

Reviewer #1, Dr. Christa Peters-Lidard

General Comments:

The general objective of this work is to extend and generalize the diagnostic soil moisture model of Pan et al, 2003 and Pan 2012 by recasting the work into an hourly time step and calibrating the parameters of the model using SCAN data. The authors then explore how hydro-climatic and edaphic similarity can be used to transfer parameters between sites.

This work is generally well described and relevant to HESS readers. My biggest concerns are the presentation of results, and particularly the lack of consistency in the evaluation metrics among the various analyses. The manuscript really needs a table showing site number, site characteristics such as hydroclimate class and soils, as well as different error metrics (ubRMSE, Bias and R^2) before and after bias correction.

We appreciate the reviewer's timely reply and positive impression of our work. We have worked to improve the charts and tables, along with the descriptions thereof, in their captions and surrounding text. The table described in the paragraph above, one which encompasses the fifteen sites used, their hydroclimatic class, their general soil characteristics, and their performance before & after bias correction has been constructed, now Table 1. This table is then referenced on lines 357-360 (after p. 2335, line 3 in the HESS-D paper).

“Table 1 presents all fifteen sites for which the diagnostic soil moisture equation has been calibrated, including information regarding their hydroclimatic class from Coopersmith et al (2012), their soil textural characteristics, and their performance before and after the KNN bias correction process.”

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 12.

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

The results from the k-NN bias correction procedure shown in Figure 2 show only the total R^2 before and after correction, and only by site number, rather than by something related to the hydro-climatic or edaphic similarity (see suggestion for table above). Further, the importance of the bias correction is likely a seasonal or diurnal bias, and this should also be discussed/shown, perhaps as separate lines on Figures 3-6 or as an add-on to Figure 2 for different types/magnitudes of bias correction. Are the sites shown in Figures 3-6 representative?

The purpose of the initial k-NN bias correction is to demonstrate the utility of that technique in improving the results of the diagnostic soil moisture equation rather than the importance of hydroclimatic or edaphic similarity. Essentially, these figures are intended to demonstrate the utility of our enhancements to the diagnostic soil moisture equation by applying those techniques to results calibrated at these sites. Regarding the importance of the bias correction in seasonal or diurnal terms, this discussion is broached with Figure 6, but should be expounded upon. In Figure 6, the improved 'shape' of the ML-enhanced soil moisture series relative to the observed measurements in blue explains, in part, the improvement in correlation observed. However, as the green line is also nearer to the actual observations than the original diagnostic soil moisture equation results (red), the improvement extends beyond the ability to introduce a diurnal cycle. The new Figures 7, 8, and 9 present the average bias correction as a function of the hour of day, the day of year, the soil moisture value estimated, and the beta series from the diagnostic soil moisture equation (the summed, decaying precipitation series). This allows a deeper assessment of how this machine learning approach is improving soil moisture estimates, forming a new section (lines 389-418) of the revised manuscript.

“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 correction 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.”

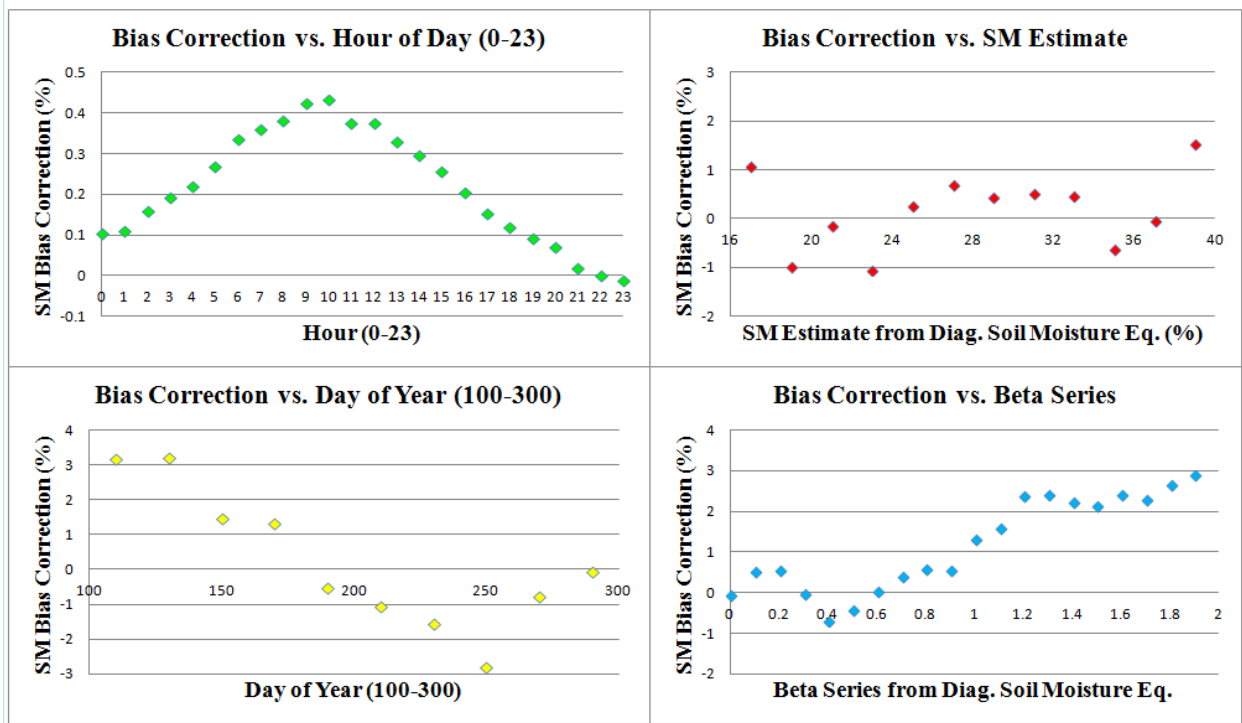


Figure 7, Bias Correction Analysis, SCAN Site 2015 (IAQ, Desert, Loamy Sand)

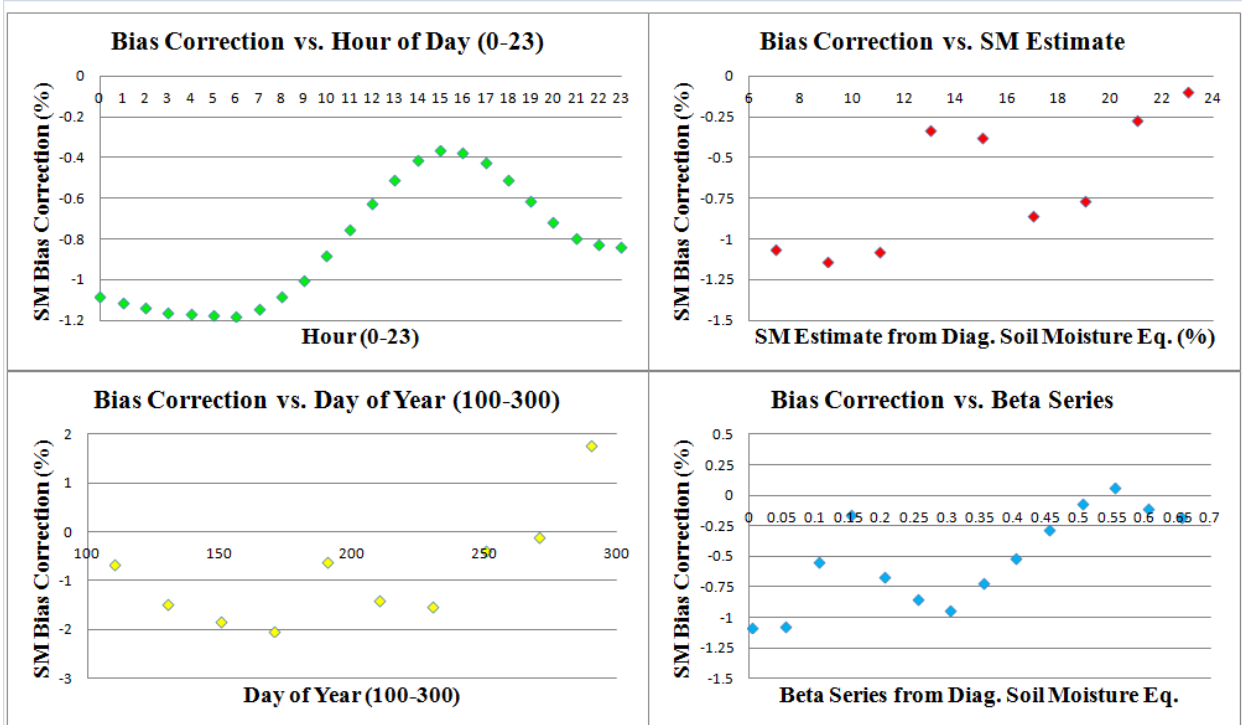


Figure 8, Bias Correction Analysis, SCAN Site 2068 (ISCJ, Plains, Silty Clay Loam)

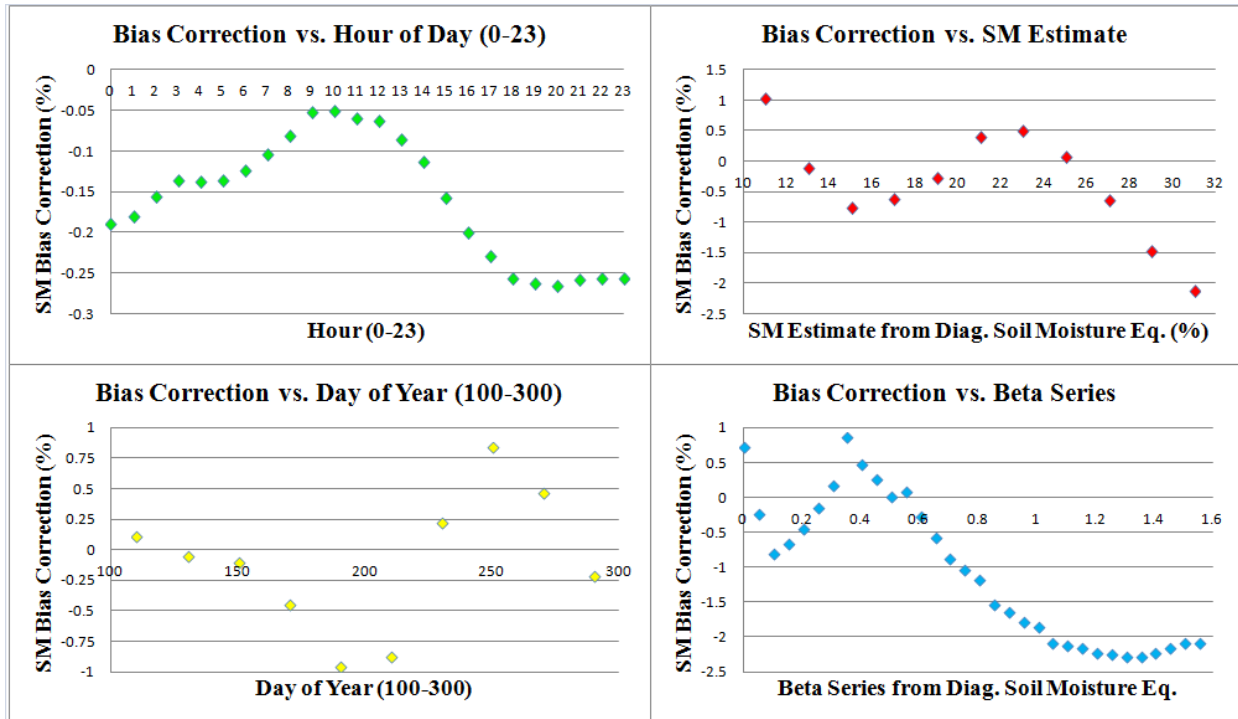


Figure 9, Bias Correction Analysis, SCAN Site 2013 (LWC, Woods, Sandy Loam)

In terms of the representation of the three catchments presented in Figures 3-5, they fall within three different hydroclimatic classes (IAQ, IS CJ, LWC), within three different regions of the continent (Southwest, Midwest, East), with differing soil types (Loamy Sand, Silty Clay Loam, Sandy Loam), and display improvements with respect to machine learning that are roughly consistent with the improvements shown over the fifteen sites examined (expanded upon on lines 323-334 of the edited manuscript). Figure 6, a zoomed-in version of Figure 3, is admittedly a more extreme example of a diurnal cycle (the conditions are very warm and dry in New Mexico, and the fluctuations are large relative to the average soil moisture value) – it was chosen primarily to illustrate the effect discussed rather than to imply that such cycles are always the dominant source of variance relative to the traditional wetting/drying processes.

“To explore these findings in more detail, three of the 15 SCAN sites, chosen to represent different hydroclimatic locations – New Mexico (#2015, hydroclimate IAQ/southwestern desert, Loamy Sand), Iowa (#2068, hydroclimate IS CJ/northern midwest plains, Silty Clay Loam), and Georgia (#2013, hydroclimate LWC/southeastern forest, Sandy Loam) are examined to illustrate how improvements from adding machine learning error models to the diagnostic soil moisture equation differ across sites. These three sites represent three distinct hydro-climatic classes, with significant differences in soil texture, seasonality of precipitation, aridity, timing of maximum precipitation, and timing of maximum runoff. Using error correction models for prediction at these sites increased R^2 -values by an average of 8.2%, which is similar to the 8.3% improvement in R^2 averaged across all fifteen sites. Thus, these three locations are representative in terms of both hydro-climatic and edaphic diversity and their responsiveness to machine learning.”

The most important results in the work are the results shown in Figure 7/Table 1 as well as the Venn diagram in Figure 11. However, similar to the suggestion above, it would be useful to examine these cross-validation results with a common set of statistics as with the initial calibration results. The boxplots do give a sense for the distribution of errors, but the description of what is actually shown in both Figure 7 and Table 1 is a bit confusing. The captions need

more information across the board. Is the loss of R^2 equivalent to the simple difference in R^2 (baseline-new)?

Yes, the difference in R^2 between the baseline values and new (bias-corrected) values is the same calculation shown in Figure 7 and Table 1. We apologize for the confusion and concur that applying a consistent metric is important – we have done so, but need to ensure that this is clear to the reader. Additional explanation appears on lines 407-409 of the edited manuscript to clarify what has been done.

“Figure 10 presents box plots illustrating the change in R^2 values for these three sets of pairs in a manner analogous to the differences shown in Figure 2. Table 2 presents the quantitative results, again averaging the deterioration of performance in terms of change in R^2 .”

The results for “similar” sites are interesting but what exactly constitutes a “similar” site is tersely defined with the discussion of Figure 7: “(different by a single split within the classification tree).” At a minimum you should refer back to Figure 1, where presumably the components of the tree algorithm are derived using MOPEX data.

With the new table constructed (Table 2, discussed in the response to the reviewer’s first point), it should be clearer which sites are considered to be within the same hydroclimatic class (column #2) or within the same edaphic group (column #3). In terms of figure 7, the yellow box-plot, referring to ‘similar’ hydroclimatic classes (different by one-split on the classification tree), additional explanation is certainly required. In this case, ISCJ/ISQJ, LWC/LPC, ISQJ/IAQ, ISCJ/IAQ, LJ/LPC, and LJ/LWC are considered ‘similar,’ as in all of these pairs, tweaking a single hydroclimatic indicator could move the catchment from one class to the other. An additional note is added to indicate that the classes chosen are those determined by more recent data, as the original classes from Coopersmith et al (2012) have undergone climate change since their original classifications using MOPEX data from 1948-2003 (Coopersmith et al, 2014).

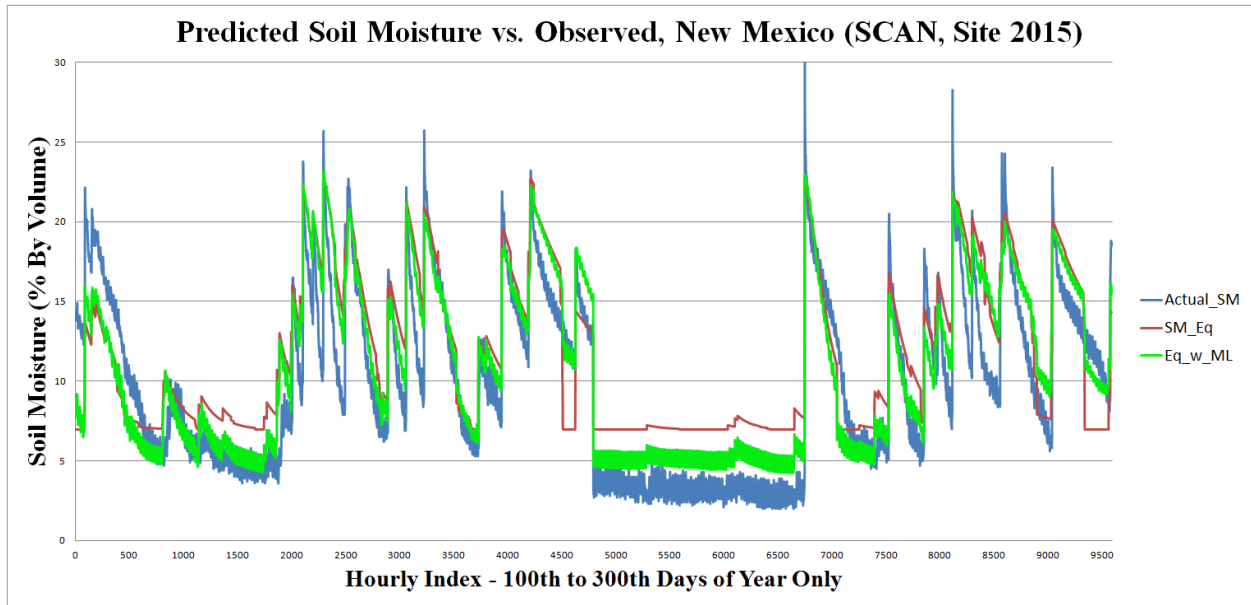
Overall, I think the work is worthy of publication, and just needs some moderate revision to standardize the statistics and also expand captions and discussion to make explicit links and provide better explanations of the results being shown.

We appreciate the reviewer’s endorsement and hope the changes made will be satisfactory.

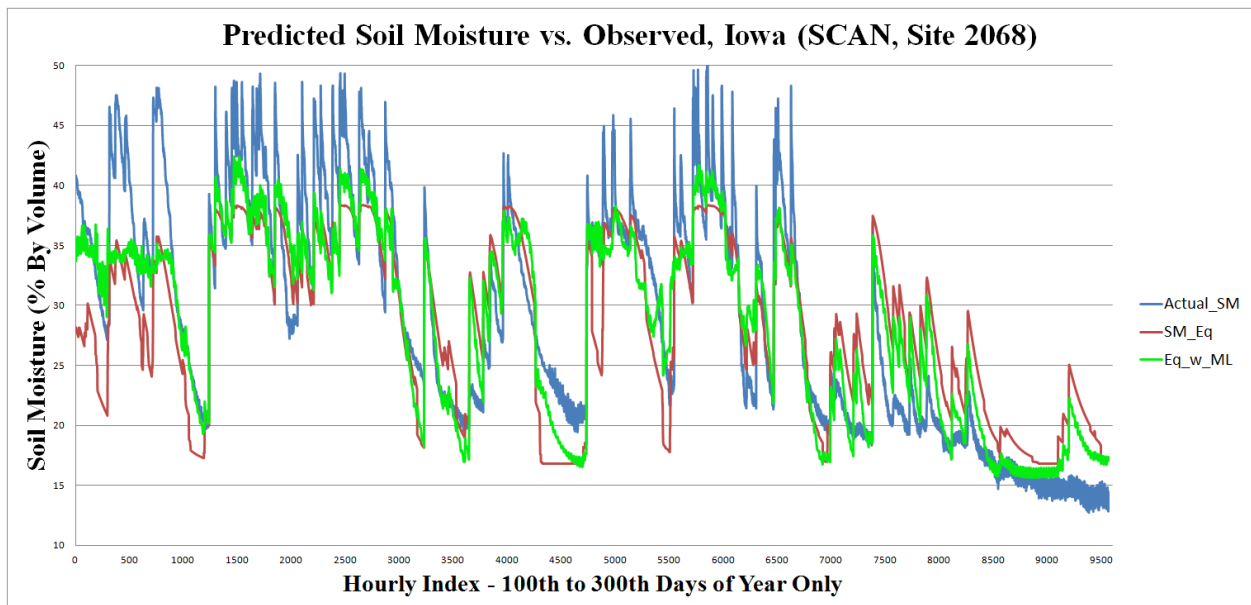
Specific Comments:

Figures 3-5 do not have legends, and appear to be cut off on the right side. The legend on Figure 6 is barely legible. The captions don’t need to be repeated. You could say “same as Figure 3 but for site xxxx, which is a <fill in type> hydroclimate and <fill intype> soil.”

