Answers to Review by Anonymous (RC3):

General Comments

This paper describes the impact of a forest parameterization on coupled WRF-CFD simulations over complex terrain for a cold front case study. Simulation results are compared to met-mast observations and UAV measurements. The test of a forest parameterization in WRF is interesting and relevant for the model community. The results are, however, described very qualitatively and the main findings should be communicated more exactly. Too often flow situations are described, which are not relevant for the main results of the paper and make the paper difficult to read. It should be focused on important results and it should be quantified what the benefit of coupling WRF with the CFD model is. Is it really necessary to run a CFD with 5 m horizontal resolution and how much better is it compared to WRF? WRF results with a spatial resolution of 90 m are mentioned, but not shown. They should be included in the manuscript and compared to observations and CFD simulations. Further, I think that the UAV observations should be included/used in a better way to quantify the model errors. In the manuscript they are only used to describe the situation qualitatively. I suggest major revision for the submitted manuscript.

We would like the reviewer for his/her very details comments and in-depth review. These very confident that these comments have helped to improve the manuscript significantly. We have improved the description of the microscale simulations, added more statistical analysis and removed some plots showing only the general flow. Please refer to the answers to each individual comment for more details.

Major Comments

1. *You are running WRF with a horizontal resolution of 12.120 km in the outer domain, which is coarser than the ECMWF data (mesh size of 9km) that is used as initial and boundary conditions. Why did you use this coarse resolution in the outer domain? This means that the flow is upscaled when interpolating from the 9 km ECMWF grid to the coarser WRF grid? Meteorological fields and synoptic events like cold fronts are strongly smoothed, which also influences the results of the inner WRF domains. Typically, mesoscale simulations are started with the same or higher grid resolutions than the driving model. I suggest that you rerun the WRF simulation by starting with domain 2 as outer domain. This should improve your results. Can you also add more information to the model setup section 2.1 about the date and time period, which is simulated: how many days were simulated, what was the time step and output interval, why was this event chosen? Why is the passage of the cold front important? What tree height did you use in the forest parameterization and where did you get the tree height from? What was the real tree height at the test site?*

We admit that this is an aspect we have not taken into account when designing the setup and a negative impact has not been observed. In order to test the impact of the outermost domain, we have removed domain 1 (dx=12.120 km) and left everything else unchanged, except that we also increased the spin-up time to 12 hours. This has interestingly very little impact on the wind speed, but a very negative impact on the development of potential temperature at the test site. In the original simulation, the potential temperature dropped by approx. 7 K in less than 30 minutes, just as observed. The modified setup lead however to a development where the potential temperature dropped by the same 7 K over a period of 3 hours! We hypothesized that this is the result of a smaller outermost domain. Thus, we increased the size of the outermost domain to the same size of the former 12 km domain an re-run the simulations. This improved the simulations with respect to potentially temperature, but the drop during the passage of the front was still smoothed out compared to the original simulation. This interesting behaviour might be due to the fact that ecmwf boundaries are updated only every 3 hours. However, given that the wind speed and most importantly the impact of the forest parametrization was very similar in all three setups. Thus, we decided to use the original setup for this publication since the sudden drop of temperature associated with the cold front is simulated best in this setup.

We have added also more information about the setup an forest to section 2.1.

2. The comparison of model results with UAV measurements has to be improved. You only use data from the met-mast to quantify the model error, whereas data from UAV flights are just used for qualitative comparisons. Model data should be interpolated in space and time to the measurement points of the UAV flights and should then be compared directly to observations. I suggest to plot correlation scatter plots (observation versus simulation) to get an impression if wind speeds are over- or underestimated. Biases, mean errors and correlation coefficients should be computed for the met-mast (already done for WRF and WRF-F) and UAV observations for all simulations: WRF, WRF-F, OF(WRF), OF(WRF-F), OF-F(WRF), OF-F(WRF-F). The description of the results is generally were qualitatively done and the effect of the forest and the coupling of WRF with CFD has to be quantified. Is it necessary to run a CFD with 5 m horizontal resolution? The UAV also measured TKE: please compare it to simulated TKE of all simulations.

