Response to Reviewer 2

General comments

The authors effectively addressed my main concerns regarding the improvements achieved with the new model version. In this regard, I find Figure 2 quite compelling. I believe that, at the current state, the manuscript is satisfactory from a scientific content standpoint and has the minimum requirements for publication on GMD.
Nonetheless, there are still many parts that require clarifications, and the study needs many improvements in terms of style, coherence, and notation. In particular, Section 2.1 needs to be greatly improved in clarity, and the authors are required to carefully review all equations.

[Response] We sincerely thank the reviewer for the constructive comments that helped us improve both the clarity of the manuscript. We have carefully revised the manuscript summarized in the following. For more details, please find our point-by-point responses to all reviewer comments below (the line numbers refer to the tracked text).

Summary:
(1) We clarified the interpretation of the drying climate scenario and revised Sect. 3.4 to ensure consistency between the experimental setup and the analysis (see reply to comment(s) #1).
(2) We improved the interpretation of residue management effects on hydrological processes, and more clearly stated the limitations of the model in Sect. 4.5 and Conclusion (see reply to comment(s) #2).
(3) We corrected the methodology description (Sect. 2.1.2 and 2.1.3 (see reply to comment(s) #3).
(4) We clarified the model parameterization (see reply to comment(s) #4).

Please find below a non-exhaustive list of required corrections. The authors are also required to extensively review the overall manuscript by carefully choosing wording and improved logical flow.

[Response] Thanks for the helpful suggestions. We carefully re-read the entire manuscript and revised the text to improve wording, clarity, and logical flow throughout the manuscript with some support of the large language model ChatGPT-5.5. All suggested changes were carefully evaluated and manually revised by the authors.
The main revisions can be found in Lines 21-29, 34-39, 60, 68-69, 78-79, 254-256, 263-264, 267, 293-295, 330-341, 359, 391-392, 516-525, 598-601, 630-634, 640, 660-667, 713-725, and 747-752. We also improved Fig 1 and the legend to simplify the flow chart.
Additional minor corrections, including improvements to prepositions, verb tense, word order, and singular/plural usage, were made throughout the manuscript but are not listed here.

Furthermore, I still have some major concerns:
1. A methodological concern remains regarding the "drying climate scenario", despite the authors' reply to my previous review. The authors mention that one meteorological year is employed to force the model, but I still fail to see how the 10-year simulations in the dry scenario differ across years. From section 3.4, it seems that results from the dry simulation are the same from one year to the next. If there is barely any difference between the years, then the pairwise comparison between "dry" and "non-dry" years is wrong: in this case, the authors should thoroughly review that section and their overall approach regarding the "drying climate scenario".

[Response] In the dry scenario, the model is forced by meteorological data from a single dry year, which is repeated over the 10-year simulation period. Therefore, we agree that there is no interannual variability in climate forcing, and model results are not expected to differ substantially from one year to the next.

The purpose of our idealized scenario is to represent persistent dry conditions and to maximize the likelihood of cumulative effects of crop residues over a 10-year period, rather than to represent future climatic conditions (including interannual climate variability). We changed the label from 'drying climate scenario' to 'dry scenario' to avoid confusion.

Following the reviewer's suggestion, we have revised Sect. 3.4 to clarify this point and remove wording related to interannual comparisons.

Line 536-549: "**In the dry scenario, the meteorological forcing corresponds to a single dry year that is repeatedly applied over the 10-year simulation. This idealized experiment is designed to explore the sustained response of the system under persistent dry conditions, rather than to represent future climatic conditions or interannual climate variability.**

Under these conditions, the presence of crop residue shows a similar increase in SWC at 0.25 m soil depth **under** current climate, with an average increase of 0.23±0.53 kg m^{-2} (3.01±5.97%) compared to simulations without residues (Fig. 5a).

Although no systematic temporal evolution is observed in near-surface SWC (at the soil depth of 0.25 m), deeper soil layers (50% soil depth) exhibit a slight gradual increase of soil moisture over the simulation period (not shown) due to sustained reduced E_{soil} which increases downward water transport. The decrease in E_{soil} by crop residues allows more water to remain in the soil, leaving more residual moisture available for the next growing season.

