Response to Comments by Referee 1 for the Manuscript gmd-2021-333
"**Optimization of Snow-Related Parameters in Noah Land Surface Model (v3.4.1) Using Micro-Genetic Algorithm (v1.7a)**"
by Sujeong Lim, Hyeon-Ju Gim, Ebony Lee, Seungyeon Lee, Won Young Lee, Yong Hee Lee, Claudio Cassardo, and Seon Ki Park

This procedure could potentially improve weather forecast in South Korea so the study has some practical implication. Regarding the optimization design I have a few questions and which should be clarified. 1) Are these six snow related parameters sensitive or not? Usually sensitivity analysis should be done first and next step is to use some schemes to optimize sensitive parameters. 2) OPT-5 and OPT-W were done separately. If only one Wmax is optimized in OPT_W, I wonder how will it behave when interacting with other parameters in OPT_6? Are these optimized values still be the best when used in OPT_6? I don't understand the rational of separating Wmax during optimization. Keep in mind that all these parameters are interrelated and together they affect the physical processes. Also one minor issue is that there are RMSDs appeared in many places. I assume that it is typo. Also Statistics in several panels in Figure 5 are hard to read. It's better to provide a table that lists optimized parameters, and their physical meanings.

⇒ We appreciate the valuable and constructive comments, which helped us improve the quality of the manuscript. An item-by-item response to the comments is provided below.

1. *Are these six snow related parameters sensitive or not? Usually sensitivity analysis should be done first and next step is to use some schemes to optimize sensitive parameters.*

   ⇒ We prepared an additional figure to explain the sensitivity of six parameters to snow variables (Figure R1 (Figure 1 in the revised manuscript)). According to the parameter ranges used in each optimization process, the variations of each snow variable (e.g., fractional snow cover (FSC), snow albedo (SA), and snow depth (SD)) are shown. We included additional description (written in blue fonts) in the revised manuscript with Figure R1 (Figure 1 in the revised manuscript) as follows:

   - L82: "It is noteworthy that P_s has a positive correlation with snow cover (Fig. 1(a))."

   - L86-87: "The SWE threshold, W_{max}, has a negative correlation with snow cover, as shown in Eq. (1) and it is more sensitive compared to P_s within a given parameter's range (Fig. 1(b))."

   - L113-115: "We optimize two empirical parameters that show positive relation to SA — $\alpha_{max,CofE}$ and C, whose default values are 0.85 and 0.5, respectively (Fig. 1(c)-(d)): SA shows similar sensitivities to both parameters within the same range but is a bit more sensitive $\alpha_{max,CofE}$."

- L126-127: "Because snow density is inversely proportional to SD, both P_1 and P_2 have negative correlations with the SD (Fig. 1(e)-(f)), where SD shows similar sensitivities to both parameters."

