10 March 2021

Response to reviewer comment on gmd-2020-371 (*Wayne Angevine, 12 Feb 2021*)

by Román Cascón and co-authors

Original comment (Wayne Angevine, reviewer):

1. *I have concerns about the use of different land surface models with the same input data. Each LSM has its own "climate", and the input data (assuming it all comes from FNL, as implied by table 2) comes from a different model with a different climate. This is strongly demonstrated by Angevine et al. (2014) in ACP for the same project. Have the model outputs been checked to be sure that there are not strong spinup effects in the soil? I particularly suspect that this is responsible from some of the behavior of the RUC LSM, which is known to have a different soil moisture baseline from Noah.*

Response from authors:

Dear reviewer,

Thank you very much for all your comments. We will take all of them into account in the next version of the manuscript, which will be provided at the end of the discussion stage. However, we would like to include here a more detailed response to your comment 1, related to the soil moisture (SM) spin up for the different land surface models (LSMs).

Indeed, we do agree with your comment and we have performed different simulations including longer spin-up times to check the differences obtained (shown later).

However, we have decided to avoid performing individual spin-up simulations for all the experiments of the paper using the different LSMs due to two main reasons:

1. The first one is related to the SM spatial heterogeneity that is obtained after spinning up the model, as commented in lines 165-169 of the current manuscript:

"*Regarding the effect of possible SM differences within the area, two aspects should be noted. On the one hand, the potential SM horizontal variability due to possible inhomogeneous precedent rainfall over the area was not taken into account. However, the SM input in the model is provided with a coarse resolution of 1° and does not show any small-scale details. This is a well-known limitation of mesoscale models which is sometimes addressed through the assimilation of SM data from satellites or with previous long simulations that serve to spin-up the surface in order to obtain the appropriate SM initial values (De Rosnay et al., 2013; Angevine et al., 2014; Santanello et al., 2016). In our case, this limitation of the mesoscale modelling is an advantage because it allows to perform a fairer model-observation comparison since this limitation also exist in the AAF.*"

That is, if we decide to spin-up the model (which is more realistic), we will lose the SM spatial homogeneity that is in the SM used to initialise the model from the NCEP-FNL data (1°). As mentioned, this homogeneity is an advantage since we evaluated the model with area-averaged fluxes (AAFs) that omit the SM spatial variability, which allowed us to perform a fairer model-observations comparison. This is well observed in Fig. 1, where we compare the SM from NCEP-FNL at 5 cm used in the simulations included in the original manuscript (note how the SM is the same for Noah (a) and RUC (b)). As shown in the figure, we have obtained a more probabilistic soil moisture field with a spin-up of 1 month (panels c and d). However, we have introduced an undesirable effect in our numerical design: soil moisture fields are now heterogeneous and therefore more difficult to be compared with AAF.

