# Evaluation of Noah-MP Snow and Soil Simulations

A final report to the DTC visitor program

**Principal Investigator**
Siwei He, Assistant Professor
Montana State University
205 Cobleigh Hall, Bozeman, MT 59717
Siwei.He@montana.edu
406-994-6113

**January 2026**

## 1 Objectives of the Visit

The purpose of this visit was to investigate the snow and soil simulations of the land surface model Noah-MP within NOAA's next-generation forecasting system, the Unified Forecast System (UFS). In addition, the visit aimed to build connections with the development teams of the UFS short-range regional weather prediction model, particularly the Rapid Refresh Forecast System (RRFS), as well as the Noah-MP development team.

## 2 Activities of the Visit

The original plan for this visit was to complete a two-month workload, split over a one-year period from May 6, 2024, to May 5, 2025. However, due to multiple unexpected changes, the visit ultimately spanned longer than originally planned, and the scope of the completed work shifted from one theme to three themes, as listed below.

1. Theme 1: Evaluate the snow and soil simulations of Noah-MP based on results from the Water-in-the-West project, a NOAA project aimed at developing prototype forecast systems for predicting atmospheric rivers (ARs) affecting the western United States. This was the originally planned theme in the proposal; however, only a preliminary evaluation was conducted due to changes in the Water-in-the-West project.
2. Theme 2: Evaluate the snow simulations of Noah-MP based on results from the Global Forecast System (GFS) v17 prototype simulations. The simulations were conducted by NOAA Environmental Modeling Center (EMC). This theme served as a temporary alternative to Theme 1.
3. Theme 3: Evaluate the snow and soil simulations of Noah-MP using offline simulations over the Montana Mesonet stations. Unlike the previous two themes, this effort required conducting offline simulations myself in order to generate evaluation results. The main reason for initiating Theme 3 was that only preliminary model simulations were available from Themes 1 and 2, with no additional simulations available for evaluation.

In addition to this report, the PI presented work on Theme 1 and Theme 2 at two conferences: the Unifying Innovations in Forecasting Capabilities Workshop 2024 (UIFCW 2024) and the American Geophysical Union Fall Meeting 2024 (AGU 2024). The PI also attended regular bi-weekly meetings of the Noah-MP development team after beginning work on Theme 3.

## 3 Main Conclusions

Through work on Themes 1 to 3, we conducted a quick assessment of the snow and soil performance of Noah-MP in the Water-in-the-West, GFS, and offline simulations. The main conclusions from this work are:

1. The model's performance can be influenced by a range of factors, including the initial conditions, forcing data, and configuration.
2. The snow data assimilation framework, Global Historical Climatology Network (GHCN) in GFSv17, improves snow simulations.
3. Offline Noah-MP simulations over Montana Mesonet stations show that the model captures the major snow and soil processes but tends to underestimate snow depth and melt out early. This work also shows that ERA5 precipitation is generally higher than in-situ observations over the Montana Mesonet stations.

4. Increasing the number of soil layers from 4 to 8 has minimal impact on snow simulations but has a noticeable effect on the simulation of surface heat fluxes, especially when the soil is unsaturated.

## 4 Theme 1: Noah-MP in the UFS

### *4.1 Water-in-the-West project*

The Water-in-the-West project is a NOAA project aimed at developing prototype forecast systems for predicting atmospheric rivers (ARs) affecting the western United States. The forecast system developed under this project is based on the UFS and is referred to as UFS-AR. UFS-AR includes two high-resolution configurations: a limited-area model (LAM) and a global nest. Both configurations have a horizontal resolution of 3 km and use the Common Community Physics Package (CCPP) physics suite including Noah-MP.

During the initial phase of the project, six AR cases from January to March 2023 were simulated: 2023010600, 2023010900, 2023011100, 2023022400, 2023030900, and 2023031300. In this study, we used simulations from nested domain 2, which has a spatial resolution of 3 km and output at a 3-hour interval.

### *4.2 Results*

This evaluation focused on Snow Water Equivalent (SWE) simulations, examining both their spatial and temporal distributions. Figure 1 compares the initial SWE from the simulations with SWE from SNODAS. The results show that the initial SWE of the simulations was substantially lower than SNODAS SWE, particularly over mountainous regions. Although Figure 1 presents only three examples, the same issue was observed across all six cases.

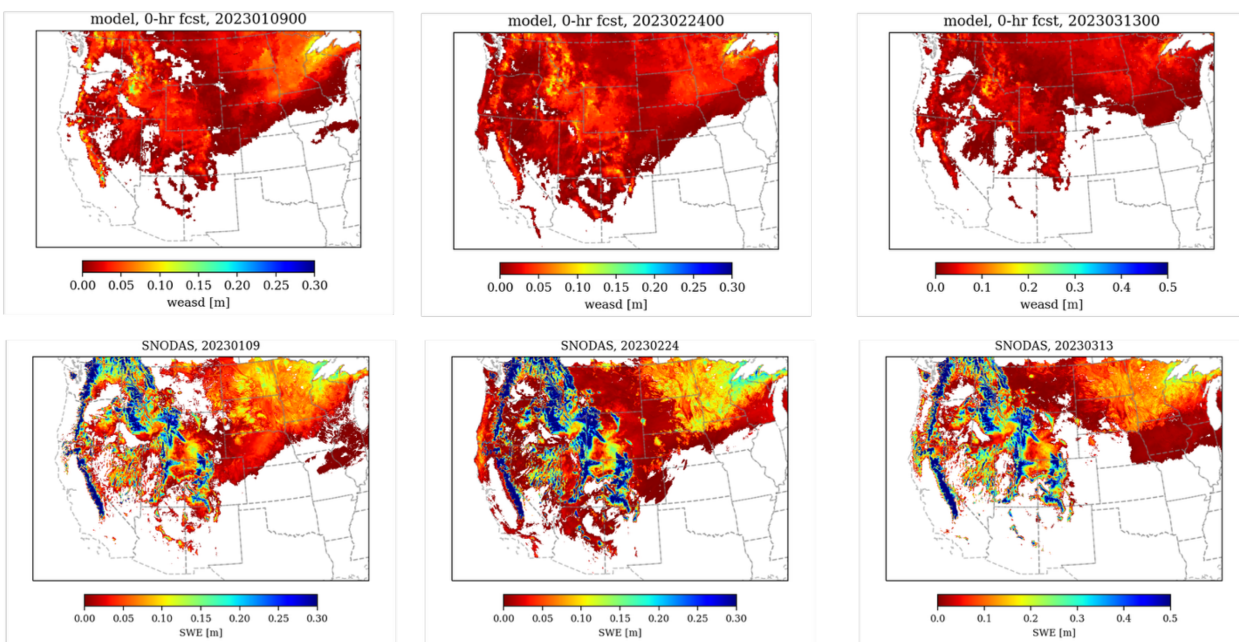


*Figure 1 Comparison of the model initial SWE (top panel) and SNODAS SWE (bottom panel) for three cases: January 9, 2023; February 24, 2023; and March 13, 2023.*

