Model signatures and aridity indices enhance the accuracy of water balance estimations in a data-scarce Eastern Mediterranean catchment

A. Gunkel a, S. Shafeed b, A. Hartmann a, T. Wagener c, J. Lange a

a Chair of Hydrology, Albert-Ludwigs-University of Freiburg, 79098 Freiburg, Germany
b Water and Environmental Studies Institute, An-Najah National University, P.O. Box 7, Nablus, Palestine
c Department of Civil Engineering, University of Bristol, Queen’s Building, University Walk, Bristol BS8 1TR, UK

Article history:
Received 6 March 2015
Received in revised form 9 August 2015
Accepted 11 August 2015
Available online 2 September 2015

Keywords:
Water balance
Semi-arid areas
TRAIN-ZIN model
Hydrological signatures
Aridity indicators

Abstract

Study region: Wadi Faria catchment, Palestine.
Study focus: The upper part of the Faria catchment (139 km²) is a typical semi-arid karst catchment in the Eastern Mediterranean, where, up to recently, data availability has hindered the accurate assessment of renewable water resources. Newly available six-year time-series of rainfall and runoff data, combined with thorough field campaigns, enabled the application of the distributed TRAIN-ZIN watershed model. The model was constrained using seven hydrological signatures derived from the available time-series.

New hydrological insights for the region: We found that the mean annual actual evapotranspiration was about 70% of precipitation, recharge was about 30% and natural runoff (excluding baseflow) 1%. Aggregated model results also supported aridity indicators that show the presence of Infiltration Excess (Hortonian) Overland Flow, as well as the importance of indirect groundwater recharge and evaporation from soil during dry months. In total, maximum annual water availability was of the same order of magnitude as actual demand estimates (23 MCM). However, high spatial and inter-annual variability, and the presence of karst features suggest that water resources in the region are highly vulnerable.

© 2015 The Authors. Published by Elsevier B.V. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

1. Introduction

Prevailing semi-arid climate conditions and increasing water consumption in the Eastern Mediterranean threaten renewable water resources, which are additionally endangered by climate change (Alpert et al., 2008). As in many semi-arid areas of the world, scarcity of reliable information about water resources impedes accurate water resources planning and management (Wheater, 2002).

In the Eastern Mediterranean, catchments are characterized by high rainfall intensities and a high variability of rainfall in space and time, which is reflected in other elements of the water balance. Moreover, the area is characterized by a strong climate gradient, for example between the Negev desert and the Galilee mountains (Cerdà, 1998) with a varying

* Corresponding author.
E-mail addresses: anne.gunkel@hydrology.uni-freiburg.de (A. Gunkel), sshafeed@najah.edu (S. Shafeed), andreas.hartmann@hydrology.uni-freiburg.de (A. Hartmann), thorsten.wagener@bristol.ac.uk (T. Wagener), jens.lange@hydrology.uni-freiburg.de (J. Lange).
degree of aridity. This degree of aridity influences several hydrological processes; higher aridity in conjunction with the associated high rainfall intensities and low vegetation cover increases the proportion of Infiltration Excess (i.e. Hortonian) compared to Saturation Excess Overland Flow (Cerdà, 1998; Kirby, 2001). It also enhances indirect groundwater recharge (i.e. through transmission losses during surface runoff) compared to direct recharge (Gee and Hillel, 1988; Simmers, 1997). Soil evaporation increases compared to transpiration (Balugani et al., 2014; Reynolds et al., 2000) in more arid areas where evapotranspiration is generally water-limited, i.e. potential is considerably higher than actual evapotranspiration.

The Eastern Mediterranean is further subject to a wide-spread karstification of carbonate rock that results in strongly heterogeneous flow and storage behaviour (Bakalowicz, 2005; Hughes et al., 2008). Water resources assessment in Eastern Mediterranean karst areas is consequently challenging in at least two aspects: the high variability of hydrological processes under Mediterranean climate conditions superimposed on the complex hydrological conditions of a karst area.

Hydrological models are frequently applied tool to interpolate in space and time between existing data to provide missing information. They are able to do so as long as they reflect main catchment functions, i.e. partitioning, storage and release of water (Wagener et al., 2007). Approaches developed for semi-arid areas cover a wide range of complexities and philosophies, from lumped to distributed models and from monthly water balance models to event-based runoff hydrograph models (see for example Jakeman et al., 1990; Lange et al., 1999; McIntyre et al., 2007; Pitman, 1973; Smith et al., 1995; Yatim (2008); Ye et al., 1997).

Modelling in the Eastern Mediterranean is frequently challenging because of a low amount of data quality and quantity (Cudennec et al., 2007) that affects the model evaluation as well. Statistical goodness-of-fit measures as the traditional approach have several practical and theoretical shortcomings under such circumstances. First and foremost they do not guarantee that the model adequately reflects catchment functioning (e.g. Gupta et al., 2008; Westerberg et al., 2011; Yilmaz et al., 2005). Alternatively, indices or time series of the response behaviour of a catchment at a given time-scale can be considered (Gupta et al., 2008). These so called signatures can serve as constraints for the selection of model parameters or good model runs, if they are adapted to local hydrology and modelling purposes. Commonly used examples include flow duration curves, runoff ratios, aridity indices or measures of discharge timing (e.g. Kapangaziwi et al., 2012).

