Response (Referee #4 comment)

Ms. Ref. No.: hess-2022-379

Title: Improving predictions of land-atmosphere interactions based on a hybrid data assimilation and machine learning method

This is a very well-written manuscript, and this reviewer enjoy reading it through. Some major comments as below:

We gratefully thank the reviewers for their review, which we believe has led to significant improvement on the original manuscript. The original reviewers' comments are reproduced below in black text and the corresponding response is shown in blue text.

The wetting/cooling effect of the oasis is interpreted in this manuscript as WRF(DA-ML) - WRF(OL). Based on figure 9, this manuscript emphasizes that this oasis effect is related to irrigation and crop growth in the midstream. However, the wetting/cooling effect of oasis is by itself there, no matter if the DA-ML framework is applied or not. As such, this reviewer found that this manuscript is lacking of certain indices to demonstrate physically the wetting/cooling effect of the oasis (for example, one can use the difference in air temperature, and relative humidity between (above) the oasis and the surrounding areas). And then you can check how this indicator will be impacted by DA-ML (e.g. Oasis_Indictor (DA-ML) - Oasis_Indictor (OL))

Thanks for your comment. The oasis-desert interactions are caused by the different hydrothermal conditions of oases and deserts. Oasis-desert interactions lead to a series of microclimate effects, including the oasis wetting/cooling island effect and oasis wind shield effect. In this study, the proposed WRF (DA-ML) method effectively reduces air warming and drying biases in simulations, particularly in the oasis region. Therefore, this method can simulate more realistic oasis-desert boundaries, including wetting/cooling effects and wind shield effects within the oasis.

To avoid ambiguity, the relevant description is revised in section 4.3: "The above mentioned findings show that the proposed WRF (DA-ML) method exhibits wetting and cooling effects in the mid- and downstream oasis. These wetting and cooling effect reduces the air warming bias and dry bias in the simulation. Therefore, the WRF (DA-ML) simulation is much closer to the observations than the WRF (OL) simulation. Two rectangular areas (blue rectangle) were selected in Figure 10 to further analyze the effect of the DA-ML on the local climate in the mid- and downstream areas. The difference between the WRF (DA-ML) and WRF (OL) methods is used to represent the enhanced cooling and wetting effects after improving the LAI and SM simulations."

The following sentence will be added to section 4.3: "The average simulated air temperature from WRF (OL) and WRF (DA-ML) methods in the midstream oasis were 293.64 K and 291.32 K, respectively. In contrast, the near-surface air temperatures over the desert are approximately 294.13 K and 293.54 K, respectively. The difference in air temperature between the oasis and desert areas indicates that the oasis areas represent a cold and wet island compared to the surrounding desert. This difference is amplified after the implementation of the DA-ML method."

The authors state that the wetting/cooling effects of the downstream oasis are due to the shallow groundwater and riparian forest growth. This reviewer can understand that the 'riparian forest growth' can be reflected via the LAI assimilation. However, it is not explicitly clear how the shallow groundwater kicks in here. Are the authors suggesting the assimilation of root zone SM could be used to reflect the effect of shallow groundwater? If that is the case, the author should demonstrate it is indeed the case using the root zone SM, groundwater table measurements, and Noah-MP GW table simulations.

Thanks for your comment. In the downstream area of the HRB, the shallow groundwater table influences the root zone soil moisture and controls the growth of the plant communities (Xu et al., 2020; Li et al., 2022). In this study, the *in situ* SM profile observations are used as target variables to train the SM surrogate model. Therefore, compared to the WRF model, the WRF (DA-ML) method can take into account root zone soil moisture variations due to the shallow groundwater.

The spatial distribution of the root zone SM (0-100 cm) estimates from the Noah-MP and hybrid model over the HRB are shown in Figure R1. Compared to the hybrid model, Noah-MP underestimates SM in the north areas of HRB. The results indicate that soil moisture values are higher in the downstream riparian forest areas, which is due to the lateral flow of the river recharging groundwater and affecting soil moisture.

