Field-based estimation and modelling of distributed groundwater recharge in a Mediterranean karst catchment, Wadi Natuf, West Bank

Clemens Messerschmid[1,2], Jens Lange[1] and Martin Sauter[2]

[1]Chair of Hydrology, University of Freiburg, Freiburg, Germany
[2] Geoscience Center, Applied Geology, University of Göttingen, Göttingen, Germany

Correspondence to: Clemens Messerschmid (clemensmesserschmid@yahoo.de)

Abstract. While groundwater recharge is one of the most prominently covered subjects in hydrogeology, the spatial distribution of recharge has been given relatively little attention, especially in semi-arid, karstic aquifers. Under conditions of highly diverse geology, relief, vegetation and land use, the complexity and variability of spatially distributed hydrological processes remains a challenge in many regions around the world, especially in hitherto ungauged basins, such as Wadi Natuf, a 103km^2 large karstic Eastern Mediterranean watershed in the Palestinian upstream mountain and recharge area of the Western Aquifer Basin (WAB), which is shared with Israel in the coastal plain. In this first in a series of two papers, distributed recharge is estimated and represented, based on seven years of extensive field observations and measurements and based conceptually on some of the lessons of research on predictions in ungauged basins (PUB research), i.e. by drawing on observable physical landscape features such as geology, land use and land cover (LU/LC) and especially soil conditions. For the first time in the WAB, a forward calculated soil moisture percolation model was set up with parameters, directly gained from field observations and in a strictly parsimonious manner, as a 'one reservoir' model ('tank' model) of saturation excess in the soil column. Following the recommendations of the PUB-community, the model design was developed by determining main active hydrological processes and by identifying patterns of linkage between different landscape features. This linkage allowed us to further simplify the model and its requirements in a realistic manner for eight soil moisture stations, chosen at six different geological formations with typical soil and LU/LC conditions. It was found that recharge varies widely at the spatial dimension, ranging between 0 % and almost 60 % of annual rainfall. The model was tested on two conceptual levels – on the level of basin form (a), e.g. observed and physical soil moisture as model parameters and on the level of basin response (b), e.g. by comparing signatures of peak deep percolation events (recharge) with peak discharge at local springs under well controlled conditions in isolated sub-catchments. The conceptual approach investigated basin form and basin response separately. Together with the parsimonious model, it is applicable in many of the scarcely gauged groundwater basins around the world, where karstic conditions and a highly diverse landscape and geology prevail, such as Wadi Natuf in the West Bank.

Keywords. Distributed recharge, landscape features, basin signatures, soil moisture, reservoir model, parsimonious approach, PUB, recharge coefficients, saturation excess.

1 Introduction

Knowledge about distributed recharge is important for the exploration, sustainable management, protection and equitable allocation of water resources. This research remains complex and challenging for hydrogeological research, especially in karstified aquifers and even more so in the many ungauged basins around the world. Karst aquifers supply about one quarter of world population (Ford and Williams, 2007), many of which remain poorly gauged or ungauged. One key issue of water negotiations in contested groundwater basins is the spatial distribution of water resources in the different zones, such as the recharge zone of an aquifer.

1.1 Background on recharge methods

The study of groundwater recharge and its spatial variability in semi-arid to sub-humid karst regions can be subdivided into direct and indirect methods of research (Lerner *et al.*, 1990) or direct and indirect approaches (Dörhöfer and Josopait, 1997) – not to be confused with direct and indirect recharge (see also Scanlon *et al.*, 2002, 2006; Bredenkamp *et al.*, 1995; Simmers *et al.*, 1988; De Vries *et al.*, 2002; Lloyd, 1980). Direct approaches determine infiltration and percolation processes in the recharge zone by surface-near observations that allow for the parameterisation of recharge governing factors (Dörhöfer and Josopait, 1997). Such approaches focus on processes such as surface runoff, interflow and soil moisture saturation. In contrast, indirect approaches usually model integrated basin behaviour based on observations of spring and river discharge and abstractions in the (often confined) productive regions, as well as of storage change, as far as it is measurable through monitoring wells in the deep saturated layers of an aquifer.

Scanlon *et al.* (2002) distinguished three groups of recharge estimation methods, for the unsaturated and the saturated zone as well as for surface water. For the unsaturated zone, Scanlon et al. (2002) listed lysimeter studies, tracer techniques, Darcy's law, numerical modelling and soil moisture balance techniques (e.g. zero flux plane method). When no other option than observations in the unsaturated zone are available, then groundwater recharge has to be equated with deep percolation from the soil cover into the aquiferous bedrock formation.

Jeannin and Sauter (1998) distinguished between two types of process-based karst models, lumped and distributed, for which different types of recharge information are required (also used by Hartmann *et al.*, 2013). Distributed models discretise the basin into cells, which require at each element, spatial information on factors that control hydrological processes, whether at the surface (such as runoff, infiltration, soil, epikarst etc.) or deep underground (such as karst properties, storage change or outflows, etc.). On the other hand, lumped models do not require spatial information but transfer input into output by a set of equations on a basin-wide scale. Still, in poorly gauged basins, lumped approaches have long been the preferred choice; yet, in most cases precisely in order to avoid problems with spatial differentiation and over-parameterization.

For ungauged basins, some of the lessons of the IAHS-led research decade from 2003 to 2012 on hydrological predictions in ungauged basins (PUB) may be useful here. Most PUB studies emphasized the need to use process-oriented models in order to represent complex spatial variability. They observed and advocated a shift from lumped and integrated models to distributed models and from indirect to direct approaches (Hrachowitz *et al.*, 2013; Sivakumar *et al.*, 2013.

1.2 Approaches to spatial variability in ungauged basins (results of PUB)

To differentiate and quantify the spatially distributed, highly complex and overlapping processes of groundwater recharge, several general lessons can be drawn from the PUB decade (Hrachowitz *et al.*, 2013), although most studies dealt with runoff rather than recharge. As Hrachowitz, *et al.* (2013) and Sivapalan *et al.* (2003a) reported, most PUB studies, both observed and suggested a shift of the research focus away from data- and calibration-focussed methods and towards the complexities of observable system understanding, which often require direct approaches to describe distributed physical in-situ processes. In order to improve the conceptual realism, Pomeroy (2011) summarized three principle aims: to demonstrate the value of observations, to reduce the reliance on calibration and to enhance the capability to predict based on process understanding.

1.2.1 Physical parameters and basins response (signatures)

The parameterization of location-specific physical landscape characteristics is a starting point for models to ensure a complete and physically realistic representation of dominant processes and their respective features (e.g. Franchini and Pacciani, 1991), and they can be loosely divided into three groups including bedrock characteristics, the characteristics of overlying soils and land use and land cover (LU/LC) at the surface: Batelaan and de Smedt (2001) accounted for land use, soil and lithology; Batelaan and de Smedt (2007) accounted for land use, soil and topography (but not geology) but calculated with parameters based on literature values, while Aish, Batelaan and de Smedt (2010) used land use and soil-type as physical parameters for their spatially distributed model of long-term annual average recharge (besides runoff and actual evapotranspiration) in the Gaza Strip. Zomlot *et al.* (2015) emphasised the importance of soil texture and vegetation cover. For an overview, see also Hrachowitz *et al.* (2013) and the discussion in Messerschmid *et al.* (2019b).

In contrast to the physical basin form, the hydrological basin function (or impact) already represents the outcome of in-situ processes. Savenije (2010) suggested assigning individual hydrological processes and distinct hydrological functions (in his case, runoff) to different landscape units by dissecting catchments in a semi-distributed way and according to a hydrologically meaningful landscape classification metric. As a challenge however, remains the correct and appropriate linkage and translation from form to function (Messerschmid *et al.*, 2019b). Among the many PUB findings on basin-specific links between physical features and basin response (i.e. between form and function), Eder *et al.* (2003) and Hartmann *et al.* (2013) suggested the use of so-called hydrological system signatures to describe emergent properties of the system – whether quantitatively or qualitatively – such as flow duration curves or spring hydrographs.

1.2.2 The need for simplification and parsimony

In their quest for a common framework for hydrological modelling many PUB researchers increasingly realized a need for simplification and parsimony with a focus on dominant local processes, such as simple water balance equation and tank models (Sivakumar *et al.*, 2013), especially under the presence of equifinality in calibrated parameters (Hrachowitz *et al.* (2013) – see also: Grayson and Blöschl (2000); Woods (2002); Sivapalan *et al.* (2003b); McDonnell and Woods (2004); Sivakumar (2004, 2008); Wagener *et al.* (2007); Young and Ratto (2009) and Olden *et al.* (2012). They observed that conventional hydrologic models were often far too complex and require too many parameters and data (Sivakumar *et al.*, 2013). They also place too much emphasis on specific, often preselected independent concepts and mathematical techniques, rather than on the coupling or integration of observations to capture the salient features of hydrologic systems. However, despite the heterogeneity of local parameters and processes, the response at the catchment-scale is often characterized by a surprising process simplicity (Sivapalan *et al.*, 2003a). Often, a minimal number of the appropriate field observations can describe the main characteristics of a process (Seibert and Beven (2009). Even in ungauged basins, few measurements and short time series data may contain most of the information found in long term extensive measurements (Juston *et al.*, 2009; Seibert and Beven, 2009), for a detailed understanding of catchment responses at the local scale (Blume *et al.*, 2008). A good empirical fit to basin responses could often be achieved with exceedingly simple models (Nash, 1957; Dooge, 1959; Lambert, 1969) focussing on processes and using physical parameters directly from the basin (Beven and Kirkby, 1979). And it should be noted that rather than diminishing, this reduction often adds to the reliability and controllability of process outcomes, particularly in karstic environments, which are known for their high spatial variability (Hartmann *et al.*, 2012b), their non-

Darcian flow, anisotropic flow fields (Dafny *et al.*, 2009, 2010) and increased problems of equifinality and multicollinearity. Keeping the steps of interpretation, conceptualisation, classification and quantification transparent, was described by Beven & Kirkby (1979) as a three way compromise between a) the advantages of model simplicity, b) the complex spatial variability of basin hydrological response and c) the economic limitations on field parameter measurements.

1.2.3 Empirical observations – the PUB paradox

Most researchers agreed that in order to ensure proper hydrological interpretation, conceptualisation and theorisation in light of complexities of hydrological processes, the research has to be grounded in observation, at least 'some' observations as Sivapalan *et al.* (2003a) pointed out, and it should be designed for specifically targeted measurements (McDonnell *et al.*, 2007). In other words, PUB research repeatedly emphasised the importance of empirical approaches. For example, in South-Africa, but also in the Israeli-Palestinian Western Aquifer Basin, Kapangaziwiri *et al.* (2012) and Hughes and Kapangaziwiri (2007) were confronted with the problem that, when model parameters were not gauged and measured on site but estimated directly and a priori from physical properties of the catchment, they may show little consistency. Therefore, these methods have to obtain process understanding from empirical experiments and furthermore should be based on fundamental physical laws and theories (Hrachowitz *et al.*, 2013), as well as local knowledge of the observable landscape and climate controls of hydrological processes (Sivapalan, 2003).

These findings point to the so-called PUB paradox, according to which data-rich catchments are needed to test methods for data-poor environments as Seibert and Beven (2009) noted. They therefore suggested taking a 'hydrologically intelligent choice', i.e. the correct interpretation of the dominant, most relevant features that govern a process while simultaneously keeping observations and measurements simple and at the necessary minimum, for reasons of practicality (e.g. measurement costs and field accessibility). Binley and Beven (2003) added that such an intelligent design remains a learning process that may start with insufficient data and adds data as required by the application and Seibert & Beven, 2009).

1.3 Western Aquifer Basin

The Western Aquifer Basin (WAB) is the largest, freshest and most productive groundwater basin in all of Israel and the oPt (West Bank and Gaza Strip). It spans from the sub-humid Mt Carmel in the North into the arid Sinai in the South and from the West Bank Mountains to the Mediterranean Sea (see Fig. 2a). In the literature, its size ranges between 9,000 and 14,167 km^2, depending on the boundaries set by different authors (for an overview, see: ESCWA-BGR, 2013). During the pre-utilization period, the WAB drained through two principle karst spring outlets in the Coastal Plain, on which its old Arabic, as well as the more recent Hebrew names are based: Aujah-Tamaseeh or Yarkon-Tanninim, respectively (SUSMAQ, 2002). The WAB is a classical transboundary water basin, subject to negotiations over equitable allocations under international water law (UN-GA, 1997). The intensive use of the basin dates back to the early 1960s, starting with deep well drilling in the Coastal Plain and adjacent foothills area; and the basin rapidly became one of the single-largest sources of Israeli water supply, potentially also one of the largest future sources of West Bank supply. On the other hand, the WAB's recharge and accumulation areas, located in the mountains, slopes and foothills of the West Bank, remain almost untouched, ungauged and unexplored, not least due to the continued restrictions imposed on Palestinian water sector and infrastructure development (World Bank, 2009). Within this recharge zone lies Wadi Natuf, a 103 km^2 large surface watershed, stretching from the crest of the West Bank Mountains, North of Ramallah, down to the Green Line (that separates Israel from the Palestinian West Bank) - see section 2.

1.3.1 Existing basin-wide recharge studies in the WAB – approaches and limitations

Recharge in the Western aquifer has been the subject of many studies, dating back to the British Mandate and the 1950's. Yet until now, most studies have approached recharge as basin-wide uniform process and by simply equating it with observable natural spring flow (Goldschmidt and Jacobs, 1958) or with discharge from springs and wells in the Israeli Coastal Plain, under the assumption of mass conservation or by accounting for storage change (Guttman *et al.*, 1988; Guttman and Zukerman, 1995; Guttman, 2000; Berger, 1999; Abusaada, 2011; Dvory *et al.*, 2016 and Abusaada and Sauter, 2017). On a regional scale such simplified, lumped water budget is permissible, yet contributes little to understand the spatial variability of distributed recharge in the outcrop areas of the West Bank Mountains, particularly on a local scale, where the hydrostratigraphy is more complex (see section 2). Abusaadah (2011) presented a three-dimensional integrated flow model, encompassing plains and slopes as separate recharge and flow zones, however with limited accuracy due to missing data and limited temporal resolution of measurements. He thus called for refined meteorological input, as well as for more details on land use, soil, geology and structural characteristics in the WAB, as many others had done before in other areas (Martínez-Santos and Andreu, 2010).

Few authors have attempted to investigate spatial recharge distribution and variability in the WAB. As an exception, Sheffer *et al.* (2010) used a hybrid of lumped basin water budgeting and spatial distribution in the upstream, albeit without independently establishing and testing the sensitive field parameters (physical form) and with little local empirical observation and insights into the complex distributed recharge processes and their dominant features. Instead, his soil percolation model parameters were taken as estimates from the literature (e.g. Dingman, 1994) and then retrofitted through repeated calibration runs. Sheffer (2009) coupled 3,300 soil moisture "tank" models in the recharge zone with a lumped flow model for basin response. He accounted only for two rock types (permeable and less permeable). Due to his lumped calibration on responses in the remote discharge zone, he had no possibility to check the plausibility of the assumed recharge zone parameters and remained exposed to problems with equifinality and the complex distributed recharge processes.

Hughes and Mansour (2005) and Hughes *et al.* (2008) published results of a fully distributed so-called "object-oriented" recharge model that separates different inputs and processes, such as runoff, spring flow, soil infiltration and deep percolation (recharge) into individually calculated compartments. However, while their model used a high spatial resolution, it did not distinguish between sub-basins but combined all distributed input together into one lumped, integrated basin-wide model. In addition, the parameters themselves were rarely grounded on local field observations. Instead, they employed assumed parameters (e.g. field capacity and wilting point in their soil moisture sub-model), estimated directly and *a priori* from physical properties of the catchment, which led to the aforementioned problems of lacking consistency between physical features and modelled basin response. Their study therefore confirms the findings from the PUB-decade, e.g. that a thorough understanding of the dominant processes at play must be firmly based on observations and descriptions of physical features, such as land use, soil, geology or structural characteristics (compare also: Abusaada, 2011 and Martínez-Santos and Andreu, 2010).

1.3.2 Local scale recharge studies in the WAB

Individual empirical field studies based on winter observations in very localised sub-catchments investigated local processes of soil saturation, runoff and infiltration on a scale of less than 1 km^2 (Steinmann, 2010 and Grodek *et al.*, 2011). However, they focussed on runoff rather than on recharge and their analysis was mainly based on differentiating surface characteristics, rather than differentiating between lithostratigraphic units (Steinman, 2010). Lange *et al.* (2003) used tracers in local sprinkler tests (of 10 x 18m plot size) near Wadi Natuf to investigate rainfall-runoff relationships and – important for the process understanding – pointed out the importance of soil saturation excess for surface runoff and recharge (Lange *et al.*, 2003). Otherwise, no regional tracer tests to investigate recharge were carried out in the WAB so far, possibly due to the above mentioned lack of intake and observation points (e.g. monitoring wells) in the upstream West Bank recharge zone. Several authors conducted cave-drip experiments in the region providing some valuable insights into percolation processes through soil and the unsaturated epikarst and delay times between event rainfall and drop formation at the roof of caves (Sheffer, 2009; Frumkin, 1994; Arbel *et al.*, 2010 and Lange *et al.*, 2010). However, the generalisation of their results is restricted because the contributing areas are unknown and caves may develop their own hydraulic environments (Ford and Williams, 2007); the results may therefore not be representative for larger areas, as pointed out by Ries *et al.* (2015) - compare also to Sheffer *et al.* (2010), Hartmann *et al.* (2012) and Lange *et al.* (2010).

