Interactive comment on “The Normalized Difference Infrared Index (NDII) as a proxy for soil moisture storage in hydrological modelling” by N. Sriwongsitanon et al.

M. van Tiel
marit.vantiel@wur.nl

Received and published: 26 October 2015

Note for the authors and editor

The following review was written by a student of the MSc programme Earth and Environment at Wageningen University. As part of the course Integrated Topics in Earth and Environment, students are asked to prepare a review of a scientific paper. The supervisor of this review process is Ryan Teuling. The manuscript by Sriwongsitanon et al. was one of the manuscripts that was selected for this exercise. The review is written as an official review in order to comply with the course guidelines, but it should
be considered by the authors as a regular comment which they can use to improve the manuscript. I hope that this comment will positively contribute to the review process and that it will help the authors to improve their manuscript for possible publication in HESS.

1 Introduction

1.1 Scope

This paper tries to relate a remote sensing product, the NDII, to the root zone soil moisture storage. The title sounds promising and also the use of remote sensing to derive root zone soil moisture is interesting. Remote sensing is a well-known technique but has more potential than where it is used for up to now and the application of it to water resources has been increasing in recent years (Melesse et al., 2007). Remote sensing can provide us with information on the temporal and spatial distribution of a certain variable which is difficult to achieve with observations. The variable addressed here is the root zone soil moisture. Soil moisture is an important component in the hydrological cycle and it determines the interaction between land surface and the atmosphere (Mahmood & Hubbard, 2007, Wang et al., 2007). Climatic and hydrologic modelling and prediction need soil moisture at a high spatial and temporal resolution to improve their performance (Wang et al., 2007). Most research has been focused on the soil moisture storage in the top layer (e.g. Baroni et al., 2013, Fensholt & Sandholt, 2003). However, the part that is especially interesting is the root zone part of the soil moisture, because it links the surface phenology and the subsurface water storages and due to evapotranspiration it influences the surface water balance and the energy fluxes (Wang et al., 2007). Next to that, it also controls the density and condition of the vegetation, which is used in this paper.
1.2 Summary

The NDII measures the shortwave infrared reflectance and is therefore related to the Equivalent Water Thickness (EWT) of the leaves. The expectation of the authors is that the NDII will only show a strong link to soil moisture in the root zone when there is a soil moisture deficit, since the NDII is an indicator of water stress. To test this hypothesis and to see if the NDII can be related to root zone soil moisture, a lumped conceptual hydrological rainfall-runoff model is used, FLEX\textsuperscript{L}. The NDII was compared to the soil moisture model output. The results show that for monthly averaged values the NDII values are higher during the wet season and lower during the dry season. When the 8 day averaged values are compared, the relationship shows a clear exponential function during moisture stress. It is concluded that the NDII and the soil moisture are highly correlated during the dry season and less during the wet season. The authors found almost no time lag between soil moisture storage and NDII and conclude that the vegetation responds fast to soil moisture variation. The findings of the study can be used in hydrological models and would according to the authors be extremely useful in predicting discharge in ungauged basins.

1.3 Recommendation

I like the integrated approach used by the authors where vegetation is used to derive soil moisture instead of omitted, because the Earth is an integrated system and should be treated according to that. The paper fits in HESS because of the interaction between a hydrological variable and vegetation investigated with remote sensing technique and it aims to improve hydrological modelling. However, I have some reservations about the use of the model, the correlation between NDII and $S_u$, the lack of detail and the structure and outline of the paper, which I will explain more in depth below. I therefore think that major revisions are necessary before the manuscript can be considered for
2 General comments

2.1 Use of model vs. observations

In this paper the NDII is compared to the soil moisture storage content ($S_u$) from the FLEX$^L$ model. My main point here is why the authors did not use observations of root zone soil moisture? A model is an approach of reality, but it depends on the quality of the model how close the model represents reality. Gao et al. (2014) show that the FLEX$^L$ model did not perform well in simulating low flows in two sub-catchments in China, which is important in this study since periods of low flow are the periods of interest. Using a distributed model instead of a lumped model sounds better when modelling a dynamical process like root zone soil moisture. However, Gao et al. (2014) showed that the distributed version of the FLEX model did not perform better, because of the inappropriate soil moisture distribution by the Thiessen polygons. How is then guaranteed that the model will give reasonable values for root zone soil moisture? I think that using observations to show how the NDII correlates with root zone soil moisture would be more convincing. I therefore suggest to test the model first with observations in the study area, or even better, test the NDII with observations. At least the performance of the model should be discussed.

