We would like to thank the referees and the editor for their constructive comments. We first provide a general answer to the criticisms received and a brief summary of the revisions made before addressing each referee's comments in detail. Please note that referees comments are quoted in black bold font while our answers are in normal black font.

GENERAL ANSWER

Three main issues were underlined by the two referees and the editor: (1) there was no justification for the mathematical models tested to investigate the point-scale relationships between soil moisture and surrogate measures; (2) the methodology for choosing the best point-scale model was unclear; and (3) there were several references to other in press or submitted papers that seemed to imply that the question investigated in the current paper was dependent upon external data and analyses.

(1) Regarding the mathematical models tested: we made no assumptions about the shape of the point-scale relationships between soil moisture and surrogate measures that we wished to investigate. Since such an exercise has not been done before, we could not postulate on the shape and form of the relationships. We tested a range of relationships from linear to nonlinear models with various degrees of complexities and especially various rates of increase (e.g. quadratic versus cubic). In the revised manuscript, we still test three increasing monotonic polynomial models (i.e. linear, quadratic and cubic) as well as a logarithmic model. In the revision, we replaced the "inverse" model with an exponential one and added a logistic model so as to consider the possibility of asymptotically reached triggering conditions.

(2) Regarding the methodology for ranking the models performance and choosing the best one: our initial strategy was to use R-square values so as to keep a methodology that is easy to replicate. In the revised manuscript, we use a three-step procedure based on Adjusted R-square values, values of the Akaike Information Criterion (AIC) and Akaike weights or probabilities associated with each of the tested models. The use of the Adj. R-square is indeed more rigorous as rightfully pointed by one of the referees. Once we have tried to fit all six mathematical models to the data (i.e. point-scale soil moisture and surrogate measure), we use the Adj. R-square as a preliminary step to label sampling locations for which none of the models can achieve a successful data fit. We then use the AIC since it quantifies how well each model reflects the data. Not only does the lowest AIC score tell us which model is most likely to be correct, but the Akaike weights are also useful in assessing how much more likely in comparison to other models tested, thus making it possible to reach a more solid conclusion about the best model among the set being evaluated. We have reproduced the whole revised "data analysis" section at the end of this document. It is worth mentioning that even though we now use a more sophisticated approach, results have barely changed; some exceptions aside in the point-scale relationships. The take home message remains the same.

(3) Regarding the multiple references to papers submitted to other journals and that are still in review: we must say that the results obtained in these other papers provide a basis for the current statistical analysis but they do not condition the analysis. Some rewriting was done in order to clarify this in the text. Out of the four papers being referred to, three are now in press and available online and one has been recommended for publication pending minor revisions. In order to clarify the situation, a table summarizing the results from these other papers has been added to the revised manuscript (Table 1).

We have worked on a revised version of the manuscript that we are now ready to submit to your evaluation. We very much appreciated the two comprehensive reviews and the editor comments and we attempted to revise the manuscript accordingly. As the whole dataset was re-processed with a revised methodology for choosing and ranking the regression models. The results of the re-analysis are not substantially different from those in the initial manuscript. Given this revised methodology and a series of contextual and interpretational issues which have been clarified, we hope that the paper is now suitable for publication in HESS.

SPECIFIC ANSWERS TO REFEREES AND EDITOR'S COMMENTS

Referee #1

General Comments

This paper describes the analysis of a detailed dataset of soil moisture measurements (121 locations, 4 depths, 16 sampling days) in order to test the ability of a variety of antecedent wetness indices to represent the soil moisture state in different parts of a small catchment. Relationships between the soil moisture at individual locations and both the mean catchment soil moisture, and catchment discharge, are also investigated.

The scientific question is relevant and will be of interest to the readership of HESS, and the dataset provides a valuable opportunity for this analysis. The paper makes good use of Figures to clearly represent complex results. I would therefore recommend the paper for publication after moderate revisions to address the choice of relationships fitted, as outlined in the following paragraph.