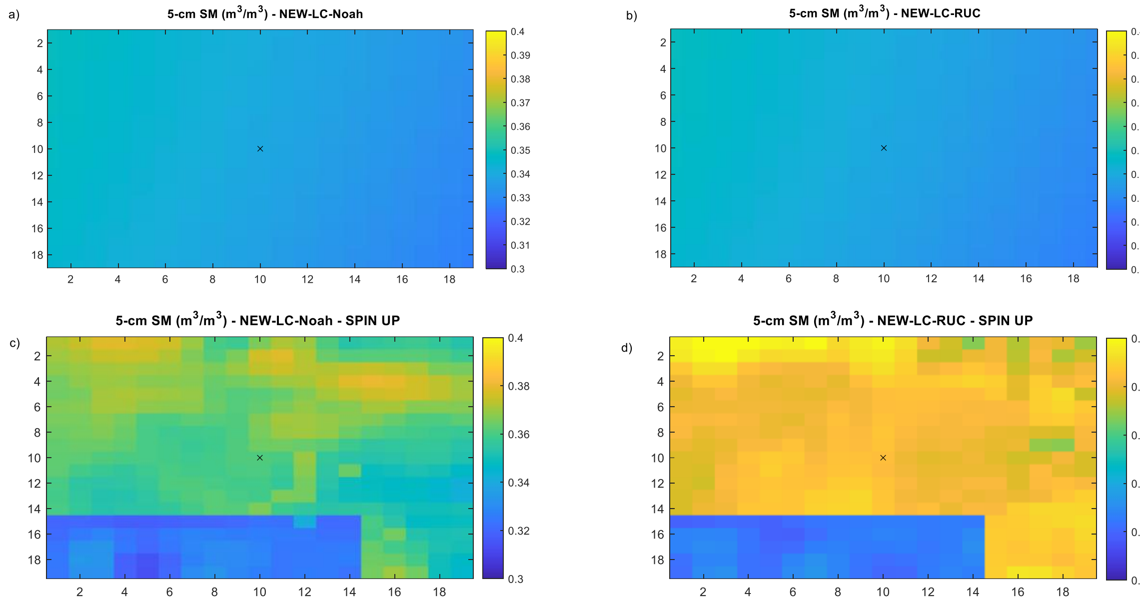

Figure 1. Soil moisture at 5-cm depth (first soil layer for Noah and the second one for RUC) valid at 12:00 UTC of 18 June 2011 in the area of interest (19x19 pixels of 1-km resolution). The central site of the BLLAST campaign is indicated with a black x. The SM is the same for all the simulations included in the original manuscript (panels a and b). Panels c and d show the SM obtained after a spin-up time of 1 month using Noah (c) and RUC (d). Note the spatial heterogeneity obtained, related to the heterogeneous previous rainfall, to the land use (see for example the urban pixels in panel c comparing to Figure 2d of the original manuscript) and, especially, to the soil type (blue box with less SM in the SW of the region (loam), in comparison with the rest of the area (clay loam, which retains more SM)). We would like to remember that, in our study, we removed this area of loam in the SW to avoid this strong effect of the soil type.

2. The second reason to avoid spinning up the model is an indirect effect. Performing long spin up simulations with different LSMs leads to SM differences caused by the differences in the WRF rainfall that are simply due to the use of different LSMs. This is well observed in Fig. 2, where we show the cumulative rainfall in each pixel after 1 month of spin-up for Noah and RUC.

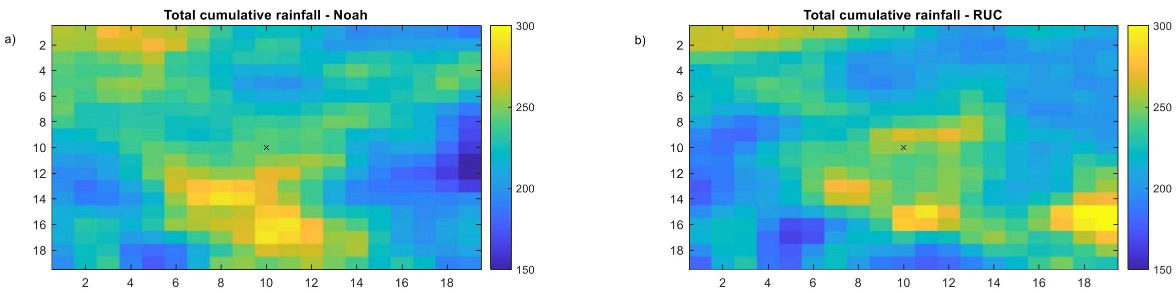

Figure 2. Total cumulative rainfall (in mm) produced by Noah (a) and RUC (b) after one month of spin-up time (from 17 May to 18 June 2011). Note the spatial differences in the rainfall patches, probably caused by differences in the position and strength of simulated convective systems (typical in the region on these dates).

Therefore, we think that for the objectives of our study it is better to avoid a period with spin up since the data used for the comparison (AAF) assumes SM spatial homogeneity (point 1), as we have in the original NCEP-FNL data. Besides, the work compares different LSMs and the application of individual spin-up periods can cause differences in rainfall among the LSMs that will affect the SM and, therefore, the fluxes (point 2).

However, we do agree that performing long spin up simulations is probably the best practice to include appropriate initial SM values in the model, which allows including a more realistic SM heterogeneity and values that agree with the dynamics (climatology) of each LSM.

Concerning the reviewer suspects about the different dynamics of RUC, we checked the differences in SM dynamics between the 4 LSMs used. Indeed, as commented by the reviewer, it seems that RUC has a remarkable different SM dynamic, which is shown in Fig. 3 for the period of the simulations analysed in the paper. For this reason, we decided to investigate these differences by performing two simulations with a spin-up time of approximately 1 month using Noah and RUC (Fig. 4).