Average T_{soil} at 0.25 m soil depth decreases with a statistically significant trend by an average of -0.42±0.80 °C (-8.42±3.58%) over years (Fig. 5b), consistent with the sustained effects of residue-induced surface cooling."

2. It seems to me that the authors claim that residual managing has uncertain implications in terms of climate change mitigation/adaptation (which is reasonable), but it could be useful as an agricultural practice. However, unless I've missed something, the authors evaluate differences between ORC-D and ORC-AE, and conclude that in terms of SWC there are no significant differences. I assume that ORC-D here would represent a configuration in which there is no residual managing (bare soil), while ORC-AE maintains the residuals. Based on just statistical significance, it seems that residual managing might be a marginally relevant agricultural practice. Regardless of how the authors intend to address this point, the conclusions should better reflect and convey the hypothetical nature and uncertainties of their evaluation.

[Response] We agree that the effect of residue management on SWC is small in our results i.e. statistically non-significant differences in SWC between ORC-D and ORC-AE. The present model configuration represents the biophysical effects of crop residues through surface albedo, roughness effects, and uses a simplified treatment of soil evaporation resistance. Tillage effects on soil hydraulic properties and longer-term changes in soil organic matter and soil physical characteristics are not represented.

Following the reviewer's suggestion, we have revised the manuscript to clarify these assumptions and related uncertainties.

Line 741-746: "**While our modelling shows minor effects of crop residues on SWC, we did not account for effects such as changes in soil texture or organic matter content (Singh et al 2018), which could change soil water availability to crops. In addition, we did not account for interactions with other managements such as tillage, which are often co-applied. In summary, the present simulations are**

particularly well suited for assessing the biophysical impacts of residue cover on the surface energy balance, while the representation of soil-water responses remains an area requiring further model development and the corresponding results should therefore be considered with caution."

In addition, we revised the conclusion to better emphasize the most robust impact of albedo changes, whereas hydrological and management implications remain more uncertain and require future model development:

Line 788-790: "**Future model developments should refine the representation of crop residue effects on soil physio-chemical properties and interactions with management practices such as tillage in order to improve simulations of plant water availability."**

3. Where does the 0.5° spatial resolution come from (e.g. L139)? The ORCHIDEE-CROP model is stated to have a 0.5° spatial resolution (L152-153), but it's unclear how this is relevant, since the model is forced by local measurement sites data. It seems to me that what is being doing is just running the model at sites location: if that's the case, shouldn't resampling gridded data at 0.5° be replaced just by interpolation at the measurement site location?

[Response] We apologize for the confusion. The 0.5° resolution is not relevant here and has been removed from the manuscript.

We have revised Sect. 2.1.2 and 2.1.3 accordingly to address this issue:

Line 150-152: "At the two ClieNFarm sites without flux observations, the meteorological variables from the hourly 0.25° ERA-interim product were **extracted** and interpolated to **exact site coordinates to provide meteorological forcing data**".

Line 165-167: "**Since the model was run at the site scale, bare soil albedo values were extracted for each site location, and linear interpolation was applied to convert the data to a daily time step, aligned with the model temporal resolution.**"

4. L301: So Hrough can be expressed as $H_{rough}=f_{(crop,max)}\times 0.5\times(1-H_{dis})$, where 0.5 basically replaces the wheat canopy height in the original formula. So, the formula would be the same as the one above and the only thing that changed is the canopy height considered? The same stands for Z0.

[Response] Yes, the overall structure of the formulation remains the same. The main change is that the canopy height used in the calculation is reduced from the winter wheat height to a representative residue height (0.5 m), and the crop fraction is replaced by the residue fraction during the residue covering period. Therefore, the formulation itself is unchanged, but the physical interpretation and parameter values differ to account for the presence of crop residues after harvest. The same reasoning applies to Z_0, where the parameterization follows the same structure but uses residue-related parameters.

We have revised Sect. 2.3.4 to clarify this point.

Line 332-333: "**These two formulations remain identical to Eqs. (7) and (8), but with adjusted parameter values for canopy height and fractional cover to represent crop residues.**"

Minor comments:

L117-118: Please add an introductory sentence, explaining that two different sources were used (ICOS and ClieNFarm)

[Response] We have added an introductory sentence at the beginning of Sect. 2.1.1 to clearly state that two different data sources (ICOS and ClieNFarm) were used.

