Response to review of: "McMillan, H.K. and Srinivasan, M.S., Controls and characteristics of variability in soil moisture and groundwater in a headwater catchment" by Reviewer 2.

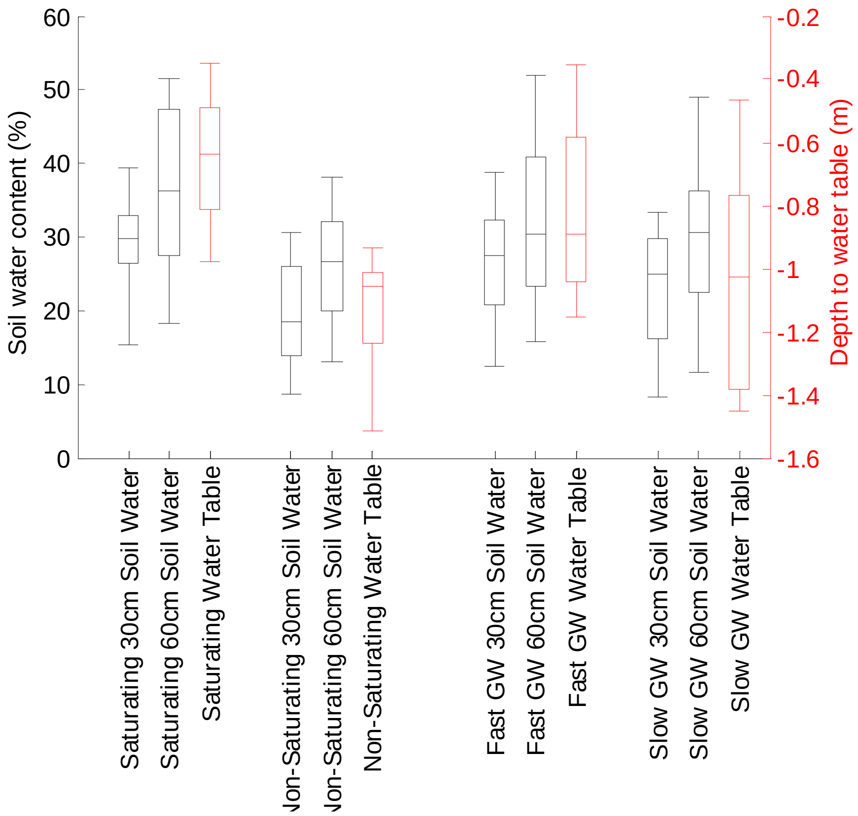
We very much appreciate the positive and constructive reviews by both referees, and their helpful suggestions on improvement of our manuscript. In our response below, we address individually each comment from Reviewer 2. For each comment we 1. Quote the comment (black font), 2. Give our response (blue font), 3. Showed the modified text (blue font, italic).

General comments

This paper addresses the soil moisture and shallow groundwater table response in an experimental catchment in New Zealand. Both soil moisture and groundwater head are sampled at a variety of topographic locations that vary in terms of proximity to stream and aspect. Aspect is found to be a significant driver of difference. Data are analysed both spatially and temporally. Soil moisture and groundwater responses are nominally considered together which is a worthwhile objective as such joint analysis has rarely been done. The paper is generally clearly written and most information required is available, although I request a series of clarifications in the specific comments below.
Thank you for this positive summary of our manuscript.

I first have a couple of overall comments to make.
I thought that the paper could be strengthened by analysing the soil moisture responses in the context of the groundwater observations, and perhaps vice versa. At present analyses are carried out by geographic grouping of sites (i.e. lower/upper slope etc) but there is substantial variation in response within each group, showing significant catchment heterogeneity. I feel that it would be informative to group the sites by water table response (shallow/deep, quick/slow) and analyse the soil moisture in the context of those groupings. This might enable more insight into the inter-relationships between soil moisture and groundwater responses, which would make the paper significantly stronger in my view.
Thank you for this suggestion. We made both the analyses suggested, grouping sites by both shallow/deep and quick/slow water table responses. The results are given in a new figure that shows in particular a strong differential between the shallow/deep classifications. In response, we added an additional paragraph to Section 4.4.2, and refer to the result in the conclusion and abstract.

