The author's answers are indicated in red color, as well as old text passages. New text passages are indicated in green color.

General comments: This manuscript describes a wide-variety of approaches to estimate subsurface and profile soil moisture from surface soil moisture data in the Qilian Mountains of China. Most of the conclusions are well-supported, but the manuscript suffers from statistical inconsistencies. A consistent set of statistical measures should be maintained throughout the manuscript (NSE, RSR, and R). The only exception could be Fig. 10 where the aim is to compare the SMAP ubRMSE to the mission's accuracy requirement. Also, 30% of the data should be withheld for validation for all three methods, not only the ANN method. The performance of the SMAP-based ExpF method for estimating profile soil moisture (i.e., SWI) is overstated, and including the NSE and RSR statistics will likely provide a much more objective view. The manuscript also suffers from poor organization in some places and is not well-written.

Response:

Thanks for your comments.

We have maintained the consistency of the statistics throughout the manuscript, expect for the use of ubRMSE in the evaluation of SMAP product.

The T_{opt} parameter of the ExpF method reflects the characteristic length of the temporal dynamics of soil moisture. Earlier studies revealed that T_{opt} is highly dependent on the sampling interval of soil moisture data (De Lange et al., 2008). In our study, we found that when using the random sampling with 70% training data as for the ANN method, T_{opt} was not suitable for the remaining data. Since it was not possible to use the same training method for ExpF method as for ANN, we used the entire soil moisture time series to estimate T_{opt}, which was also the standard procedure in earlier studies (e.g. Wagner et al., 1999; Albergel et al., 2008; De Lange et al., 2008; Ford et al., 2014; Wang et al., 2017).

In the revised manuscript, the performance of SMAP-based ExpF method for estimating profile soil moisture has been reanalyzed using the metrics of NSE and RSR.

In addition, we have deleted some content that is not important for the analysis and we have reorganized the text to make the revised manuscript easier to understand.

The revised manuscript is now divided into three parts. Firstly, we evaluated the different methods for estimating subsurface soil moisture (SM). The ExpF method was found to be the most suitable method for the further application in the study area.

Secondly, our results indicate that the median value of T_{opt} can be used for application of the ExpF method in the study area.

Finally, the ExpF method derived with the median value of T_{opt} was combined with the SMAP_L3 surface SM product to estimate the subsurface SM. The subsurface SM was also compared to the SMAP_L4 root zone SM product (a widely used large-scale root zone SM product). Results indicated that the combination of the ExpF method with the SMAP_L3 surface SM product can significantly improve the estimation of profile SM in mountainous areas. Furthermore, the combination of the SMAP_L3 and ExpF method (with the median value of Topt) were applied to estimate the temporal and spatial distribution of profile SM in the study area.

Specific comments:

1. line 126. Clarify the meaning of “to consider the assumption of uniform vertical profiles of soil temperature and soil dielectric properties”.

Response:

We have deleted it in the revised manuscript:

2. Table 1. Results are being reported with too many significant figures. I doubt that the laboratory is able to measure sand and silt content with adequate precision to justify four significant figures. Reduce to a more appropriate level, perhaps three significant figures.

Response:

We have reduced the level to three significant figures in the revised manuscript:

Table 1. Statistics of the soil physical characteristics at the 35 soil moisture stations: mean (standard deviation)

Layer	Depth (cm)	Bulk Density (g/cm^3)	K_S (cm/hour)	SOC (g/100g)	Sand (%)	Silt (%)	Clay (%)
Layer 1	0-10	1.13(0.28)	3.87(4.11)	4.35(4.11)	26.6(11.9)	66.2(10.9)	7.2(1.6)
Layer 2	10-20	1.14(0.24)	4.61(4.53)	3.9(3.87)	24.5(11.9)	68.6(11.2)	6.9(1.2)
Layer 3	20-30	1.18(0.32)	4.78(6.22)	3.63(3.54)	27.0(15.2)	66.5(14.3)	6.5(1.4)
Layer 4	30-50	1.29(0.3)	3.94(4.68)	2.21(2.28)	29.5(15.3)	63.8(14.5)	6.5(1.6)
Layer 5	50-70	1.34(0.3)	1.85(2.35)	2.34(2.47)	26.9(17.1)	66.5(15.9)	6.7(1.9)

Note: K_S is the Saturated Hydraulic Conductivity; SOC is the Soil Organic Carbon.

