June 13, 2013

Dear HESS Editorial Board and Reviewers,

The authors would like to thank you for your time spent in reviewing our manuscript "Improving simulation of soil moisture in China using a multiple meteorological forcing ensemble approach (hess-2013-103)" submitted to *Hydrology and Earth System Sciences.* Because of your constructive comments and suggestions, the revised manuscript was improved greatly. It is appreciated very much.

Sincerely,

Zhenghui Xie
(On behalf of all authors)

Response to Referee #1:

We thank the referee for the constructive comments and suggestions, which are in plain text below. Our response is bold text.

General comments:

I recommend that this paper either be rejected or sent back to the authors for major revision. It has several flaws that would need to be addressed before it would be ready for publication:

1. What is the real bottom line? That if you test a land surface model, remove the bias in the simulations, and then weight simulations by the ones that already do the best, that you can get a good simulation? This is just curve-fitting and is obvious. What is the new science? That you give this procedure a fancy name, Bayesian Model Averaging?

Response: In present work, our main aim is how to reduce the land surface simulation uncertainty arising from meteorological forcings, and how to improve soil moisture simulation. Since none of individual simulation by a forcing do the best at all subregions, and multiple forcing ensemble approach can reduce the uncertainty arising from individual forcing, we applied BMA method to improve the simulation further and to investigate the role of forcing in the soil moisture simulation by a land surface model, which can provide a reliable description of the modeling uncertainty, and can reduce the uncertainty. The BMA ensemble simulation reproduced anomalies and seasonal variations in observed soil moisture, and simulated the mean soil moisture. It is presented here as a promising way for reproducing long-term, high-resolution spatial and temporal soil moisture data.

State Key Laboratory of Numerical Modelling for Atmospheric Sciences
and Geophysical Fluid Dynamics (LASG)
Institute of Atmospheric Physics, Chinese Academy of Sciences

2. The authors use four different input data sets, but do no analysis of those data sets. Why are they different? They each need to be validated with in situ observations before they can be trusted. They also need to be compared, with time series and maps. Some have identical data going into them, such as 2 and 3 both using surface station data. So what differentiates them? Obviously some are better than others. By forcing the model with some data sets that are not very good, and then given less weight to the bad results, you get a better answer – but how could you expect this not to be the case?

Response: Following the reviewer's suggestions, we will add a proper analysis (Figures) and more discussions on the used meteorological forcings in the revised manuscript.

3. The paper needs to investigate why some simulations did better than others. This will follow from an analysis of the input data in its biases. If you did not force the model with poor data in the first place, you would not need to do corrections to the output.

Response: We will do it as suggested by including an analysis of the input data in its biases and relationship between simulated soil moisture and forcing (precipitation，temperature and radiation) in the revised manuscript.

4. The simulations were done incorrectly. The authors first did a long run with one of the forcings to 2011, but then used that balanced output to initialize simulations in 2005 with the four different forcings. This will obviously cause a drift due to spin-up for two reasons – the model will start with the incorrect soil moisture values, and the forcing (for the three not used in the first place) will produce a drift toward a different climate. Yet it seems that the authors used three years of this simulation to train the Bayesian model, without correcting for drift in at least several months at the beginning of the run. You can clearly see this drift in Fig. 5 for the 70-100 cm layer for several of the simulations.

Response: The spin-ups and initializations on the simulations were not described clearly in the section 3.1 of the manuscript. We will revise it in the revised manuscript. Here, we spin-up in order to achieve an equilibrium state of the CLM3.5 and to get reasonable initial fields. Due to the four different forcings spanned different time periods, we first did a long run using ITP forcing from 1979 to 2010, and then applied the balanced output (1 January 2011 or 1 July 2010) to initialize all four simulations at the initial time of four different forcings (1 January 1979 for ITP and JRA; 1 January 2004 for TIAN; 1 July 2005 for FY), not in 2005 for all four simulations. At last, we chose the common time periods (July 2005 – December 2010) to study. We think that the drifts in the first place of Fig. 5 are caused by different forcings.

5. Why was the CLM model used? How would the results depend on the land surface model used? Has this model done well in other intercomparisons? Why not use the well-tested VIC or NOAH models?

