Review of

Do convection-permitting regional climate models have added value for hydroclimatic simulations? A test case over small and medium-sized catchments in Germany

By Oakley Wagner, Verena Maleska, and Laurens M. Bouwer

This manuscript presents an evaluation of the convection-permitting regional climate model ICON-CLM 2.6.4 at 3 km resolution (ICON3km) compared to its driving model at 11 km with parameterized convection (ICON11km), focusing on the Weiße Elster basin in East Central Germany. The study assesses the ability of both models to reproduce key meteorological variables (air temperature, radiation, humidity, wind speed, and precipitation) and evaluates their suitability for hydrological impact modeling using the distributed hydrological model WaSiM.

The analysis is performed for a 10-year period, from 2005 to 2014, focusing on verifying different atmospheric variables and the capability of the atmospheric models to drive a hydrological model over a series of small or medium-size catchments.

A comparison between the two models and against observed data from different sources is carried on, with the final aim to address the potential added value of the convection-permitting model.

The manuscript is well written and structured, presents new data and in general deserves to be published, although some revisions are needed to ensure that the results are better substantiated with figures/tables and that the conclusions are fully supported by the presented evidence.

General comments:

The abstract and introduction are clear, well-structured, and scientifically sound. They effectively present the study objectives, methods, and key results.

Section 2.

A more detailed discussion of the limited time period considered for a climatological analysis needs to be included in the manuscript, in particular in view of the fact that several results are characterized by statistically not significant differences between model’s results.

> Thank you for the comment. We will expand the manuscript by a discussion on the limitations posed by the short time series.

While subsection 2.2 (regarding the observational data is well described and comprehensive) , Section 2.3 (Climate data model) should be expanded. I would suggest adding some more information on the ICON model, of the model setup and of the main parametrizations used.

> We will follow this recommendation in the revision of the manuscript.

Apart from precipitation, the other atmospheric variables are verified against observations from the nearest ground station using the closest model grid value, which can be particularly problematic for temperature. In complex terrain, the altitude of the model grid can differ significantly from that of the station.

I believe it is worth adding some comments on this aspect in the subsection 2.4 and/or in the results section in the discussion of the biases, in particular for temperature.

We looked into the effect that the mismatch of the station elevations and the elevations of the associated climate model grid cells has on the temperature estimations. As a digital elevation model, we used SRTM 1 Arc-Second Global provided by USGS at 30 m resolution. We upscaled in a two-step-process, once to 1 km keeping the native grid orientation, and then to 3 km, resp. 11 km while regridding to the ICON3km, resp. ICON11km grid. We employed the environmental lapse rate (0.65 °C/ 100 m) and found elevation-induced biases in the range of ± 1 K (see Fig. 1 below). A set of stations, particularly in the west of the study area, are located in river valleys and are therefore at lower elevation than the climate model cells. These deviations in elevation appear as a cold bias by the climate models. Other stations are located at hill tops and therefore at higher elevation than the climate model cells, which appears as a warm bias by the climate models. We will discuss this in the revised paper.