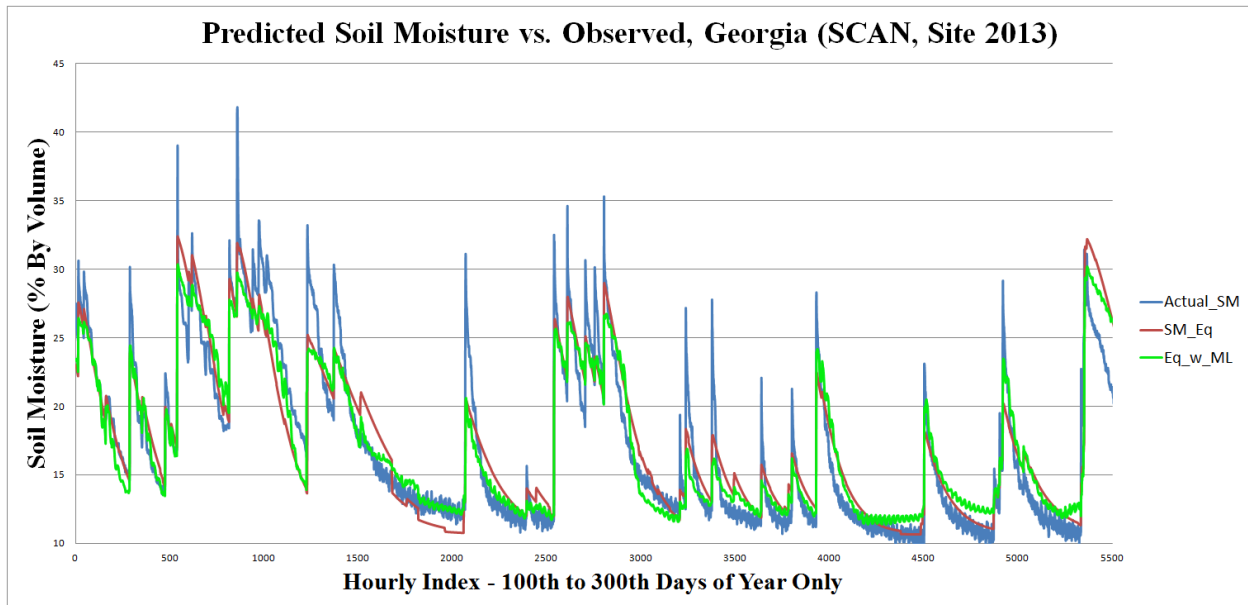
Regarding the right edge, this is an idiosyncrasy of the margin settings within the manuscript file sent to HESS – this has been corrected. Legends are now included and/or enlarged. The captions have been altered to provide the full description in the first of the three figures and refer back to Figure 3 thereafter. Hydroclimate information and soil textural characteristics are now listed in all cases.



**Figure 3, Soil Moisture Time Series, SCAN Site 2015, New Mexico (USA), Actual Soil Moisture (Blue Line), Diagnostic Soil Moisture Equation Estimate (Red Line), and Diagnostic Soil Moisture Equation with Machine Learning Error Correction (Green Line).
Hydroclimate: IAQ (Intermediate Seasonality, Arid, Summer Peak Runoff)
Soil Texture: Loamy Sand**



**Figure 4, SM Time Series, SCAN Site 2068, Iowa (USA), line colors from Fig. 3
Hydroclimate: ISCJ (Intermediate Seasonality, Semi-Arid, Winter Peak Runoff, Summer Peak Precipitation)
Soil Texture: Silty Clay Loam**



**Figure 5, SM Time Series, SCAN Site 2013, Georgia (USA), line colors from Fig. 3
 Hydroclimate: LWC (Low Seasonality, Winter Peak Precipitation, Winter Peak Runoff)
 Soil Texture: Sandy Loam**

Figure 9 and Figure 10 are redundant. Only Figure 10 is necessary.

Agreed, Figure 9 has been removed.

Figure 11 is a nice summary of the net effect, but I'm left wanting to see more results for the four cases illustrated in the diagram. I would think the type of errors due to soil and hydroclimate are different. Some more insightful discussion here is warranted.

This is an important question, but one best left for future work. Lines 515-519 of the revised manuscript discuss this question in further detail.

"It is likely that the types of errors made when parameters are cross-applied between sites of different hydroclimates will differ from the types of errors that appear when the sites differ edaphically. Further research into the specific conditions under which models err, along with the magnitude and bias of those errors, would be highly useful."

When one begins to analyze types of errors under various cross-applications, the computational complexity reaches a level that suggests a separate inquiry would be required. Essentially, there would be 210 (x,y) pairs to evaluate in detail – this would make for an excellent follow-up paper, but the computational complexity of understanding how ~40,000 examples per site calibrate roughly 10,000 examples in 210 cases with extensive analysis regarding how and why the error occur would likely fall beyond the scope of this paper.

Technical Issues:

Throughout the manuscript, the references to Pan et al., 2012, need to be corrected to Pan 2012, as there are no other coauthors.

Good catch. The relevant changes have been made.

Reviewer #2, Anonymous

General Comments:

This model attempts to represent soil moisture storage by modifying the model of Pan et al. The model has 6 parameters and includes two exponential functions, which should be sufficient to represent a wide variety of soil moisture response to precipitation, if hydrology is a model. Yet, the model does not consistently reproduce soil moisture patterns at a single depth with this limited information. It is not clear how this is an improvement in modeling soil moisture. The paper does give rise to several questions and concerns.

We thank the reviewer for his/her commentary, and appreciate the opportunity to discuss our findings. The performance obtained by our approach is comparable to the results published by Pan in 2012, despite the fact that we have attempted to generate an hourly estimate (Pan, 2012 predicts daily soil moisture values) and chosen numerous sites where wetting/drying behavior is more nuanced than the predominately drier locations chosen by Pan (2012). Additionally, the fundamental objective of this paper is not to build a better soil moisture model than any of the precursors mentioned in the literature review, but rather, to develop an approach to soil moisture prediction that requires only a precipitation time series for ungauged location.

Major concerns. The authors correctly state that soil heterogeneity poses a substantial challenge for soil moisture modeling. This restricts model application to the relatively homogenous soils. Yet, even in homogenous soils bulk density is commonly regarded to decrease with depth. In the model ϕ is a singular value for the soil profile, how is ϕ determined or chosen? Is this physical parameter subject to change from the genetic algorithm, if so, is it a physical parameter or a “degree of freedom” parameter? The authors state that the prediction is made for a specific soil depth, but none of the demonstration figures identify sensor depth, not do they compare performance for multiple depths at a single site. Such comprehensive analysis would be of interest to the reader and perhaps give the authors insight into the model performance, especially near the upper and lower boundaries of the soil.

The reviewer correctly notes that soil density can be a function of soil depth, yet the parameter, ϕ , is a constant by location. Given the general limitation of our datasets and the fact that shallow-depth soil moisture is most relevant to decision-support, all of our analyses occur with measurements of 2in (~5cm) depth. A note to this effect has been added following equation two, to avoid any subsequent confusion. ϕ , as discussed on p.2327, lines 22-26 of the original manuscript, is fit by genetic algorithm. With calibration sets consisting of tens of thousands of hours of data, with only six total parameters, the ‘degree of freedom’ parameter issue is minimal. Comparisons of performance by depth at each site is a relevant question, but one best left for future research.

The soil moisture conditions of greatest concern to agriculture are excess moisture, which limits soil strength and trafficability in the spring, and excessively wet or dry soils during various stages of crop development. The vertical fluxes to drainage and evaporation differ dramatically under these conditions and would seem to require greater control than a precipitation decay function coupled with a soil water flux resistance term. It will be beneficial to the reader for the authors to explain clearly how their model accomplishes a water balance through a growing

season without separate representations of percolation and evaporation. The causal dismissal of the need for farmers to know extent of saturation is disappointing. In this model, and in reality, the time until a farmer can resume field operations is largely dependent on the extent of saturation.

Though percolation and evaporation are not explicitly included, the Eta series, presented in equation three, describes a sinusoidal loss function over a year. This approximates the phenomena discussed in this comment. The diagnostic soil moisture equation, a parsimonious model by design, has been peer-reviewed and published in 2003 and 2012. It was chosen specifically because of its limited data requirements – no doubt there are other hydrologic and agricultural features that would be relevant to soil moisture predictions, but to include them in a model is to require their availability wherever the model is applied.

The KNN correction is intended to allow consistent model biases, but the results show that the model consistently over predicts or under predicts for some case studies. This is not a convincing demonstration.

The KNN approach does achieve results in smoothing systematic biases from the diurnal cycle (see Fig. 6). The following sets of images (created at the behest of reviewer #1) demonstrate the effect of bias correction for time-of-day, day-of-year, soil moisture conditions, or even antecedent precipitation. However, despite all of these adjustments, it is still possible that a validation year is notably wetter/drier than the training data, causing over or under-predictions. Moreover, sensor calibrations are imperfect and dynamic – it would be difficult for any model to avoid such errors. Even NASA’s SMAP mission (Soil Moisture Active Passive), during which soil moisture will be remotely sensed and validated with in situ sensors, targets root-mean-squared errors of 4%. Without the satellite images or other information, this approach (after bias correction), is only marginally worse; see the newly created Table 1 (also created at the behest of reviewer #1).

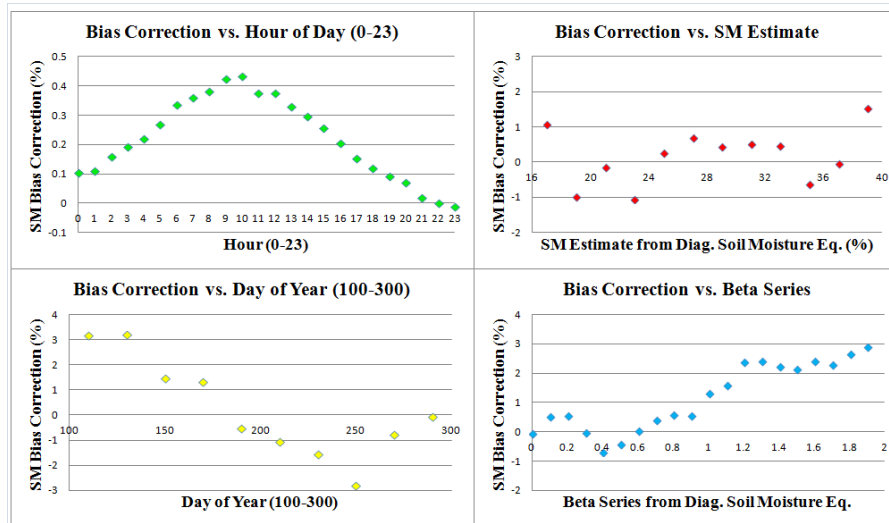


Figure 7, Bias Correction Analysis, SCAN Site 2015 (IAQ, Desert, Loamy Sand)

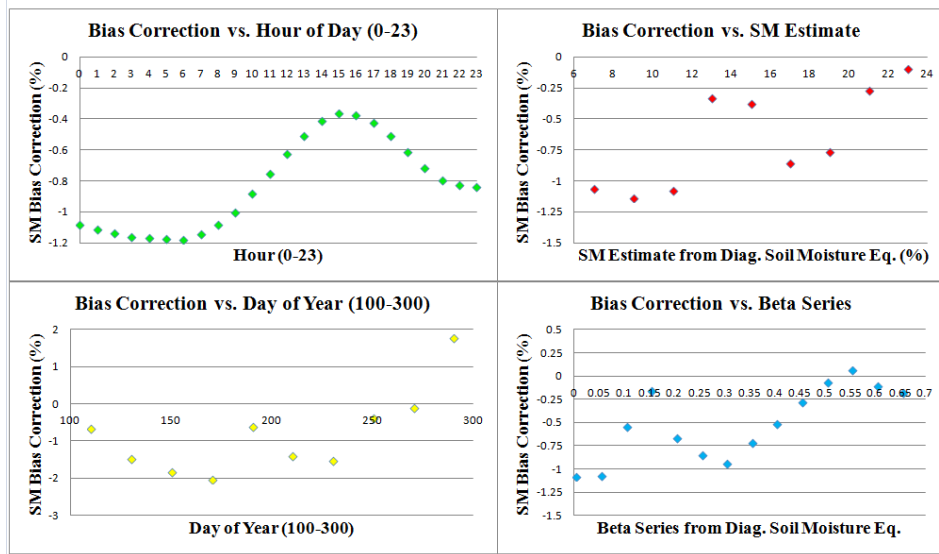


Figure 8, Bias Correction Analysis, SCAN Site 2068 (ISCJ, Plains, Silty Clay Loam)

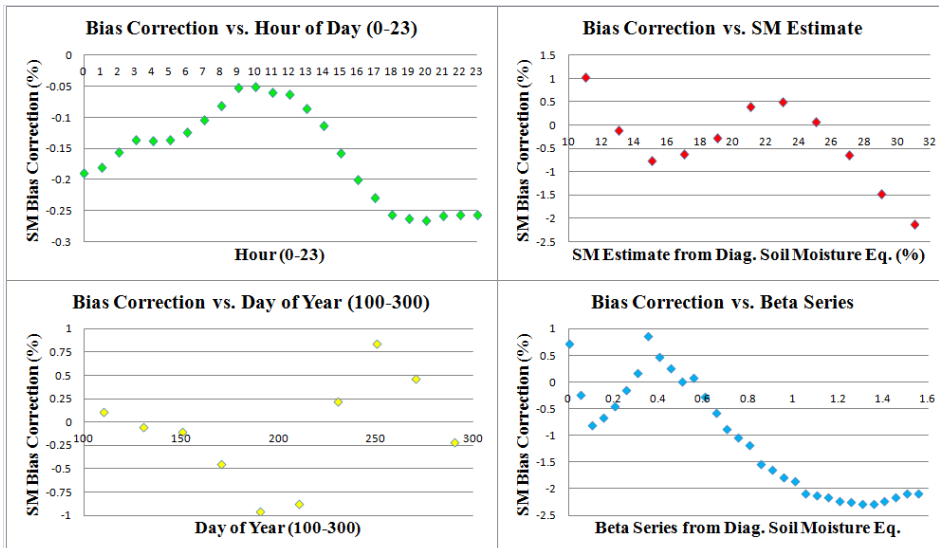


Figure 9, Bias Correction Analysis, SCAN Site 2013 (LWC, Woods, Sandy Loam)

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

Minor concerns The basis of adding a diurnal cycle to soil moisture is not well supported if prediction for agricultural management is the goal. The benefit of using LT (presumably local time) rather than simply stating the 24 h time is not clear.

We presume the reviewer is referring to p.2328, lines 15, 24, and 28 of the original manuscript. The “LT” can easily be removed. It is superfluous. We believe that including a diurnal cycle is relevant insofar as it may guide agricultural decision-makers to choose one hour rather than another to irrigate their fields or eschew traffic due to the wetter ground conditions.

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. Clearly, 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.”

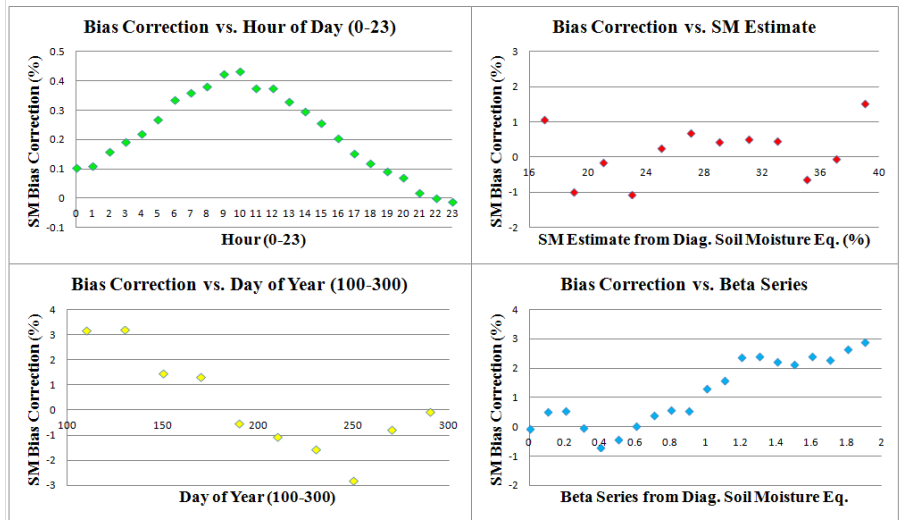


Figure 7, Bias Correction Analysis, SCAN Site 2015 (IAQ, Desert, Loamy Sand)

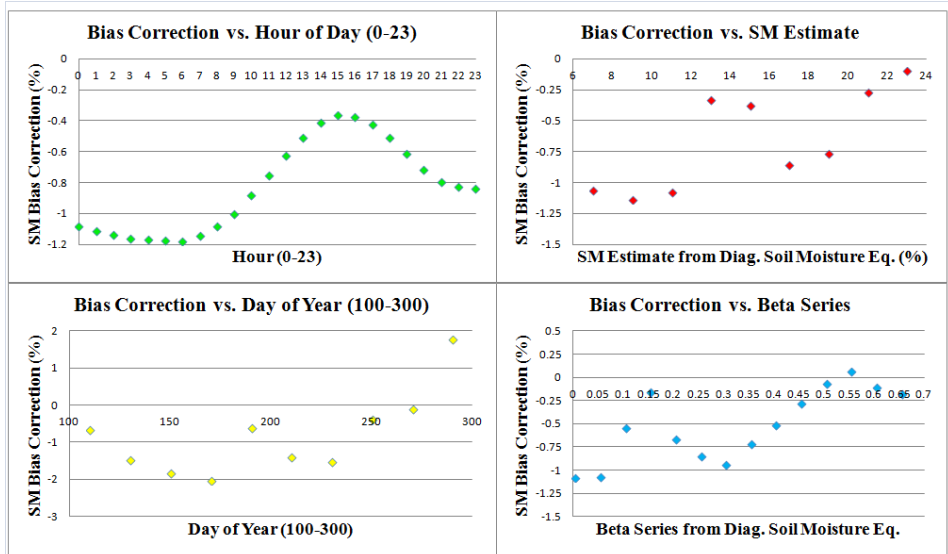


Figure 8, Bias Correction Analysis, SCAN Site 2068 (ISCJ, Plains, Silty Clay Loam)

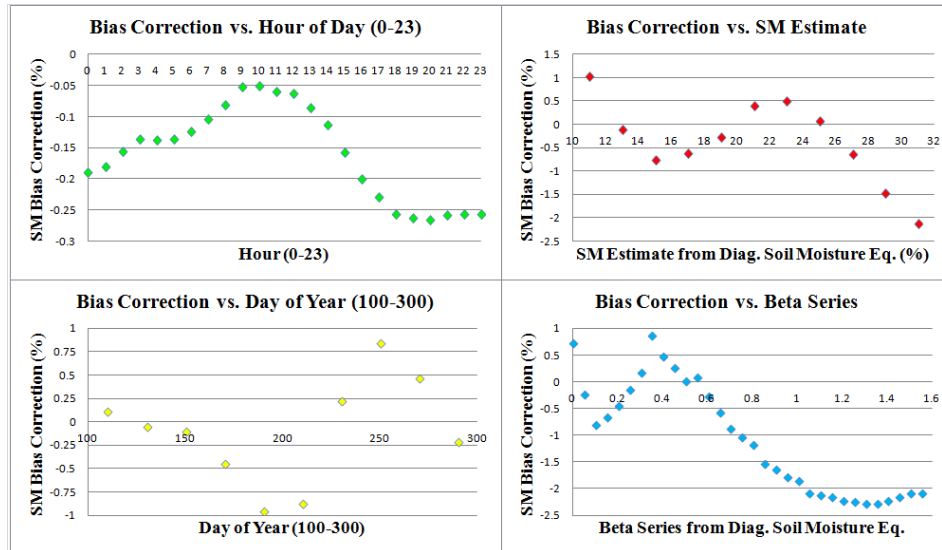


Figure 9, Bias Correction Analysis, SCAN Site 2013 (LWC, Woods, Sandy Loam)

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 Entekhabi 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 can 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.