3. Section 3.4. Provide more explanation. Was a separate ANN model developed for every depth combination and every site?

Response:

Yes. In this study, a separate ANN model was developed for every depth of every site. We have added more details about the set-up of the ANN method in the revised manuscript (Section 3.3, Line 153-165).

4. Equation 6 is the wrong equation.

Response:

We have corrected Equation 6 in the revised manuscript. (Line 173)

$$\hat{\Delta} = K_0 + K_1 \cdot \theta_1 + K_2 \cdot {\theta_1}^2 + K_3 \cdot {\theta_1}^3 \quad (9)$$

Where $\hat{\Delta}$ is the predicted difference between surface and subsurface SM, and K_i (i=0,1,2,3) are parameters.

5. line 196. not “persistent” but “consistent”

Response:

As suggested by referee 3, we have deleted this part.

6. line 205-208. This should be moved to the methods section.

Response:

As suggested by referee 3, we have deleted this part.

7. line 257. You have not provided any convincing evidence that “For most hydrological researches, the correct temporal variation of SM is more crucial than the exact value, suggesting that more emphasis should be given to R when selecting the most appropriate estimation method.” You have presented three statistical measures to evaluate these methods (RSR, R, and NSE). For two out of the three statistical measures (RSR and NSE) the ANN method had the best performance. Therefore, you should include a clear statement that the results from the ANN method were statistically superior to those from the other two methods. You are still free to prefer the ExpF approach if it is simpler to apply than the ANN method. Just don’t try to justify that choice on a statistical basis.

Response:

Thank you for your comment and suggestion. Yes, the ANN method is statistically superior to those from the other two methods. However, we prefer the ExpF approach as it is simpler to apply and more process-based than the ANN method. We have deleted the statement “For most hydrological researches, the correct temporal variation of SM is more crucial than the exact value, suggesting that more emphasis should be given to R when selecting the most appropriate estimation method” in the revised manuscript. Furthermore, we have included the statement “The results suggested that for two out of the three statistical measures (RSR and NSE), the ANN method was statistically superior to those from the other two methods.” in the revised manuscript. (Line 213-223)

As expected, all metrics showed that the performance decreased with depth. The results indicate that for two out of the three statistical measures (RSR and NSE), the ANN method was statistically superior to the other two methods. Specifically, the ANN method resulted in the lowest estimation error, while the ExpF method was better able to capture SM dynamics. A similar finding was reported by Zhang et al. (2017a), who found that the ExpF method had a significantly higher correlation coefficient along with a higher mean bias compared to the ANN method. Results also demonstrated that both ANN and ExpF method were able to provide accurate estimates of subsurface SM for layer 2, layer 3 and profile SM.

Overall, both the ANN and ExpF method are useful for estimating subsurface SM from surface SM in our study area, and ANN was the statistically superior method. However, the ExpF method is a simpler approach as it only needs one parameter (T_{opt}), and can thus be easily applied in data-scarce mountainous areas. Therefore, the ExpF method was used to estimate subsurface SM in the remainder of this study.

8. Figure 7a is unnecessary and should be deleted. Your results show that "Year" does not have a significant effect, so the data should be presented including all years as done in Fig. 7b.

Response:

Thank you for your comment and suggestion. We have now merged Figure 5 into Figure 4 in the revised manuscript (Fig. 4).

