In this study, the simple soil moisture model developed by Pan et al. (2003; 2012) was calibrated using in-situ soil moisture measurements from the Soil Climate Analysis Network (SCAN). The optimised parameters are transferred to ungauged locations via a hydro-climatic classification system and the decrease in soil moisture prediction capability is analysed.

*We appreciate the reviewer's thoughtful remarks and his appropriate synopsis of the work that has been done. We recognize that this work's limitations must be more explicitly stated within the text to ensure that we do not overstate the accomplishments of this analysis. We hope our responses to the comments and the changes made to the text will be satisfactory.*

The authors claim that the extrapolation of model parameters to ungauged locations via a hydro-climatic classification system will enable near real-time irrigation decision making. They showed that the decrease in soil moisture prediction capability is less within a hydro-climatic class than between different classes. However, from existing 200 SCAN sites they selected only 15 for this analysis of which most are not located in regions with significant irrigation demand. If the main goal of the study is to support irrigation decision making this study should focus on more relevant sites.

*The reviewer is correct that irrigation is unnecessary in many of the 15 chosen sites are energy-limited rather than water-limited. The following can be added to the final paragraph of the introduction to clarify the value of these soil moisture estimates:*

> "With respect to agricultural decision-support, for energy-limited sites, the value of hourly soil moisture estimates is found in the determination of whether or not a field is trafficable – whether heavy equipment will damage fields or become mired. With respect to water-limited sites, the value of soil moisture estimates is found in devising optimal irrigation strategies that utilize limited water resources most efficiently. Of the fifteen SCAN sites examined, the three sites in New Mexico, the site in Colorado, the site in Nebraska, the site in Wyoming, and the two in Iowa are all water-limited (8 in total). The remaining sites (7 in total), located in Pennsylvania (2), Arkansas, Georgia, South Carolina, North Carolina, and Virginia, are all energy-limited.

The authors applied the model developed by Pan et al. (2003; 2012) at an hourly resolution without further adapting the method. The model was explicitly developed for daily resolution as it assumes that the daily soil water loss can be described with a sinusoidal function describing the inter-annual change in evapotranspiration (ET) demand. Cleary, at hourly resolution, the changes in ET demand during the day need to be considered as well. This needs further model development. In addition, the model results should be analysed in more detail. For instance, it should be tested whether the sinusoidal function is able to capture the seasonal variation in soil loss (also after application of the machine learning algorithm).

*We agree that, when a daily model is refit for hourly estimation, diurnal patterns of potential evapotranspiration become important. Figure 6 does illustrate the capacity of the model to generate a diurnal pattern of soil moisture despite the fact that the sinusoidal parameters fit in the diagnostic soil moisture equation do not include one. The new figures presented below, along with a substantial explanation to be inserted into the manuscript, help explain the capacity of the machine learning model to account for various features. We hope these explanations and images will help assuage some of the concerns regarding the transition from an hourly model to a daily model.*

> "Figure 7, 8, and 9 present these results in more detail for each of the three SCAN sites presented in Figures 3, 4, and 5. In each figure, the upper-left image presents the average bias correction (change in % soil moisture) for each hour of the day (0-23). At all three sites, bias corrections display a clear diurnal pattern – that is to say the removal of a diurnal cycle is a substantial role of machine learning under a variety of hydroclimatic and edaphic conditions. The upper-right image of each figure presents the bias correction as a function of the unadjusted soil moisture estimate – essentially, whether there exists a systemic over- or underestimation when values are high or low.
>
> The first two sites (Figures 7 and 8) do not present a clear pattern, but Figure 9 displays a trend suggesting that the highest estimates of soil moisture tend to be overestimates and the lowest estimates of soil moisture tend to be underestimates – but these biases are removed via machine learning. The lower-left image presents bias correct as a function of the day of the year (from 100-300, the days of the year when the model is applied). At all three sites, the seasonal cycle does appear in terms of the patterns of bias correction, but the pattern is noisier than the diurnal cycle. The magnitude of the adjustments are largest in the monsoon-affected desert of New Mexico, a bit smaller in the Midwestern plains characterized by less extreme seasonal behavior, and smallest in the Southeast where seasonal variations are low.
>
> Finally, the lower-right image relates bias correction to the beta series from the diagnostic soil moisture equation (Pan, 2012), a convolution of a decaying precipitation time series working backwards temporally from the current time. Stated differently, these charts relate bias correction to the amount of antecedent precipitation (with more recent precipitation weighted more heavily). In Figure 7 (Plains, Silty Clay Loam), the model tends to underestimate moisture when large quantities of antecedent rainfall are present, where in Figure 9 (Woods, Sandy Loam), once antecedent precipitation becomes non-trivial, displays the opposite pattern. This is consistent with the finer Midwestern soils' proclivity for ponding/flooding due to larger proportions of clay. In these cases, larger amounts of rain will soak soils from above, and capillary rise might further soak sensors from below, leading to underestimation from the diagnostic soil moisture equation and subsequent machine learning correction. By contrast, with sandier soils, drainage occurs easily, leading to higher rates of loss than the eta series (Pan, 2012) would predict (there is more available water to lose), leading to overestimation with large amounts of antecedent rainfall."