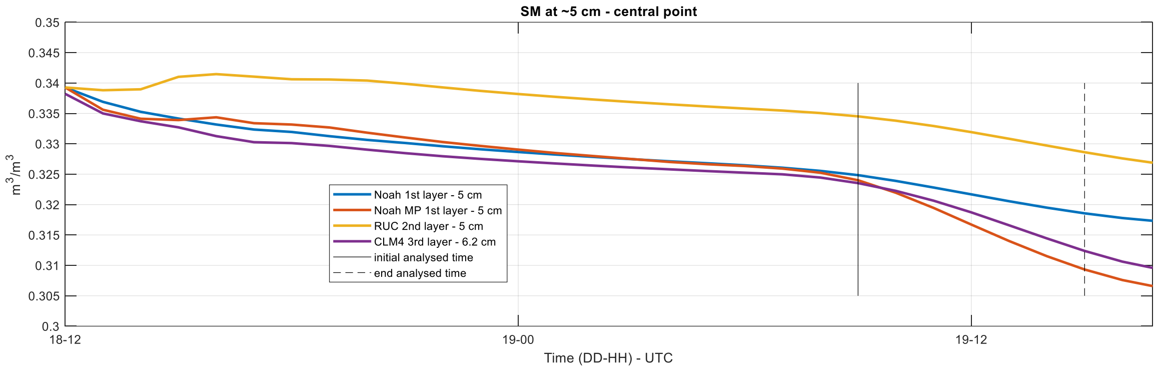

Figure 3. SM at approximately 5 cm depth simulated by the different LSMs in the original simulations of the manuscript (without spin-up). Note the different dynamics of RUC.

.

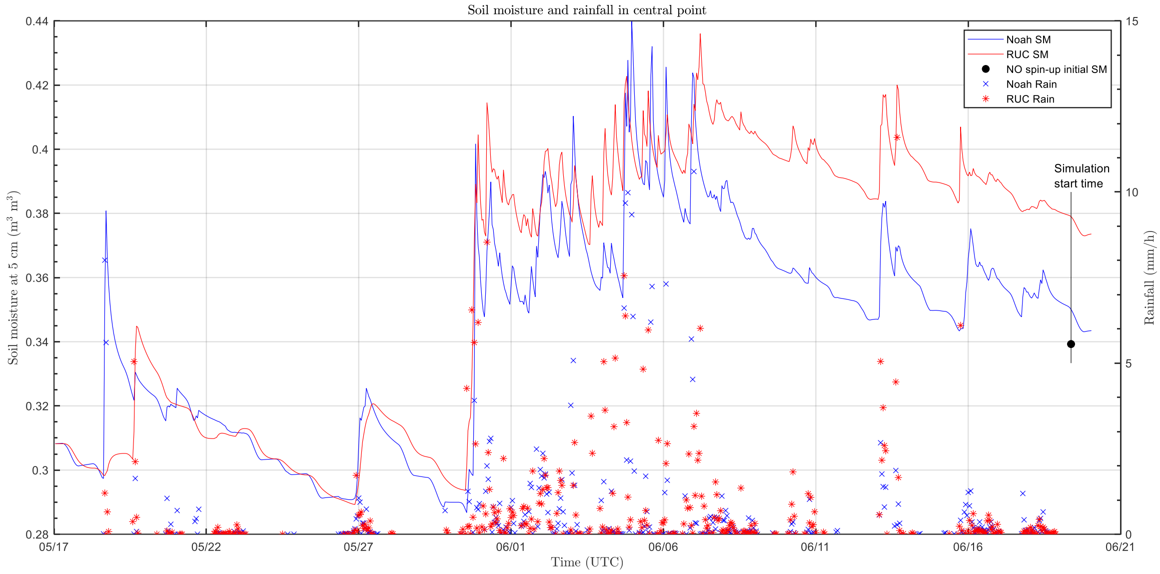

Figure 4. SM at 5 cm (solid lines) simulated by Noah (blue) and RUC (red) in the central pixel. The SM included in the original simulations (without spin up), is indicated with a black point (~0.34 m^3/m^3), to be compared to the value obtained after the spin-up with Noah (~0.35 m^3/m^3) and with RUC (~0.38 m^3/m^3). Rainfall (mm/h) is included with symbols (total of 242.5 mm with Noah and 252.2 mm with RUC).

Fig. 4 demonstrates the suspects of the reviewer for the different behaviour of RUC in comparison with the other LSMs. Indeed, the SM dynamics in this LSM is different in comparison with the others and, therefore, the initialisation with the default SM initial data from NCEP-FNL is not appropriate. For the case of Noah, the value originally used, obtained from the NCEP-FNL data (~0.34 m^3/m^3) is quite similar to the value obtained after the spin-up period (~0.35 m^3/m^3). However, the differences are larger for RUC (~0.38 m^3/m^3).