The paper would benefit from more careful consideration and explanation of the relationships fitted between the various surrogates for AMC and the measured soil moisture data. The authors fit in each case five different regression models (linear, quadratic, cubic, semi-logarithmic and inverse (not clear what 'inverse' is)). There is no physical justification for any of these models so I assume they are chosen purely to give a range of possible curves, with varying complexity.

Details have been added in the text. Six models rather than five are now evaluated and we include the equations used for data fitting. We also mention that there was no physical basis underlying the choice of the six models. As mentioned by the reviewer, the models are chosen mainly to give a range of possible linear and non-linear relationships.

It is not clear whether the quadratic and cubic curves are constrained to be monotonic increasing.

They are constrained to be monotonic increasing. This was clarified in the revised manuscript.

I suggest that to quantify only the strength of the relationship between each surrogate and the soil moisture, that a distribution-independent correlation measure such as rank correlation would be more appropriate.

Thank you for this relevant suggestion. While appealing to sort the relationships in terms of strength, the use of the rank correlation would not have given us the likelihood that the model ranked in top place was significantly "better" than the one ranked in second place. We therefore decided to rely on an information criterion instead.

Although the authors state in Section 2.4 that nonlinear relationships will only be chosen where the data justifies it, no method of making that decision is identified, only that the improvement in the proportion of variation explained should be "tangible" (P10L8). A clear method for identifying when the higher-order relationships are justified is needed.

A three-step, more rigorous procedure has now been devised. Decision on the best model is now made based on Adjusted R-square values, corrected Akaike Information Criterion values and Akaike weights or probabilities. Details can be found in the text.

Specific Comments

On P3 L9 and P5 L27 Soil moisture is described as a 'major control on catchment response' and a 'critical hydrological state variable whose variation indicates active areas'. It would useful here to bring in and comment on the findings by Tromp-van-Meerveld, and the comment-and-reply by Western and Tromp-van Meerveld (citations below) on the question of whether soil moisture is the control on streamflow response or whether it is a passive signal, with transient saturation driving the streamflow response.

Tromp-van Meerveld, H. J., and J. J. McDonnell (2006), Threshold relations in subsurface stormflow: 2. The fill and spill hypothesis, Water Resour. Res., 42, W02411, doi:10.1029/2004WR003800.

Western, A. W., S. L. Zhou, et al. (2004). "Spatial correlation of soil moisture in small catchments and its relationship to dominant spatial hydrological processes." Journal of Hydrology 286(1-4): 113-134.
Tromp-van Meerveld, H. J. and J. J. McDonnell (2005). "Comment to "Spatial correlation of soil moisture in small catchments and its relationship to dominant spatial hydrological processes, Journal of Hydrology 286 : 113-134"." Journal of Hydrology 303(1-4): 307-312.
Western, A. W., S. L. Zhou, et al. (2005). "Reply to comment by Tromp van Meerveld and McDonnell on Spatial correlation of soil moisture in small catchments and its relationship to dominant spatial hydrological processes." Journal of Hydrology 303(1-4): 313-315.

You are right for pointing out that we only presented the rationale on which we based ourselves for our study; that is that soil moisture is a major control on catchment response. If soil moisture was proven to be a passive signal in the Hermine catchment, then the relationships patterns would have to be revisited or re-interpreted. Hence, we took your comment into consideration and added a few sentences not in the introduction but rather at the end of the discussion section.

P3L21 - Mention here that the difficulty in collected soil moisture data is tied largely to the heterogeneity of soil moisture across a typical catchment.

This is now mentioned in the revised manuscript.

P4L15 - You say that baseflow indices are less useful as they are dependent on the season of the year - however that is also a problem with the APIs you are using (stated at L5).

That is right. That similarity between the two types of indices is underlined in the revised manuscript.