"Our results suggest that relative groundwater levels, and the classification of sites as fast or slow groundwater responses, are consistent between events. Previous work reviewed in the introduction (Section 1.3) showed that groundwater level can influence soil moisture distribution. We therefore hypothesise that groundwater behaviour might help to define distinct spatial zones of the catchment. To test this, we firstly classified sites by maximum groundwater level, separating sites where the water table rose as high as the 30 cm soil moisture probe at any point during the study period ('Saturating sites'), against those where it did not ('Non-Saturating sites'). We only used near-stream sites to remove the influence of distance to stream. Secondly we classified sites by speed of groundwater response, as described in the previous paragraph. Other sites where groundwater rarely responds were not included as only the peaks of groundwater responses are measured, and therefore these sites could not be easily classified. We calculated the distributions of the soil moisture and water table level for each classification (Figure 10). The results show that the Saturating vs. Non-Saturating classification clearly delineates two zones with consistent differences in soil moisture content at 30 cm and 60 cm, and water table level. The fast vs. slow groundwater response classification is much less distinct, with the two zones having similar soil moisture distributions. The slow groundwater response zone has slightly deeper water tables, although this is partly because it includes two far-stream sites."



Additional Figure 10: Summary statistics of soil moisture and groundwater values, classified as Saturating/Non-saturating sites, and Fast/Slow groundwater responses sites. Saturating sites were defined as those where the water table rose as high as the 30 cm soil moisture probe at any point during the study period. Fast/Slow sites were classified according to the speed of groundwater response as described in Section 4.4.2 and Figure 9C.

I think it would also be worth the authors considering the definition of soil moisture and groundwater briefly in the paper. They are terms we use easily but often people are referring to different things. Soil moisture can be near surface, root zone, profile, at a specific depth, etc. Groundwater can be shallow, perched, regional, etc. Perched water tables in the soil profile might be considered either groundwater or soil water and no doubt there is grey in between. Distinguishing carefully might be useful in interpreting the literature a little more.
We added a paragraph explaining these definitions in our new section on Soil moisture – groundwater interactions and variability (see later). The new text is as follows:

"The dividing line between stored water that is considered as soil moisture or groundwater is not well defined. Soil moisture is typically measured as volumetric water content at specific depths in the unsaturated zone, although soil moisture sensors can be subsumed by perched or deeper water tables. Here, we use groundwater synonymously with water table, referring to saturated subsurface layers, which may be above or below any soil/bedrock interface. Piezometers or shallow wells to measure groundwater level can be screened along their whole length (as in our study) or at specific depths if multiple perched or confined layers are suspected. Where the geology includes fractured rock or buried lenses of gravels, groundwater levels may be highly heterogeneous."

The analysis (e.g. in Figure 3) often relies on grouping sites and averaging which is fine but consider showing the spread between sites somehow – it is not clear how "different" these groupings actually are without knowing the within-group variability. The same can be said for other figures.
We used the inter-site groupings to make the figures more readable, and to enable the reader to interpret the effects of different controls, despite the high variation observed. In response to this suggestion, we tested the option to include ranges in Figures 4&5 showing saturation behaviour (example below in Figure Part B). However we felt that showing the ranges made the figures slower to

interpret, and therefore we did not keep this change in the paper. We added a sentence into the paper to explain this in Section 4.2:

"Note that the statistics describing the extremes of the data are highly variable between locations (e.g. some locations are saturated much of the time; others almost never), however we show averages by location to assist interpretation of the spatial control."