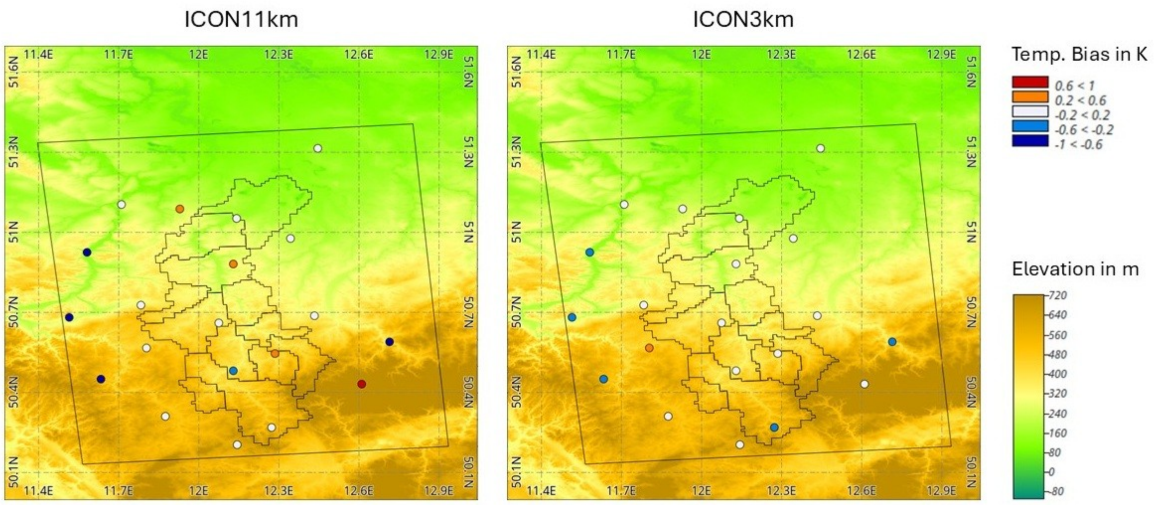

Fig. 1: Air temperature biases induced by the offset between mean grid cell elevation and station elevation for the 20 stations in the study area (dots), plotted over a digital elevation model of 30 m resolution (SRTM 1 Arc-Second Global by USGS)

Section 3.

Overall, this section provides a useful comparison of ICON3km and ICON11km, but the analysis is uneven across variables. While the analysis of wind speed and precipitation is well-supported and convincing, the sections on air temperature, global radiation and relative humidity require additional figures.

Several results are not supported by any figures/table, which in my opinion should be added at least as supplementary material.

Thank you for the comment. At the core of the paper is the study of the convection-permitting regional climate model for hydrological impact modelling. To ensure that there is a balanced focus on the climatological and hydrological analyses, we can only show a selection of plots in the main body of the paper. We will however take the comment into account and include additional supporting figures in the appendix, as outlined below.

In Section 3.1.1 (Temperature), several detailed results are discussed (frequency distributions, diurnal cycle, seasonal variability of biases, DJF vs JJA differences). However, only the monthly mean biases are shown. The absence of a figure for the diurnal cycle or the frequency distribution makes it difficult for the reader to evaluate the stated findings and much of the text remains purely descriptive without visual support.

The diurnal cycles of air temperature for the four seasons as calculated by ICON11km and ICON3km and as observed are shown in Fig. 2 below.