During the last decade, various studies addressed different aspects of the hydrology of the Eastern Mediterranean Karst areas. Groundwater recharge mechanisms were studied in the Nahal Oren watershed, Israel, by monitoring cave drips under natural rainfall conditions (Arbel et al., 2010) and with a sprinkling experiment (Lange et al., 2010) or for the Sif Cave, 100 km south (Sheffer et al., 2011). Ries et al. (2015) investigated recharge quantities and soil dynamics near Jerusalem. Large scale modelling of recharge aspects was undertaken for the Western Mountain basin by Sheffer et al. (2010).

Several studies modelled the hydrological conditions in the Upper Jordan basin, for example with the Hydrological Model for Karst Environment (HYMKE) (Rimmer and Salinger, 2006), or climate change impacts in the same basin (e.g. Samuels et al., 2010, 2009; Smiatek et al., 2011). The water balance of the Lower Jordan River basin was modelled with the TRAIN-ZIN model (Gunkel and Lange, 2012).

For two of the major springs of the Jordan River, Hartmann et al. (2013a) developed and tested model realism of four hydrological model structures. They also modelled the actual runoff and climate change scenarios for the main spring in the Wadi Faria basin (Hartmann et al., 2012) and five different karst systems in Europe and the Middle East, two of them in the Eastern Mediterranean (Hartmann et al., 2013b). Another modelling of spring and Karst aspects was conducted for the Auja spring (Schmidt et al., 2014). The groundwater regime in the western Dead Sea escarpment was for example addressed by Gräbe et al. (2013).

However, relatively little is known about the water balance of the basins lining the western side of the Lower Jordan River. Moreover, the aforementioned studies are either hydrological, i.e. focus on the surface water balance, or hydrogeological, i.e. deal with an aquifer water balance. None of them simulates generation of overland flow and percolation to groundwater simultaneously.

The main purpose of this study is to obtain holistic knowledge of the entire water balance of a typical Eastern Mediterranean karst catchment. We chose the Faria catchment located in the West Bank, Palestine, and applied a distributed hydrological model that was constrained with a signature approach. We used the recently developed TRAIN-ZIN model (Gunkel and Lange, 2012) that can represent the relevant hydrological processes of the study region at an adequate temporal and spatial scale. The model is operated to study the surface water-balance including land surface processes like evaporation and surface runoff generation. Based on seven runoff based signatures that can be computed with the available data we are able to select the most realistic model runs. Percolation output from TRAIN-ZIN is then used as input to a parsimonious lumped karst model ensemble at the main spring in the basin, as developed in a previous simulation study (Hartmann et al., 2012). Additionally, output from TRAIN-ZIN is compared to empirical precipitation recharge relationships. Evaluated model results allow us to analyse the water balance and to apply three aridity indicators, the proportion of Infiltration Excess runoff, indirect percolation and soil evaporation.

2. Study area and available data

The Faria catchment is one of the major arteries of water draining into the Lower Jordan River downstream of Lake Kinneret (also known as Lake Tiberias). It covers an area of approximately 320 km² of the West Bank, Palestine (Fig. 1). Surface runoff is mainly generated in the wetter upper part of the catchment, which is also more relevant in terms of population, water
management and data availability (Shadeed, 2008). Therefore, we chose the two main subbasins that make up the upper part, i.e. Al-Badan (AB; 83 km²) and Al-Faria (AF; 56 km²), as study area (Fig. 1).

The main land use types in the Upper Faria (68–900 m above mean sea level) are agriculture, hillslopes with natural grassland or olive plantations and about 9% built-up areas (Shadeed and Lange, 2010). Its dominant soils are Terra Rossas, Brown Rendzinas and Alluvial Grumusols. Geologically, the area is characterized by carbonate lithology (Flexer, 1968). Dolomites and limestones of Cenomanian-Turonian age (Judea group, Ajlun Series) are overlain by a chalky unit of Senonian to Paleocene age (Mount Scopus group) that separates a lower Cenomanian aquifer from an upper aquifer consisting of limestone, chalk and chert strata of Eocene age (Jenin subseries, Avedat group) (Flexer, 1968; Ghanem, 1999; Guttman, 2000). Karst features and fractures can be found in limestones and dolomites of both, Eocene and Cretaceous strata (Ghanem, 1999).

Climate is semi-arid and highly variable, characterized by mild rainy winters and dry, hot summers. The mean annual temperature is 18 °C. Potential evaporation is particularly high in the Mediterranean summer from April to October. Meteorological data (daily values for average temperature, wind speed, relative humidity and global radiation) for the model was obtained from the Nablus Meteorological Station (Palestinian Meteorological Department, PALMET).