State Key Laboratory of Numerical Modelling for Atmospheric Sciences
and Geophysical Fluid Dynamics (LASG)
Institute of Atmospheric Physics, Chinese Academy of Sciences

Response: The community land surface model CLM3.5 (Oleson et al., 2007) is released by the National Center for Atmospheric Research, which is a modified version of CLM3.0 (Oleson et al., 2004), and has significantly improved the simulation of many variables in the hydrological cycle. CLM3.0 is the land surface component of the community climate system model (CCSM3.0). One advantage of a community model is that there are a significant number of scientists willing to scrutinizing its scientific contests, offer constructive criticism, and improve its performance. Many previous studies have demonstrated that CLM3.5 simulated soil moisture in mainland China is proper (Li and Ma, 2010, 2011; Wang et al., 2011a, 2011b). The quality of soil moisture simulation also depend on the land surface model (LSM) used. We compared the results of soil moisture simulations using different LSMs (CLM3.0，CLM3.5，VIC, BATS, NOAH et al) in mainland China，the results showed that none of LSMs simulated accurately soil moisture values in all subregions of mainland China. CLM3.5 simulated soil moisture values were closer to observed soil moisture values, and we can get soil moisture simulations at 10 uneven soil layers. Furthermore, the quality of soil moisture simulation by VIC and NOAH models depends on regional parameter estimations. In present study, we did not consider the uncertainty of CLM3.5, we will add some discussions about the uncertainty of CLM3.5 in the revised manuscript.

6. Why are there gaps in the results for three of the regions in Figs. 3-5? Even if the observations are missing, why are the simulations missing? Or were there gaps in the forcing data? If so, was the model restarted each time? From what initial conditions?

Response: Because the soil moisture was measured in the warm season, and no observations in the frozen soil. Hence there are gaps in the observations for three of the regions over northern China (China I, II, VI) in Figs. 3-5. In order to compare the observation and the simulations, they are consistent in time. In fact, there are no simulations missing over these subregions.

7. The English is pretty good, but there are a few mistakes. I attach an annotated manuscript with corrections, but also with comments on a number of other items, all of which have to be addressed in any revision.

Response: We will correct the mistakes as suggested in the revised manuscript.

8. What time step was used for the simulations?

Response: The time step is 1800s. We added the time step in section 3.1 Experiment design in revised manuscript.

9. There needs to be a figure comparing the model levels and the observed levels, showing on a linear depth scale each of the layers of each. The text says observations were taken at 50 and

State Key Laboratory of Numerical Modelling for Atmospheric Sciences
and Geophysical Fluid Dynamics (LASG)
Institute of Atmospheric Physics, Chinese Academy of Sciences

100 cm, but then they use data for 70-100 cm. How is that possible? I know that actually observations are taken for each 10 cm layer from 0 to 100 cm. The question is how they were saved and put into the archive – with what averaging.

Response: We added a figure including hydrologic process simulated by CLM3.5 and the soil layers formation. In previous manuscript, we missed the soil depth 70 cm. This sentence has been changed to "The soil moisture was measured three times on the 8th, 18th and 28th days of every month in the warm season, and no observations in the frozen soil, and at the soil depths of 10 cm, 20 cm, 50 cm, 70cm and 100 cm".

10. This is a very valuable soil moisture data set. How can other scientists access the data to repeat these experiments? If they are not generally available, this is a problem for science. Are the data in the International Soil Moisture Network, http://www.ipf.tuwien.ac.at/insitu/? If not, why not?

Response: We have obtained data of in situ measurements of soil moisture during 2005.7-2010.6 at 411 stations from the China Meteorological Administration (CMA) National Meteorological Information Center (NMIC) operational network through personal communications from C.X. Shi (2011). It is constantly updated. Some of 411 stations are not included in International Soil Moisture Network. They will be included in the International Soil Moisture Network in the future.

Please also note the supplement to this comment: http://www.hydrol-earth-syst-sci-discuss.net/10/ C1361/2013/hessd-10-C1361-2013-supplement.pdf

Response: According to the supplement to this comment, we corrected some mistakes and revised the manuscript.