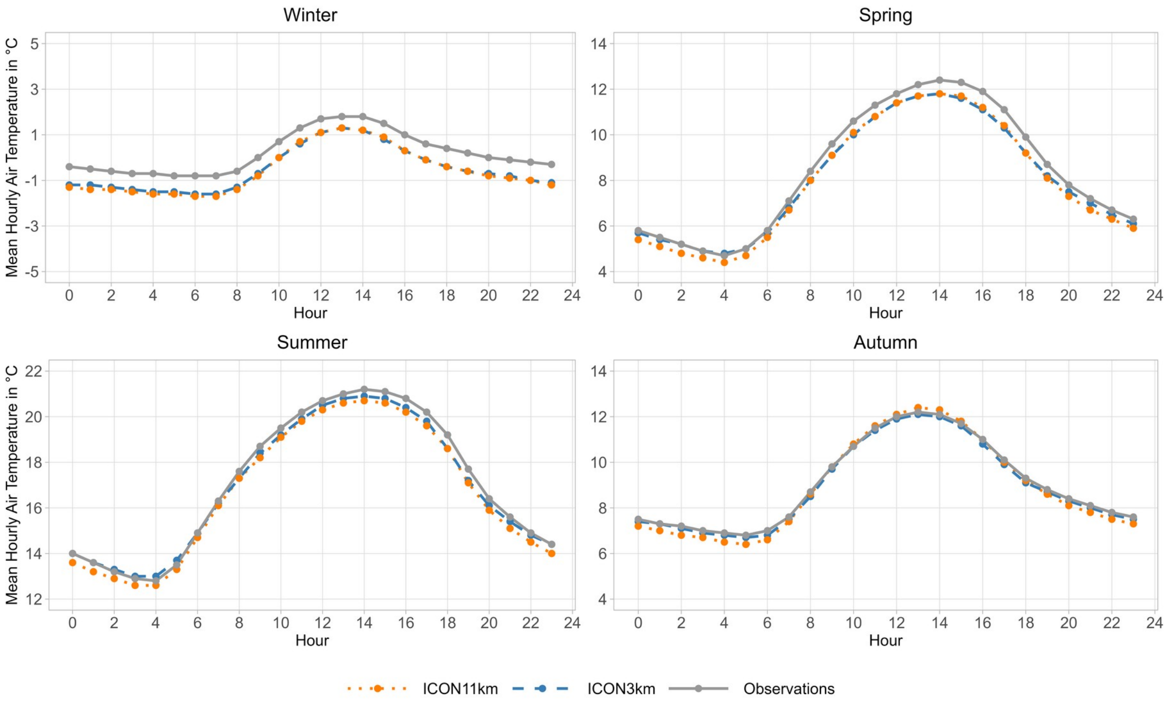

Fig. 2: Seasonal diurnal cycles of mean hourly air temperature computed by ICON11km and ICON3km, as well as the observations for the time period of 2005 to 2014

The frequency polygons for hourly air temperature as gained from ICON11km, ICON3km and observations are shown in Fig. 3 below.

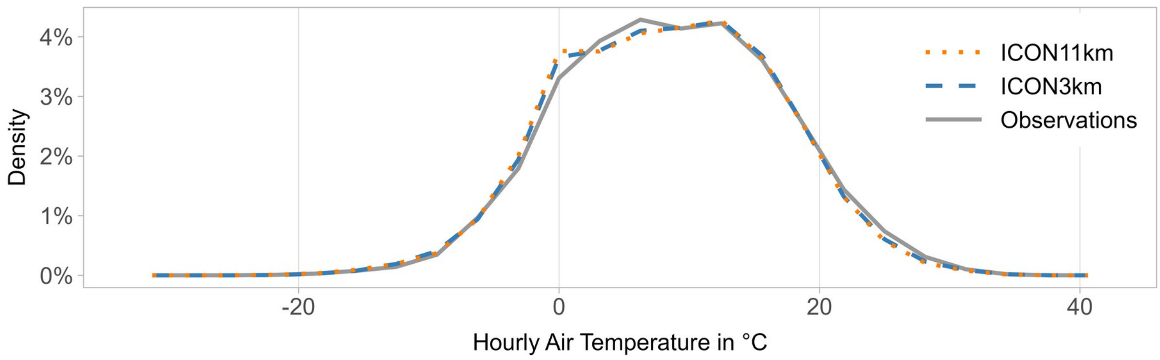

Fig. 3: Frequency polygons for hourly air temperature as calculated by ICON11km and ICON3km and as observed for the period of 2005 to 2014

We will add these figures to the appendix of the revised paper.

Similarly, in Section 3.1.2 (Global radiation) the only figure presented relates to the diurnal cycle, claims about the frequency distribution of daily mean global radiation and the monthly bias (e.g., the July improvement of 2.5 J/cm² for ICON 3km) are not supported by any figure or table . Without such evidence, this part remains insufficiently substantiated.

The frequency polygons of daily mean global radiation are shown in Fig. 4.

