | R_{\max} | τ_{\max} (days) | h_w^{\max} [m] | MAD (m) | F_g^{\max} [m] | D_r^{d*} | k_{\log} | k_{lin} | k_{lin}/k_{\log} | NSE | NSE_{lin} |
|-----|------------|----------------------|------------------|---------|------------------|------------|------------|------------------|---------------------------|-------|---------------------------|
| P1 | 0.807 | 0 | 291.1 | 0.497 | 0.012 | 3.26 | 0.786 | 1.87 | 2.38 | 0.661 | 0.542 |
| P2 | 0.826 | 0 | 277.68 | 0.394 | 0.0108 | 4.02 | 0.507 | 3.541 | 6.99 | 0.648 | 0.602 |
| P3 | 0.758 | 0 | 282.95 | 0.36 | 0.0108 | 4.18 | 0.295 | 4.367 | 14.79 | 0.569 | 0.555 |
| P4 | 0.921 | 9 | 298.49 | 1.943 | 0.012 | 2.21 | 1.325 | −0.902 | −0.68 | 0.826 | 0.443 |
| P5 | 0.91 | 4 | 298.26 | 2.118 | 0.012 | 2.47 | 1.263 | −0.422 | −0.33 | 0.815 | 0.479 |
| P6 | 0.778 | 0 | 456.91 | 1.773 | 0.0098 | 2.41 | 1.005 | 0.234 | 0.23 | 0.592 | 0.388 |
| P7 | 0.686 | 0 | 158.02 | 0.654 | 0.0071 | 2.95 | 0.688 | 1.383 | 2.01 | 0.467 | 0.337 |
| P8 | 0.764 | 0 | 154.2 | 0.422 | 0.0071 | 3.79 | 0.623 | 2.558 | 4.1 | 0.523 | 0.425 |
| P9 | 0.788 | 0 | 211.14 | 0.62 | 0.0077 | 2.56 | 1.113 | −0.952 | −0.86 | 0.573 | 0.183 |
| P10 | 0.833 | 0 | 239.08 | 2.703 | 0.0091 | 1.66 | 0.945 | −1 | −1.06 | 0.605 | 0.173 |

where k_{\log}^m and k_{lin}^m are the monthly weighting coefficients. The numerical values of these coefficients (Table 5) have been obtained by the nonlinear least squares method.

As might be expected, k_{\log}^m and k_{lin}^m values are very close to those of the daily quantities. This feature is even more evident in terms of the ratio $k_{\text{lin}}^m/k_{\log}^m$, which confirms the distinction between the logarithmic and linear behavior of the piezometers. In Fig. 12, the simulations obtained by means of Eq. (18) are shown for all the selected piezometers, sorted again from left to right in increasing order by the ratio $k_{\text{lin}}^m/k_{\log}^m$. The distinction between the piezometers whose behavior is simulated by a logarithmic (on the left) or a linear relationship (on the right) is even clearer than that shown in Fig. 10. The inverse relation between MAD and the ratio k_{lin}/k_{\log} , and the direct relation between D_r^{m*} and the ratio k_{lin}/k_{\log} is also verified.

Both the NSE and correlation coefficient of the monthly values strongly increase with respect to the daily ones, as shown in Table 5.

Moreover, the NSE obtained by a linear relation, NSE_{lin} , is always smaller than the one obtained by using the log–linear approach. However, this decrease is not as drastic as for daily values, indicating that also a linear approach could work well on a monthly scale.

The large increase of the NSE efficiency coefficient is confirmed also by Fig. 13, where the time behavior of the simulated monthly water table elevation is plotted against the observed one. Note that the proposed relationship makes it possible to capture very well the main features of the observed values.

5. Conclusions

In this paper, a method for simulating the water table elevation of shallow unconfined aquifers is proposed based on the use of the soil moisture time series from atmospheric global datasets (reanalysis) and water table measurements by means of piezometers.

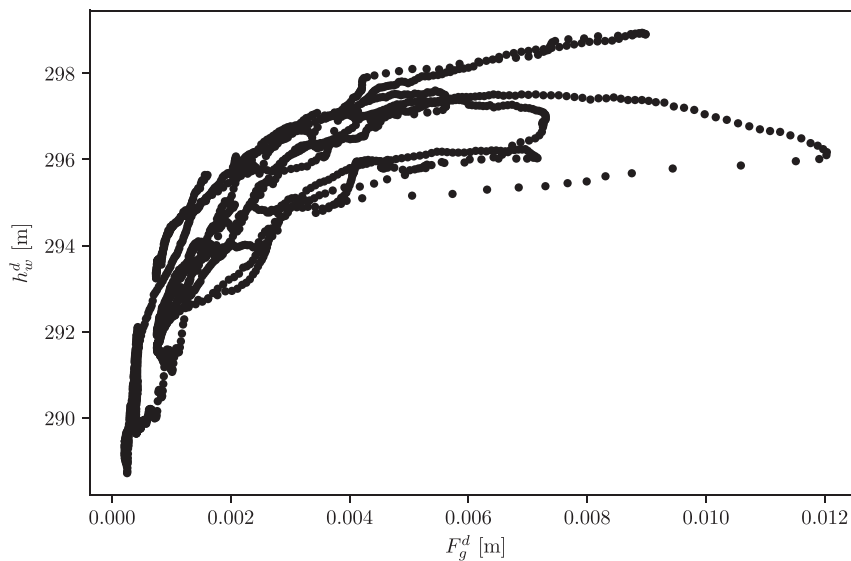


FIG. 9. Scatterplot of the daily integral fluxes F_g^d vs the daily water table elevation h_w^d for piezometer P5 from 2010 to 2015.