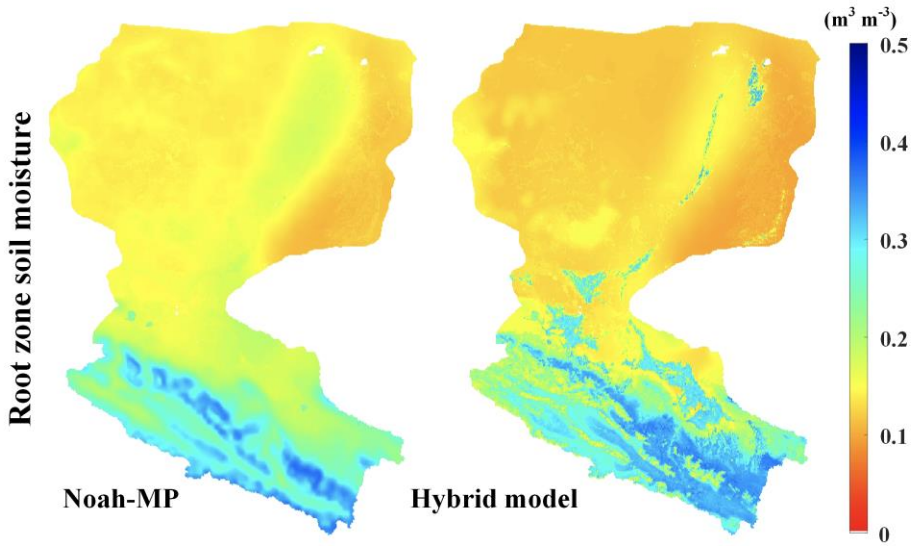

Figure R1. Spatial distribution of the root zone soil moisture (0-100 cm) estimates from the Noah-MP and hybrid model over the HRB in 2015.

Xu et al. (2020) analyzed the changes in groundwater table in the downstream of HRB based on the Heihe integrated observatory network. Figure R2 shows the groundwater table variations in the downstream oasis area in 2015. The groundwater was very shallow in the downstream area (approximately 1–3 m). Therefore, shallow groundwater will enhance the root zone soil moisture and accelerate vegetation transpiration.

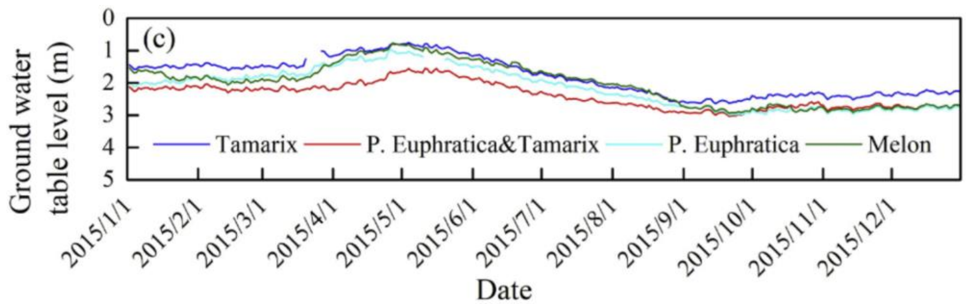

Figure R2. The groundwater table variations in natural oasis in the downstream area in 2015.

The following sentence will be added to section 3.2: "Compared with the direct assimilation of coarse-resolution remotely sensed SM, this method can improve the estimation of SM and ET on the heterogeneous land surface. This is because *in situ* SM profile observations are used to construct an ML-based surrogate model to improve SM and ET estimation on complex underlying surfaces."

The following sentence will be revised in section 3.2: "The SM surrogate model can consider the effects of midstream irrigation events and downstream shallow groundwater tables on SM and improve Noah-MP ET estimates."

Although the SM surrogate model development has been published in another paper. This reviewer strongly suggested the author illustrate how these surrogate SM models were constructed with workflow/flowchart etc.

Thank you for your good comment. More details about the soil moisture surrogate model will be added to section 3.2: "In the ML part, the normalized soil texture (ST), land cover (LC), air temperature and humidity (Ta and RH), wind speed (*U*), precipitation (*P*), solar radiation (Rs), LAI, and SM observations were used to construct the SM surrogate model. ST, LC, Ta, RH, *U*, *P*, Rs, and LAI are the predictor variables. The *in situ* SM profile observations (from 19 automatic weather stations) and SMAP SM products in the HRB are used as target variables to train and test the SM surrogate model. The extreme gradient boosting (XGBoost) method was chosen in the SM surrogate model to improve multi-layer SM simulations."

To address the reviewer's comment, the flowchart of the hybrid model coupled with the WRF is shown in Figure 2.