I see a confusing circular reasoning in the title of the paper, apparently root zone soil moisture is not well presented in hydrological modes and the paper wants to fill in this gap by using NDII. However, the usefulness of NDII is tested with a hydrological model (which does not present root zone soil moisture well)! By stating this the authors are actually not convinced by their own methods of using a model. The authors tone down their title statement a bit by mentioning once that the NDII could be used for calibrating...
models in ungauged basins. However, also in that case you want to make sure that the values used for calibration represent reality. An example showing that model root zone soil moisture is still a major challenge is the study of Das & Mohenty (2006) where a vadose flow model did not model the root zone soil moisture perfectly. Also this point is therefore solved by using observations. Though, if the authors assume that modelling root zone soil moisture is possible with their model, the use of NDII for calibration in ungauged basins should be made more clear in the beginning, starting from the title.

Another issue of only using a model is that the water stress periods, where NDII performs best, can be virtual water stress periods. It is not sure that with low values for logarithmic Kling-Gupta efficiencies (but still reasonable), the moisture stress periods are modelled well. Also here a validation with observations would help.

In the end of the paper, the authors conclude that there is almost no time lag between NDII and $S_u$ and that vegetation shows a fast response to soil moisture changes. However, this cannot be concluded because the evaporation is not modelled in a dynamical way in the FLEX\textsuperscript{L} model. The changing vegetation is not taken into account for the evapotranspiration, so the vegetation is assumed to be constant. Since evapotranspiration is related to soil moisture, a static value ($S_u$) and a dynamical (NDII) value are compared. This does not give relevant information and is crucial for this research where vegetation is used instead of omitted. This means that a higher NDII because of a higher density of plants will cause a lower $S_u$ than reality and the other way around for a lower NDII. If the dynamical vegetation is taken into account it will probably improve the correlation. As suggested in the paper itself, NDVI could be used to model the actual transpiration of an ecosystem. Next to that, Wang et al. (2007) concludes that there is a time lag of 10 days and a lag of 5 days for humid and semi-arid sites, respectively, of vegetation to respond to soil moisture changes. Therefore I think the conclusion about no time-lag is not grounded and the analysis should be done again, taking a dynamical vegetation into account.
2.2 Seasonal correlation of NDII and $S_u$

My second major point is the way in which NDII and $S_u$ are correlated. In my view everyone knows that in the wet season the soil has a higher soil moisture content than in the dry season (Wang et al., 2007). A seasonal correlation does not add new information. It is general known that seasonal correlations always give high R² values, since a lot is correlated with that (Schaefli & Gupta, 2007). The use of umbrellas would in that case also be a reasonable indicator of soil moisture. What would be of interest is to know how root zone soil moisture behaves over different years. If different August months are compared with each other, would the NDII be able to detect a wetter and drier August month? That are the dynamics we are interested in. Therefore I suggest to get the seasonality out of the dataset, the same as Wang et al. (2007) did. If the deseasonalized dataset shows no correlation between $S_u$ and NDII, the NDII would not be a better indicator of soil moisture than just the day of the year.

2.3 Calibration, validation and correlation

The Kling-Gupta efficiency for flow duration curve showed the best result, the efficiency for low flow the worst but still reasonable result. Since the whole research is based on periods where soil moisture is limited it is better to achieve the best efficiency on low flows. I suggest to define a different parameter set that has the highest Kling-Gupta efficiency for low flows.

Next to that, the model is not validated, only calibrated. The sub-basins wherefore the model is calibrated with runoff data from the sub-basin are used to determine the correlation between NDII and $S_u$. No validation is needed here probably, because the model is only used to derive values for the root zone soil moisture. However, it should be stated more clear in the paper that NDII is compared to $S_u$ only for the calibration periods. The authors should also mention that this has some consequences for the
correlation in the different basins. Some dry or wet years are not taken into account for specific sub-basins because of a shorter calibration period. It would be better to have the same comparison period for all the basins.

Furthermore, the authors claim that this NDII can be used in the future especially for predicting discharge in ungauged basins. I think it is a missed change that the authors did not show this. It is not clear what the authors mean with predicting discharge in ungauged basins. Will it be used for calibration or not? If NDII will be used for calibration, the authors could have use the NDII to calibrate the model for one of the runoff stations and see if this gives reasonable results compared with the runoff data. If the NDII is not used for calibration in ungauged basins but just as input data, they could have test the model with NDII input for one of the other non-calibrated six basins and see if the results are reasonable. Or use the NDII as input for the model without calibrating in one of the basins with a runoff station and compare the outcome with observations. However, Gao et al. (2014), shows that the FLEX model did not perform well without calibrating. This adds questions again for the use of the model, especially by saying that it can be used for ungauged basins, without saying that NDII as proxy for soil moisture is used for calibration, so that a model needs to do all the work.