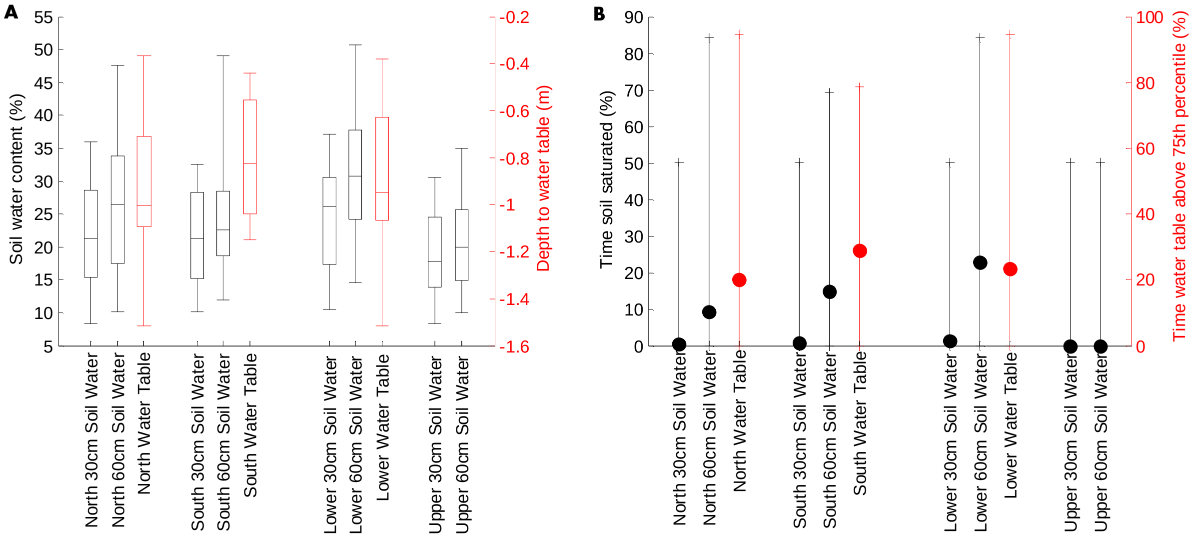


Figure 1: Demonstration of the effect of including ranges into the Figure showing spatial controls on soil moisture and water table extremes

In the discussion the reasons for the observation that aspect is a strong differentiator of behaviour could be discussed in more depth. Is this a direct effect through ET or is it due to impacts on soil development, issues such as geological bedding (strike-dip effects), vegetation differences, etc.
We added an extra paragraph into the discussion to elaborate on the aspect differences:
"Our finding that aspect is an important control on soil moisture echoes the results of previous studies in NZ hill country (e.g. Bretherton et al., 2010; Lambert and Roberts, 1976). The mechanisms linking aspect with soil moisture are varied. For example, Lambert and Roberts (1976) found complex interactions between air temperature, soil temperature and ET, driven by wind direction and aspect-induced radiation differences. They note that the specific heat capacity of soil drops as it dries, leading to a positive feedback cycle. In the Langs Gully catchment, the South facing slopes are also steeper than the North facing slopes. This is not obviously due to geological bedding – the main trend of syncline-anticline pairs in the wider Waipara catchment is Northwest-Southeast (transverse to catchment slopes), and in the immediate area of Langs Gully, known dip directions are highly variable. However, feedbacks are likely to exist between slope angle, vegetation (denser shrub cover on South-facing slopes), soil depth (thinner on South-facing slopes) and downslope sediment transport. Shading by denser vegetation and increased lateral flow are possible causes of the increased number of wetting events on the South-facing slope. Typical hydrological models do not account for aspect, but our results suggest that this is an important factor to consider in hillslope runoff generation."