We then consider that the initial SM at 5 cm used in RUC should be higher than the value used in Noah (according to Figures 3 and 4), being a value that must be approximately 0.03-0.04 m^3/m^3 higher. However, this is not the case at deeper levels, since each LSM has a different physics within the soil. Indeed, Fig. 5 shows the SM at 150 cm (for Noah, 4th level) and 160 cm (for RUC, 5th level). In this case, the SM is higher in Noah than in RUC.

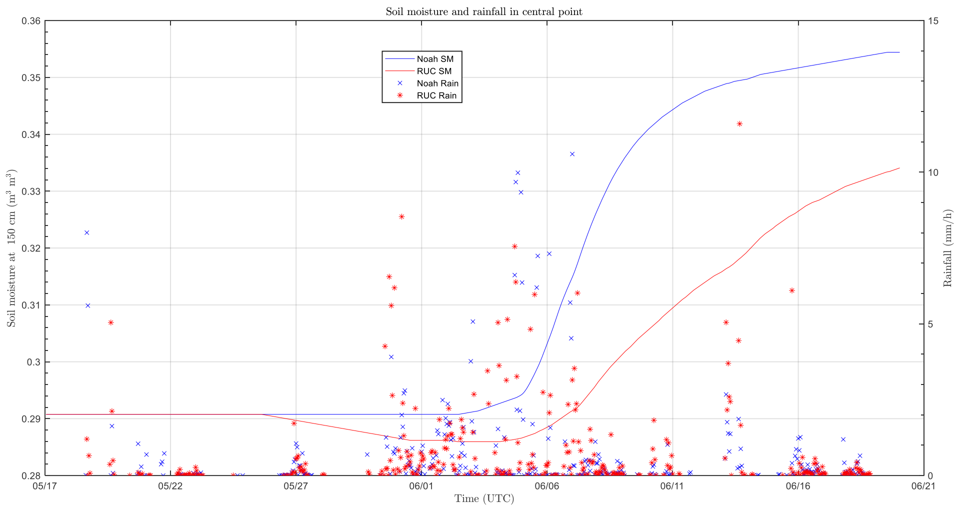

Figure 5. As in Fig. 4 but for a deeper soil level: 150 cm for Noah (4th level), 160 cm for RUC (5th level).

Therefore, the SM differences between LSMs also depends on the soil levels, which makes it more difficult to apply a SM correction for RUC to initialise the model (e.g., it is not correct to apply +0.3 or +0.4 m^3/m^3 in all the levels to the NCEP-FNL initial SM values used in our case). Hence, we cannot easily correct the SM initial value for RUC if we want to maintain the SM horizontal homogeneity we wanted to compare with the AAF.

In any case, it is demonstrated that the superficial SM should be higher for RUC, which will act correcting the systematic underestimation of latent heat flux (Le) and overestimation of sensible heat flux (SH) observed in our study for RUC in all the land cover categories (see panels b and e of Figure 6 of the original manuscript).

We have computed the scores that would be obtained using the 1-month spin up with RUC, obtaining improvements in both surface fluxes (Table I).

Table I. Comparison between scores for NEW-LC-RUC with and without spin up.

	NEW-LC-RUC (NO spin-up)	**NEW-LC-RUC (spin-up)**
Total bias SH	77 W/m^2	47 W/m^2
Total bias Le	-70 W/m^2	-46 W/m^2
Total RMSE SH	90 W/m^2	68 W/m^2
Total RMSE Le	85 W/m^2	68 W/m^2

Nonetheless, we do agree that the impact of the SM initialization should be, somehow, better addressed or commented in the paper. We then suggest to include in the paper an additional experiment named “SPIN-UP”, commenting on this point, adding a link to this online discussion and including the scores obtained after spinning up the LSMs: Noah and RUC results have been shown here; the results for Noah-MP does not show significant differences with or without spin up; for CLM4, we have found some technical issues with the spin-up simulation, leading to unrealistic radiation values in the morning transition that affect the fluxes*.

We think that adding the information obtained from these experiments completely fits with the main objective of the paper, which is to investigate the surface representation impacts on the fluxes (including now the soil moisture initialisation).

* We continue investigating this issue with CLM4.