P4L16-19 A Key advantage of using the ABFI is that it is a proxy for a combination of shallow soil moisture and deeper groundwater conditions. Flow is extensively used as a proxy for antecedent wetness in the data-based mechanistic approach by Young (example citation below).
Young, P. C. and K. J. Beven (1994). "Data-based mechanistic modeling and the rainfall-flow nonlinearity." Environmetrics 5(3): 335-363.

Thank you for pointing that reference to us. It has been added to the revised manuscript.

P6L17 "the sole reference to AMCs measures" I don't understand what this means

This has been rephrased in the revised manuscript.

P8L22-29 Explain why these indices were chosen over the other options mentioned (e.g. API)

The indices evaluated were especially chosen for their computational simplicity. A sentence stating to this effect was added to the manuscript.

P10L26 Over what timescale was PET calculated - that may strongly influence the results

PET was calculated on a diurnal timescale. This is now clarified in the revised manuscript.

P12L5 Explain what CD_DISCH means

CD_DISCH was defined at the end of section 2.3. It refers to current-day discharges, namely mean daily discharges for each of the 16 soil moisture surveys.

P14L16 "locations for which soil moisture is strongly related with AMCs can be labelled as source areas" - why?

In the initial version of the manuscript, we tried to label sampling locations as potentially active and contributing areas. 'Active' areas were hypothesized to be locations where soil moisture content was significantly controlled by AMCs (surrogate measures) while "contributing" areas were hypothesized to be locations where soil moisture content influence catchment discharges. We understand that this hypothesis has no clear physical basis and may be perceived as far-fetched. Because its omission does not change the message conveyed by the paper, we do not refer to it anymore in the revised manuscript.

P15L2 Nonlinear relationships are attributed to amount and timing of precip as well as ET, but the nonlinear relationship may just be a feature of the way that water is transmitted through the soil - this does not need to be linear.

A brief sentence echoing that idea/hypothesis has been added to the text.

Referee #2

The authors describe a study where a soil moisture data set was used to assess the appropriateness of surrogate antecedent moisture indices. This soil moisture data set has a relatively high spatial extent of 121 locations on a 15 by 15 m grid covering a catchment of 5 ha. At each location measurements were taken in 4 depths at 16 occasions. The antecedent moisture indices focus on meteorological indices such as the antecedent precipitation index for various time spans and the days-since-precipitation index. These indices were correlated with the point measurements of soil moisture, resulting in maps of degrees of correlation as well as maps of the correlation model types which produced the best fit (i.e. linear, quadratic, cubed, etc). In a next step the potential influence of topographical characteristics was

analyzed by using by further correlation tests. The authors found that the standard surrogates, the API for 7 or 10 days did not satisfactorily describe catchment soil moisture state. Furthermore the relationships between data and surrogates were found to be spatially heterogeneous (not very surprising given the known strong variability of soil moisture and soil physical characteristics), possibly indicating active or contributing areas and spatially variable threshold processes.

Overall I am most impressed by the soil moisture data set and the question of the appropriateness of surrogates for antecedent moisture conditions is an important one; however, the analysis of "best fit model types" remains vague and does not seem to offer a lot of additional information.

The analysis of "best fit model types" has been modified and thoroughly described in the revised manuscript. The aim of the paper has also been clarified so as to better ground and justify the analysis.

In the discussion references are made to other publications of the same authors, two of which are in press and two of which are only listed as "submitted". These publications seem to contain data that would make the here described analysis more interesting and better founded. This could also be discussed as part of the introduction. However, there also seem to be some contradictions between these data sets and the here presented one, as mentioned by the authors. This should be discussed in more detail in the conclusions section.