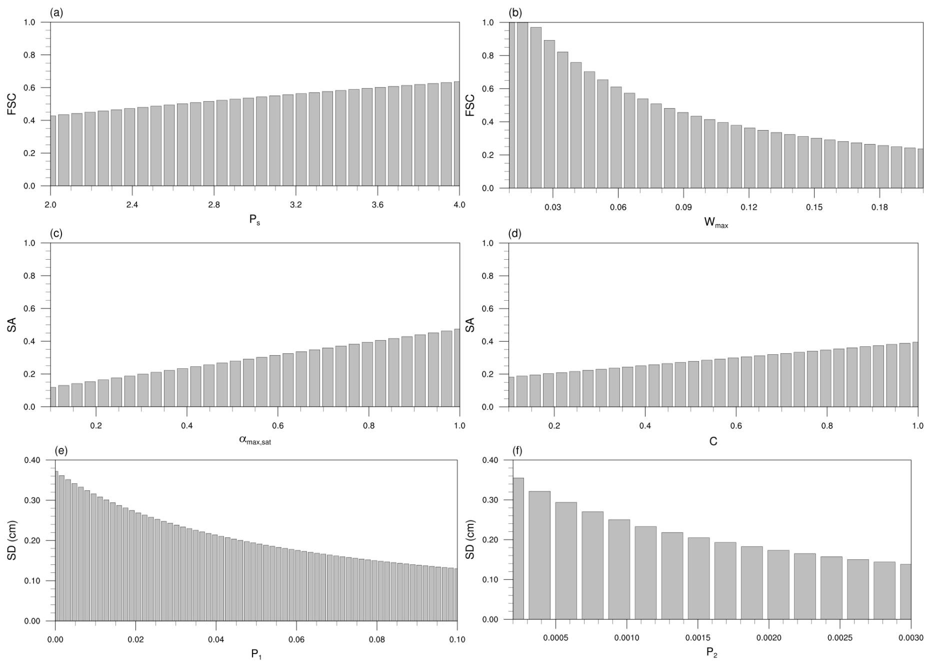

Figure R1 (Figure 1 in the revised manuscript): Responses of the snow variables to the variations in the snow-related parameters for given ranges: (a, b) Responses of FSC, for $W_s = 0.02$, to variations in P_s (with $W_{max} = 0.08$) and in W_{max} (with $P_s = 2.6$), respectively; (c, d) Responses of SA, for $\alpha_{max,sat} = 0.2$ and $t = 10$ days, to variations in $\alpha_{max,CofE}$ (with $C = 0.5$) and in C (with $\alpha_{max,CofE} = 0.85$), respectively; and (e, f) Responses of SD (in cm), for $W_s = 0.02$ and $T_{air} =$ -5 °C), to variations in P_1 (with $P_2 = 0.0017$) and in P_2 (with $P_1 = 0.05$ g cm^{-3}).

2. *OPT_5 and OPT_W were done separately. If only one Wmax is optimized in OPT_W, I wonder how will it behave when interacting with other parameters in OPT_6? Are these optimized values still be the best when used in OPT_6? I don't understand the rational of separating Wmax during optimization. Keep in mind that all these parameters are interrelated and together they affect the physical processes.*

   ⇒ W_{max} is the only parameter that relies on the land cover types (LCTs) while the other parameters (i.e., P_s, C, $\alpha_{max,sat}$, P_1, and P_2) are independent to LCTs. In optimization, we do not need to consider the dependence on LCTs for the five parameters other than W_{max}; however, we should definitely consider the dependence of W_{max} on LCTs. Thus, we have designed the experiments OPT_5 and OPT_W separately.

In addition, we have conducted new verification experiments by increasing the number of observation stations from 10 (stations per LCT) to 25 (5 stations per LCT). In the revised manuscript, the GA optimization experiments are expressed with "OPT" while the verification experiments are expressed with "VRF". We have conducted the following verification experiments for the 25 observation stations: 1) CNTL using non-optimized (i.e., default) parameters; 2) VRF_5 using the five optimized parameters obtained from OPT_5; and 3) VRF_6 using the six optimized parameters obtained from both OPT_5 and OPT_W. With the increased number of stations, we expect to have more reasonable verification statistics over South Korea. We modified the experimental design and related figure or table caption in the revised manuscript as follows:

- L225-233: "We have designed the following two GA optimization experiments: 1) OPT_5 that optimizes five snow parameters (P_s, $\alpha_{max,CofE}$, C, P_1, and P_2); and 2) OPT_W that optimizes W_{max}. Among the six parameters, W_{max} is the only parameter that depends on the LCTs; thus, we conducted OPT_5 and OPT_W separately. Note that SK is represented by five different LCTs considering the sufficient days of snowfall and ASOS observation (see Table 1). Because OPT_5 optimizes with more parameters and generations, we have selected 10 stations (i.e., 2 stations per LCT) based on snowfall amount to reduce the computation time. To investigate the performance of snow prediction through optimized snow parameters, we have designed the following three verification experiments for the 25 observation stations: 1) CNTL using non-optimized (i.e., default) parameters; 2) VRF_5 using the five optimized parameters obtained from OPT_5; and 3) VRF_6 using the six optimized parameters obtained from both OPT_5 and OPT_W (see Fig 3(b)). "

- Table 1 caption: "Five representative LCTs over SK, following the IGBP classification — DBF, MF, WS, CL, and UB. For each LCT, five selected stations are shown with the station name (abbreviation in parenthesis), location in latitude (°N) and longitude (°E), ratio of LCT in 2.5 km buffer (%), soil type, and missing ratio (%). The experiment OPT_5 employs only the stations highlighted in bold while the other experiments use all the stations"

- Figure 3 caption: "Stations used for the experiments (a) OPT_5 and (b) OPT_W, CNTL, VRF_5 and VRF_6. Different colors in the station acronyms represent different LCT: DBF (black), MF (blue), WS (green), CL (yellow), and UB (red). See Table 1 for the acronyms of stations and LCTs."