The figures are generally clear but quite a few have lines distinguished by colour but lack a legend. Also the captions could probably give a little more detail to make the figures a little more stand alone.
We added additional legends and improved figure captions as needed

Specific comments
9477:24-9478:2. Variation in soil properties is important for similar reasons. It is worth mentioning that it is not just the water status that needs to be thought about.
We added the following comment:
"Similarly, averaging of soil texture or water-holding properties should take spatial organisation into account."

Section 1.1. The authors would find the following paper interesting for this section (and it does a little comparison of soil water-groundwater interplay too).

Rosenbaum, U., Bogena, H.R., Herbs, M., Huisman, J.A., Peterson, T.J., Weuthen, A., Western, A.W., Vereecken, H., 2012. Seasonal and event dynamics of spatial soil moisture patterns at the small catchment scale. Water Resour. Res., 48: W10544.
Thanks for the suggestion. We added the results from the paper into this section on causes of high soil moisture variability, and also added it into the summary section:

'Rosenbaum et al. (2012) similarly found that seasonal differences between groundwater influenced and groundwater distant locations had a strong effect on soil moisture standard deviation.'

9482:4. I was expecting a section on soil moisture-groundwater inter-relationships here given the promise of the abstract.
Thanks for this suggestion. We added a new section on Soil moisture – groundwater interactions and variability at the end of the introduction section, as follows:

"There are many processes by which soil moisture and groundwater interact. As soil water drains downwards, layers of low hydraulic conductivity may create perched water tables. Such layers could include clay pans, or the soil/bedrock interface (Tromp-van Meerveld and McDonnell, 2006a). Macropores provide a fast route for surface and soil water to recharge groundwater (Beven and Germann, 2013). They may allow water to bypass confining layers (e.g. see the discussion of perceptual models in McMillan et al., 2011) or to flow quickly along them (e.g. lateral preferential flow along the bedrock interface found by Graham et al., 2010). If groundwater rises into upper soil layers, large increases in soil matrix porosity or macropores may 'cap' water table levels, as additional water is quickly transported to the stream (Haught and Meerveld, 2011). Lana-Renault et al. (2014) found in a Mediterranean catchment that patterns of near-surface saturation and transient water tables were affected not only by topographic controls but also soil properties and previous agricultural land use. The riparian zone facilitates mixing between soil water and groundwater, and tracers, temperature, electrical conductivity, flow gauging and head differences may all be used to quantify the interactions (Unland et al., 2013). Using modelling and tracer data, Binley et al. (2013) found that in a 200 m river reach the upper section was connected to regional groundwater, but lower section inflows were from local lateral and down-river flow paths.

Interactions between soil moisture and groundwater provide possible explanations for relationships between the two. Results from three Nordic catchments showed a consistent negative correlation between soil moisture content and depth to water table, so that soil moisture distributions could be described as a function of depth to water table (Beldring et al., 1999). Kaplan and Munoz-Carpena (2011) studied soil moisture regime in a coastal floodplain forest in Florida, finding that groundwater and standing surface water elevations were successful predictors of soil moisture using dynamic factor analysis and regression models. Model-based studies demonstrate how capillary-rise can lead to dependencies between groundwater level and soil moisture. Kim et al. (1999) used a hillslope model to show how gravity-driven downhill groundwater flow creates downslope zones with high water tables. In those areas, capillary rise keeps soil moisture content and evaporation rates high. Similarly, the model developed by Chen and Hu (2004) showed that soil moisture in the upper 1 m of soil was 21% higher when exchange between soil moisture and groundwater was included. They inferred that groundwater variability may drive soil moisture variability."

9482:13. Is there any orographic effect over the catchment (given the 250m relief)?
We do not have good information on this, as there is only one raingauge in the catchment. The nearest other raingauge is at the Melrose farmhouse, 2km from the Langs Gully gauge, and at an altitude of 460 m compared to 580 m at Langs Gully. During two years of coincident record, in one year Melrose records more rainfall, in the other year Langs Gully records more rainfall. On average, Langs Gully receives 8% more rainfall. As these results are not conclusive, we did not add any information to the paper.