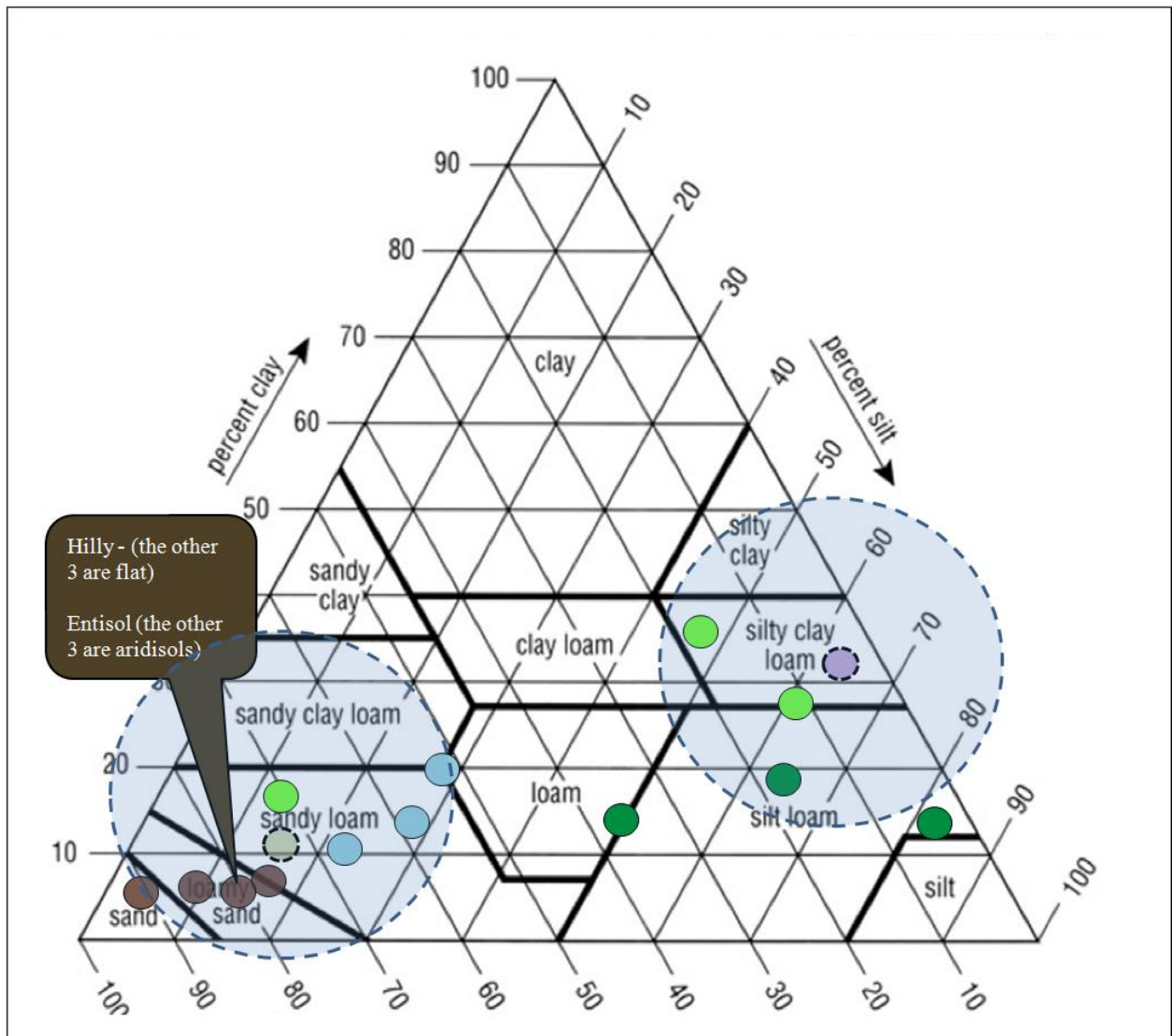


Fig. 11 should be presented as a Table.

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.

This paper investigates how a simple soil moisture model can be used at sites with no soil moisture measurements available for model training, but with similar climate and/or soil type. Given the sparsity of soil moisture measurements this is an important contribution as it allows to spatially generalize (soil moisture) model calibrations.

We thank the reviewer for his/her compliments, as this was precisely our intention with this work. We hope that our responses to the comments below improve the manuscript and satisfy the reviewer.

General comments:

The paper needs minor revisions. In my opinion the paper is very well written. It is straightforward to understand and clearly structured. There are other simple soil moisture models that do not require soil moisture information for calibration. (Koster and Mahanama 2012, JHM; Orth et al. 2013, JHM) A reference to such approaches (e.g. in the discussion section) could indicate another possible direction in which to apply the derived results. I appreciate the tables and information the authors provided in response to reviewer #1, I agree that these will improve the manuscript.

We thank the reviewer for drawing our attention to the two listed papers. Indeed, these works also provide estimates of soil moisture without antecedent soil moisture information. Following line 13 on p. 2338, the following will be added:

“The diagnostic soil moisture equation used in this paper (Pan et al, 2003; Pan, 2012) was an appropriate choice due to its ability to generate soil moisture estimates without the need for knowledge of antecedent soil moisture conditions. Koster and Mahanama (2012) and Orth et al. (2013) have developed approaches to estimate soil moisture at the watershed scale by leveraging hydroclimatic variability and long-term streamflow measurements in a water-balance model – also without employing previous soil moisture conditions. If the parameters calibrated and then generalized in this work produce point estimates of soil moisture at a diversity of locations, integration with a water balance approach could help with the up-scaling process.

Specific comments:

Title: Maybe you want to consider simplifying your title such that a broader audience can understand it. I would think of e.g. "Using climate and soil information to generalize soil moisture prediction"

Perhaps a different title would lead to wider appeal. Consider:

“Using similarity of soil texture and hydroclimate to enhance soil moisture estimation”

page 2323, line 7: remove "(precipitation"

Agreed. The change will be made.

page 2326: From equation 1, soil moisture content would never increase. I guess you add (possible) precipitation at each time step?

In equation 1, the β term represents the convolution of previous hours of precipitation. One can observe that if β is equal to zero (no precipitation during the relevant historical window), θ_{est} is

set to θ_{re} , the residual soil moisture of the soil. Should β grow large (saturating the soil), θ_{est} approaches the porosity, ϕ_e . At each time step, β changes, as its temporal window is fixed – at each time step, the oldest hour of precipitation data used to calculate β is replaced by the most recent hour. In this manner, precipitation in the most recent hour (weighted more heavily in the convolution) increases θ_{est} .

page 2326, lines 21-25: Why do you use different metrics (objective functions) that are minimized/maximized here?

Does the reviewer refer to lines 21-25 of p.2327? If this is the case, the reviewer's question regards the fact that we maximize correlation for the first three parameters fit by genetic algorithm, yet minimize the sum of squared errors for the three parameters fit during the second stage. The reason for this is that the β -series is characterized by a wholly different numerical scale than the soil moisture series were are ultimately attempting to estimate. Moreover, it is still (at that point in the analysis) missing the three soil-specific parameters. Thus, choosing optimal values for the first three parameters entails developing a β -series whose shape follows the shape of the measured soil moisture values. Once this 'shape' is modeled, choosing a lower-bound for θ_{est} (residual soil moisture, θ_{re}) and an upper-bound (porosity, ϕ_e), along with a rate of drainage (c_4), allows the generation of a soil moisture series that should have a minimal total sum of square errors with respect to the observed soil moisture series.

page 2333: Please mention that this error correction approach cannot deal with trends in the soil moisture data.

The following sentence will be added, following line 18 of p. 2333:

“This approach to error correction, as it relies on previous errors to predict future errors, will not address long-term trends within the soil moisture record.”

page 2333, line 13: add "when considering the entire time series" before "but without flooding events ..."

These sentences will amended at the behest of this and another reviewer. The paragraph that begins on line 7 of p. 2333 will now read:

“During the validation period, specifically 2010, wetter conditions were observed than were present during calibration. At this SCAN, before 2010, the average soil moisture value observed as 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...”

page 2333, lines 16-18: In terms of droughts this shortcoming has more serious consequences. Whereas it may not matter much if it is wet or very wet, it is important if it is dry (plants may survive) or very dry (plants may die), especially in the context of irrigation management.

Agreed. Small errors in terms matter more during dry conditions than during wet conditions. Generally, the model does make smaller errors, in absolute terms, during dry conditions. Following line 18 of p.2333, the following sentence will be added.

“It is important to note that small errors are more significant in terms of decision support (specifically when and where to irrigate) during dry conditions. Generally, the model’s errors are smaller, in absolute terms, during drier conditions.”

page 2336, lines 15-18: Would you say hydro-climate and soil type are about equally important or is it too little data to make such a statement here?

*This is an important question – the data seem to suggest that hydroclimatic characteristics are slightly more important than edaphic features (the model performs better when hydroclimates align but soil textures do not than the converse). However, what is not rigorously analyzed (though it could be, in a subsequent paper) is the extent of the hydroclimatic differences vs. the extent of the edaphic differences within those groups respectively. As a result, it would be speculative to say which is **more** important – one can only state that both are important. To this end, after line 18 of p. 2336, the following will be added:*

“...future modeling work, in which the relative importance of hydroclimates and soil textures can be examined in greater detail.”

page 2337: The model may not only benefit from accounting for overland flow but also for subsurface flow/runoff, especially in hilly areas.

Absolutely. Line 19 of p. 2337 will read:

“...by considering overland and subsurface flows, specifically in areas characterized by more complex topography.”

page 2338: What is the soil depth considered in the model? Satellite data represents only the upper centimeters of the soil and may therefore be of limited use to improve total column soil moisture model estimates.

Another reviewer (#2) has raised this question as well. The response to that reviewer is reproduced below:

“Given the general limitation of our datasets and the fact that shallow-depth soil moisture is most relevant to decision-support, all of our analyses occur with measurements of 2in (~5cm) depth. A note to this effect has been added following equation two, to avoid any subsequent confusion.”

Figure 2: Please label the x-axis. You can cut the range of the y-axis such that it starts at 0.3 or so.

The new Figure 2 appears below:

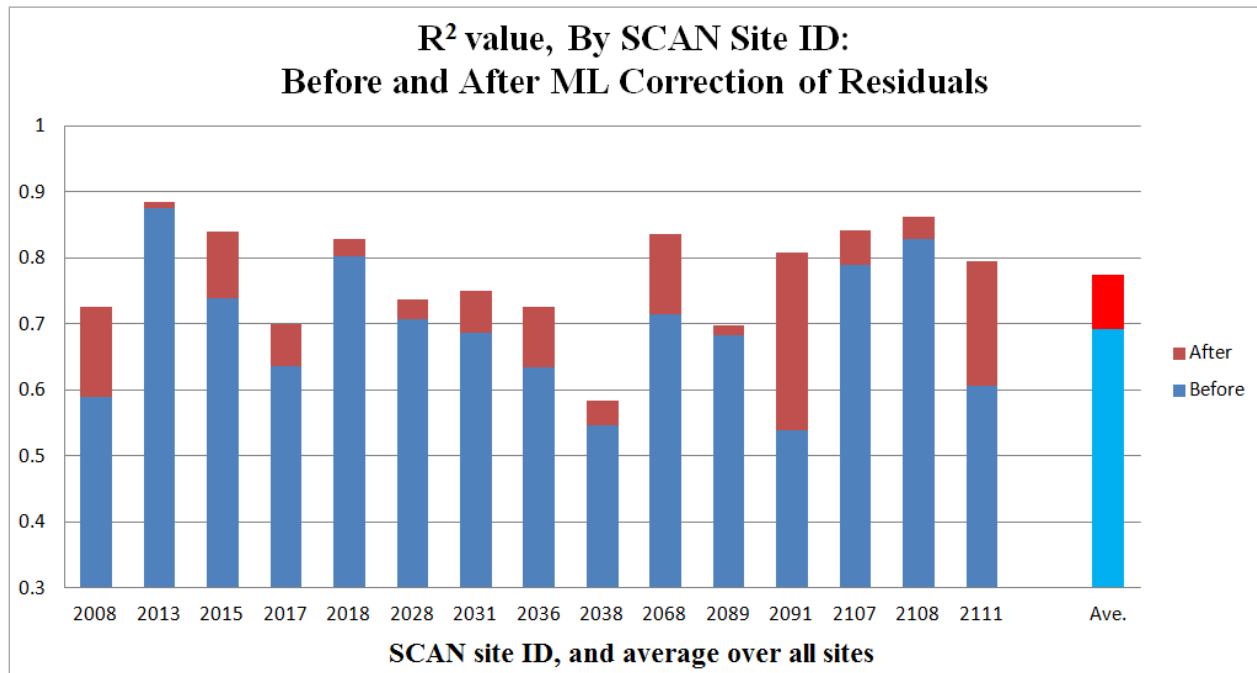
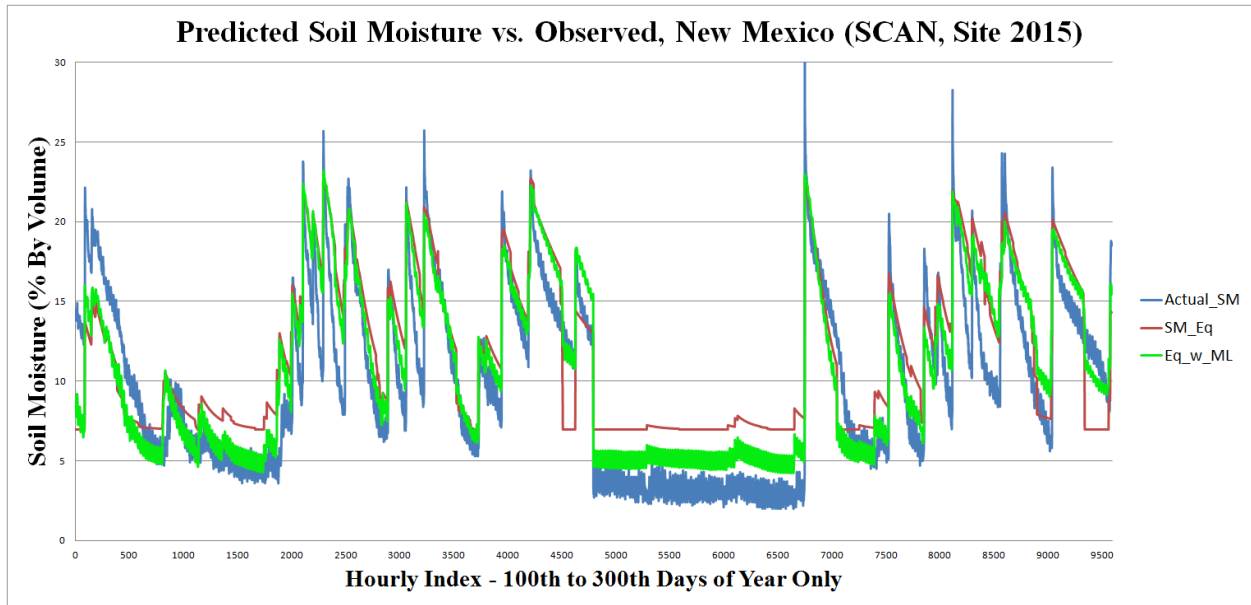


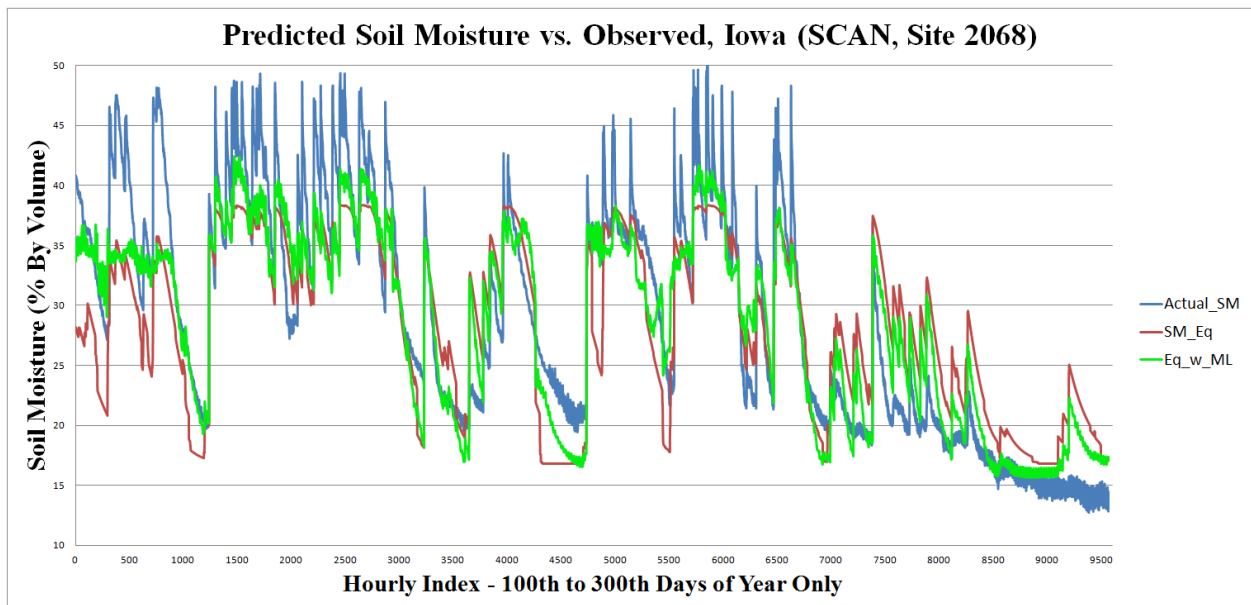
Figure 2, Improvements from machine learning (KNN) models of residuals.

Figures 3-6: Put exact dates/times on x-axis.

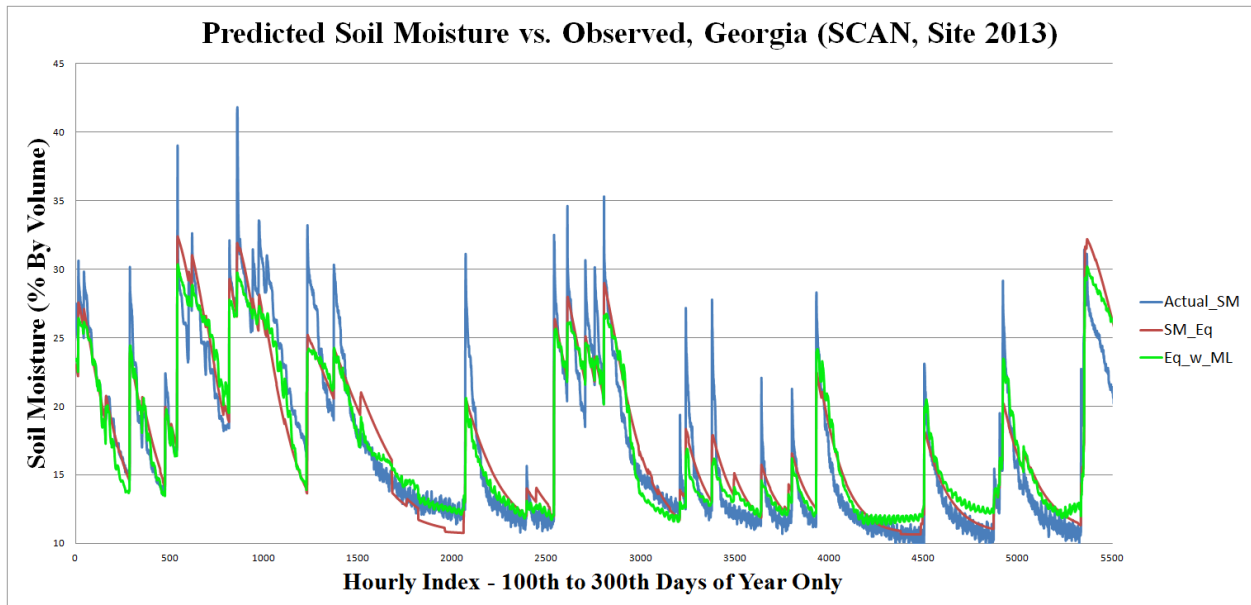
In figures 3-6, the only hourly time stamps that appear are those for which the date falls between the 100th and 300th days of the year (to ensure analysis of unfrozen ground) and for which the precipitation and soil moisture values from the relevant sensor are available. Thus, these are not wholly continuous time-series and consequently, it could confuse readers were dates to appear at inconsistent intervals. Figures 3-6 have been updated slightly at the behest of other reviewers, appearing below.