3. *Also one minor issue is that there are RMSDs appeared in many places. I assume that it is typo.*

   ⇒ We corrected the "RMSDs" to "RMSEs" in the revised manuscript.

4. *Also Statistics in several panels in Figure 5 are hard to read. It's better to provide a table that lists optimized parameters, and their physical meanings.*

   ⇒ We appreciate this comment. Following the reviewer's suggestion, we listed the statistics in the Table R1 (Table 5 in the revised manuscript) and we removed the statistics in Figure 5 in the revised manuscript (see Fig. R2 (Figure 6 in the revised manuscript)). These changes are described in the revised manuscript in L335-352.

   "To understand more details of the improvements due to the optimization, we analyzed the scatter plots that compare the observations and the model results in Figure 6 and listed their RMSE and R^2 in Table 5. Since the observation patterns are different for different stations, we selected the representative station for each LCT. For FSC, it is relatively hard to recognize the explicit bias patterns, as shown in Fig. 6 (left panels); however, compared to CNTL, the RMSE decreased in VRF_5 and further reduced in VRF_6 (see Table 5). The VRF_6 revealed the largest R^2 values over most LCTs, except WS (station NG) and CL (station BR). In particular, VRF_6 produced the highest FSC over MF (station GM) (see Fig. 6(d)) with the smallest RMSE and the largest R^2, which significantly alleviated the underestimation problem. For SA, its overestimation in CNTL has been prominently reduced in both VRF_5 and VRF_6 — see Fig. 6 (middle panels). For instance, SA decreased over DBF (station UL) in both VRF_5 and VRF_6, with a larger decrease VRF_6 (Fig. 6(b)). The performance statistics of both VRF_5 and VRF_6 demonstrated improvements over most LCTs except UB (station SL) (see Table 5). For SD, the parameter optimization brought about remarkable improvement compared to FSC and SA — see Fig. 6 (right panels). Note that SD is optimized using the hourly in-situ observations (i.e., larger amount of data) while both FSC and SA are optimized using the daily satellite observations. For example, VRF_6 with DBF produced notably large SD values (Fig. 6(c)) with the lowest RMSE and the highest R^2 (Table 5), diminishing the underestimation problem in CNTL. It is hard to say which verification experiment gives the best results (i.e., VRF_5 versus VRF_6), but the performance with optimized parameters is usually better than CNTL in terms of RMSE (e.g., for most LCTs such as DBF, MF, WS, UB) and R^2 (e.g., for LCTs including DBF, MF, and CL). Overall, both VRF_5 and VRF_6 produced snow variables that are closer to observations than CNTL for most LCTs (i.e., stations), and VRF_6 generally showed the lowest RMSE and the highest R^2 in all the snow variables."

Table R1 (Table 5 in the revised manuscript): Statistics of model performance using non-optimized parameters (CNTL) and optimized parameters (VRF_5 and VRF_6) over different LCTs represented by different stations — DBF represented by UL, MF by GM, WS by NG, CL by BR, and UB by SL. The RMSEs and R^2 values are shown for three snow variables — FSC, SA, and SD.

Statistics		RMSE			R^2		
LCT	Snow Variable	CNTL	VRF_5	VRF_6	CNTL	VRF_5	VRF_6
DBF (UL)	FSC	0.328	0.327	0.252	0.248	0.215	0.256
	SA	0.218	0.197	0.159	0.157	0.157	0.176
	SD	15.763	13.640	12.616	0.764	0.781	0.796
MF (GM)	FSC	0.208	0.206	0.178	0.388	0.408	0.520
	SA	0.105	0.103	0.103	0.411	0.421	0.460
	SD	1.789	1.526	1.542	0.435	0.502	0.493
WS (NG)	FSC	0.279	0.269	0.249	0.354	0.333	0.341
	SA	0.196	0.160	0.156	0.314	0.328	0.324
	SD	9.836	8.231	8.009	0.895	0.887	0.888
CL (BR)	FSC	0.163	0.160	0.160	0.363	0.385	0.384
	SA	0.132	0.122	0.122	0.443	0.457	0.456
	SD	2.542	2.583	2.590	0.478	0.540	0.539
UB (SL)	FSC	0.255	0.252	0.242	0.184	0.195	0.195
	SA	0.071	0.070	0.073	0.150	0.148	0.124
	SD	4.790	4.286	4.699	0.484	0.449	0.385

