Authors’ response to Reviewer 2

General comment

The manuscript describes a modeling study of the spatial and temporal variation of recharge in a 2.16 km2 upland catchment in a semi-arid region. Recharge in semi-arid regions constitutes a small fraction of precipitation and is subject to a large temporal and spatial variability. Studies of this hydrological component under semi-arid conditions are relatively few although the references provided by the authors are all more than 10 years old and should thus be updated when revising the manuscript. Nevertheless, I believe that the presented study expands research on recharge in semi-arid regions and that the manuscript deserves publication after revision.

Author’s response: We thank Reviewer 2 for the thorough review of the paper that highlighted, in particular, a lack of clarity in the calibration section. We think that, replying to the comments, we explained our rationale for the proposed approach.

Major comments

1) Comment from Referee: My major concern of the presented work relates to the calibration of the MIKE SHE model, which is inadequately carried out and described. Calibration of a hydrological model should preferably be carried out using an autocalibration method (e.g. PEST) in order to (1) identify the sensitive parameters, (2) calibrate the parameters selected for calibration using an objective method, (3) identify non-uniqueness issues and correlation among the parameters, and (4) identify uncertainty intervals of the calibrated parameter values. The process can be carried out in a more or less sophisticated procedure but in any case it makes the process transparent. The authors do not describe which parameters have been subject to calibration and it is not discussed if the resulting parameters values are reasonable based on prior knowledge of the characteristics of the site. I will encourage the authors to carry out a sensitivity and calibration analysis using an autocalibration method.

Author’s response: Thanks to the reviewer’s comment we realized that the text provided in the calibration section was not enough to explain our approach. We will make substantial edits to this section to better describe the rationale and to provide more information about the sensitivity and uncertainty. Here we provide a more in-depth description of our approach.

The parameters involved in the calibration process were surface roughness, detention storage, imperviousness, rooting depth, Leaf Area Index, crop coefficient, hydraulic conductivity and water content parameters of alluvium, hydraulic conductivity and water content parameters of weathered bedrock. We understand that an “autocalibration” approach would have provided additional information and transparency on parameter sensitivity and the uncertainty in recharge estimates. However, we consider our calibration approach to have been rigorous, whereby we tested a range of parameters supported by a large set of field data, against an objective function comprised of groundwater level and stream flow measurements. Results of multiple calibration simulations were compared against these observations and other data (e.g. chloride mass balance) which helped to further validate the calibration.

The calibration process proceeded in an iterative manner. After each calibration run, the primary calibration parameters were examined with a variety of metrics including:

Streamflow Calibration Metrics

- Simulated vs Observed Average Annual flow
    - Mean Error
- Simulated vs Observed Average Monthly Flow:
    - Mean Error
    - Root Mean Squared Error
    - Correlation
    - Nash Sutcliffe Efficiency
- Graphical Plots of Simulated Streamflow Versus Observed Streamflow and Precipitation
    - Provided a qualitative measure of event correlation to observed precipitation and streamflow

Groundwater Level Calibration Metrics

- Simulated versus observed water levels
    - Mean Error
    - Mean Absolute Error
    - Root Mean Squared Error
    - Normalized Root Mean Squared Error
- Graphical Plot of Simulated Vs Observed Water Levels (1:1 residual plot)
    - Provided a quantitative and qualitative assessment of the residual error present at observation wells throughout the domain
- Spatial Plot of Groundwater Residuals (map)
    - Provided a quantitative assessment of water level residuals plotted in the model domain
    - Spatial patterns of fit or misfit of the model were compared against other spatial data (e.g. hydraulic conductivity, boundary conditions, land uses, surface geology) to evaluate potential correlations.

Following an assessment of these calibration targets, model parameters were revised to improve representation of the calibration metrics. In instances where the results were not consistent with the site conceptualization consideration was given as to whether an alternative conceptualization would explain the results predicted by the model. From a semi-quantitative assessment of the calibration process, the values of hydraulic conductivity and water content parameters of alluvium and weathered bedrock had the strongest impact on the calibration targets. These deposits represent the upper layers of our model domain and variations in hydraulic conductivity or unsaturated zone properties control the rate of infiltration, evapotranspiration drainage and, therefore, recharge.