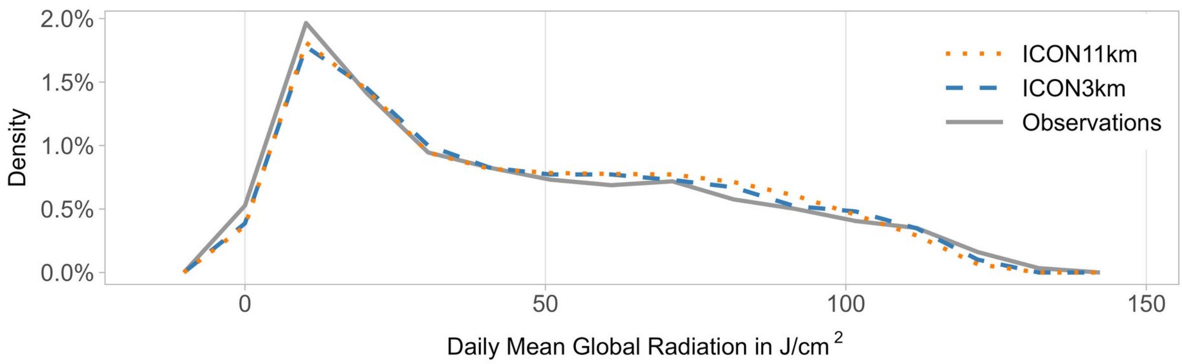

Fig. 4: Frequency polygons for daily mean global radiation as calculated by ICON11km and ICON3km and as observed for the period of 2005 to 2014

The annual variability of the monthly mean bias of daily mean global radiation is shown in Fig. 5.

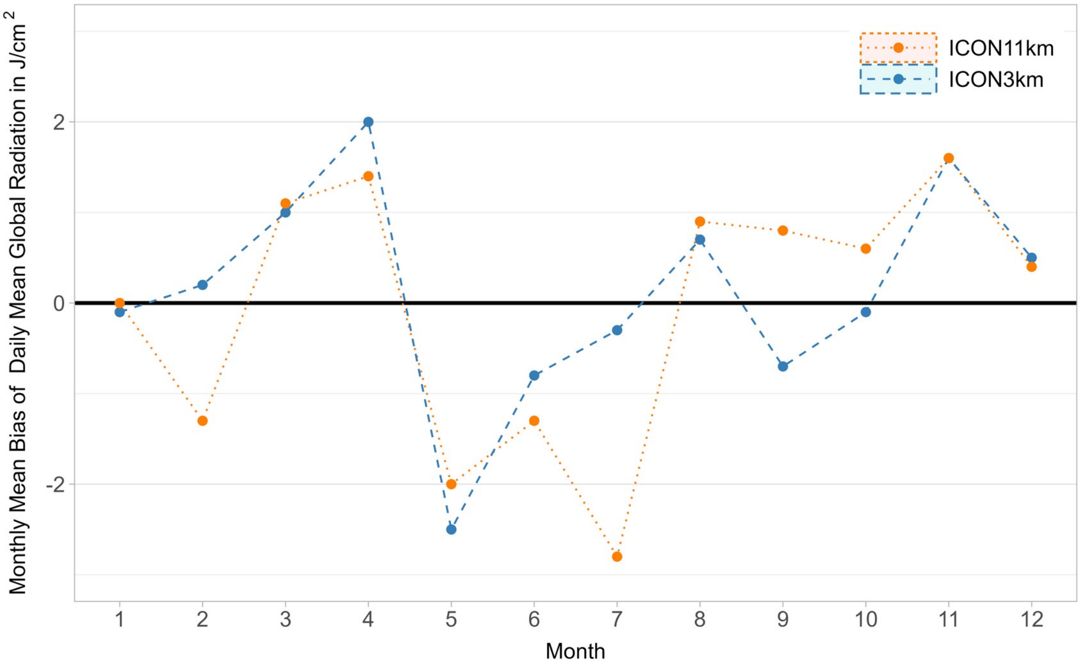

Fig. 5: Monthly mean bias of daily mean global radiation of ICON11km and ICON3km for the period of 2005 to 2014

We will expand the supplementary material by these graphs.

As regards the diurnal cycle, Figure 3 shows a one-hour shift between the models and observations throughout the entire diurnal cycle (not only for the peak) and for all seasons.

This perhaps deserves further investigation, or at least a verification of potential data misalignment, if this has not already been done.

