Response to the comments of referee#2 (anonymous) on the manuscript “A distributed soil moisture, temperature and infiltrometer dataset for permeable pavements and green spaces”

We thank referee#2 for the valuable comments that will help us in improving the quality and readably of the manuscript. We are deeply grateful for that. We assigned the comments into the three categories text errors, provided data and data uncertainty.

Text errors

R2 C1: *The specific language seems quite awkward and potentially distracting in places. I itemize some of those errors below but I have no doubt that I missed many of them. These arise at least in part from German-to-English mis-translations. The journal / publisher will pick up some of these errors at the proof-reading step but I think that responsibility for these corrections lies with authors, not the journal. I strongly recommend that the authors engage a scientific technical editor to read and revise this text.*

We will revise the whole manuscript and have it checked by an English native speaker to improve its readability.

P2, L9: “alternated” should be altered

Corrected

P6, L7: “Thereby”

Changed into this

P7, L2: “see chapter data availability”, ‘chapter’ as used here refers to a book or thesis, not to this paper

According to the author guidelines of ESSD, the abbreviation Sect. is used throughout the manuscript instead of “chapter”

P11, L20: flashy?

Explanation added in the manuscript (fast rise and recession of soil moisture)

Provided data

R2 C2: *Page 7 line 14: Authors mention evapotranspiration here (and provide two data files, one daily and one hourly) but then make no further mention or use of them or the data. Again, a residual remaining from a separate publication or thesis?*

In the following, we explain the reasons for providing reference crop evaporation (et_0) with different temporal resolutions. We will add these reasons in the manuscript.

Although not analyzed within the manuscript, we decided to provide et_0 since it is a key variable for most hydrological studies. Since a high temporal resolution might be desirable for further users, we provide et_0 with an hourly temporal resolution. However, the time step recommended by Allen et al. (1998) for the calculation of et_0 is one day. Therefore, we decided to include also daily values for et_0.

R2 C3: *One often finds in this text file as well as in several others, very strange formatting errors, e.g air temperatures of 4.0489999999999995 or, in the metaPlot.txt file, GPS values of 7.85091691909999994.*
Note errors in metaPlots.txt file: latitude and longitude apparently erroneously reversed for stations G and H

Thank you for pointing out these formatting errors, which will be corrected in the data files. The plot coordinates will be provided with 6 decimal digits, while e.g. air temperature will be provided with a precision of 2 digits.

Uncertainties

R2 C4: *ESSD, according to it guidelines (https://www.earth-syst-sci-data.net/10/2275/2018/) requires explicit detailed description of uncertainty factors plus careful validation. I understand that, due to the unique nature and scale of these urban measurements, validation may prove difficult. However, the manuscript as presented remains woefully deficient on uncertainties.*
The authors seem to assign uncertainty solely to sensor performance. For example, at page 10 lines 10 to 12, the authors merely recite manufacturer's performance data. But in fact they have a whole cascade of uncertainties among which manufacturer sensor performance may prove small.
A rigorous uncertainty analysis necessitates careful accounting of the full range of uncertainty factors. I do not contend that users should consider any of these data as 'wrong' but neither should we consider them - as these authors apparently do - as absolute. Soil moisture, soil temperature, saturated water content, etc. all have associated uncertainties. Readers need to know those uncertainties, need to know that the data providers recognize those uncertainties, and need to know - as we currently can not - how large an impact those uncertainties might or might not have on the validity of these data.
The authors hope to see these data useful in the context of model calibration or validation, but most models require quantified uncertainty ranges.

Preliminary notes:
Indeed, there are different sources of uncertainties affecting the measurements and the parameters presented within this manuscript. We are grateful for the comments of reviewer#2 highlighting these uncertainties. Since the data originate from point measurements, we focus on the uncertainties of the measured and derived quantities, while problems of scale and time are not discussed.

