Referee #1

Much work is still needed to improve the Introduction section, see comments below.
We acknowledge these shortcomings and have addresses as stated for each point below.
Besides, the authors have still not addressed the issue with limited area coupling (only catchment uses 3D groundwater model, while rest of the domain uses 1D land surface model), except stating that that there is significant change in land surface variables for the coupled domain, which is what one would expect and some discussion on scale of variables in Section 4.5. How does this discontinuity in the landsurface physics introduce uncertainty in the atmospheric variables analyzed in this study for the coupled runs?
The addition to the present version reviewed in this round on this subject was (section 2.5 – lines 196-206):

In the current version of the coupling LE and Ts (and therefore H) calculated by MIKE SHE directly replaces the corresponding variables within HIRHAM one-to-one over the shared domain, whereas outside of the domain the simple land surface scheme embedded in the regional climate model is preserved. Atmospheric fields are then updated based on the modified surface energy balance from MIKE SHE. In this study no means are implemented to assure ensuing internal physical consistency of fields within HIRHAM. Therefore, effects directly related to differences in spatial and temporal scales and in the physical formulation of the land surface scheme may be found along the boundary of the hydrological catchment. The boundary effects seen here are however relatively small, which again to a large degree is due to differences in spatial and temporal scales, i.e. to cell averaging and cancellation of errors when feeding the MIKE SHE surface back to HIRHAM. In this work we address primarily the effect of the temporal scale differences on the coupled system i.e. by varying DTI.

Therefore, we agree that the discontinuity in the land surface representation between the area described by the HIRHAM 1 D land surface model and the area described by the MIKE SHE 3D groundwater model may have some effect on the atmospheric variables. As our experimental setup did not allow distinguishing between this discontinuity effect and the effect of the improved feedback in the coupled domain, we are not able to quantify it. However, there are just as large differences between the land surface atmosphere fluxes between neighboring HIRHAM grids outside as well as inside the MIKE SHE coupled domain due to differences in soil and vegetation characteristics, so we do not expect the discontinuity just on the border between the two domains to be particularly important.
We have however added a section in the discussion also where we address this further along the lines of the above and relate it to possible uncertainties, since we sense a wish to address it in this section specifically as opposed to in the method section only.

<u>Minor Comments:</u>
Pg. 1, Ln 20: computational interaction or coupling? We use the term ‘coupling’ throughout the paper as a means of describing both the simultaneous operation and, in this process, the data exchange between associated models and therefore interdependency in terms of delivering driving data both ways.

Pg. 2, Ln 39: coupled to what? We agree that the term 'coupling' is imprecise in this context and have revised the sentence to state that hydrological processes have been 'added' to an existing vegetation model.
Pg. 2, Ln 41: Change stand-alone to offline hydrological model, and begin with new sentence. Yes – and sentence revised according to the above also.
Pg. 2, Ln 42: simulated. We do not understand this comment – the word 'simulated' is already in this line. We however revised the sentence according to the above and therefore it may be revised.
Pg. 2, Ln 44: Explain what model they used. Good suggestion. This is done for both studies in this sentence.
Pg. 3, The authors mention very little about the previous studies, only emphasizing the number of hours of simulation carried out by different coupled modeling studies but fails completely to expand on the important findings of these studies, which is more relevant in context of this paper. We agree that this could be explained better and have included this accordingly.
Pg. 4, Ln 82: replace long-term with annual simulation. At the same time, since the coupling is done over a catchment only, not the entire domain of the atmospheric model, this should be mentioned here as a limitation as there is a discontinuity in model physics on how the lower boundary condition is computed for the atmospheric model. This is revised accordingly.
Pg. 8, Ln 189: Three dimensions variable exchange? That is not what we are doing no, nor trying to state, so this sentence is revised for improved understanding.
Pg. 12, Ln 296: The sentence does not correspond to the figure. In a previous sentence we mention that the largest variability for PRECIP and Ta is seen for the DTI runs. This sentence should therefore reflect only the variability between the CV and HUV runs, which it clearly fails to do. We are glad this was discovered and have revised to properly explain this.
Pg. 13, Ln 298 to 301: RMSE or variability? RMSE – it is correct as it is.
Pg. 13, Ln 303 to 305: Does not follow above paragraph and heading, confusing. We agree and have tried integrating it better with the previous paragraph also in terms of description.
Pg. 13, Ln 311: But observed precipitation also decreases right? Yes – really good point which should of course be included in this sentence. Revised.
Pg. 14, Ln 323-334: Give background on why these time periods were chosen? The figures do not illustrate much either. The periods were chosen as a middle week in the hottest month resulting in a high degree of dynamics. For precipitation however more could be seen for august. The one-week (and one-month for precipitation) durations was chosen according to how busy the resulting plot was. This is explained in the text also.
Pg. 14: Ln 345: The response in MIKE SHE from the coupling is generally low and the connection between DTI rate and model performance has opposing directions as also shown in figure 8.
Pg. 15: Ln 347-Ln 364: Again why this period chosen, give background. Is it even compared with observations? The period is chosen due to the same reasons as above. The results are indeed compared to observations both in time-plots and in terms of statistics – the statistics is included in e.g. table 2. We have explained the reason for not including the observations in these figures earlier (maybe another reviewer), and we take the liberty to reuse the answer here (covering several aspects on this), as the reviewer may be changed:
We have chosen not to add the observations to figure 5, 7 and 9. The figure will be (even) more incomprehensible and voluminous and we would have to discuss the specific dynamics of observations against simulations (for the shown period only) which would make an already extensive