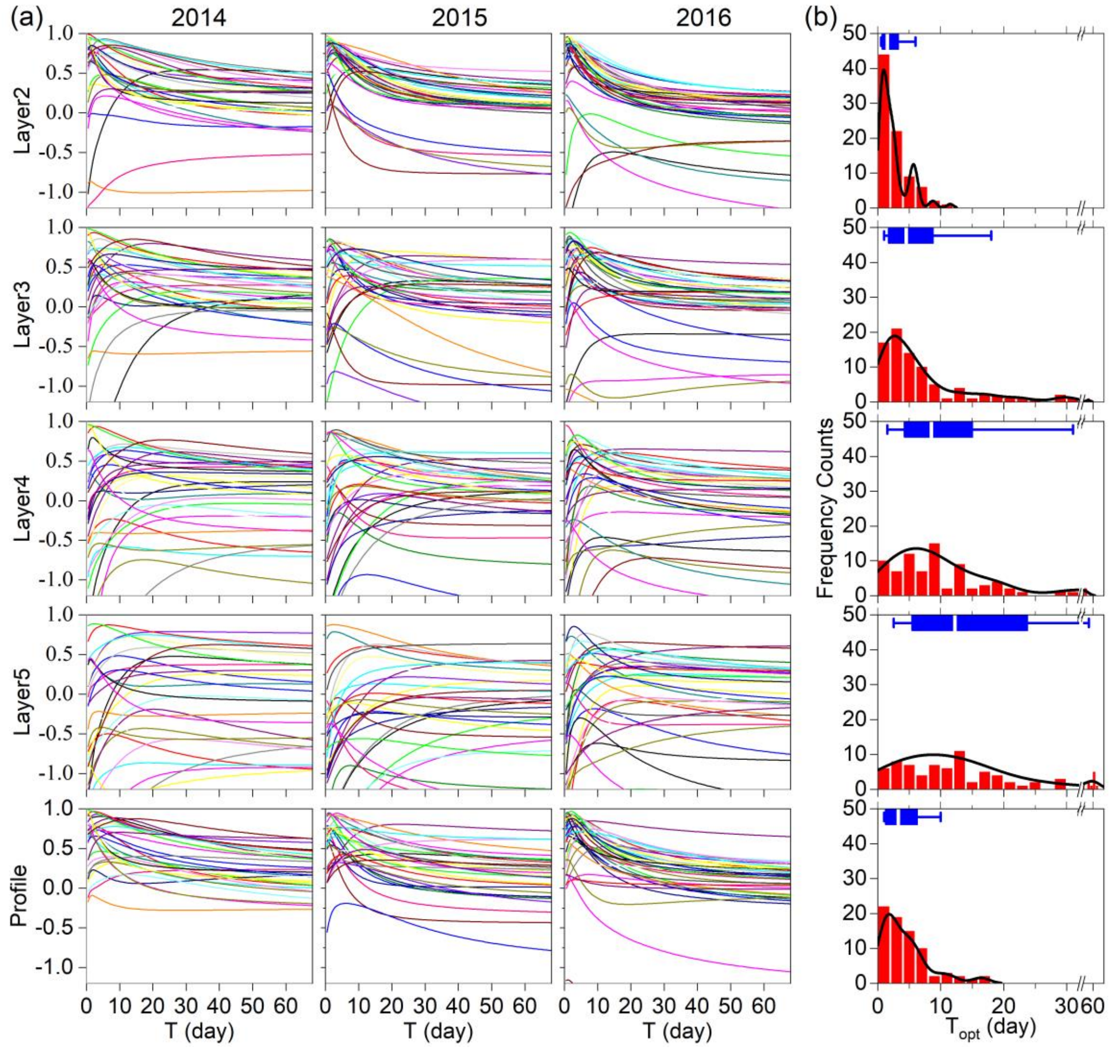

Fig. 4. Variation of NSE with T of the exponential filter method at different layers of each stations during the growing season of 2014, 2015 and 2016. Y axis is the NSE value. Frequency distribution curve and the boxplots to show the distribution of T_{opt} with depth for all stations.

9. line 290-297. This should be moved to the methods section.

Response:

Thank you for your comment and suggestion. As suggested in Comment 10, the correlation among variables (e.g. Ln-transformed LAI and precipitation) are strong, and the results of partial correlation analysis among T_{opt} and the variables indicated that the relationship among T_{opt} and variables are not valid. Therefore, the relationships between T_{opt} and variables are not applied in the further analysis. As a result, this part has been deleted in the revised manuscript.

10. line 298-305. This section is not convincing. How strong is the correlation between ln-transformed LAI and precipitation? Perhaps the apparent relationship between Topt and LAI is a spurious result of the correlation between LAI and precipitation.

Response:

Thank you for your comment and suggestion. The correlation between ln-transformed LAI and precipitation are significant (Pearson's R=0.80, P<0.01). Furthermore, we tested the partial correlation analysis of the ln-transformed LAI, precipitation and T_{opt}. The results showed that the relationships between ln-transformed LAI and T_{opt} are non-significant under the control of precipitation. Meanwhile, the relationships between precipitation and T_{opt} under the control of ln-transformed LAI are not valid for all layers. Thus, this section on the control factors of T_{opt} is not convincing and was removed.