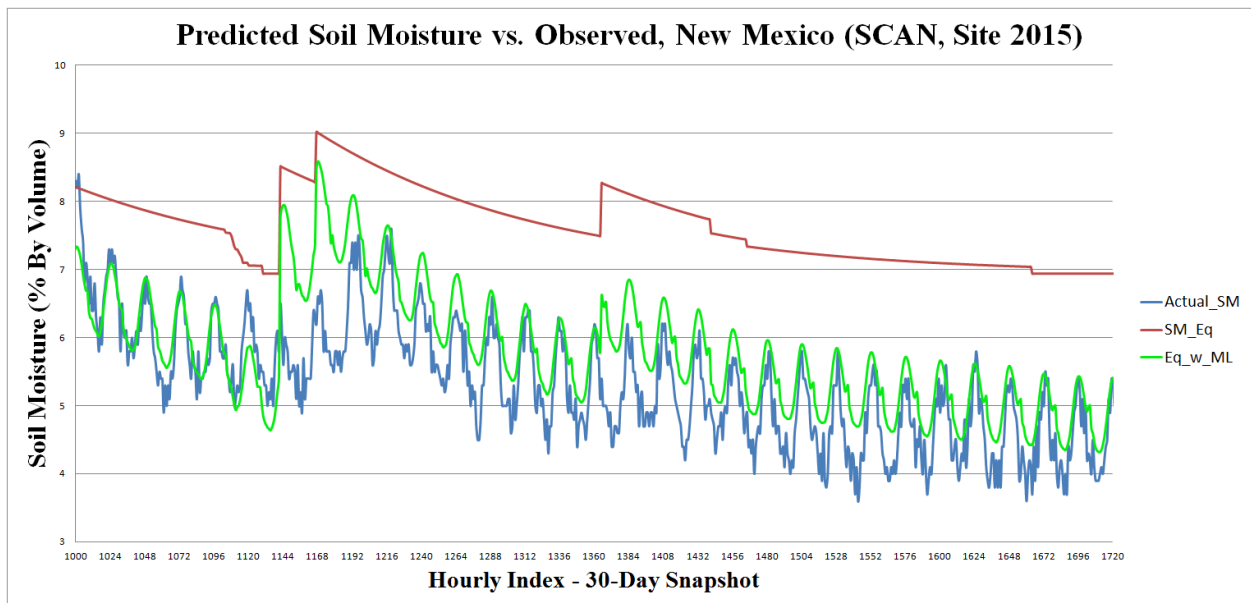
**Figure 3, Soil Moisture Time Series, SCAN Site 2015, New Mexico (USA), Actual Soil Moisture (Blue Line), Diagnostic Soil Moisture Equation Estimate (Red Line), and Diagnostic Soil Moisture Equation with Machine Learning Error Correction (Green Line).
Hydroclimate: IAQ (Intermediate Seasonality, Arid, Summer Peak Runoff)
Soil Texture: Loamy Sand**



**Figure 4, SM Time Series, SCAN Site 2068, Iowa (USA), line colors from Fig. 3
Hydroclimate: ISCJ (Intermediate Seasonality, Semi-Arid, Winter Peak Runoff, Summer Peak Precipitation)
Soil Texture: Silty Clay Loam**



**Figure 5, SM Time Series, SCAN Site 2013, Georgia (USA), line colors from Fig. 3
 Hydroclimate: LWC (Low Seasonality, Winter Peak Precipitation, Winter Peak Runoff)
 Soil Texture: Sandy Loam**



**Figure 6, Soil Moisture Time Series, SCAN Site 2015, New Mexico (USA), Actual Soil
 Moisture (Blue Line), Diagnostic Soil Moisture Equation Estimate (Red Line), and
 Diagnostic Soil Moisture Equation with Machine Learning Error Correction (Green Line)**

Figure 10: Some text is missing in the brown box.

The correction has been made, see below:

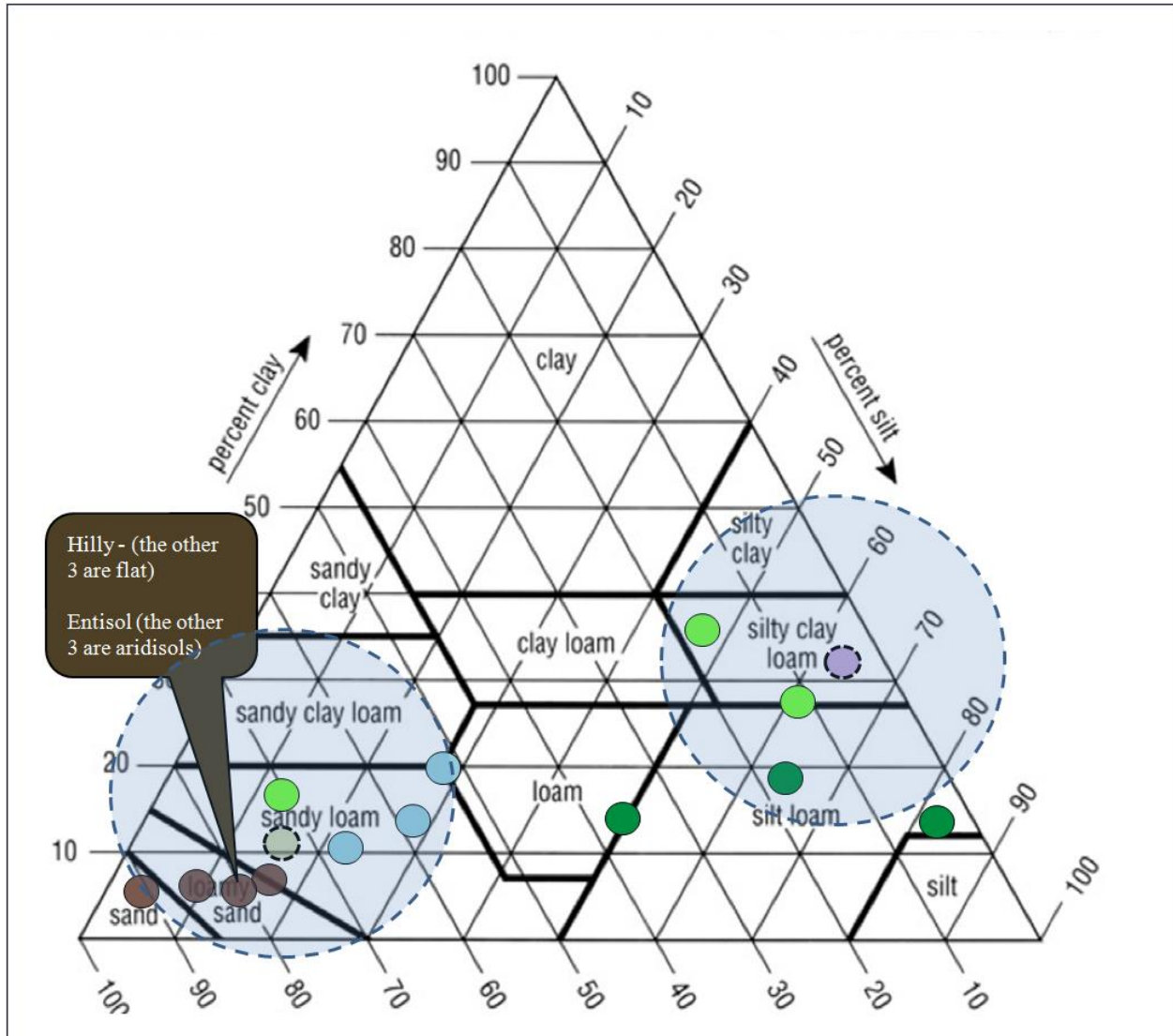


Figure 12, The 15 SCAN sensors, color-coded to match their hydro-climatic class, with similar soil textures shaded.

Figure 11: Has it been referred to in the text? Use different colors for soil texture circle and hydroclimate circle.

Figure 11 is referenced on line 12 of p. 2336. It is reproduced below (it will be Figure 13 in the revised manuscript). Green shades denote hydroclimatic similarity, brown shades denote similarity with respect to soil texture.

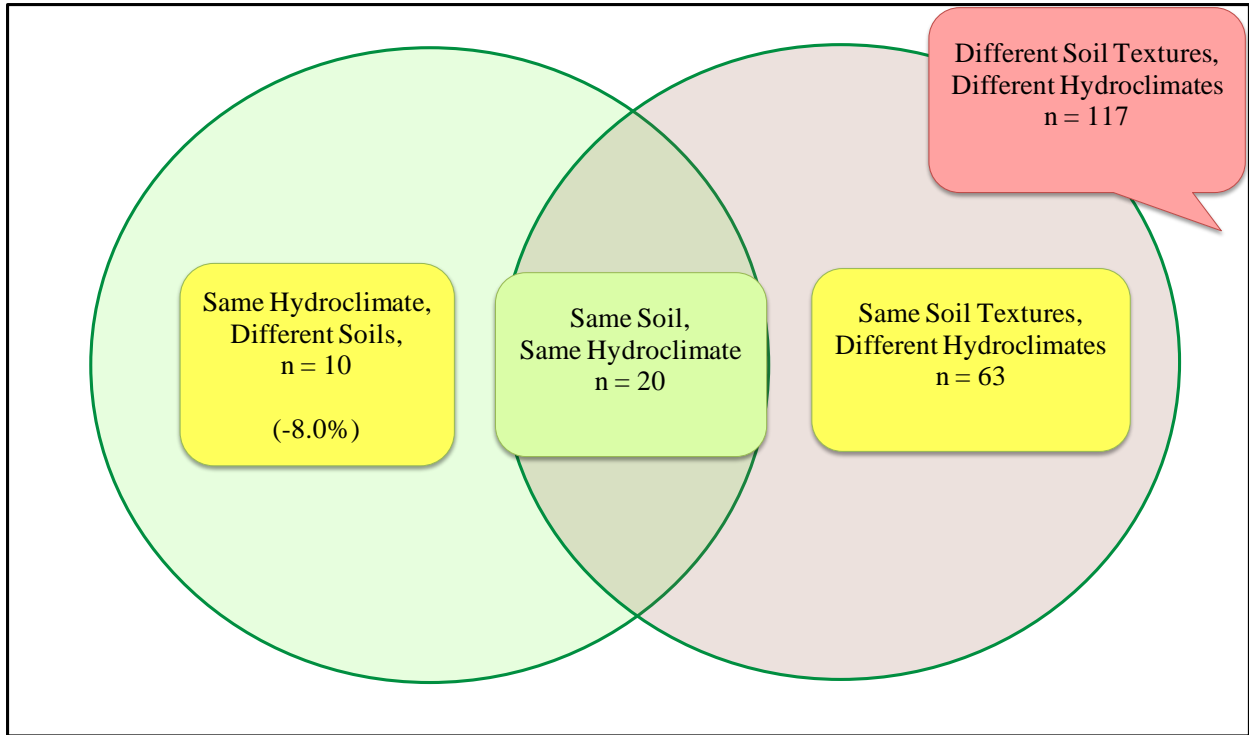


Figure 13 Venn-Diagram of Modeling Errors with Similar and Different Soils and Hydro-climates

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26

~~Using hydro-climatic and edaphic similarity to enhance soil moisture prediction~~ Using similarity of soil texture and hydro-climate to enhance soil moisture estimation

E. J. Coopersmith,¹ B. S. Minsker,¹ and M. Sivapalan^{1,2}

¹ *Department of Civil & Environmental Engineering, University of Illinois at Urbana-Champaign, Urbana, IL 61801, USA*

² *Department of Geography and Geographic Information Science, University of Illinois at Urbana-Champaign, Urbana, IL 61801, USA*

Formatted: Numbering: Continuous

27 **Abstract**

28 Estimating soil moisture typically involves calibrating models to sparse networks of *in situ*
29 sensors, which introduces considerable error in locations where sensors are not available. We
30 address this issue by calibrating parameters of a parsimonious soil moisture model, which
31 requires only antecedent precipitation information, at gauged locations and then extrapolating
32 these values to ungauged locations via a hydro-climatic classification system. Fifteen sites
33 within the soil climate analysis network (SCAN) containing multi-year time series data for
34 precipitation and soil moisture are used to calibrate the model. By calibrating at one of these
35 fifteen sites and validating at another, we observe that the best results are obtained where
36 calibration and validation occur within the same hydro-climatic class. Additionally, soil texture
37 data are tested for their importance in improving predictions between calibration and validation
38 sites. Results have the largest errors when calibration/validation pairs differ hydro-climatically
39 and edaphically, improve when one of these two characteristics are aligned, and are strongest
40 when the calibration and validation sites are hydro-climatically and edaphically similar. These
41 findings indicate considerable promise for improving soil moisture estimation in ungauged
42 locations by considering these similarities.

43 **Keywords:** soil moisture, hydro-climatic, edaphic, similarity, prediction, operational models

44

45

46

47

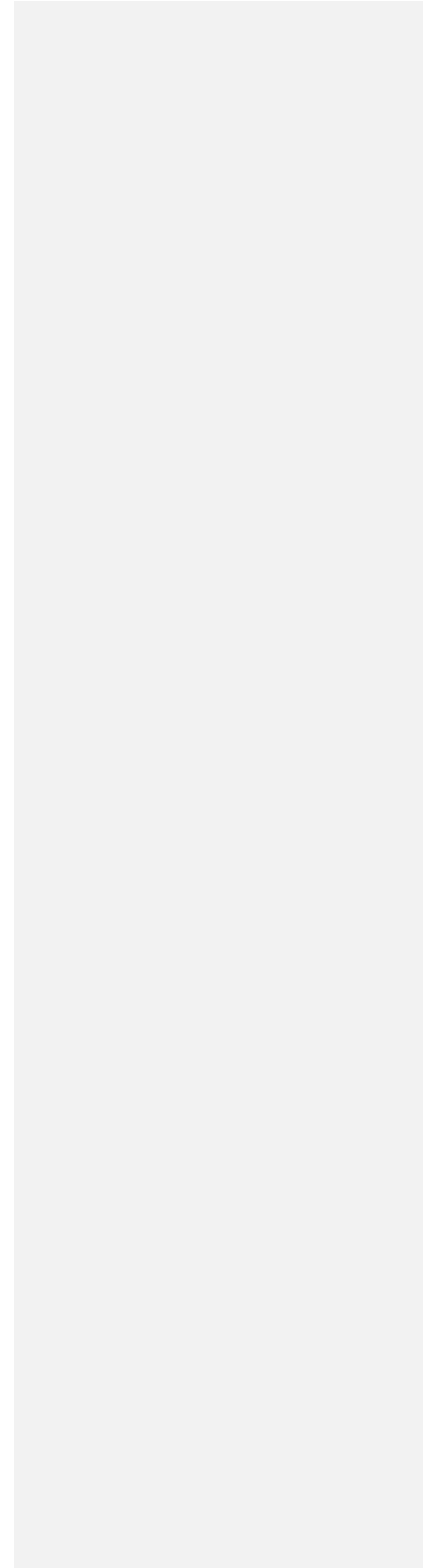
48

49

50

51

52



53 **1. Introduction**

54 Soil moisture estimates are needed routinely for many practical applications, such as irrigation
55 scheduling and operation of farm machinery. They are typically produced either through remote
56 sensing or sparse networks of *in situ* sensors. Although recent remote sensing studies have
57 confirmed that such measurements approximate *in situ* sensor networks (Jackson et al, 2012),
58 satellite-based sensors provide measurements at a spatial resolution of several kilometers – too
59 large for daily agricultural decision making. On the other hand, *in situ* sensor networks produce
60 values that are difficult to generalize to locations with no proximal sensors. Under these
61 circumstances, dynamic soil moisture evolution models are typically used for soil moisture
62 estimation at the desired location, using information from the nearest available sensors. This
63 method of soil moisture estimation immediately raises the issue regarding the type of model that
64 is most appropriate for such an application. One could think of several different types of models
65 that may be suitable.

66 _____
67 The first group of soil moisture models considers only the variability of precipitation, as it
68 has been shown that precipitation variability is the primary mechanism for wetting/drying
69 (Entekhabi and Rodriguez-Iturbe, 1994). Many subsequent models employed an “antecedent
70 precipitation index” (API), defining a pre-established temporal window for antecedent rainfall.
71 This index is then used to estimate current levels of soil moisture (Saxton and Lenz, 1967) and
72 has been implemented with recession modeling for soil water in agriculture (Choudhury and
73 Blanchard, 1983) and also in weather prediction (Wetzel and Chang, 1988). Other precipitation-
74 focused approaches utilize stochastic models to estimate the distributions of soil moisture values
75 using an initialization of daily rainfall (Farago, 1985). Both the stochastic and API approaches
76 require some initial condition for soil moisture at the forecast location – requiring either
77 professional judgment or a sensor. While these issues can be addressed using a soil water
78 balance model, this type of model must be recalibrated frequently, which most soil moisture
79 models are not, -as its errors are cumulative (Jones, 2004).

80 The second group of models adopts a more hydrologic process-based approach, estimating
81 soil moisture from surface radiation and precipitation (Capehart and Carlson, 1994). These
82 process-based models are typically forced by evapotranspiration demand and precipitation at

Formatted: Space After: 0 pt

83 | their upper boundary and, if applicable, by groundwater at their lower boundary. More
84 | sophisticated models of this type, such as HYDRUS (Simunek et al, 1998), ~~require attempt to~~
85 | improve predictions via detailed knowledge of hydraulic soil parameters, information regarding
86 | root structures, soil temperature readings, ionic chemistry, CO₂ concentrations, solute transport
87 | data, and detailed atmospheric/meteorological information, which are not widely available,
88 | especially for routine applications envisaged here.

89 | The third group of models ~~is-are~~ agriculturally-focused, building model projections
90 | outward from existing instrumentation and additional measurements. Gamache et al (2009)
91 | developed a soil drying model for which cone penetrometers and soil moisture sensors are
92 | required. At most remote sites, these data sources are not currently accessible. Another similar
93 | approach employs specific soil type information (theoretically, publicly available data), but
94 | ultimately requires proximal sensors to provide the needed soil moisture estimates (Chico-
95 | Santamaria, et al, 2009).

96 | Pan et al (2003,-) and Pan 2012) addressed many of the shortcomings of the existing
97 | modeling approaches reviewed above by developing what they called a “diagnostic soil moisture
98 | equation” (i.e., model) in the form of a partial differential equation representing the lumped
99 | water balance of a vertical soil column, and representing the soil moisture at any moment in time
100 | as a function of the sum of a temporally decaying sequence of observed past rainfall events. The
101 | model has the advantage that initial soil moisture conditions are not required (only antecedent
102 | precipitation data), nor must the model be recalibrated periodically. However, this approach
103 | does require a soil moisture sensor at the relevant location for initial calibration of the model’s
104 | parameters. This method has the disadvantage that the presence of soil heterogeneity could
105 | necessitate a large number of sensors to account for the spatial variation of soil moisture (Pan
106 | and Peters-Lidard, 2008). Furthermore, decision support often requires estimation at locations
107 | lacking sensors.

108 | The aim of this paper is to present and test an approach that can help overcome the
109 | ~~problems-issues of calibration at ungauged locations~~ associated with the Pan et al. soil moisture
110 | estimation model. The proposed solution involves calibrating the Pan ~~et al.(2012)~~ diagnostic soil
111 | moisture equation (model) at gauged sites and then extrapolating the calibrated model to
112 | ungauged sites by invoking similarity. Similarity here is defined on the basis of hydro-climatic

113 characteristics, using a classification system developed by Coopersmith et al (2012), as well as
114 edaphic (soil) properties. The proposed new scheme maintains the advantage of Pan et al.'s
115 parsimonious soil moisture model in that it does not require specification of initial soil moisture
116 condition, and also there is no need to recalibrate periodically. The model's simplicity also
117 permits implementation of the model in a manner that can easily be refit with new parameters,
118 where necessary. Section 2 provides more details on the approach.