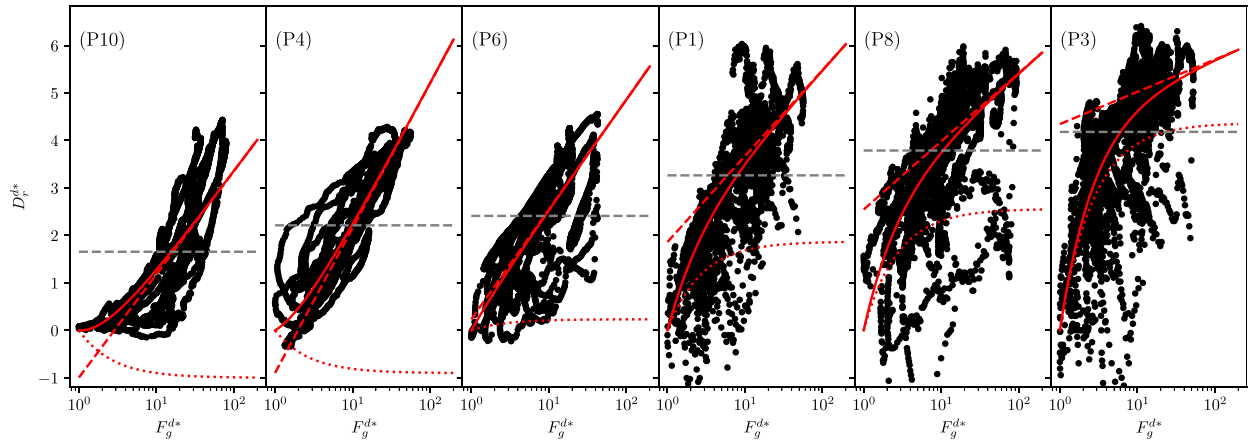


FIG. 10. Relation between the standardized relative depth D_r^{d*} and relative fluxes F_g^{d*} (in logarithmic scale): observations (black dots) and regression curve (in red) based on Eq. (14) and numerical values in Table 4; logarithmic component raised by k_{lin} (dashed red line); linear component (dotted red curve); mean standardized relative depth D_r^{d*} (dashed gray horizontal line).

The assumed groundwater recharge mechanism—that is, the vertical infiltration—is corroborated by the verified strong link between the water flux through the vadose zone, evaluated by means of the Richards equation, and the water table elevation, measured at some piezometers in the Umbria region of Italy.

In the proposed approach, the shallow water table dynamics has been simulated by means of a conceptual but physically based model in which the role of the water flux through the vadose zone is prominent. It is worth noting that a preliminary

check has shown that the water flux toward the aquifer can be properly simulated, that is, with a negligible global mass error, by considering the soil moisture data given by the H-TESSEL model used within the ERA5 reanalysis provided by ECMWF.

Two relationships for simulating the water table elevation have been derived on a daily and monthly scale that are of interest for the groundwater management. The structure of these relationships—with a linear and a logarithmic term—reflects the