Figures 2 and 3 show comparisons of SWE with SNOTEL in-situ observations for the initial conditions and the 60-h forecasts. In these figures, histograms represent SWE bias between the model results and SNOTEL observations, while scatter plots show direct comparisons of SWE between the model and SNOTEL. The results clearly indicate that model initial SWE is substantially lower than the observations, with larger biases in March than in January. Furthermore, Figures 2 and 3 demonstrate that the SWE bias introduced in the initial conditions is not eliminated during the model simulations but persists through the 60-h forecast.

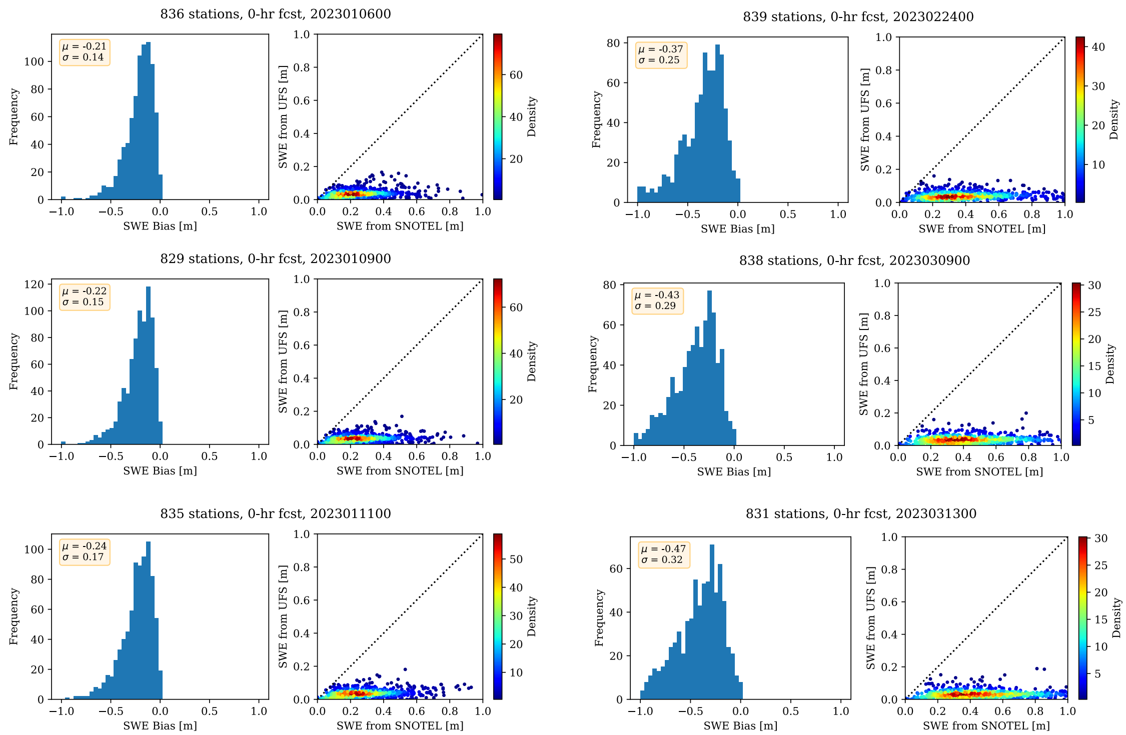


*Figure 2 Comparison of SWE from model simulations and in-situ SNOTEL observations for the initial conditions of six cases. In each panel, the histogram shows the SWE bias (model − SNOTEL), with mean bias and variance listed in the upper-left corner. The scatter plot shows simulated vs. observed SWE, with the 1:1 line as reference. Point color represents density, with higher density indicated by warmer colors (i.e., more frequent values).*