The status of all papers has changed and this is now reflected in the revised manuscript. Out of the four being referred to, three are now in press and available online and one has been recommended for publication pending minor revisions. It must, however, be said that the contradictions underlined are not between the datasets but rather concern the ability of surrogate measures for AMCs to explain the hydrologically-relevant observations derived from different datasets taken from the same catchment. We took your advice and now mention the other studies in the introduction.

p.3334 l.14 "Considering that both "active" and "contributing" areas are important in assessing a catchment initial state, do surrogate measures for AMCs reflect these dynamics? From a spatially-distributed point of view, the fact that all catchment areas are not "activated" at the same time may indicate that they are responsive to different hydro-meteorological factors. Similarly, the non-uniform contribution of source areas to streamflow may point towards different triggering hydro-meteorological factors." – It seems like a more straight forward hypothesis would be that the main factors controlling the activation of contributing areas are structure, topology and - as a result - connectivity, which is also a threshold process depending on antecedent conditions and event characteristics.

That sentence has been rephrased in the revised manuscript.

What exactly are you referring to as "hydro-meteorological factors"? Please explain.

The phrase "hydro-meteorological factors" was meant to refer to antecedent conditions in general and storm characteristics. Given its lack of clarity, it has been changed throughout the revised manuscript.

You are also not looking at event response in this study – wouldn't it make sense to include the response in the analysis and look at what type of response is most likely connected with which antecedent conditions and then see if these conditions can or cannot be estimated by the proposed surrogate measures? Another possibility would be to classify catchment states in wet and dry and see for what conditions the surrogates are suited best.

I assume that by referring to "event response", you suppose that soil moisture surveys systematically took place just prior to and after storm events. That was not the case, and the only way for us to link observed soil moisture patterns to catchment response was using mean daily discharge values (CD_DISCH). We decided not to classify catchment states as it can be rather subjective and based on a wide variety of individual or combination of variables (discharge, soil moisture statistical distribution, soil moisture central tendency measures, MSMC, depth-specific mean soil moisture contents, etc.).

p.3336, l. 20: how is the probe pushed into the soil? Manually? Does the soil have to be soft in order to do this? How problematic are stones and roots? Or are you using permanently installed access tubes? If not, how do you assure you are measuring the exact same spot on all 16 occasions?

The probe is pushed manually into the soil. It does, indeed, pose some difficulties in the Hermine catchment given the stony/bouldery soils. However, it must be said that prior to the very first soil moisture survey in 2007, our preliminary investigations allowed us to assess the depth to the confining layer around each sampling location and the maximum soil moisture measuring depth that could be reached with the instrument. Since we were unable to use permanently installed access tubes on the whole catchment area, we do not measure exactly the same spot on all 16 occasions. From week to week, we constrained ourselves to taking soil moisture measurements in the same small radius (75-100 cm) around each sampling benchmark. There again, prior to the first soil moisture survey in 2007, we tested the consistency between multiple measurements taken in the 75-100 cm radius and found that the largest discrepancies occurred in dryer conditions. Hence, for the driest patterns, each soil moisture measurement is the field-calculated average of three separate measurements.

p.3337, l. 15: is the mean soil moisture content calculated over all depths and sampling points?

Yes. This has been clarified in the revised manuscript.

p.3338, l.28: what does "tangible" mean in this case? Using only R^2 means that there is no "punishment" for the use of more complicated models. More complicated models however, are always likely to produce better R^2 values. Wouldn't it make sense to use the adjusted R^2 in this case?

You are right about the use of the Adjusted R-square. In the revised version of the manuscript, we go even further since the decision on the best model is now made based on Adjusted R-square values, corrected Akaike Information Criterion values and Akaike weights or probabilities. Details can be found in the text.

p.3339, l.3: please explain the non-parametric Kruskal-Wallis test. Please also show some exemplary data and exemplary regressions.

Some details have been added with regards to the statistical test. Exemplary regressions are now shown in the new Fig. 7.

p.3339, l.10: please also show some of the original soil moisture data – how different are the moisture contents for the different depths? Is there a clear pattern here? Is the main variability found from sampling time to sampling time, from location to location or from depth to depth?