1.3.3 Studies in adjacent catchments

Finally, three studies from aquifers adjacent to Wadi Natuf shall be presented. Two empirical field studies in a nearby catchment of the Eastern Aquifer Basin (EAB) fitted a soil moisture saturation excess model to soil moisture field measurements (once with three and another time with 4 soil moisture plots) and identified seasonal recharge pulses on an event basis (Ries *et al.*, 2015 and 2017). In line with the aforementioned findings by Lange *et al.* (2003 and 2010), Ries *et al.* (2015 and 2017) showed that soil moisture saturation was the principle process of runoff and recharge generation in the specific field conditions of the mountain areas of the EAB adjacent to Wadi Natuf under sub-humid climate conditions and an abundance of epikarst. Schmidt *et al.* (2014) in another nearby catchment on the Eastern slopes that stretches from of the West Bank Mountains to the Jordan Valley, successfully employed a parsimonious recharge and groundwater flow model with two components: they coupled a soil moisture saturation model, similar to that of Sheffer *et al.* (2010) for deep percolation (input) with a karstic double porosity aquifer component for groundwater flow towards a field-monitored spring outlet. However, again, their soil parameters and time series data of soil saturation were not based on empirical field measurements but taken from the general literature and refitted through model calibration against the basin wide response in form of the hydrograph of the principle drainage point (Al-‘Aujah spring).

1.4 Research gaps

Two of the three general venues to estimate recharge, cited by Scanlon *et al.* (2002) are not applicable to Wadi Natuf: As discussed in section 1.3 and 2, observations of the saturated zone are severely limited. Equally, no surface water based approaches are a possible venue for recharge studies in Wadi Natuf. Only the unsaturated zone is accessible for field measurements and observations – see also Sheffer *et al*. (2010). As in most, if not all ungauged basins worldwide, also in the WAB, the general problem of a severe lack of primary field data from observation and measurements has to be addressed and cannot be circumvented by ever more sophisticated advances in modelling. This remains a challenge, as the IAHS-led decade on PUB concluded. Instead, new methods are needed that focus on the understanding of the spatially variable and complex recharge processes and find ways for their conceptual description and classification as well as their parameterisation. Surface- and process-based direct methods in the recharge area and on the catchment scale are called for. The description and quantification of dominant parameters should be firmly anchored in field observation and where possible, field measurements. In ungauged and poorly gauged basins, the basin classification and process understanding should start with observable physical features and their patterns and interactions. At the same time, models should adhere to the goal of parsimony and avoid over-calibration as well as the problems of equifinality.

1.5 Aims, motivation and approach

The objective of this study is to investigate spatially distributed recharge in a largely ungauged karst aquifer catchment in order to develop formation-specific recharge coefficients (RC) in the central WAB. A simplified and parsimonious “tank” model for laterally disconnected one-column elements (Hrachowitz and Clark, 2017) was performed in daily steps and based on observable physical features (basin from) and empirical field measurements for core parameters of the soil water budget – thus avoiding the need for retrofitting. The model was examined and tested in two different ways: On the one hand, and at the level of physical parameters governing recharge (basin form), the modelled soil moisture content was compared to the field-measured readings. On the other hand, the signature of measured spring flow events with their periods of peak discharge (basin response) was compared to the modelled periods of deep percolation (DP). In other words, this study used empirical (inductive) approaches to determine spatial variations in groundwater recharge and then tested the model by both, semi-qualitative and quantitative, measured basin observations alike. Finally, the model results of spatially distributed annual recharge coefficients were compared to literature results from former recharge studies in the WAB.

2 Study area

Wadi Natuf is a 103 km^2 catchment stretching from the mountain plateau at the crest of the West Bank at 816 m asl down to its foothills at 138 m asl (Fig. 1). In the Eastern mountain region, the topography is characterized by the rocky mountain crests and high plains with steep terraced slopes towards the Central Wadi Natuf further west. Here, in the midstream catchment, undulating hills with deeply incised ephemeral rivers (Wadis) dominate, descending further down to the gentle slopes and plains in the lower Natuf region near the outlet of the main Wadi bordering the Coastal plain. Within the Natuf catchment, all sub-aquifer formations of the WAB and some of the overlying cover series crop out (Fig. 2) and are therefore exposed to direct infiltration. The climate is typically Eastern Mediterranean with rainfall amounts monotonously rising from the semi-arid Western foothills to sub-humid conditions in the Eastern Mountains. Wadi Natuf drains westwards towards the Mediterranean Sea and has a remarkably low overall runoff coefficient of approximately 0.11 %, mainly due to high percolation rates into the bedrock under shallow soils and due to considerable transmission losses (Messerschmid *et al.*, 2018) into the karstified carbonate materials underlying the wadis.

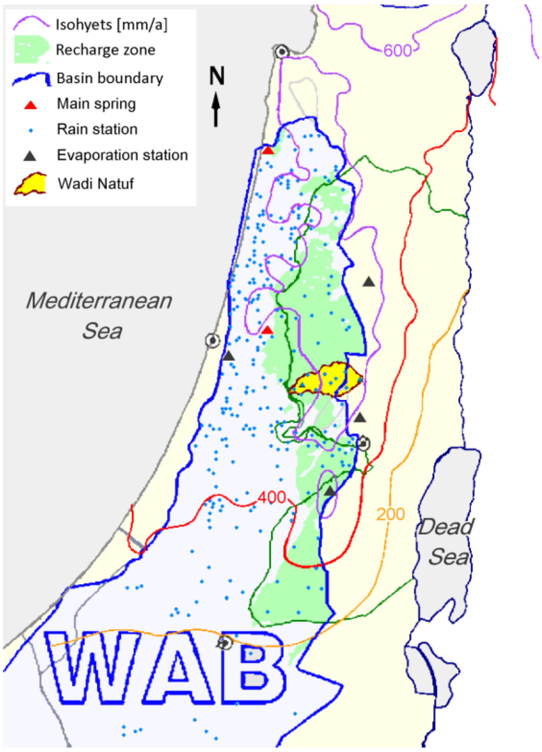

Figure 1: Western Aquifer Basin (WAB) overview and Natuf location map. Wadi Natuf (in yellow) lies in the uphill recharge zone (shaded green) of the WAB, and the easternmost part of its surface catchment belongs to the Eastern Aquifer Basin. Isohyets (different colours) indicate semi-arid to sub-humid climate in Wadi Natuf. The former principle spring outlets of the WAB (red triangles) have all but dried up, due to excessive pumpage in Israel. Four evaporation stations (black triangles) around Wadi Natuf were complemented by two Automatic weather stations within the catchment by this study. Source: modified after Ettinger (1996).

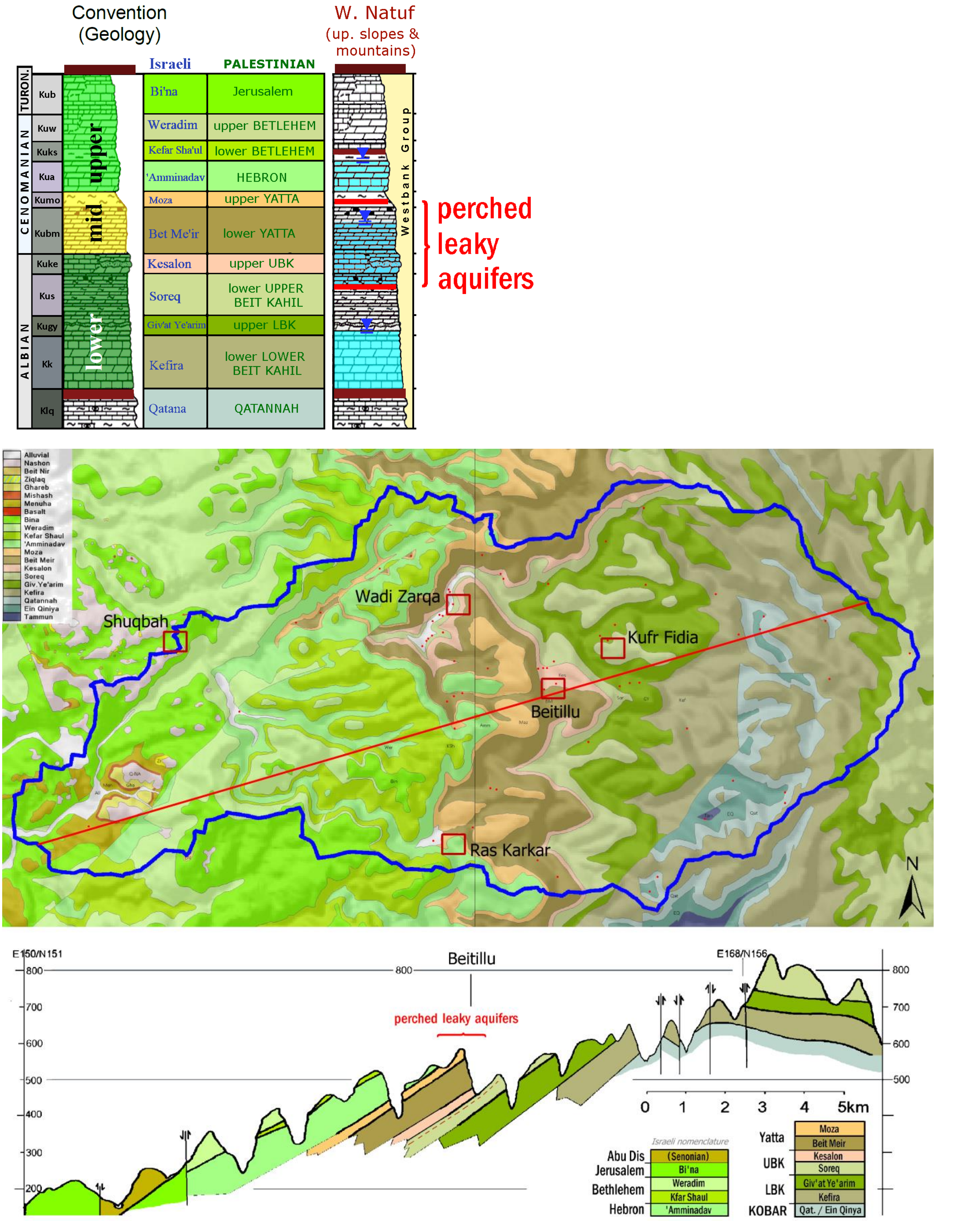

Figure 2: Stratigraphic column (2a) and geological relief map with cross section (2b). The stratigraphic column shows the regional hydrostratigraphy with only Lower and Upper Aquifers (left side); on the local scale, they can be differentiated several sub-units of aquiferal and aquitardal series with three separate perched aquifers in the middle (l-Yat, u-UBK and the top of l-UBK formations). The geological relief map indicates formation names of the Israeli nomenclature in the top left corner of the map (2a) and inside the boxes below the cross section beneath. Palestinian formation names, as used in this study, are on the side of the boxes. Red boxes inside the map indicate the five areas of the eight soil moisture stations. Sources – stratigraphic column: modified after Dafny (2010); geological map: modified after GSI (2000 and 2008); cross section: this study.

Conventionally, the WAB is subdivided into two main regional aquifer units of mostly karstified carbonates – the 'Lower Aquifer' of Lower Albian age (ca. 113-106 million years) and the 'Upper Aquifer' of Upper Cenomanian to Turonian age (ca. 95-89 million years) and cover around two thirds (64.4 %) of Wadi Natuf (Fig. 2a, Table 1). In between, the regional 'Middle Aquitard' can be further subdivided into an aquitard or even impermeable aquiclude section of yellow soft marl (u–Yat) and more carbonatic, and in parts karstified perched aquifer horizons (l–Yat) of local distribution that give rise to over 100 small local springs (Messerschmid *et al.*, 2003a, 2003b). As shown in Fig. 2a, stratigraphically they stretch between mixed series of carbonates, intercalated with chalk and marl and some chert (l-Yat) and rhythmic alterations of dm-thick dolomites interbedded with cm-thick marl layers (l-UBK formation), ending at its bottom with a conspicuous twin marl layer (see Fig. 3a). L-UBK is overlain by a cliff-forming very permeable reefal limestone formation (u-UBK, see Fig. 3b). While small contact springs (Fetter, 1994) emerge in large numbers from the isolated perched local aquifers, the recharge areas of the regional Upper and Lower Aquifers on the western flanks of the West Bank are almost entirely void of springs.

Figure 3. Lithostratigraphic formations. Twin marl band, located 2.5 km SE of Beitillu (a). It acts as confining layer beneath the top of lower UBK formation (top l-UBK). Cliff-forming coral reef limestone of upper UBK formation (u-UBK) at Wadi Zarqa with high primary porosity but also signs of karstification; the inlet photo shows a 1.5 cm thick remnant of coral branches (b). Colourful, thinly plated limestone of lower Betlehem formation (l-Bet) with fine marl intercalations (c).

Table 1: Outcrop (recharge) area, rainfall and classes of recharge potential of the perched aquifers and the main regional aquifers.

formation	u–Yat	l–Yat	u–UBK	l–UBK	"Upper Aquifer"*	"Lower Aquifer"*
area (km^2)	4.931	10.182	2.442	8.441	36.732	29.556
avg. rain 2003–10[1] (mcm/a)	2.917	6.145	1.502	5.262	20.696	18.452
class of recharge potential	V	IV	I, II	III	(I-III)	(I-III)

[1] average precipitation between 2003/04 and 2009/10; * "Upper Aquifer" stands for Hebron, Betlehem and Jerusalem formations, "Lower Aquifer" for upper and lower UBK formations.

The soil cover in Wadi Natuf consists of terra rossa, mostly with thicknesses in the decimetre range (Fig. A2, Appendix A). Although regional maps (Dan *et al.*, 1975) indicate small portions of rendzina in the central part (near Ras Karkar, Fig. 2b), no such types of soils were found during intensive field reconnaissance and measurement campaigns. The soils have high silt and clay contents and form desiccation cracks during the dry season disintegrating the soil into fist-size crumbs. During the early half of the winter season, these cracks can act as preferential pathways for infiltration, until the soils swell and cracks close during winter and early spring. Except for small plains over the aquitard series (like upper Yatta formation, see Fig. A1), no soils with thicknesses larger than one metre were found. Table 2 show the typical soil thickness for different formations in Wadi Natuf (see also Appendix F). An important and recurring observation during the field work was a pattern where soil thickness is associated to certain types of both, underlying bedrock lithology of formations and to the type of land use and land cover (LU/LC), where carbonates show thin soils, while over argillaceous rock, thicker soil covers develop. Equally, soil cover over grassland and terraces were found thin (see Table 2), whereas agricultural cultivated fields and garden plots were located over thick soil stratum, see Fig. A2. (This association of soil depths will be further discussed in the follow-up paper on regionalisation, Messerschmid *et al.*, 2019b).

Table 2: Soil moisture measurement data by location (in mm and %).

	Beitillu, gardens	Shuqbah, grassland	Wadi Zarqa, fields upper terrace	hothouse	Kufr Fidiah, field KF-W	KF-E	Ras Karkar, terraces RK-W	RK-E
# years measured	1	1	1	1	2	1	4	2
years measured	3/4	3/4	5/6	4/5	5/6–6/7	6/7	5/6–8/9	7/8–8/9
formation	l–Yat	Jerus	u–UBK (top l–UBK)		Top l–UBK		Heb	
soil depth (mm)	500	320	770	400	650	940	400	400
sensor depth (mm)	110/110/430	110/190	420/660/750	170/360	130/180/390	50/75/175	50/140/320	100/200
SM_{max} (mm)	179	60	172	132	253	237	129	129
SM_{min} (%)	8 %	10 %	6.5 %	16.1 %	10.5 %	0.5 %	4.2 %	4.2 %
SMm_{in} (mm)	40	20	50.05	64.4	68.25	4.7	16.8	16.8
θ^s (mm)	**139**	**40**	**121.5**	**67.5**	**185**	**232.5**	**112.5**	**112.5**
θ_{peak} (mm)	149	44	175	93	232	234	119	127

Note: θ_{peak} is a brief, episodically occurring available water storage above saturation storage capacity (θ^s); SM_{max} and SM_{max} are highest and lowest recurring soil water contents measured in the field (see also section 3.2.4).

3 Material and Methodology

As already mentioned in the introduction, this study presents a distributed recharge assessment method for a karst catchment based on hydrological field measurements (soils and springs) and parsimonious modelling (Fig. 4). In the following, the materials used and developed will be presented and the methodology described in more detail.

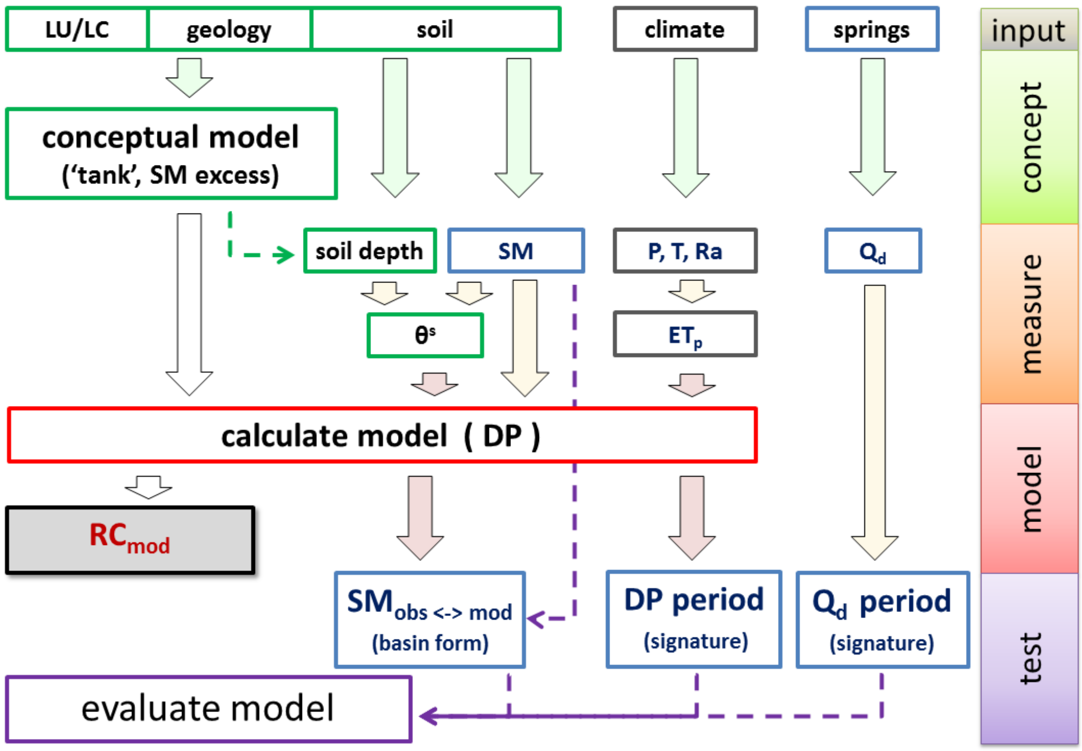

Figure 4: Conceptual flow chart of the work steps for the model. First row: measured and observed input such as climate, spring flow, land use and land cover (LU/LC), etc.; second row: conceptualisation of the one-reservoir 'tank' model; third part: actual measurements such as soil moisture (SM), precipitation (P), temperature (T), solar radiation (Ra), daily spring flow (Q_d), etc.; fourth step: modelling deep percolation (DP) and calculating recharge coefficients of the modelled formations (RC_{mod}); last part: model evaluation by different tests such as, quantitatively, comparing observed and modelled soil moisture levels (SM_{obs} and SM_{mod}), as well as semi-qualitatively by comparing signatures (periods) of percolation events (DP) and peak discharge at the daily read springs (Q_d).