In order to improve the comparison of UAS and WRF model, we implemented a routine into WRF to follow the trajectory of the UAS and output u,v,w and tke at each time step at the grid box closest to the current position of the UAS. Given that the length of the UAS track is only 1.5 km long, an interpolation of WRF-data at dx=150 m to the UAS track won't lead to a fair comparison of UAS and WRF Model. Thus, we average data from the UAS leg-by-leg for all datapoints which location resides within a corresponding WRF-datapoint. WRF-data, being available at each time step is then averaged over the period of time the UAS needs to cross each respective grid cell. This way, we can directly compare UAS and WRF and create correlation scatter plots. Such plots are also generated for the tower (100 m and 45 m) for all simulations. Error statistics are calculated and presented in tables for each simulation and all tower levels as well as UAS measurements. We decided to leave out comparisons of TKE for this work since we wish to focus on the mean flow and the impact of the forest. The turbulence structure will be studied in the future.

3. You mention the WRF run with 90 m horizontal resolution, but don't show the results. These simulations should be included in the paper and compared to OF simulations.

All simulations have to be compared quantitatively to both met-mast and UAV observations.

This short remark has been added to the manuscript to demonstrate that we explored alternative setups as well. However, we decided not to include data from the dx=90 m simulation to this manuscript. The reason is, that it is based on a rather unusual setup, following Muñoz-Esparza et al. (2017), who used five grids with mesh sizes 8910/2970/990/90/30 m. Mark the interesting jump by a factor 11 between the third and fourth domain. Muñoz-Esparza et al. (2017) suggested that some random perturbations must be added to the boundary data between the two domains 3 and 4 to help the spinup of turbulence and wrote a code to add this feature to WRF. However, this feature has not yet been included to the official version of WRF, so we have no access to this variant. We wanted to test this approach nevertheless and ran a simular setup, but removed the 30 m domain. To provide the flow enough space to spin-up turbulence before it reaches the test-site, we had to use a very large 90 m domain. This lead to very expensive simulation that were inadequate for testing sensitivities as required in our research. Furthermore, including this simulation in the manuscript would require a substantial increase of the methods section 2.1 since the setup is quite different. Also, the results were comparable to the ones of the coarser WRF simulations except that smaller scale feature and turbulence were somewhat better resolved. For this reason, we decided not to include results from this simulation in this manuscript.

4. Some figures should be left out: Figs. 7 and 11 don't add additional value to the manuscript and corresponding passages in the text are difficult to read and understand, as they only describe meteorological situations in a qualitative way and try to explain how the flow situation was in some valleys (e.g. Simonsbach valley), whose location is not clear/described. As observations were only available at the WINSENT test site the description of the flow should focus on that location. It's also not necessary to show 3 hours of streamline plots in Fig. 4, as streamlines in two different levels are confusing. Please only show streamlines on the lower level for one hour (e.g. 14 UTC). The streamline plots in Fig. 12 can be left out as it is nearly impossible to detect differences between OF and OF-F. Correlation plots should be included for all model runs instead, which make clear how much wind speeds are over- or underestimated. These numbers have to be summarized in a respective table.

The Simonsbach valley is now marked in Figure 1 and we kept the part of the description of the model simulations inside this valley since it is a rather interesting side-effect of the forest. Given that the main wind directions at this site are east and north-west, one has to expect that similar flow features can be found in observations. It is also important for the inflow of smaller scale simulations. We have nevertheless focused on the test-site itself with regards to statistical evaluations and have replaced Figs 7 and 11 with correlation scatter plots and tables summarizing the findings.

Minor Comments

1. P2L41: Please add: ... and the CFD model in the order of tens of meters...

done

2. P2L47: Can you add an overview of the paper like: "The paper is organized as follows: section 2 describes the used methods..."

done

3. P3L55: Can you please add the mesh sizes of all model domains. D1 has 12150m, D4 450m and D5 150m resolution. What about D2 and D3? Probably 4050m and 1350m as you use a factor of 3? Please add this to the text.
done

4. P3L55: Why do you use such a coarse horizontal resolution of 12.125 km for domain D1? You are initializing WRF with ECMWF, which has a horizontal resolution of 9 km. This means that you strongly smooth ECMWF data in space before they are used in WRF.