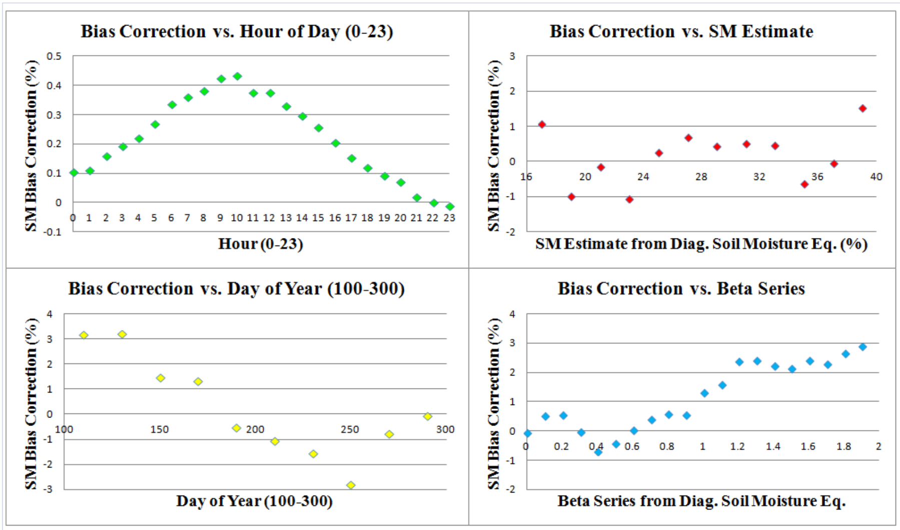


**Figure 7, Bias Correction Analysis, SCAN Site 2015 (IAQ, Desert, Loamy Sand)**

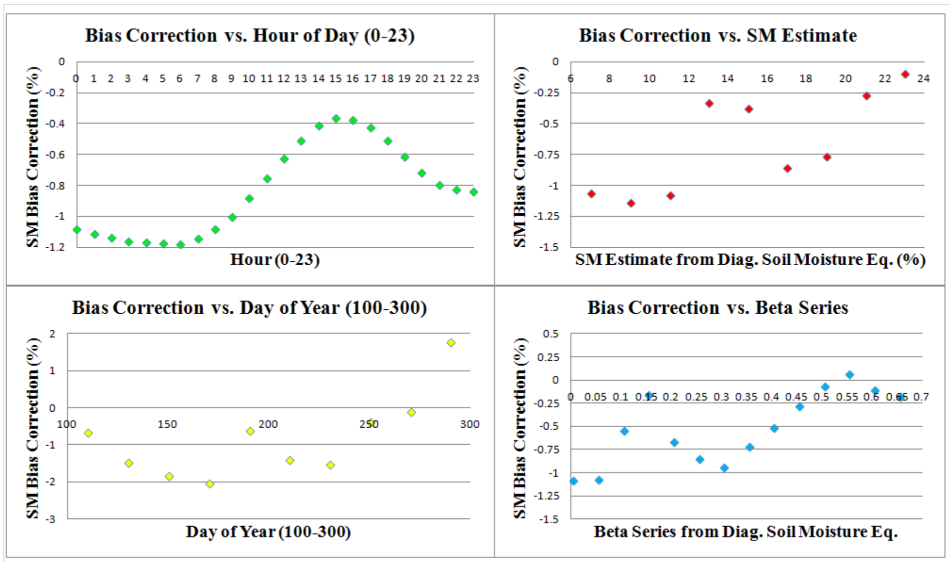


**Figure 8, Bias Correction Analysis, SCAN Site 2068 (ISCJ, Plains, Silty Clay Loam)**

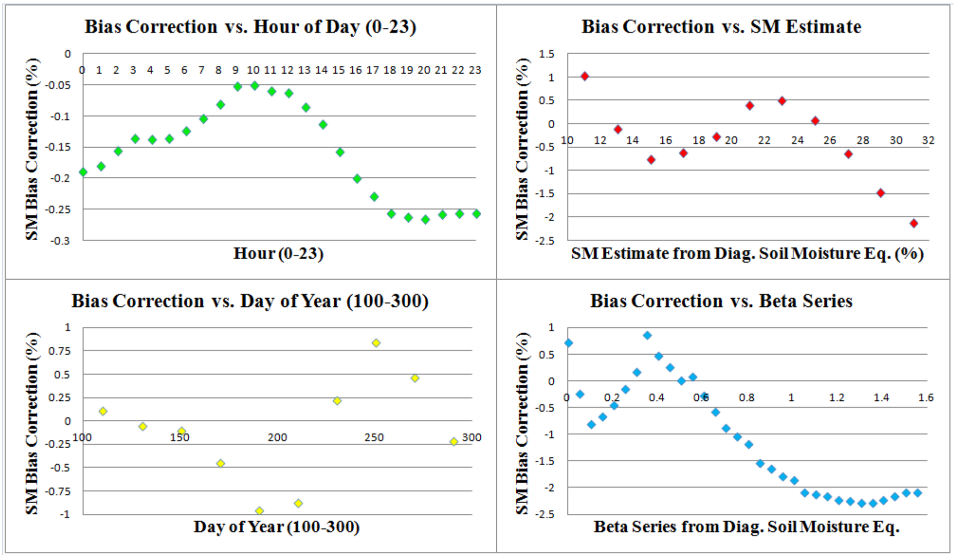


**Figure 9, Bias Correction Analysis, SCAN Site 2013 (LWC, Woods, Sandy Loam)**

The machine learning algorithm that was used to further optimise the model results (bias correction) actually produced also worse results, especially during dry periods (see specific comments for examples). Furthermore, the model even adopts measurement artefacts (see New Mexico SCAN station). These problems need to be critically discussed and the usefulness and limits of the machine learning approach should be challenged.

*Though the reviewer rightly notes that machine learning, at times, causes deterioration in model performance, the new table below (created at the behest of this and the previous reviews) illustrates that machine learning is beneficial at all sites. The following will be added following the table to help address this issue, following Figure 6:*

"Any corrective algorithm will, over thousands of validation points, push the estimate away from the observed value in some cases. However, the results from Table 1 demonstrate that its overall performance represents an improvement at all sites, and thereby justifies its use. Regarding the issue of 'measurement artifacts,' whether the diurnal cycle is genuine or an idiosyncratic sensor output, the model is tasked with calibrating itself and correcting biases as defined by the empirically-reported data. Figure 6 illustrates its ability to do so. Were the sensors to no longer report such a diurnal pattern (i.e. it is merely a measurement artifact, and subsequently corrected), the machine learning step would no longer observe those biases, and consequently, no longer introduce such a pattern. The accuracy of the SCAN network is a relevant inquiry, but unfortunately, not within the scope of this paper."