9482:23-27. Are these sand/clay fractions specified including or excluding the coarse fraction?
Excluding – now noted in the caption to Table 2.

9484:5. How were soil moisture sensor calibrations obtained?

We used factory calibration settings, as the Acclima TDT sensors we used are recommended for use in loam soils without calibration (http://acclima.com/wd/acclimadocs/agriculture/SDI-12%20TDT%20Sensor%20Data%20Sheet.pdf). This is now noted in the paper.

9484:8. Were the groundwater wells open across their whole length? If not what was the screen interval?
Yes they were open across the whole length. This is now noted in the paper.

9486:1-17. I was unclear about these calculations. The monitoring only covers a very small part of the catchment. Are you implying that you extrapolated to the whole catchment (that seems courageous given the extent of the extrapolation and the fact that monitoring seems concentrated on lower slopes) or were you averaging over the monitored area. I would suggest concentrating on the monitored area.
The reviewer is correct that we estimated the total water storage in the catchment by extrapolating from the 32 soil moisture sensors to the whole catchment based on soil type. We included this estimate based on a request from the topical editor "*it would be useful to quantify the total storage in soil moisture and groundwater to see what variability has the largest effect on total variability in storage*". We agree that this is a significant and uncertain extrapolation, although perhaps not unusual in hydrological studies. The bias towards lower slopes will be in part corrected for as we considered average soil depth for each soil type. We also note the usefulness and explanatory power of the estimate of stored water as an antecedent soil moisture index, as shown in the new Figure 5B. We did not change the calculation, but added some additional commentary as follows:
"We recognise that this calculation involves a significant and uncertain extrapolation from the 32 soil moisture time series to the remainder of the catchment. However, given that the sensor locations sampled across aspect, distance from stream, and landscape position and depth, we anticipate that the estimated storage dynamics are a reasonable guide to true behaviour."

Figure 3. Red and magenta are hard to distinguish.
We replaced pink with orange to improve the contrast

9487:19. It was mentioned in the methods that you determined saturation thresholds for soil moisture sensors by visually examining the traces. You could and should use the groundwater data to confirm this given that a strength of your study is the co-located measurement of both.
This is a good point. We took the average soil moisture reading at the point where the water table reached 30 cm/60 cm respectively and used this as the saturation threshold. We compared these with the thresholds that we obtained visually. In most cases agreement was good and we used the water table value; in two cases there was discrepancy and we used an intermediate value: this is probably due to uncertainties in the GPS elevation readings used to determine the relative positions of soil moisture sensors and groundwater wells when they were installed on uneven/sloping ground. We noted the change as follows in the text:
"Soil saturation points were defined individually for each sensor, using the co-located groundwater well record to determine times when the water table intersected the sensor, and taking the average soil moisture reading at those times. These values were confirmed (and in two cases adjusted) based on visual inspection of the soil moisture time series."
The changes made small differences to graphs 5 & 8 and we replotted them accordingly.

9487:26. Figure 6 – a common behaviour (but not that often reported). It would be nice to see this plotted following Detty and McGuire 201a, Fig 4 with antecedent moisture incorporated.
As per the reviewer's suggestion, we incorporated an antecedent soil moisture index into the x axis of this rainfall against runoff figure. The soil moisture index was taken as the Total Soil Moisture value from Eq 1. The results show a nice relationship (although not linear as per Detty and McGuire) and were added into the figure as a second subplot. Extra text was also added to the paper:
"To demonstrate the effect of antecedent wetness on storm runoff depths, we plotted runoff depth against the sum of antecedent soil moisture storage (ASM) and storm precipitation (Figure 6b), following Detty and McGuire (2010b; their figure 4a). Antecedent soil moisture storage was taken as the Total Soil Moisture value from Eq 1. The results show a threshold relationship between ASM + precipitation and runoff depth, although it is not linear as was found by Detty and McGuire (2010b)."