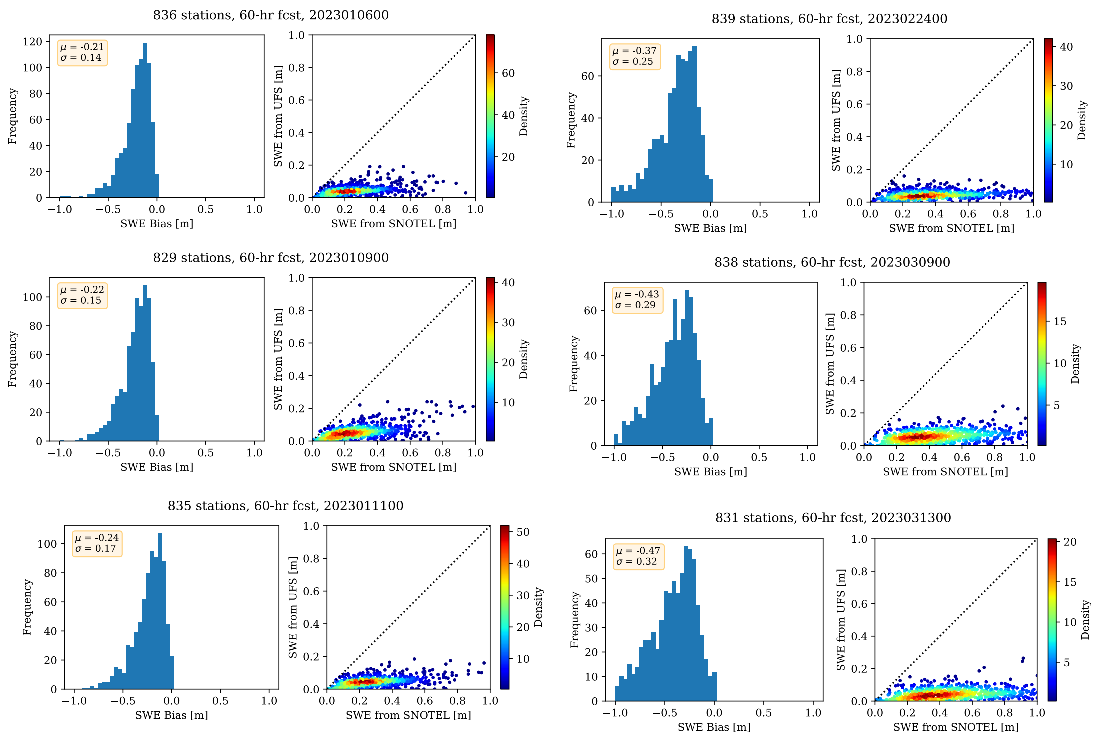


*Figure 3 The same as Figure 2, but for 60-h forecast.*

## 5 Theme 3: Offline NoahMP

### *5.1 Experiment setup*

In this study, we evaluated snow and soil simulations from the offline Noah-MP simulations over the Montana Mesonet stations. The version of NoahMP is from the UFS CCPP branch (de2f058d; https://github.com/NOAA-EMC/ufs-land-driver), and it is driven by ERA5 reanalysis forcing data. The simulation period spans from January 1, 2010, to December 31, 2024. The first few years of the simulation were treated as the model spin-up period, and the analysis focuses on the period from October 1, 2023, to September 30, 2024. In total, three experiments were conducted, and descriptions of these experiments are provided in Table 1.

*Table 1 Different configurations of the experiments for standalone Noah-MP simulations in this project*

| Experiments | Description |
|---|---|
| control | The default configuration |
| exp2 | Same as the control experiment, but with precipitation updated using Montana Mesonet observations. |
| exp3 | Same as the control experiment, but with the number of soil layers increased from 4 to 8 while keeping the total soil depth the same. |

### *5.2 Statistical indices*

We evaluated the snow and soil simulations of the offline Noah-MP model based on results from the control simulations. Two statistical indices were used for the evaluation: the correlation coefficient and the bias ratio:

$$\text{correlation coefficient} = \frac{\sum_{j=1}^{n}(x_{f,j}-\mu_{sim})(x_{o,j}-\mu_{obs})}{\sqrt{\sum_{j=1}^{n}(x_{f,j}-\mu_{sim})^2}\sqrt{\sum_{j=1}^{n}(x_{o,j}-\mu_{obs})^2}} \quad (1)$$

$$\text{bias ratio} = \frac{\mu_{sim}}{\mu_{obs}} \quad (2)$$

where $x_{o,j}$ is the j-th observed value and $x_{f,j}$ is the j-th simulated value, $n$ is the number of pairs in the comparison, $\mu_{obs}$ is the observation mean, and $\mu_{sim}$ is the simulation mean.

### *5.3 Results*

Snow depth simulations were evaluated at 48 Montana Mesonet stations from October 1, 2023, to May 31, 2024. Figure 4 presents the correlation coefficients and bias ratios of the simulated snow depth relative to in situ observations. The results show that correlation coefficients are reasonably high (greater than 0.8) at most stations, while the bias ratios indicate that the model systematically underestimates snow depth across the majority of stations.

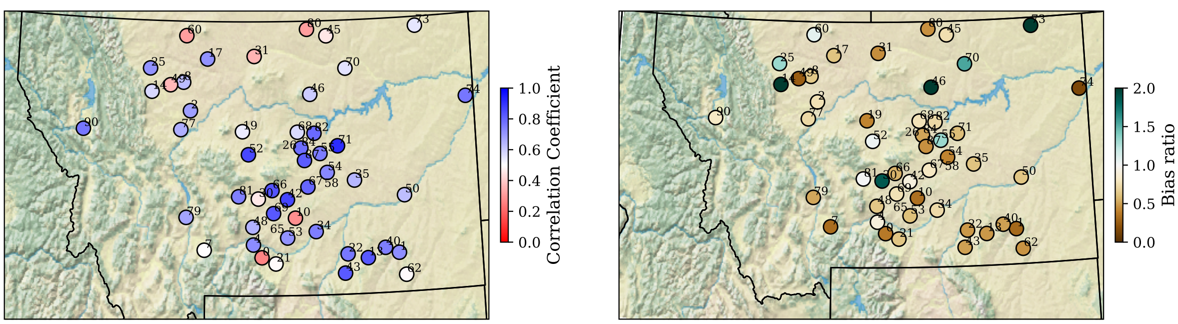


*Figure 4 Correlation coefficients (left) and bias ratios (right) of the simulated snow depth at Montana Mesonet stations. The evaluation period spans October 1, 2023, to May 31, 2024, which generally corresponds to the snow season in Montana.*

Figure 5 shows four examples of time series comparisons between observations and simulations of snow depth, which confirm the underestimation of snow depth in the model results. These four stations are well representative of the Montana Mesonet network, as they are distributed across a wide range of geographic settings. The stations include Bozeman (acebozem; western foothills; Station 7 hereafter), Busby South (acebusby; southern plains; Station 13 hereafter), Roundup (aceround; central plains; Station 67 hereafter), and Turner North (aceturne; northern plains; Station 80 hereafter). A camera view of each station is shown in Figure 6.

At the Station 7, the simulated snow depth is typically about 5 cm lower than the observations and melts out noticeably earlier. At Station 13, the simulated snow depth is slightly lower than observed, with melt-out dates occurring somewhat earlier. At Station 67, the modeled snow depth also melts out slightly earlier than observed. At Station 80, the melt-out dates are again noticeably earlier in the simulations.

To better understand the model's performance in snow simulations, the following sections present event-based analyses. In total, three events were selected for analysis.