Besides that, the graphs showing the correlation do not all allow to plot this exponential function. Especially in the wet season, low values for root zone soil moisture correspond with several values of the NDII. This shows that soil moisture is insensitive to the NDII. The second part of the wet season plots, for high NDII, show a big scatter. If the NDII is high, soil moisture could be any value. Fitting these exponential functions is therefore not appropriate here. My suggestion is to give the uncertainty in soil moisture prediction for a certain NDII value. From this it would immediately be clear that the annual and wet season correlations are not well correlated and could not be used to derive the root zone soil moisture content. This would also change the conclusion that NDII could be used to derive dry conditions in wet seasons since this is absolutely not clear from the graphs.
The saturation effect on the NDII is discussed but could be elaborated more. The graphs in figure 6 should be used to indicate where saturation occurs. It is clearly visible that in the wet season the same NDII could have different values of root zone soil moisture, indicating that increasing soil moisture does not influence the NDII anymore. This part should be omitted from the dataset to make a meaningful function for lower soil moisture values. The way it is presented now, shows that an exponential function is used to glue two different functions, for high and low soil moisture, together. Also the study of Yilmaz et al. (2008) shows that the relationship between NDII and VWC changes under water stress conditions. Since VWC is related to the root zone soil moisture, it should also be taken into account here. The changing relationship cannot be presented by one correlation line in the graph.

What is also remarkable is that the values of NDII and $S_u$ on the axes do not differ that much for the annual, dry and wet seasons. If one looks at the three correlation graphs for one basin, the maximum soil moisture is somewhat higher for the wet season, but in general, the soil moisture conditions that occur in the dry season do also occur in the wet season. The division in dry and wet season is based on rainfall amounts but it is strange this this is not reflected in soil moisture storage values. A value of $S_u$ from 200 mm shows a much larger spread in NDII for the wet season than the dry season. This is also true for lower values and I do not think that saturation already plays a role here. I expect a more thorough discussion here between the dry and wet season and the possible reasons for the differences.

2.4 Lack of detail

Another major point is the lack of detail in the whole paper. First, the choice for the model is not explained and also the reliability of the model to simulate root zone soil moisture is not discussed. This is quite essential for thrusting the results of the paper and is not general knowledge since the model cannot be found easily on the internet.
Second, the choice for the study area is not explained. As I already mentioned I do not understand why no observations are used. If the authors choose an area without observations, explain at least why this area was chosen. I get the impression that this area is chosen because of its clear seasonality with wet and dry seasons. This would improve the R2 values, but this is no good reason to choose a study area. What is also missing in the area description is the type of vegetation. The type of vegetation determines the root zone moisture uptake and can therefore result in spatially different root zone soil moisture values (Jayawickreme et al., 2008). It would be interesting to see if NDII is able to see these differences.

Another example is the relationship between soil moisture and vegetation, which is very important in this research, but not explained. In the results one sentence is included about the suction pressure, which relates the vegetation water content to the soil moisture, but this should be made really clear in the introduction. What also misses is the derivation from canopy water content to vegetation water content. An allometric equation is needed to convert the canopy water content to the vegetation water content which includes also the stems of the vegetation. I suggest to describe this equation, or if this was not used, not to use vegetation water content but canopy water content instead. Next point is the use of NDII instead of NDVI. Wang et al. (2007) already studied the connection between NDVI and soil moisture. It is not clear what extra information the NDII could give or why the NDII is used instead of the NDVI. I suggest to explain this or compare the results with Wang et al. (2007) to see if NDII is a better indicator.

Furthermore the hypothesis is not clearly formulated and this makes it difficult to understand what the objectives and aim of this paper are. The aim indicated in the paper is to prove the effectiveness of the NDII as indicator for root zone soil moisture. As I said earlier this should not be proven only by showing seasonal correlation. A clearly formulated hypothesis would overcome this problem: Is NDII a better explanatory variable than the day of the year for explaining root zone soil moisture storage. However also this word ‘storage’ is questionable since the research is rather about the variability
in root zone soil moisture. It is not clear if the authors are interested in the variability of root zone soil moisture over the year and spatially or how much the storage is itself. If this is stated both clear in the introduction it will help improving the clarity.