R2 C5: *The climate source data (from WBI) must have substantial uncertainties. At a quick glance one sees many RH values near or at 100%, values in the highly-uncertain range for most humidity sensors.*

There are four different climate stations available for the study area, which are all operated by different institutions. In the manuscript, we provide a link to the data of each station. The focus of the manuscript is on soil moisture, soil temperature and infiltrometer data. Hence, we think that a comprehensive discussion of climate data uncertainty is beyond the scope of this manuscript. However, we will include the following remarks into the manuscript:

Data of the individual climate stations differ in resolution, documentation, provided variables and vicinity to soil moisture clusters. Therefore, the selection of the climate data should be purpose-specific. Documentation is best for the DWD climate station, which is operated according to guidelines of the World Meteorological Organization. For the WBI climate station, online available data is limited to an hourly temporal resolution, while data with a higher temporal resolution is available only upon request. To facilitate the use of high resolution data for the WBI climate station and to ensure its long-term availability, we asked for the permission to include this data in our data repository.

Note that the link to DWD data has changed and will be updated in the manuscript. Furthermore the station-ID will be added.

R2 C6: *Need to add variability in specific locations and PP types*

Indeed, this variability may be important for interpreting the data. We will provide this information by:

- Adding images of the PP surfaces to the data repository which show the variability of the PP surface. These images cover an area of 1 m^2 and consist of digitized paving stones (black) and joints (white).

- Adding a file metaClusters.txt to the data repository. For each cluster, this file will contain a column with the fraction of different urban structures (buildings, asphalt, PPs and green spaces) within a 5 m and 10 m radius around the clusters. This data will be obtained by means of a GIS analysis and will capture the variability in urban structures in the surrounding of each cluster.

R2 C7: *Need to add uncertainties in the infiltration measurements*

For the infiltration data, uncertainties comprise the measurement accuracy (approx. 0.5 mm for the visual observations) and the parameter uncertainties of the fitted Philip infiltration model. While measurement errors over the entire infiltration course are small (measurement of a cumulative quantity), the latter is included in table 3 of the manuscript and will be further included in the file metaPlots.txt. For sake of the clarity, we decided to not include uncertainty bands for the fitted Philip model in Figure 10.

R2 C8: *Need to add uncertainties in the CRIM equation*

Given that the CRIM equation correctly describes the physics of the system, uncertainties arise from measurement errors (dielectric permittivity and temperature) and parameter uncertainties (porosity, permittivity of the gaseous and solid phase). By estimating these input uncertainties, Roth et al. (1990) calculated their effect to not exceed 1.3 vol.%. We will add this uncertainty in the manuscript.

R2 C9: *Figures 6 through 9, which ought to help us understand the value of the data, have no indications of uncertainty.*

Uncertainty bands will be added to Fig.6-9

R2 C10*: Files of permititivity, soil moisture, soil temperature, etc., have no indications of uncertainty.*

Indeed, the uncertainties should be indicated in the data files. We will add the uncertainty of each measured/derived variable in the comment line of the corresponding data file. Thereby, the error of the permittivity measurement, is estimated from the data reported by Bogena et al. (2017) who tested 701 SMT100 sensors in reference liquids with known dielectric properties. Their results indicate that the error of the permittivity measurement is below 1.5.

R2 C11: *Several times the authors mention "means" of all locations or all depths, but we never read nor see anything about standard deviations, standard errors, etc.*

Thank you for this comment. We will add the standard deviation for the median of the derived hydrologic parameters in Table 3 and in the text. Furthermore, Fig. 6-8 show mean values of soil moisture and soil temperatures over different sensors. Including the standard deviation in these Figures would be confusing. Instead, we will provide the standard deviation in a separate table.

R2 C12: *At the top of page 19 (lines 2 thru 4), the authors write "The plots E1 and E2 are equal in terms of joint properties and proportions, which leads to the assumption that infiltration measured at E1 might be applied as representative for E2." I appreciate that the authors used the cautionary word 'might' but this reader find no basis elsewhere in the text, particularly assurances on uncertainties, that would allow me to accept similarity of E1 and E2.*

We thank reviewer#2 for scrutinizing the assumed similarity in the infiltration patterns between plots E1/E2 and plots F1/F2. Due to the fact, that these plots were constructed during the same field campaigns (same age, similar soil material used for base and bedding layers), have the same proportion of joints and are exposed to similar microclimatological conditions, we think that this assumption is reasonable. However, since infiltration patterns may vary on small spatial scales, we will remove this assumption in the text and for the infiltration parameters (A, S and i_{cap}) in the file metaPlots.txt.