119 To calibrate and validate the model, we use data from the U.S. Department of
120 Agriculture's (USDA) Soil Climate Analysis Network (SCAN). This national array of soil
121 moisture sensors (with co-located precipitation sensors) delivers hourly data at a variety of
122 publically-accessible sites throughout the United States. Fifteen sensor locations with numerous
123 years of high-quality, minimally-interrupted data were selected for further analysis. These sites
124 display considerable hydrologic diversity, which aids in demonstrating that the nationwide
125 application of the proposed soil moisture model using precipitation data represents a feasible
126 goal. With respect to agricultural decision-support, for energy-limited sites, the value of hourly
127 soil moisture estimates is found in the determination of whether or not a field is trafficable –
128 whether heavy equipment will damage fields or become mired. With respect to water-limited
129 sites, the value of soil moisture estimates is found in devising optimal irrigation strategies that
130 utilize limited water resources most efficiently. Of the fifteen SCAN sites examined, the three
131 sites in New Mexico, the site in Colorado, the site in Nebraska, the site in Wyoming, and the two
132 in Iowa are all water-limited (8 in total). The remaining sites (7 in total), located in Pennsylvania
133 (2), Arkansas, Georgia, South Carolina, North Carolina, and Virginia, are all energy-limited.
134 Results of the analysis are given in Section 3, followed by discussion in Section 4 to suggest
135 further improvements and conclusions are presented in Section 5.

136 2. Methodology

137 The proposed modeling approach involves four steps, summarized in Figure 1 and described in
138 more detail in the sections below. First, the diagnostic soil moisture model of Pan et al.(2012) is
139 calibrated at locations with ample data. Given that the focus of this study is on soil moisture
140 estimation for agriculture, we only consider prediction during the growing season, which is
141 appropriate given that the model does not address snow melt processes. Second, the predictions
142 at these locations are improved using machine learning techniques for error correction. Third,

143 the classification system proposed by Coopersmith et al. (2012) is used to generalize the
144 parameters calibrated at each location, enabling its application at other sites characterized by the
145 same hydro-climatic class. Fourth, sites are examined for edaphic (soil property) similarity in
146 addition to hydro-climates. The results of these four steps are then examined to identify which
147 approach to regionalization performs best.

148

149 **Step 1: Calibration Using a Two-Layer Genetic Algorithm**

150 Unlike the original diagnostic soil moisture calibrations, the ultimate objective of this work is to
151 enable agricultural decision support in near real time. To this end, the daily model from Pan ~~et~~
152 ~~al.~~(2012) is first modified to yield an hourly model within the same framework. Genetic
153 algorithms are then deployed to calibrate the model, enabling more efficient exploration of the
154 parameter search space than the traditional Monte Carlo search, which was the approach taken
155 by Pan ~~et al.~~(2012).

156 Genetic algorithms (GAs), a subset of evolutionary algorithms, were originally developed
157 by Barricelli (1963) and have become increasingly common in environmental and water
158 resources applications, including the calibration of hydrologic model parameters (e.g., Cheng et
159 al, 2006; Singh et al, 2008; Zhang et al, 2009).

160 In this work, a simple genetic algorithm uses the operations of selection, crossover, and
161 mutation (for reference, see Goldberg 1989) to search for parameters that minimize prediction
162 errors from the diagnostic soil moisture equation (Pan ~~et al.~~, 2012):

163

$$164 \quad \theta_{est} = \theta_{re} + (\phi_e - \theta_{re})(1 - e^{-c_4\beta}) \quad (1)$$

165

166 Here θ_{est} represents the best estimate of soil moisture during a given hour. θ_{re} denotes residual
167 soil moisture, the minimum quantity of moisture that is present regardless of the length of time
168 without precipitation. ϕ_e , the soil's porosity, signifies the maximum possible soil moisture value,
169 at which point the soil becomes saturated ~~and cannot increase its moisture content~~. Finally, c_4 is

170 a parameter related to conductivity and drainage properties, essentially defining the rate at which
 171 soil can dry. If c_4 assumes a value of zero, the soil is permanently at its residual soil moisture
 172 value, θ_{re} - a soil that dries infinitely rapidly. Conversely, as c_4 becomes large, the soil will
 173 permanently assume the value of its porosity, ϕ_e - a soil that dries infinitely slowly. The β term
 174 in Equation 1 is calculated in Equation 2 below:

175

$$176 \quad \beta = \sum_{i=2}^{i=n-1} \left[\frac{P_i}{\eta_i} \left(1 - e^{-\frac{\eta_i}{z}} \right) e^{-\sum_{j=1}^{j=i-1} \left(\frac{\eta_j}{z} \right)} \right] + \frac{P_1}{\eta_1} \left(1 - e^{-\frac{\eta_1}{z}} \right) \quad (2)$$

177

178 Here, P_i denotes the quantity of rainfall during hour i (day in the original presentation in Pan et
 179 al.). The soil depth at which an estimation occurs is given by z . This convolution summation
 180 has a temporal window of size- n hours for considering past precipitation. ~~To wit, yesterday's~~
 181 ~~rainfall affects today's soil moisture, last week's rainfall is relevant, but less so, and rainfall from~~
 182 ~~ten years ago is probably not relevant~~ For instance, today's soil moisture is strongly influenced by
 183 yesterday's rainfall, influenced to a lesser degree by last week's rainfall, and not influenced at all
 184 by rainfall from ten years previous. Given the general limitation of our datasets and the fact that
 185 shallow-depth soil moisture is most relevant to decision-support, all of our analyses occur with
 186 measurements of two inch (~5cm) depth.

187 To choose the appropriate value for n , the value of β is calculated at each hour throughout
 188 the dataset - setting n to a very large value (2000 hours, denoted by M) initially. Next this "beta
 189 series" (where $n = M$) is correlated with a separate beta series, calculated where $n \ll M$. If the
 190 correlation coefficient between these two time series approaches unity, then the smaller value of
 191 n is selected. Otherwise, n is increased incrementally until the correlation between the $n \ll M$
 192 beta series and the $n = M$ beta series approaches unity.

193 Finally, ~~the η_i terms signify~~ the estimated ~~potential evaporation / drainage~~ soil water loss at
 194 hour i ~~of the calendar year, e.g. due to evapotranspiration or deep drainage, is expressed by the~~
 195 term, η_i . As this algorithm does not presume any more detailed knowledge of potential
 196 evaporation/drainage behaviors, this "eta series," representing losses due to evapotranspiration

197 | and deep drainage. is modeled as a sinusoid (Pan-et-al., 2012) with period 8,760 (the number of
198 hours in a year). The eta (η) series is required to calculate the beta (β) series (Eq. 2), which is
199 required to use the diagnostic soil moisture equation (Eq. 1). Thus, before any other parameters
200 are chosen, a generalized sinusoidal form of η is estimated as given in Equation 3:

201

$$202 \quad \eta = \alpha \sin(i - \delta) + \gamma \quad (3)$$

203

204 Here, α represents the sinusoid's amplitude, γ denotes the vertical shift, and δ signifies the
205 necessary phase shift. These three parameters are fitted via the genetic algorithm such that the
206 correlation between the beta series (using the eta series implied by α , γ , and δ) and the observed
207 soil moisture series (θ_{obs}) is maximized. Once values for the eta series are established, the
208 remaining three parameters of Equation 1 (θ_{re} , ϕ_e , and c_4) are then fitted by a second application
209 of the genetic algorithm, this time minimizing the sum of squared errors between the estimated
210 soil moisture series (θ_{est}) and the observed values (θ_{obs}).

211

212

213 **Step 2: Error Correction Using The k-Nearest Neighbors Machine Learning Algorithm**

214 After the parameters of the diagnostic soil moisture equation (Eq. 1) have been calibrated, the
215 hourly precipitation time series is used to generate a soil moisture time series during the growing
216 season months of interest. Discrepancies between the observed soil moisture values (θ_{obs}) and
217 the estimated values (θ_{est}) are computed as shown in Equation 4:

218

$$219 \quad \theta_{obs} = \theta_{est} + \varepsilon \quad (4)$$

220

221 where ε represents the error associated with any hour's soil moisture estimate.

222 To correct biases in these errors, the k-Nearest Neighbor algorithm (Fix and Hodges,
223 1951) is employed to predict ε using the characteristics from the training data. More
224 specifically, the data are searched for the most similar matches in terms of time of day, day of
225 year, θ_{est} , $\beta(n)$, and $\beta(M) - \beta(n)$. For example, if the model returns a prediction of $\theta_{est} =$
226 0.35 at 2:00pm during July when rainfall has been heavy recently but drier over a longer period,
227 KNN will search the training set for other estimates near 0.35 made on mid-summer afternoons
228 where a similar recent rainfall pattern has been observed. Next, the algorithm averages the value
229 of the error, ε , associated with those types of conditions, producing an estimated error, ε_{est} .
230 Each validation estimate is then adjusted to be $\theta_{est} + \varepsilon_{est}$. This technique allows consistent
231 model biases, such as underestimating wetter days and overestimating drier days, to be corrected.

232 This error correction model also accounts for diurnal soil moisture variations that were not
233 considered in developing the diagnostic soil equation, which was designed to deliver daily soil
234 moisture estimates. Consider a soil moisture estimate at 4pm, after soil has had a full day of
235 sunlight (theoretically) to dry. As the diagnostic soil moisture equation only considers drainage
236 and evapotranspiration losses on a daily basis, θ_{est} will be larger than θ_{obs} . Yet, because this
237 type of mistake presumably occurred frequently throughout the training data, the algorithm will
238 locate other 4pm estimates, each of which will be biased in the same direction, and our final soil
239 moisture estimates will take this bias into account, improving the results as shown subsequently.

240 To assess the performance of the soil moisture models with and without machine
241 learning, an R^2 value as defined in Eq. 5) is used, as this value represents the proportion of
242 variance in soil moisture explained by the developed model.

$$243 \quad R^2 = 1 - \frac{SSR}{SST} \quad (5)$$

244 where SSR denotes the sum of squared residuals and the SST term signifies the total sum of
245 squares, i.e. the sample's variance.

246

247 **Step 3: Estimation by Hydro-climatic Similarity**

248 This step tests the hypothesis that the classification system by Coopersmith et al. (2012) can be
249 used to generalize the calibrated parameters for the diagnostic soil moisture equation using
250 hydro-climatic similarity. If two locations are assigned the same hydro-climatic classification,
251 then the calibrated parameters from one SCAN sensor within that class will be assumed to
252 perform well at another.

253 This hypothesis was tested at fifteen SCAN sensors for which soil moisture and
254 precipitation data are available hourly for a period of several years. These sensors are located in
255 diverse geographic locations and hydro-climatic classes in Iowa, North Carolina, Pennsylvania,
256 New Mexico, Arkansas, Georgia, Virginia, South Carolina, Nebraska, Colorado, and Wyoming.
257 The data at each of these locations were divided into training/validation sets and parameters were
258 calibrated using training data only. Next, these parameters were employed on the validation sets
259 at the locations for which they were calibrated. The subsequent R^2 values (proportion of
260 variance in soil moisture explained by the machine-learning-enhanced diagnostic soil moisture
261 equation, see Steel and Torrie, 1960, for reference) defined a baseline level of performance for
262 that site.

263 The process of cross-validation is detailed below:

- 264 1. Consider two sites, x and y , chosen from the fifteen available calibrated locations.
- 265 2. Estimate the soil moisture values in the validation dataset of site y , using the parameters
266 calibrated from the training dataset at site x .
- 267 3. Record the difference between the R^2 baseline value at site y (obtained using parameters
268 calibrated at site y) and the performance obtained at site y using parameters calibrated at
269 site x .
- 270 4. Repeat steps 1-3 for all 210 possible (x, y) pairs where $x \neq y$
271 Note: (x, y) and (y, x) are not equivalent. One signifies the performance of parameters
272 calibrated at site x making predictions at site y , the other signifies the performance of
273 parameters calibrated at site y making predictions at site x .

274
275 At this point, three types of (x, y) pairs emerge. If the hypothesis is correct, then the first
276 type, when x and y fall within the same hydro-climatic class, should display limited losses in

277 predictive power. The second type, when x and y fall within a “similar” hydro-climatic class
278 (two classes differing by a single division of the classification tree developed in Coopersmith et
279 al., 2012) should display greater losses of predictive power. Finally, the third type, when x and
280 y fall in two unrelated classes, should display the largest loss of predictive power.

281

282 **Step 4: Estimation by Hydro-climatic and Edaphic Similarity**

283 The final step extends the hypothesis proposed in Step 3 by evaluating the impacts of soil texture
284 and type on soil moisture predictive power. The fifteen sites from the SCAN network are
285 examined based upon the soil textural information available from the Pedon soil reports that
286 SCAN provides, as well as data from NRCS’s soil survey database¹.

287 This information allows sites already deemed hydro-climatically similar to be further sub-
288 divided into sites that are and are not edaphically similar. Analogous to the previous section, we
289 consider pairs of sites, x and y , where parameters are calibrated at site x and validated at site y .
290 In this case, four groups can be defined – the first, where x and y are hydroclimatically similar,
291 the second, where x and y are hydroclimatically similar, but differ edaphically, the third, where
292 x and y are edaphically similar, but differ hydroclimatically, and finally, where x and y are
293 hydro-climatically and edaphically dissimilar.

294

295 **3. Results**

296 This section begins by presenting the results of the machine learning approach used in error
297 correction during the initial calibration step (Section 3.1). Next, Section 3.2 presents results for
298 the hydro-climatic similarity analysis, illustrating the performance of calibration/validation pairs
299 within the same class and without. Finally, Section 3.3 shows how the predictive power
300 improves when both hydro-climatic and edaphic similarity are considered.

301

¹ <http://websoilsurvey.sc.egov.usda.gov/App/WebSoilSurvey.aspx>

302 **3.1 Testing the Value of Machine Learning Error Correction for Soil Moisture Prediction**
303 **Using the Diagnostic Soil Moisture Equation**

304 Figure 2 shows the performance of the calibrated parameters for the 15 SCAN sites using only
305 the diagnostic soil moisture equation (Step 1 of the methodology) along with the subsequent
306 improvement in performance following machine learning error correction (Step 2). In each case,
307 the six parameters required for the implementation of the diagnostic soil moisture equation are
308 calibrated using training data from before 2010. Sensors with hourly precipitation and soil
309 moisture time series data between 2004 and 2009 (inclusive) provide four to six years of training
310 data (some sites are missing one or two years of data). Only days of the year where snow cover
311 is unlikely are used to train the algorithm (from the 100th to 300th day of the year in all
312 locations, for consistency). Validation data consist of days 100-300 for 2010 and 2011.

313 The results illustrate that in all fifteen test cases, performance within the validation
314 sample is improved by machine learning modeling of residuals from the training set, in some
315 cases, as much as 26.9% of the unexplained variance (site 2091) in soil moisture is corrected
316 from by this technique. The average results (far right column, Fig. 2) illustrate that the
317 diagnostic soil moisture equation explains just 69.2% of the variance in soil moisture ($\rho = 0.83$)
318 before machine learning corrections occur, but explains 77.5% of the variance in soil moisture (ρ
319 = 0.88) thereafter.

320 To explore these findings in more detail, three of the 15 SCAN sites, chosen to represent
321 different hydro-climatic locations – New Mexico (#2015, hydroclimate IAQ/southwestern
322 desert, Loamy Sand), Iowa (#2068, hydroclimate ISCI/northern midwest plains, Silty Clay
323 Loam), and Georgia (#2013, hydroclimate LWC/southeastern forest, Sandy Loam) are examined
324 to illustrate how improvements from adding machine learning error models to the diagnostic soil
325 moisture equation differ across sites. These three sites represent three distinct hydro-climatic
326 classes, with significant differences in soil texture, seasonality of precipitation, aridity, timing of
327 maximum precipitation, and timing of maximum runoff. Using error correction models for
328 prediction at these sites increased R2-values by an average of 8.2%, which is similar to the 8.3%
329 improvement in R2 averaged across all fifteen sites. Thus, these three locations are
330 representative in terms of both hydro-climatic and edaphic diversity and their responsiveness to
331 machine learning.

332 The base soil moisture model results from applying Step 1 at the three sites are displayed
333 in Figures 3-5. These predictions are shown with the results produced by deploying the machine
334 learning algorithm (KNN) in Step 2 to remove bias and correct errors. In each image, the blue
335 line represents the observed soil moisture readings, the red line represents the estimates
336 generated by the diagnostic soil moisture equation, and the green line represents those
337 predictions after the machine learning algorithm has removed biases and corrected errors. Soil
338 moisture values (y-axis) are ~~measured in percentage terms presented as volumetric percentage~~
339 (0-100).

340 In Figure 3, the diagnostic soil moisture equation is able to trace the general trend of the
341 soil moisture time series ($\rho = 0.860$). However, during the middle of the time series, in which
342 the observed soil moisture values fall below 5%, the benefits of machine learning error
343 correction are most noteworthy. There are other hours scattered throughout the dataset where the
344 green line (ML prediction) follows the blue line (observed values) much more closely than the
345 red line (diagnostic soil moisture equation). The green line ($\rho = 0.917$) not only improves upon
346 the correlation value of Pearson's Rho (the square root of the R^2 value in Eq. 5), but also displays
347 marked improvement for those cases in which the diagnostic soil moisture equation produces
348 significant errors.

349 ~~During the validation period, specifically 2010, wetter conditions were observed than~~
350 ~~were present during calibration. At this SCAN site, before 2010, the average soil moisture value~~
351 ~~observed was 28.55%, with only 25% of values exceeding 35% volumetric soil moisture.~~
352 ~~However, in 2010, the average soil moisture value measured was 33.16% with 45% of values~~
353 ~~exceeding 35%. The machine learning driven error correction improves the diagnostic soil~~
354 ~~moisture equation ($\rho = 0.846$) significantly ($\rho = 0.915$), but fails to raise its forecasts to reach~~
355 ~~some of the wetter conditions experienced in validation. Underestimations of this nature,~~
356 ~~although detrimental, during the validation period, considerable flooding occurred in Iowa~~
357 ~~(Figure 4). Flood events of this nature were not experienced during calibration, and the porosity~~
358 ~~parameter, ϕ_e (Eq. 1) was set at 38.4%. While this was appropriate for the training data (during~~
359 ~~which soil moisture did not exceed this level), extreme flooding events caused moisture levels~~
360 ~~for which the diagnostic soil moisture equation was not properly calibrated. The machine~~
361 ~~learning driven error correction improves the diagnostic soil moisture equation ($\rho = 0.846$)~~

362 significantly ($\rho = 0.915$), but without flooding events in the training set is unable to correct for
363 the errors due to exceedingly wet (flooded) soil. Underestimations due to floods, although
364 detrimental in terms of numerical errors, are not necessarily a problem for decision support of
365 agricultural or construction activities, for example. If a model warns that a site is very wet and in
366 reality, it is even wetter than predicted, the user has still been given adequate warning not to
367 attempt activity at that site. It is important to note that small errors are more significant in terms
368 of decision support (specifically when and where to irrigate) during dry conditions. Generally,
369 the model's errors are smaller, in absolute terms, during drier conditions. This analysis's
370 approach to error correction, as it relies on previous errors to predict future errors, will not
371 address long-term trends within the soil moisture record.

372 In Figure 5, a soil moisture series from Georgia is modeled by the diagnostic soil
373 moisture equation. Even before adding any error correction, the equation performs well ($\rho =$
374 0.936) and the machine learning approach yields a smaller improvement ($\rho = 0.941$). It is worth
375 noting that machine learning does not damage an already excellent performance, offering slight
376 improvements when possible and essentially no correction when training data suggest the model
377 has already performed adequately.

378 Table 1 presents all fifteen sites for which the diagnostic soil moisture equation has been
379 calibrated, including information regarding their hydroclimatic class from Coopersmith et al
380 (2012), their soil textural characteristics, and their performance before and after the KNN bias
381 correction process.

382 **3.2 Bias Correction – More Detailed Results**

383 In addition to generalizing the parameters calibrated in the diagnostic soil moisture
384 equation, the error correction approach allows for systematic biases to be removed by searching
385 training data for similar conditions and then predicting the types of mistakes most likely to occur.
386 Figure 6, by zooming in upon a 30-day period from Figure 2, illustrates how machine learning
387 reduces errors by introducing a diurnal cycle into a model that previously lacked one. The
388 remaining bias is likely explained by a slightly wetter training dataset as compared with the
389 validation data. It is possible that the diurnal cycle at some locations reflects a soil moisture
390 probe's dependency on electromagnetic properties driven by temperature change (apparent

Formatted: List Paragraph, Indent: First line: 0.5"

Formatted: Font: Bold

391 permittivity) rather than hydrologic processes (Rosenbaum et al, 2011). However, the model's
392 ability to respond to these nuances would not compromise its performance were these nuances
393 subsequently removed.