We changed the setup (also in accordance to your major comment 1) and run now a setup with 4 domains, leaving out the coarsest domain. For the results of this test, please refer to our response to Major Comment 1.

5. P3L59: What is the vertical level distance of the coarser domains? I guess dz=15m at 10m above ground level is valid for the LES domains?

Yes, dz=15 m is valid for the innermost domain. For the mesoscale domains, dz=50 m and for domain 4 it is 25 m. This information has been added to the manuscript.

6. P3L63: I think this is wrong: "subgrid-scale turbulence is parameterized using the revised MM5 surface layer scheme."
Subgrid-scale turbulence is parameterized by the Deardorff TKE scheme. The MM5 surface scheme parameterizes the exchange processes at the surface. Please correct this in the text.

Line corrected accordingly.

7. P3L68: What time step do you use? Do you use adaptive time stepping? In Fig. 4 and 5 you say 10 minute averages are shown. Are these really averages or snapshots? What's the time interval used for averaging over 10 minutes? Every time step?

The time step is 60/20/5/1/0.25 in domain 1/2/3/4/5. The 10 minute averages of the time series are based on every time step. For the maps, we averaged data online every 30 s. The reason for this choice is a practical one. We wanted to calculate standard deviation and turbulence intensity for the whole domain as well. This required saving the relevant data over a period of 10 minutes. Without increasing the sampling interval, this process required too much memory.

8. P3L70: Please add an article: ...“the ASTER topography data set...“

done

9. P4L74: Please change to: "... were initialized at ...“

done

10. P4L74: Please change the time format: "...21 September 2018, 00:00 UTC“

done

11. P4L74: Please change to: "... considered as model spin-up...“

done

12. P4L81: I think the formulas for LAD are taken from Lalic and Mihailovic (2004). Please cite this paper, when you describe the formulas.

We have added the citation.

13. P4L82: Where do you have the formula for Lm=LAI(0587h-0.124) and hm=0.6m from? Please cite the corresponding paper. I guess it should be 0.587h in the brackets?

The Formula for L_m is a linear fit for a range of vales for h of the integral when solving the Equation for L_m for a given LAI. Since we did this ourself, there is no paper to cite. However, we recently found that Mohr et al. (2014) used $L_m = LAI/h1.69$, which leads to virtually identical results. However, since we have implemented the version above, we shall keep the formula as is, together with a brief explanation. The decimal point was missing indeed and has been added.

14. L4P84: I don't understand the sentence: "... are classified as forest and lie below the maximum tree height“. What does this mean? Which tree heights do you use? What is the real tree height in your modelling domain?

We have improved the syntax. We tried to state that the forest parametrization is activated over all tiles classified as any kind of forest. The tree height at this site can be estimated my subtracting the data of a digital elevation model from a digital object model (that contains trees and buildings). This yields a broad spectrum of tree heights anywhere from 10 to 35 m. The value 30 ± 5 m is larger than the average thee height at this site, but we wanted to make sure that at least 2 data points are affected in WRF.

15. L4P85: Change "lowest 2-3 data points“ to "lowest 2-3 model levels“

done

16. P5L88: Add a reference to the dashed white box in Fig. 1b): "... along the borders of a 10x10x2.5km large box (see dashed white box in Fig. 1b).“

done

17. P5L89: Can you add the dates and time intervals that will be simulated with WRF and the CFD model? The simulation started at 21 September 2018 at 00:00 UTC. When did it finish? What was the output interval? When was the CFD model started/finished?

The WRF model ran from 21.September 2018 00 UTC to 23 September 00 UTC. The CFD model ran from 21.September 2018 09 to 18 UTC. The WRF model ran for a second day to test the performance of the forest parametrization over a longer period (especially during the night). We decided to focus on the frontal passage with the CDF as these simulations are much more expensive. Times are added to the manuscript.

18. P5L93: Please add the acronym URANS: "An unsteady Reynolds Averaged Navier-Stokes (URANS) approach...“

done

19. P5L103: I think it should be: "Vegetation is discretized into finite...“

done

20. P5L108: Can you explain what a PIMPLE algorithm is?

PIMPLE is a large time-step transient solver which combines the PISO (Pressure Implicit with Splitting of Operators) and SIMPLE (Semi-Implicit Method for Pressure Linked Equations) algorithms. PIMPLE enables to loop over the PISO algorithm within one time step.