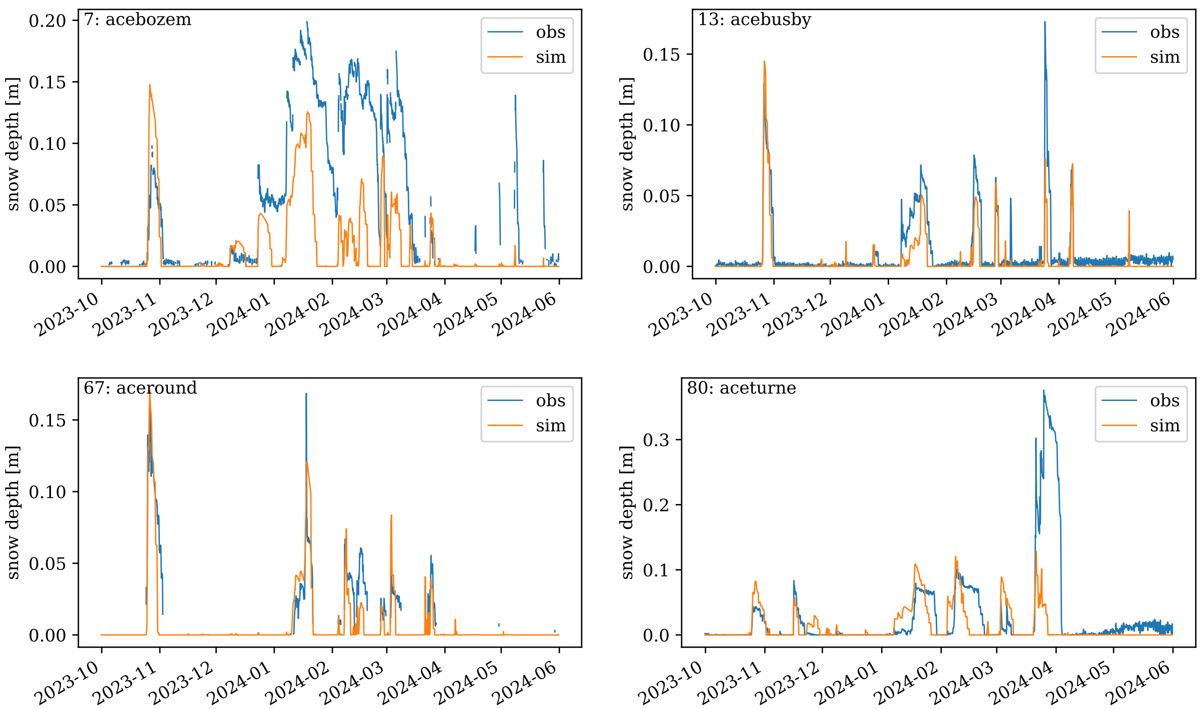


*Figure 5 Examples of simulated snow depth time series at four different stations. The station name is labeled in each subplot.*

*Figure 6 camera view of different stations on Oct 1, 2023. From left to right: Bozeman (acebozem; western foothills), Busby South (acebusby; southern plains), Roundup (aceround; central plains), and Turner North (aceturne; northern plains).*

### *5.4 Event-based analysis: 2023/10/23 to 2023/11/4*

A snow event happened from October 23 to November 4, 2023. This event took place during the early snow season and is expected to be representative of the model's performance under early-season snow conditions. Figure 7 shows the bias ratios over the 48 Montana Mesonet stations for the three experiments during this event.

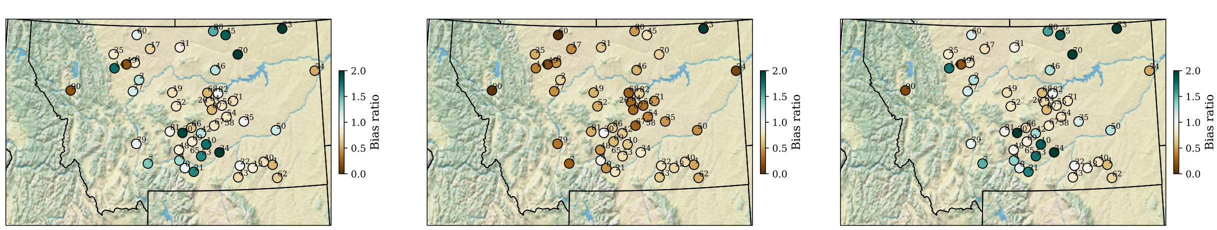


*Figure 7 Bias ratios of snow depth from three different experiments for the snow event from 2023/10/23 to 2023/11/4: control (left), exp2 (middle), and exp3 (right).*

As a reminder, the difference between exp2 and the control experiment is the use of in-situ precipitation, while the difference between exp3 and the control experiment is the use of eight soil layers. Figure 7 shows that snow depth is systematically underestimated at nearly all stations when in-situ precipitation is used (exp2), whereas the bias varies from overestimation to underestimation when ERA5 precipitation is used (control and exp3) although there is generally underestimation at the majority of stations. This result suggests that ERA5 precipitation magnitudes are larger than those from in situ observations at these stations. No major visual differences are observed between the bias ratios of the control experiment and exp3, which means increasing the number of soil layers generally does not affect the overall performance of the snow depth simulations in this event.