Rainfall events predominantly occur in autumn and winter accounting for 80% of the total annual precipitation (Shadeed and Almasri, 2007). Rainfall data for the six simulated years (2004/05–2009/10) was available from four newly installed tipping bucket rain gauges (Tubas, Tamhum, Talluza, Salim; Fig. 1) with a rainfall depth of 0.2 mm per tip. Based on data from these stations, spatial rainfall patterns were calculated by inverse distance weighting for time steps of 5 min. In the interpolation routine, we considered an elevation gradient of 10%/100 m, as calculated from observed data. Mean annual catchment rainfall over the six years was 516 mm for AB and 497 mm for AF.

Runoff was measured with two Parshall Flumes that were installed at the outlets of AB and AF (Fig. 1) in 2004 as continuous 10 min record for the rainy seasons. These ephemeral riverbeds contain occasionally some baseflow originating from karst springs or wastewater. At the Faria spring, monthly discharge observations were available for the first three hydrological years (2004/05–2006/07). The spring with its long term mean annual discharge of 5.2 MCM (Shadeed, 2008) has dried out due to overpumping since 2007 (Hartmann et al., 2012).

![Fig. 1. Location of the Faria Catchment and the two upper subbasins.](image-url)
3. Methods

3.1. Models

TRAIN-ZIN is a rainfall-runoff model that was developed for semi-arid catchments and represents all relevant hydrological processes in this climate with an adequate temporal and spatial resolution. The model combines the concepts of two hydrological models, the TRAIN model by Menzel (1997) and the ZIN model by Lange et al. (1999). Its output comprises runoff hydrographs for each sub-catchment, as well as daily estimates of all water balance components (see Fig. 2 for a schematic overview and Gunkel and Lange (2012) for a further description of the model). It is used to simulate natural wadi flow excluding human abstractions or baseflow from springs and wastewater. Former applications in the Middle East include the Lower Jordan River Basin (Gunkel and Lange, 2012), Wadi Kafrein (Alkhoury, 2011), Mount Carmel (Kohn, 2008) and the application of an early version of the model to three rainy seasons in the Faria catchment (Shadeed and Lange, 2010; Shadeed, 2008).

In a previous study, the outflow of Faria spring was modelled by the MOSST model using climate change scenarios (Hartmann et al., 2012). MOSST simulates spring outflow of the karstic Faria spring using an ensemble of five lumped hydrological models that consider karst hydrological processes in various ways and complexities (Fig. 2). They represent the vadose and the groundwater zone through different combinations of linear reservoirs with three to six parameters. Their structure was chosen to be parsimonious due to the limited data for model calibration and evaluation. In this study, MOSST was driven by percolation output from TRAIN-ZIN.

3.2. Calibration strategy and model parameterisation

In TRAIN-ZIN, evapotranspiration was modelled at daily time steps, runoff generation processes at five minutes and channel routing at one minute time steps. Spatial data layers were interpolated and re-projected to congruent grids of 50 × 50 m² (55,453 grid cells). Runoff generated in the grid cells was aggregated to 257 sub-catchments and routed through 257 channel segments. As the present study aims at simulating direct runoff from rainfall, event flow was identified through baseflow separation applying a widely used numerical filter method (Lyne and Hollick, 1979).

Six hydrological response units (Fig. 3a, Table 1) were defined based on land cover classes derived from satellite images (Google Maps, 2011) and field surveys (Shadeed, 2008), a soil map of Israel (Schacht et al., 2011) and a geological map (Sneh et al., 1998) (Fig. 3b). Parameter values required for calculating evapotranspiration (i.e. albedo, leaf area index (LAI) and vegetation canopy height) were adapted for vegetation periods and followed literature values (see Gunkel and Lange, 2012 for further details).
Based on its physical nature, TRAIN-ZIN was calibrated by attributing parameters from field observations followed by manual adjustments (Lange et al., 1999). Soil parameters were the main focus of our manual model calibration as they control the separation of rainfall into surface runoff and vertical percolation; while most other parameters were kept unchanged. Parameters were chosen to similarly represent soil and epikarst (Table 1) and are summarized in Table 2. In total, 22 parameter combinations were realized as different model runs, considering the long computing times of the model. Parameters of these runs vary within the ranges given in Table 2 (single values indicate parameter values kept constant). A particular case is the saturated conductivity: departing from the values in Table 2, it is set to 0.1 for all grid cells with the Mount Scopus group with marl and clay strata (Guttmann, 2000) (Fig. 3b) as dominant geology.

Runoff concentration from grid cells to the small sub-catchments was parameterized by a Unit Hydrograph approach applying a Fisher–Tippet distribution for extreme values (Fisher and Tippett, 1928). For the calculation of channel flow and transmission losses, parameters for the 257 channel segments (average length: 783 m) were derived from field surveys, aerial photographs and GIS tools (Shadeed, 2008).

All parameters of the MOSST model were kept as published in Hartmann et al. (2012). The critical step for interaction between the models was the determination of the contributing recharge area of the Faria spring. Results from Hartmann et al. (2012) suggested a recharge area of 27–33 km². Under the assumption that underground and surface catchments correspond well in the steep terrain, we delineated a recharge area of 30 km² according to topography and surface catchments, local geology (i.e. the extent of Eocene strata) and the recharge area estimation of Hartmann et al. (2012) (see Figs. 1 and 3). TRAIN-ZIN percolation output for this area was summed up and used as an input to the MOSST model routines for calculating recharge and spring flow. Thereby, percolation output from the three best model runs were used to force the five variations of MOSST.
Table 2
Hydrological response units for the Upper Faria and parameters for runoff generation in TRAIN-ZIN with sources of parameters and final parameters values or ranges resulting from different model runs.