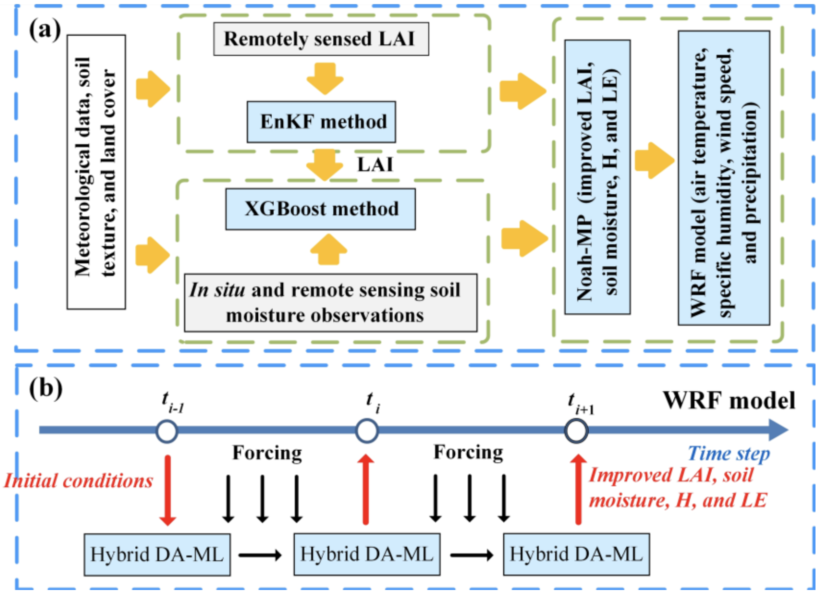

Figure 2: (a) Details of the hybrid DA and ML method, and (b) flowchart of the coupling with the WRF model.

Minor comments:

L93: improving the representation of soil and vegetation processes in affecting regional climate via the coupled DA and ML ....?

Thank you for your comment. The sentence is revised as: "However, the advantages of improving the representation of soil and vegetation processes in affecting regional climate via the coupled DA and ML framework have not been fully exploited."

L165: It would be wise to illustrate how this DA-ML framework is constructed with a workflow/flowchart.

Thank you for your comment. The flowchart of the hybrid model coupled with the WRF is shown in Figure 2.

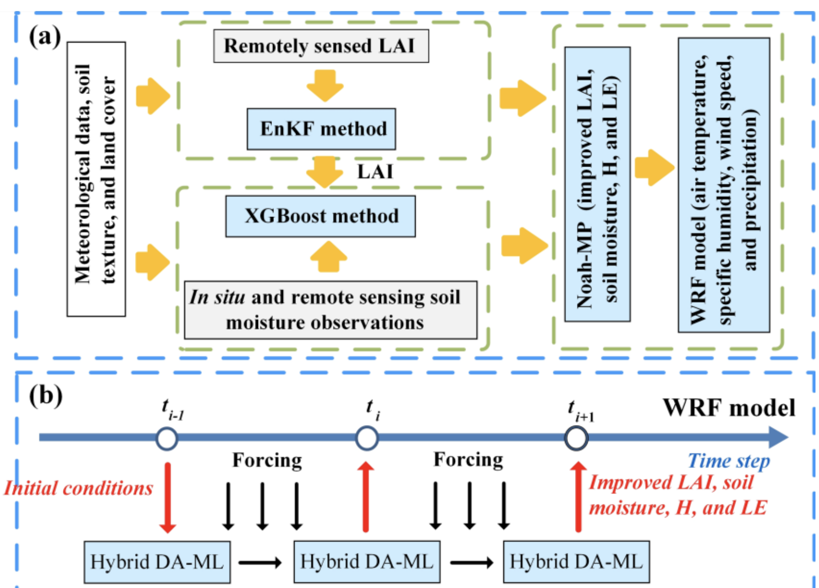

Figure 2: (a) Details of the hybrid DA and ML method, and (b) flowchart of the coupling with the WRF model.

L183: what are predictors? and what is the target variable? A diagram illustrating how these two surrogate models were constructed would be helpful for readers, and spare their need to read He et al. 2022 paper, in order to have this paper fully understood.