Some of the original soil moisture data is illustrated via three contrasted surveys in Fig. 2. To answer some of your questions, there is no clear pattern except that the catchment northern slope wets up "faster". The main variability is found from sampling time to sampling time. Also, soil moisture content is usually higher at the 5 and 15 cm depths in comparison to the 30 and 45 cm depths. A few sentences were added at the end of section 2.2 in the revised manuscript.

p.3343, l. 12: "locations for which soil moisture is strongly related with AMCs can be labeled as source areas" – please explain why

We here copy the answer provided to the first referee earlier.

> In the initial version of the manuscript, we tried to label sampling locations as potentially active and contributing areas. 'Active' areas were hypothesized to be locations whose soil moisture content was significantly controlled by AMCs (surrogate measures) while "contributing" areas were hypothesized to be locations whose soil moisture content influenced catchment discharges. We understand that this hypothesis has no clear physical basis, may be perceived as far-fetched, and its omission does not change the message conveyed by the paper. Therefore we do not refer to it anymore in the revised manuscript.

when you are referring to AMCs here, do you mean surrogate measures for AMCs?

Yes. This has been clarified where necessary in the text.

p.3343, l.17: " in order to substantiate this hypothesis" – which hypothesis are you referring to here?

On the contrary to topographic variables for which we have detailed spatial maps to compare with Fig. 4, 5 and 6, we only relied on non systematic field observations to say that canopy density is lower at locations where PET has a weak influence on the relationships patterns. This is why we refer to this as an hypothesis in the previous version of the manuscript, however that sentence was deleted in the revised manuscript because it is perhaps too speculative.

p.3343, l.19: what do you mean by much of the catchment is "activated by short-term AMCs"? Please rephrase.

That sentence has been rephrased.

p.3343-p.3344: two main hypotheses are listed here:

1. linear relationships with AP1, AP2, AP5 at depths of 5 and 15 cm and non-linear relationships at greater depths are explained by the fact " that the soil storage capacity is a function of the amount and timing of precipitation in addition to evapotranspiration, hence the nonlinear relationships." How do you explain the linear relationships at 5 and 15 cm?

It may be an issue of soil transmissivity governing the vertical drainage of water. A brief sentence echoing that hypothesis has been added to the text.

2. "Locations for which soil moisture is strongly related with catchment discharge can be considered as contributing areas." However, no statistically significant results were obtained that confirmed this hypothesis and the authors go on to say that "the obtained regression models may not necessarily reflect causal relationships".
In my opinion these statistical relationships might have the potential to help formulate hypotheses that then can be tested by additional data, data analysis or modeling but, as you say, they do not have to reflect causal relationships. However, keep in mind in your discussion that hypotheses can generally only be falsified. As to the hypotheses formulated as a result of this statistical analysis: maybe you could discuss how these hypotheses could be tested?

We totally agree with you when it comes to the usefulness of our statistical approach for hypothesis formulating. Hypothesis testing is another story. A few sentences have been added to the text to make that distinction.

p.3345, first paragraph: did you check the relationship between mean soil moisture and the surrogate measures for AMCs? How do they compare? Is AP5 also the best surrogate measure in this case?

A new figure (Fig. 7 in the revised manuscript) has been added for discussion. AP5 also seems as the best surrogate measure when MSMC is used as an indication of the catchment macrostate; the relationship between the two variables is exponential but rather weak (Adj. R-square of 0.4).

p.3345, l.18 and p.3346, l.1: you mention that your findings conflict with the findings of Kohler and Lindsey – do you have any suggestion why you obtained different results?

We suspect that it might be a matter of catchment size as Kohler and Lindsey (1951) refer to large river basins. A sentence has been added to the text.

p.3346, l.6: the references to Ali et al. 2010 are unclear as there are 4 references this could refer to. Please clarify.

The "referencing" system has been modified.

p.3346, l.8: what do you mean by broad and very large characteristic scale? This sentence is unclear. Please explain.