| SiteID | Hydro-climate | Soil Information | RMSE | RMSE w/ KNN | $R^2$ | $R^2$ w/ KNN |
|---|---|---|---|---|---|---|
| 2008 | LJ | Sandy Loam | 8.38 | 7.69 | 0.590 | 0.726 |
| 2013 | LWC | Sandy Loam | 2.16 | 2.06 | 0.876 | 0.885 |
| 2015 | IAQ | Loamy Sand | 3.29 | 2.37 | 0.740 | 0.841 |
| 2017 | ISQJ | Sandy Loam | 3.62 | 3.27 | 0.637 | 0.701 |
| 2018 | IAQ | Loamy Sand* | 2.23 | 2.16 | 0.803 | 0.828 |
| 2028 | LPC | Loam | 4.89 | 4.71 | 0.707 | 0.738 |
| 2031 | ISQJ | Silty Clay Loam | 5.46 | 6.00 | 0.687 | 0.750 |
| 2036 | LPC | Silt Loam | 4.61 | 3.95 | 0.635 | 0.726 |
| 2038 | LJ | Sandy Loam | 4.81 | 4.51 | 0.546 | 0.584 |
| 2068 | ISCJ | Silty Clay Loam | 5.28 | 4.03 | 0.716 | 0.837 |
| 2089 | LJ | Sandy Loam | 6.7 | 6.31 | 0.682 | 0.697 |
| 2091 | LPC | Silt | 8.12 | 6.89 | 0.539 | 0.808 |
| 2107 | IAQ | Loamy Sand | 1.98 | 1.85 | 0.790 | 0.843 |
| 2108 | IAQ | Loamy Sand/Sand | 1.26 | 1.12 | 0.828 | 0.863 |
| 2111 | ISQJ | Silty Clay Loam | 5.38 | 5.01 | 0.607 | 0.796 |

*Not similar to other sandy soils, see Figure 9.

**Table 1, The Fifteen SCAN Sites: Class & Soil Information and Performance**

The paper is mostly well written and structured. However, given its numerous problems concerning its contents, the parameter optimisation approach as well as the interpretation of the results of the analysis I cannot recommend the publication of this paper. Nevertheless the topic is relevant and well suited for HESS. I therefore recommend a major revision.

*We thank the reviewer for his appreciation of the topic's relevance and hope our revisions will earn his approval for publication.*

Specific comments:
The introduction section has several errors and inconsistencies

- The analytical model proposed by Entekhabi and Rodriguez-Iturbe [1994] does not belong to the API based model group.

*We apologize if this was unclear. The sentences in question, p. 2323, lines 6-9, currently read:*

> "The first group of soil moisture models considers only the variability of precipitation as the primary mechanism for wetting/drying (precipitation (Entekhabi and Rodriguez-Iturbe, 1994). These models often employ an "antecedent precipitation index" (API), defining a pre-established temporal window for antecedent rainfall."

*The intention was to cite the paper from '94 which argued that the primary mechanism of wetting/drying is precipitation, with the subsequent papers as examples of the API approach. To avoid such confusion, these lines will now read:*

> "The first group of soil moisture models considers only the variability of precipitation as it has been shown that precipitation variability is the primary mechanism for wetting/drying, (Entekhabi and Rodriguez-Iturbe, 1994). Many subsequent models employed an "antecedent precipitation index" (API), defining a pre-established temporal window for antecedent rainfall."

- Soil moisture models are typically not subject to recalibration.

*We agree – most soil moisture models are not periodically recalibrated. Our purpose was to state, via Jones (2004), that without recalibration, models based on soil water balance experience cumulative error propagation, which can diminish utility for decision-support. We will amend the offending sentence to read:*

> "While these issues can be addressed using a soil water balance model, this type of model must be recalibrated frequently, which most soil moisture models are not, as its errors are cumulative (Jones, 2004).

- The second group should refer to "process based model approaches".

*Agreed, now p. 2323, line 19, will read:*

> "The second group of models adopts a process-based approach, …"

- It should be mentioned that process based models are typically forced by evapotranspiration demand and precipitation (upper boundary) and, if applicable, by groundwater (lower boundary).

*Agreed. This can be included in the paragraph following the one presented in response to the reviewer's preceding comment.*

> "These process-based models are typically forced by evapotranspiration demand and precipitation at their upper-boundary and, if applicable, by groundwater at their lower boundary."

-HYDRUS does not necessarily needs soil temperature, ionic chemistry etc. for the simulation of soil water dynamics.

*This is true, the intended argument is that HYDRUS is a far more complicated model, which, at a minimum, benefits from having access to numerous pieces of detailed information that are less widely available than precipitation. On p. 2323, lines 21-22 will now read:*

> "…models of this type, such as HYDRUS (Simunek et al., 1998), attempt to improve predictions via detailed knowledge of hydraulic parameters…"

- The third group is actually not model-based, but refers to a field instrument for the characterization of near surface soil properties. Therefore, this group should be omitted.