Furthermore, as the control factors and regression of T_{opt} are not applied any more to the further estimation of subsurface soil moisture from the SMAP_L3 product. Thus, we have deleted the content on the control factors and regression of T_{opt} in the revised manuscript.

11. line 319-322. Move to methods.

Response:
We have moved it to the Methods section (Section 3.3).

12. line 319-322. What steps were taken to prevent problems due to collinearity of the predictor variables?

Response:
Firstly, the regression equation was not used in the further analysis. Secondly, it is difficult to obtain independent variables for the regression equations in data-scarce mountainous areas. Thus, the regression equation for T_{opt} should not be used in such environments. Therefore, we have deleted this part in the revised manuscript.

13. Table 4. Is "ln ln (sand)" correct in the last row?

Response:
As stated in the reply for comment 10, this part has been deleted in the revised manuscript.

14. Fig. 9. Present RSR instead of RMSE to be consistent with the rest of the manuscript.

Response:
We have changed the RMSE with RSR in the revised manuscript. (Fig. 5)

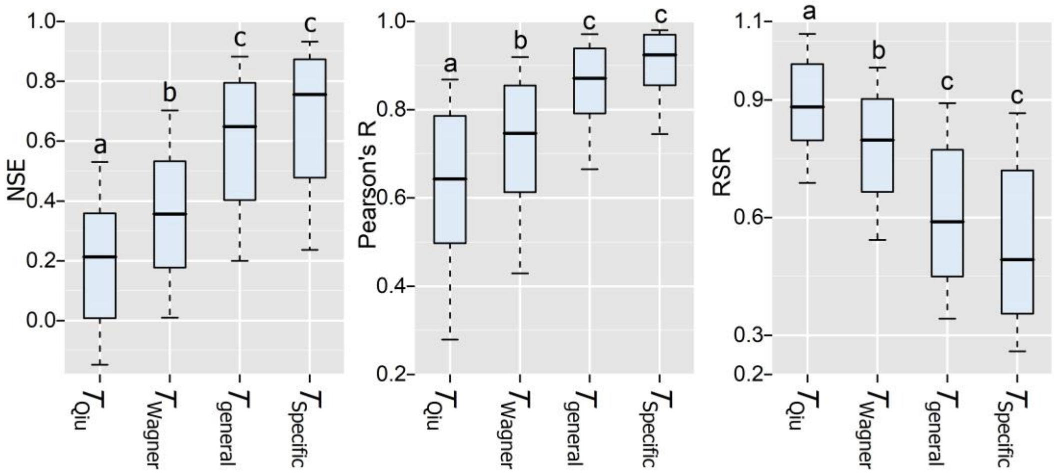

Fig. 5. The boxplot of NSE, Pearson's R, and RMSE for the T_{opt} generated from different schemes. The different letters above box indicate the significant difference for different schemes.

15. line 380. Not "persistent" but "consistent".

Response:

We have changed it. (Line 283)

16. Section 4.4.1. You should note an important limitation of this analysis. There is a huge scale mismatch between the 9 km SMAP data and the in situ sensors which measure at a single point. This will likely degrade the agreement between the two data sets.

Response:

Due to the harsh natural environment of the Qilian Mountain, there is still a lack of high-precision soil property data in the study area (Li et al. 2017; Tian et al. 2017; Li et al., 2018; Zhao et al. 2018), which is necessary to downscale SMAP soil moisture data (Montzka et al. 2018). In general, large research gaps still exist with respect to scaling issue in mountainous areas with complex topography (Jin et al. 2017; Fan et al. 2019). Such scale issues are beyond the scope of this manuscript. We have added a discussion about the limitations in the manuscript. (Line 297- 300)

Notably, the SMAP_L3 product has a spatial posting of 9 km×9 km, while the in-situ measurements are point-based and soil moisture has a strong spatial variability in mountainous areas (Tian et al., 2019). Thus, the disparity of spatial scales between points and satellite footprints will introduce additional errors in the validation of the satellite products (Jin et al. 2017).

17. line 394-399. Move this to methods.

Response:

We have moved it to Data and Methods section (Section 3.1). (Line 122-124)

SMAP_L3 surface soil moisture was also used to estimate subsurface soil moisture (Layer 2: 10-20 cm, Layer 3: 20-30 cm, Layer 4: 30-50 cm, Layer 5: 50-70 cm, Profile: 0-70 cm) during the growing seasons of 2015 and 2016 in the study area.