Thank you for your comment. In this study, the normalized soil texture (ST), land cover (LC), air temperature and humidity (Ta and RH), wind speed (*U*), precipitation (*P*), solar radiation (Rs), LAI, and SM observations were used to construct the SM surrogate model. ST, LC, Ta, RH, *U*, *P*, Rs, and LAI are the predictor variables. The soil moisture profile observations and SMAP SM products in the HRB are used as target variables to train and test the soil moisture surrogate model (machine learning model).

The following paragraph will be revised in section 3.2: "In the ML part, the normalized soil texture (ST), land cover (LC), air temperature and humidity (Ta and RH), wind speed (*U*), precipitation (*P*), solar radiation (Rs), LAI, and SM observations were used to construct the SM surrogate model. ST, LC, Ta, RH, *U*, *P*, Rs, and LAI are the predictor variables. The *in situ* SM profile observations (from 19 automatic weather stations) and SMAP SM products in the HRB are used as target variables to train and test the SM surrogate model."

In addition, more details about the SM surrogate model are illustrated in Figure 2.

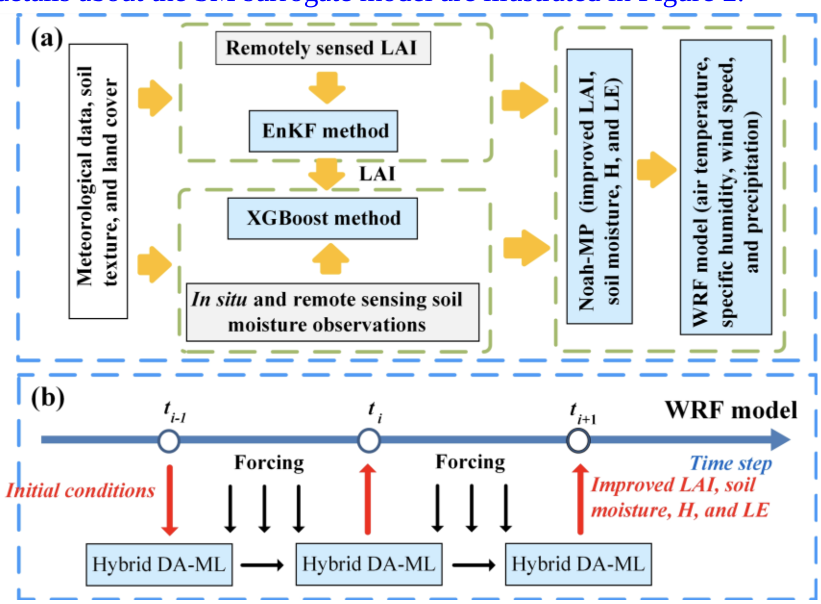

Figure 2: (a) Details of the hybrid DA and ML method, and (b) flowchart of the coupling with the WRF model.

L193: The LAI is updated via Noah-MP, and the SM is updated via the surrogate model?

Thank you for your comment. Yes, the LAI is updated via Noah-MP, and the SM is updated via the surrogate model. The sentence will be added to section 3.2: "In this step, the LAI is updated by the DA method, and the SM is updated via the ML-based surrogate model."

L360: but this is kept unchanged in either WRF(OL) or WRF (DA-ML), right?

Thank you for your comment. The bulk transfer coefficient is changed after the implementation of the DA-ML method. The mean bulk transfer coefficient shown in Figure 11 is to compare the surface roughness in the oasis with that of the surrounding desert.

The sentence is revised in section 4.4: “In Figure 13, the bulk transfer coefficient (C_h) from the WRF (DA-ML) was used to compare the surface roughness in the oasis with that of the surrounding desert.”

References:

Li, Xin, Cheng, G., Fu, B., Xia, J., Zhang, L., Yang, D., Zheng, C., Liu, S., Li, Xiubin, Song, C., Kang, S., Li, Xiaoyan, Che, T., Zheng, Y., Zhou, Y., Wang, H., Ran, Y., 2022. Linking Critical Zone With Watershed Science: The Example of the Heihe River Basin. Earth’s Future 10. https://doi.org/10.1029/2022EF002966

Xu, Z., Liu, S., Zhu, Z., Zhou, J., Shi, W., Xu, T., Yang, X., Zhang, Y., He, X., 2020. Exploring evapotranspiration changes in a typical endorheic basin through the integrated observatory network. Agricultural and Forest Meteorology 290, 108010. https://doi.org/10.1016/j.agrformet.2020.108010