paper even larger. This is not the intent of the figure. The intent is instead to show that the CV, HUV and TI fall in distinct groups of comparable patterns related to their simulation condition (being coupled (perturbed), coupled (varying DTI) or uncoupled (perturbed)). A discussion on hour-to-hour simulation performance against observations would also be limited by the issue of comparing an 11 km grid output with gridded observation data from point stations (is this really feasible?) and this would have to be addressed again increasing the paper volume. Instead we address observations against simulations by longer term run statistics.

Pg. 18: Ln 441: performance . In … Yes. Revised.
Pg. 18, Ln 443: tend to underestimate Yes. Revised.
Pg. 18, Ln 445, e.g., Revised.
Pg. 19, Ln 450, higher degree or low frequency coupling? Higher degree as stated. We refer to the degree of coupled cells as seen in figure 1 and have revised the sentence to more clearly reflect this.

Referee #2

2nd review of the manuscript „Results from a full coupling of the HIRHAM regional climate model and the MIKE SHE hydrological model fora Danish catchment" by Larsen et al., Hydrol. Earth Syst. Sci. Discuss., 11, 3005–3047, 2014

The authors did a reasonable job of addressing my questions. However, to finalize the review, there are two remaining issues:

1) The authors often claimed that they considered my recommendations in the manuscript. While I appreciate this very much, it is impossible for me to find and to evaluate these changes, if these are neither marked nor cited in the response. It is essential that the modifications are clearly indicated in the revised manuscript. This should easily be possible by using track changes, citation of the insertions/changes or using different colors. We have provided the editor (also uploaded as supplement) with a version where each previously revision point addressed in the text is given a number and the corresponding revision in the text is highlighted with that number.

2) I do not agree that downscaling of the HIRHAM results to higher spatial resolution adapted to MIKE SHE would not have a positive impact (see Point 2). In this connection, I did not state that HIRHAM is "too coarse" for driving a hydrological model but I pointed out that a better representation and interaction of physical processes can be achieved, if further dynamical downscaling of HIRHAM is applied. Ok. We agree that resolutions in the order of the 11 km, as used here, are not closing the resolution gap between coupling atmospheric and land-surface hydrological processes, even though being in the higher resolution range of RCMs. We therefore also agree that further downscaling could be beneficial especially for regions with a high degree of localised/convective rainfall.

I do not expect that the authors are adding additional model runs to demonstrate this in order to avoid extensive work. However, recent results in regional downscaling using convection-permitting resolution, which clearly demonstrated the improved linkage between forcing data and hydrological output, should be considered in the discussion and the summary. This is a good suggestion and we have added results from three recent studies on this in both sections (Klüpfel et al. 2011; Berg et al. 2012; Xue et al. 2014).

As requested in point 1 we point out that these additions are at lines 74-77 and 505-510 in the new version.

References

- Berg, P., Feldmann, H. and Panitz, H.-J. (2012): Bias correction of high resolution regional climate model data
- Klüpfel, V. Kalthoff, N., Gantner, L. and Kottmeier, C. (2011): Evaluation of soil moisture ensemble runs to estimate precipitation variability in convection-permitting model simulations for West Africa.
- Xue, Y., Janjic, Z., Dudhia, J., Vasic, R. and De Sales, F. (2014): A review on regional dynamical downscaling in intraseasonal to seasonal simulation/prediction and major factors that affect downscaling ability