<table>
<thead>
<tr>
<th>Hydrological response unit</th>
<th>URB</th>
<th>GRASS</th>
<th>GRASSSTONE</th>
<th>MAN_OLV</th>
<th>UNMAN_OLV</th>
<th>AGR</th>
</tr>
</thead>
<tbody>
<tr>
<td>Final infiltration capacity (mm/h)</td>
<td>7–15</td>
<td>25–50</td>
<td>35–70</td>
<td>150</td>
<td>150</td>
<td>150</td>
</tr>
<tr>
<td>Initial losses (mm)</td>
<td>10–15</td>
<td>8</td>
<td>8</td>
<td>8</td>
<td>8</td>
<td>8</td>
</tr>
<tr>
<td>Depth of soil (m)</td>
<td>0.68–1.0</td>
<td>0.15</td>
<td>0.36–0.58</td>
<td>0.3</td>
<td>0.31</td>
<td>0.63</td>
</tr>
<tr>
<td>Porosity (-)</td>
<td>0.54</td>
<td>0.54</td>
<td>0.53</td>
<td>0.54</td>
<td>0.54</td>
<td>0.54</td>
</tr>
<tr>
<td>Permanent wilting point (-)</td>
<td>0.34</td>
<td>0.34</td>
<td>0.31</td>
<td>0.34</td>
<td>0.34</td>
<td>0.33</td>
</tr>
<tr>
<td>Field capacity (-)</td>
<td>0.45</td>
<td>0.45</td>
<td>0.43</td>
<td>0.45</td>
<td>0.45</td>
<td>0.44</td>
</tr>
<tr>
<td>Bulk density of soils (g/cm³)</td>
<td>1.22</td>
<td>1.22</td>
<td>1.25</td>
<td>1.22</td>
<td>1.22</td>
<td>1.23</td>
</tr>
<tr>
<td>Brook-Corey pore-size distribution index λ (-)</td>
<td>0.33</td>
<td>0.33</td>
<td>0.33</td>
<td>0.33</td>
<td>0.33</td>
<td>0.33</td>
</tr>
<tr>
<td>Saturated conductivity (cm/h)</td>
<td>1.4</td>
<td>1.75–3</td>
<td>1.25–3</td>
<td>2</td>
<td>3</td>
<td>3</td>
</tr>
</tbody>
</table>

Field measurements
Lange and Leibundgut (2003)
Schacht et al. (2011)
SPAW hydraulic properties calculator (Saxton et al., 1986)

Maidment (1994)
Saxton et al. (1986),
Singhal and Gupta (2010)

Table 3
Hydrologic signatures and the applied time steps.

<table>
<thead>
<tr>
<th>Evaluation Criterion (abbreviation, time step)</th>
<th>Equation</th>
<th>Explanation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bias in runoff ratio (%BiasRR, days)</td>
<td>$\sum_{t=1}^{N} \frac{Q_{\text{sim}}(t) - Q_{\text{obs}}(t)}{Q_{\text{obs}}(t)} \times 100$</td>
<td>$Q_{\text{sim}}$ and $Q_{\text{obs}}$ are simulated and observed runoff values, $N$ is the number of days</td>
</tr>
<tr>
<td>Bias median flow (%BiasFMM, days)</td>
<td>$\frac{\log(Q_{\text{median, sim}}) - \log(Q_{\text{median, obs}})}{\log(Q_{\text{obs}})} \times 100$</td>
<td>$Q_{\text{median, sim}}$ and $Q_{\text{median, obs}}$ are simulated and observed median runoff values</td>
</tr>
<tr>
<td>Bias high flows (%BiasHighQ, days)</td>
<td>$\frac{\sum_{t=1}^{N} Q_{\text{sim}}(t) - \sum_{t=1}^{N} Q_{\text{obs}}(t)}{\sum_{t=1}^{N} Q_{\text{obs}}(t)} \times 100$</td>
<td>$Q_{\text{sim},t}$ is the flow of timestep $t$ above a flow threshold for high flows (exceedance probability of $\sim 0.1$ in the observed quick flow duration curve), $N$ is the number of days</td>
</tr>
<tr>
<td>Bias medium flows (%BiasMiddleQ, days)</td>
<td>$\frac{\sum_{t=1}^{N} Q_{\text{sim}}(t) - \sum_{t=1}^{N} Q_{\text{obs}}(t)}{\sum_{t=1}^{N} Q_{\text{obs}}(t)} \times 100$</td>
<td>$Q_{\text{sim},t}$ is the flow of timestep $t$ within flow thresholds for medium flows (exceedance probability of between 0.1 and 0.9 in the observed quick flow duration curve), $N$ is the number of days</td>
</tr>
<tr>
<td>Bias low flows (%BiasLowQ, days)</td>
<td>$\frac{\sum_{t=1}^{N} Q_{\text{sim}}(t) - \sum_{t=1}^{N} Q_{\text{obs}}(t)}{\sum_{t=1}^{N} Q_{\text{obs}}(t)} \times 100$</td>
<td>$Q_{\text{sim},t}$ is the flow of timestep $t$ within flow thresholds for low flows (exceedance probability of $\sim 0.9$ in the observed quick flow duration curve), $N$ is the number of days</td>
</tr>
<tr>
<td>Bias time lag (%BiasTimeLag, hours)</td>
<td>$\frac{\text{LagTime}(Q_{\text{sim}}) - \text{LagTime}(Q_{\text{obs}})}{\text{LagTime}(Q_{\text{obs}})} \times 100$</td>
<td>LagTime($Q_{\text{sim}}$) and LagTime($Q_{\text{obs}}$) are lag times calculated for simulated and observed flows resp.</td>
</tr>
<tr>
<td>Bias rain class (%BiasRainClass, days)</td>
<td>$\frac{\sum_{h=1}^{H} \left( \frac{\text{BiasRainClass}(t) - \text{BiasRainClass}(t)}{\text{BiasRainClass}(t)} \right)}{\sum_{h=1}^{H} \left( \frac{Q_{\text{sim}}(t) - Q_{\text{obs}}(t)}{Q_{\text{obs}}(t)} \right)} \times 100$</td>
<td>$h = 1, 2, \ldots, H$ are the 60 combinations of daily max. rainfall intensity and rainfalls sum of the foregoing ten days and $t = 1, 2, \ldots, T$ are the days within each group</td>
</tr>
</tbody>
</table>