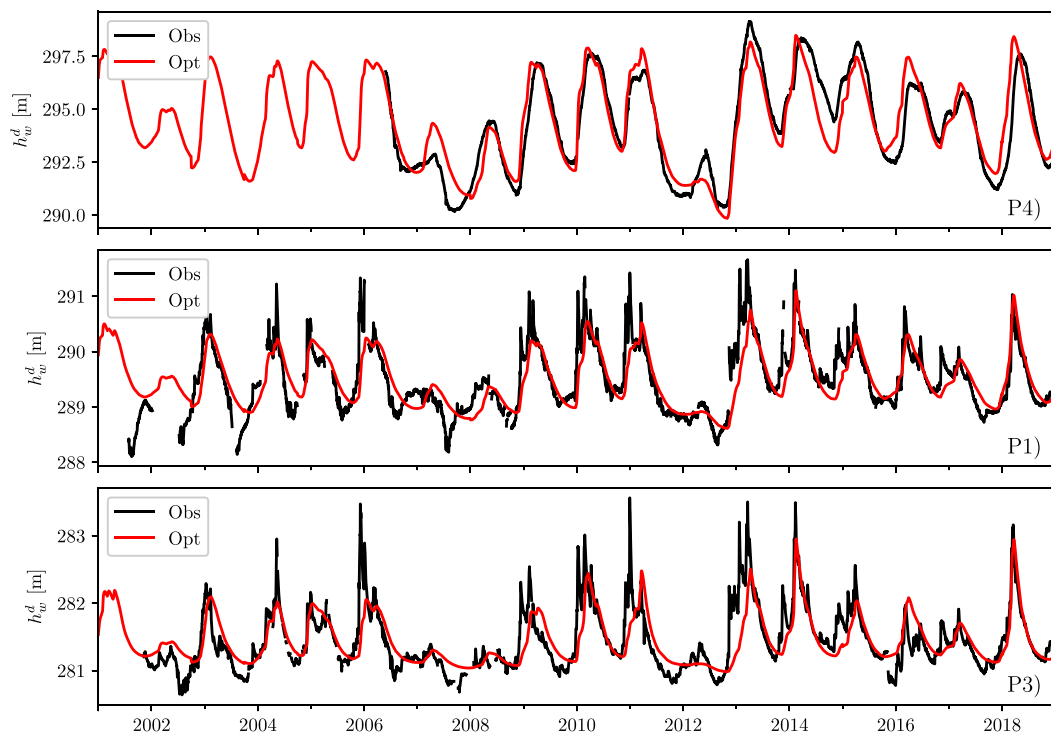


FIG. 11. Water table elevation simulation (Opt; red line) vs observed values (Obs; black line) from (top) piezometer P4, (middle) piezometer P1, and (bottom) piezometer P3.

TABLE 5. Monthly statistical properties of the water table observations and interpolation parameters. Columns are as in Table 4.

| P | R_{\max} | k_{\log}^m | k_{lin}^m | $k_{\text{lin}}^m/k_{\log}^m$ | NSE | NSE _{lin} |
|-----|------------|--------------|--------------------|-------------------------------|-------|--------------------|
| P1 | 0.818 | 0.791 | 1.863 | 2.35 | 0.984 | 0.983 |
| P2 | 0.839 | 0.509 | 3.547 | 6.97 | 0.983 | 0.972 |
| P3 | 0.77 | 0.316 | 4.312 | 13.65 | 0.981 | 0.979 |
| P4 | 0.913 | 1.31 | -0.849 | -0.65 | 0.979 | 0.979 |
| P5 | 0.91 | 1.247 | -0.366 | -0.29 | 0.98 | 0.977 |
| P6 | 0.788 | 1.008 | 0.226 | 0.22 | 0.983 | 0.979 |
| P7 | 0.692 | 0.691 | 1.375 | 1.99 | 0.994 | 0.984 |
| P8 | 0.77 | 0.629 | 2.551 | 4.06 | 0.995 | 0.983 |
| P9 | 0.788 | 1.106 | -0.934 | -0.85 | 0.986 | 0.972 |
| P10 | 0.837 | 0.946 | -1 | -1.06 | 0.981 | 0.98 |

fact that, according to the considered time scale, a different regime characterizes the shallow aquifers with the infiltration being the main mechanism of recharge.

In the writers' opinion, the good quality in terms of the water table elevation of the simulations provided by the proposed method encourages extending in the future the analysis to further areas characterized by different climate conditions and soils. In fact, soil moisture data from reanalysis are available throughout the world, and water table measurements, easy to execute, are the most popular method for groundwater control.

Acknowledgments. This research has been funded by Italian Ministry of University and Research (MIUR) and University of Perugia within the program Dipartimenti di Eccellenza 2018–2022. The support of Mr. M. Nucci of ARPA Umbria for providing the water table measurements is highly appreciated.

Data availability statement. The water table data that support the findings of this study are available upon request

(<https://apps.arpa.umbria.it/acqua/contenuto/Livelli-Di-Falda>), and the ERA5 reanalysis data are freely available on the Copernicus Climate Data Store (C3S 2017).

APPENDIX

Soil Water Budget

In H-TESSSEL (ECMWF 2016b; Balsamo et al. 2009), the water flux through the unsaturated zone is simulated by means of Eq. (1).

As mentioned, according to the proposed approach, the root uptake term is neglected since in the two deepest layers it is much smaller than the drainage. As a consequence, the discretized-in-time form of Eq. (1) is

$$\frac{\theta^{t+1} - \theta^t}{\Delta t} = -\frac{\partial}{\partial z} \left[\lambda(\theta^t) \frac{\partial \hat{\theta}}{\partial z} - \gamma(\theta^t) \right]. \quad (\text{A1})$$

On the RHS, the coefficients λ and γ are assumed as a function of the soil moisture at the current time step θ^t [and then in the following: $\lambda^t = \lambda(\theta^t)$ and $\gamma^t = \gamma(\theta^t)$], whereas, to compute the vertical gradient $\partial \hat{\theta} / \partial z$, the soil moisture value $\hat{\theta}$, given by

$$\hat{\theta} = \beta \theta^{t+1} + (1 - \beta) \theta^t, \quad (\text{A2})$$

is used. According to the value of β , the diffusion term in Eq. (1), $-\partial / \partial z (\lambda \partial \hat{\theta} / \partial z)$, is solved by an implicit ($\beta = 1$), explicit ($\beta = 0$), or semi-implicit interpolation ($\beta = 0.5$). Note that H-TESSSEL uses the implicit interpolation and then $\hat{\theta} = \theta^{t+1}$. In this paper, we kept β varying, in order to also analyze the sensitivity to the used discretization approach. In the case of the implicit interpolation, Eq. (A2) can be written as