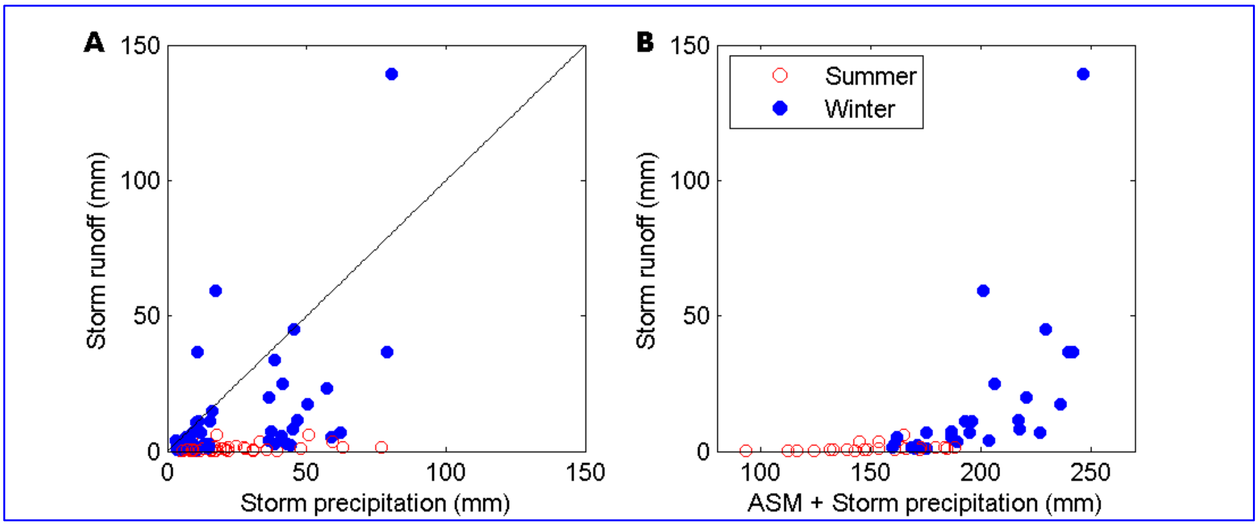


Figure 5: (A) Storm Runoff against Storm Precipitation, split by season. This figure was created after pre-processing of the data to define storm and inter-storm periods, based on the method of McMillan et al (2014) using thresholds for precipitation depth and inter-storm duration, and without baseflow separation. (B) Storm Runoff against the sum of Storm Precipitation and Antecedent Soil Moisture storage (ASM), split by season. ASM was taken as the Total Soil Moisture value from Eq 1.

9488:7. Not clear if the moisture is averaged across sites before counting the events – please clarify. Figures 7 and 8 don't have upper water table – why?
We calculated events on a per-site basis, and then averaged across sites afterwards. We clarified that as follows:
"We calculated events on a per-site basis, and then averaged across sites"
Regarding the upper water tables, we added an explanation:
"We did not include water table statistics for the upper rows as water tables only rarely occurred during event peaks and therefore distribution estimates would not be accurate."

9488:23+. I think the different datums are used for soil water and groundwater (lowest observed level for groundwater vs zero moisture for soil moisture). How does this affect comparison of absolute values?
We added an explanatory note:
"The difference may be further enhanced given that the part of the soil moisture volume below wilting point is not likely to be mobilised."

9489:4. Which standard deviation – I realised latter that it was spatial but clarify here. Clarify in figure caption too.
Done

9489:7-8. There is quite a lot of literature that shows variability decreases with drying, so I disagree with this statement. See Rosenbaum (2012) referenced above.
The reviewer is correct and we replaced the sentence:
"Previous studies have shown that the relationship between soil moisture and soil moisture standard deviation varies by catchment (Section 1.1)."

9489:12-14. What does this imply about preferential flow?
This effect is probably not related to preferential flow, as instead it is caused by rising water tables that envelop some soil moisture sensors but not others.