3.3. Hydrologic signatures

The selection of appropriate hydrologic signatures for model evaluation (see Table 3) was based on the modelling objectives, i.e. sound water balance estimation, and the availability of observed data. Several of the seven selected signatures denote long term input–output behaviour of the basin (%BiasRR) or runoff of different magnitudes (%BiasHighQ, %BiasMiddleQ, %BiasLowQ, %BiasFMM). Another measure (%BiasRainClass) compares measured and modelled runoff for combinations of rainfall conditions, i.e. rainfall of the ten foregoing days as a proxy for antecedent moisture and the maximum rainfall intensity per 5 min. The timing of runoff was analysed as time-shift at which the cross-correlation between mean areal rainfall and stream flow time series is maximum (%BiasTimeLag).

Using all these signatures, we selected the less-biased runs through a two-step approach: First, for the 22 realized model runs, we defined behavioural runs assuming acceptable bias ranges for each signature ($\pm 25\%$, $\pm 50\%$ or $\pm 100\%$, depending
on signature and subbasin). Among those, the three best model runs for each subbasin were selected calculating the lowest total of the absolute bias value of all signatures.

### 3.4. Model evaluation and aridity indicators

We used empirical recharge equations applied in previous studies in the region (e.g. Gräbe et al., 2013; Sheffer et al., 2010) to benchmark our modelled percolation rates. In 1958, Goldschmidt and Jacobs published estimates of annual recharge based on annual precipitation:

$$R = 0.8 \times (P - 360)$$

(1)

where $R$ is the recharge in mm/year and $P$ is the annual rainfall (mm/year).

Later modifications include the work of Zukerman and Shachnai (1999):

$$R = 0.97 \times (P - 463) \text{ if } P > 1000 \text{ mm/a},$$

(2)

$$R = 0.88 \times (P - 410) \text{ if } 650 \text{ mm/a} < P \leq 1000 \text{ mm/a},$$

(3)

$$R = 0.45 \times P \text{ if } 200 \text{ mm/a} < P \leq 650 \text{ mm/a},$$

(4)

and of Guttman (2000):

$$R = 0.8 \times (P - 360) \text{ if } P \geq 650 \text{ mm/a},$$

(5)

$$R = 0.534 \times (P - 216) \text{ if } 650 \text{ mm/a} < P \leq 300 \text{ mm/a},$$

(6)

$$R = 0.8 \times P \text{ if } P < 300 \text{ mm/a},$$

(7)

Additionally, we evaluated the TRAIN-ZIN percolation estimations with the MOSST model ensemble. The main disadvantage of MOSST is the simplification of hydrological processes at the soil surface. TRAIN-ZIN delivers percolation from its soil storage in a more realistic way; we therefore use the TRAIN-ZIN percolation to feed the groundwater routines of MOSST. Thereby, data from a distributed model replaces the conceptual soil-epikarst storage of MOSST used in previous studies. Finally, we used our model to calculate three monthly aridity indicators:

$$I_{\text{IEOF}} = \frac{Q_{\text{EOF}}}{Q_{\text{Tot}}}$$

(8)

$$I_{\text{Indir}} = \frac{\text{Perc}_{\text{Indir}}}{\text{Perc}_{\text{Tot}}}$$

(9)

$$I_{\text{Evap}} = \frac{\text{Evap}}{\text{ETP}}$$

(10)

where $I_{\text{IEOF}}$ is the indicator for Infiltration Excess Overland Flow, $Q_{\text{EOF}}$ its monthly sum and $Q_{\text{Tot}}$ the total overland flow in the same period. $I_{\text{Indir}}$ is the indicator for indirect percolation, with $\text{Perc}_{\text{Indir}}$ as monthly indirect percolation and $\text{Perc}_{\text{Tot}}$ as sum of direct and indirect percolation. $I_{\text{Evap}}$ is the indicator for evaporation from soil, Evap its monthly sum of soil evaporation and ETP is the total of all evapotranspiration processes.