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

404 By addressing such systematic biases, machine learning enables model performance to
405 improve with each successive growing season as the training dataset expands. For instance,
406 although the fields in Iowa endured flooding during the validation period and subsequently made
407 errors, such errors would eventually populate the training data. The next time such flooding
408 occurs, the model is likely to recognize the occurrence of those same conditions and adjust the
409 diagnostic soil moisture equation's predictions accordingly. In this vein, model performance is
410 likely to improve over time, especially with the models already showing reasonable accuracy
411 using only a few years of training data.

412 Figure 7, 8, and 9 present these results in more detail for each of the three SCAN sites
413 presented in Figures 3, 4, and 5. In each figure, the upper-left image presents the average bias
414 correction (change in % soil moisture) for each hour of the day (0-23). At all three sites, bias
415 corrections display a clear diurnal pattern – that is to say the removal of a diurnal cycle is a
416 substantial role of machine learning under a variety of hydroclimatic and edaphic conditions.
417 The upper-right image of each figure presents the bias correction as a function of the unadjusted
418 soil moisture estimate – essentially, whether there exists a systemic over- or underestimation
419 when values are high or low.

420 The first two sites (Figures 7 and 8) do not present a clear pattern, but Figure 9 displays a
421 trend suggesting that the highest estimates of soil moisture tend to be overestimates and the

Formatted: Indent: Left: 0"

Formatted: Font: (Default) Times New Roman, 12 pt, Font color: Text 1

422 lowest estimates of soil moisture tend to be underestimates – but these biases are removed via
423 machine learning. The lower-left image presents bias correct as a function of the day of the year
424 (from 100-300, the days of the year when the model is applied). At all three sites, the seasonal
425 cycle does appear in terms of the patterns of bias correction, but the pattern is noisier than the
426 diurnal cycle. The magnitude of the adjustments are largest in the monsoon-affected desert of
427 New Mexico, a bit smaller in the Midwestern plains characterized by less extreme seasonal
428 behavior, and smallest in the Southeast where seasonal variations are low.

Formatted: Font: (Default) Times New Roman, 12 pt, Font color: Text 1

429 Finally, the lower-right image relates bias correction to the beta series from the diagnostic
430 soil moisture equation (Pan, 2012), a convolution of a decaying precipitation time series working
431 backwards temporally from the current time. Stated differently, these charts relate bias
432 correction to the amount of antecedent precipitation (with more recent precipitation weighted
433 more heavily). In Figure 7 (Plains, Silty Clay Loam), the model tends to underestimate moisture
434 when large quantities of antecedent rainfall are present, where in Figure 9 (Woods, Sandy
435 Loam), once antecedent precipitation becomes non-trivial, displays the opposite pattern. This is
436 consistent with the finer Midwestern soils' proclivity for ponding/flooding due to larger
437 proportions of clay. In these cases, larger amounts of rain will soak soils from above, and
438 capillary rise might further soak sensors from below, leading to underestimation from the
439 diagnostic soil moisture equation and subsequent machine learning correction. By contrast, with
440 sandier soils, drainage occurs easily, leading to higher rates of loss than the eta series (Pan, 2012)
441 would predict (there is more available water to lose), leading to overestimation with large
442 amounts of antecedent rainfall.

Formatted: Font: (Default) Times New Roman, 12 pt, Font color: Text 1

Formatted: Normal, Indent: First line: 0"

445 **3.2.3 Cross-Validation Results for Hydro-climatic Similarity: Qualitative Findings and** 446 **Significance Testing**

447 To test the hypothesis that models calibrated in one location can be used in a hydro-climatically
448 similar location, cross-validation was used as described in Step 3 of Section 2. The fifteen SCAN
449 sites yield $15^2 = 225$ possible (x, y) pairs. Fifteen of these 225 pairs occur when $x = y$,
450 establishing the baseline level of performance for a given site (validation performed using the

451 parameters calibrated at that same location). Of the 210 remaining (x, y) pairs, 120 of them
452 consist of paired catchments in which x and y are located in unrelated classes, 60 consist of
453 paired catchments in which x and y are located in a “similar” class (different by a single split
454 within the classification tree), and 30 consist of paired catchments in which x and y fall within
455 the same hydroclimatic class (but x and y do not represent the same catchment). Figure 7-10
456 presents box plots illustrating the ~~results change in R^2 values of for~~ these three sets of pairs in a
457 manner analogous to the differences shown in Figure 2, and Table 2 presents the quantitative
458 results, again averaging the deterioration of performance in terms of change in R^2 .

459 These findings show that calibrating the model at one location and applying those
460 parameters elsewhere within the same class (green) is preferable to applying those parameters in
461 a similar, but not identical class (yellow) and vastly superior to applying those parameters in an
462 unrelated class (red). The differences between any two clusters (same-class, similar-class,
463 unrelated class) are all significant at the $\alpha = 0.01$ level ($p < .001$ in all cases) as calculated by a
464 two-sample, heteroscedastic t-test (Welch, 1947).

465

466 **3.3-4 Impact of Soils: Cross-validation Results for Edaphic and Hydro-climatic Similarity**

467 To isolate the impacts of soil types (edaphic similarity) on soil moisture prediction,
468 groups of sensor locations among the 15 SCAN sites that are hydro-climatically similar were
469 analyzed, shown in Figure 8-11. The soil textural data for each of these fifteen sensors are plotted
470 on a soil texture pyramid diagram in Figure 9-12. These data were obtained from either Pedon
471 Soil Reports available through the SCAN network (which provide precise percentages of clay,
472 silt, and sand), or, where this information was unavailable, from soil information in the national
473 soil Web database².

474 Of the thirteen sensors from the four hydro-climatic classes with multiple SCAN sensors
475 (light green, blue, dark green, and brown in Figures 8-11 and 9-12), 30 (x, y) pairs exist where the
476 model can be calibrated at site x and its parameters applied at site y . Note that (x, y) is not
477 equivalent to (y, x) as the sites for calibration and validation are reversed. Of these 30 pairs, 20

² <http://websoilsurvey.nrcs.usda.gov/app/WebSoilSurvey.aspx>

478 pairs are edaphically similar as well. However, 10 of them include a pair of points where the soil
479 | types or terrain types are notably misaligned (for example, light green dots in Figure [40-12](#) where
480 | two of the three sensors are in silty clay loam and the third is in sandy loam– a notably different
481 | soil). A similar analysis to the one presented in Figure [7-10](#) and Table 2 has been reproduced,
482 | comparing the loss in predictive power (R^2) for the 20 pairs with similar hydro-climates and soils
483 | against the loss for the 10 pairs in which either the soil texture (Figure [912](#)) or type do not align.
484 | The average loss of 1.0% for the 20 very similar pairs is a much smaller decline than the 8.0%
485 | average decline observed for the 10 pairs for which soil/terrain information suggests
486 | dissimilarity. These results are significant with a p-value of approximately 0.02. Additionally,
487 | the upper-most two green dots in Figure [107](#), where calibrated parameters at one location
488 | perform poorly at another of similar hydro-climatic class, fall within these 10 cases.

489 | These observations show the importance of soil information, or edaphic similarity. While
490 | pairs of calibration/validation locations with similar hydro-climates, but dissimilar soils, show a
491 | decline in performance as compared with pairs of locations where both are similar, so too do
492 | locations with similar soils, but dissimilar hydro-climates. The shaded circles in Figure [40-12](#)
493 | illustrate groups of sensors that are quite similar in terms of soil textures. However, despite their
494 | soil similarities, differences in hydro-climates hinder cross-application, showing a decline in
495 | performance of 10.9% for all (x, y) pairs within the shaded regions of Figure [40-12](#) for which x
496 | and y are not from the same hydro-climatic class.

497 | As summarized in Figure [4413](#), these results suggest that in cases where both soil type
498 | and hydro-climate align, very little performance is lost when parameters are re-applied (1.0%),
499 | moderate declines in performance are observed when one of these two factors are aligned (8.0%
500 | if hydro-climates align and soil types do not; 10.9% if soil types align, but hydro-climates do
501 | not), and large declines in performance appear when neither align (20.5%). Clearly both types of
502 | attributes are important and should be considered in future modeling work [in which the relative](#)
503 | [importance of hydroclimates and soil textures can be examined in greater detail.](#)-

504

505 | **4. Discussion: Future Work to Improve Predictions**

506 This section discusses other approaches that could be used in the future to improve and broaden
507 the applicability of the methods developed in this work. First, we will consider micro-
508 topographic effects on soil moisture, as local peaks and valleys can cause soils to dry more or
509 less rapidly. Second, we will discuss a conceptual omission within the diagnostic soil moisture
510 equation – infiltration excess. Finally, we will discuss the role of future satellite data on soil
511 moisture modeling.

512 **4.1 Estimates Enhanced By Topographic Classification**

513 Ultimately, the combination of a hydro-climatic classification system and the diagnostic soil
514 moisture equation demonstrates a generalization of calibrations, facilitating predictions at any
515 location where a viable sensor exists within a similar hydro-climatic class and soil type.
516 However, the lumped, bucket model is not ideally-suited for landscapes with complex
517 topography. Conveniently, the majority of SCAN sites are placed on relatively flat surfaces.
518 Integration of topographic insights is a fertile area for future research. One possible approach to
519 further improving predictive accuracy is to disaggregate the soil moisture estimates as a function
520 of local topography. While SCAN sites used for soil moisture data are generally located on flat
521 surfaces, predictions may be needed at locations located on ridges or in valleys where the soils
522 are likely to be wetter or drier than their surroundings. This requires the notion of regional
523 topological classification. In this manner, the notion of similarity is extended to include hydro-
524 climatology, soil characteristics, and topographic designation (ridge, slope, valley, etc).
525 Preliminary analyses suggest that small-scale topography does play a meaningful role in the
526 wetting/drying process. Future research with more extensive datasets in locations with more
527 complex topological contours could improve soil moisture predictions by enabling the models
528 developed in this work to be adjusted as a function of local topographic classification.

529

530 **4.2 An Enhanced Diagnostic Soil Moisture Equation**

531 The diagnostic soil moisture equation could also be improved in future modeling efforts by
532 considering overland and subsurface flows, specifically in areas characterized by more complex
533 topography. Currently, the model assumes that, in the absence of saturation, all rainfall will
534 ultimately infiltrate, as the porosity parameter serves as an upper bound on soil moisture levels.

535 The diagnostic soil moisture equation was designed originally as a daily model, and it is
536 probably rare that on any given day, a significant fraction of precipitation does not infiltrate.
537 However, at the hourly scale it is quite possible that the water from an intense rainfall event will
538 not make its way into the soil at the location of the sensor. To address this lateral transfer
539 phenomenon, additional parameters can be introduced into the diagnostic soil moisture equation
540 that place an upper bound on the quantity of rainfall that can be infiltrated during any hour (or
541 other interval) of the convolution calculation for any particular soil type. Agricultural decision-
542 support includes trafficability when wet and irrigation support when dry. While overland flow is
543 perhaps an unneeded component in water-limited catchments where irrigation schemes represent
544 the most significant soil-moisture-related decision, in wetter catchments, in which trafficability is
545 a real concern, such an addition could improve the model. While this approach would require
546 the fitting of additional parameters, it is likely that predictions would be improved. These
547 additional parameters could also be considered in assessing cross-site edaphic similarity using
548 the methods described above, although they may be highly correlated with existing parameters
549 such as porosity, residual soil moisture, and drainage.

550

551 **4.3 NASA's Soil Moisture Active Passive (SMAP) Mission**

552 With NASA satellite data for soil moisture available at the 36 km, 9 km, and 3 km scales
553 throughout the United States, and with the SMAP satellite scheduled to launch during 2014
554 (O'Neill et al, 2011), the models developed in this work will have ample measurements against
555 which to test and improve their results, and can be used to help check the accuracy of satellite
556 measurements. Future research in LiDAR-driven disaggregation, proposed above, could also be
557 used to improve satellite soil moisture estimates by accounting for smaller-scale topography.

558

559 **4.4 Water Balance Models and Up-Scaling**

560 The diagnostic soil moisture equation used in this paper (Pan et al, 2003; Pan, 2012) was an
561 appropriate choice due to its ability to generate soil moisture estimates without the need for
562 knowledge of antecedent soil moisture conditions. Koster and Mahanama (2012) and Orth et al.

Formatted: Font: Bold

563 (2013) have developed approaches to estimate soil moisture at the watershed scale by leveraging
564 hydroclimatic variability and long-term streamflow measurements in a water-balance model –
565 also without employing previous soil moisture conditions. If the parameters calibrated and then
566 generalized in this work produce point estimates of soil moisture at a diversity of locations,
567 integration with a water balance approach could help with the up-scaling process.

568

569 **5. Conclusions**

570 This work has demonstrated the feasibility of estimating soil moisture at locations where soil
571 moisture sensors are unavailable for calibration, provided they fall within hydro-climatically and
572 edaphically similar areas to gauged locations. By calibrating the diagnostic soil moisture
573 equation via a two-part genetic algorithm, improving its performance via a machine learning
574 algorithm for error correction, then validating that algorithm at the same location in subsequent
575 years, a baseline level of predictive performance is established at fifteen locations. Next, these
576 results are cross-validated – deploying parameters calibrated at a given site at sites of similar and
577 different hydro-climatic classes, demonstrating that parameters can be re-applied elsewhere
578 within the same class, but not without. Finally, by incorporating edaphic information, we
579 observe the strongest cross-validation results when hydro-climatic and edaphic characteristics
580 align. As only 24 hydro-climatic classes describe the entire nation (and only 6 describe a
581 significant majority), it is entirely possible that a couple dozen well-placed soil moisture sensors
582 can enable reasonably accurate soil moisture modeling at any location within the continental
583 United States.

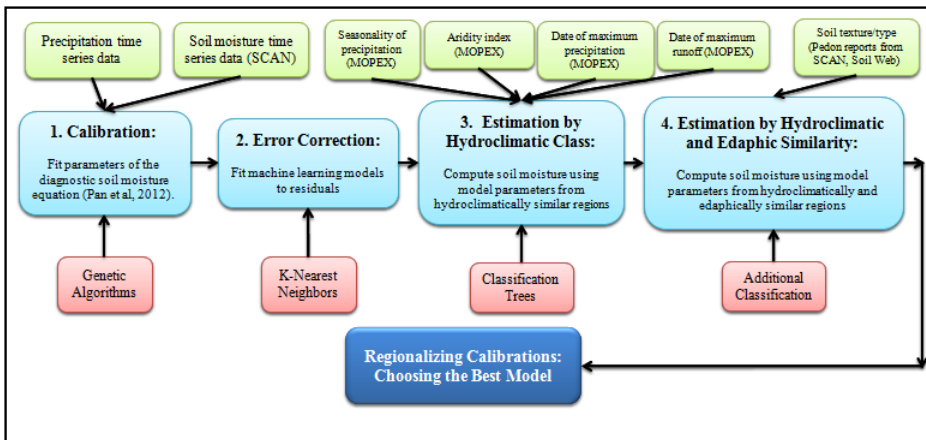
584 It is likely that the types of errors made when parameters are cross-applied between sites
585 of different hydroclimates will differ from the types of errors that appear when the sites differ
586 edaphically. Further research extending beyond model performance into the specific conditions
587 under which models perform less effectively along with the magnitude and bias of those errors
588 would be highly illustrative for future researchers.

589 This analysis can improve agricultural decision-support by offering insight into locations
590 that can benefit from targeted irrigation in drier conditions, or conversely, by minimizing risks of
591 ruts and damaged equipment when fields are no longer trafficable during wetter conditions.

592 Scaling the results of these models upward can assist with larger-scale assessments of flood risks
 593 or as calibration/validation tools for satellite estimates of soil moisture. Scaling these results
 594 downward can help maximize yields. Given the ubiquity of precipitation data, which are the
 595 only inputs these models require, better understanding of the transferability of modeled
 596 parameters is a step towards far wider availability of soil moisture estimates.

597 Leveraging these findings, the discussion section also presented the results of preliminary
 598 analysis that illustrates how further improvements in soil moisture predictions could be gained
 599 by disaggregating based on local topography. This would enable more accurate predictions at
 600 sites characterized by peaks and valleys that dry faster or slower than the relatively flat locations
 601 at which soil moisture algorithms are generally calibrated. Incorporating overland flow into the
 602 diagnostic soil moisture equation and integrating satellite data into the approach could also
 603 improve predictions in the future.

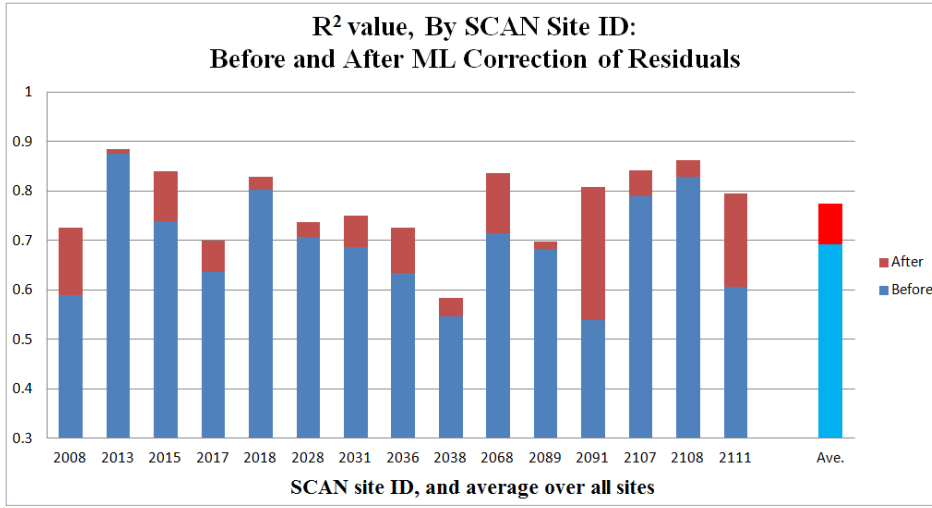
604
605
606



607
608
609

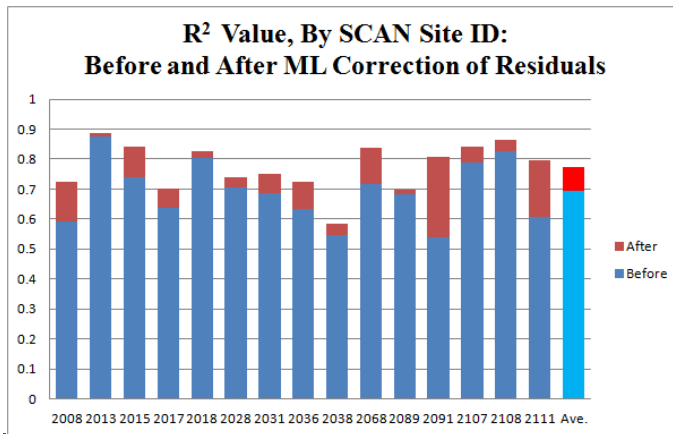
Figure 1, Methodological flow chart

Formatted: Font: (Default) Times New Roman, 12 pt, Italic



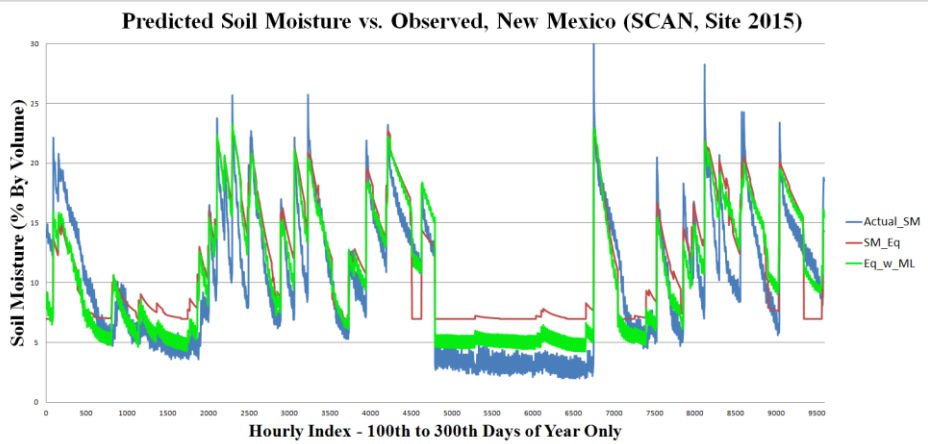
610
611

Figure 2, Improvements from machine learning (KNN) models of residuals.