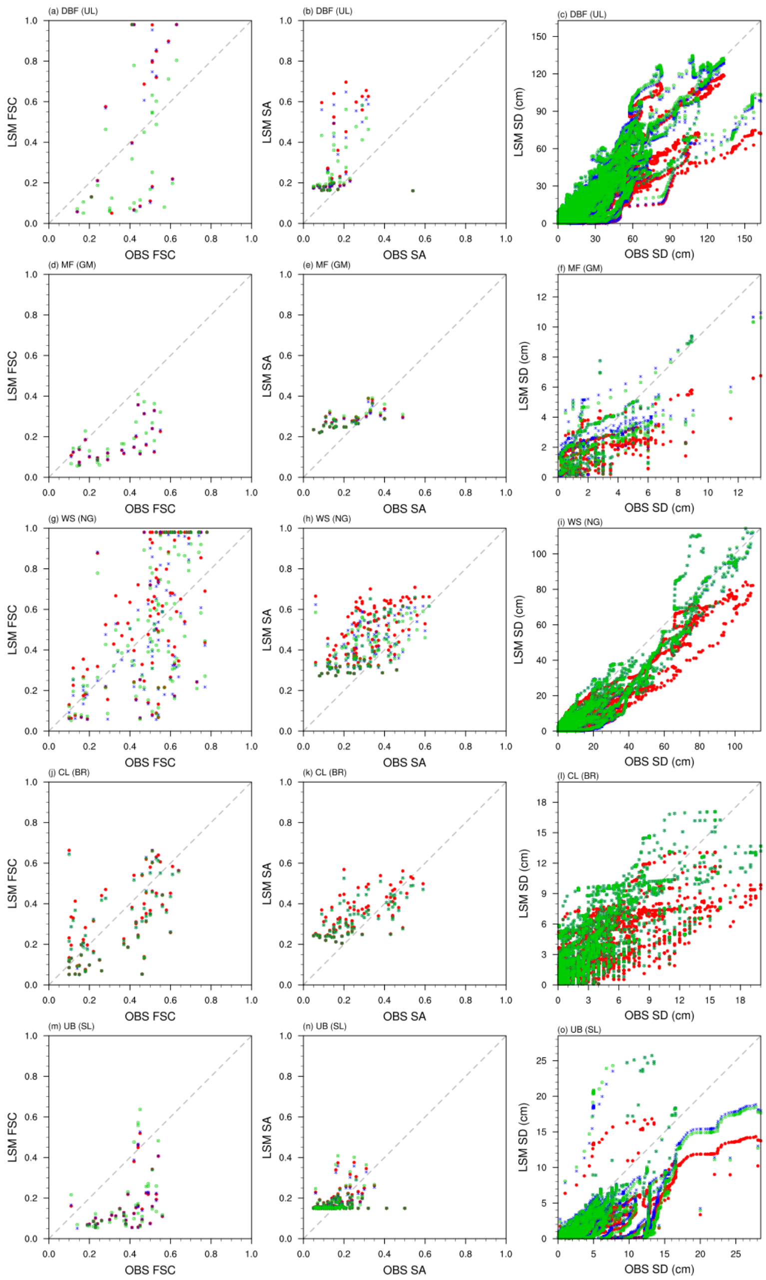

Figure R2 (Figure 6 in the revised manuscript): Scatter plots of observations (OBS) and model results (LSM) for snow variables FSC (left panels), SA (middle panels), and SD (in cm; right panels) from the verification experiments — CNTL (red dots), VRF_5 (blue dots), and VRF_6 (green dots), which are evaluated over different LCTs; (a–c) DBF represented by the station UL, (d–f) MF by GM, (g–i) WS by NG, (j–l) CL by BR, and (m–o) UB by SL.