3.1 Material

The material used and developed under this study belongs to three types - climatic drivers, like precipitation and evaporation, hydrological measurements such as soil moisture (SM) and five daily read springs, and finally physical features (soil depth, geology and LU/LC) to set up the recharge or deep percolation model (DP model).

3.1.1 Climatic drivers

The measurement period for precipitation and evaporation spans over nine seasons, from 2003/04 until 2011/12. A complete set of rainfall data was obtained from a previous study, originally measured in ten stations and equipped with automated rain gauges (tipping buckets) with Eijkelkamp loggers, run in event mode; the runoff analysis in this previous study (Messerschmid *et al.*, 2018) was based on five-minute steps. Here, the data were used in daily steps. The set of stations consists of three stations upstream and downstream, respectively and of four stations midstream (in central Wadi Natuf) with weighted area rainfall for the respective sub-catchments (Table 3).

Daily minimum and maximum temperature and solar radiation data for the calculation of potential evaporation with the formula of Hargreaves *et al.* (1985) - see section 3.2.2 were collected at two stations in Wadi Natuf; one station was installed upstream at Bir Zeit University at the mountain crest. Another downstream station at Shibteen well No. 4, was located at the transition between the foothills and Coastal Plain to the West. Missing records were filled up with corrected temperature data from Jerusalem (IMS, 2015), which represented intermediate values between the relatively cool conditions in Bir Zeit and the relatively warm conditions downstream in Shibteen. The readings were taken with an Automatic Weather Station (AWS) of the Type Campbell Scientific with a CR10X data logger and a HMP45D (Y3520067) Sensor.

3.1.2 Hydrological measurements

Soil moisture was monitored at eight different stations over a period of seven years. However, due to limitations in budget and field access (particularly severe during the height of the 2. Intifada), the measurements were kept at a minimum (in line with the recommendations by Beven, 2002 and by Seibert and Beven, 2009 – see section 1). Therefore, under a lack of sufficient measurement instruments and manpower to maintain and read the SM stations, it was decided to increase the number of different SM stations and to decrease the period of years each station was run. Of the eight stations, none was read throughout the entire period; instead, the number of measured years at a given station was in a range between one and four years. Table 2 documents the measurement period for each station. Only a maximum of three sets of equipment was available in parallel during any given moment. All in all, 13 years of soil moisture measurements were carried out at the eight locations. Some stations were only measured during one season, others during up to four seasons. The period of measurements at each respective station is documented in Table 2 (section 2).

Initially two sets of ML2-Theta probes (Delta-T Devices Ltd., Cambridge, GB) were used. Due to malfunction and vandalism they were replaced consecutively by two ECH2O loggers with five sensors (Decagon Devices, Inc., Pullman, USA). Each station was equipped with 2-3 sensors installed at different depths from 5 cm below surface down to 75 cm at the thickest soils (Table 2). Target of the soil moisture (SM) measurements were daily time series data of overall available water content (θ), measured originally in 30-minute steps, then averaged to daily values. SM was measured at different depths by each sensor and then normalised to total water content of the respective soil column.

In addition to the logging, hundreds of soil samples (150-550 g) were collected throughout the rainy seasons (with an emphasis during the first two seasons) and particularly after rainy events, both, in the surroundings of the SM-stations, as well as at unmonitored sites of different formations and mostly from surface-near, but also some from deep soil layers. Their water content was oven-weighed at Bir Zeit laboratory. The target here was to obtain control values to approve the reliability and representativeness of the selected soil stations. Technical problems with the equipment or interpretation of the data (e.g. exaggerated moisture readings) occurred during 8.09% of measured days, mostly during autumn, when the dried-up soils were picking up moisture (and once in spring, after the recharge period). The statistics and possible reasons of these misreadings are discussed in Appendix B (Tables B1-B3). No equipment problems occurred during the periods of intensive precipitation and recharge.

The second set of hydrological field work was the daily spring flow measurement. Taking into account practical considerations such as field accessibility and minimisation of costs and efforts, five springs were selected, which were then continuously measured throughout almost the entire period and on a daily basis (spring data coverage see, Table C1). All selected springs lie near the SM measurement stations and issue from the perched aquifers recharged at the SM plots with very short travel times (see section 4.1.4). Three of these springs belong to the so-called 'Beitillu spring group'; the two other springs to the 'Wadi Zarqa spring group' (see Fig. 2b). The measurements were carried out by hand with stop watch and bucket (an initial installation of water meters proved unreliable). The target of the spring readings was to establish, for the first time in the WAB Mountains, a complete spring hydrograph of high temporal resolution in order to detect the signatures, i.e. the temporal patterns of low and peak discharge for comparison with model results of recharge. The procedure of these calculations is detailed under section 3.2.

3.1.3 Physical features

The mapping, measurement and recording of physical features was field-based, as well as based on previous work in the literature. The soil in Wadi Natuf was found to consist only of one prevailing soil type, terra rossa (see section 2); field investigations, sampling and granulometric lab tests showed that most of these soils are silty-clayey, with some samples also showing higher clay contents (Messerschmid *et al.*, 2018). This allowed for an important simplification of the requirements for the DP model, which only accounted for variable soil depth but a uniform soil type throughout the study area. The focus of the field work hence was laid on the spatial distribution of soil thickness (b), by digging up the soil down to the soil-rock interface. In many places, a transition zone between soil and bedrock was encountered, with rock content gradually increasing downward. Here, digging usually stopped when in the transition zone the volumetric portion of rock fragments outweighed the soil content. This depth was then taken as the depth of the soil-rock contact horizon. Since soil thickness (b) is a quantifiable parameter and plays a crucial role in the soil moisture - percolation model, much effort was spent on probing the soil depth at many sites for each outcropping formation and on the establishment of a soil thickness matrix (Table D1), where the distribution of soil depth is documented for different LU/LC-types and different lithostratigraphic units (see also Table 2). The target was to create a simplified but realistic categorization of different geological formations and LU/LC-types representing typical soil thicknesses (compare also Messerschmid, *et al.*, 2019b).

3.2 Methodology

3.2.1 Rainfall

In order to estimate average recharge coefficients, it is necessary to first verify that the observation period represents the normal range of long-term annual rainfall variations ($\geq$30 years). Since no 30-year record for Wadi Natuf is available, we based our examination on the documented area rainfall of the WAB. The calculations are documented in Table E1 and Fig. E1. As a result, we can state with confidence that the seven years covered in this study happen to match very closely (102%) with the overall range of long-term annual rain heights in the WAB.

3.2.2 Evapotranspiration

As already mentioned, the sensitive parameter of actual evapotranspiration (ET_a) was approached through calculations of potential evapotranspiration (ET_p), based on the Hargreaves formula. This approach had already been applied by Schmidt *et al.* (2014) on the Eastern Mountains and Slopes of the West Bank, near Wadi Natuf. The use of the Hargreaves-formula and the equation of actual evapotranspiration with potential evapotranspiration during the winter season is based on the assumption that, as shown by Ryu *et al.* (2008) in a semi-arid grassland and very similar climate in California, for most of the cold and wet winter months, i.e. the '*energy limited period*' and during the initial spring period, ET_a is nearly equal to potential evapotranspiration (ET_p), especially in fine-grained soils such as the silty terra rossa soils in Wadi Natuf, which retain moisture for longer periods (Rushton *et al.*, 2006). ET_p is calculated by the equation:

$$ET_p = 0.0023 * Ra * (T_{mean} + 17.8) * (T_{max} - T_{min})^{0.5} * \lambda^{-1} \quad (1)$$

where Ra is extra-terrestrial solar radiation (MJ d^{-1}), T_{mean} is the mean air temperature during the respective time interval (calculated as average of daily minimum and maximum temperature, in ^{0}C), $T_{max} - T_{min}$ is the daily temperature range, and λ is the latent heat of vaporisation in order to obtain ET_p in units of mm d^{-1}.

3.2.3 Springs

The hydrographs of the daily read springs allowed us to identify periods of rise, peak and decline in spring flow for each year and each spring (or spring group) individually and in high temporal resolution (daily steps) for the sake of comparison with model results. The representativeness of the flow pattern of the springs was tested and is presented in Table F1. These periods are also referred to as basin signatures (Eder *et al.*, 2003 and Hrachowitz *et al.*, 2013). Winsemius *et al.* (2009) suggested the use of combinations of quantitative and qualitative information from the local or basin scale. While peak flow and recharge periods can act as semi-qualitative information (signatures), quantitative information on physical parameters such as SM) was available from the SM recordings, as discussed in the next section.

3.2.4 Soil moisture (SM)

The sensor readings of available SM contents from different soil depths were added up to an overall water content of the soil column (as m^3/m^3, % or area normalised to mm). The continuous soil hydrograph was further analysed and for each location separately, a typical and annually returning minimum (SM_{min}) and maximum water content (SM_{max}) was found. These values were temporally stable parameters but differed spatially. The results were also confirmed by the many hand collected samples, tested for soil moisture from the same formations but at other locations. The maximum soil water storage capacity or storage capacity at saturation (θ^s) was then calculated for each SM station and according to the formula:

$$\theta^s = ((SM_{max,1} - SM_{min,1}) * b_1) + ((SM_{max,2} - SM_{min,2}) * b_2) + ((SM_{max,3} - SM_{min,3}) * b_3) \quad (2)$$

whereby SM_{max} and SM_{min} were used as measured water content per soil segment volume (m^3/m^3), $SM_{max,1}$ indicates the maximum water content at the first surface-near soil layer and deeper layers are indicated by successive numbers (2 and 3), and where b indicates the soil thickness (m) of the respective soil layers (1, 2, 3). Over a unit area, the cumulative θ^s therefore is the normalised effective maximum storage content in m (or mm) of the whole soil column (see Sheffer, 2009 and Sheffer *et al.*, 2010). The field-read moisture contents included brief periods of water storage beyond the field capacity, such as water in desiccation cracks (especially in heavy storm events during the late autumn and early winter season); in the hydrograph, these peak levels can be noticed as brief overshots of SM-readings (here indicated as SM_{peak}) above the otherwise stable levels of SM_{max}. It should however be emphasized here that these additional water amounts are fully accounted for in the daily budgets of the percolation model (see sections 4.1.1 and 5.3).

3.2.5 Parsimonious soil moisture saturation and percolation modelling – design and processes

The model implies a classical soil moisture balance approach; soils dry up to minimum water content during summer and then accumulate and store water with successive rainfall during autumn and winter. SM content above SM_{min} is subject to direct evaporation and plant transpiration and has to be deducted in daily steps from the SM content accumulated successively by precipitation. When soil conditions reach saturation (observed SM_{max} or full effective storage capacity θ^s), deep percolation (DP) into the bedrock is triggered, here equated with groundwater recharge (Fig. 5). When the shallow soils of Wadi Natuf were sufficiently saturated, the model equated actual evapotranspiration with potential evapotranspiration; but when the available water content (above SM_{min}) fell below daily potential evapotranspiration ET_p (mm), actual evapotranspiration ET_a (mm) was limited by water availability. The model thus follows the function:

$$ETa_{i+1} = \begin{cases} ETp_{i+1} & \text{if } (\theta_i + P_{i+1}) \geq ETp_{i+1} \\ \theta_i + P_{i+1} & \text{if } (\theta_i + P_{i+1}) < ETp_{i+1} \end{cases} \quad (3)$$

where θ_i is available soil moisture (mm), P daily rainfall (mm), ET_a daily actual evapotranspiration (mm) and ET_p daily potential evapotranspiration (mm).

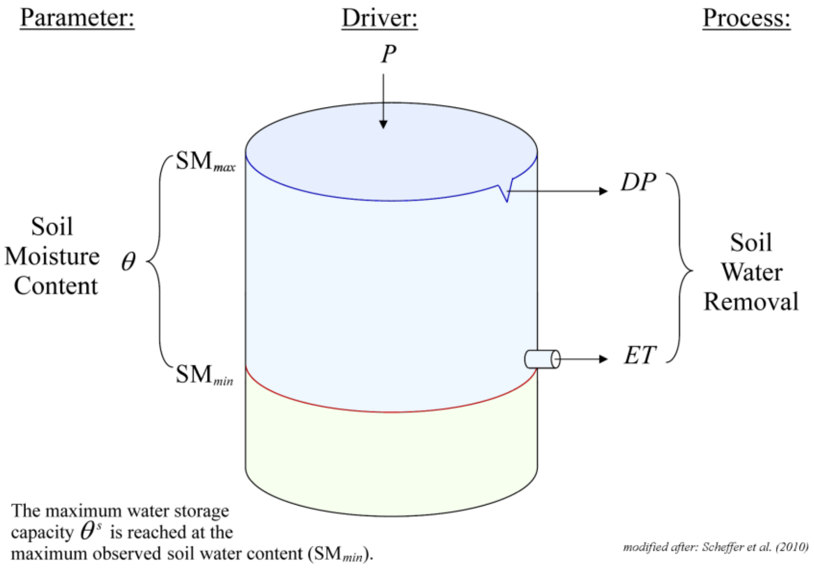

Figure 5: Parsimonious SM-percolation model of '*one reservoir saturation excess*' (modified after Sheffer *et al.*, 2010)**.** Parameters and drivers (top) were field-measured for every station and model. The available water content, shown as θ (between SM_{max} and SM_{min}), is seasonally variable but location-specific. Deep percolation (DP) or recharge is triggered above θ^s. Precipitation (P) was measured; potential evapotranspiration was calculated from weather station data and transformed into actual ET values (see section 3.2.2).

Secondly, all additional rainfall infiltrating from the surface and beyond the daily evapotranspiration losses can either be added to the available water storage, or, when limits of θ^s are exceeded, is considered to percolate into the bedrock and represent recharge:

$$DP_{i+1} = \begin{cases} 0 & \text{if } (\theta_i + P_{i+1} - ETa_{i+1}) \leq \theta_{max} \\ \theta_i + P_{i+1} - ETa_{i+1} - \theta_{max} & \text{if } (\theta_i + P_{i+1} - ETa_{i+1}) > \theta_{max} \end{cases} \quad (4)$$

$$\theta_{i+1} = \theta_i + P_{i+1} - ETa_{i+1} - DP_{i+1} \quad (5)$$

where θ_{max} is the effective maximum soil water storage and DP is daily deep percolation, i.e. groundwater recharges (mm). The testing of the model was reached in two independent ways, by a quantitative comparison of modelled and observed SM values, and by a semi-qualitative comparison of recharge periods (DP events) with times of observed spring discharge peaks;

in other words, one examination was made on the reliability of the model to produce the observed physical feature of SM and the other by using the hydrologic response signature of spring flow responding to recharge.

3.2.6 Annual and mean recharge coefficients

In line with the PUB recommendations, we again kept it simple and used the individual annual rainfall coefficients (RC) for a direct, somewhat extrapolative calculation of long-term average RCs. First, the annual DP-rates (mm/a) divided by annual station rainfall delivered recharge coefficients (RC, in %) for each modelled station; secondly, DP-rates from the model were referred to the outcrop sizes of the respective formations in Wadi Natuf to obtain annual recharge amounts (R, in m^3/a). Then, at each station (or for each formation) the seven different annual RC-values were transformed into an average recharge coefficient (RC_{avg}). This approach is based on two assumptions, i.e. that our seven year rain period is a fair representation of long-term averages of both inter-annual and seasonal distribution of precipitation. The long-term inter-annual rainfall distribution is also discussed in Appendix E. The seasonal distribution however, cannot be sufficiently addressed in this study on spatial recharge distribution. It remains therefore an approximation that remains difficult to be narrowed down and be truly reflective of all scenarios of an event resolution of long-term precipitation and recharge processes. However, it can already be stated here that it appears not very likely that a true reflection of seasonal distribution would lead to a significant change in overall annual average recharge coefficients.

4 Results

4.1 Soil Moisture modelling

4.1.1 Seasonal soil moisture patterns

Table 2 presents the eight locations of soil measurements, used for the percolation models, with their respective bedrock formations and soil depths as well as the recurring maximum and minimum soil moisture levels, from which the maximum water storage capacity (θ^s) of each location was derived.

In summer, soil usually dried up to a SM content of somewhere between 5 and 10 percent. (Two exceptions are found here: In one case, the plot near the hothouses in Wadi Zarqa (Zar-hh), some summer irrigation of crops cannot be excluded. However, it was assured that no irrigation took place during the crucial winter modelling period. In the second case, the extremely low moisture content in the bushy landscape near Kufr Fidiah (KF-E) was a result of vandalism. In any case, such low soil moisture is not representative and the data of this station had to be taken with caution. In addition, the soil depth encountered at KF-E, with almost one metre, is a rare exception, and according to our soil depth survey not representative for usual conditions at the top of l-UBK formation (compare soil depth matrix, Table D1). Ultimately, the results of KF-E were excluded from the regionalisation analysis (see sections 4.1.3 and 4.1.4).