The fact that the values are 1) in the same range of those present in literature (Canadell et al., 1996; Scurlock et al., 2001; Chin et al., 2000), 2) similar to those used by the Surface Water Expert Panel (https://www.boeing.com/principles/environment/santa-susana/technical-reports.page) to model surface water flow, 3) in the range of those measured in the groundwater zone during on-site investigations conducted for 20 years (Cherry et al., 2009) helps to significantly constrain the calibration.

2) Comment from Referee: My second major concern relates to the conceptualization of the system being studied. The subsurface consists of densely fractured bedrock with parallel beddings and vertical joints and faults leading to preferential flow as also emphasized by the authors at several places in the manuscript. For interpreting chloride and isotope concentration measurements preferential flow appears to be important. Furthermore, the authors have developed a conceptual model for recharge, where distribution between matrix and fractures is described (l. 469-479). The flow processes in and between the two domains are mainly based on speculation and not documented by modelling. The authors need to substantiate why two domains are not considered in their modeling approach.

Author's response: Actually, in a previous published paper, the roles of matrix and preferential flow were examined in detail. Analyzing the different average Cl concentration in the vadose zone and in groundwater, Manna et al. (2017) estimated that 80% of the recharge occurs as intergranular flow in the porous matrix block and 20% as fracture flow. Therefore, we think that an EPM model, such as MIKE SHE would reproduce accurately the bulk (matrix -predominantly- and fracture) flow. In addition, the spatial resolution (20 by 20 m cells) is such that the hydrogeological system can be approximated by an EPM model. Our confidence regarding this latter point comes also from the good comparison between the simulated recharge and the observed water level at monitoring wells.

The "conceptual model" section includes findings of previous studies that are incorporated and analyzed in the light of the outcome of the present paper to create indeed a conceptual model. This is why we mention the possible occurrence of preferential flow in the deeper vadose zone, which is not simulated with MIKE SHE but analyzed in previous studies. The text on flow mechanisms at line (469-479) is there to complete the discussion about recharge and unsaturated zone dynamics but not to describe processes that are embedded in our modeling.

Specific comments

1) Comment from Referee: l. 66-75: Please update literature review with newer references

Author's response: We updated the literature following also the suggestions of reviewer 1. However, we want to highlight the surprisingly lack of integrated spatially distributed models for semi-arid catchments in recent years.

2) Comment from Referee: l. 103-104: As fracture flow is stated to be an important flow process the authors need to substantiate why this flow process is not considered in the modelling.

Author's response: see response to major comment 2.

3) Comment from Referee: l. 153-156: Is the lateral boundary condition a closed boundary? Is the lower boundary condition based on field measurements? To which extent will it impact the modeling results? Do I understand correctly that groundwater does not contribute to stream flow and that all recharge will to deeper aquifer systems? Please elaborate on the model conceptualization.

Author's response:

There is a fixed head boundary conditions applied to the base and along the lateral faces of the model representing the deep groundwater flow system. The shallow water table and perched systems within the alluvium and weathered bedrock are well above this deeper water table. These heads are based on observed groundwater levels at the site and simulations based on a detailed groundwater flow model. Given that the groundwater heads associated with deep aquifer system are generally observed at relatively large depths below ground surface throughout the domain, it is expected that variations in these specific values assigned would not have a significant effect on predicted recharge values. In areas where the groundwater is observed to be closer to ground surface, the alteration of these values could potentially have a more direct effect on groundwater recharge in that a groundwater table close to the surface could rise to meet the ground surface given sufficient recharge.

It is correct that groundwater contribution to streamflow is minimal and only occurs after rainfall event at the farthest downstream location of the catchment. At the outlet of the catchment where the groundwater table rises close to the ground surface and there is intermittent and limited (~ 0.1 mm y^{-1} for the period of 1995-2014) contribution of groundwater to streamflow leaving the system.

4) Comment from Referee: l. 178-179: What are the thicknesses of the two groundwater zone layers?

Author's response: Layer 1 has a thickness variable from 24 to 185 m (average: 109 m) whereas layer has a uniform thickness of 5 m. While layer 1 may appear very thick the 'active' part from a numerical perspective begin only when the water table is reached. Flow above that occurs in the unsaturated zone that features a finer discretization.