*We apologize for the confusion, these approaches are 'models,' simply models that require proximal instrumentation and measurement to build out near-surface soil moisture. To help minimize this confusion, line 26 of p.2323 will read:*

> "The third group of models are agriculturally-focused, building model projections outward from existing instrumentation and additional measurements."

- The model of Pan et al. (2003) is a simplification of the linear stochastic partial differential equation proposed by Entekhabi and Rodriguez-Iturbe [1994].

*Agreed. This can be mentioned when introduced on p. 2324. Lines 8-9 can now include:*

> "…observed past rainfall events. This model is a simplification of the linear stochastic partial differential equation originally proposed by Entekabi and Rodriguez-Iturbe (1994).

- Clearly the model of Pan et al. (2003) does not address all shortcomings of other existing modelling approaches.

*We will change the offending sentence to read:*

> "addressed many of the shortcomings of the existing modeling approaches…"

The discussion section has several weaknesses:

- The model of Pan et al. (2003) ignores lateral water flow. Therefore, an application of this model to sites with significant topography is not appropriate without considering lateral flow processes, which cannot be implemented in meaningful way in such a simple model approach.

*This is a fair critique. To address it, the following will be used to augment section 4.1, which discusses enhancement of these estimates via topographic classification:*

> "The lumped, bucket model is not ideally-suited for landscapes with complex topography. Conveniently, the majority of SCAN sites are placed on relatively flat surfaces. Integration of topographic insights is a fertile area for future research."

- Similarly, the enhancement of the model with respect to overland flow is not reasonable. In addition, since the main goal of the model is to support irrigation management overland flow is not important.

*The reviewer makes a fair point, extending a previous comment about agricultural decision support's utility in wetter (non-irrigated) locations. In addition to that response, we will add:*

> "Agricultural decision-support includes trafficability when wet and irrigation support when dry. While overland flow is perhaps an unneeded component in water-limited catchments where irrigation schemes represent the most significant soil-moisture-related decision, in wetter catchments, in which trafficability is a real concern, such an addition could improve the model.

- The link to the SMAP mission is unclear and should be skipped.'

*We believe it is important to place this work in terms of its larger content. If the reviewer has specific concerns regarding the clarity of this paragraph, we will do our best to address them.*

- Instead, problems and uncertainties involved in the modelling and in the parameter transfer should be addressed in greater detail.

*We hope the new analysis of the bias correction process along with the table displaying results at each location, listing hydroclimatic class, edaphic characteristics, and results pre/post machine learning will help flesh out some of the uncertainties the reviewer mentions.*

- Last but not least the benefit of this study should be presented more clearly.

*We agree. The conclusions section can include, as a penultimate paragraph:*

> "This analysis can improve agricultural decision-support by offering insight into locations that can benefit from targeted irrigation in drier conditions, or conversely, by minimizing risks of ruts and damaged equipment when fields are no longer trafficable during wetter conditions. Scaling the results of these models upward can assist with larger-scale assessments of flood risks or as calibration/validation tools for satellite estimates of soil moisture. Scaling these results downward can help maximize yields. Given the ubiquity of precipitation data, which are the only inputs these models require, better understanding of the transferability of modeled parameters are a step towards far wider availability of soil moisture estimates."

P2323 L7: Delete "precipitation"

*Agreed. The change will be made.*

P2324 L17: Which problems are you referring to?

*The 'problems' are the challenges of calibration at locations lacking sensors, as mentioned in the previous sentence. The sentence in question will now read,*

> "…overcome the issues of calibration at ungauged locations associated with the Pan et al…"

P2326 L11: The index “4” of parameter “c” is unnecessary.

*The equation is taken from Pan et al. (2012), in which the first three constants ($c_1$, $c_2$, $c_3$) are used to fit a sinusoid. Thus, while in this case, the index ‘4’ may be unnecessary, allowing readers to consider the larger derivation presented in Pan et al. without confusion seems important in this case.*

P2326 L16: Delete: “and cannot increase its moisture content”

*Agreed. The change will be made.*

P2326 L27: “n hours”

*Agreed. The change will be made.*

P2327 L1-2: This sentence should be rewritten in a more scientifically way.