This has been rephrased.

l.10: "the dynamics illustrated in Fig.5" – change dynamics to "patterns".

Done

l.14-18: what do you make of this contradiction?

The contradiction only concerns AP2 here and the reason for it is unclear.

p.3346, ll.3-18: the whole discussion of the influence of AP measures on different characteristic scales is slightly confusing as the type of analysis carried out in this study does not become clear. Also if the applicability of surrogate measures was already tested here, in what way does the here presented study give additional information? This could also be discussed in the introduction where reference to the prior studies should also be made.

References to characteristic scales have been rephrased and much simplified in the revised manuscript. The applicability of surrogate measures was not tested in the Ali *et al*. 2010 WRR paper. However, the influence of some of these surrogate measures on soil moisture was scale-dependent, which is of interest while looking at some of the relationships patterns shown in Fig. 4 and 5.

p.3346, l.27: what do you mean by "scenarios"? event types?

Scenarios meant hydrological conditions defined by discharge, AP2 and AP7 values. However, this explanation has been rephrased and much simplified in the revised manuscript.

p.3347, l.26: I would suggest changing "to understand the hydrologic behavior of a catchment" to "to describe the moisture conditions within a catchment".

The change has been made.

p.3348, l.6-8: "soil moisture was not related to cumulative rainfall amounts over antecedent temporal windows of 7 or 10 days" – this conclusion does not match with some of the findings of your other studies and you explain this by the type of relationship/regression model used – this should also be mentioned here or you should weaken the conclusion by limiting it to the types of regression models tested in the here presented study.

That is true. That sentence has been modified to nuance our previous statement.

Figures:

Fig 1: the 15x15m grid cells do not match with the scale bar of 100 m – there should be about 7 pixels per 100m.

Right. The scale bar has been modified.

Fig.2: this figure is also presented in Ali et al. 2010 in WRR; however, I still think it is useful here.

It is, indeed, the same figure as in the WRR paper.

What depth do the measurements presented here refer to?

For each measurement depth (indicated on the left in the figure), three patterns associated with three soil moisture surveys are presented.

Fig.4 and 5 – maybe the types of relationships would be easier to make out if the figures would be in color.

Fig. 4, 5 and 6 are now in color.

Fig. 7: please add the total number of points which make up each boxplot for the different relationships. I assume the data contained in the boxplots is the number of significant correlations found for each elevation and relationship type? Please clarify this in the figure caption.

The figure as well as the caption were modified. It is now referred to as Fig. 8.

Fig. 8: It seems like the plots showing the log of the contributing areas are not discussed in the text. You should either discuss the findings presented here or remove the plots.

Right. That figure has been redrawn with different variables in the revised manuscript. It is now referred to as Fig. 9.

Technical corrections

p.3341, l.16: correction – "…however, the higher the elevation, the higher the order of the relationship."

That sentence no longer stands in the text.

Editor

Both reviewers agree that this manuscript presents a very interesting soil moisture data set and that it adresses an important science question (appropriateness of surrogates for antecedent moisture conditions). However, it is clear from the reviews that in its current form, the manuscript does not satisfactorily explore the data set, some methodological choices are questionable (e.g. the type of fitted curves), it is not entirely clear how the data has been measured and that the obtained results should be further discussed.

The methodology has been revised and the data re-processed accordingly.

Both reviewers give detailed indications of how the analysis could be improved. I invite the authors to carefully answer all the points before revising their manuscript accordingly. The analysis part will need some considerable improvement, both in terms of the chosen methods (see e.g. the suggestion of using rank correlation) and in terms of reference to existing work and process hypotheses (see reviewer 2). This will certainly result in some major revisions but I am confident that the additional work will result in a manuscript of high interest for the readers of HESS.

All points have been answered, except the one concerning the use of rank correlation to choose the "best" models. Justification for our choice of Adj. R-square and AIC was previously provided.

Final comment: please make sure that the revised manuscript contains all relevant details without refering to unpublished (submitted) material.