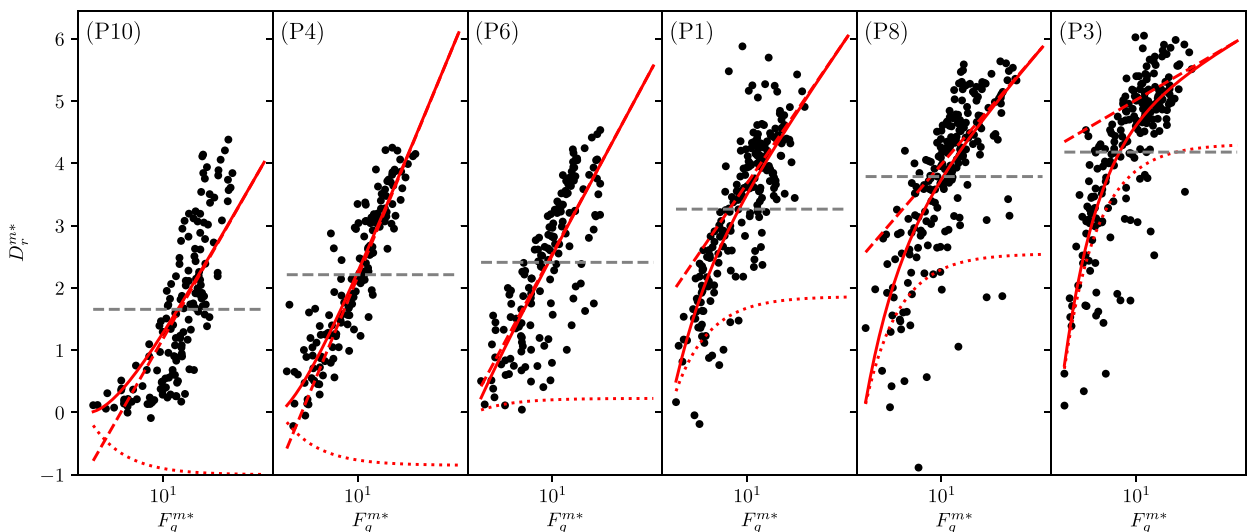


FIG. 12. Relation between monthly standardized relative depth D_r^{m*} and monthly mean relative fluxes, F_g^{m*} (in logarithmic scale): observations (black dots), regression curve (in red) based on Eq. (14) and numerical values in Table 4; logarithmic component raised by k_{lin}^m (dashed red line), linear component (dotted red curve); mean standardized relative depth D_r^{d*} (dashed gray horizontal line).

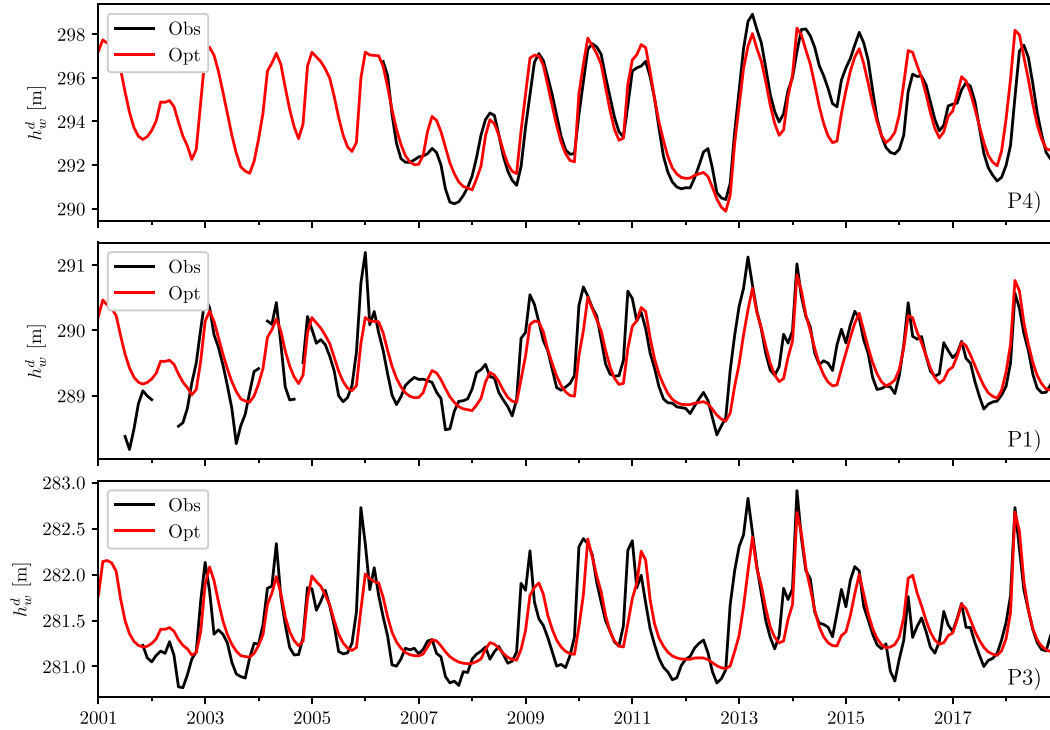


FIG. 13. Monthly water table elevation simulated (Opt; red line) vs observed values (Obs; black line) for (top) piezometer P4, (middle) piezometer P1, and (bottom) piezometer P3.