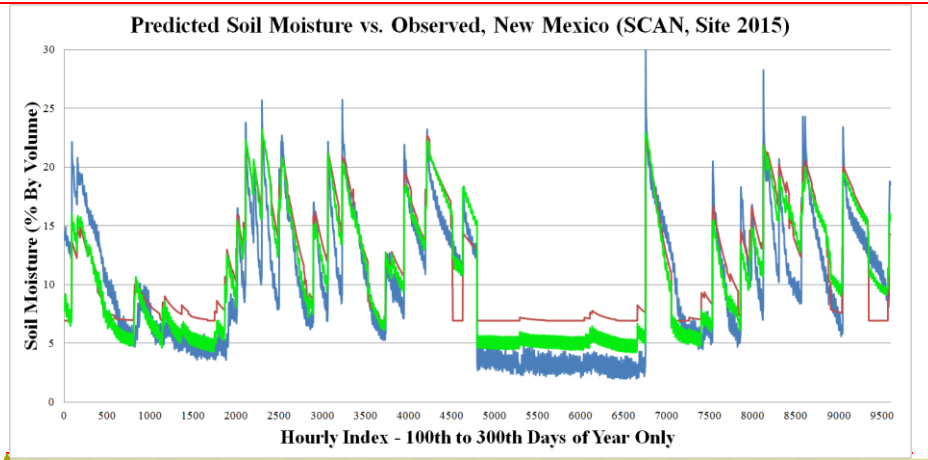


612
613

Figure 2, Improvements from machine learning (KNN) models of residuals.



614



615

Formatted: Font: (Default) Times New Roman, 12 pt, Font color: Text 1

616

Figure 3, Soil Moisture Soil Moisture Time Series, SCAN Site 2015, New Mexico (USA),
 617 Actual Soil Moisture (Blue Line), Diagnostic Soil Moisture Equation Estimate (Red Line),
 618 and Diagnostic Soil Moisture Equation with Machine Learning Error Correction (Green
 619 Line). Hydroclimate: IAQ (Intermediate Seasonality, Arid, Summer Peak Runoff)
 620 Soil Texture: Loamy Sand

Formatted: Space After: 0 pt

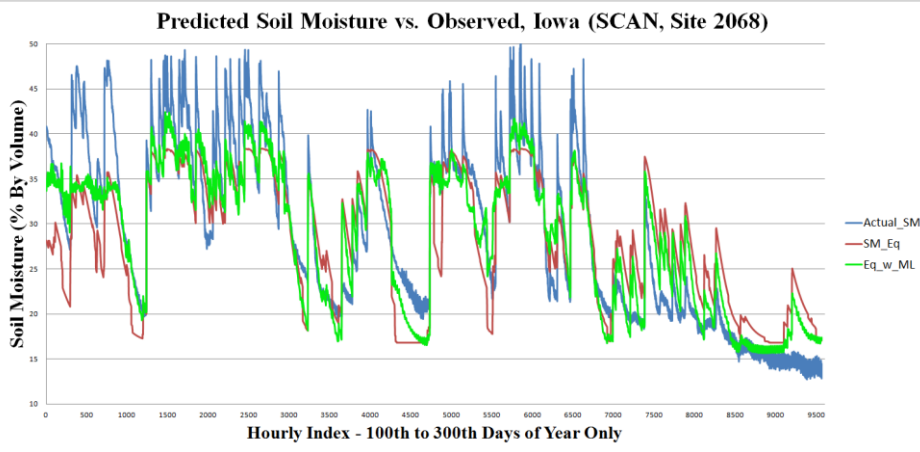
621

Formatted: Left, Space After: 0 pt

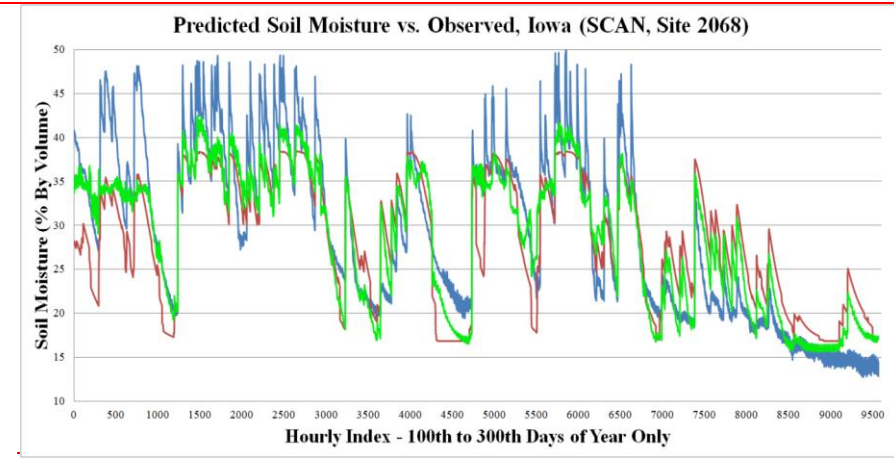
622

Formatted: Space After: 0 pt

623



624



625

626

627

628

629

630

631

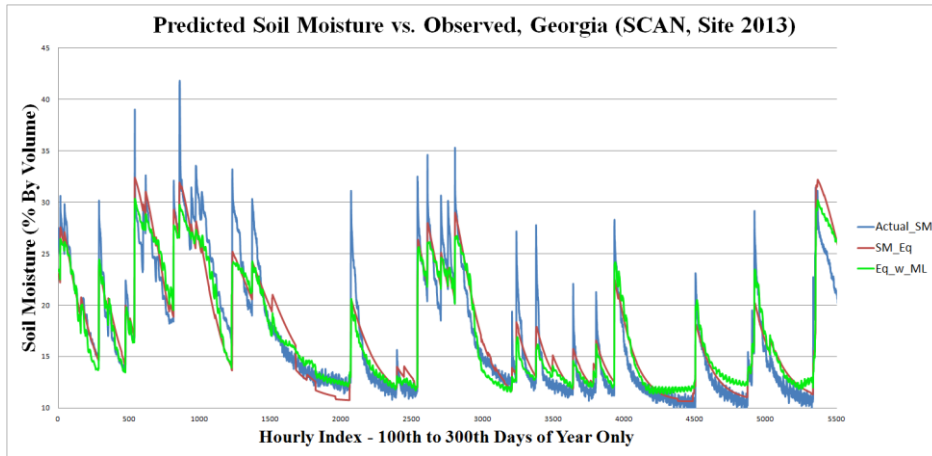
632

633

634

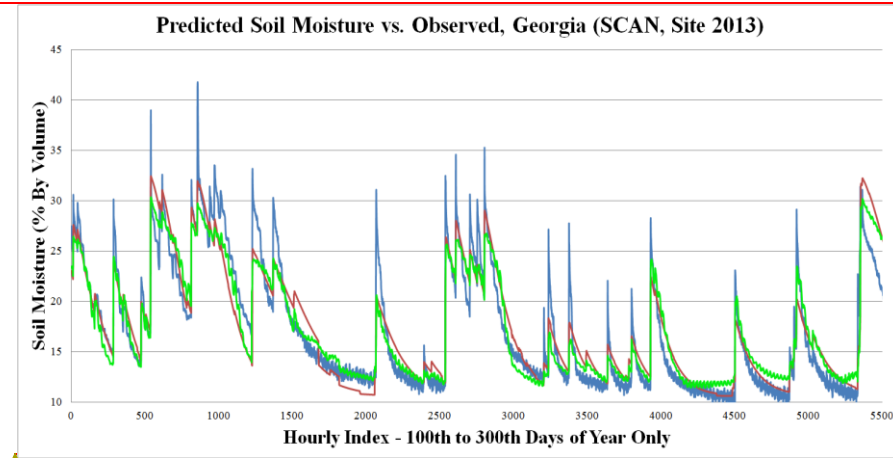
Figure 44, Soil MoistureSM Time Series, SCAN Site 2068, Iowa (USA), Actual Soil Moisture (Blue Line), Diagnostic Soil Moisture Equation Estimate (Red Line), and Diagnostic Soil Moisture Equation with Machine Learning Error Correction (Green Line) line colors from Fig. 3
Hydroclimate: ISCJ (Intermediate Seasonality, Semi-Arid, Winter Peak Runoff, Summer Peak Precipitation)
Soil Texture: Silty Clay Loam

Formatted: Space After: 0 pt
 Formatted: Centered, Space After: 0 pt



Formatted: Space After: 0 pt

635



Formatted: Font color: Text 1

636

Figure 55, Soil Moisture SM Time Series, SCAN Site 2013, Georgia (USA), line colors from Fig. 3

637

638

639

640

641

642

643

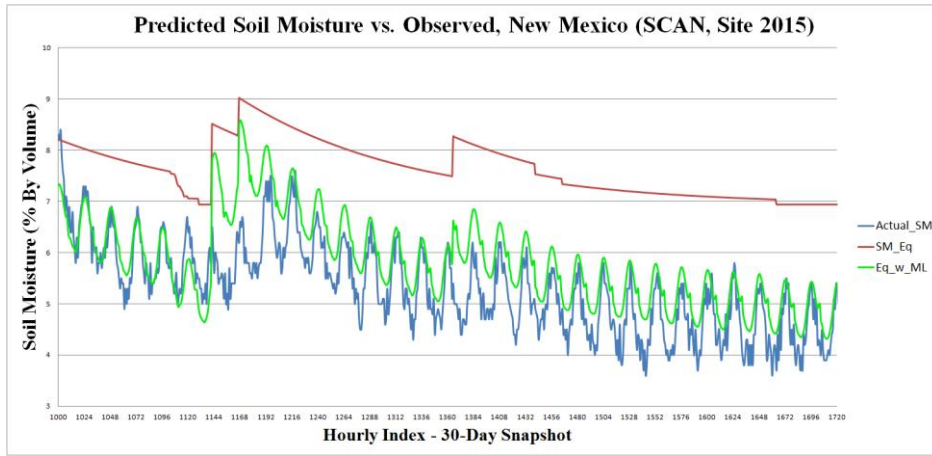
644

645

Actual Soil Moisture (Blue Line), Diagnostic Soil Moisture Equation Estimate (Red Line), and Diagnostic Soil Moisture Equation with Machine Learning Error Correction (Green Line)

Hydroclimate: LWC (Low Seasonality, Winter Peak Precipitation, Winter Peak Runoff)
Soil Texture: Sandy Loam

Formatted: Left, Space After: 0 pt



646

647

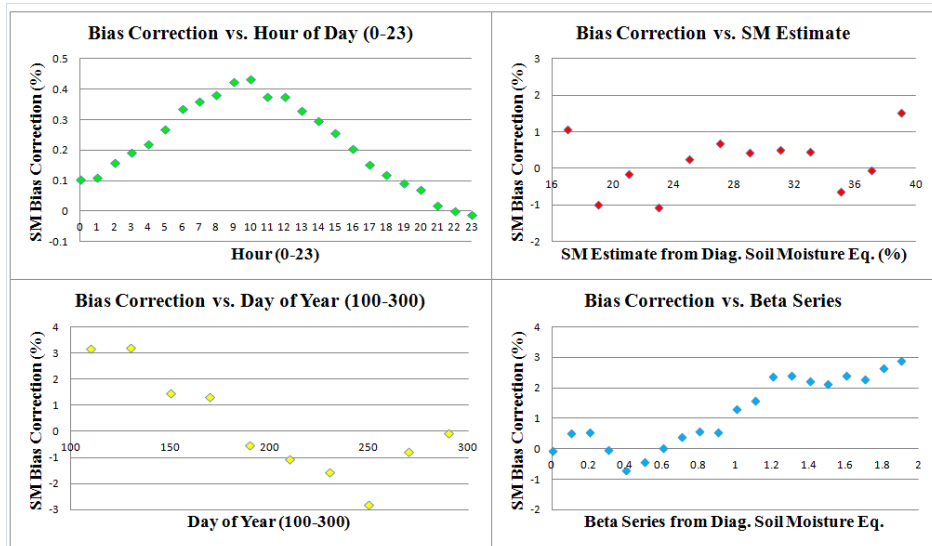
648

649

Figure 6, Soil Moisture Time Series, SCAN Site 2015, New Mexico (USA), Actual Soil Moisture (Blue Line), Diagnostic Soil Moisture Equation Estimate (Red Line), and Diagnostic Soil Moisture Equation with Machine Learning Error Correction (Green Line)

Formatted: Normal, Line spacing: single

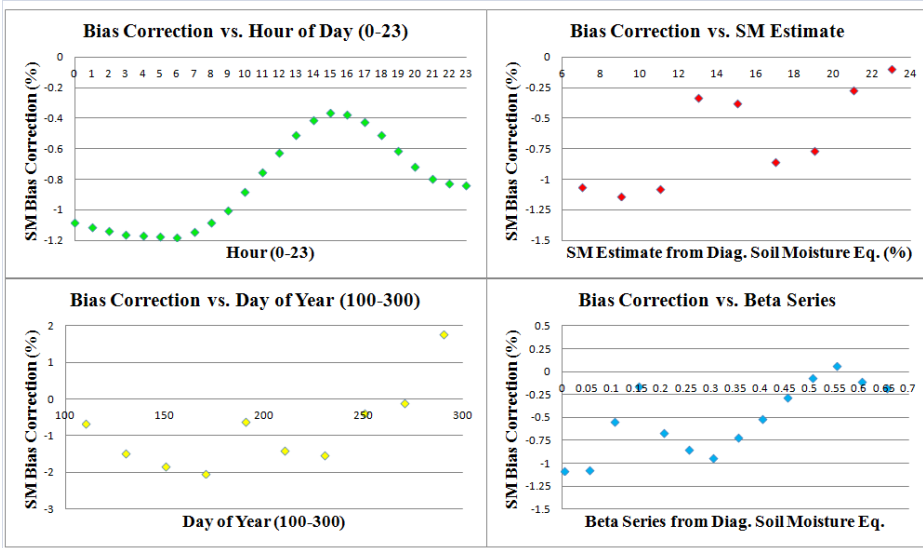
Formatted: Font: Bold



650

651

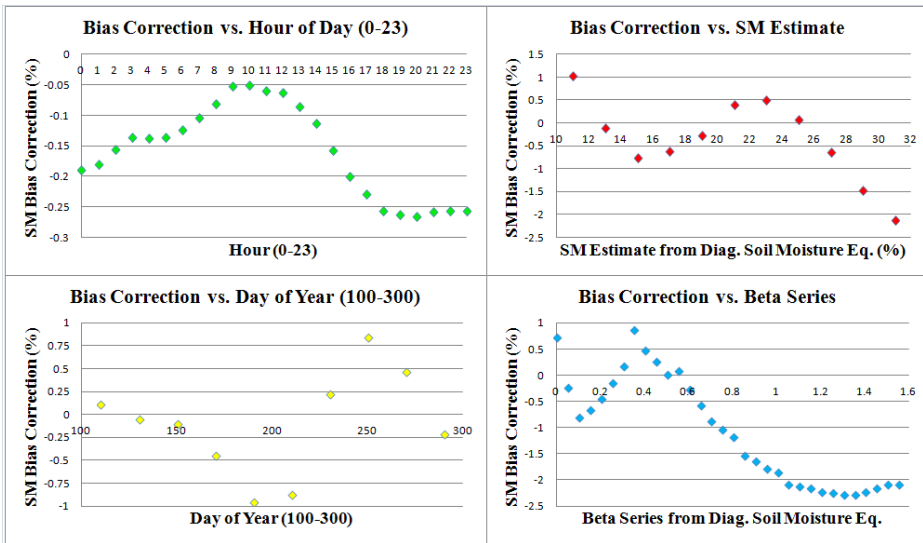
Figure 7, Bias Correction Analysis, SCAN Site 2015 (IAQ, Desert, Loamy Sand)



652

653

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

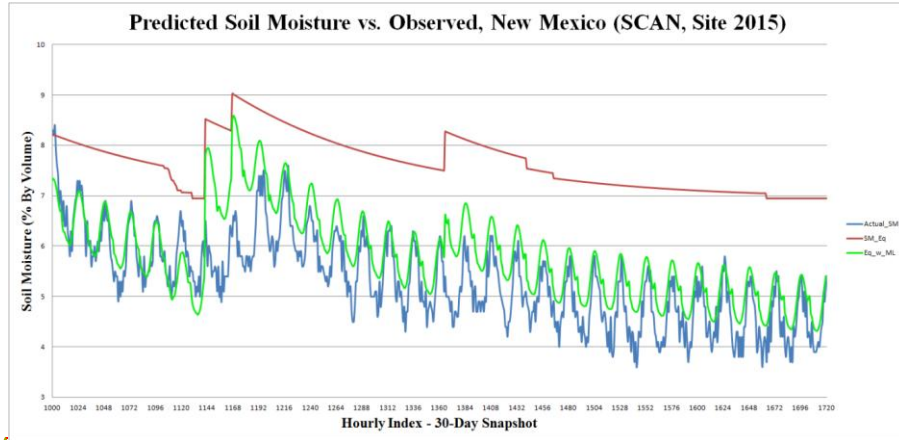


654

655

656

Figure 9, Bias Correction Analysis, SCAN Site 1213 (LWC, Woods, Sandy Loam)



Formatted: Font color: Text 1

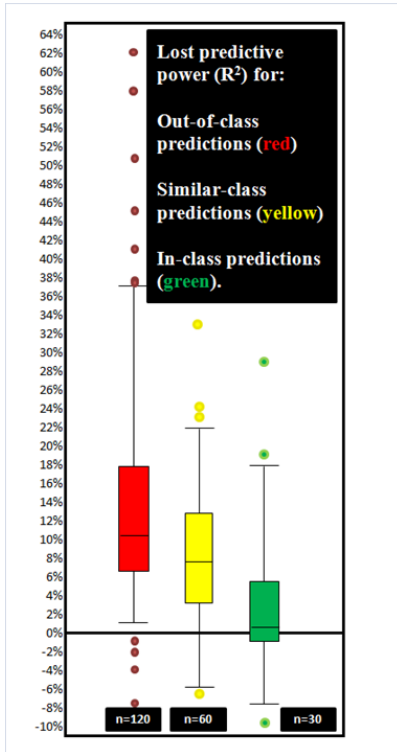
657

658

659

660

Figure 6, Soil Moisture Time Series, SCAN Site 2015, New Mexico (USA), Actual Soil Moisture (Blue Line), Diagnostic Soil Moisture Equation Estimate (Red Line), and Diagnostic Soil Moisture Equation with Machine Learning Error Correction (Green Line)



661

662 | **Figure 710, Loss of Predictive Power (R^2) (y-axis) Between Baseline Predictions (model**
 663 **calibrated in the same watershed) and Cross-Validation Predictions (model calibrated in**
 664 **other watersheds)**