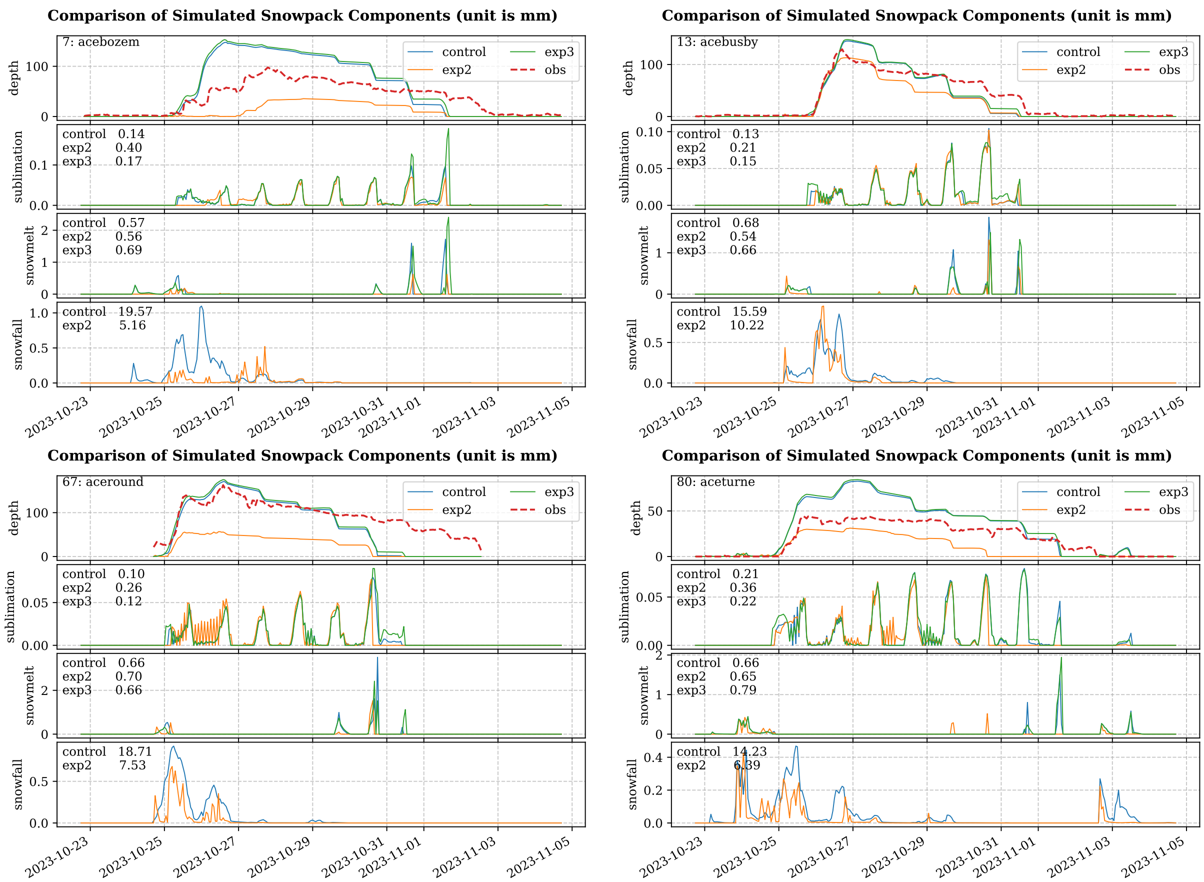


*Figure 8 Simulated snowpack components, namely snow depth (top), snow sublimation (second), snowmelt (third), and snowfall (bottom), during the snow event from October 23 to November 4, 2023, across the four stations. In the snow sublimation and snowmelt subplots, the annotations indicate the ratio of each component to the total snowfall in the corresponding simulation. In the snowfall subplot, the annotations show the total snowfall amount for each simulation.*

Figure 8 compares simulated snow-related components, namely snow depth, snow sublimation, snowmelt, and snowfall, from the three experiments at four representative stations for the snow event from October 23 to November 4, 2023. In the figure, we showed the ratios of snow sublimation and snowmelt to the total snowfall, as well as the total snowfall, for each experiment and station.

The results in Figure 8 show that simulated snow depths using ERA5 precipitation are higher than observations, while those using in situ precipitation are lower than observations at Stations 7 and

80. At Station 13, simulated snow depths agree reasonably well with observations during the accumulation period for both ERA5 and in situ precipitation. At Station 67, simulated snow depths during the accumulation period are reasonably well represented when using ERA5 precipitation but not when using in situ precipitation. Several common features are observed across the four stations: (1) modeled snow melts out earlier than observed; (2) snow sublimation exhibits clear diurnal cycles and accounts for about 15% of total snow loss; (3) snow sublimation shows oscillatory behavior during certain periods; (4) snowmelt accounts for more than 60% of total snow loss, although it generally occurs over a much shorter duration compared to snow sublimation; and (5) the snow mass balance (snowfall = snow sublimation + snowmelt) is not strictly conserved in the simulations.

### *5.5 Event-based analysis: 2024/01/05 to 2024/01/31*

Here, we conducted the same analyses for snow events that occurred from January 5 to January 31, 2024, as for the previous period. The key difference between this period and the previous one is that it occurred during the mid-snow season, meaning that antecedent snowpack could be presented on the ground and could influence the simulations. Figure 9 shows the bias ratios of snow depth from the three experiments during this event period.

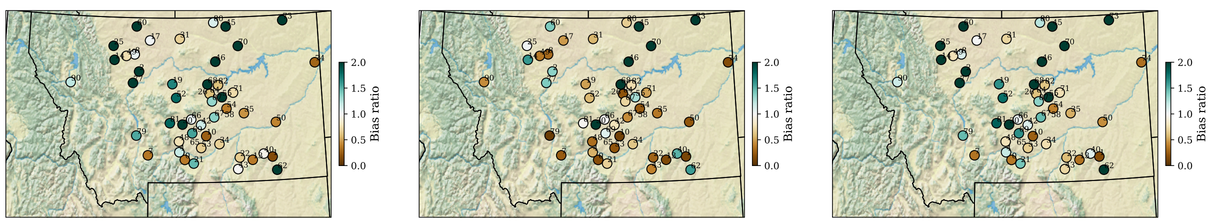


*Figure 9 Bias ratios of snow depth from three different experiments for the snow event from 2024/01/05 to 2024/01/31: control (left), exp2 (middle), and exp3 (right).*

Compared to Figure 7, Figure 9 shows that underestimation and overestimation of snow depth in the control experiment are more evenly distributed across stations. As in the previous event, exp2 tends to underestimate snow depth more than the other experiments, indicating in-situ precipitation was smaller than ERA5 precipitation during this period. The impact of increasing the number of soil layers is minimal.

Similarly, Figure 10 compares simulated snow-related components, namely snow depth, snow sublimation, snowmelt, and snowfall, from the three experiments at four representative stations for the snow event from January 5 to January 31, 2024. In the figure, we show the ratios of snow sublimation and snowmelt to the total snowfall, as well as the total snowfall, for each experiment and station.

Figure 10 shows that simulated snow depths are lower than observations at Stations 7 and 13 for both ERA5 and in-situ precipitation, but for different reasons. At Station 7, antecedent snow prior to this event was not captured by the model, whereas at Station 13, snow melted and did not accumulate. Simulated snow depths at Station 67 agree reasonably well with observations using ERA5 precipitation but are underestimated when using in-situ precipitation. At Station 80, the model overestimates snow depth with ERA5 precipitation and slightly underestimates it with in-situ precipitation. The percentage of snow sublimation relative to total snow loss exceeds 20% for this period, representing a noticeable increase compared to the previous period. Several common features are observed across the four stations: (1) Modeled snow melts out earlier than observations; (2) Snow sublimation exhibits diurnal cycles and accounts for more than 20% of the total snow loss; (3) Snow sublimation shows oscillations during certain periods; (4) Snowmelt accounts for less than 60% of the total snow loss, although it generally occurs over a much shorter duration than snow sublimation; (5) The snow mass balance (snowfall = snow sublimation + snowmelt) is not strictly conserved in the simulations.