$$\frac{\theta^{t+1} - \theta^t}{\Delta t} = - \underbrace{\frac{\partial}{\partial z} \left(\lambda^t \frac{\partial \theta^{t+1}}{\partial z} - \gamma^t \right)}_{\text{Interlayer fluxes}}. \quad (\text{A3})$$

Equation (A3) is solved over a soil column that is discretized into four layers (Fig. 3) reaching a depth of 2.89 m. As mentioned, interlayer fluxes F_w , on the RHS of Eq. (A3), are evaluated at the layer interfaces, whereas θ , and then γ and λ , are defined at the layer center. Hydraulic conductivities and diffusivities at the layer interfaces are defined in H-TESEL by using the largest soil moisture value between the two adjacent layers (max $_{\theta}$ approximation):

$$\lambda_{k+1/2} = \lambda[\max(\theta_k, \theta_{k+1})] \quad \text{and} \quad \gamma_{k+1/2} = \gamma[\max(\theta_k, \theta_{k+1})]. \quad (\text{A4})$$

As mentioned, in this paper, for the sake of completeness, we have compared the above approach with an alternative one (Fig. 5) in which the interlayer properties are evaluated by using the soil moisture value at the interface as obtained by a linear interpolation of the soil moisture profile (lin $_{\theta}$ approximation):

$$\lambda_{k+1/2} = \lambda[\text{lin}(\theta_k, \theta_{k+1})] \quad \text{and} \quad \gamma_{k+1/2} = \gamma[\text{lin}(\theta_k, \theta_{k+1})]. \quad (\text{A5})$$

Based on the above considerations, the discretized-in-space form of Eq. (A3) is

$$\frac{\theta_k^{t+1} - \theta_k^t}{\Delta t} = - \left[\left(\lambda_{k+1/2}^t \frac{\theta_{k+1}^{t+1} - \theta_k^{t+1}}{\Delta z_{k+1/2}} - \gamma_{k+1/2}^t \right) - \left(\lambda_{k-1/2}^t \frac{\theta_k^{t+1} - \theta_{k-1}^{t+1}}{\Delta z_{k-1/2}} - \gamma_{k-1/2}^t \right) \right] \frac{1}{\Delta z_k}, \quad (\text{A6})$$

where $\Delta z_k = z_{k+1/2} - z_{k-1/2}$ and $\Delta z_{k+1/2} = z_{k+1} - z_k$ are the distance between two layer interfaces and centers, respectively. By using the following notation for discrete water fluxes:

$$F_{k+1/2}^{t+1} = \lambda_{k+1/2}^t \left(\frac{\theta_{k+1}^{t+1} - \theta_k^{t+1}}{\Delta z_{k+1/2}} \right) - \gamma_{k+1/2}^t, \quad (\text{A7})$$

we can reshape Eq. (A6) into

$$\frac{\theta_k^{t+1} - \theta_k^t}{\Delta t} = - \frac{F_{k+1/2}^{t+1} - F_{k-1/2}^{t+1}}{\Delta z_k}. \quad (\text{A8})$$

If we define the water integral over layer k at time t as

$$W_k^t = \int_{z_{k-1/2}}^{z_{k+1/2}} \theta^t(z) dz = \theta_k^t \Delta z_k, \quad (\text{A9})$$

we get

$$W_k^{t+1} - W_k^t = (F_{k+1/2}^{t+1} - F_{k-1/2}^{t+1}) \Delta t. \quad (\text{A10})$$