In contrast to minimum recorded SM-values, maximum moisture levels showed a more consistent pattern, which is important, because soil saturation is crucial for the percolation model. In line with our saturation excess model, it should be noted that once this maximum soil moisture level was reached, it remained stable for days and weeks and did not rise any further. Throughout the measured period and at almost all stations, it was found that maximum soil moisture is spatially highly variable and can be associated with the local soil depth, as apparent from Table 2 and in Fig. 6, and as will be discussed below. Another special case however is the occurrence of brief periods with extreme moisture (SM_{peak}) above saturation storage capacity (θ^s). They only occurred during extreme storms or during the early winter months, when desiccation cracks still enabled preferential flow and rapid infiltration into deeper soil layers (see also section 5.3. and discussion of Fig. 7) and lasted usually less than one day.

Temporal (seasonal) variation of recharge did not stand at the focus of this study. However, some brief remarks on the results of the daily percolation model shall be added here. At locations with thinner soil cover, such Ras Karkar (RK), Wadi Zarqa hothouses (WZ-hh) and especially in Shuqbah (Shu), full saturation (θ^s) was usually reached by mid or late November. In the deeper soils (such as in Kufr Fidiah and Wadi Zarqa, upper terrace, or WZ-upT) saturation was usually reached later, by mid-December. In most years, full saturation conditions prevailed during January and February, on rare occasions until April. The season 2005/06 experienced the most intensive rain event of the entire measurement period and at the same time occurring very late in winter, between 1st and 5th April 2006. Figure 7 gives an example of the hydrographs of observed and modelled soil moisture (station RK-W). As can be seen, recharge (deep percolation, DP) only occurs when full saturation is reached and in form of relatively distinct events of a few days per year, depending on the high-resolution temporal rain distribution of a given season.

4.1.2 Annual soil moisture and recharge calculations

Although in each year the hydrograph of calculated soil moisture was the result of a unique set of different local parameters and different physical features (and thus of different overlapping processes), certain common patterns could be determined: The levels of available and effective soil moisture, i.e. the water content above the minimum observed soil moisture (SM_{min}) usually dropped to zero within one or two months after the last major rainfall event. The hydrographs of course were location-specific, which means, formation-specific. But they were also temporally variable. Instead of a single seasonal rainfall-percolation threshold over the years, we rather found accentuated individual patterns of soil moisture accumulation durations for every rainy season, depending on the respective temporal distribution of precipitation and evaporation, as shown in Fig. 6. In this diagram, the years on the x-axis are sorted by precipitation depth from left to right and from the driest to the wettest seasons. Unlike average areal rainfall, the annual recharge does not rise monotonously from left (dry) to right (wet years). Again, this demonstrates the event character of recharge and the importance of daily time step calculations. Other implications of Fig. 6 are discussed in section 5.2.

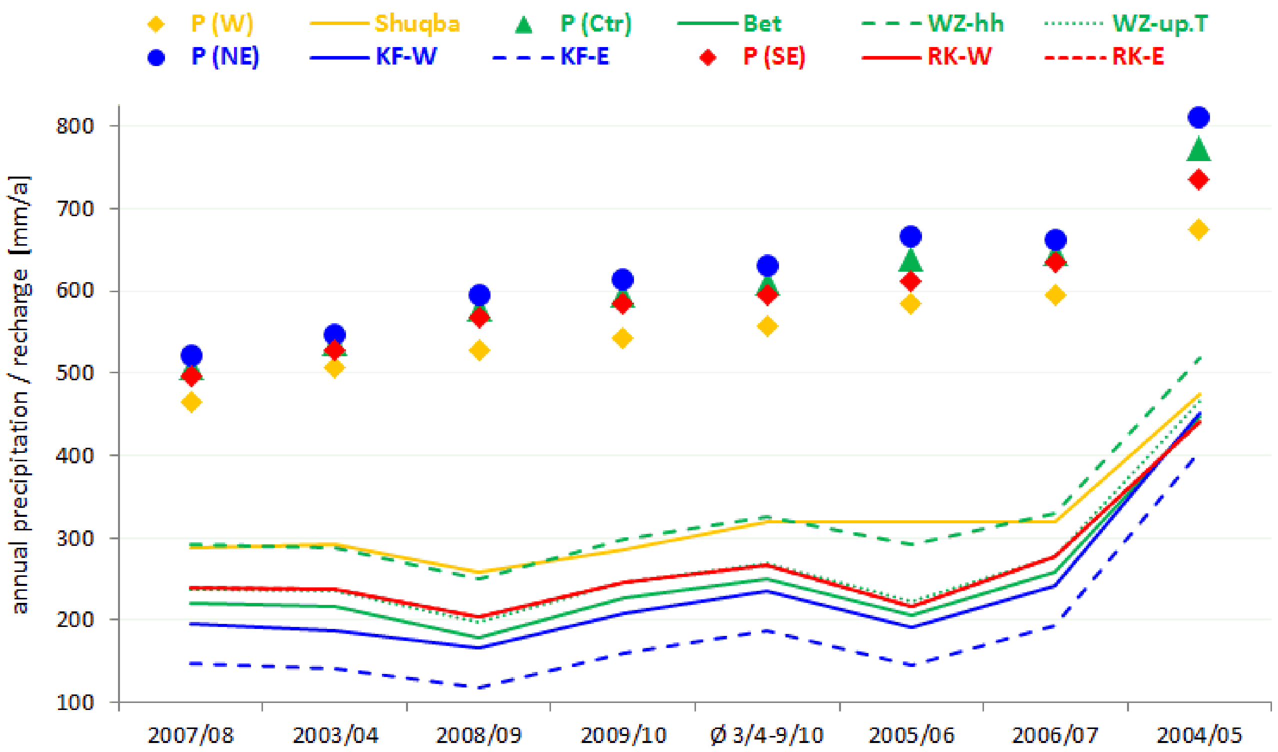

Figure 6: Annual percolation (DP) and precipitation (P) rates for 8 SM stations and 4 respective precipitation sub-catchments. The years (x-axis) are ordered by annual rain levels. Towards the right, spatial variability of P increases, but that of DP diminishes (see sections 4 and 5 and in Appendix G).

4.1.3 Quantitative examination of the model (by physical parameters)

Figure 7 also allows checking the soil moisture percolation model. Testing the model results was performed on two independent levels, first quantitatively on the level of physical features (basin form), e.g. observed versus modelled soil moisture (Fig. 7), and second qualitatively on the level of basin response, e.g. periods of deep percolation or recharge, compared with the periods of peak spring flow (Fig. 8).

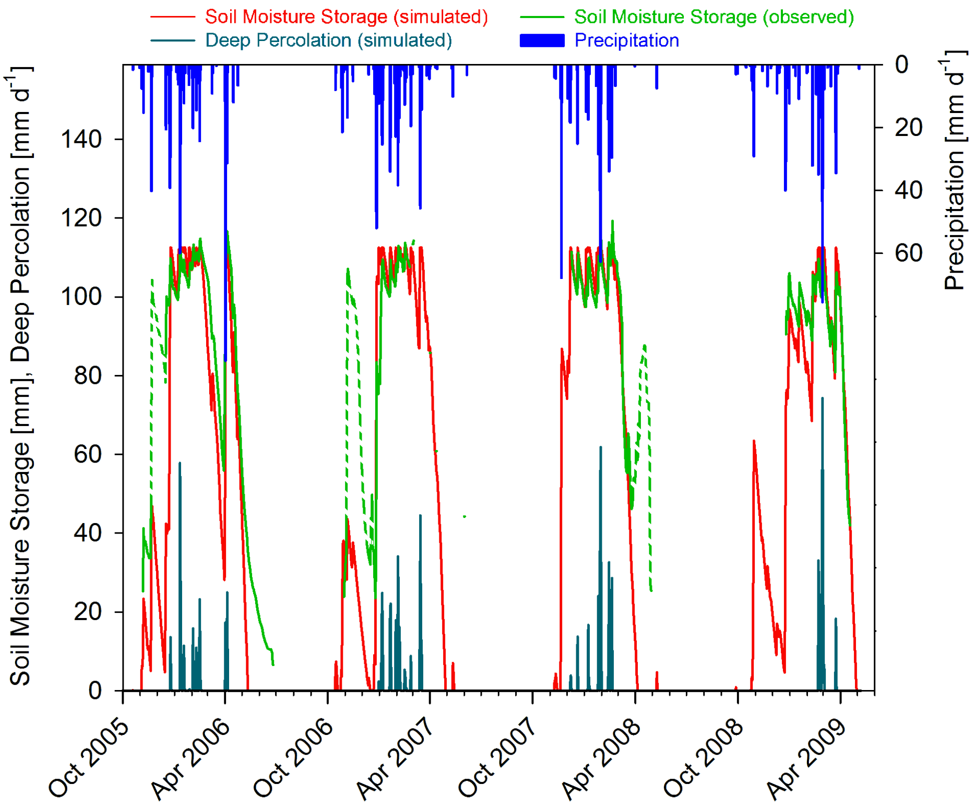

Figure 7: Comparison between observed and modelled available soil moisture at Ras Karkar West (RK-W), 2005/06-2008/09. Maximum water storage capacity (θ^s, at 112.5 mm) of the modelled moisture level (red) matched well with the observed maximum moisture (green). Brief periods of dotted green lines show questionable and erroneous soil moisture readings (see also section 5 and Appendix B). Deep percolation (DP, triggered above θ^s) is indicated as dark blue column (bottom of the graph), daily rainfall as blue columns (at the top).

Some of the readings indicate a mismatch between observed and modelled SM values; these periods are indicated as dotted green line, where recorded soil moisture exceeded modelled SM-values by far. In some cases, recorded moisture levels even surpassed accumulated seasonal rain (which clearly hinted to equipment failure). Both issues shall be further discussed in section 5. Appendix B also presents some statistical analysis on this issue. However, this statistical analysis on the comparison between observed and modelled soil moisture (Fig. 7) resulted in a very good to excellent correlation, with an average Nash-Sutcliffe efficiency correlation of 0.73 for all stations, except RK-E (which had a negative correlation of -0.35 and was excluded from further analysis). In Shuqbah and Beitillu, NSE was found between 0.8 and 0.87, in Wadi Zarqa at 0.63, at KF-W and KF-E between 0.79 and 0.96 and at RK-W station at 0.7 (Table B3). These results include all periods, dry and wet alike. For winter values alone, where soil moisture lied above the minimum water content (SM_{min}), the NSE was found slightly lower at 0.69.

4.1.4 Semi-qualitative examination of the model (by peak spring flow signatures)

Since basin form does not necessarily translate linearly into basin function (see section 1.2), the model was additionally examined on the level of basin response: We compared periods of deep percolation (DP) with the signature periods of peak spring discharge, as shown in Fig. 8.

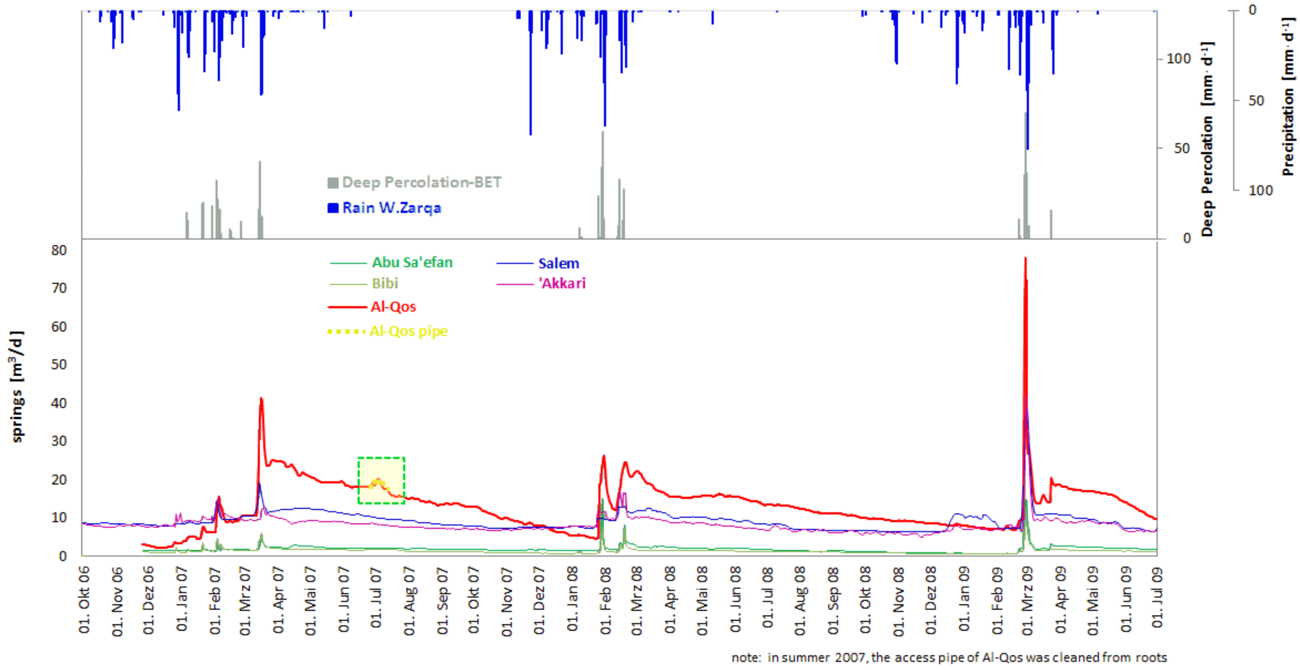

Figure 8: Signatures of peak spring flow and DP-events. Yellow shaded box: temporary increase in spring discharge, caused by cleaning of the access pipe from plant roots at Al-Qos spring (July 2007). Daily read springs at Beitillu spring group in red, olive and green colours (bottom); at Wadi Zarqa group in blue and purple colours. Daily deep percolation (DP) as grey columns (upper half), precipitation as blue columns (top).

Here, a relatively uniform overall pattern was observable; in most stations, recharge occurred during a period between 11 and 21 days per year, during very wet winters a maximum of up to 31 days per year (e.g. in WZ-hh); during very dry winters a minimum of only five days per year (e.g. KF-E) were recorded. Rainfall thresholds that trigger DP vary strongly between the different years and especially stations, e.g. from 74 mm of accumulated rainfall (Shu) as a minimum to 470 mm (KF-E) as the highest threshold (Fig. 9). For the different years, the gaps between station thresholds ranged between 206 mm (in 2004/05) and 396 mm (in 2008/09), with a multi-annual average of 290 mm.

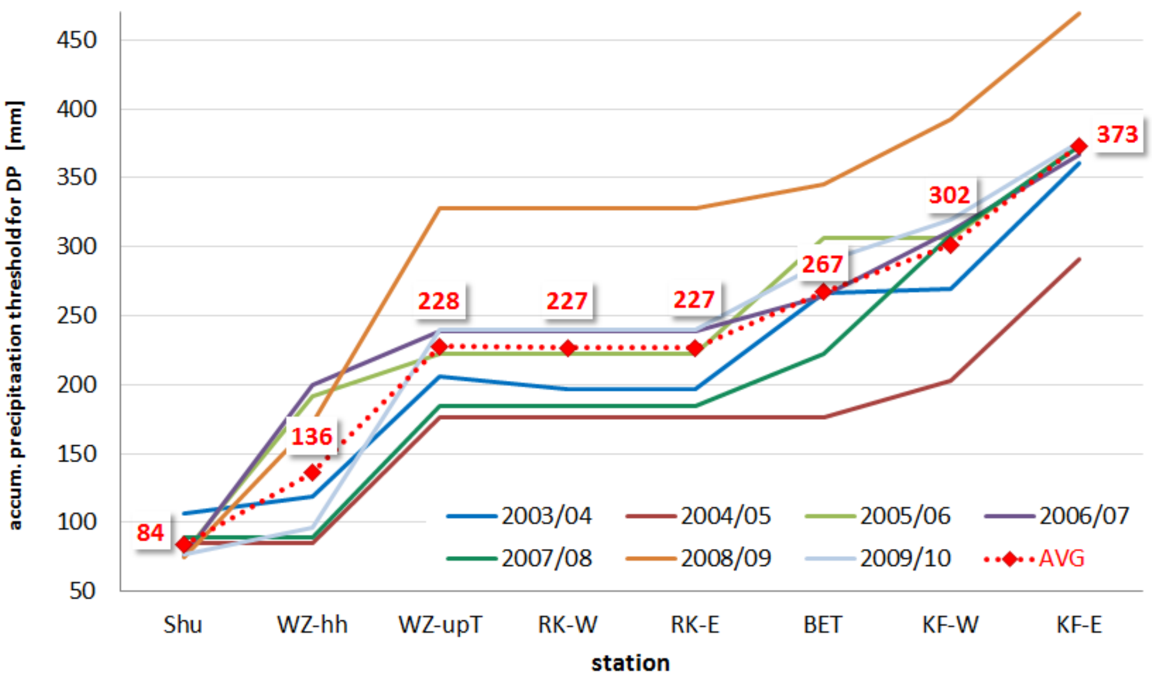

Figure 9: DP thresholds for different years and stations. Preceding accumulated seasonal precipitation to trigger DP shown in mm.

Figure 8 also exhibits a very close match between the dates: Observed spring flow in the local perched aquifers picks up with almost no delay, usually within one day after the first modelled recharge event (again, the representativeness of the springs was tested, see Appendix F). This demonstrates both, the karstic nature of the aquifers with rapid flow connections, as well as the very local recharge conditions on the isolated hillsides.

The most productive springs of the respective groups are Salem (Wadi Zarqa spring group) and Al-Qos (Beitillu spring group). As can be seen in Fig. 8, the average of spring flow was lying between 1.3 (Bibi) and 12.4 (Qos) m^3/d. The weakest spring (Bibi) had a maximum discharge of 9 m^3/d, the strongest spring (Al-Qos) of 78 m^3/d, i.e. a maximum of ca. six times or 600% of the annual average levels of daily discharges.

As our semi-qualitative analysis revealed, every slight rise of observed daily spring flow (even very slight increases, by only 8% the amount of the average discharge rates) was connected to our modelled deep percolation or recharge event into the perched aquifers. Comparing the dates of peak spring flow with that of DP events, a complete congruence of the respective dates is apparent. (This analysis was performed for the period until September 2009, during which, all springs had been read reliably, see Table C1). All days with more than 1 m^3/d of increase in discharge at Al-Qos spring (= 8% of average Q) were also found to show percolation in our model. In Salem spring, the strongest spring in Wadi Zarqa spring group, all days with over 0.7 m^3/d rise (= 7.95% of average Q) were recorded during the periods, where the model indicated deep percolation. For Akkari spring, all days with an increase larger than 0.66 m^3/d (= 8.3% of average Q) showed modelled recharge at the respective soil moisture model station in Wadi Zarqa.