### 4. Results

#### 4.1. Runoff measurements and hydrographs

In total, 49 mm and 4 mm of runoff derived from overland flow were measured in the wadis of AB and AF, respectively. The annual runoff coefficients ranged between 0.9% and 3.5% in AB and 0.1% and 0.6% in AF. Fig. 4 illustrates the observed daily quickflow hydrographs plus the simulated runoff. Generally, more surface runoff was generated in AlBadan (Fig. 4a) than in AlFaria subbasin (Fig. 4b). In some cases, e.g. at the end of the season 2005/06, the simulated runoff events were outside the gauging period. The first peaks of the season were partly overestimated by the model. Apparently, the worst fit was reached for the last season.

#### 4.2. Hydrologic signatures

A comparison of the 22 model runs across all seven signatures (Fig. 5) showed a great diversity in bias magnitudes, with maximum bias values of 65% for AB and more than 100% in AF. While differences between the runs were high for some of the signatures (e.g. %BiasHighQ), they were low in other cases (e.g. %BiasTimeLag). In AB, only values for %BiasMiddleQ and %BiasTimeLag exceeded 25% for the three best model runs. In AF, the bias in low flows (%BiasLowQ) did not reach acceptable values, while other bias values mostly remained below +/-50%. Overall, signatures of the three best model runs were more different in AF than in AB; at the same time, several signatures were close to zero in both subbasins.
4.3. Water balance analysis and model evaluation

4.3.1. Inter-annual variation

Inter-annual changes in water balance showed slight, but recognizable differences between AB and AF (Fig. 6 and Table 4). The part of the rainfall input that evaporated was higher in AF than in AB, at the expense of both percolation and runoff. Relative change between the seasons was not the same for all water balance elements. For example, rain in AB differed by more than 150 mm between the seasons, i.e. by about 40% of its average value of 516 mm. Both evapotranspiration and percolation varied by about 100 mm, which corresponded to relative changes of 30% and 70%. Similar effects were found for AF. As illustrated by the dashed lines in Fig. 6, years with very similar rainfall amounts yielded considerably different amounts of percolation both in AB and AF.

4.3.2. Precipitation-recharge relationship

A linear relation could be established between annual values of rainfall and recharge for the observed precipitation range (Fig. 7). The percolation–precipitation ratio was generally higher in AB than in AF. When we combined all values into a single linear relationship, the empirical equations of Gutman (2000) and Zukerman and Shachnai (1999) fell within our confidence band, while the simpler equation of Goldschmidt and Jacobs (1958) did not.

4.3.3. Percolation and simulated spring outflow

For the assumed catchment of Faria spring, TRAIN-ZIN estimated a mean percolation of 170 mm to 201 mm for the period 2004/05–2009/10, depending on the chosen model run. Fig. 8 shows the simulated spring discharge as output from the five different variations of the MOSST model (Models 1–2 vs. 3–5) driven by percolation output from the three best model runs of TRAIN-ZIN (the range resulting from the different TRAIN-ZIN model runs is given as coloured area). Measured spring discharge is given as well. Because abstractions were not included in the model concept, the declining discharge in the measurements that started in 2006 caused by overpumping could not be represented. Differences in the MOSST model structure and the chosen TRAIN-ZIN runs both influence the result and represent the uncertainty in modelling results.
4.4. Aridity indicators

Fig. 9 illustrates the relationship between the monthly rainfall amounts and the aridity indicators for the total Upper Faria (AB and AF) based on our model results. As expected, drier months were mainly dominated by Infiltration Excess runoff generation ($I_{\text{IEOF}}$, Eq. (8)), indirect percolation processes in the form of transmission losses account for a recognizable proportion of total percolation in dry months only ($I_{\text{IDir}}$, Eq. (9)), and the percentage of soil evaporation on the total evapotranspiration was highest in the driest months ($I_{\text{Evap}}$, Eq. (10)). Intercomparison between the three aridity indicators (Fig. 10) suggested that a high $I_{\text{Indir}}$ as well as a high $I_{\text{Evap}}$ corresponded to a high $I_{\text{IEOF}}$. However, a high $I_{\text{Evap}}$ occurred for very low $I_{\text{Indir}}$ values, whereas the proportion of evaporation from initial losses to total evapotranspiration was associated with high $I_{\text{Indir}}$ values. Colour code is used to illustrate the influence of the period of the year (beginning or end of wet season) on the results.