At a daily scale, the percentage error in the water budget is defined as

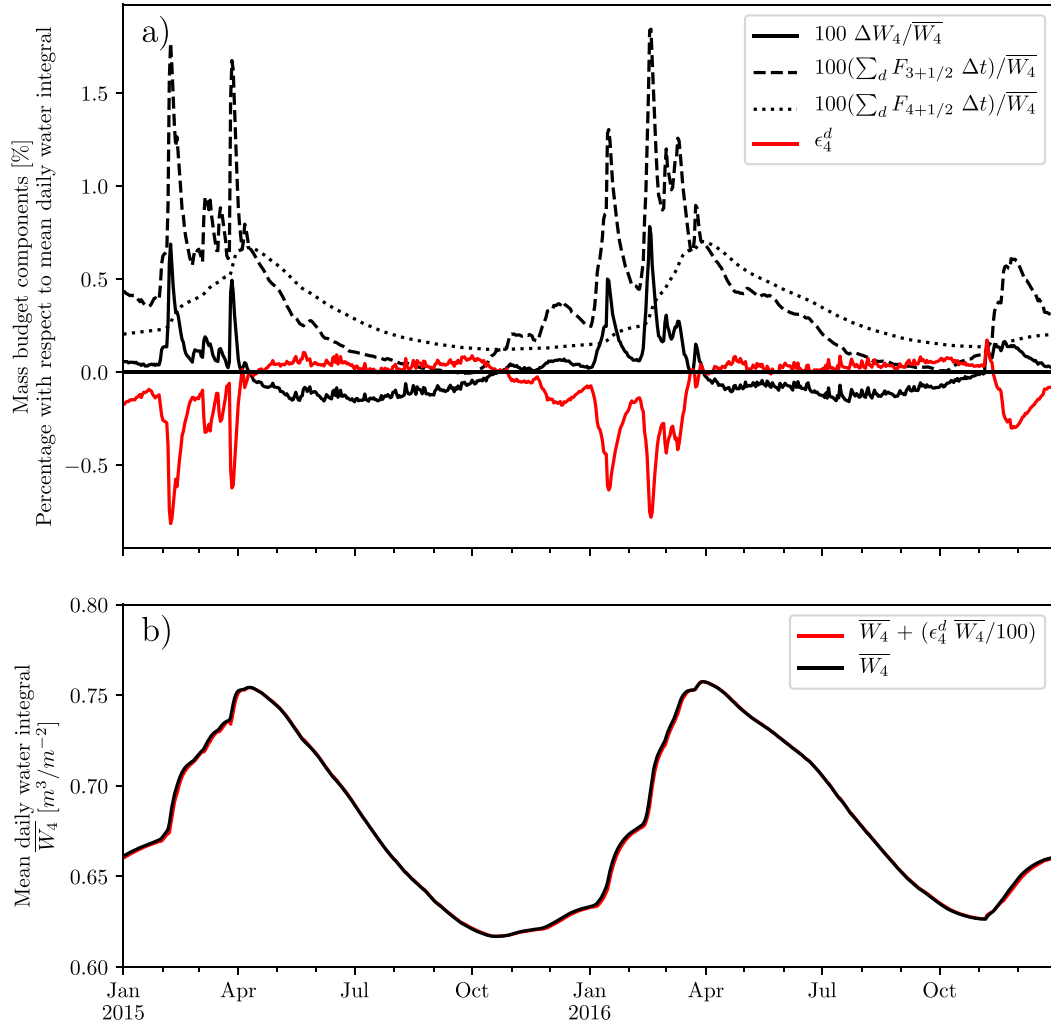


FIG. A1. (a) Water mass budget components and error evaluated for the soil layer 4 at grid point G2 in the Umbria region from 2015 to 2016 for the semi-implicit case ($\beta = 0.5$). Each component has been expressed in percentage with respect to the mean daily water integral \bar{W}_4 . The different components are the soil moisture daily integral change, $\Delta W_4 = \sum_{t \in d} (W_k^{t+1} - W_k^t)$ (solid black line); incoming accumulated water fluxes (dashed black line); outgoing accumulated water fluxes (dotted black line); mass error percentage (red line). (b) Comparison between the mean daily water integral time evolution (black solid line) and the same integral incremented by the mass error (red solid line). A large zoom is needed to visualize the difference between the two curves, confirming the negligible influence of mass error on the mean water integral.

$$\epsilon_k^d = \frac{100}{\bar{W}_k} \left[\sum_{t \in d} (W_k^{t+1} - W_k^t) - \sum_{t \in d} (F_{k-1/2}^{t+1} - F_{k+1/2}^{t+1}) \Delta t \right], \quad (\text{A11})$$

where \bar{W}_k is the daily mean water integral [note that Eq. (A11) is equal to Eq. (12) for the implicit case ($\beta = 1$)]. The daily scale has been selected since that groundwater management usually work at large temporal scales according to the large inertia of water bodies. Changes of the water integral and water table elevation are almost negligible at temporal resolutions smaller than the daily one.

According to Eq. (A11), the daily percentage error, ϵ_k^d , is a measure of the relative magnitude of the mass error to the total amount of water (per unit surface) contained in the k th soil

layer. We did not use water fluxes for the error adimensionalization because they very often get closer to zero.

In Fig. A1a, the mass budget with all the components is shown for the grid point G2 in the semi-implicit case ($\beta = 0.5$).

It can be noticed that the magnitude of all the components is very small, when compared with the average water integral \bar{W}_4 (Fig. A1). Usually, the mass error is null from April to October, when incoming fluxes are smaller and the water integral decreases in time (negative values of ΔW_4 in Fig. A1a). However, especially during the winter season, when the soil moisture increases (Fig. 6a), and so does the water integral with positive values of ΔW_4 (Fig. A1b), the mass error increases. Despite the above shortcomings, the magnitude of the mass error is very

small when compared with the mean water integral and it never exceeds, as absolute value, 2%, as indicated by the boxplots in Fig. 7. This feature is confirmed by Fig. A1b plots where the difference between the mean water integral (solid black line) and the same integral incremented by the mass error (solid red line) can be appreciated only by zooming significantly.