*It can be reworded as:*

> “For instance, today’s soil moisture is strongly influenced by yesterday’s rainfall , influenced to a lesser degree by last week’s rainfall, and not influenced at all by rainfall from ten years previous.”

P2327 L10-11: Please change the sentence into: “Soil water loss at hour i, e.g. due to evapotranspiration or deep drainage, is expressed by coefficient n.”

*Agreed, the change will be made.*

P2327 L13: The term “eta series” is not appropriate because it neglects loss due to drainage.

*Eta refers to the nomenclature chosen in Pan et al, not simply potential evapotranspiration. A note to this effect can be added the line the reviewer mentions:*

> “…‘eta series,’ representing losses due to evapotranspiration and deep drainage…”

P2327 L17: In the original model of Pan et al. (2003) this sinusoidal wave function is used to represent the changing evaporation demand during the year. I wonder whether the hourly resolution used in this paper is not violating the assumption of Pan et al. (2003). For instance an additional function to represent the daily fluctuations of evaporation demand could be added.

*The new figures illustrating bias correction as a function of day-of-year suggest that, at least some of this seasonal variation is addressed in the machine learning step. An additional, superimposed diurnal cycle of potential evaporation is an excellent idea for subsequent research.*

P2332 L26: 100 % would mean pure water, therefore change “are measured in percentage terms (0–100)” into “are presented as volumetric percentage”.

*Agreed. The change will be made.*

P2333 L7-11: The authors claim that the high soil water contents are due to flood events occurring only during the validation years. This is, however, not true. First, saturated soil conditions can in principle also happen also during times without flooding. Second, I checked the data from this SCAN Site and found that soil moisture exceeds the given porosity value frequently during the training period (e.g. Nov 2009 soil moisture reached 50 Vol.%). Please revise this part accordingly.

*The reviewer is correct that saturated conditions con occur without flooding. The reviewer is also correct that the SCAN site does exceed the porosity value during calibration. The lines in question will now read:*

> “During the validation period, specifically 2010, wetter conditions were observed than were present during calibration. At this SCAN site, before 2010, the average soil moisture value observed was 28.55%, with only 25% of values exceeding 35% volumetric soil moisture. However, in 2010, the average soil moisture value measured was 33.16% with 45% of values exceeding 35%. The machine learning driven error correction improves the diagnostic soil moisture equation ($\rho = 0.846$) significantly ($\rho = 0.915$), but fails to raise its forecasts to reach some of the wetter conditions experienced in validation. Underestimations of this nature, although detrimental…”

P2333 L20-24: This is not entirely true. During several recessions (e.g. 4300 – 4500, 4700 – 4800) the machine learning correction clearly degrades the previous result (overprediction). This is especially alarming since the degradation happens during dry soil states, making the model more unreliable for irrigation management.

*As with the previous comment to this effect, any machine learning model for error correction will, occasionally, push estimates away from the observed values. The improvement in terms of correlation and RMSE for all fifteen sites suggests that, on the whole, bias correction is beneficial and appropriate.*

P2333 L25-29: First, the machine learning correction code seems to fit the diurnal cycles in the New Mexico data by transforming the annual sinus function into a daily one. This is clearly a violation of the original model of Pan et al. (2003). Second, the reason for the diurnal cycles is not a water related process. Actually it reflects the dependency of the electromagnetic soil properties to temperature change. Therefore, the apparent permittivity, which is measured by the soil moisture probe to infer soil moisture, decreases with temperature (from 88 at 0 C to 76 at 30 C). This is a well-known problem; see e.g. Rosenbaum et al. (2011).

*We will cite Rosenbaum et al (2011) and introduce the possibility that observed diurnal patterns are a property of soil moisture sensors. The following will be added to the discussion following Figure 6:*

"It is possible that the diurnal cycle at some locations reflects a soil moisture probe's dependency on electromagnetic properties driven by temperature change (apparent permittivity) rather than hydrologic processes (Rosenbaum et al, 2011). However, the model's ability to respond to these nuances would not compromise its performance were these nuances subsequently removed."