21. P5L112: In line 88 on page 5 you say that the domain size of the CFD model is 10x10x2.5km? Which altitude is correct (here you say it's 2km)?

There is a missunderstanding. The CDF model has its ceiling at an altitude of 2.5 km. With the terrain at its lowest point at 500 m msl in the CDF model, this leads to a 2 km high domain. We ensured that this is formulated more clearly in the manuscript.

22. P5L114: What does "finer near the ground“ mean? Can you add the vertical mesh size near the ground?

The mesh size near the ground is 1.6 m vertically.

23. P5L115: Does "conforming to the site orography“ mean a terrain following vertical coordinate?

Yes.

24. P6L118: Which roughness length did you use for forested areas when no forest parameterization was used?

A roughness length of 0.5 m was assigned in this case. The values is stated in the manuscript.

25. P6L131: Typo: "an inertial navigation...“

corrected

26. P6L130: When did the UAS measurements take place (at which time)?

The UAS measurements took place at 21. September 2019, from 10:55 to 12:29 UTC.

27. P6L140: Please change the date format.

done

28. P7L47: Typo: "is characterized“

corrected

29. P7L158: Can you add a reference for the used boundary layer height definition?

This is one pbl-height definition that is implemented in the MYNN-PBL-scheme in WRF. Reference added.

30. P8L171: Change the synax: "At 12 UTC the average wind speed...“. What does "over forest mean? Why don't you compare wind speeds at the location of the metmast?

We see that the syntax must be improved in this sentence. Since we discuss the impact of the forest parameterization on a larger scale in this paragraph, we decided to calculate a spacial average of the wind speed over all datapoints with forest at the two levels 187 and 54 m above ground. This is much more relevant if one speaks of a single point in time. The comparison to the met mast follows later on anyway.

31. P8L175: I don't understand this sentence: "Thus, the impact of the topography on the flow is stronger due to the strong correlation of the land use categories with the *terrain." As I understand land use categories are arbitrary numbers/classes that cannot correlate with terrain. Can you explain this differently?*

The term "correlate" is used in a sloppy way here, so we have removed the sentence.

32. P8L165: Why do you show streamlines in 58.5m and 190m AGL in Fig. 4 (by the way in the caption of Fig. 4 you say it's 187m AGL)? Wouldn't it be better to use 100m, which is the height of the met-mast? Is it necessary to show 3 hours of streamline plots in Fig.4? To me it would be more interesting to show the plot at 14:00 UTC and one after the passage of the front at 17:00 UTC. As the flow is not very different in 190m AGL for the WRF and WRF-F run I suggest to plot streamlines only at one altitude, as it's difficult to distinguish between all these lines.

The heights 58.5 and 190 m above ground are the height of the model levels from which data is shown. We understand that this plot is rather busy, but decided to keep both levels in order to show the strong difference in wind direction. We show only the time step 14:00 as also suggested in major comment 4.

33. P8L175-L180: It's difficult to understand what you want to explain here, as the location Simonsbach is not shown in Fig.4. Are these explanations relevant? To me this passage can be left out.

The Simonsbach valley is now marked in Figure 1b. This passage is however relevant to understand the impact of the forest on the flow approaching the turbine. Thus, we decided to keep it.

34. P10L182: Add: "... a northern wind component..."

done

35. P10L206: You mention the effect of the forest on the vertical wind component. Can you plot w in Fig. 5 as additional contour plot?

The w-component has been added to Fig. 5. A few lines discussing the new plots have been added to the text.

36. P10L210: Improve the syntax in: "The geostrophic wind at 5km height is with values..."

done

37. P11L221: You say that the increase in wind speed during the frontal passage was not observed. As I can see there is a strong increase in wind speed in Fig.6at about 17 UTC *on 21 September (blue curve). Can you please correct this in the text?*

A misunderstanding. We meant that it is not very well observed in the model (black and red), but observed by the tower measurements. We improved the sentence.