REFERENCES

- Albergel, C., P. De Rosnay, G. Balsamo, L. Isaksen, and J. Muñoz-Sabater, 2012: Soil moisture analyses at ECMWF: Evaluation using global ground-based in situ observations. *J. Hydrometeorol.*, **13**, 1442–1460, <https://doi.org/10.1175/JHM-D-11-0107.1>.
- , E. Dutra, S. Munier, J.-C. Calvet, J. Muñoz-Sabater, P. de Rosnay, and G. Balsamo, 2018: ERA-5 and ERA-Interim driven ISBA land surface model simulations: Which one performs better? *Hydrol. Earth Syst. Sci.*, **22**, 3515–3532, <https://doi.org/10.5194/hess-22-3515-2018>.
- ARPA, 2008: Monitoring recording of the groundwater levels in Umbria (in Italian). Umbria ARPA Tech. Rep., 88 pp., <http://www.arpa.umbria.it/articoli/il-monitoraggio-continuo-dei-livelli-di-falda-in-u>.
- Balsamo, G., A. Beljaars, K. Scipal, P. Viterbo, B. van den Hurk, M. Hirschi, and A. K. Betts, 2009: A revised hydrology for the ECMWF model: Verification from field site to terrestrial water storage and impact in the Integrated Forecast System. *J. Hydrometeorol.*, **10**, 623–643, <https://doi.org/10.1175/2008JHM1068.1>.
- Bardsley, T., A. Wood, M. Hobbins, T. Kirkham, L. Briefer, J. Niermeyer, and S. Burian, 2013: Planning for an uncertain future: Climate change sensitivity assessment toward adaptation planning for public water supply. *Earth Interact.*, **17**, <https://doi.org/10.1175/2012EI000501.1>.
- Bjerklie, D. M., T. J. Trombley, and R. J. Viger, 2011: Simulations of historical and future trends in snowfall and groundwater recharge for basins draining to Long Island Sound. *Earth Interact.*, **15**, <https://doi.org/10.1175/2011EI374.1>.
- Bonta, J. V., and M. Müller, 1999: Evaluation of the Glugla method for estimating evapotranspiration and groundwater recharge. *Hydrol. Sci. J.*, **44**, 743–761, <https://doi.org/10.1080/02626669909492271>.
- Brunone, B., M. Ferrante, N. Romano, and A. Santini, 2003: Numerical simulations of one-dimensional infiltration into layered soils with the Richards equation using different estimates of the interlayer conductivity. *Vadose Zone J.*, **2**, 193–200, <https://doi.org/10.2136/vzj2003.1930>.
- C3S, 2017: ERA5: Fifth generation of ECMWF atmospheric reanalyses of the global climate. Copernicus Climate Change Service, accessed 20 June 2020, <https://cds.climate.copernicus.eu#!/home>.
- Cameron, D. R., 1978: Variability of soil water retention curves and predicted hydraulic conductivities on a small plot. *Soil Sci.*, **126**, 364–371, <https://doi.org/10.1097/00010694-197812000-00009>.
- Celia, M. A., and P. Binning, 1990: A mass conservative numerical solution for two-phase flow in porous media with application to unsaturated flow. *Water Resour. Res.*, **28**, 2819–2828, <https://doi.org/10.1029/92WR01488>.
- Cerlini, P. B., S. Meniconi, and B. Brunone, 2017: Groundwater supply and climate change management by means of global atmospheric datasets. Preliminary results. *Procedia Eng.*, **186**, 420–427, <https://doi.org/10.1016/j.proeng.2017.03.245>.
- Chinnasamy, P., J. A. Hubbard, and G. Agoramoorthy, 2013: Using remote sensing data to improve groundwater supply estimations in Gujarat, India. *Earth Interact.*, **17**, <https://doi.org/10.1175/2012EI000456.1>.
- Dee, D. P., and Coauthors, 2011: The ERA-Interim reanalysis: Configuration and performance of the data assimilation system. *Quart. J. Roy. Meteor. Soc.*, **137**, 553–597, <https://doi.org/10.1002/qj.828>.
- de Rosnay, P., M. Drusch, D. Vasiljevic, G. Balsamo, C. Albergel, and L. Isaksen, 2013: A simplified extended Kalman filter for the global operational soil moisture analysis at ECMWF. *Quart. J. Roy. Meteor. Soc.*, **139**, 1199–1213, <https://doi.org/10.1002/qj.2023>.
- , G. Balsamo, C. Albergel, J. Muñoz-Sabater, and L. Isaksen, 2014: Initialisation of land surface variables for numerical weather prediction. *Surv. Geophys.*, **35**, 607–621, <https://doi.org/10.1007/s10712-012-9207-x>.
- de Vries, J. J., and I. Simmers, 2002: Groundwater recharge: An overview of processes and challenges. *Hydrogeol. J.*, **10**, 5–17, <https://doi.org/10.1007/s10040-001-0171-7>.
- Drusch, M., and P. Viterbo, 2007: Assimilation of screen-level variables in ECMWF's Integrated Forecast System: A study on the impact on the forecast quality and analyzed soil moisture. *Mon. Wea. Rev.*, **135**, 300–314, <https://doi.org/10.1175/MWR3309.1>.
- ECMWF, 2016a: Part II: Data assimilation. IFS Doc. CY41R2, 103 pp., <https://www.ecmwf.int/node/16666>.
- , 2016b: Part IV: Physical processes. IFS Doc. CY41R2, 213 pp., <https://www.ecmwf.int/node/16648>.
- Fairbairn, D., P. de Rosnay, and P. A. Browne, 2019: The new stand-alone surface analysis at ECMWF: Implications for land–atmosphere DA coupling. *J. Hydrometeorol.*, **20**, 2023–2042, <https://doi.org/10.1175/JHM-D-19-0074.1>.
- FAO, 2003: Digital soil map of the world and derived soil properties. Land and Water Development Division, accessed 10 March 2020, <http://www.fao.org/geonetwork/srv/en/main.home?uuid=446ed430-8383-11db-b9b2-000d939bc5d8>.
- Gogolev, M. I., 2002: Assessing groundwater recharge with two unsaturated zone modeling technologies. *Environ. Geol.*, **42**, 248–258, <https://doi.org/10.1007/s00254-001-0494-7>.
- Healy, R. W., and P. G. Cook, 2002: Using groundwater levels to estimate recharge. *Hydrogeol. J.*, **10**, 91–109, <https://doi.org/10.1007/s10040-001-0178-0>.
- Hersbach, H., and Coauthors, 2020: The ERA5 global reanalysis. *Quart. J. Roy. Meteor. Soc.*, **146**, 1999–2049, <https://doi.org/10.1002/qj.3803>.
- Huet, M., R. Chesnaux, M.-A. Boucher, and C. Poirier, 2016: Comparing various approaches for assessing groundwater recharge at a regional scale in the Canadian Shield. *Hydrol. Sci. J.*, **61**, 2267–2283, <https://doi.org/10.1080/02626667.2015.1106544>.
- Iqbal, J., J. A. Thomasson, J. N. Jenkins, P. R. Owens, and F. D. Whisler, 2005: Spatial variability analysis of soil physical properties of alluvial soils. *Soil. Sci. Soc. Amer. J.*, **69**, 1338–1350, <https://doi.org/10.2136/sssaj2004.0154>.
- Jinno, K., A. Tsutsumi, O. Alkaeed, S. Saita, and R. Berndtsson, 2009: Effects of land-use change on groundwater recharge model parameters. *Hydrol. Sci. J.*, **54**, 300–315, <https://doi.org/10.1623/hysj.54.2.300>.
- Karamouz, M., E. Goharian, and S. Nazif, 2013: Reliability assessment of the water supply systems under uncertain future extreme climate conditions. *Earth Interact.*, **17**, <https://doi.org/10.1175/2012EI000503.1>.
- Leach, B., 1982: The development of a groundwater recharge model for Hong Kong. *Hydrol. Sci. J.*, **27**, 469–491, <https://doi.org/10.1080/02626668209491127>.