Response to Comments by the Referee 2 for the Manuscript gmd-2021-333
"**Optimization of Snow-Related Parameters in Noah Land Surface Model (v3.4.1) Using Micro-Genetic Algorithm (v1.7a)**"
by Sujeong Lim, Hyeon-Ju Gim, Ebony Lee, Seungyeon Lee, Won Young Lee, Yong Hee Lee, Claudio Cassardo, and Seon Ki Park

The manuscript "Optimization of Snow-Related Parameters in Noah Land Surface Model (v3.4.1) Using Micro-Genetic Algorithm (v1.7a)" by Lim et al. This is my second time review. Authors have addressed some of the concerns, however I think the manuscript needs serious improvements before it is accepted for publication. The following are the comments, which may improve the manuscript. Still, the results are not promising.

⇒ We appreciate the valuable and constructive comments, which helped us improve the quality of the manuscript. An item-by-item response to the comments is provided below.

1. *Table 4: In terms of correlation, all improvements are in second decimal. For example for FSC R^2=0.219 (r 0.467) and 16.4% improvements in OPT_6 will be R^2=0.255 (r 0.50), which is the highest improvements. Is change of r by about 0.037 significant? I am not sure. Unfortunately authors have tried to mislead their study by not bringing out meaningful results.*

   ⇒ We appreciate the reviewer pointing this out. In the revised manuscript, we have increased the number of stations from 10 (Table R1 (Table 4 in the original manuscript)) to 25 (Table R2 (Table 4 in the revised manuscript)) for verification experiments (VRF_5 and VRF_6). The optimization still affects the second decimal places, but the difference between CNTL and the verification experiments has 95% statistical significance, as evaluated with a two-tailed t-test. Rather than showing the improvement ratio only, we have included the original values of RMSE, MB, and R^2 and put the improvement ratio in the parentheses of Table R2 (Table 4 in the revised manuscript) and explained that the optimization is mostly effective in RMSE and MB. We have included the related descriptions (written in blue font) in L312-317 with a replaced Table R2 (Table 4 in the revised manuscript) in the revised manuscript.

   "We also investigated R^2, which measures the proportion of variation for a dependent variable that can be explained by an independent variable. Although the R^2 values are low in FSC and SA, the difference between CNTL and verification experiment (e.g., VRF_5) has 95% statistical significance, as evaluated with a two-tailed t-test. After optimization, the R^2 values in VRF_5 improve by 3.3 % and 1.5 % for FSC and SD, respectively. However, these changes are insignificant compared to the other statistics such as RMSE and MB."

Table R1 (Table 4 in the original manuscript): The RMSE, MB, R^2 of snow variables and improvement ratio (%) in parentheses from CNTL to OPT_5, and OPT_6 over the ten representative stations. The difference between CNTL and OPTM (i.e., OPT_5 and OPT_6) experiments has 95% statistical significance, as evaluated with a two-tailed t-test.

EXP	CNTL_5			OPT_5			OPT_6		
Snow variable	FSC	SA	SD	FSC	SA	SD	FSC	SA	SD
RMSE	0.270	0.155	10.599	0.267 (1.3 %)	0.145 (6.7 %)	9.141 (13.8 %)	0.253 (6.5 %)	0.142 (8.5 %)	8.722 (17.7 %)
MB	-0.107	0.0513	-5.38	-0.125 (-16.8 %)	0.0381 (25.7 %)	-3.46 (35.7 %)	-0.130 (-21.5 %)	0.0359 (30.0 %)	-2.93 (45.5 %)
R^2	0.219	0.183	0.806	0.226 (3.1 %)	0.179 (-2.4 %)	0.819 (1.6 %)	0.183 (16.4 %)	0.428 (-0.2 %)	0.830 (3.0 %)

Table R2 (Table 4 in the revised manuscript): The RMSE, MB, R^2 of snow variables and improvement ratio (%) in parentheses from CNTL to VRF_5, and VRF_6 over the 25 representative stations. The difference between CNTL and verification experiments (i.e., VRF_5 and VRF_6) has 95% statistical significance, as evaluated with a two-tailed t-test.