38. P14L235-L250: It would make more sense to interpolate both WRF and WRF-F data in space and time to the UAV measurement points and plot them the same way as the observations are shown in Fig. 8. instead of comparing observed snapshots with WRF data averaged over 1.5 hours.

A new plot (replacing Figs 8 and 9) has been generated. We have interpolated data from the UAS to 4 points along the line of the UAS. Then, UAS data at common heights is averaged to arrive at the profiles. Corresponding data of WRF, WRF-F, OF and OF-F is shown as well. This, way we compare data at common times and locations.

39. P14L245: WRF data have to be interpolated in space and time to the flight tracks of the UAV. These data can then be compared directly with UAV observations. Can you change Fig. 9 and make a scatter correlation plot by plotting observed versus simulated wind speed and wind direction. This would give a better impression if wind speeds are over- or underestimated. What is the bias and correlation between simulations and UAV observations? Please add this information to Table 1.

To better compare data from model and UAS, we decided to follow the UAS along its track virtually in the model and extract data at every time step along this trajectory. Since the flight track is only 1500 m or 10 WRF-datapoints long, we decided to average UAS-data leg-by-leg within the bounds of the WRF-grid to match WRF-data spatially. Thus, we ensure that the data we compare is representative for an area that is resolved by the WRF-model. Time-steps of WRF-data are then averaged for grid-points as long as the UAS is within that grid point. This leads to 10 data points per flight leg minus those that had to be discarded due to a low signal-to noise ratio. We plotted model results versus UAS for WRF and WRF-F in a new figure and quantified the error.

40. P15L266-P16L271: This discussion including Fig. 11 should be left out as it is just a qualitative comparison. Can you please quantitatively compare WRF, WRF-F, OF-F and OF simulations driven by either WRF or WRF-F to show what is qualitatively the impact of the forest and the different boundary conditions? Please interpolate all the simulations to the flight track of the UAV and compare with these observations.

Figure 11 has been discarded and replaced by a new figure (see previous answer). The model error is quantified.

41. P16L272-P17L283: The streamline plots in Fig. 12 are difficult to interpret and differences are hard to see. The comparison is too qualitatively done and the simulation *results should be quantified by computing correlations and biases.*

Figure 12 has been removed as well.

42. P17L279: Where is Simonsbach valley and why is this location important? It's not shown anywhere in the plots.

It is now marked in Figure 1.

43. P19L285: Wrong section reference.

corrected.

44. P19L300: Which resolution? Probably horizontal mesh size? Why don't you show these results?

The setup we tested for the 90 m resolution requires a very large inner domain which leads to rather expensive runs that were sufficiently similar to the 150 m runs for which reason we decided to use these. For this reason, we never ran the whole simulation chain with 90 m as boundary conditions and do not want to show results for only parts. A small remark to demonstrate that we did perform some sensitivity tests should suffice.

45. P19L304-L311: You only show qualitative comparisons. Please quantify the model errors for both WRF and OF. What is the benefit of OF?

Model error quantification have been added for OF.

46. P19L314: "The model indicates...". Which model do you mean? WRF or OF?

We have re-written this paragraph partially to better distinct between WRF and OF.

47. P19L316: You have to interpolate model data in space and time to the flight track and can then compare simulated with observed values.

See answer to Q39.

48. P20L319: Can you plot measured turbulence of the UAV and compare it to simulated TKE?

We decide to focus on the mean flow in this work and shall focus on TKE later on. This is also because sonic anemometers were not yet installed during the simulation period.

49. P20L323: Typo: "even further"

corrected

50. P20L329-L330: I don't understand this point. Why does the forest drag prevent upper level winds from disturbing lower level winds?

Under calm conditions, one would expect the development of a valley wind during the day. Given that this valley is very small, the resulting wind is also weak. The increased drag forces the flow to go partially around the hill and enter the valley from the north; the same direction the valley wind would have. Thus, one can observe a distinct shift in wind direction (indirectly) due to forest drag.

51. P20L331: Where can I see this? Please confirm your conclusion by quantifying model errors.

52. P20L339: Quantify your conclusions.

Answer to Q51 and 52: We modified the conclusions and refer to the statistical analysis.