Next to that it is also not clear on which timescales we are looking when explaining the relation between $S_u$ and NDII. On the short timescales, an increase in soil moisture causes an increase in evapotranspiration. This causes the EWT to go up and to measure a high NDII. At the same time the soil moisture will go down, because it is used by the roots. In this case a high NDII corresponds to a low soil moisture. This effect is nowhere explained and should be mentioned. A time lag of 10 to 5 days from Wang et al. (2007) show that this effect could be visible in the 8 day averaged NDII and $S_u$ values, if the vegetation would be modelled dynamically.

What is also missing is the explanation of landscape characteristics. These are shown in the figure 2 and mentioned in the conclusion, but is not taken into account in the analysis. Soil type, vegetation and topography determine the spatial distribution of soil moisture (Baroni et al., 2013). The occurrence of bare soil between vegetation can influence the NDII (Dasgupta & Qu, 2008). All these variables are interesting to compare with the spatial distributed maps of the NDII. It can then be discussed if the NDII is able to detect the spatial distribution according to the landscape characteristics.

2.5 Structure and outline

In several places, important information is missing. Also the outline of the paper is not clear at all. Results are not clearly results but also show interpretation. The discussion does not contain the shortages of the study and the use of results. I suggest to explain in the discussion if the NDII only works in the growing season (Yilmaz et al., 2008) and what other papers found about saturation. Next to that it is also important to discuss what would happen if the NDII detects bare soil and if the whole catchment or sub-basin could be used to derive an averaged NDII value. Can the results of this paper
be of use in the tau-omega model (Yilmaz et al., 2008)? Also a more elaborate study on the future where hypothesis for other climates and regions are given is needed and perspectives should be given on how the results could be implemented in hydrological model and what it would improve.

The introduction does not contain all the relevant information, mentioned above in lack of detail. The conclusion is more a summary of the research and does not give an in depth answer on the objectives and hypothesis (which are missing or vague formulated). Furthermore, the definitions of root zone soil moisture is changing in the paper which is confusing. Sometimes just soil moisture is used, then soil moisture storage or soil moisture deficit. Also the reservoir representing the root zone soil moisture in the model is confusing. The name of this reservoir is unsaturated zone, while the root zone is only part of this unsaturated zone (see figure 1). From this it is not clear if the dynamical part is explicitly modelled. I think it is only a definition issue, but it should be made clear what is modelled and where the NDII is assumed to show a correlation with. If the results would be compared to observations it should be clear what part of the soil is modelled. I suggest to rename the reservoir to root zone reservoir.

Lastly the title is not well chosen. The proxy for hydrological modelling is only discussed in the discussion part of the paper and is therefore not the main content of the paper. The paper does actually show that the NDII can be used as a drought indicator for soil moisture droughts, because it only reflects the root zone soil moisture content under water stress conditions. The amount of root zone soil moisture is not determined, whereas the variability over the seasons and over spatial scales is. Also, the hydrological model part in the title sounds very promising, but is not because of the use of models as methods in the paper, as explained by my first major point. The title should better reflect the actual contents of the research. An example is naming the study area in the title.
3 Specific comments

Page 8419 L6: deficit is not monitored in the study, the occurrence of deficit is monitored, there is no threshold defined that could be used to quantify the deficit.

Page 8419 L8: Northern Thailand.

Page 8419 L14: if NDII proves to be a very strong proxy for moisture storage ‘deficit’ in the root zone, why does the title suggest something else? Title should be based on this information.

Page 8419 L15: a crucial component, the soil moisture is, but the moisture storage deficit in the root zone is more crucial for the vegetation.

Page 8419 L15: in addition,…, if the NDII is an indicator for storage deficit it is actually an indicator of soil moisture drought, so what is the addition?

Page 8419 L21-22: repetition, is already stated in Lines 14-16.

Page 8419 L23: sentence or explanation is not finished, the leaf-water deficit increases in the dry season. … The next sentence starts with “Once leaf-water is close to saturation”, this is not a logical following up from an increasing deficit.

Page 8419 L27-28: I miss the hydrological modelling aspect here, while I think that what is mentioned here fits more to what is done in the paper, the perspectives are way less strong than the title and better, so see my general comments to change the title.

Page 8420 L7-8: I miss a reference here that the root zone is the key part, I suggest Shukla & Mintz (1982).