- Lerner, D. N., 1997: Groundwater recharge. *Geochemical Processes, Weathering and Groundwater Recharge in Catchments*, O. M. Saether and P. de Caritat, Eds., AA Balkema, 109–150.
- , A. S. Issar, and I. Simmers, 1990: *Groundwater Recharge. A Guide to Understanding and Estimating Natural Recharge*. IAH International Contributions to Hydrogeology, Vol. 8, Heinz Heise, 345 pp.
- Omorinbola, E. O., 1986: Empirical equations of groundwater recharge patterns. *Hydrol. Sci. J.*, **31**, 1–11, <https://doi.org/10.1080/02626668609491023>.
- Seibert, J., A. Rodhe, and K. Bishop, 2003: Simulating interactions between saturated and unsaturated storage in a conceptual runoff model. *Hydrol. Processes*, **17**, 379–390, <https://doi.org/10.1002/hyp.1130>.
- Sekhar, M., M. Shindekar, S. K. Tomer, and P. Goswami, 2013: Modeling the vulnerability of an urban groundwater system due to the combined impacts of climate change and management scenarios. *Earth Interact.*, **17**, <https://doi.org/10.1175/2012EI000499.1>.
- Touhami, I., J. M. Andreu, E. Chirino, J. R. Sánchez, A. Pulido-Bosch, P. Martínez-Santos, H. Moutahir, and J. Bellot, 2014: Comparative performance of soil water balance models in computing semi-arid aquifer recharge. *Hydrol. Sci. J.*, **59**, 193–203, <https://doi.org/10.1080/02626667.2013.802094>.
- van Genuchten, M. T., 1980: A closed-form equation for predicting the hydraulic conductivity of unsaturated soils 1. *Soil. Sci. Soc. Amer. J.*, **44**, 892–898, <https://doi.org/10.2136/sssaj1980.03615995004400050002x>.
- Varni, M., R. Comas, P. Weinzettel, and S. Dietrich, 2013: Application of the water table fluctuation method to characterize groundwater recharge in the Pampa plain, Argentina. *Hydrol. Sci. J.*, **58**, 1445–1455, <https://doi.org/10.1080/02626667.2013.833663>.
- Yeh, P. J., and E. A. Eltahir, 2005: Representation of water table dynamics in a land surface scheme. Part I: Model development. *J. Climate*, **18**, 1861–1880, <https://doi.org/10.1175/JCLI3330.1>.
- Yin, L., G. Hu, J. Huang, D. Wen, J. Dong, X. Wang, and H. Li, 2011: Groundwater-recharge estimation in the Ordos Plateau, China: Comparison of methods. *Hydrogeol. J.*, **19**, 1563–1575, <https://doi.org/10.1007/s10040-011-0777-3>.