Thank you for the comment. The climate models use the proleptic gregorian calendar and the unit of the dimension time is days since 2004-07-01T00:00:00**Z** for ICON3km and days since 1979-01-01T00:00:00**Z** for ICON11km. The time format is conform to ISO-8601, with **Z** indicating zero UTC offset. It is therefore in keeping with the observational data used, which is also in UTC. The plot of the diurnal cycle of temperature (c.f. Fig. 2 above) does not show a one-hour-shift, which suggests a correct representation of the climate model data. We did not identify a possible cause of the one-hour shift in the interpretation of the observational global radiation data and its peak around noon is plausible. Hence, we conclude that the temporal shift must indeed be intrinsic to the climate models.

Section 3.1.3 (Relative Humidity), similar to the previous section regarding temperature and global radiation, describes frequency distributions, monthly biases, and relative model performance (ICON3km vs ICON11km), but no figures are provided. As a result, the reader cannot verify whether the reported differences are meaningful or fall within observational uncertainty.

The frequency distribution of hourly relative humidity is shown in Fig. 6 below, the monthly mean biases in Fig. 7, and the QQ-plots as a reflection of model performance in Fig. 8. We will add the plots to the appendix of the revised paper.

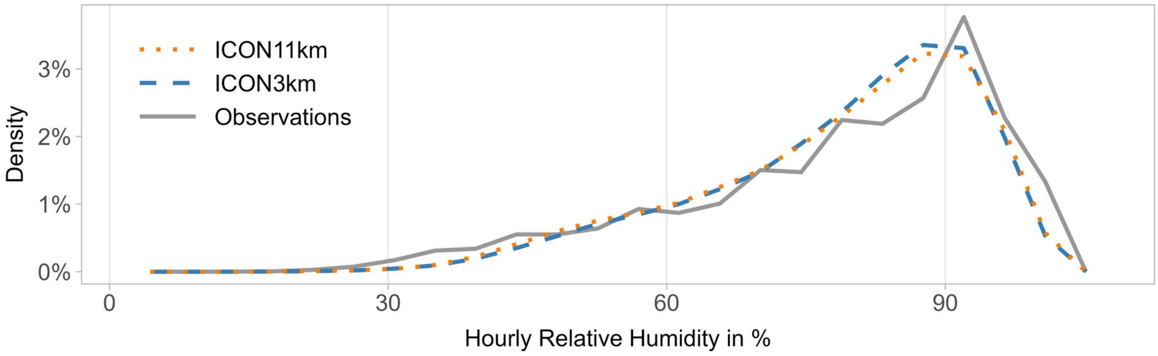

Fig. 6: Frequency polygons for hourly relative humidity as calculated by ICON11km and ICON3km and as observed for the period of 2005 to 2014