Page 8420 L14-16: Not clear why observations at certain depths are not the right indicators for storage in the root zone and this sentence does not give at all an explanation why correlations between remote sensing products and observed soil moisture...
at different depths are not that good

Page 8420 L22: Not clear why NDVI is explained here, the explanation is not connected to the NDII and is very confusing, it makes the whole introduction unclear and it does not read as a coherent part of the paper

Page 8421 L16: They also discovered...canopy EWT and NDII. This sentence is too long. That VWC is the most successful parameter for retrieval of soil moisture should be explained or it needs a reference. Next to that Yilmaz et al. (2008) show that VWC and NDII relationships change during waterstress, but this is not mentioned here.

Page 8421 L20: Fensholt and Sandholt (2003) is not a strong reference since they found a strong correlation between NDII and in situ top layer soil measurements, where it is known that top layer measurements cannot be used directly for the root zone soil moisture storage (Dumedah et al., 2015)

Page 8421 L27: the relationship between NDII and root zone soil moisture storage was evaluated at 8 sub-basins, not 14

Page 8422 L4-9: This whole part is already mentioned in the introduction earlier, so should be deleted

Page 8423 L6-8: This is already mentioned in the introduction

Page 8423 L11-14: reflectance the other way around, a higher rho 0.85 means a value above zero

Page 8424 L1: elevation corrected daily average rainfall, where is this data derived from? Monthly precipitation data, or daily precipitation data? In the latter case, it is not the average rainfall but the sum of rainfall. Or is it daily rainfall data averaged from several years?

Page 8424 L2: How is the potential evaporation derived?

Page 8425 L11: title of 2.3 does not need the word ‘method’, it is a paragraph of
methods

Page 8426 L2-4: Not clear why ‘n’ is used, is this not a fixed number based on the number of model parameters?

Page 8426 L5: Not very useful to have a title ‘Study site and data used’ and two subtitles ‘study site’ and ‘data collection’. I suggest to think of another title or to put section 2 and 3 together under the title Materials and methods, so that explaining about NDII can be followed by the part about satellite data, and that this part starts with the study site

Page 8426 L8: Northern Thailand

Page 8426 L19: not clear why land cover in 1988 is relevant for this study

Page 8427 L8-11: Strange sentence, make it shorter and as it, as if in one sentence makes it difficult to read

Page 8427 L22: What were the criteria in the quality control? Unusual rainfall data was excluded, when is rainfall defined unusual? And from this piece of text it is not clear where rainfall data is used for, mention it here

Page 8428 L5: How is the reliability checked?

Page 8428 L6: data period, make clear that this is the data period for calibration

Page 8428 L15: each station, clarify, each of the eight runoff stations

Page 8428 L26: dry season is from November to April, not November and April

Page 8429 L5: average for whole UPRB basin, not mentioned in text but is included in caption of table

Page 8429 L15 a wet year and a dry year do not occur in the same year, strange sentence, should be dry and wet season

Page 8429 L17: The 8 day NDII values were also computed, what were the results?
This sentence has no follow up part, so it should be deleted or more information should be added.

Page 8429 L21: make clear that higher moisture content is related to the vegetation.

Page 8430 L6: Why are these two stations chosen to put in the figure? Not only the duration curves are shown but also some hydrographs. It would be good to indicate the volume error in runoff by the model (Krause et al., 2005).

Page 8430 L7: remove ‘that’.

Page 8430 L26: - - should be –

Page 8431 L4: .. once the rainfall reduces significantly. .. Where can this information seen in the graph? Also in L 8, dry spells in the rainy season, where does this information come from?

Page 8432 L16-17: key to sensing the storage of moisture, indicate that this is only under deficit conditions.

Page 8432 L17: the water content in the leaves is directly connected to the suction pressure of the roots, needs a reference.

Page 8433 L9: several studies, but only one reference?

Page 8434 L6: NDII used to investigate drought, it was used to investigate vegetation water content? If it is used to investigate drought, which value of NDII is the threshold for a drought?

Page 8434 L20: is should be are.

Page 8451 Figure 4: Explain in caption that three wettest and three driest basins are presented in the graph.

Page 8453 Figure 6: the graphs are too small and the axes are not readable, it is also not mentioned that x-axis changes scale in the different plots.
Page 8456 Figure 7 and 8: indicate how well the correlation is because I see clear difference in figure 7 and 8 in the dry season, while in the text it is mentioned that for both figures the correlation is high

4 References


"Hydrologic Processes in the Unsaturated Soil Layer." Soil Moisture Regime: E-

Interactive comment on Hydrol. Earth Syst. Sci. Discuss., 12, 8419, 2015.
Fig. 1.