5. Discussion

The chosen signatures revealed differences in the model performance between the two subbasins. In AB, time lag and medium flows showed highest deviations between observed and simulated signatures. In AF, time lag was critical as well, but low flows, not medium flows had a high bias. However, small runoff volumes can have high relative bias values even
if absolute errors are minor, as is the case for AF in general and specifically for low flows. The smaller total runoff volume in AF (4 mm compared to 49 mm in AB) made accurate modelling more difficult. Overall, the chosen signatures suggest that the correct timing of runoff generation was a shortcoming which may be attributed to deficiencies in model structure, e.g. the overestimation of Infiltration Excess Overland Flow mechanisms, or input data, e.g. the correct representation of precipitation pattern. Runoff volumes were represented reasonably well to deal with water balance issues. The three best model runs selected for each sub-basin spanned the uncertainty range of the given modelling strategy, which was not too high for our subsequent analysis. So far, no commonly accepted threshold values exist for hydrological signatures. We chose values of mostly below 25% as acceptable given the conditions of the study area.

An optimal fit between measured and simulated daily hydrographs, as it is often the goal of hydrological model calibration, was not the main aim of this study for several reasons: First, low runoff coefficients indicate that a calibration based on runoff considers only a small portion of the total water balance. Second, runoff measurements are often subject to considerable errors. Third, the high spatial variability leads to problems simulating exact runoff responses, e.g. if the real rainfall distribution is not met by the existing measurement devices (e.g. Goodrich et al., 1995). Underestimations of measured runoff by the model (marked event No. 1 in Fig. 4a) could be caused by higher rainfall amounts not detected by existing gauges, whereas overestimations (marked event No. 2 in Fig. 4a) could result from overestimation of regional rainfall due to the position of rain gauges. Partly, the measurement period did not span the entire rainfall season (e.g. marked event No. 3 in Fig. 4a) and overestimation of Infiltration Excess Overland Flow could be the reason for modelled runoff at the beginning

**Fig. 6.** Annual water balance for AB (left) and AF (right) for the six simulated seasons; the width of the colour band indicates range in values for each water balance elements caused by differences in the three selected model runs; dashed lines indicate years with similar precipitation totals.

**Fig. 7.** Annual rainfall vs. annual recharge for the Upper Faria with annual average values of the three best three model runs for AB and AF, 95% confidence band and 95% prediction band and empirical recharge estimates.
of the season not met by measurements (e.g. events marked with No. 4 in Fig. 4a). Whereas the local relevance of Infiltration Excess Overland flow needs further investigations, Saturation Excess Flow processes were better reproduced.

The correct representation of soil moisture dynamics is not only crucial for saturation related runoff processes, but also for percolation modelling. This component of the water balance is important because it accounts for about 30% of the rainfall, whereas surface runoff is only responsible for 2%. Our simulation results have the same slope as the frequently applied empirical estimations of Guttman (2000) and Zukerman and Shachnai (1999), but not as the formula of Goldschmidt and Jacobs (1958). The relation between annual rainfall and recharge is linear in the range of annual rainfall investigated (>400 mm/a). However, a certain amount of initial cumulative rainfall has to be retained before recharge and runoff processes take place, leading to a non-linear relation for smaller rainfall amounts (e.g. Ben-Zvi, 1988). Rimmer and Salingar (2006) showed for the Hermon Karst area that the drying-out of the top soils during summer leads to a delay in runoff production and a non-linear relation for very dry years. Similar results were obtained by Samuels et al. (2009) in the same region and by Hartmann et al. (2014) for a Spanish karst spring. Really dry years were not included in our data, but overestimations of the season's first peaks may also be a result from underestimating this threshold for runoff generation.

Moreover, simple annual rainfall–recharge relationships neglect the spatial and temporal variability within the seasons. Percolation rates, as well as surface flow rates, do not only depend on the yearly rainfall amount of the season, but also on the distribution of the rainfall during the rainy season (Ries et al., 2015). In addition, our results illustrate that different space-time patterns of rainfall may lead to identical mean annual precipitation inputs, but to a wide range of possible catchment
reactions such as different runoff generation mechanisms. The resulting differences in annual percolation rates are neglected in empirical linear equations, although the variability in recharge is possibly more relevant for water management than the more obvious, but quantitatively less important variances in wadi runoff. Furthermore, our results show that spatial variability may occur within small distances, since more water percolated in AB than in AF although precipitation input was similar. This difference in water redistribution may be caused by a combination of geology, land use and terrain. Evapotranspiration is the greatest component of the water balance, but is not always adequately represented in dryland hydrological modelling. Potential evapotranspiration is often estimated with empirical methods as the Hargreaves formula (e.g. Hartmann et al., 2012; Schmidt, 2014) or the Penman–Monteith-approach (e.g. Portoghese et al., 2008; Smiatek et al., 2014). However, the treatment of the canopy as a single uniform cover in the latter approach is questionable for drier areas (Zhou et al., 2006). In this study, the detailed consideration of different evaporation processes, applying the Shuttleworth–Wallace approach suitable for non-closed canopies (Shuttleworth and Wallace, 1985), allowed for realistic water budget calculations.