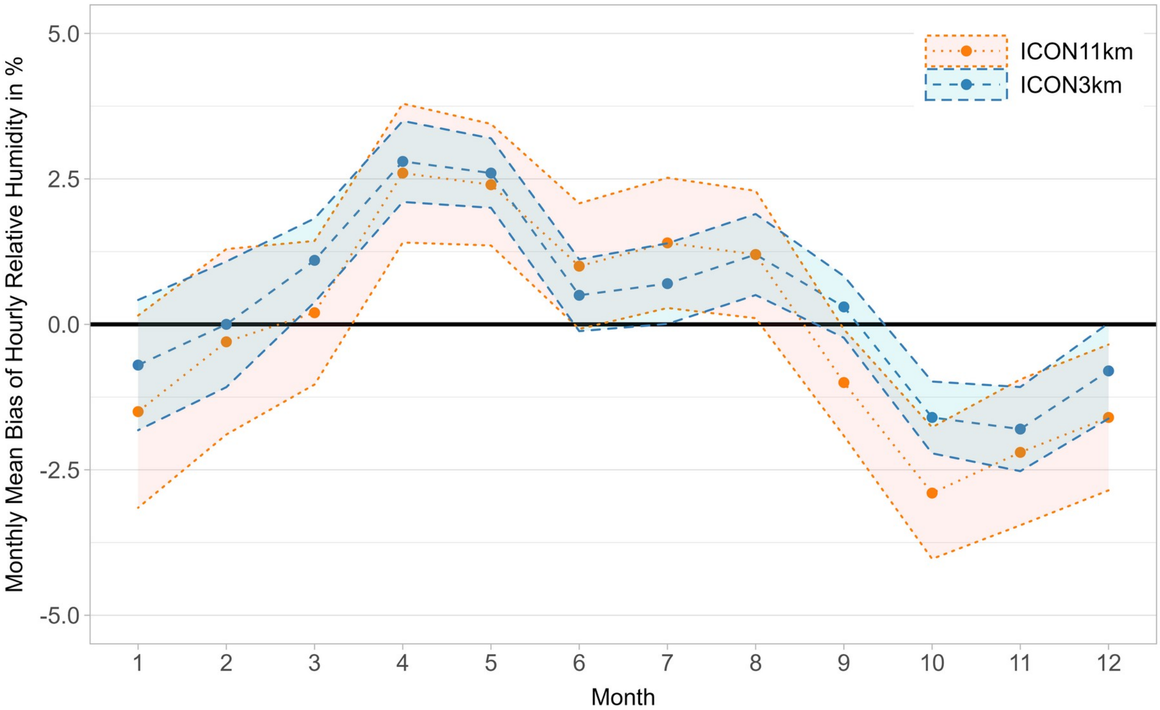

Fig. 7: Monthly mean bias of hourly relative humidity of ICON3km and ICON11km for the period of 2005 to 2014, as well as the 95%-confidence intervals from the station means

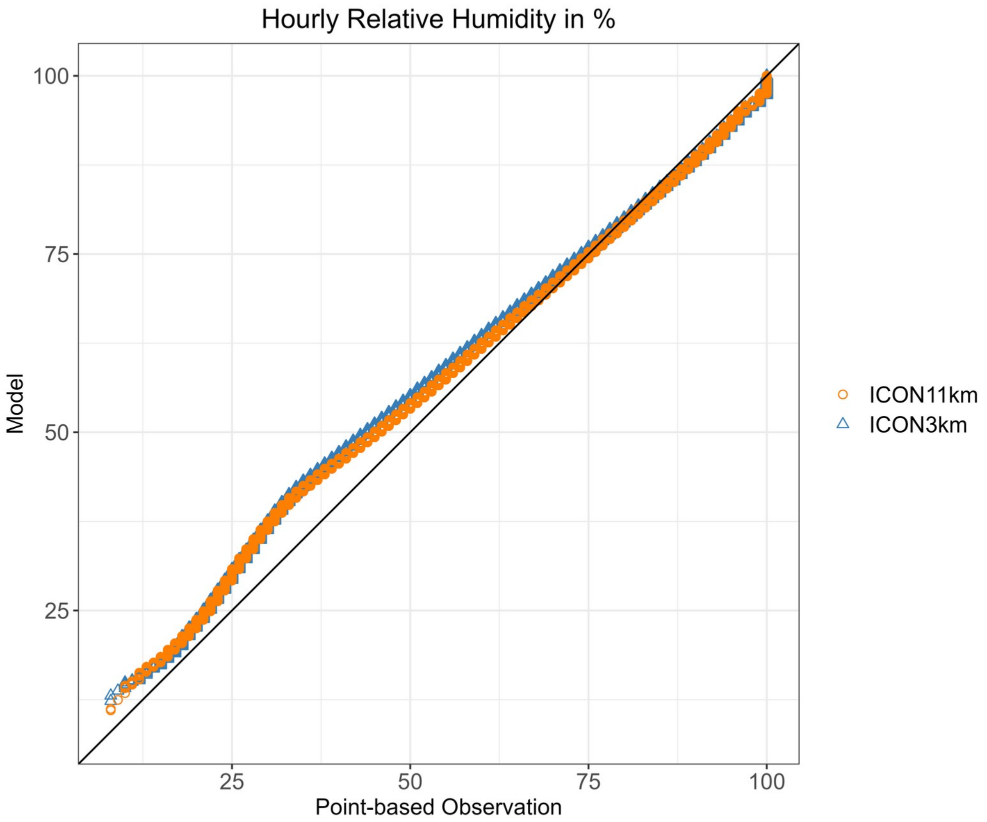

Fig. 8: QQ-plots for hourly relative humidity for ICON11km and ICON3km to the point-based observations for the period of 2005 to 2014

Section 4

This part will probably need a slight revision after the revision of section 3.

