Replies to Referee 2

The authors have suggested a method for creating a soil moisture product for hydrological applications using multiple data sources retrieved from three sources (SAC-SMA land surface modelling product, MODIS satellite-retrieved land-surface temperature, and Soil Moisture and Ocean Salinity (SMOS) project data) using the Gamma-test and the Local Linear Regression techniques. The accuracy of the produced soil moisture data was evaluated against the Xinanjiang (XAJ) hydrological model's soil moisture simulations. The authors have concluded that "together with the chosen data inputs can be used with high confidence to estimate an unintermitted hydrological soil moisture product, and the proposed method could be easily applied to other catchments and fields". The topic is of current scientific interest and the manuscript is overall well prepared. However, there are some general points that need to be clarified and at some points more detailed information or analysis is necessary. The following general and specific comments should be addressed before accepting this manuscript for publication.

General Comments

1. Duration of the study period is two years: from January 2010 to December 2011. In my opinion, the presented calibration/validation results don't allow a reader to evaluate the model applicability and the aforementioned conclusion looks too optimistic. The point is that the study period is too short, and the presented results of the model validation are deficient. It means that the overall model performance based on these results is very sensitive to the meteorological conditions of the study period and the performance assessments are rather casual. This fact limits opportunities for application of the proposed method "to other catchments and fields". The conclusion on the model applicability would be more convincing if the authors evaluated the model against hydrological data for a longer period. According to the USGS Water Resources webpage, streamflow data series exceed 10 years for the Vermilion River at Pontiac, IL.

Reply: We agree with the referee's suggestion on using a longer period for model evaluation. However this study has been constrained due to the following two reasons: first although the selected catchment has flow data over 10 years, most of them are discontinuous (e.g., frequent data gaps), and the selected period provides the most complete flow observations. This is essential for the XAJ model's calibration and validation (it is a continuous simulation model instead of an event-based model); second the SMOS satellite was launched in late 2009, so its data products are only available since then. In addition, since this study is a proof of concept for a new method in optimal fusing of multi-source data for soil moisture estimation, it is more effective to inform the hydrological community about such an approach and trial it in a variety of locations than testing it at one location for a few more years.

2. The Pontiac catchment is characterized by frequent soil freezing events in winter seasons. During freezing events, soil moisture transfer fundamentally differs from the unfrozen conditions (e.g. Gelfan, 2006). To my knowledge, the lumped XAJ model does not consider soil freezing, thus SMD simulations can be inaccurate for winter seasons. Please clarify.

Gelfan A. N. (2006) Physically based model of heat and water transfer in frozen soil and its parametrization by basic soil data. IAHS Publ., 303, pp. 293-304.

Reply: The lumped XAJ model has been frequently used in frozen soil conditions (e.g., see Application of Xin'anjiang model in severe cold region of Niqiu River, 2008 (Zhou et al., 2008)). The XAJ model's evapotranspiration component plays a vital role in the model's flow generation. The component constitutes three soil storages as seen in Fig.2 (i.e., the SMD simulations). The accuracy of SMD can therefore largely affect the accuracy of the model's flow calculation. As seen in Fig.3, XAJ model is very good at simulating the flow variations in the Pontiac catchment with high NSE values, even during the winter season, which agrees well with the literature. This indicates the model's SMD simulations are quite realistic.

Zhou, S., Li, Y., Zhu, J., 2008. Application of Xin'anjiang model in severe cold region of Niqiu River. Water Resources & Hydropower of Northeast China, 26(9). DOI:10.3969/j.issn.1002-0624.2008.09.016

3. The authors argued that "only the surface SMD referring to the vegetation and the very thin topsoil, is utilised as a hydrological soil moisture target". Does the XAJ model allow one to simulate SMD in the "very thin topsoil"? If no, this should be clearly pointed out in the manuscript, and the simulation results' interpretation should be corrected.

Reply: XAJ model has three soil layers in its modelling structure, and yes the top layer represents the very thin topsoil. We will make this clearer in the updated manuscript.

4. I fully agree with the authors that the results are "model parameter dependent" (line 486). But I disagree that the proposed NHSMS indicators allow one to obtain independent results. I think that Figure 13 cannot be considered as an evidence of such independence because of at least two reasons: (1) only one parameter has been changed; (2) the obtained closeness of the two curves is shown for only 4 months (of 2 years) with mostly high SMD values. Thus I believe that the results presented in the manuscript are dependent on the XAJ model structure, parameters and inputs. Please give a comment.

Reply: It is true the proposed NHSMS only considers the model parameter factor, and other factors such as model structure is required in our future studies. As a result, the proposed indicator is only the very first step into creating a universal soil moisture product, and a lot of case studies adopting different hydrological models will need to be carried out. This has been pointed out in the original manuscript Lines 493-500 ("In the future it is planned to use the same process on other hydrological models to test if the normalised soil moisture indicators are not only model parameter independent but also model structure independent. Since all hydrological models are driven by the same hydrological inputs (precipitation, evapotranspiration and flow), their normalised soil moisture indicators should respond in a similar way (soil becomes wetter when it rains and drier when there is no rain). If this is true a new soil moisture product based on NHSMS could be generated as a routine product by the operational organisations such as NASA and ESA.").

The reason of showing the 4 months results is for a better visualisation. The selection of a dry period (i.e., high SMD values) is because it is the most critical period of time for the need of accurate soil moisture values for hydrological modelling. A good analogy is a hydrological model's water storage can be seen as a tea cup, during wet season with a lot of rain, it is easy to calculate how much water comes out of the cup as we know the maximum storage amount (for a hydrological model, it is via parameter calibration), so it is the total rainfall minus the maximum storage. However during dry season, it is impossible to know the outflow amount.

Therefore during the real time flood forecasting, after a long period of dryness, the accumulation of error in the hydrological models can become larger and larger with time. With accurate soil moisture information, the error could be corrected. This clarification will be added in the updated manuscript.

5. Conclusion, lines 539-541. I do not share the optimistic view of the authors on the perspective of the proposed fusion technique. Yes, the data sources contain part of useful (and probably independent) information. However, these data contain their own large measurement errors and error's synergy can result in dramatically increase of the presented results' uncertainty. I would like to read the authors' comment on this topic.

Reply: It is true that all data sources have their own errors and the errors' synergy could increase the uncertainty of the merged data. This is why a good data fusion scheme should be explored and adopted. In this study, when combining them together using the data fusion method, the total error is controlled by the desired target (the XAJ model's SMD, which has been regarded as quite realistic due to the model's good flow simulations results).

Specific Comments

1. Lines 133-136: The phrase beginning from the words "It is worth noting that. . ." looks unnecessary in scientific (non-popular) hydrological text

Reply: Agreed. It will be removed.

2. I suggest removing Fig. 3. This figure has been already demonstrated in three (at least) recently published papers of the authors (Zhuo and Han, 2016, 2017; Zhuo et al., 2015)

Reply: Agreed. It will be removed.

3. Line 165: There are 17 parameters in Table 1

Reply: Agreed. It will be corrected.

4. Eq. 1: M is not defined

Reply: The definition will be added.

5. Lines 495-496: The statement "all hydrological models are driven by the same hydrological inputs (precipitation, evapotranspiration and flow)" is misconception. Please be more precise.

Reply: This statement will be modified as "all hydrological models are driven by the same physics laws on the conservation of mass…"