53. P20L342: You say that the flight legs should be over the slope. According to Fig. 8 the flight legs were over the slope and I don't understand what you want to explain here.

We need to improve the sentence here. We meant over the slope and terrain following, as the UAS was flown with constant altitude. However, the strong shift in wind direction over the slope lies mostly below the UAS track.

Figure comments

Fig. 1: Caption: Please change "Setup of ..." to "Model domains of the WRF simulations...". The green dot in b) for the met-mast is hard to see. Can you change the colour maybe and make the dot larger?

Caption changed as suggested. The color of the dot is changed to yellow and its size has been doubled.

Fig. 2: Can you add the dot of the met-mast in both figures, please? Can you increase the axes and colorbar labels? Can you add a grid to the right figure showing the OpenFoam landuse? To me it would make more sense to plot the forest distribution in the left Figure instead of the CORINE land cover classes.

Plots modified as suggested.

Fig. 3: Please add the date that is shown (21 September 2018).

The date has been added to the caption.

Fig. 4: Please add the date that is shown (21 September 2018). Can you increase the dot and the flight path of the UAS?

The date has been added to the caption. We modified the picture and show only 14:00 UTC as suggested in the major comment 4. We increased the thickness of flight path slightly, but kept the size of the dot as is. Given that only one hour is shown now, the figure itself should be big enough.

Fig. 5: Can you add the date that is shown. What is the "first day“? Can you change the range of wind direction from 0 to 360 degrees, which is the common meteorological range of wind direction. Please include contour lines of potential temperature.

The date is added to the caption and "first day" (refering to the 21. September) is removed. Wind direction is now shown in the traditional range. Contour lines have not been added to the plot, since the atmosphere is neutral at the shown time and height. Contour lines of potential temperature do not provide additional information.

Table 1: Please add units for bias, MAE and r.

Units added for bias and MAE. The Correlation coefficient is dimensionless.

Fig. 7: I think this figure brings no additional value and can be left out.

Done.

Fig. 8: Please add the date. Why don't you show results for the WRF run? Can you interpolate the WRF and WRF-F data in space and time to the measurement points of the UAV and plot these data the same way as it is done for the UAV? This would make more sense than comparing the UAV observations with 1.5 hour averages of WRF simulations. Can you change the range for wind direction to 0 to 360 degrees and increase the contour interval for both wind speed and wind direction, as the gradients can be better seen, when the contour interval is coarser.

Since all Reviewers have criticized this plot, we decided to re-design it. Instead of contour-plots, which are more difficult to compare, we decided to show profiles of UAS, WRF, WRF-F, OF and OF-F at four different locations.

Fig. 9: Interpolate WRF data to the flight track and compare profiles directly with observations. I would suggest to make a scatter plot (observation versus simulation) instead/in addition to the profiles.

We decided to write a piece of code that follows the position of the UAS and writes out wind data at every time step. This data is used to create a new plot.

Fig. 10: Caption: the description of used colors is wrong: "... taken directly from WRF (dashed black;+) and WRF-F (dashed red)...". Please add the date.

Date added, description corrected.

Fig. 11: This figure should be left out, as it does not show any relevant information. Can you instead make a correlation plot, which shows the benefit of the OF-F simulation compared to WRF simulations and the impact of WRF and WRF-F boundary conditions?

Figure 11 removed. We have created correlation plots for WRF, WRF-F, OF and OF-F.

Fig. 12: This figure can be left out, as differences between simulations are hard to see and the model comparison should be done in a more quantitative way.

Figure removed.

References

Mohr, M., Jayawardena, W., Arnqvist, J., and Bergström, H.: Wind energy estimation over forest canopies using WRF model, European Wind Energy Association, EWEA, 2014.

Muñoz-Esparza, D., Lundquist, J. K., Sauer, J. A., Kosović, B., and Linn, R. R.: Coupled mesoscale-LES modeling of a diurnal cycle during the CWEX-13 field campaign: From weather to boundary-layer eddies, Journal of Advances in Modeling Earth Systems, 9, 1572–1594, https://doi.org/10.1002/2017MS000960, 2017.