R2 C13*: A large uncertainty factor, at least for this user/reviewer, relates to solar exposure. How much direct solar radiation or shading by buildings or vegetation occurred at any site? For these latitudes, shade can influence soil temperatures by 10°C or more, e.g. 50% or more of total diurnal ranges described here. Intensity of shade, diurnal pattern of shade, seasonal pattern of shade - we get none of this information and - apparently - no hints about how we might retrieve such data. Clearly the authors know more about solar radiation and local exposure factors than any users will ever know, but we get nothing?*
On page 17 line 12 one reads about station D as located "an east-west orientated urban canyon within the city center." Using lat lon coordinates from metaPlots.txt file to locate the stations in Google Earth, and then applying the GE 'street view' function, I confirm the narrow streets and tallish buildings around station D, but I also find more dispersed but taller (5 or 6 stories?) buildings around station H, albeit with different E-W N-S orientations. From those two explorations (which I might have done wrongly, see note about lat lon below), this reader remains just as concerned and perhaps more concerned about insolation and shading effects. Authors must have recognized insolation effects, must have assessed and selected locations with solar exposure in mind, but they have shared none of that information with readers? They offer readers neither tools nor information needed to assess such a large uncertainty factor?

Indeed shading and insulation have a decisive effect on ground surface and subsurface temperatures. Adding the necessary information to the publication will certainly improve the usability of the data. Since shading and insulation is variable over time, the effect cannot be quantified by a single number. To enable for a time-dependent quantification of the effect, we will include hemispherical photos of each cluster in the data repository. From these images, the sun path, but also the mean radiant temperature can be calculated by using a suited model. One example for such a model is 'RayMan' (Matzarakis et al., 2007).

C14: *Page 11 lines 6 thru 12: Here the authors describe uncertainties related to freezing conditions and possible salt applied as anti-freeze, e.g. reasons for not using winter-time data, but we never find any cautions about uses of the data they do provide!*

The soil moisture data set does not contain data for freezing conditions (see P11 L6). However, the permittivity data set contains data for these periods. In the manuscript, we will emphasize that a suited model is required when relating permittivity to soil moisture for these periods (see answer to comment 3 of referee#1).

The other point concerns the usage of salt as an anti-freeze, which affects electromagnetic soil moisture measurements. Indeed, a further note on the uncertainty arising from this practice may be useful for other users and will be included in the manuscript.

R2 C15: *P11, L25-26: Characterization as vegetated, restricted or free. But, according to Table 2, they only analyzed 3 vegetated sites and 4 restricted drainage sites. Given many other sources of spatial variability and uncertainty, can the authors provide any quantitative basis that we should accept these categorizations?*

Indeed, there is only a small number of plots for the categories “vegetated” and “PPs with restricted drainage”. The small number of vegetated sites is due to the fact that only four of the clusters were in the close vicinity of urban green spaces.

The quantification into the categories free drainage and restricted drainage is based on a combination between a visual classification and an analysis of the empirical frequency distribution of soil moisture records during rain events. As illustrated in Fig. 7 the mode of these frequency distributions can be used as a quantitative measure for this classification. However, using a single threshold would neglect the effect of soil properties on soil moisture. In order to still provide some quantitative measure, we will add the mode of the empirical frequency distributions to the file metaPlots.txt.

Literature

Allen, R. G., Pereira, L. S., Raes, D. and Smith, M.: Crop evapotranspiration - Guidelines for computing crop water requirements, Rome, Italy., 1998.

Bogena, H. R., Huisman, J. A., Schilling, B., Weuthen, A. and Vereecken, H.: Effective calibration of low-cost soil water content sensors, Sensors (Switzerland), 17(1), doi:10.3390/s17010208, 2017.

Matzarakis, A., Rutz, F., Matzarakis, A., Rutz, F. and Mayer, H.: Modelling radiation fluxes in simple and complex environments - Application of the RayMan model Modelling radiation fluxes in simple and complex, Int J Biometeorol, 51, 323–334, doi:10.1007/s00484-006-0061-8, 2007.

Roth, K., Schulin, R., Flühler, H. and Attinger, W.: Calibration of time domain reflectometry for water content measurement using a composite dielectric approach, Water Resour. Res., 26(10), 2267–2273, doi:10.1029/WR026i010p02267, 1990.