EXP	CNTL			VRF_5			VRF_6		
Snow Variable	FSC	SA	SD	FSC	SA	SD	FSC	SA	SD
RMSE	0.249	0.132	9.094	0.247 (0.7 %)	0.125 (5.4 %)	7.847 (13.7 %)	0.124 (3.3 %)	0.125 (6.2 %)	7.547 (17.0 %)
MB	-0.133	0.0408	-4.39	-0.145 (-9.1 %)	0.0298 (26.9 %)	-2.81 (35.9 %)	-0.149 (-11.9 %)	0.0281 (31.0 %)	-2.45 (44.2 %)
R^2	0.257	0.281	0.808	0.265 (3.3 %)	0.276 (-1.7 %)	0.821 (1.5 %)	0.277 (8.0 %)	0.274 (-2.2 %)	0.834 (3.0 %)

2. *Similarly in Figure 6, one would be interested to see the difference between observations and model simulation, and not only the difference between control and improved version of model. This makes the improvements further questionable.*

   ⇒ We added a new figure that compares the time series of snow variables for DBF, represented by UL, in the observations (black dots) and in the model simulations — CNTL (red dots) and VRF_6 (green dots). The CNTL shows positive or negative biases in FSC, positive bias (overestimation) in SA, and negative bias (underestimation) in SD: these biases are all reduced down in VRF_6. These bias patterns are consistent with Figure 5 of the original manuscript, which represents the scatter plots of observation and model results. We modified the manuscript in L353-356 and L357-361 as follows:

   ⇒ (L353-356): "Figure 7 compares the time series of snow variables between the observations and the model simulations — CNTL and VRF_6 — for DBF represented by UL. The CNTL shows positive or negative biases in FSC, positive bias (overestimation) in SA, and negative bias (underestimation) in SD: these biases are all reduced down in VRF_6. The bias patterns in Fig. 7 are consistent with those in Fig. 6."