Therefore, it can be concluded that the check of the model on two independent levels was very good (see also discussion, section 5); the two methods of examination were a) quantitative on the level of physical basin form (soil moisture levels) and b) semi-qualitative on the level of basin signature, i.e. by periods of DP and peak spring flow. In addition, our resulting recharge coefficients were compared to literature values, see Messerschmid *et al.*, 2019b). The results of the model are summarised in Table 3. This table shows the annual DP-rates, the recharge coefficients (RC) for each year and the average of these RC-values (between 30% and 57%) for the entire period, together with the soil thickness at each station, ranging between 30 cm and 90 cm.

Table 3: Annual recharge rates (DP), modelled at representative SM locations.

year	rainfall area avg.	DP Bet	DP Shuqbah	DP WZ-hh	DP WZ-upT	DP KF-W	DP KF-E	DP RK-W	DP RK-E
2003/04	529	217	291	289	235	188	140	238	238
2004/05	742	448	475	519	465	452	405	442	442
2005/06	621	206	319	293	223	192	145	218	218
2006/07	632	259	319	331	277	241	194	278	278
2007/08	496	221	287	293	239	196	149	240	240
2008/09	565	180	258	251	197	165	118	204	204
2009/10	582	228	286	299	245	208	160	247	247
average DP	–	251	319	325	269	235	187	266	266
avg. P (‘3/4–’9/10)	–	601	557	601		631		588	
P – sub-catchment		avg. WZ/S1	avg. Nea/S1	avg. WZ/S1		Wadi Zarqa		avg. Ayb/S2	
DP_{avg} / P_{avg}	**–**	**42 %**	**57 %**	**54 %**	**45 %**	**37 %**	**30 %**	**45 %**	**45 %**

Note: For reasons of overview, this table shows only the average area rainfall of entire Wadi Natuf (595 mm/a between 2003/04 and 2009/10). However, the daily rates of deep percolation (DP) were modelled with the respective sub-catchment precipitation. Area averages for the five sub-catchments of Wadi Natuf (Ne’alin, Shibteen 1, Shibteen 2, Wadi Zarqa and Ain Ayoub) were taken from Messerschmid *et al.* (2018). All units are in mm/a, except for the recharge coefficients (expressed as DP_{avg}/P_{avg}, in %).

4.1.5 Summary of model results on annual levels

In sum, for most years, a rise in annual recharge coincides with a rise in total annual rainfall. The wettest winter (2004/05 with a rainfall of 742 mm/a) showed the highest recharge in each of the stations. However, the driest year (2007/08 with 496 mm/a) did not generate the lowest recharge. This demonstrates the necessity to first calculate and model recharge in daily time steps in order to reflect the event character, i.e. the temporal distribution of the thunderstorm (Sheffer *et al.*, 2010, Cheng *et al.*, 2017), before these results can then be used for annual recharge calculations. It is also important to note the strictly reciprocal relationship between soil thicknesses and recharge coefficients. At all stations except for one (WZ-upT), the thinnest soils showed the highest recharge and the thickest soils the lowest recharge. This general pattern confirms our conceptual approach for a basin classification framework, grounded on a close link between basin form (soil depth) and basin response (recharge) – different classes of recharge potential are indicated in Table 1, as will be further discussed in the follow-up article by Messerschmid *et al.* (2019b) on recharge regionalisation for the entire Natuf catchment.

The diagram in Fig. 10 plots different measurement stations on the x-axis against the respective soil depths and annual recharge rates on the y-axis with a marked and almost consistent linkage between the latter two (except for WZ-upT). This pattern of association forms the basis of our extrapolation and attribution approach for entire formations and their recharge coefficients (for the representativeness of the soil depths at our SM stations, see Table D1).

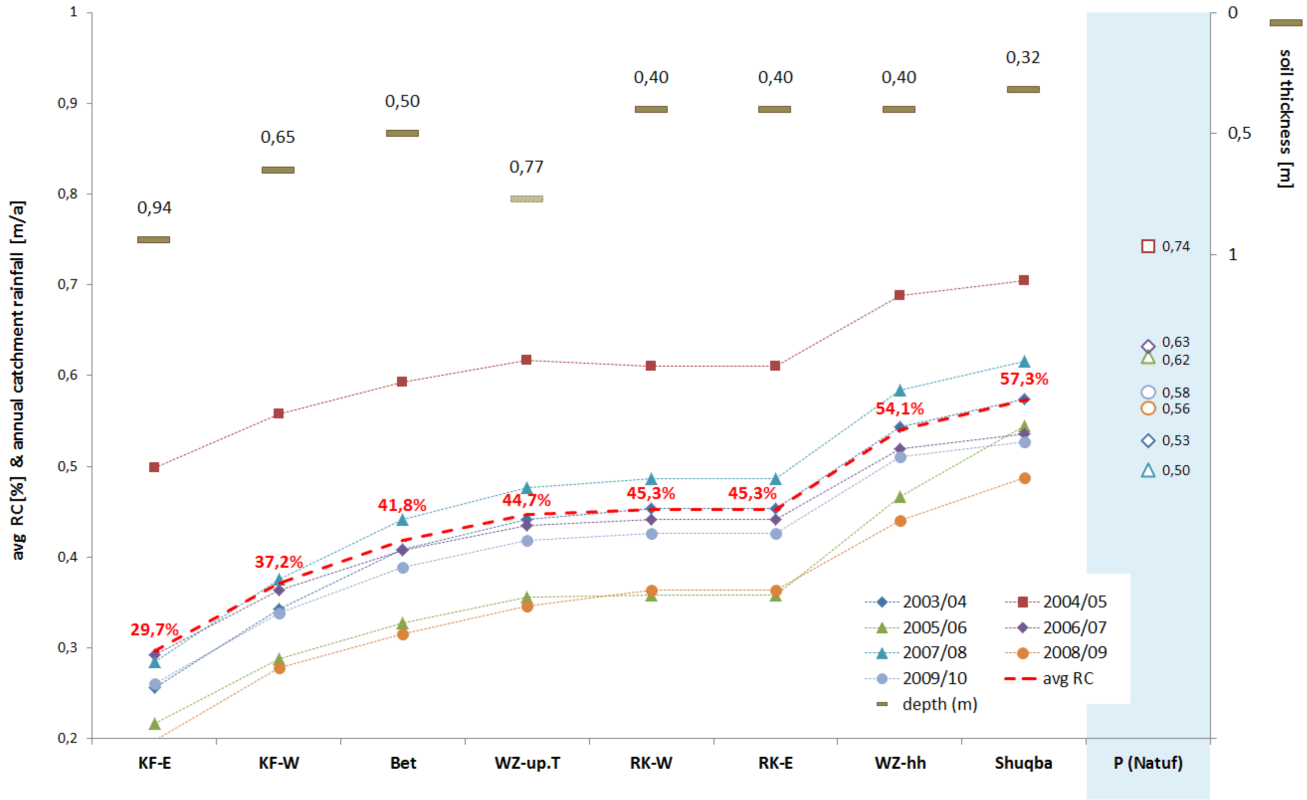

Figure 10: Annual recharge coefficients, annual precipitation and soil depth (stations ordered by increasing RC). The coloured lines present annual recharge coefficients by station (x-axis), increasing from left to right (in %). Average annual RC-values (red numbers) are reciprocal to the respective soil depths (as brown bars, with exception of WZ-upT as beige bar). The blue-shaded column at the right indicates annual precipitation - shown as P (Natuf) - as area average (in m/a).

5 Discussion

5.1 General approach of process representation

Our results clearly demonstrate that even in ungauged basins a realistic model of distributed groundwater recharge can be obtained through limited field observations of key parameters of the complex percolation process. Hereby, two points can be noted. The direct and location-specific assessment of water storage capacity from long-term soil moisture readings is a new approach in the determination of recharge of the WAB. Previous studies only assumed maximum storage capacity by drawing on the general literature but without sufficient 'local knowledge', i.e. site-specific field evidence. Instead they manly concentrated on retro-fitting the models through repeated calibration runs. This is particularly important because the rate of modelled percolation (recharge) is a direct function of SM storage, which is highly location-specific. This highlights the importance of robust and realistic input data. Our values were simple but field-based and therefore realistic findings. In addition, the literature often only provides values for permanent wilting points (pwp) and field capacities (Fc), which are similar but not identical with our measured maximum and minimum SM values. The θ^s-value is a direct indication of the mobile, available water inside the soil column, subject to accumulation, evapotranspiration, saturation and deep percolation (recharge) and without the need for further assumptions. As a second issue, it shall be noted here that in reality, the processes of soil moisture accumulation, deficit and saturation are complex and depth-dependent. In fact, our sensors installed at different depths indicated slightly different hydrographs of temporal distribution, i.e., the deeper soil segments picked up moisture content later in the season, showed the arrival at a relatively stable maximum plateau value at about the same times as the shallower sensors but then also dried up at a slower pace in summer. This is in line with the generally known and expected dominant processes of accumulation, storage and loss through evaporation and transpiration from roots. It should be added that most of our soils were only about half a metre thick or less and therefore, even the lower soil segments lie within the reach of many of the typical plants (trees, bushes and shrubs). Our calculation method of adding up the measured water contents for each layer separately therefore accounts for the different conditions and processes at varying soil depths and by this provides a simplified, yet realistic representation of the overlay of all depth-dependent soil processes in total.

In addition and in line with the findings of the PUB decade, the intention of this study was to adhere to the goal of parsimony. Our simple SM saturation and percolation model therefore only accounted for two climatic drivers (P & ET_p) and one spatially distributed physical parameter (θ^s), each set-up in daily time steps. The model is a simple 'tank' or 'combined reservoir' model, as used by Sheffer *et al.* (2010) and Schmidt *et al.* (2014) – see Fig. 5. The applicability of this modelling approach is based on three conceptual but field-observed assumptions of dominant processes. Firstly, the model only accounts for direct recharge (as DP) from in-situ soil infiltration. Whereas, some local lateral water movement on and inside the soils and within the range of a few metres distance has to be expected, this nonetheless is conceptually included here under local processes and therefore quantitatively accounted for as direct recharge. This particularly applies to soil pockets in formations with more developed epikarst (see Lange *et al.*, 2003). Otherwise, as discussed in section 2, no signs of indirect recharge processes were observed in the field (with one notable exception). Ponding as a sign of lateral surface accumulation was restricted to small puddles of a few metres in diameter at maximum and within clearly discernible very small topographic depressions. Lange *et al.* (2003) in their runoff generation sprinkler and tracer tests also observed such surface near rainwater regrouping, mobilisation and accumulation and found them to occur in a radius of decimetres, at maximum metres (see also Ries *et al.*, 2015). The notable exception where indirect recharge occurs concerns the second process to be discussed here – the transmission loss of surface runoff through Wadi gravel beds. Indeed, this transmission loss could frequently be observed to reach 100% of runoff, with Wadis found dry only a few kilometres downstream of an observed runoff event site in real time. However, as shown in Messerschmid *et al.* (2018), overall runoff generation in Wadi Natuf was remarkably low at only 1% of annual area precipitation. Secondly, for recharge calculations, runoff can therefore be neglected as a significant part of the water budget (and under the PUB goal of simplification and parsimony). Thirdly, our model accounts for only one dominant recharge generation process – deep percolation from the soil into the bedrock under SM saturation excess. This is in line with the already discussed findings from other studies in the WAB and adjacent basins that deep percolation or recharge, as much as runoff, must be described as a process of soil moisture saturation excess (see sections 1.3.2 and 1.3.3). For further discussions on the applicability of soil moisture balance models in the WAB, see Hughes *et al.*, (2005 and 2008).

5.2 Spatial and temporal variability of precipitation and recharge

The focus of our approach was clearly the spatial differentiation, although our model also considered the temporal distribution of recharge. Our annual coefficients were based on a model, which was run in daily steps and therefore fully accounts for the temporal distribution of a storm event for groundwater recharge calculation. Further analysis of the spatial recharge distribution is provided in Messerschmid *et al.*, (2019b). The model results of annual recharge coefficients (see Fig. 6) indicate an unobserved relationship between spatial and temporal variability. If the standard deviations for annual recharge and annual precipitation are plotted against each other, we can find a clear relation for each year of modelling (except for the year 2005/06), where in wetter winters with larger standard deviations for rainfall, the spatial variability of recharge diminishes (see Fig. G1). Although seven years of modelling may not be enough to draw far-reaching conclusions, we can still conclude that the question of spatial variability of recharge deserves greater attention than given in most recharge studies, particularly in the Western Aquifer Basin.

5.3 Some inconsistencies between modelled and observed SM-values

Another point to discuss is the deviation of observed SM from the modelled SM in several of the stations, particularly during the beginning of the measurement seasons, before the first large rainfall and recharge events. Some of the SM-levels obtained during measurements were unreasonably high and at times reached or even surpassed cumulative rainfall. A deviation of SM_{obs} from SM_{mod} can occur for three reasons: either because the model is inadequate (1) or because the interpretation of recorded field data is complicated by additional factors (2). The third possibility (3) is malfunctioning of the measurement equipment, and it seems clear that at least in the few instances, where alleged levels of SM_{obs} surpassed preceding seasonal precipitation, this latter case was responsible for the deviation from SM_{mod}.

All in all, three of the eight SM stations (RK-W, KF-W and WZ-upT) faced such measurement issues (Table B1). During summer and late autumn and before soil moisture slowly started to accumulate in the soil column, a lumpy to blocky, aggregate structure was encountered in the silty and clayey soils that had dried up and shrunk during summer. The soil formed about fist-size lumps, separated by desiccation cracks. Therefore, preferential flow paths forming along these cracks may lead to increased rates of rapid infiltration and soil moisture accumulation that may cause some of the sensors to read unrealistic moisture levels, if the sensors happen to cross such desiccation crack. It is difficult to assess such a possibility in hindsight since the desiccation cracks are a transient soil pattern and their temporary position remains unknown. However, such unrealistic SM readings did not occur in winter during the recharge events (possible because the temporary desiccation cracks had already closed). In addition, hand collected soil samples during the peak rainy season with independent control readings in the laboratory confirmed the validity of field readings at the automatic SM gauges. In only one instance, unrealistic SM readings were also encountered during the end of the rainy season (April 2008) at station RK-W. Equipment failure is the most probable explanation here. However, this isolated late event happened long after the period of SM saturation and recharge; it does not affect the results of the model calculations (Appendix B).

The overall reliability of SM readings is shown in Tables B1 and B2; all stations except one, RK-E, show an excellent Nash-Sutcliffe efficiency (NSE) coefficient. Therefore modelled SM can be considered a valid representation of actual SM developments over time and for different spatial units.

5.4 Annual recharge coefficients – comparison with other studies

As observed by Lerner *et al.* (1990), recharge is a temporally highly variable process that is most adequately investigated on the event level (in WAB a time span of half a day to five days). While the model with its daily steps fully reflects the event character during the measurement period, a generalisation of annual recharge coefficients always carries some range of uncertainty, since two years with identical annual rainfall can result in very different recharge values, depending on the seasonal distribution of precipitation and hence SM accumulation.
However, as presented already, the individual recharge coefficients for the different formations cropping out in Wadi Natuf range between a minimum of 0% (for non-recharging formations like u-Yat) and a maximum of 57%. These overall recharge values fall well within the range, usually quoted for the WAB. Table H1 lists the regional and other reported recharge coefficients, both, for annual and event-based calculations and together with the methods applied therein.

6 Conclusions

This study contributes to the assessment of distributed recharge based on soil moisture measurements and other field observations in Mediterranean karst areas with highly variable lithostratigraphy. Its findings extend those of PUB in general - and of Abusaada (2010) for the WAB in particular - that empirical and spatially differentiating field recordings are called for to produce model input parameters that are both, truly spatially distributed and realistic at the same time. Prior work has documented the importance and sensitivity of separate physical features in determining realistic parameters for distributed recharge models (Batelaan and de Smedt, 2001 and 2007, and Savenije, 2010). However, studies on distributed recharge often employed untested, virtual parameters or general data from the literature and performed only semi-distributed models, or models tested only by lumped basin responses, especially in the WAB.

In this study we based our parameters on long-term field measurements and observations and tested soil moisture hydrographs as well as the basin response on small, local, well-controlled sub-catchments of perched hillside aquifers. We found that for virtually all cases, the location-specific parameters such as soil water storage capacity were temporally stable, and the main transfer formula, linking the percolation process to physical soil properties, remained valid, irrespective of seasonal and inter-annual variations of climatic drivers. Furthermore, we could show that, as suggested by Binley and Beven (2003), an intelligent design and determination of the main processes at hand could enable a strictly parsimonious model to represent the key factors of complex hydrological processes such as recharge and thereby address the needs for distributed recharge models across a wide range of different climates and land forms. Thus, point measurements of soil moisture for various physical patterns of basin form may help to understand dominant mechanisms of deep percolation (recharge) even in hitherto scarcely gauged mountainsheds and deliver the basis for a realistic representation of distributed recharge. This application of parsimonious modelling and empirical field measurements is a novel step in WAB but also for most existing distributed recharge studies in ungauged basins worldwide. Unlike previous SM models in the WAB (Hughes *et al.*, 2008 and Sheffer, 2009), this model was fully distributed, not lumped and based on actual field observations and measurements.

Since our research was restricted to surface observations of the soil cover, we equated groundwater recharge with deep percolation from the soil into the unsaturated aquifer bedrock. Based on field observations of scarce runoff and lateral overland flow as well as the results by Messerschmid *et al.* (2018), we were able to neglect indirect recharge from runoff (1% in Wadi Natuf) and accounted only for direct recharge (in-situ percolation). Lateral water movement on a scale of decimetres and metres however, as from surrounding rock outcrops into soil pockets, was accounted for and included in our percolation model.