We will adjust our discussion section (Section 4) accordingly.

Section 5

Overall, the conclusions are well written, clear, and consistent with the results presented in the manuscript. The authors provide a balanced discussion of both strengths (e.g., improvements in summer temperature, radiation, and wind speed representation) and limitations (notably the overestimation of heavy rainfall and its implications for discharge modelling).

Specific comments:

l145: The bandwidth of temperature is not clear.

> Thank you, we will rewrite the temperature measurement uncertainty as ± 0.08 K to ± 0.76 K.

l150-153: The sentence is too long and could be split for clarity.

> We suggest the following rephrasing: „Monthly mean biases of hourly air temperature of ICON11km and ICON3km differ only slightly (Fig. 2) and these deviations fall within the uncertainty bandwidth of the observations. Nevertheless, the results hint towards two points of improvement by the CPRCM. Firstly, ICON3km most strongly reduces the bias of its driving model ICON11km in the summer months (JJA). Secondly, the spatial bias variability throughout the year is lower for ICON3km than for ICON11km, as reflected by the 95% confidence interval from the station means."

l160-163: The sentence is too long and could be split for clarity.

> We will split the sentence as follows: „Both ICON3km and ICON11km compute daily mean global radiation significantly higher than observed according to a Welch two sample t-test ($G_{ICON3km}$ = 45.9 J/cm², $G_{ICON11km}$ = 45.9 J/cm², G_{Obs} = 44.5 J/cm²). However, if it is assumed that the observations are systematically too low by 3%, which is also the pyranometers' measurement uncertainty (DWD, n.d.), the apparent overestimation by the climate models is not significant any more."

l192: The sentence 'Overall, for most months ICON 3km was found to outperform ..' seems inconsistent with the preceding part of the paragraph which states that " ICON3km does not seem to offer noticeable improvement in the frequency distribution" (l190-191) and " neither of the climate models shows significant difference to the observations.

> Thank you for the comment. Indeed, no improvement is visible in the frequency distribution of hourly relative humidity (see Fig. 6 above). However, results suggest lower monthly mean bias in the estimation of hourly relative humidity by ICON3km than by ICON11km for most months of the year (see Fig. 7 above). It should however be noted that these improvements are small and fall within the measurement uncertainty bandwidth.
>
> We will clarify this in the revised paper and include the plots as suggested in the comment to section 3.1.3. We intend to modify the corresponding paragraph as follows:
> „For most months of the year, monthly mean bias of hourly relative humidity is lower with ICON3km than with ICON11km. Only in the months of spring (MAM) was ICON11km found to outperform ICON3km. However, it should be noted that the largest absolute

difference in the monthly means of hourly errors between the two models occurs in September and October and is of only 1.3 % relative humidity. These deviations fall within the measurement uncertainty bandwidth.“

l247 : I suggest simplifying the phrase “in the summer month of July” to just “in July” .

Thank you. We will follow the recommendation.

l445-448: The sentence is too long and could be split for clarity.

We suggest to split the sentence as follows: „In the work presented here, the skill of estimating meteorological variables on the catchment scale is analysed for ICON-CLM 2.6.4 in its convection-permitting setup at 3 km resolution (ICON3km), and for its driving model ICON-CLM 2.6.4 at 11 km resolution with parametrised convection (ICON11km, forced with ECMWF-ERA5). Analyses are conducted exemplarily over the Weiße Elster basin in East Central Germany for the historical period of 2005 to 2014.“