18. line 394-399. Why did you even bother all the effort to determine Topt from the in situ stations in the prior sections? Now you are not using those Topt values but instead finding new ones based on comparison of the SMAP data with the in situ data. This does not make sense in the flow of the manuscript.

Response:

Firstly, the ExpF method only need one parameter (T_{opt}) to estimate the subsurface soil moisture. In order to use the exponential filter method in areas with limited or no soil moisture observations, especially in high and cold mountainous areas, it is necessary to understand local controls on the estimation of T_{opt}. Therefore, we put more effort on analyzing T_{opt}.

Secondly, the T_{opt} from in-situ stations were used to test the alternative methods for estimation of T_{opt} in the study area. This analysis showed that the median value of T_{opt} is robust in the study area.

Finally, earlier studies revealed that T_{opt} highly depended on the sampling interval of soil moisture data (De Lange et al., 2008). As data with the longer time interval typically have higher T_{opt} values, T_{opt} of in-situ data are mostly lower than satellite retrievals (Albergel et al. 2008; Ford et al., 2014; González-Zamora et al., 2016). In this study, the in-situ soil moisture observations are daily scale (the mean and standard deviation of the time interval for the 35 sites are 1 days and 0.3 days, respectively). However, as SM estimation from spaceborne sensors is still challenging for the high and cold mountainous regions, data gaps exist for the SMAP_L3 data in our study area (the mean and standard deviation of the time interval for the 35 sites are 2.3 days and 1.6 days, respectively). Thus, the time interval of the SMAP_L3 data is larger than the in-situ observations. Thus, the T_{opt} of SMAP is different from the T_{opt} of in-situ observations. Consequently, to get the best estimation of subsurface soil moisture at large scale, the subsurface SWI was estimated through combination of SMAP surface soil moisture with the ExpF method (with the site-specific T_{opt} based on the best match between observations in terms of NSE).

19. line 400-406. Include NSE and RSR measures here. They are crucial for quantifying the mismatch between the SMAP SWI and the observed SWI values as shown in Fig. 11.

Response:

We have included the NSE and RSR here. (Line 310-317)

As expected, the estimation accuracy of subsurface SM decreased with depth. The ANOVA results showed that the subsurface SM estimation accuracy for layer 2 (median value of RSR=0.92, R=0.69, NSE=0.18) and profile SM (RSR=0.92, R=0.65, NSE=0.14) were significantly higher than for layer 4 (RSR=1.12, R=0.31, NSE=-0.13) and layer 5 (RSR=1.17, R=0.34, NSE=-0.15) ($p<0.05$). The NSE values were positive for layer 2 and profile SM, while the NSE values for the other layers were negative. The negative MBE showed that subsurface SM was underestimated. The relationship between SMAP-derived and in-situ observed subsurface SM for layer 2 and profile SM was significant ($p<0.01$) at all but one station (D15). Thus, the SMAP surface product and ExpF method can be used to estimate the subsurface SM in the study area, especially for layer 2 (10-20 cm) and profile (0-70 cm) SM.

20. line 425-428. This point should also have been made in Section 4.4.1.

Response:

We have added the error caused by the scale mismatch between point measurements and satellite footprints in Section 4.3.1 of the revised manuscript. (Line 291-294)

Here, it is important to note that the SMAP_L3 product is provided at a 9 km × 9 km resolution while the in-situ measurements are point-based and soil moisture has a strong spatial variability in mountainous areas (Tian et al., 2019). Thus, part of the variability in Fig. 6 is due the disparity of spatial scales between the point-scale and the satellite footprint (Jin et al. 2017).

21. Tables 5 and 6. Replace RMSE with RSR. Add NSE.

Response:

We have replaced the RMSE with RSR and added the NSE in Tables 4 and 5. The analysis of the results was also changed according to the metrics of RSR, R and NSE in the revised manuscript. (Table 4 and Table 5, as the original Table 4 is the results of regression equation for T_{opt}, it has been deleted in the revised manuscript)

Table 4. Statistics of the metrics (RSR, R, NSE) of the comparisons of SMAP estimated and observed SWI at different layers for the 35 stations during the growing seasons of 2015-2016.