Based on previous PUB recommendations, we kept the actual soil moisture saturation model strictly parsimonious, empirically based and inductive; the employment of such a forward-calculating model is possible even in previously ungauged basins, if the dominant process can be identified and crucial input parameters for the model be gained through direct field observation and multi-annual measurements. Our model was based on the assumption that groundwater recharge can be simplified and described as soil saturation process in a variety of local conditions with a wide range of land features. Another assumption in line with PUB lessons was that specific maximum water storage capacities for the different soil moisture measurement plots can be determined and considered representative for specific aquifer formations. Thirdly, in order to build average coefficients, the measured time series had to be long enough (in our case seven years) to cover the prevailing climatic variations of wet & dry years. Last not least, and as our main assumption in the PUB approach, the recharge potential of the different formations, tested at different SM station, conceptually depends on a few select and interrelated key factors (lithology, land form, soil thickness) that drive and dominate the hydrological process of groundwater recharge.

As already mentioned, the temporal distribution of rainfall has a strong effect on event and seasonal recharge amounts and a modelling frequency of daily steps is appropriate under the particular climatic conditions of the WAB recharge areas in the Eastern Mediterranean mountainsides. This climate is characterised by a pronounced two-season annual rainfall pattern and by semi-arid to sub-humid overall annual precipitation depths. More importantly, this study for the first time in the WAB used a truly distributed approach for a great variety of different physical land forms. Instead of introducing untested proxy values as input to the model, and in line with the findings of the PUB decade, it was possible to solidly base our basin classification for dominant recharge processes on observations of the physical form and based on fundamental laws of physics (and for each of the basin form groups, such as LU/LC, soil depth and lithology, respectively).

However, some limitations are worth noting. Although our statistical work indicated very good to excellent NSE values, the reasons for some of the discrepancies between observed and modelled SM could not fully be explained. Furthermore, the examination of spring response signatures, although semi-qualitative, was restricted only to the local hillside aquifers. Yet, it should be emphasized that hardly any study on distributed recharge, and certainly no previous study in the WAB performed such testing on two independent conceptual levels of form and of function.

Further research may take two different directions: on the one hand to focus on the understanding of the effect of temporal differentiation of inter-annual recharge variations for each of the formations. On the other hand, this research provides a basis for the regionalisation of the local formation-specific recharge values and coefficients in order to estimate total recharge of all litho-stratigraphic in the catchment or on the WAB basin scale. (This regionalisation will be the subject of a follow-up article – see Messerschmid *et al.*, 2019b).

Acknowledgements: The initial concept and design and the first two measurement seasons of this project were conducted by the Palestinian-British research and development project SUSMAQ, the Sustainable Use of West Bank and Gaza Aquifers, between the Palestinian Water Authority and University of Newcastle upon Tyne, funded by DFID. This work would not have been possible without the unselfish efforts and invaluable contributions by many, particularly the daily spring readings by Eng. Majdi Zaydah, Ahmad Bajes and Ahmad Falah. The authors would like to thank Sebastian Schmidt for his version of the spread sheet model and the ET_p input, as well as the reviewers for their detailed and helpful comments. Gratitude is also owed to Ivonne Mansbach for many field trips and unwavering support throughout. This article is dedicated to Almut Hoffmann, my beloved partner who tragically was torn out of life in March 2015.

References

Abusaada, M. J.: Flow Dynamics and Management Options in Stressed Carbonate Aquifer System, The Western Aquifer Basin, Palestine, PhD Thesis, University of Göttingen, 2011, https://d-nb.info/1042263574/34, last access: 19 April 2017.

Abusaada, M. and Sauter, M.: Recharge Estimation in Karst Aquifers by Applying the Water Level Fluctuation Approach, Int J Earth Sci Geophys 3:013, 2017.

Aish, A.M., Batelaan, O. and De Smedt, F.: Distributed recharge estimation for groundwater modeling using WetSpass model, case study—Gaza strip, Palestine, *Arabian Journal for Science and Engineering 35*(1), 155, 2010.

Allocca, V., Manna, F., and De Vita, P.: Estimating annual groundwater recharge coefficient for karst aquifers of the southern Apennines (Italy), *Hydrol. Earth Syst. Sci., 18,* 803–817, doi:10.5194/hess-18-803-2014, 2014.

Allocca, V. , De Vita, P., Manna, F., and Nimmo, J.R.: Groundwater recharge assessment at local and episodic scale in a soil mantled perched karst aquifer in southern Italy, *J. Hydrol., 529,* 843–853, doi.org/10.1016/j.jhydrol.2015.08.032, 2015.

Arbel, Y., Greenbaum, N., Lange, J., and Inbar, M.: Infiltration processes and flow rates in developed karst vadose zone using tracers in cave drips, *Earth Surf. Process. Landf.*, *35*, 1682–1693, doi:10.1002/esp.2010, 2010.

Batelaan, O. and De Smedt, F.: WetSpass: a flexible, GIS based, distributed recharge methodology for regional groundwater modelling, IAHS Publ., *11*(18), 2001.

Batelaan, O. and De Smedt, F: GIS-based recharge estimation by coupling surface–subsurface water balances, *J. Hydrol. 337*(3-4), 337-355, 2007.

Berger, D.: Hydrological Model for the Yarqon-Taninim Aquifer, (in Hebrew), Mekorot Ltd., Tel Aviv, 50pp., 1999.

Beven, K.: Towards a coherent philosophy for modelling the environment, *Proc. Royal Soc.*, London, 458, 2465–2484, 2002.

Beven, K.J. and Kirkby, M.J.: A physically based, variable contributing area model of basin hydrology, Hydrological Sciences Bulletin *42*(1), 43–69, 1979.

Binley, A. and Beven, K.: Vadose Zone Flow Model Uncertainty as Conditioned on Geophysical Data, *Ground Water*, *41*, 119–127, 2003.

Blume, T., Zehe, E. and Bronstert, A.: Investigation of runoff generation in a pristine, poorly gauged catchment in the Chilean Andes I: a multi-method experimental study, *Hydrological Processes*, *22*(18), 3661-3675, 2008.

Bradford, R.B., Ragab, R., Crooks, S.M., Bouraoui, F. and Peters, E.: Simplicity versus complexity in modelling groundwater recharge in Chalk catchments, *Hydrology and Earth System Sciences*, *6*(5), 927-937, 2002.

Bredenkamp, D.B., Botha, L.J., van Tonder, G.J. and van Rensburg, H.J: Manual on quantitaive estimation of groundwater recharge and aquifer storativity: based on practical hydro-geological methods, Report (TT 73/95), Water Research Commission, Pretoria, South Africa, 1995.

Cheng, Y., Zhan, H., Yang, W., Dang, H. and Li, W.: Is annual recharge coefficient a valid concept in arid and semi-arid regions?, *Hydrol. Earth Syst. Sci*., *21*, 5031-5042, doi.org/10.5194/hess-21-5031-2017, 2017.

Dafny, E.: Groundwater flow and solute transport within the Yarqon-Taninim aquifer, Israel, *Diss. PhD Thesis*, Hebrew University of Jerusalem, Israel, 2009.

Dafny, E., Burg, A., and Gvirtzman, H.: Effects of Karst and geological structure on groundwater flow: The case of Yarqon-Taninim Aquifer, Israel, *J. Hydrol.*, *389*(3), 260-275. doi: 10.1016/j.jhydrol. 2010.05.038, 2010.

Dan, Y., Raz, Z., Yaalon, D.H. and Koyumdjisky, H.: Soil map of Israel 1: 500,000, Ministry of Agriculture, Agricultural Research Organization & Soil Consveration and Drainage Dept., Jerusalem, 1975, http://esdac.jrc.ec.europa.eu/images/Eudasm/Asia/images/maps/download/IL3002_SO.jpg, last access: 16 January 2019.

De Vries, J.J. and Simmers, I.: Groundwater recharge: an overview of processes and challenges, *Hydrogeol. J.*, *10*, 5–17, 2002.

Dingman, S.L.: *Physical/ Hydrology*, Macmillan, New York, USA, 1994.

Dooge, J.CI.: A general theory of the unit hydrograph, *Journal of Geophysical Research*, *64*(2), 241-256, 1959.

Dörhöfer, G., and Jesopait, V.: Grundwasserneubildung und ihre Ermittlung–eine Anmerkung zum Beitrag von Hölting: Modellrechnungen zur Grundwasserneubildung, Grundwasser, *2*(2), 77-80, doi.org/10.1007/s767-1997-8523-3, 1997.

Dvory, N.Z., Livshitz, Y., Kuznetsov, M., Adar, E., and Yakirevich, A.: The effect of hydrogeological conditions on variability and dynamic of groundwater recharge in a carbonate aquifer at local scale, *J. Hydrol.*, 535, 480-494, doi.org/10.1016/j.jhydrol.2016.02.011, 2016.

Eder, G., Sivapalan, M., and Nachtnebel, H.P.: Modelling water balances in an Alpine catchment through exploitation of emergent properties over changing time scales, *Hydrological Processes*, *17*, 2125–2149, 2003.

ESCWA-BGR: Inventory of shared water resources in Western Asia, Chapter 16 Western Aquifer Basin, United Nations Economic and Social Commission for Western Asia (*UN-ESCWA*) and Federal Institute for Geosciences and Natural Resources (BGR), Beirut, 23 pp., 2013, http://waterinventory.org/groundwater/western-aquifer-basin, last access: 14 May 2018.

Ettinger, H.: Spatial Analysis of extreme Storms and Floods in the Ayalon-Yarqon Catchment, *Phys. Geogr. Dept.*, Inst. Earth Sci., Hebrew Univ, Jerusalem, 55 pp., 1996.

Ford, D. and Williams, P.: Karst hydrology and geomorphology, John Wiley & Sons Ltd., West Sussex, England, 2007

Franchini, M. and Pacciani, M.: Comparative analysis of several conceptual rainfall–runoff models, *J. Hydrol.*, *122*, 161–219, 1991.

Freeze, R. A.: The mechanism of natural ground-water recharge and discharge. L: One-dimensional, vertical, unsteady, unsaturated flow above a recharging or discharging ground-water flow system, *Wat. Resour. Res. 5*(1), 153-171, 1969.

Frumkin, A., Schwarz, H.P. and Ford D.C.: Evidence for isotopic equilibrium in stalagmites from caves in a dry region: Jerusalem, Israel, *Isr. J. Earth Sci., 43,* 221-230, 1994.

Goldscheider, N. and Drew, D. (eds.): Methods in Karst Hydrogeology, International Contribution to Hydrogeology, IAH, vol 26. Taylor and Francis/Balkema, London, 264 pp., 2007.

Goldschmidt, M. J. and Jacobs, M.: Precipitation over and replenishment of the Yarqon and Nahal Hatteninim underground catchments, Hydrological Paper 3, Hydrological Service of Israel, Jerusalem, 1958.

Grayson, R. and Blöschl, G.: Summary of pattern comparison and concluding remarks. Chapter 14 *In*: R. Grayson, *et al*., eds. *Spatial patterns in catchment hydrology: observations and modelling*, Cambridge: Cambridge University Press, 355–367, 2000.

Grodek, T., Lange, J., Lekach, J., and Husary, S.: Urban hydrology in mountainous middle eastern cities, *Hydrol. Earth Syst. Sci.*, *15*, 953–966, doi:10.5194/hess-15-953-2011, 2011.

GSI: Geological Map of Israel, Sheet 8-III: Lod, Geological Map by: Yechieli, Y., GSI - Geological Survey of Israel, Jerusalem, 2008, http://www.gsi.gov.il/eng/?CategoryID=253&ArticleID=778, last access: 16 March 2018.

GSI: Geological Map of Israel, Sheet 8-IV: Ramallah, Geological Map by: Shachnai, E., GSI - Geological Survey of Israel, Jerusalem, 2000, http://www.gsi.gov.il/eng/?CategoryID=253&ArticleID=777, last access: 16 March 2018.

Gunkel, A., Shadeed, S., Hartmann, A., Wagener, T. and Lange, J.: Model signatures and aridity indices enhance the accuracy of water balance estimations in a data-scarce Eastern Mediterranean catchment. J. Hydrol.: Reg. Stud., 4, 487-501, doi.org/10.1016/j.ejrh.2015.08.002, 2015.

Guttman, J.: Multi-Lateral Project B: Hydrogeology of the Eastern Aquifer in the Judea Hills and Jordan Valley, Mekorot Water Company, Report No. 468, Tel Aviv, Israel, 2000.

Guttman, J., Goldshtoff, I., Baida, A. and Mercado, A.: A two-layer model of flow regime and salinity in the Yarqon-Taninim Aquifer, Tahal Consult. Eng., (Rep. 01/88/23), Tel Aviv, Israel, 19 pp., 1988.

Guttman, J., and Zukerman, H.: Yarqon-Taninim-Beer Sheeba Basin, flow model, Tahal Consult. Eng., Tel Aviv, Israel, 37 pp., 1995.

Hargreaves, G.H. and Samani, Z.A.: Reference crop evapotranspiration from temperature, *Appl. Eng. Agr.*, *1*, 96–99, 1985.

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

Hrachowitz, M., Savenije, H.H., Blöschl, G., McDonnell J.J., Sivapalan, M., Pomeroy, J.W., Arheimer, B., Blume T., Clark, M.P., Ehret, U. and Fenicia, F.: A decade of Predictions in Ungauged Basins (PUB) - a review, *Hydrol. Sci. J.*, *58*(6), 1198-255, 2013.

Hrachowitz, M. and Clark, M.P.: HESS Opinions: The complementary merits of competing modelling philosophies in hydrology, *Hydrol. Earth Syst. Sci.*, *21*(8), 3953-3973, doi.org/10.5194/hess-21-3953-2017, 2017.

HSI: Development of utilization and status of water resources in Israel until Autumn 2014, (in Hebrew), State of Israel, Water commission, Hydrological service, Annual report, Jerusalem, 509 pp., 2016, http://water.gov.il/Hebrew/ProfessionalInfoAndData/Data-Hidrologeime/Pages/water-resources-2014.aspx, last access: 17 May 2018.

Hughes, A.G. and Mansour, M.M.: Recharge modeling for the West Bank aquifers, *BGS* - British Geological Survey, Commissioned Report, CR/05/087, Keyworth, Nottingham, UK, 90 pp., 2005.

Hughes, D.A. and Kapangziwiri, E.: The use of physical basin properties and runoff generation concepts as an aid to parameter quantification in conceptual type rainfall–runoff models. *In*: E. Boegh, *et al*.., eds. *Quantification and reduction of predictive uncertainty for sustainable water resources management*, Wallingford: IAHS Press, IAHS Publ. *313*, 311–318, 2007

Hughes, A. G., Mansour, M. M. and Robins, N. S.: Evaluation of distributed recharge in an upland semi-arid karst system: the West Bank Mountain Aquifer, Middle East, *Hydrogeol. J.*, *16*, 845–854, doi:10.1007/s10040-008-0273-6, 2008.

IMS – Israel Meteorological Service: online database of the Israel Meteorological Service, Bet Dagan, http://data.gov.il/ims/3, last access: 1 October 2017.

Jeannin, P. Y. and Sauter, M.: Analysis of karst hydrodynamic behaviour using global approaches: a review, *Bull. Hydrogéol. (Neuchâtel)*, *16*, 31-48, 1998.

Juston, J., Seibert, J. and Johansson, P.O.: Temporal sampling strategies and uncertainty in calibrating a conceptual hydrological model for a small boreal catchment, *Hydrological Processes*, *23*(21), 3093-3109, 2009.

Kapangaziwiri, E., Hughes, D.A. and Wagener, T.: Constraining uncertainty in hydrological predictions for ungauged basins in southern Africa, *Hydrological Sciences Journal*, *57*(5), 1000–1019, 2012.

Lambert, A.O.: A comprehensive rainfall-runoff model for an upland catchment, *J. Instn Wat. Engrs*, *23*, 231-238, 1969.

Lange, J., Greenbaum, N., Husary, S., Ghanem, M., Leibundgut, C. and Schick, A. P.: Runoff generation from successive simulated rainfalls on a rocky, semi-arid, Mediterranean hillslope, *Hydrological Processes*, *17*, 279–296, doi:10.1002/hyp. 1124, 2003.

Lange, J., Arbel, Y., Grodek, T. and Greenbaum, N.: Water percolation process studies in a Mediterranean karst area, *Hydrological Processes*, *24*, 1866–1879, 2010.

Lerner D.N., Issar, A.S. and Simmers, I.: Groundwater recharge: A guide to understanding and estimating natural recharge, *International Contributions to Hydrogeology*, Vol. *8*, Verlag Heinz Heise, Hannover, 375 pp., 1990.

Lloyd, J.W.: A review of various problems in the estimation of groundwater recharge, in: Proc. Groundwater Recharge Conf., AWRC, Townsville, July 1980, pp. 1-25, 1980.

Martínez-Santos, P. and Andreu, J. M.: Lumped and distributed approaches to model natural recharge in semiarid karst aquifers, J. Hydrol., 388 (3-4), 389–398, doi.org/10.1016/j.jhydrol.2010.05.018, 2010.

McDonnell, J.J. and Woods, R.: On the need for catchment classification, *J. Hydrol.*, *299*, 2-3, 2004.

McDonnell, J. J., Sivapalan, M., Vaché, K., Dunn, S., Grant, G., Haggerty, R., Hinz, C., Hooper, R., Kirchner, J., Roderick, M. L., Selker, J. and Weiler, M.: Moving beyond heterogeneity and process complexity: a new vision for catchment hydrology, *Water Resources Research*, *43*, W07301. doi: 10.1029/2006WR005467, 2007.

Messerschmid, C.: Deep-seated aquicludes in a fractured carbonate environment, West Bank, in: Krásný, J., Hrkal, Z. and Bruthans, J. (eds.): International Conference on Groundwater in fractured rocks, IAH, 15-19 September 2003, Prague, Czech Republic, 8 pp., 2003.

Messerschmid, C., Kalbouneh, A., Zayed, O. and Sayfi, E.: Conceptual Model of Wadi Natuf Groundwater Recharge, Final Report – SUSMAQ-NAT # 64 V 0.3, Prepared by: SUSMAQ TEAM, Palestinian Water Authority, Palestine; Water Resource Systems Research Laboratory, University of Newcastle upon Tyne, UK; November 2003, Ramallah, 2003a.