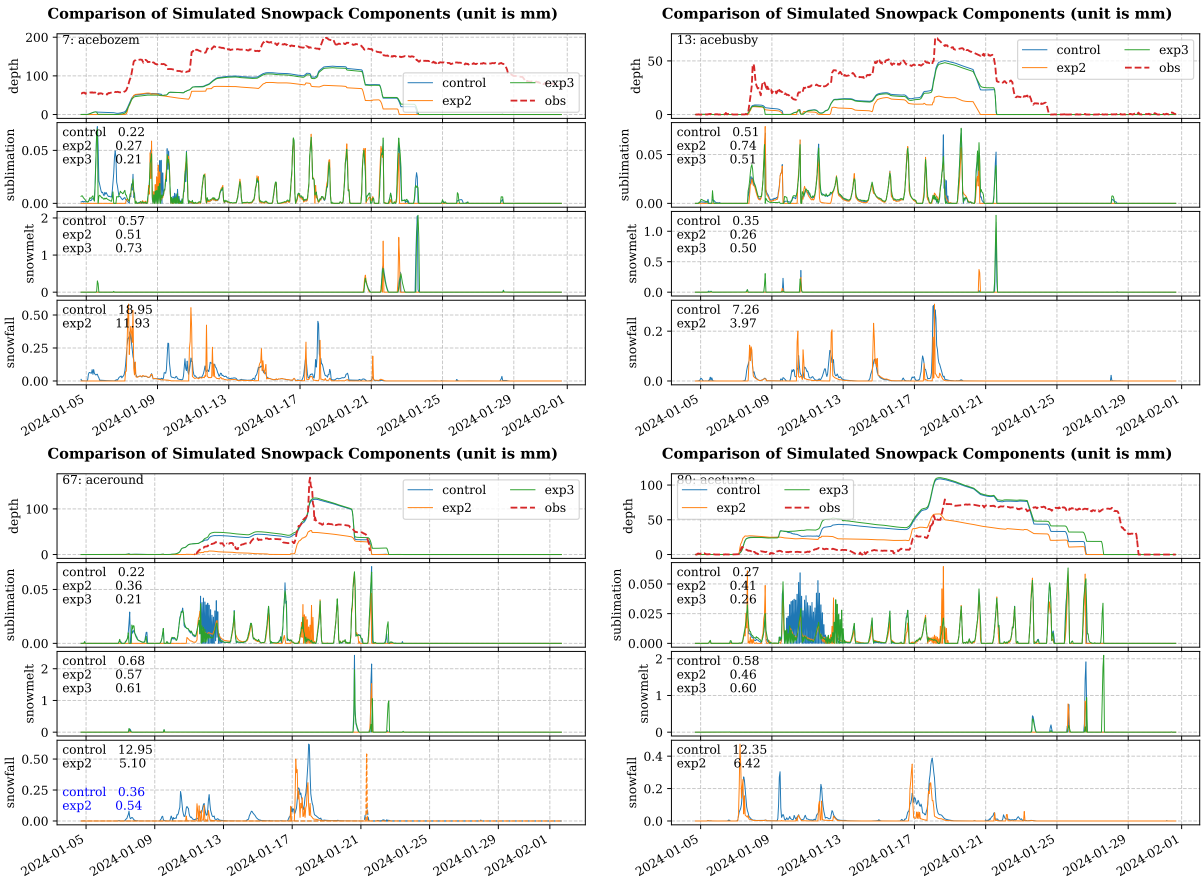


*Figure 10 Same as Figure 8, but for the snow event from January 5 to February 1, 2024.*

### *5.6 Event-based analysis: 2024/04/15 to 2024/04/20*

We also conducted a similar analysis for the snow event that occurred from April 15 to April 20, 2024. The key difference between this event and the previous two is that it occurred near the end of the snow season, when precipitation was often a mix of rainfall and snowfall, which can complicate simulations. Figure 11 shows the bias ratios of snow depth from the three experiments during this event period.

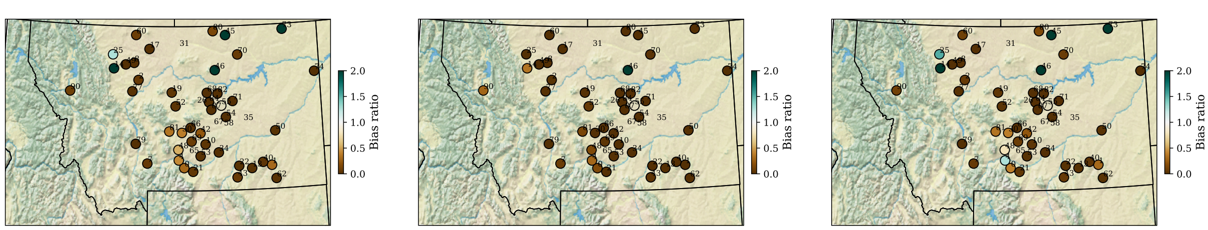


*Figure 11 Bias ratios of snow depth from three different experiments for the snow event from 2024/04/15 to 2024/04/20: control (left), exp2 (middle), and exp3 (right).*

Figure 11 shows that the model underestimates snow depth at nearly all stations. This underestimation may be due to the separation of rainfall and snowfall in the precipitation forcing. Again, the impact of increasing the number of soil layers is minimal, although it appears to reduce underestimation at a few stations.

Similarly, Figure 12 compares simulated snow-related components, namely snow depth, snow sublimation, snowmelt, and snowfall, from the three experiments at four representative stations for the snow event from April 15 to April 20, 2024. In the figure, we showed the ratios of snow sublimation and snowmelt to the total snowfall, as well as the total snowfall, for each experiment and station. In addition, the time series of rainfall and its total amount are also shown in the figure.