One of the papers is still in the review process and will be re-submitted soon.

Appendix
Data analysis section as written in the revised manuscript

N.B. #1: Text in normal red font signals changes with reference to the original manuscript
N.B. #2: Tables and figures numbering have changed give the addition of one table and one figure.

Our methodology intended to answer two research questions (Fig. 3). Firstly, we aimed to determine the nature and the strength of the relationships between point-scale soil moisture (i.e. soil moisture measured at each sampling point) and each of the AMCs and catchment response surrogates previously described. We hypothesized that the identified relationships would illustrate the variety of point-scale hydrologic behaviours that can be encountered within the Hermine catchment. Secondly, we examined the spatial organization of the nature and the strength of these point-scale relationships to link them with possible topographic controls.

For the determination of point-scale relationships, data cases were soil moisture survey dates (n = 16). When the aim was to evaluate the ability of AMCs measures to describe soil moisture patterns, the independent variable was the chosen surrogate for AMCs and the dependent variable was the depth-specific, point-scale soil moisture content. In order to assess the potential of $MSMC$ to represent the Hermine catchment macrostate, we rather considered $MSMC$ to be an independent variable while point-scale soil moisture was the dependent one. Lastly, in order to identify catchment areas that might contribute to streamflow discharge, the statistical procedure detailed below was also applied using point-scale soil moisture as the dependent variable and CD_DISCH as the independent one. No postulate could be made on the form of the relationship between the dependent and the independent variables since no such exercise has been done before. Six regression models (i.e. linear, quadratic, cubic, exponential, logarithmic and logistic), which represent six different types of possible relationships, were fitted to the data and compared so as to select the one with the best fit. Model equations can be written as follows:

Linear model
$$\mathrm{Y} = a \cdot \mathrm{X} + b \tag{3}$$

Quadratic (monotonic increasing) model
$$\mathrm{Y} = a \cdot \mathrm{X}^2 + b \cdot \mathrm{X} + c \tag{4}$$

Cubic (monotonic increasing) model
$$\mathrm{Y} = a \cdot \mathrm{X}^3 + b \cdot \mathrm{X}^2 + c \cdot \mathrm{X} + d \tag{5}$$

Exponential model
$$\mathrm{Y} = a \cdot e^{b \cdot \mathrm{X}} \tag{6}$$

Logarithmic model
$$\mathrm{Y} = a + b \cdot \ln(\mathrm{X}) \tag{7}$$

Logistic model

$$Y = \frac{c}{1 + a \cdot e^{-b \cdot X}} \quad (8)$$

Where Y is the dependent variable, X is the independent variable, and *a*, *b*, *c* and *d* are model parameters. There was no physical basis for the choice of the six mathematical models; the aim was rather to explore the dataset using different models with various degrees of complexity. In each case, a least squares-like regression method was used for all models, which means that the fitting of each model to the data had to minimize the squared differences between observed and predicted values. Selection of the best mathematical model among the six tested was then performed in three steps.

First, the adjusted coefficient of determination (Adj. R-square) was used to discard any model that would only explain a small proportion of the variance in the data. Throughout this paper, we refer to R-square as the proportion of variance in the dependent variable that is explained by the chosen regression model. It can be computed for any linear or nonlinear model:

$$\text{R-square} = 1 - \frac{\text{residual SS}}{\text{total SS}} \quad (9)$$

where SS refers to a sum of squares. Total SS is the total amount of variability in the dependent variable while residual SS is the amount of variability that still cannot be accounted for after the regression model is fitted to the data. Given that the value of R-square often increases when a nonlinear model is used instead of a linear relationship, the use of the Adj. R-square is more adequate in the context of multiple models evaluation and comparison as it assesses the goodness of fit while taking into consideration the numbers of degrees of freedom of the numerator and the denominator of R-square (Legendre and Legendre, 1998):

$$\text{Adj. R-square} = 1 - (1 - \text{R-square})\left(\frac{\text{total d.f.}}{\text{residual d.f.}}\right) \quad (10)$$

Hence, the Adj. R-square "penalizes" models bearing a large number of parameters. For the current analysis, if all six models failed to produce an Adj. R-square value exceeding 0.3, then the relationship between point-scale soil moisture and the surrogate measure being evaluated was labelled as "not significant". Otherwise, only the models with an Adj. R-square exceeding 0.3 were kept for further consideration towards choosing the best fitting model.