9489:14-15. This sentence doesn't seem to follow on from the previous "Accordingly...." – I could not see the relationship.
We meant that rises in spatial standard deviation are caused by saturation of some sensors but not others. But, on the N-facing slope at 30 cm, none of the sensors typically saturate, so standard

deviation does not rise. We clarified "*Accordingly, 30 cm North facing soil moisture has smaller rises in spatial standard deviation, as none of those sensors typically saturate*."

9489:24-25. Yes you saw significant spatial variability in a small part of the catchment – and presumably there is in other parts of the catchment – but you can't actually say how large or small that might be from your sampling scheme.
We updated the text to emphasise that we only know about variability in locations where we have sensors:
"there is significant spatial variability between different parts of the catchment as represented by the range of sensor locations"

9490. When discussing particular events, please give the rainfall depths.
We added rainfall depths to the event descriptions.

9491:20-22. Maybe I am misinterpreting comparison of Fig. 10 and 11 but only about half the sites showing saturation had a WT response in summer.
We reworded this comparison more carefully to make sure it was clear and correct:
"3 out of 4 locations where saturation occurred at the 60 cm probes in this event were locations that showed a water table response during the summer event described in the previous section. All locations that had a water table response in the summer event also has a water table response during this event."

9492:4-5. This observation about gravel should be in the site description.
We added the information to the site description:
"During installation of soil moisture sensors (Section 3.2), at 6 out of 16 locations there were found to be distinct gravel-rich layers within the soil profile."

9493:20. It would be good to see these recessions together with the deeper sensor and groundwater.
We tried adding an addition panel to the figure with the deeper sensor responses (groundwater was not used as no response occurred in many of the locations). The figure is shown below: it shows that the shape of the lower sensor typically mirrors that of the upper sensor (apart from one case where the sensor is saturated), but the shape is perhaps less well defined. We didn't feel it added sufficiently to our results to include in the paper. Instead we added an extra sentence explaining as follows:
"We also found that the shape (i.e. convex or concave) of the corresponding 60 cm soil moisture response was typically the same as the 30 cm sensor (not shown)."

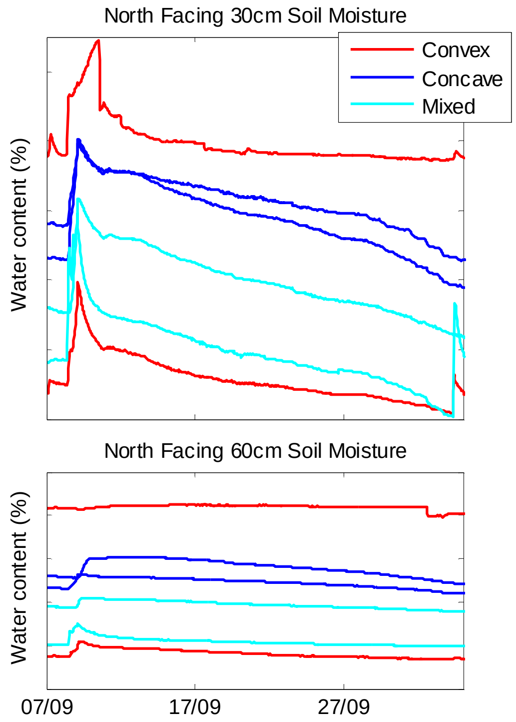


Figure X: Demonstration of adding extra panel showing responses of deep soil moisture sensors

Technical corrections

9476:18-20. The wording seems awkward.

The paragraph was re-structured as follows:

“Co-measurement of soil moisture and water table level allowed us to identify interrelationships between the two. Locations where water tables peaked closest to the surface had consistently wetter soils and higher water tables. These wetter sites were the same across seasons. However, temporary patterns of strong soil moisture response to summer storms did not correspond to the wetter sites.”