Closed water balance estimations are rarely found in the in the western tributaries of the Lower Jordan River. If so, they agree with the findings of this study where mean annual actual evapotranspiration and recharge are about 70% and 30% of mean annual rainfall respectively. Hartmann et al. (2012) found similar numbers for long term simulations of the Faria spring. Comair et al. (2012) derived similar ratios for the entire West Bank from satellite data and Sheffer et al. (2010) for the Western Mountain Aquifer Basin. However, this study is the first high-resolution water balance model in the West Bank we are aware of and the first model that is validated by hydrological signatures and checked by an independent karst groundwater model. It has to be kept in mind that the TRAIN-ZIN values are percolation, not recharge values. The use of the percolation output from the distributed TRAIN-ZIN model to drive the lumped conceptual MOSST model and convert percolation into recharge proved useful in this context.

The six-year average of channel flow measured for the total upper Faria was about 0.8 MCM and we calculated annual percolation as potential groundwater recharge to reach up about 22 MCM. Even if all this water were available for human use, it would only be slightly higher than the yearly water demand for irrigation (15.3 MCM) and for domestic water use (5.7 MCM), as estimated for the beginning of the 21st century (Shadeed, 2008). Considering lower water availability due to issues of water quality, infrastructure and an expected increase in consumption following population growth, a considerable supply-demand gap exists today and will increase in the near future. In addition, natural variability limits a secure annual renewal of water resources even without severe drought conditions. Drier years might also intensify effects like the drying out of local springs.

Aridity indices calculated from our modelled aggregated monthly water balance components yielded consistent results. Low monthly rainfall sums were related to higher fractions of Infiltration Excess Overland Flow and indirect percolation. Our results were less clear for the relation of soil evaporation to total evapotranspiration. Redistribution of water between evaporation and transpiration is apparently more complex, and might be influenced by aspects of timing and vegetation development as well (Lauenroth and Bradford, 2006; Nicholson, 2011). However, only process-oriented models allow establishing such general relations between climate and hydrology that may provide additional information to water managers, e.g. volumes of transmission losses to be expected for managed aquifer recharge measures.

Likewise, the combination of TRAIN-ZIN with the MOSST model ensemble has further potential for water management, despite its critical aspects like the possibility of transboundary flows. Water balance problems in data-scarce regions are often addressed with lumped monthly or even yearly water balance approaches. TRAIN-ZIN works on higher spatial and
temporal scale and thereby allows evaluating model results with the help of runoff based signatures that provide insights into model shortcomings (Gupta et al., 2008). However, the signature approach cannot solve the general problem of limited data quality and quantity in semi-arid and arid areas. Runoff is often the only recorded measure to compare model results to, although it is only a small percentage of the total water balance. Hence, also in our case, future studies are desirable to improve the model, to confirm the results and to gain experience in selecting signatures that are non-redundant and relevant. However, the large variability shown in the results emphasizes the importance of temporally distributed instead of annual models.

6. Conclusions

The combination of karstic geology and semi-arid climate in the Eastern Mediterranean creates specific hydrologic conditions that are best addressed by a model specifically developed for these environments. Scarcity and low quality of data further increase the difficulty of model evaluation. The distributed hydrological model TRAIN-ZIN was designed for these conditions and applied for water balance estimation in the upper part of the Wadi Faria catchment. Seven hydrologic signatures helped to constrain the possible model outcomes and provided insights into shortcomings in model structure or parameterization, for example problems in modelling the correct timing of runoff. Model results could be validated by recharge estimation methods and by a karst hydrological model.

Our water balance estimation corroborated the order of magnitude from previous studies as well as some frequently applied empirical formulae for estimating groundwater recharge. As in previous studies, the non-linearity of processes and the importance of the spatial and temporal variability became obvious, which is still often neglected in practical applications. Model results allowed investigating the influence of aridity on several processes (Infiltration Excess vs. Saturation Excess Overland Flow, direct vs. indirect groundwater recharge, evaporation vs. transpiration). However, these results are model output that needs to be supported in future studies.

The often stated gap between water supply and demand could be confirmed in this study for the Faria catchment. Even without climate change or extreme droughts, the replenishment of renewable water supplies varies considerably between consecutive years, through a combination of climatic conditions and karst aspects. This variability has practical implications for water management decisions, e.g. for planning of rainwater harvesting or managed aquifer recharge infrastructures. A detailed model like TRAIN-ZIN estimates this variability and offers more guidance in these questions than simple annual models. The growing demand for water in the region makes future studies necessary that combine reliable data, adapted models and advanced model strategies to create the best possible basis for water management.

Acknowledgments

This study has been partly funded by the German Ministry for Education and Research (BMBF) in the framework of the GLOWA Jordan River Project. We are grateful to Lucas Menzel for providing the original TRAIN model code and to the EXACT project for their help with runoff and rainfall gauging data. The article processing charge was funded by the German Research Foundation (DFG) and the Albert Ludwigs University Freiburg in the funding programme Open Access Publishing. We are grateful to an anonymous reviewer and to Alon Rimmer who improved our manuscript.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at http://dx.doi.org/10.1016/j.ejrh.2015.08.002.

References