5) Comment from Referee: l. 189: Table 2 is incomplete, unsaturated zone characteristics should also be listed.

Author's response: The table has been completed with porosity, field capacity, residual water content and the Van Genuchten parameters (α, n) used in the model.

The model uses three separate sets of Van Genuchten parameter to represent the pressure-saturation-hydraulic conductivity relationships; 1) alluvium, 2) weathered bedrock, 3) un-weathered bedrock. The parameters used reflect our understanding that the rock matrix transmits the largest volume of recharge, while recharge through the fractures is faster. The relationships used are biased towards the matrix response. These values were further calibrated using the groundwater level responses and the stream flow. Further rock core samples indicate a high moisture content (~80%) indicating that K is often close to K_s and the hydraulic conductivity-saturation curve reflects this understanding.

					Van Genuchten parameters		
Hydrogeologic unit	**K_s (m s^{-1})**	**Saturation (θ_s)**	**Field capacity (θ_{fc})**	**Residual Water content (θ_{fc})**	**α**	**n**	**l**
Alluvium	1×10^{-6}	0.4	0.25	0.05	0.021	1.61	0.5
Weathered bedrock	2×10^{-7}	0.2	0.11	0.01	0.033	1.49	0.5
Unweathered bedrock	4.1×10^{-10} to 2.3×10^{-7}	0.13	0.1	0.025	0.01	1.23	0.5
Unweathered bedrock	1×10^{-10} to 1×10^{-5}	0.13	0.09	0.01	0.01	2	0.5
Unweathered bedrock	1×10^{-9} to 1×10^{-6}	0.13	0.1	0.025	0.01	2	0.5

6) Comment from Referee: l. 205-211: Could you please be a bit more clear on how the land use are estimated.

Author's response: Land use classes were identified and delineated based on aerial imagery and local land cover datasets (Davis et al., 1998). Descriptions of vegetation classes and species were used in conjunction with literature values for vegetation rooting depth and leaf area indices to describe local vegetation within the model.

7) Comment from Referee: L140 – l. 280- : The calibration procedure needs to be elaborated and revised as described above.

Author's response: see main comment 1.

8) Comment from Referee: l. 301: Generally, I would consider a mean absolute error of 4.5 m to be rather high. Perhaps you mean root mean square error?

Author's response: We agree that 4.5 might be seen as high error. However, we are in a recharge area, on a topographic high with hundreds of meters of head potential. In addition, given the complex structural setting (faults located in the deeper system -not modeled), the heterogeneity of the media (porosity ranging between 2 and 20% within meters observed in rock cores, hydraulic conductivities between 1×10^{-5} and 1×10^{-10} m s^{-1}) and the vertical discretization of the model around the water table, we think that 4.5 m is a reasonable mean error.

9) Comment from Referee: l. 303-: To me it would make more sense to compare simulated and observed hydraulic heads directly?

Author's response: Unfortunately, given the subsurface heterogeneity (see response to comment 8) and the spatial variability of recharge is difficult to reproduce reliable transient head time series. The goal of showing this comparison is to validate the ability of the model to reproduce transient recharge conditions over the studied catchment and we think that the good match between simulated recharge and observed water level provides this confidence.

10) Comment from Referee: l. 316- 318: Perhaps the equivalent porous medium approach is suitable for simulation of water flow but for solute transport and the interpretation of chloride and isotopes I am not sure.

Author's response: Agree but this is truer for the saturated zone than for the vadose zone. As explained in the response to the main comment 2, a previous study found that at the site recharge (and Cl transport) occurs mainly as intergranular matrix flow in the vadose zone. Therefore, we think that our EPM model can be corroborated by recharge studies based on the Chloride Mass Balance method and that the isotopic composition of groundwater can be interpreted under an EPM conceptual model (especially because the Et zone is made of alluvium and weathered bedrock).

11) Comment from Referee: l. 352: Fig. 8a and 8b.

Author's response: Ops! We replaced 7b with 8b.

12) Comment from Referee: l. 373: Check consistency with lines 216-217.

Author's response: Thanks. We made it consistent.