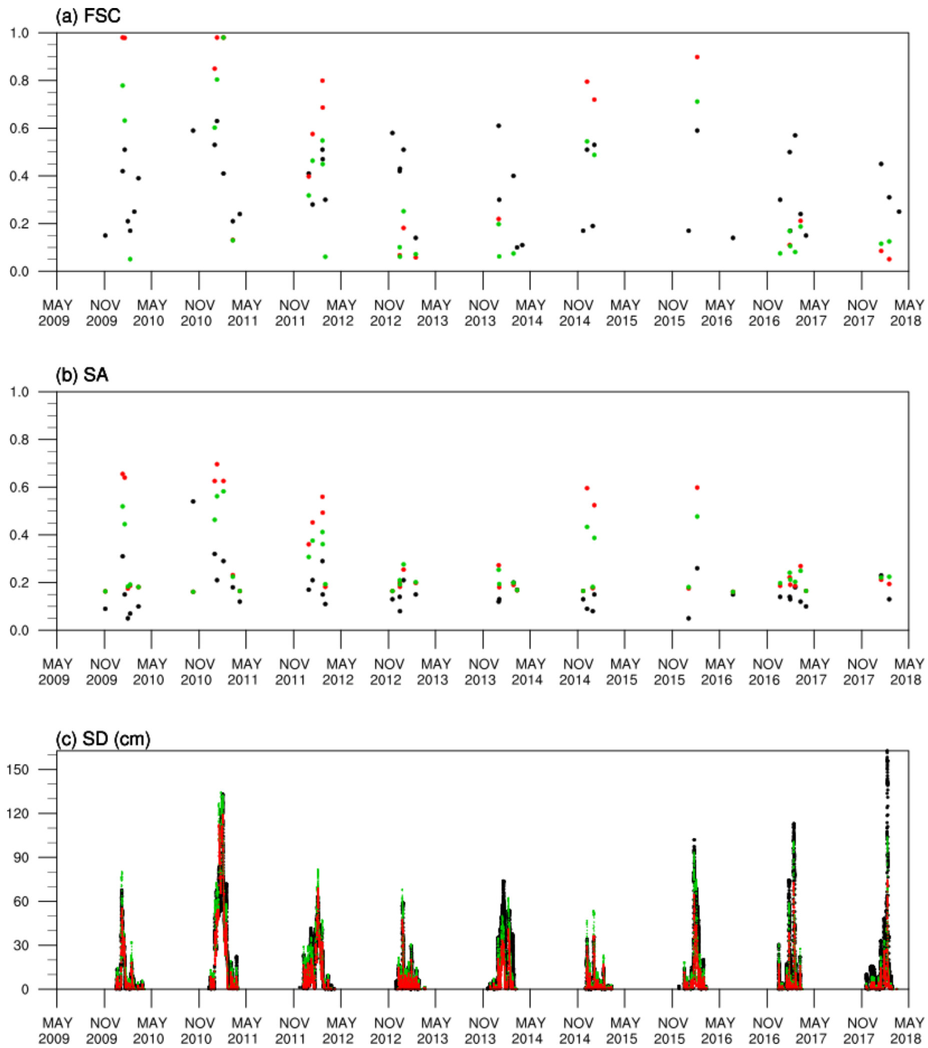

Figure R1 (Figure 7 in the revised manuscript): Time series of the snow variables for DBF from May 2009 to April 2018: (a) FSC, (b) SA, and (c) SD (in cm). Observations are in black dots and model results are in red dots for CNTL and in green dots for VRF_6.

(L357-361): "Lastly, we have investigated how the optimized snow parameters can affect the other variables in LSM. Figure 8 depicts the time series of the differences of LSM variables (soil temperature, sensible heat flux, and soil moisture) between VRF_6 and CNTL (i.e., VRF_6 minus CNTL) following the changes in SD. Although the LSM variables here are not directly optimized, they respond to the optimized snow parameters through associated physical processes. Note that the underestimation of SD in CNTL has been alleviated in VRF_6 by using the optimized snow parameters (see Fig. 7(c) and Fig. 8(a))."

3. *How the evolution of snow/land parameters (e.g. snow depth) looks like in model and observations. As shown in Figure 6, how the mean state varies with time (daily/monthly average) in model and observations. I can see/guess existence of a seasonal cycle of snow depth in Figure 6a, however authors in their reply have argued that "Snow parameters do not have the observations; thus, it is impossible to compare the snow-related parameters between model and observations."*

   ⇒ First, we would like to define the variables and parameters. A variable represents the model state (e.g., fractional snow cover (FSC), snow albedo (SA), and snow depth (SD)), thus it can vary with time. On the other hand, a parameter (P_s, W_{max}, C, $\alpha_{max,CofE}$, P_1, and $P2$) is a constant that composes the equation in the simulation, thus it is usually fixed during the simulation. The parameter can adjust when the user wants to change the equation in the simulation. Furthermore, a parameter may or may not be measured whereas a variable is a measurable quantity.

   Therefore, we cannot prepare a time series of the snow parameter (i.e. constant) due to absent observation, but we have included the time series of the snow variable in the revised manuscript as in the reply of #2 above (Fig. R1 (Figure 7 in the revised manuscript)). To clarify the variable and parameter concept, we included their definition in the revised manuscript (L39-41):

   "Here, the parameter is a constant that makes up the equations, which is usually fixed during the simulation and differs from the variable representing the time-varying state of the model."

4. *Finally, authors need to show some improvements/results to be worth publication. I do not agree that offline Noah LSM takes so much time that one cannot do grid-point simulation over SK region. However, I am not insisting on this, but pointing out one of the possibilities to bring out some positive impact of this method/study. At this stage the results are not encouraging.*

   ⇒ We have conducted the following new verification experiments by increasing the number of observation stations from 10 (stations per LCT) to 25 (5 stations per LCT): 1) CNTL using non-optimized (i.e., default) parameters; 2) VRF_5 using the five optimized parameters obtained from

OPT_5; and 3) VRF_6 using the six optimized parameters obtained from both OPT_5 and OPT_W. In the revised manuscript, the GA optimization experiments are expressed with "OPT" while the verification experiments are expressed with "VRF". Although the 25 stations in South Korea (SK) are not evenly distributed, they cover SK similarly to the 0.5° resolution (e.g., 29 stations). With the increased number of stations, we expect to have more reasonable verification statistics over SK. We modified the experimental design and related descriptions including the figure or table in the revised manuscript as follows:

- L12-15 (Abstract): "Optimization of these six snow-related parameters has led to improvement in the root-mean squared errors by 17.0 %, 6.2 %, and 3.3 % on snow depth, snow albedo, and snow cover fraction, respectively. In terms of the mean bias, the underestimation problems of snow depth and overestimation problems of snow albedo have been alleviated through optimization of parameters calculating the fresh snow by about 44.2 % and 31.0 %, respectively."