Layer	RSR		R		NSE	
	Mean±Std	Median	Mean±Std	Median	Mean±Std	Median
Layer 2 a	1.21±1.27	0.90	0.64±0.24	0.73a	0.04±0.49	0.19
Layer 3 ab	1.27±0.93	1.06	0.55±0.31	0.64ab	-0.05±0.46	0.05

Layer 4 b	1.53±1.49	1.10	0.43±0.38	0.52b	-0.22±0.53	-0.08
Layer 5 b	2.03±3.48	1.16	0.41±0.39	0.55b	-0.27±0.61	-0.14
Profile a	1.18±0.74	0.88	0.63±0.27	0.72a	0.12±0.38	0.19

Note: the different letters after the layers indicate that the difference is significant at $p<0.05$ (Kruskal-Wallis ANOVA)

Table 5. Statistics of the metrics (RSR, R, NSE) of the comparisons of estimated-observed profile SWI datasets, SMAP_L3-observed surface SWI datasets, SMAP_L3-observed profile SWI datasets, and SMAP_L4-observed profile SWI datasets for the 35 stations during the growing season of 2015-2016.

Comparisons	RSR		R		NSE	
	Mean±std	Median	Mean±std	Median	Mean±std	Median
Estimated-observed PSWI	0.86±1.00	0.68	0.88±0.11	0.9	0.56±0.32	0.64
SMAP_L3-observed SSWI	1.13±0.49	1.01	0.57±0.17	0.59	-0.09±0.52	-0.07
SMAP_L3-observed PSWI	1.18±0.74	0.88	0.63±0.27	0.72	0.12±0.38	0.19
SMAP_L4-observed PSWI	1.42±0.76	1.25	0.47±0.31	0.55	-0.49±0.68	-0.3

Note: e.g. Estimated-observed PSWI means the comparison of the estimated profile SWI and observed profile SWI.

Fig. S8 and Table 5 show that the SMAP-observed SWI comparisons had lower performance metrics for surface SWI (median value of RSR, R and NSE are 1.01, 0.59 and -0.07, respectively) than for profile SWI (median value of RSR, R and NSE are 0.88, 0.72 and 0.19, respectively). A possible reason for this is that the profile SWI was estimated based on the SMAP surface SWI and T_{opt}. The latter was determined by optimization using the maximum NSE, which improved the performance of profile SWI estimation. In addition, the performance metrics for SMAP–observed SWI comparisons for both surface and profile SWI were significantly ($p<0.001$) lower than those of estimated–observed profile SWI (median value of RSR, R and NSE are 0.68, 0.90 and 0.64, respectively). Thus, the major error in SMAP-based profile SWI estimates stems from the SMAP satellite product and is not derived from ExpF method, which is also supported by other researches (e.g. Ford et al., 2014; Pablos et al., 2018). Notably, the scale mismatch between point measurements and satellite footprints will introduce additional errors in the validation of the satellite estimated subsurface products (Jin et al. 2017).

22. line 436. Your results (Fig. 11) show that the performance of the SMAP profile SWI estimates is relatively poor. This is being partly obscured by the omission of the NSE and RSR statistics.

Response:

Thank you for your comments and suggests. We have added NSE and RSR in the revised manuscript. As shown in Tables 4 and 5, the SMAP estimated profile SWI has a median value of RSR=0.88, R=0.72 and NSE=0.19, respectively. Although the NSE is low, its performance is significantly (p<0.001) higher than the SMAP_L4 profile SWI in our study area.

23. line 451-452. Here on the 22nd page of the manuscript a completely new data set is introduced. This is inappropriate. If this section is important to the manuscript, then take the time to justify it in the introduction and describe it in the methods.

Response:

We have deleted this part in the revised manuscript.

24. line 465. Interpolated how? What evidence do you have that the interpolation is statistically valid? What is the associated uncertainty? This again should be justified in the introduction and described in the methods.

Response:

As suggested in comment 25, we can use the median value of T_{opt} instead of using the site-specific T_{opt} value. Thus, the median value of T_{opt} (T_{opt}=10 days) was used to derive ExpF method in estimating the profile SWI in the study area in the revised manuscript. We have deleted the content related to the interpolation of T_{opt} and the Fig. S14 (spatial distribution of T_{opt}).

25. line 465. Also, why bother to spatially interpolate Topt? You have just argued that Topt defined in one region (Heihe) is valid in another region (Maqu).