Figures
Fig. 9 and 10 should be merged.

*Agreed. This is done.*

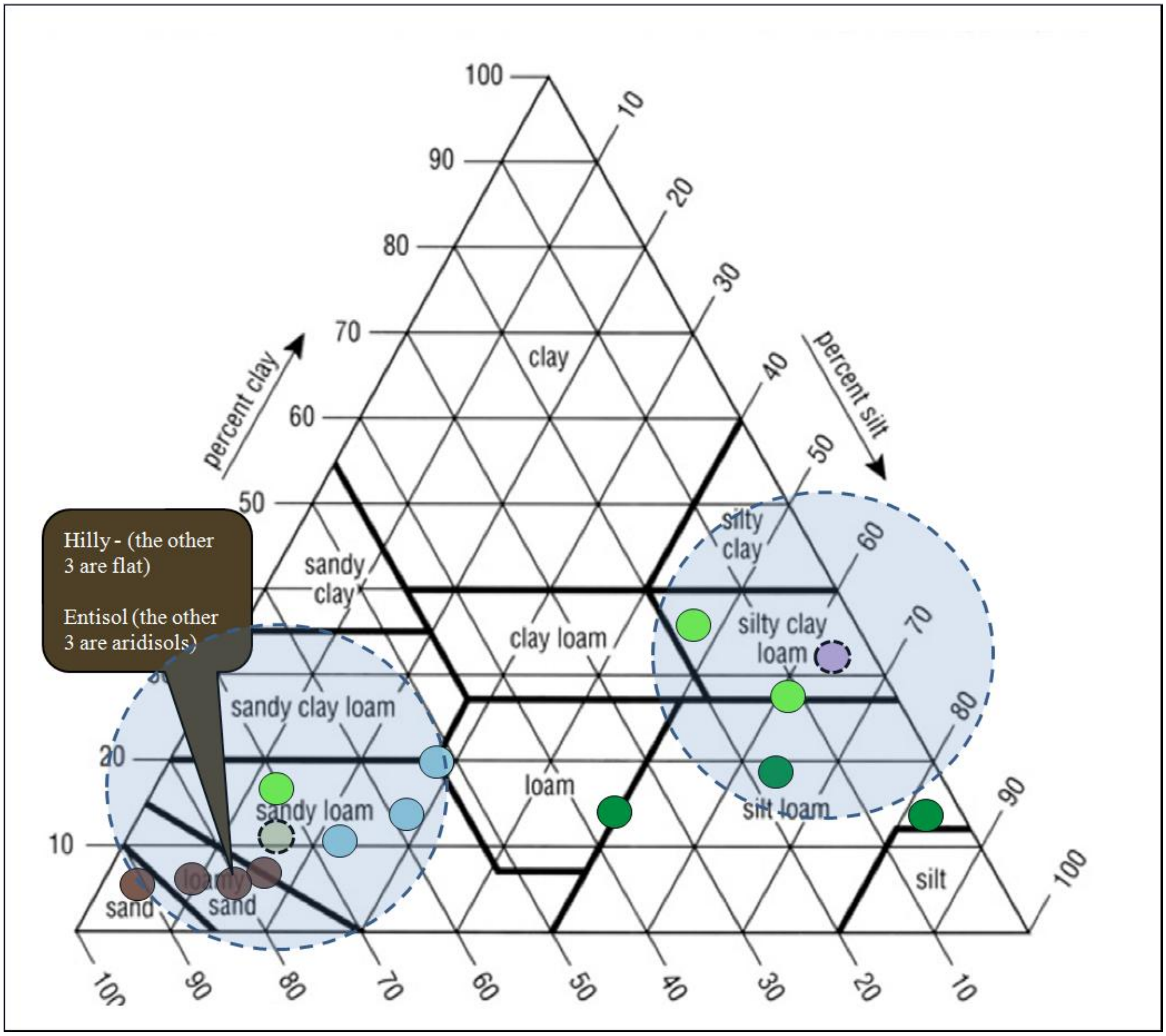


Fig. 11 should be presented as a Table.

*See the table produced above. A larger, multi-layer table would be cumbersome for readers trying to observe every conceivable pair and their relative similarities.*