Messerschmid, C., Aliewi, A., Kalbouneh, A., Sayfi, E. and Ziada, M.: Field trips to the springs in Wadi Natuf, Final Report – SUSMAQ-NAT # 65 V 0.1, Prepared by: SUSMAQ TEAM, Palestinian Water Authority, Palestine; Water Resource Systems Research Laboratory, University of Newcastle upon Tyne, UK; July 2003, Ramallah, 2003b.

Messerschmid, C., Lange, J. and Sauter, M.: Assessment of transmission loss in a Mediterranean karstic watershed (Wadi Natuf, West Bank), *Hydr. Proc*., *32*(10), 1375-1390, doi.org/10.1002/hyp.11496, 2018.

Messerschmid, C., Sauter, M. and Lange, J.: Regionalization of distributed groundwater recharge and leakage calculations in a Mediterranean karst catchment, Wadi Natuf, West Bank, *Hydrol. Earth Syst. Sci., - in preparation*, 2019b.

Nash, J. E.: The form of the instantaneous unit hydrograph, *International Association of Scientific Hydrology, Publ*, *3*, 114-121, 1957.

Nativ, R., Adar, E., Dahan, O. and Geyh, M.: Water recharge and solute transport through the vadose zone of fractured chalk under desert conditions, *Water Resources Research*, *31*(2), 253-261, 1995.

Radulović, M.M., Stevanović, Z. and Radulović, M.: A new approach in assessing recharge of highly karstified terrains – Montenegro case studies. Environ Earth Sci., 65 (8), 2221-2230, doi.org/10.1007/s12665-011-1378-0, 2011.

Olden, J. D., Kennard, M. J. and Pusey, B. J.: A framework for hydrologic classification with a review of methodologies and applications in ecohydrology, *Ecohydrology*, *5*(4), 503–518, 2012.

Pomeroy, J.: IAHS Decade for Predictions in Ungauged Basins, PUB, Presentation at IAHS-PUB 11 May, 2011, Canmore, Canada, 2011, www.usask.ca/hydrology/pub2011/presentations/Pomeroy_PUB2011_Weds11May.pdf, last access: 17 December 2018.

Ries, F., Lange, J., Schmidt, S., Puhlmann, H., and Sauter, M.: Recharge estimation and soil moisture dynamics in a Mediterranean, semi-arid karst region, *Hydrol. Earth Syst. Sci.*, *19*, 1439-1456, doi.org/10.5194/hess-19-1439-2015, 2015.

Ries, F., Schmidt, S., Sauter, M. and Lange, J.: Controls on runoff generation along a steep climatic gradient in the Eastern Mediterranean, *J. Hydrol.: Reg. Stud.*, *9*, 18–33, doi.org/10.1016/j.ejrh.2016.11.001, 2017.

Rosenzweig, A.: Study of the difference in effects of forest and other vegetative covers on water yield, Final report, Project A-10-FS-13, State of Israel, Ministry of Agriculture, Soil Conservation and Drainage Division Research Unit, Rishon LeTsiyon, Israel, 1972.

Rushton, K.R., Eilers, V.H.M. and Carter, R.C.: Improved soil moisture balance methodology for recharge estimation, *J. Hydrol.*, *318*(1-4), 379-399, doi.org/10.1016/j.jhydrol.2005.06.022, 2006.

Ryu, Y., Baldocchi, D.D., Ma, S. and Hehn, T.: Interannual variability of evapotranspiration and energy exchange over an annual grassland in California, *J. Geophys. Res. 113*, D09104, dx.doi.org/10.1029/2007JD009263, 2008.

Sanz, E., Menéndez Pidal de Navascués, I., and Távara, C.: Calculating the average natural recharge in large areas as a factor of their lithology and precipitation, *Hydrol. Earth Syst. Sci. Discuss.*, *8*, 4753–4788, doi:10.5194/hessd-8-4753-2011, 2011.
Savenije, H.H.G.: Topography driven conceptual modeling, FLEX-Topo. *Hydrol. Earth Syst. Sci*., *14*, 2681–2692, 2010.
Scanlon, B.R., Healy, R.W., and Cook, P.G.: Choosing appropriate techniques for quantifying groundwater recharge, *Hydrogeol. J.*, *10*, 18–39, doi:10.1007/s10040-002-0200-1, 2002.
Scanlon, B.R., Keese, K.E.A., Flint, L., Flint, L.E., Gaye, C.B., Edmunds, W.M. and Simmers, I.: Global synthesis of groundwater recharge in semiarid and arid regions, *Hydrol. Proc.*, *20(*15), 3335–3370, doi:10.1002/hyp.6335, 2006.
Schmidt, S., Geyer, T., Guttman, J., Marei, A., Ries, F., and Sauter, M.: Characterisation and modelling of conduit restricted karst aquifers – example of the Auja spring, Jordan Valley, *J. Hydrol., 511*, 750–763, 2014.
Seibert, J. and Beven, K.J.: Gauging the ungauged basin: how many discharge measurements are needed?, *Hydrol. Earth Syst. Sci*., *13*(6), 883–892, 2009, www.hydrol-earth-syst-sci.net/13/883/2009/, last access 17 December 2018.
Shachori A., Michaeli A. and Rosenzweig D.: Hydrologic Studies on a Representative Karst Catchment in Israel, Symposium of Budapest, International Union of Geodesy and Geophysics, 333-346, 1965.
Sheffer, N.A.: Variable scale recharge measurement and modeling using the hydrometeorological DREAM, Ph.D. dissertation, Hebrew Univ. of Jerusalem, Israel, 97 pp., 2009.
Sheffer, N.A., Dafny, E., Gvirtzman, H., Navon, S., Frumkin, A., and Morin, E.: Hydrometeorological daily recharge assessment model (DREAM) for the Western Mountain Aquifer, Israel: Model application and effects of temporal patterns, *Water Resour. Res.*, *46*, W05510, doi:10.1029/2008WR007607, 2010.
Simmers, I., (Ed.): Estimation of Natural Groundwater Recharge, D. Reidel Publishing Company, Dordrecht, Nederland, 510 pp., 1988.
Sivakumar, B.: Dominant processes concept in hydrology: Moving forward, *Hydr. Proc.*, *18*(12), 2349–2353, 2004.
Sivakumar, B.: Dominant processes concept, model simplification and classification framework in catchment hydrology, *Stochastic Environ. Res. Risk Assess.*, *22*(6), 737–748, 2008.
Sivakumar, B., Singh, V.P., Berndtsson, R. and Khan, S.K.: Catchment classification framework in hydrology: challenges and directions, *Journal of Hydrologic Engineering 20*(1), A4014002, 2013.
Sivapalan, M.: Prediction in ungauged basins: a grand challenge for theoretical hydrology, *Hydr. Proc. 17*(15), 3163-3170, 2003.
Sivapalan, M., Takeuchi, K., Franks, S.W., Gupta, V.K., Karambiri, H., Lakshmi, V., Liang, X., McDonnell, J.J., Mendiondo, E.M., O'Connell, P.E., Oki, T., Pomeroy, J.W., Schertzer, D., Uhlenbrook, S. and Zehe, E.: IAHS Decade on Predictions in Ungauged Basins (PUB), 2003–2012: Shaping an exciting future for the hydrological sciences, *Hydr. Sci. J.*, *48*(6), 857-880, DOI:10.1623/hysj.48.6.857.51421, 2003a.
Sivapalan, M., Blöschl, G., Zhang, L., and Vertessy, R.: Downward approach to hydrological prediction, *Hydrol. Proc.*, *17*(11), 2101–2111, 2003b.
Steinmann, A.: Runoff generation in Mediterranean areas, Dissertation, University of Freiburg i.Br; Hydrology Dept., Freiburg, Germany, 159 pp., 2010.
SUSMAQ: Compiled Base Data for the Numerical Groundwater Flow Model of the Western Aquifer Basin. Volume 1 - General Background, SUSMAQ-MOD Report No. #07 V0.4, The Sustainable Management for the West Bank and Gaza Aquifers, Palestinian Water Authority and University of Newcastle upon Tyne, UK, Ramallah, Palestine, 306 pp., 2002.
UN-GA: Convention on the Law of the Non-navigational Uses of International Watercourses, adopted by the General Assembly of the United Nations on 21 May 1997, New York, United Nations, 1997, http://legal.un.org/ilc/texts/instruments/english/conventions/8_3_1997.pdf, last access: 19 January 2019.
Wagener, T., Sivapalan, M., Troch, P. and Woods, R.: Catchment classification and hydrologic similarity, *Geography compass*, *1*(4), 901-931, 2007.
Weiss, M. and Gvirtzman, H.: Estimating ground water recharge using flow models of perched karstic aquifers, GroundWater, 45, 761–773, 2007.
Winsemius, H.C., Schaefli, B., Montanari, A. and Savenije, H.H.G.: On the calibration of hydrological models in ungauged basins: A framework for integrating hard and soft hydrological information, *Water Resources Research*, *45*(12), 2009.
Woods, R.A.: Seeing catchments with new eyes, *Hydrological Processes*, *16*(5), 1111–1113, 2002.
World Bank: West Bank and Gaza - Assessment of restrictions on Palestinian water sector development, Washington, DC: World Bank, 2009, http://documents.worldbank.org/curated/en/775491468139782240/West-Bank-and-Gaza-Assessment-of-restrictions-on-Palestinian-water-sector-development, last access: 17 December 2018.
Young, P.C. and Ratto, M.: A unified approach to environmental systems modeling, *Stochastic Environmental Research and Risk Assessment 23*(7), 1037-1057, 2009.
Zomlot, Z., Verbeiren, B., Huysmans, M. and Batelaan, O.: Spatial distribution of groundwater recharge and base flow: Assessment of controlling factors, J. Hydrol. Reg. Stud., 4, 349-368, doi.org/10.1016/j.ejrh.2015.07.005, 2015.

Figure A1: Formation outcrops. Rare outcrops of yellow marls, forming the main regional aquiclude (upper-Yatta ftn.) near Ras Karkar (a), reefal limestone cliff at Wadi Zarqa (upper-UBK ftn.) with high primary porosity, development of karst and epi-karst (b), karstic karren landscape (upper-Betlehem ftn.), a relatively rare example of discontinuous soil cover (soil pockets) in Wadi Natuf (c).

Figure A2: Soil cover. Soil covering marly limestone (bottom Hebron/top Yatta ftn.), with some white caliche-type 'Nari' crust (a, b, c), Ras Karkar soil measurement station (RK-W, lower Hebron ftn.) on natural terraces of thin terra rossa soil (d).

Appendix B

Some technical equipment an interpretation problems were faced at three of the eight SM stations. During brief periods, the devices read out different soil moisture levels than those modelled (and in some cases, SM higher than preceding accumulated rainfall or a rise of SM during a dry period). The following tables show the results of the statistical analysis. Table B1 counts the days with erroneous and questionable readings (together 147 days).

Table B1: Days with misreadings.

soil station	periods	no. of misread days	sum [days]
WZ-up.T	31.10.-5.11.2005	6	6
KF-W	9.-20.11.'05	12	42
	17.-25.12.'05	+9	
	29.10-13.11.'06	+16	
	18.-22.11.'06	+5	
RK-W	6.-15.11.'05	10	55
	22.11.-25.12.'05	+24	
	4.-24.11.'06	+21	
		sub-total	**103**
RK-W (Apr-'08)	17.3.-29.4.'08	44	44
all stations		**total sum**	**147**

Table B2 compares the days with such "misreadings" to the total number of days measured and modelled, and it indicates their respective shares. Misreadings here are understood as a general term for strong deviations between SM_{obs} and SM_{mod} or as unreasonable measurement values, either with alleged SM-contents above cumulative rain heights or moisture increase during dry spells; the reasons are therefore manifold and misreadings can occur due to desiccation cracks (mostly in autumn) or malfunctioning of the instruments.

Together, the 147 day of questionable data in three stations represent 8.1% of all days with SM records (out of 1,818 days of total SM readings), and 0.7% of the total model period (20,456 days for the eight stations and over seven years), marked grey in Table B2.

- In WZ-upT, only 6 days of failure occurred (Nov-2005), equivalent to 2.9% of recording days.
- In KF-W, 42 days (or 14.9%) occurred over the entire read-out period.
- In RK-W, 55 days occurred in early winter 2005 and 2006 (with $SM_{obs} > P_{cumulative}$), equivalent to 8.9% of recorded days, in addition to 44 days in April 2008 (together 16% failed readings).

Table B2: periods and shares of days with misreadings.

soil station	reference period	no. of days days misread	days misread in 7-year period (2,557 d) – [%]	misread days as share of recorded days – [days, (%)]
WZ-up.T	w/o Apr-'08	6	0.23%	207 (2.90%)
KF-W	w/o Apr-'08	42	1.64%	281 (14.95%)
RK-W	w/o Apr-'08	55	2.15%	617 (8.91%)
RK-W	incl. Apr-'08	99	3.87%	617 (16.05%)

soil station	reference period	no. of days with SM>P	days misread over total period of 7 years & 8 stations (20,456 d) – [%]	no. of all days read [-] and share of misread days [(%)]
all stations	incl. Apr-'08	147	**0.72%**	1818 (**8.09%**)
all stations	w/o Apr-'08	103	0.50%	1818 (5.67%)
RK-W	only Apr-'08	44	0.22%	617 (7.13%)

Table B3 presents the Nash-Sutcliffe efficiency coefficients for the eight SM stations. All stations (except one, RK-E) show an excellent NSE between 0.40 and 0.96. The one station with a bad correlation, RK-E (with an NSE of -0.35) was excluded from further calculations; its results were not used for the analysis of annual runoff coefficients (RC). As a result, we can state that such periods of malfunctioning soil moisture probes are extremely short and rare; thus, they do not put into question the large bulk of soil moisture results, particularly during the important periods of actual deep percolation (recharge) after strong rainfall events and under sufficient preceding pre-wetting of the soils. In addition, these quantities of additional soil water above the maximum storage capacity are fully accounted for by the daily budgets of the percolation model.

Table B3: NSE statistics of SM_{obs} versus SM_{mod}.

station	**NSE**
BET	**0.87**
Shu	**0.80**
WZ-upT	**0.63**
WZ-hh	**0.40**
KF-W	**0.79**
KF-E	**0.96**
RK-W	**0.70**
RK-E	-0.35
avg. all (w/o RK-E)	**0.73**
avg. all (incl. RK-E)	**0.60**

Appendix C

Table C1: Five spring data coverage.

Year	Abu Sa‘efan	Bibi	Qos	Salem	‘Akkari	∑ Coverage
2003/04	**57 %**	**35 %**	**54 %**	**58 %**	**58 %**	35–58 %
2004/05	**2 %**	**2 %**	**2 %**	**2 %**	**2 %**	2 %
2005/06	**5 %**	**5 %**	**5 %**	**79 %**	**79 %**	5 % & 79 %
2006/07	100 %	100 %	100 %	100 %	100 %	**full**
2007/08	100 %	100 %	100 %	100 %	100 %	**full**
2008/09	100 %	100 %	100 %	100 %	100 %	**full**
2009/10	**42 %**	**42 %**	**42 %**	**68 %**	**68 %**	42 % + 68 %

Daily spring readings were carried out by hand (with stop watch and bucket) and by different personnel in the respective spring groups. Beitillu spring group was fully covered for three years. In the first season (2003/04), measurements started late into the season, then were read most of the days in winter and scarcely in summer. In season 2009/10, continuous reading continued until February 2010. The two springs at Wadi Zarqa spring group (Salem & ‘Akkari) had the same coverage as Beitillu group in 2003/04 and then were almost uninterruptedly read from December 2005 (season 2005/06) until February 2010. Hence, high coverage was achieved during almost six seasons in Wadi Zarqa and during nearly five seasons in Beitillu.

Appendix D

Table D1: Soil depth matrix.

Hydrostratigraphy	**Upper Aquifer**	**Yatta**	**Lower Aquifer**
Geological formation	Heb., Bet., Jer.	up UBK, Yatta	LBK, low UBK*
A) Terraces			
1. terraces with olives	65, 40 cm	**42, 25 cm**	**(49), 26 cm**
2. terraces with other types cultivation	NA	**(25)**	NA
3. formerly used but now uncultivated terraces	NA	12, 47	46, 51, 22
4. natural terracing with shrubs and grass cover	(26) 40	10	**22, 29**
B) Plains			
1. arable plains with olive orchards	63, 67	**94**	26
2. arable plains with other types of cultivation	36	**(56), 50, 65, 77**	40
3. arable plains without cultivation	(58)	56	NA
4. rock plastered plains (karstification)	32	NA	(5), 15
5. dry plains with shrubs and grass cover	**19**	NA	19
C) Slopes			
1. non-terraced slopes with olives	NA	56	49
2. non-terraced slopes with other types of cultivation	**58**	50	NA
3. slopes with shrubs and grass cover	**11, 26, 32**	40	**18**
4. rock plastered slopes	32	NA	**11**
D) Pure rock cover and cliffs			

The soil thickness survey was carried out at representative locations of every outcropping formation, such as different types of vegetation, relief, land use and natural land cover. The table simplifies the results for an overview over different typical soil depths for the regional units of Upper and Lower Aquifers and the regional Middle Aquitard (with the individual formations indicated in line 2). Values in red colour were also encountered at the SM stations. Main land form types are shaded grey. The values for l-UBK formation are representative for the main body of the formation but not for the top of l-UBK (see *). Boxes marked NA, represent untypical vegetation and land form types for the respective formations.

The seven years covered in this study (2003/04-2009/10) are representative for long term rainfall in the basin. The Hydrological Service of Israel presents rainfall records for different sub-basins (cell groups) of the Western Aquifer Basin. HSI (2016) divides the WAB into six cells from North to South (Fig. E1). Wadi Natuf lies at the border between the Northern and Central cell group of the WAB and nearby the Jerusalem cell group and was found to be best represented by a mixture of the central, northern and Jerusalem cells of WAB.