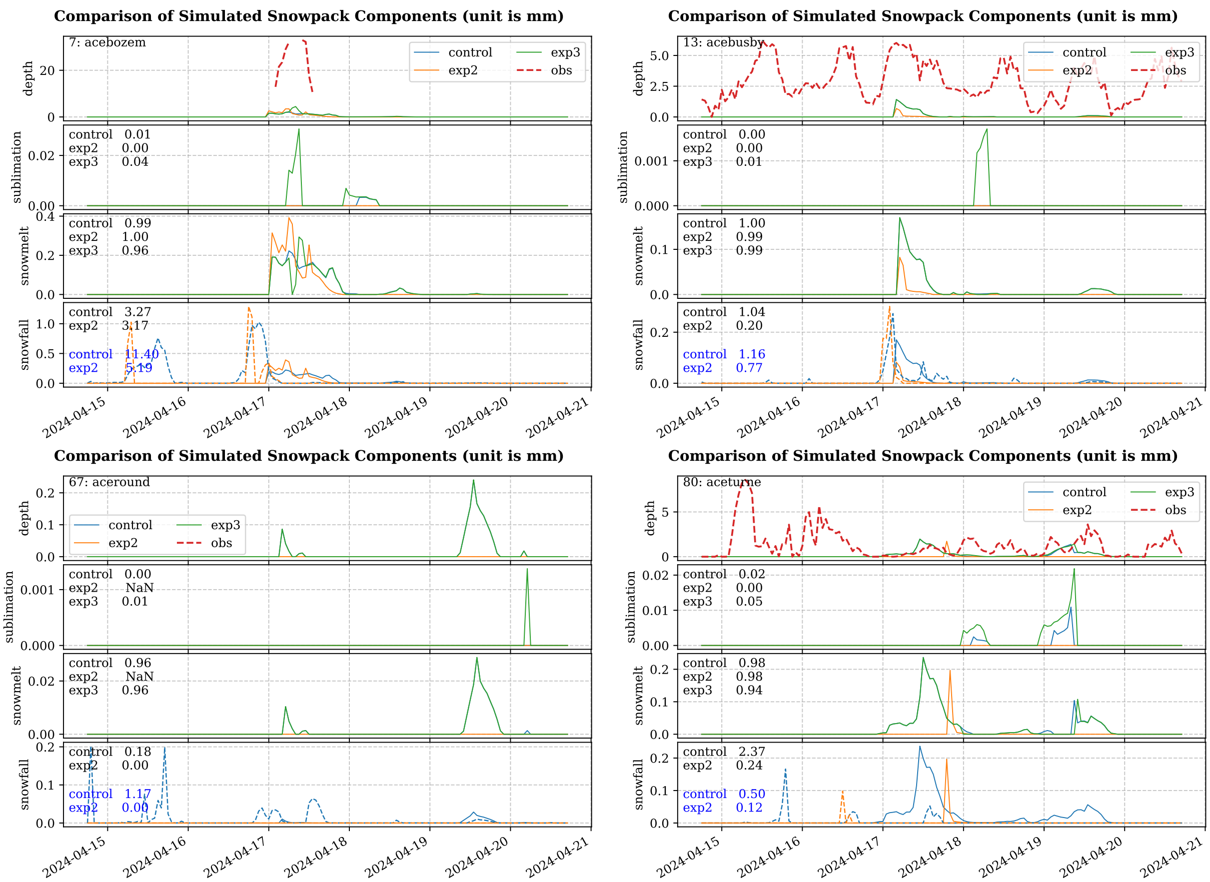


*Figure 12 Same as Figure 8, but for the snow event from April 15 to 20, 2024. The only difference is that the time series of rainfall (dash lines) and its total amount (blue color) are also shown in the figure.*

Figure 12 shows results from the snow event occurring from April 15 to 20, 2024, a period when precipitation is typically a mixture of rainfall and snowfall. Snow depth observations from the different stations were visually validated using camera views. Observations from Station 7 (Bozeman) are considered reliable, whereas observations from the other stations have uncertainties. For example, only a shallow layer of snow was present on the ground in the April 16 afternoon photo of Station 80 (Turner North), while Stations 13 (Busby) and 67 (Roundup) were snow-free. During this period, snow depth tends to be shallow and remain on the ground for only a short time (usually less than one day). In addition, the determination of whether precipitation falls as snow or rain is also complicated when air temperature is near 0 °C. Consequently, the model underestimates snow depth at Station 17 and overestimates it at the other stations. An additional factor contributing to these discrepancies is the separation of precipitation into rainfall and snowfall, as illustrated in the bottom subplot of Figure 12 for each station.

### *5.7 Impact of soil layers*

For exp3 of this project, the number of soil layers was increased from 4 to 8 while keeping the total soil depth the same. This section presents the impact of this change.

Figure 13 shows a comparison of soil moisture from the 5-cm in situ observations, the first layer of the control experiment, and the first and second layers of exp3 from April 1 to September 30, 2024 over four Montana Mesonet stations.

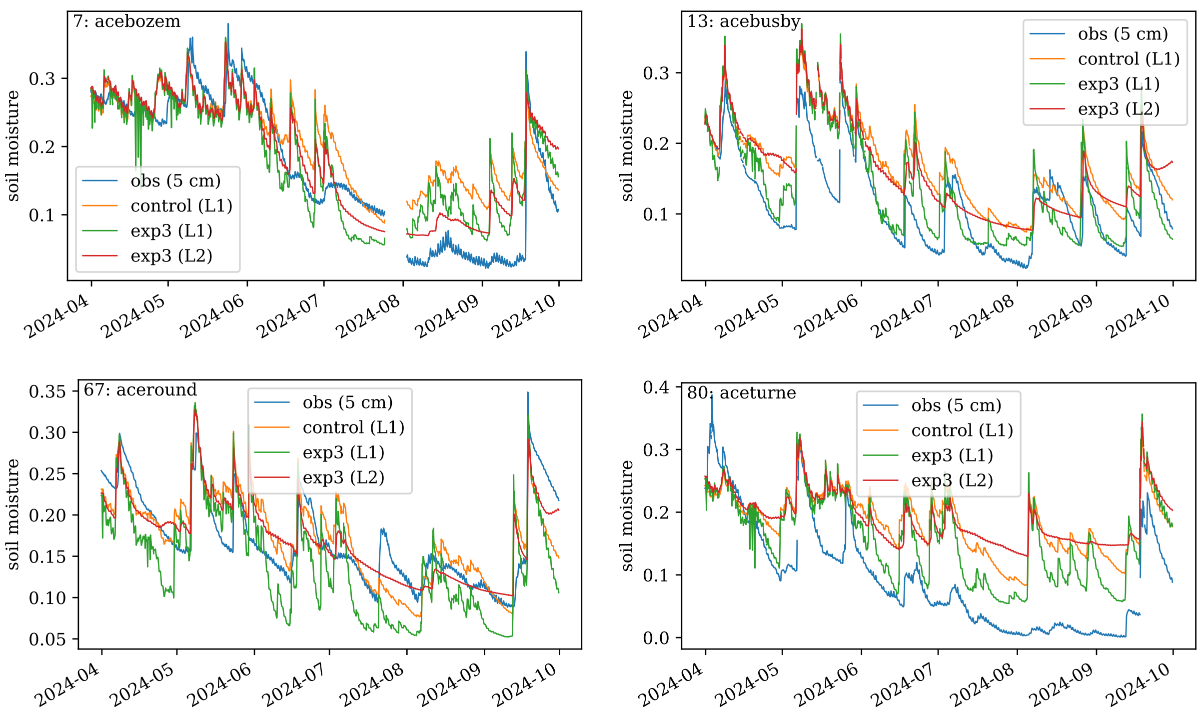


*Figure 13 Comparison of soil moisture from the 5-cm in-situ observations, the first layer of the control experiment, and the first and second layers of exp3 from April 1 to September 30, 2024 over four Montana Mesonet stations.*