665

666

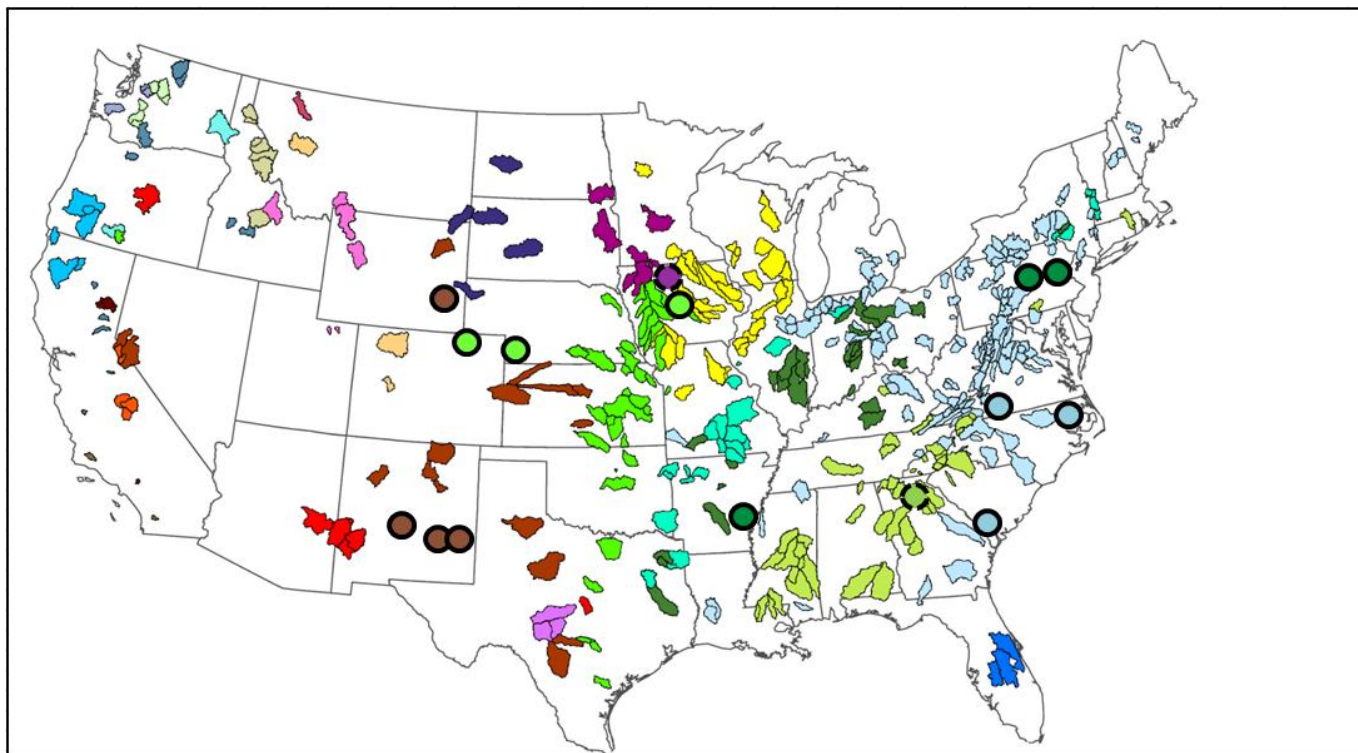


Figure 811, 428 MOPEX catchments colored by hydro-climatic class (Coopersmith et al, 2012).
 15 SCAN sensors (for which the Diagnostic Soil Moisture Equation is calibrated) are shown as colored circles.
 Circle colors correspond to the hydro-climatic class of the point in question.
 Circles with dotted borders are unique (no other sensor for calibration is available within that class)

Formatted: Font: (Default) Times New Roman, 12 pt, Bold, Font color: Text 1

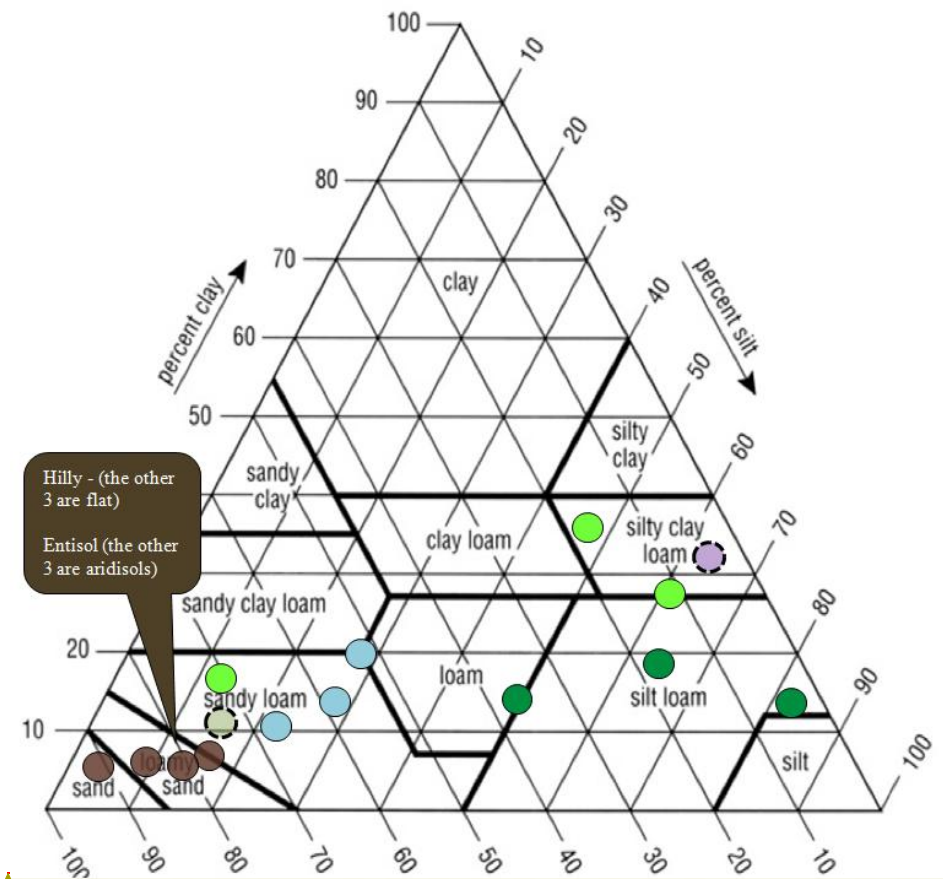


Figure 9, The 15 SCAN sensors, color-coded to match their hydro-climatic class.

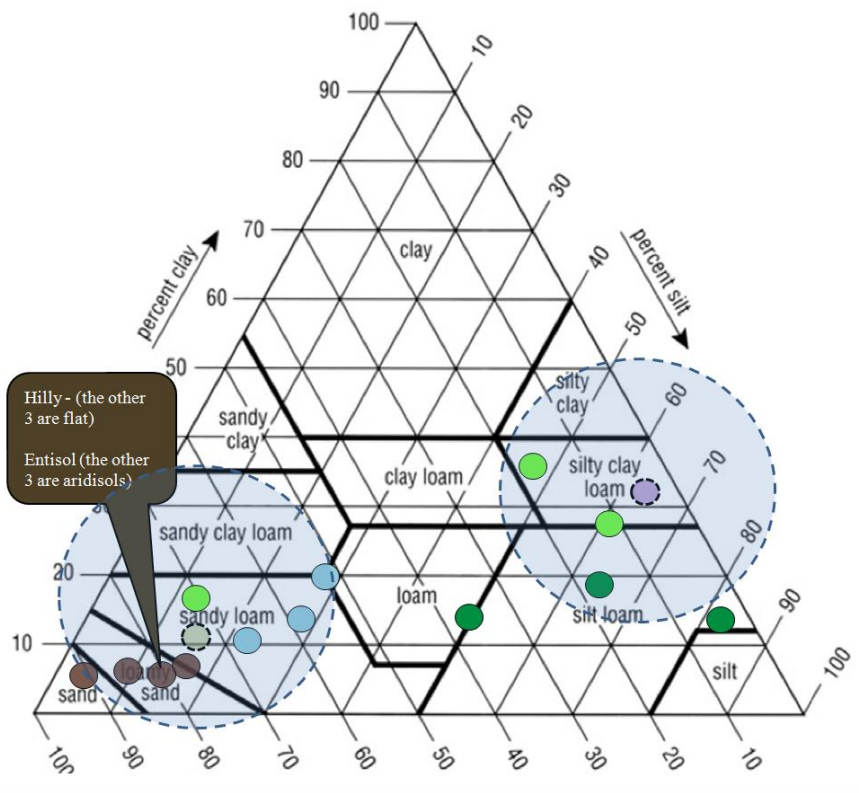


Figure 12. The 15 SCAN sensors, color-coded to match their hydro-climatic class, with similar soil textures shaded.

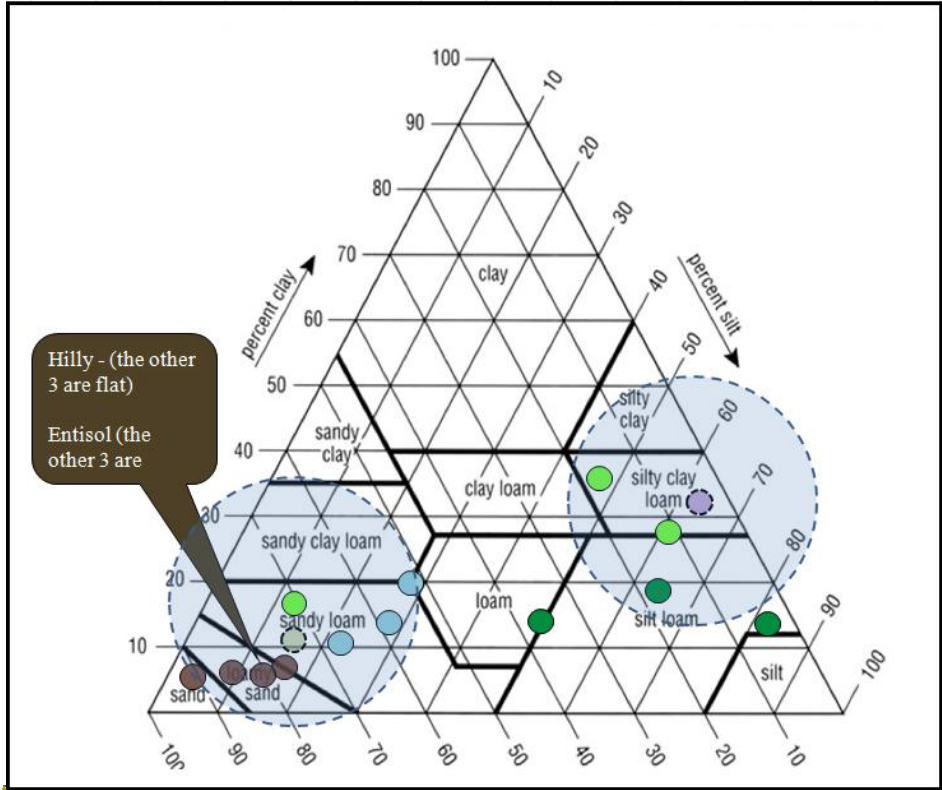


Figure 10, The 15 SCAN sensors with similar soil textures shaded.

Formatted: Font: (Default) Times New Roman, 12 pt, Font color: Text 1

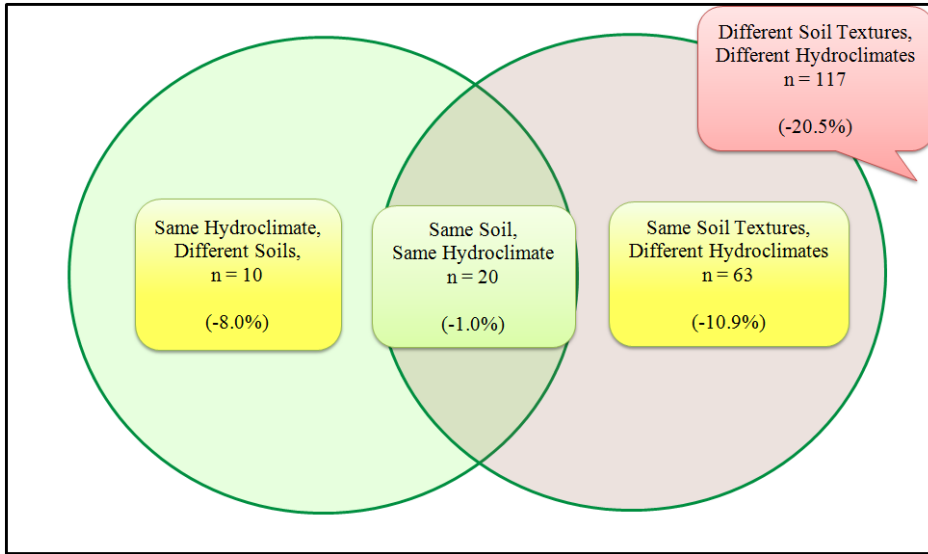


Figure 13 Venn-Diagram of Modeling Errors with Similar and Different Soils and Hydro-climates

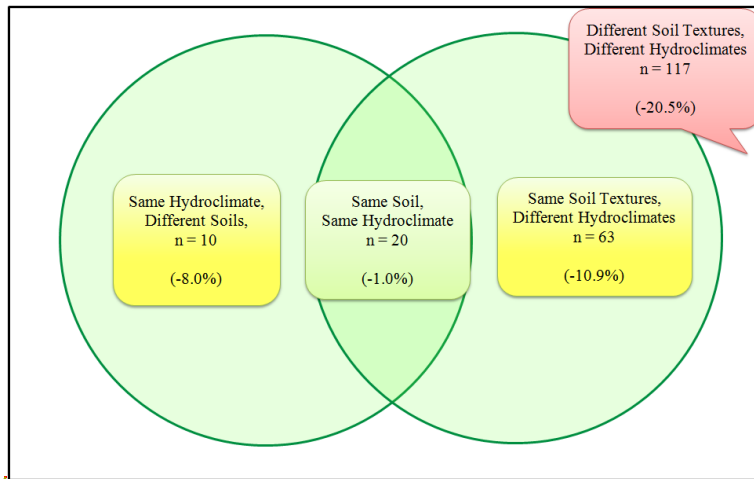


Figure 11 Venn-Diagram of Modeling Errors with Similar and Different Soils and Hydro-climates

Formatted: Font: (Default) Times New Roman, 12 pt, Font color: Text 1

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

	Unrelated Class	Similar Class	Same Class
Median	-10.5%	-7.3%	-0.8%
Mean	-13.7%	-7.7%	-3.4%
Standard Deviation	1.0%	1.1%	1.4%

Table 2, Cross-Validation Results

~~Works Cited~~ References

Barricelli, N.A. "Numerical testing of evolution theories. Part II. Preliminary tests of performance, symbiogenesis and terrestrial life". *Acta Biotheoretica* (16): 99–126. 1963.

~~Berghuijs, W.R., Sivapalan, M., Savenije, H.H.G., and Woods, R.A. "The seasonal water balance as a window to explore catchment similarity at various time scales." In preparation, 2014.~~

Capehart, W. J., and Carlson, T.N. "Estimating near-surface soil moisture availability using a meteorologically driven soil water profile model," *J. Hydrol.*, 160, 1– 20, 1994.

Cheng, C.T., Zhao, M.Y., Chau, K.W., and Wu, X.Y. "Using Genetic Algorithm and TOPSIS for Xinanjiang model calibration with a single procedure." *Journal of Hydrology*. Vol. 316, Issue 1-4, pp. 129-140, 2006.

Chico-Santamarta, L., Richards, T., and Godwin, R.J. "A laboratory study into the mobility of travelling irrigators in air dry, field capacity and saturated sandy soils." *American Society of Agricultural and Biological Engineers Annual International Meeting 2009*. Volume 4, 2009, Pages 2629-2646.

Choudhury, B. J. and Blanchard, B.J., "Simulating soil water recession coefficients for agricultural watersheds," *Water Resour. Bull.*, 19, 241–247, 1983.

Coopersmith, E., Yaeger, M., Ye, S., Cheng, L., and Sivapalan, M. "Exploring the physical controls of regional patters of flow duration curves – Part 3: A catchment classification system based on regime curve indicators." *Hydrol. Earth Syst. Sci.* doi:10.5194/hess-16-1-2012.

Entekhabi, D. and Rodriguez-Iturbe, I., "Analytical framework for the characterization of the space-time variability of soil moisture," *Adv. Water Resour.*, 17, 35–45, 1994.

Farago, T., "Soil moisture content: Statistical estimation of its probability distribution," *J. Clim. Appl. Meteorol.*, 24(4), 371– 376, 1985.

Fix, E., Hodges, J.L. Discriminatory analysis, nonparametric discrimination: Consistency properties. Technical Report 4, USAF School of Aviation Medicine, Randolph Field, Texas, 1951.

Gamache, R.W., Kianirad, E., and Alshawabkeh, A.N. "An automatic portable near surface soil characterization system." Geotechnical Special Publication, Issue 192, 2009, Pages 89-94.

Goldberg, D.E. Genetic Algorithms in Search, Optimization, and Machine Learning. Addison-Wesley Professional. 1989.

Jackson, T.J., Bindlish, R., Cosh, M.H., Zhao, T., Starks, P.J., Bosch, D.D., Seyfried, M., Moran, M.S., Goodrich, D.C., Kerr, Y.H., Leroux, D., "Validation of soil moisture and ocean salinity (SMOS) soil moisture over watershed networks in the U.S." IEEE Transactions on Geoscience and Remote Sensing, Volume 50, Issue 5, Part 1, 1530-1543. May 2012.

Jones, H. G. (2004). "Irrigation scheduling: advantages and pitfalls of plant-based methods." *Journal of Experimental Botany*, 55, 2427-2436

[Koster, R.D. and Mahanama, S.P.P. 2012: "Land Surface Controls on Hydroclimatic Means and Variability." *J. Hydrometeor.* 13, 1604–1620. 2012. doi: <http://dx.doi.org/10.1175/JHM-D-12-050.1>](#)

O'Neill, P., Entekhabi, D., Njoku, E, and Kellogg, K. "The NASA Soil Moisture Active Passive (SMAP) Mission: Overview". NASA. Goddard Space Flight Center, Jet Propulsion Laboratory. http://ntrs.nasa.gov/archive/nasa/casi.ntrs.nasa.gov/20110015242_2011016052.pdf (Retrieved 14 May 2013.)

[Orth, R.A., Koster, R.D.B., Seneviratne, S.I.A. "Inferring soil moisture memory from streamflow observations using a simple water balance model." *J. Hydrometeor.* Vol.14, Iss. 6, 1773-1790, 2013.](#)

Pan, F., "Estimating daily surface soil moisture using a daily diagnostic soil moisture equation. *Journal of Irrigation and Drainage Engineering*, 138(7), 625-631, 2012.

Pan, F., and Peters-Lidard, C. D. (2008). "On the relationship between the mean and variance of soil moisture fields." *Journal of the American Water Resources Association*. 44(1), 235-242.

Pan, F., Peters-Lidard, C. D., and Sale, M. J. (2003). "An analytical method for predicting surface soil moisture from rainfall observations." *Water Resources Research*, 39(11), Art.

[Rosenbaum, U., J.A. Huisman, J. Vrba, H. Vereecken and H.R. Bogaen. 2011. Correction of temperature and electrical conductivity effects on dielectric permittivity measurements with ECH2O Sensors. *Vadose Zone J.* 10: 582-593. doi:10.2136/vzj2010.0083.](#)

Saxton, K. E., and Lenz, A.T, "Antecedent retention indexes predict soil moisture," *J. Hydraul. Div. Proc. Am. Soc. Civ. Eng.*, 93, 223– 241, 1967.

Sharifat, K., and Kushwaha, R.L., "Sinkage simulation model for vehicles on soft soil." 2000 ASAE Annual International Meeting, Technical Papers: Engineering Solutions for a New Century 1, pp. 2549-2553

Simunek J., Sejna, M., van Genuchten, M. "The HYDRUS-1D software package for simulating water flow and solute transport in two-dimensional variably saturated media. Version 2.0, IGWMC –TPS –70, International Ground Water Modeling Center, Colorado School of Mines, Golden, CO, 1998.

Singh, A., and Minsker, B.S., "Uncertainty-based multiobjective optimization of groundwater remediation design," *Water Resources Research*, 44, W02404, doi:10.1029/2005WR004436.

Silva, R.B., Lancas, K.P., Miranda, E.E.V., Silva, F.A.M., and Baio, F.H.R. "Estimation and evaluation of dynamic properties as indicators of changes on soil structure in sugarcane fields of Sao Paulo State – Brazil." *Soil and Tillage Research* 103 (2), pp. 265-270. 2009.

Steel, R.G.D, and Torrie, J. H., *Principles and Procedures of Statistics with Special Reference to the Biological Sciences.*, McGraw Hill, 1960, pp. 187, 287

Wetzel, P. J., and Chang, J. T., "Evapotranspiration from nonuniform surfaces— A 1st approach for short-term numerical weather prediction," *Mon. Weather Rev.*, 116, 600–621, 1988.

Zhang, X., Srinivasan, R., and Bosch, D. "Calibration and uncertainty analysis of the SWAT model using Genetic Algorithms and Bayesian Model Averaging." *Journal of Hydrology*, 374, pp. 307-317 (2009).

Formatted: Space After: 0 pt