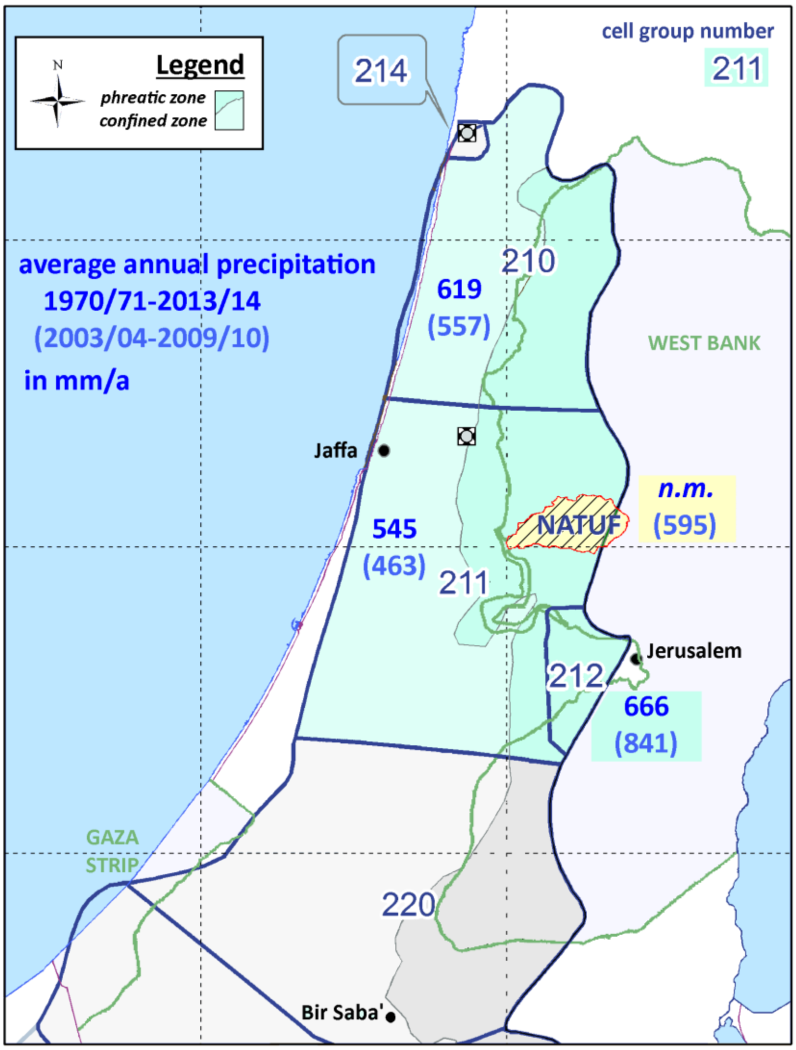

Figure E1: WAB cells of area precipitation (modified after: HSI, 2016). The map indicates area precipitation for two periods, long-term (1970/71-2013/14) and for the seven years of this study. Long-term area rainfall data for Wadi Natuf are not available; the 7-year period was measured by tipping buckets. Precipitation data for the three cells groups, no. 210, 211 and 212 were taken from HSI (2016).

The period of rainfall documented records spans 44 years from 1970/71 to 1013/14. Of the 44 years, only two years were drier and only 12 years were wetter than the seven years studied here. Out of these 12 wet years, only five years (11% of the 44 years) were wetter by more than 20% than the maximum in the range of seven years. And if we allow for a 10% margin of rainfall deviation, 82% of all 44 years (36 of 44 years) lie within the total range.

The following analysis was carried out. For each cell group (North, centre and Jerusalem) the arithmetic mean precipitation was calculated over 44 years and over the seven year period, respectively. Then the average of the three mean values for each cell group was formed, again, for the entire and the 7-year period. This average lies at 610mm/yr over the entire 44 years and at 621mm/yr during the seven years studied here. In other words, the seven years were wetter but only slightly; the average of seven years is 101.7% the average of all 44 years (for the spatial averages of three cell groups).

Especially in highly variant climates, it is recommended, to use the median value instead of the arithmetic mean, because it indicates the most probable rainfall. Here, the median was formed for 44 years and seven years, independently and for each cell group, respectively. Again, the average of the three geographic cell groups was formed and resulted in slightly different median rain heights, of 600 mm/yr long term and of 624 mm/yr for the 7-year period. The short- and long-term median values were compared in each cell group independently; their average lies at a ratio of 99.6% between long- and short-term rainfall observations. The ratio between the averages (of cell group medians) of seven years and of all 44 years is 99.6%. Table E1 shows the two different calculation methods and the end results of the analysis, together with the individual results for each cell group.

Table E1: Comparison of average rainfall in WAB cell groups, long-term and for the 7-years 2003/04-2009/10.

cell item	**type**	**period**	**# 210 North**	**# 211 Centre**	**# 212 Jerus.**	**avg. of all three cell groups**	**item**
avg. of all years	long term	1970/71-2013/14	619	545	666	610	avg. of averages (all years)
average 7 years	7-year period	2003/04-2009/10	557	463	841	621	avg. of averages (7years)
ratio			**90,0%**	**85,0%**	**126,3%**	**101,7%**	* *ratio between averages*
median of all years	long term	1970/71-2013/14	597	504	698	600	avg. of medians (all years)
median of 7 years	7-year period	2003/04-2009/10	593	460	820	624	Avg. of Medians (7 years)
Ratio			**99,3%**	**91,3%**	**117,6%**	**99,6%**	* *average of 3 groups*

* The above calculation uses the average between resulting means (610 mm and 621 mm per, respectively). The below calculations show the average of the ratios, formed in each cell group. WAB annual precipitation was based on data from the Hydrological Service (HSI, 2016). All rain heights are in mm/yr, the ratios in percent.

Thus, it can be shown that the seven years of rainfall observation in Wadi Natuf cover almost the entire range of long-term annual rain distribution. Only one single year, the winter of 1991/92 with its century rainfall clearly lies outside the rainfall range covered by our measurement period. As a result, we can state with confidence that the seven years covered in this study happen to match very closely with the overall range of long-term annual rain heights in the WAB.

Appendix F

When signatures of peak spring flow periods were compared with those of recharge (DP), it had to be assured that the hydrographs of the daily read springs were representative for the entire spring groups draining the respective aquifers. This could be affirmed by statistical analysis as shown below. In addition to the five daily read springs, all springs in Wadi Natuf were measured during four key date campaigns (as detailed in Messerschmid *et al.*, 2019b). The tables below show that the ratios between the combined daily measured springs and the entire spring group discharge were very stable; their relative shares deviated only slightly, as shown by the low standard deviation values in Table F1 and Table F2.

Table F1: Spring flow comparison and standard deviation between daily read springs and entire spring group flow during key date campaigns – spring group Beitillu (Harat al-Wad).

Spring flow (m^3/d)	03 summer	03/04 winter	04 summer	07 summer	**St. Dev.**
Ein Abu Sa'efan	4.1	3.8	2.0	2.0	
Ein Al - Bibi	3.0	1.6	1.4	1.3	
Ein Al -Qos	12.4	16.7	9.6	13.2	
sum Beitillu spring group	**98**	**110**	**61**	**73**	
Relative share of group flow (%)					
Ein Abu Sa'efan	4%	3%	3%	3%	
Ein Al - Bibi	3%	1%	2%	2%	
Ein Al -Qos	13%	15%	16%	18%	
Sum (of 3 springs)	**19.9%**	**20.2%**	**21.4%**	**22.4%**	**1.15%**

At the four key date campaigns, total spring outflow at Beitillu spring group (Table F1) was lying between 61 and 110 m^3/d. The three daily monitored springs combined made up 20-22% of total spring group flow with a standard deviation of 1.15%. At the four key date campaigns, total spring outflow at Wadi Zarqa spring group (Table F2) was lying between 165 and 493 m^3/d. The two daily monitored springs combined made up 6.8-8.5% of total spring group flow with a standard deviation of 0.72%.

Table F2: Spring flow comparison and standard deviation between daily read springs and entire spring group flow during key date campaigns – spring group Wadi Zarqa.

Spring flow (m^3/d)	03 summer	03/04 winter	04 summer	07 summer	**St. Dev.**
Ein Al-'Akkari 1	13	17	14	7	
Ein Salem 1 (upper)	20	25	10	4	
sum Wadi Zarqa spring group	**414**	**493**	**303**	**165**	
Relative share of group flow (%)					
Ein Al-'Akkari 1	3.1%	3.5%	4.5%	4.3%	
Ein Salem 1 (upper)	4.8%	5.1%	3.2%	2.4%	
sum (of 2 springs)	**7.8%**	**8.5%**	**7.7%**	**6.8%**	**0.72%**

In most years, the variability of precipitation was reciprocal to the variability of recharge – in other words: the wetter the year, the higher the spatial variability of rainfall, but the lower the spatial variability of recharge. Variability here is expressed as the standard deviation (σ) for the values of the different stations. Although one year (2005/06) was excluded from the trend line, the otherwise quite pronounced pattern of the diagram shows that indeed spatial variability of recharge deserves more attention that it was given so far, at least but not only in the WAB, where so far, almost all studies focussed mostly if not solely on temporal recharge variability. The trend is in line with general observations worldwide that basin responses in relatively dry climates are more variable than those in wet climates. A similar response pattern was observed for runoff in Messerschmid *et al.*, (2018), where strong precipitation events triggered runoff in all areas alike, whereas weaker rainfall events showed different responses in different areas.

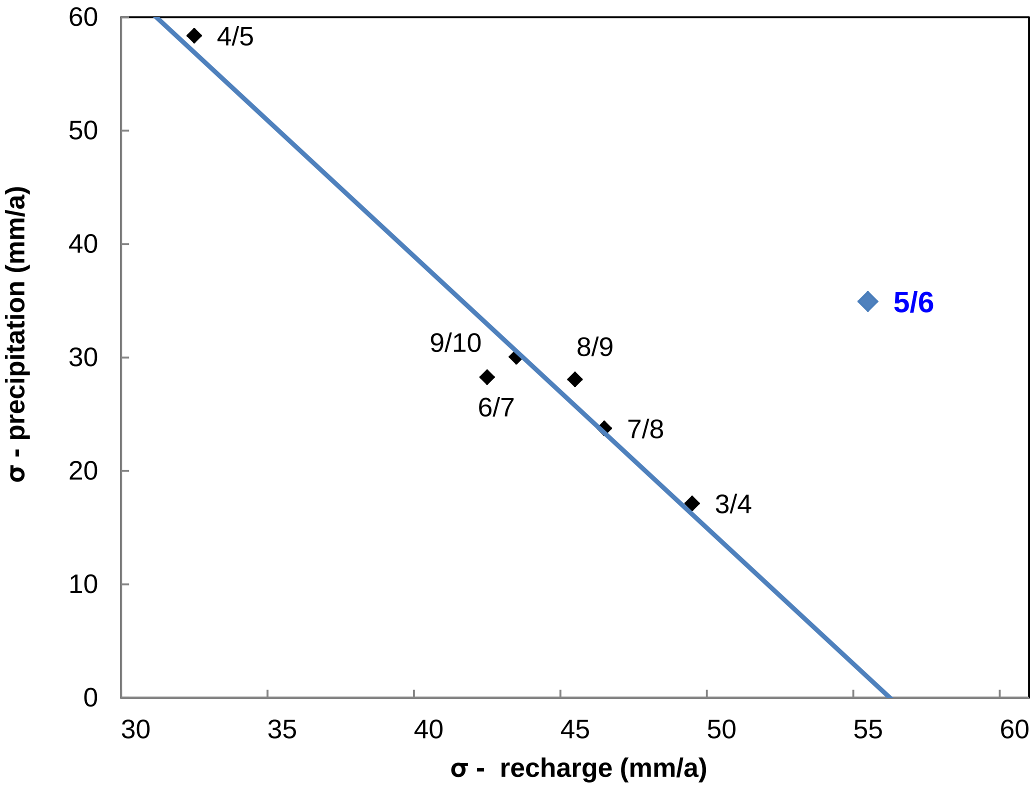

Figure G1: Scaling - standard deviations (σ) of P & DP. Note: the linear trend line excludes the year 2005/06, which deviates from the trend of the other years. In this year, spatial variability of recharge was quite high but nonetheless, spatial variability of recharge was high as well. 2005/06 was a special year, because it had an exceptionally intensive storm event (and the strongest during the entire measurement period) very late in the season (early April 2006). The different stations were affected differently by the relatively long dry period before the event and the high precipitation during the event, depending on the soil thickness and therefore on the speed of drying up of the soils and consequent soil saturation during the event.

Appendix H

Recharge rates (in mm/a, or calculated as coefficients in %) have been reported in numerous studies within the WAB and adjacent basins. Table H1 presents a collection of findings in the literature from both, the WAB as well as other regions of the world and from a large variety of different research methods. The table shows that with an average area recharge coefficient between 41% and 47% (for the WAB portion of Wadi Natuf, see Messerschmid *et al.*, 2019b), and with RC-values between 30% and 57% for the individual aquiferous formations (not the aquitards), the findings of the model match very well with the ranges from the literature. While the lowest figures reported only 9-40% of RC, the highest figures reached over 50% and up to 60% of recharge coefficients (in Mt. Carmel basin with a very similar geology and climate as in Wadi Natuf). For a small groundwater catchment overlapping with Wadi Natuf surface catchment, Weiss and Gvirtzman (2007) found a long-term average of 47% (see note below Table H1).Both, the temporal and spatial variability of reported values are impressive (temporal variability can be detected by comparing average with maximum RC-values; spatial variability emerges from the differences between different basins and sub-basins, both, within and outside the WAB). In this respect, Wadi Natuf is highly representative for many basins with similar climatic, geological, soil type and landscape conditions.

Table H1: Recharge coefficients in the literature.

author	year	avg	max	area	time	based on	how	method	miscellaneous
Sheffer *et al.*	2010	9-40 %	(52 %)	**Mt.Aq.**	annual	model	calibrated numerical water budget model	daily soil saturation	non-empirical assumed FC (low) and soil thickness (high)
Gunkel *et al.*	2015	30 %	-	**Mt.Aq.**	annual	model	distributed, field-based, water balance model	TRAIN-ZIN, MOSST	EAB, Wadi Fari'a + Jordan Valley, 6-year measurement period, event-based
Sanz *et al.*	2011	34.3 %	34.3 %	Spain	annual	lithology	clay 1.3 %; silt 3.3 %; lst+dol 34.4 %	large areas, SMD + HEP	as factor of lithology & precipitation
Cheng *et al.*	2017	-	42 %	China	event	thick soil	max annual: 32.4 %	>2m thick soils	in one event
Bradford *et al.*	2002	39-45 %	45 %	UK Chalk	annual	models	with input data at varying spatial resolutions	SMD + HEP methods	UK Chalk, 4 years
Nativ *et al.*	1995	8-55 %	-	N Negev	annual	bromide	in boreholes in fractured chalk	bromide profiles	arid desert climate (P: 200mm/a)
Abusaadah	2011	**36 %**	53.2 %	**Mt.Aq.**	annual	flow model	WAB basin-wide water budget	3-dimensional flow model	max. recharge in 2003/04
Dvory *et al.*	2016	34 %	54.3 %	**Mt.Aq.**	annual	model	1-D dual porosity model (linear)	based on rock properties, karst	local, W of Jerusalem, spatial litho-based differences (abs. maxima: 73.7 %)
Schmidt *et al.*	2014	25-50 %	55 %	**Mt.Aq.**	annual	spatial	Rain, Etp (spring flow) data	SMD+GW model, spring response	EAB at 'Auja spring
Sheffer *et al.*	2010	20-60 %	-	**Mt.Aq.**	annual	formula	quotes literature on WAB-RC	single-cell flow model	calibrated numerical annual water budget (non-empirical), linear equation
Ries *et al.*	2015	33 %	66 %	**Mt.Aq.**	annual	model	max. in wet years (avg. in Aujah)	SM model, budget	EAB (mountains and slopes)
Guttman *et al.*	1988	**30.6 %**	75.3 %	**Mt.Aq.**	annual	model	WAB basin-wide water budget	quasi-3-dimensional flow model	max. recharge in 1991/92 (*single-cell water balance model*)
Weiss & Gvirtzman	2007	**6-65 %**	**-**	**Mt.Aq.**	annual	formula	of Guttman ('00), G. & Zuckerman ('95)	quasi 3-D flow model	applying formula to 4 spring areas, one of which overlaps with Wadi Natuf
Radulovic *et al.*	2011	66.3-76.2 %	**76.2 %**	Montenegro	annual	spatial	water budget & other methods (Montenegro)	ranking system and matrix	karst, lithology, structure, landform, runoff, plants, climate, soil, slope
Allocca *et al.*	2014	50-79 %	**79 %**	S-Italy	annual	lithology	precipitation & air temperature	empir. model, linear regression	correlated with other factors: karst outcrops, morphology, land use, soil type
Goldscheider & Drew	2007	-	**>80 %**	-	event	spatial	bare karstified limestones	-	summary from literature
Weiss & Gvirtzman*	2007	**<10->80 %**	**91 %**	**Mt.Aq.**	annual	lithology	max. in 1988 at l-LBK ftn.	model, budget, spring response	28-a avg. RC of **47.2 %** at Ein Al-Harrasheh spring groundwater catchment
Allocca	2015	35-97 %	**97 %**	Apennine	event	model	storm-by-storm, Apennine	air temp., ET & rain	
Shacori	1965	**47 %**	-	**Mt.Aq.**	annual	spatial	Mt Carmel avg. (undifferentiated)	-	
Mero	1958	**53 %**	-	**Mt.Aq.**	annual	spatial	Na'aman spring	-	(quote after Lerner et al. 1990)
Rosenzweig	1972	**60 %**	-	**Mt.Aq.**	annual	landform	Mt Carmel avg. (by landform)	annual on pasture/grassland	while at forest 0 %
Messerschmid *et al.*	2018	**40.8-47.3 %**		**Mt.Aq.**	annual	3 groups	here: alt.-2: soil-depth group	SM-model, ranking	WAB in Wadi Natuf

* for the groundwater catchment of Ein Al-Harrasheh spring, which overlaps with Natuf surface catchment, Weiss & Gvirtzman (2007) reported a 28-year average recharge coefficient of 47.2 %.