As a second step in the best model selection, the models with an Adj. R-square exceeding 0.3 were ranked according to their corrected Akaike Information Criterion value. The Akaike Information Criterion or AIC (Akaike, 1974) is also a measure of the goodness of fit of a mathematical model but on the contrary to the Adj. R-square, it is not grounded in the statistical theory of hypothesis testing but rather in the information theory. The AIC estimates the Kullback–Leibler information loss by approximating the observed data with the fitted model (details regarding model selection using information theory can be found in Burnham and Anderson, 2002). The fit of any regression model to any dataset can be summarized by the Akaike Information Criterion (AIC) defined by the equation:

$$\text{AIC} = \text{N} \cdot \ln\left(\frac{\text{residual SS}}{\text{N}}\right) + 2\text{K} \qquad (11)$$

where N is the number of data points and K is the number of parameters fit by the regression plus one. The definition of K as the number of parameters plus one is justified by the fact that the regression is "estimating" not only the values of the parameters but also the sum of squares. It is worth noting that the computation equation of the AIC consists of two additive terms, namely one term representing the lack of model fit to the data and another term related to the number of parameters; hence, the AIC can be seen as a measure of both the accuracy and the complexity of the chosen model. In cases where N/K < 40 as in this study, a second order corrected AIC, hereafter referred to as AIC_c, is used:

$$\text{AIC}_c = \text{AIC} + \frac{2\text{K}(\text{K}+1)}{\text{N-K-1}} \qquad (12)$$

When comparing several mathematical models, it is the one with the lowest AIC_c that is the best or that is most likely to be correct. Hence, in this study, mathematical models with an Adj. R-square exceeding 0.3 were ranked by sorting their associated AIC_c scores in ascending order, and the top-ranked model chosen as the best one.

The third and last step in the best model selection process consisted in confirming the choice made at the end of step 2. Indeed, if the AIC_c scores between the top two-ranked models are very close, there is not much evidence to choose one model over the other. We therefore used the following equation to compute the probability that the top ranked model is indeed the best one:

$$\text{probability} = \frac{e^{-0.5(\text{AIC}_{c_2} - \text{AIC}_{c_1})}}{1+e^{-0.5(\text{AIC}_{c_2} - \text{AIC}_{c_1})}} \qquad (13)$$

This probability can thus be seen as an uncertainty measure as it expresses the likelihood that the top-ranked model is the best among the set of models being evaluated.

The possible influence of catchment topography, both surface and subsurface, was studied with regards not only to the nature (e.g. linear versus quadratic, versus cubic, etc.) but also to the strength of the point-scale relationships between actual soil moisture and surrogate measures. Nonparametric Kruskal-Wallis tests were run to assess whether the different types of point-scale relationships were spatially associated with specific topographic properties. The Kruskal-Wallis test is identical to one-way analysis of variance except that the data are replaced by their ranks. Hence, it is used to compare samples from three or more groups. The null hypothesis states that all group medians are equal, while the alternative hypothesis states that at least one group median is different from the others. In this study, each mathematical model is a group and we compare the topography underlying the locations subjected to different relationships between point-scale relationships between actual soil moisture and surrogate measures. When the p-value associated with the statistical test is less than 0.05, we reject the null hypothesis and suggest that the differences in relationship types can be explained by topography. Spearman correlation coefficients were also computed between the strength of the point-scale relationships (i.e. R-square values) and the values of the terrain attributes.