Response:

As argued in Section 4.3.3, the median value of T_{opt} can be used to derive the ExpF method for estimating profile soil moisture. For the estimation of profile SWI from SMAP surface product, we used the median value of T_{opt} (T_{opt}=10 days) instead of the interpolated T_{opt} to derive the ExpF method in the revised manuscript.

26. line 495. The data in Fig. 11 show that the accuracy is relatively poor. Relying on the R value alone is clearly misleading in this case where there is a substantial bias. Including the NSE and RSR as suggested above will likely show that the performance is not very good.

Response:

We have included NSE and RSR in the analysis (Tables 4 and 5 in the revised manuscript). This is now mentioned in the revised manuscript accordingly. (Line 382-384)

4) The ExpF method is useful and potential for estimating profile SM from SMAP_L3 surface product. As it can significantly improve the estimation of profile SM in cold arid mountainous areas (e.g. compared to the SMAP_L4 root zone product), and the main error stems from the satellite product.

Reference:

Albergel C, Rüdiger C, Pellarin T et al. From near-surface to root-zone soil moisture using an exponential filter: an assessment of the method based on in-situ observations and model simulations. Hydrology and Earth System Sciences, 2008, 12(6): 1323-1337.

Fan L, Al-Yaari A, Frappart F et al. Mapping Soil Moisture at a High Resolution over Mountainous Regions by Integrating In Situ Measurements, Topography Data, and MODIS Land Surface Temperatures. Remote Sensing, 2019, 11(6): 656.

Ford T W, Harris E, Quiring S M. Estimating root zone soil moisture using near-surface observations from SMOS. Hydrology and Earth System Sciences, 2014, 18(1): 139-154.

González-Zamora Á, Sánchez N, Martínez-Fernández J, Wagner W. Root-zone plant available water estimation using the SMOS-derived soil water index. Advances in Water Resources, 2016, 96: 339-353.

Jin R, Li X, Liu S M. Understanding the Heterogeneity of Soil Moisture and Evapotranspiration Using Multiscale Observations From Satellites, Airborne Sensors, and a Ground-Based Observation Matrix. IEEE Geoscience and Remote Sensing Letters, 2017, 14(11): 2132-2136.

Lange R d, Beck R, Giesen N v d, Friesen J, Wit A d, Wagner W. Scatterometer-Derived Soil Moisture Calibrated for Soil Texture With a One-Dimensional Water-Flow Model. IEEE transactions on Geoscience and remote sensing, 2008, 46(12): 4041-4049.

Li J, Zhang L, He C, Zhao C. A Comparison of Markov Chain Random Field and Ordinary Kriging Methods for Calculating Soil Texture in a Mountainous Watershed, Northwest China. Sustainability, 2018, 10(8): 2819.

Li X, Liu S, Xiao Q et al. A multiscale dataset for understanding complex eco-hydrological processes in a heterogeneous oasis system. Scientific data, 2017, 4: 170083.

Montzka C, Rötzer K, Bogena H R, Sanchez N, Vereecken H. A New Soil Moisture Downscaling Approach for SMAP, SMOS, and ASCAT by Predicting Sub-Grid Variability. Remote Sensing, 2018, 10(3): 427.

Pablos M, González-Zamora Á, Sánchez N, Martínez-Fernández J. Assessment of Root Zone Soil Moisture Estimations from SMAP, SMOS and MODIS Observations. Remote Sensing, 2018, 10(7): 981.

Tian J, Zhang B, He C, Han Z, Bogena H R, Huisman J A. Dynamic response patterns of profile soil moisture wetting events under different land covers in the Mountainous area of the Heihe River Watershed, Northwest China. Agricultural and Forest Meteorology, 2019, 271(15): 225-239.

Tian J, Zhang B, He C, Yang L. Variability in Soil Hydraulic Conductivity and Soil Hydrological Response Under Different Land Covers in the Mountainous Area of the Heihe River Watershed, Northwest China. Land Degradation & Development, 2017, 28(4): 1437-1449.

Wagner W, Lemoine G, Rott H. A Method for Estimating Soil Moisture from ERS Scatterometer and Soil Data. Remote Sensing of Environment, 1999, 70(2): 191-207.

Wang T, Franz T E, You J, Shulski M D, Ray C. Evaluating controls of soil properties and climatic conditions on the use of an exponential filter for converting near surface to root zone soil moisture contents. Journal of Hydrology, 2017, 548: 683-696.