Figure 13 shows that the first layer (5-cm thick) of exp3 is more dynamic than the first layer (10-cm thick) of control simulations while the second layer (5-cm thick) of exp3 is less dynamic than the first layer of control simulations. By splitting the original 10-cm layer into two 5-cm layers, the modeled soil moisture exhibits greater variability than in the original configuration. Figure 13 also shows that observed soil moisture is generally drier from July to September. Since the observations are from a 5-cm depth, direct comparison with the model results is not straightforward. At Station 7, the second layer of exp3 is closer to the observations than the other layers, whereas at Station 13, the first layer of exp3 is closer. For this reason, model performance is evaluated primarily through its correlation with observations in the following discussion.

Figure 14 shows the correlation coefficient of the first-layer soil moisture from the control simulations with observations for each station, as well as the correlation coefficient differences between the first layer and second layer of exp3 and the first layer of the control. The reason for presenting the data this way is to highlight the changes in correlation coefficients from the first layer of the control experiment, since these values are similar.

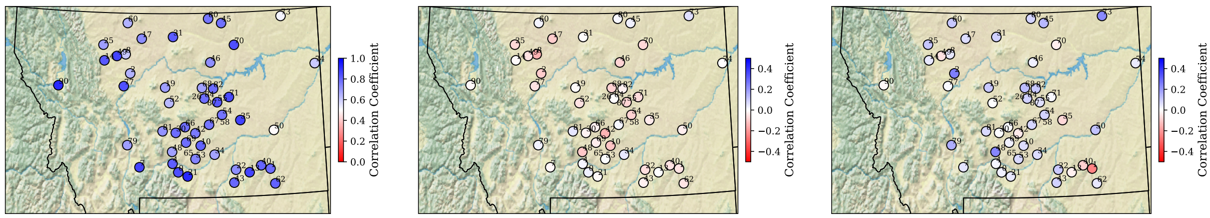


*Figure 14 Correlation coefficient of the first-layer soil moisture from the control simulations with observations for each station (left), as well as the correlation coefficient differences between the first layer (middle) and second layer (right) of exp3 and the first layer of the control from April 1 to September 30, 2024.*

The correlation coefficients of soil moisture between the first layer of the control and in-situ observations in Figure 14 show that the model reasonably captures the trend of soil moisture during the period. By increasing the number of soil layers from 4 to 8, the correlation of the first layer of the new simulations (i.e., exp3) is lower than that of the original first layer, whereas the correlation of the second layer of the new simulation is higher than that of the original first layer.

We also investigated the impact of increasing the number of soil layers on surface heat fluxes, namely, sensible heat flux, latent heat flux, and ground heat flux, in Noah-MP simulations. The results are given in Figure 15.

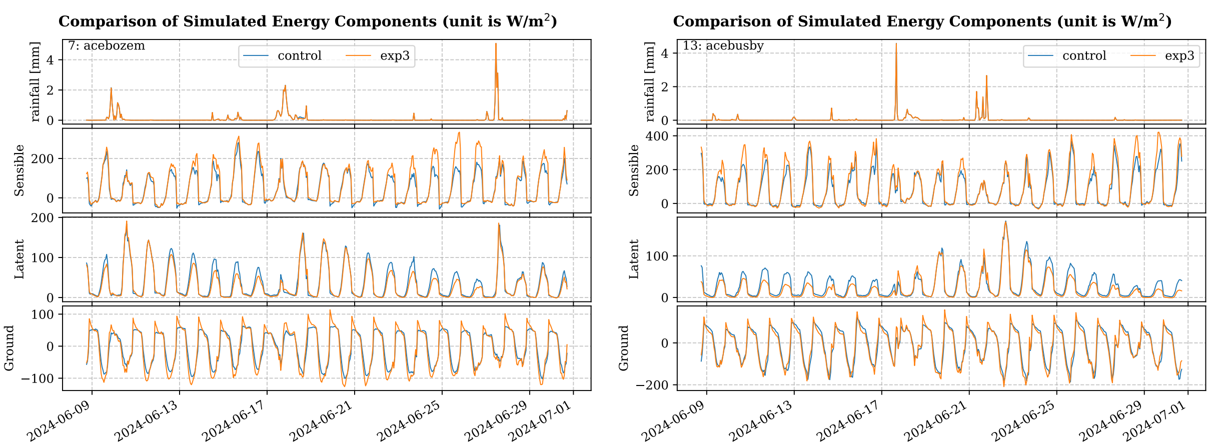


*Figure 15 Comparison of the simulated sensible heat flux, latent heat flux, and ground heat flux from control and exp3 simulations.*

From Figure 15, it can be seen that increasing the number of soil layers from 4 to 8 has a noticeable impact on the simulation of surface heat fluxes. Rainfall events can be identified using the rainfall data shown in the figure. The simulated sensible and latent heat fluxes are similar between the control and exp3 simulations for a short period after rainfall. However, when the soil is dry, the simulated latent heat flux tends to be higher in exp3 than in the control, while the simulated sensible heat flux tends to be lower in exp3 than in the control. The simulated ground heat flux exhibits larger diurnal variations in exp3 compared to the control.

**6 Acknowledgments**

This project would not have been possible without the support and contributions of many individuals. I would like to sincerely thank Dr. Man Zhang and Kathryn Newman for hosting me and providing the support needed to move this project forward. I also thank Drs. Michael Barlage, Cenlin He, Anders Jensen, Joseph Olson, and Ligia Bernardet for their valuable discussions, technical guidance, and insightful feedback on the project. I am also grateful to Drs. Jeff Beck, Louisa Nance, Weiwei Li, and Samuel Trahan for their assistance in helping facilitate collaboration across teams and data resources.

I also appreciate the support from NSF NCAR's Derecho supercomputing system, which provided the computational resources and environment necessary to carry out this work.

Any remaining errors or interpretations are my own.