- Table 1 caption: "Five representative LCTs over SK, following the IGBP classification — DBF, MF, WS, CL, and UB. For each LCT, five selected stations are shown with the station name (abbreviation in parenthesis), location in latitude (°N) and longitude (°E), ratio of LCT in 2.5 km buffer (%), soil type, and missing ratio (%). The experiment OPT_5 employs only the stations highlighted in bold while the other experiments use all the stations."

- Figure 3 caption: "Stations used for the experiments (a) OPT_5 and (b) OPT_W, CNTL, VRF_5 and VRF_6. Different colors in the station acronyms represent different LCT: DBF (black), MF (blue), WS (green), CL (yellow), and UB (red). See Table 1 for the acronyms of stations and LCTs. "

- L225-233 (3. Experimental design): "We have designed the following two GA optimization experiments: 1) OPT_5 that optimizes five snow parameters (P_s, $\alpha_{max,CofE}$, C, P_1, and P_2); and 2) OPT_W that optimizes W_{max}. Among the six parameters, W_{max} is the only parameter that depends on the LCTs; thus, we conducted OPT_5 and OPT_W separately. Note that SK is represented by five different LCTs considering the sufficient days of snowfall and ASOS observation (see Table 1). Because OPT_5 optimizes with more parameters and generations, we have selected 10 stations (i.e., 2 stations per LCT) based on snowfall amount to reduce the computation time. To investigate the performance of snow prediction through optimized snow parameters, we have designed the following three verification experiments for the 25 observation stations: 1) CNTL using non-optimized (i.e., default) parameters; 2) VRF_5 using the five optimized parameters obtained from OPT_5; and 3) VRF_6 using the six optimized parameters obtained from both OPT_5 and OPT_W (see Fig. 3(b)). "

- L290-334 (4. Results): We modified the result parts with the verification experiments over 25 stations. Because there are many correction in the manuscript, we have not included them here. Please check these paragraphs in the revised manuscript.

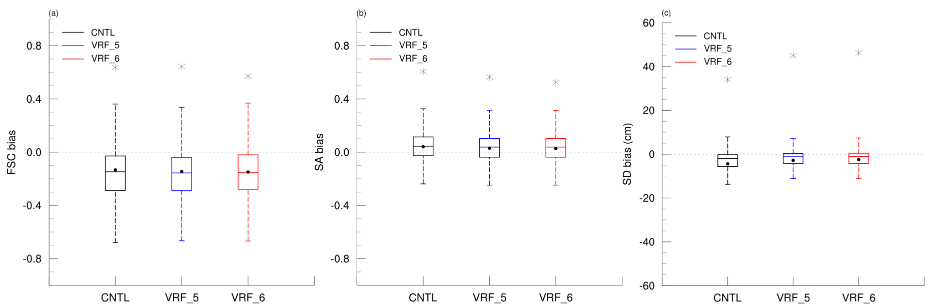

Figure R2 (Figure 5 in the revised manuscript): Box plots of (a) FSC bias, (b) SA bias, and (c) SD bias (cm) for CNTL, VRF_5 and VRF_6. The maximum differences are indicated with the black star symbol (e.g., 0.637 (CNTL), 0.643 (VRF_5), 0.570 (VRF_6) for FSC, 0.605 (CNTL), 0.563 (VRF_5), and 0.525 (VRF_6) for SA, and 34.1 cm (CNTL), 45.1 cm (VRF_5), and 46.3 cm (VRF_6) for SD). Each mean of snow variables is indicated as a black circle (e.g., -0.133 (CNTL), -0.145 (OPT_5), and -0.149 (VRF_6) for FSC, 0.0408 (CNTL), 0.0298 (VRF_5), and 0.0281 (VRF_6) for SA, and -4.39 cm (CNTL), -2.81 cm (VRF_5), and -2.45 cm (VRF_6) for SD).

- L376-377 (5. Discussion) "Our results showed improvement in all snow variables in terms of RMSE by 3.3 %, 6.2 %, and 17.0 % for FSC, SA, and SD, respectively."