Line 124-127: "To derive surface albedo (α_{surf}, unitless) from site observations, **we combined datasets from two sources: (1) eddy-covariance measurements from the Integrated Carbon Observation System (ICOS) Data Portal (Dumont et al., 2023; Schmidt et al., 2023; Buysse et al., 2023; Brut et al., 2023; Brümmer et al., 2023; Tallec et al., 2023; Bernhofer et al., 2023), and (2) satellite-derived products at farms belonging to the ClieNFarm project.** "

L13 of supplementary: title would be "Quality filtering and preprocessing of surface albedo data at ICOS sites"? To differentiate them from the ClieNFarm sites

[Response] We replaced the previous title in SI with "Quality filtering and preprocessing of surface albedo data at ICOS sites" to differentiate them from the ClieNFarm sites"

Line 13: "Text 1.1 **Quality filtering and preprocessing of surface albedo data at ICOS sites**"

L128-129: should this sentence be "Across the selected 10 sites, we chose only the years in which winter wheat was cultivated, and obtained a total of 33 winter wheat years"?

[Response] We revised this sentence in Sect. 2.1.1:

Line 139-140: "**For the selected 10 sites, we chose only the years during which winter wheat was cultivated, and obtained a total of 33 winter wheat years.**"

L142: I suggest removing "including precipitation, air temperature, wind speed, humidity, and radiation" – stating "All meteorological variables" would already be enough

[Response] We removed the repeated words according to this suggestion.

Line 155-156: "All meteorological variables from this year were then used to force the simulation."

L151: If "S2 α_surf" is referring to the Sentinel-2, I suggest using "S2-α_surf" for clarity

[Response] We replaced S2 α_{surf} with S2-α_{surf} in Sect.2.1.3.

L158: "LAI was used primarily to constrain", isn't the term 'primarly' redundant? Based on what the authors said it was used just to constrain the onset, correct?

[Response] We have revised the description in Sect. 2.1.3.

Line 172-173: "MODIS-derived *LAI* was used to constrain the onset of crop maturity period, rather than to reproduce absolute *LAI* values."

L166: why winter wheat and maize instead of just winter wheat?

[Response] The reference to both winter wheat and maize relates to the calibration of the original ORCHIDEE-CROP model version used as a starting point in this study. In the present work, we only simulate winter wheat. We have clarified this point in the revised manuscript to avoid confusion.

Line 182-184: "This study employs a **previously developed model** version calibrated **against** *LAI*, flowering and harvesting dates, and crop yield for **both** winter wheat and maize

(https://github.com/yangsssu/ORCHIDEE-CROP.git). **In the present study, only winter wheat is considered.**"

L178: "SW_IN↓" is used in the formula, but shouldn't it just be either "SW↓" or "SW_IN"?

[Response] We deleted "↓" in equation 3.

L196-200: is this formula correct? At first glance the units do not add up. What are the units for the beta terms?

[Response] We apologise for the lack of clarity in the original manuscript. The time step was not explicitly included in the equation. The bulk formulation on the right-hand side yields evaporation in kg m^{-2} s^{-1}, which is integrated over the model time step (dt) to give kg m^{-2} per time step.
In the model, this quantity is then expressed in mm day^{-1} using the equivalence 1 mm = 1 kg m^{-2}.
In addition, the terms β_1, β_4 and β_5 are dimensionless scaling factors representing soil moisture and resistance effects, and therefore do not affect the dimensional consistency. We have revised the equation and clarified the unit definitions in the manuscript.
Line 216-222: "$E_{soil} = (1-\beta_1) \times (1-\beta_5) \times \beta_4 \times r_{au} \times v \times C_d \times (q_{surf} - q_{air}) \times dt$ (5)
Where β_1, β_4 and β_5 **are dimensionless scaling factors representing soil moisture and surface resistance effects** for snow sublimation (β_1), soil evaporation (β_4) and floodplain evaporation (β_5), respectively. r_{au} (kg m^{-3}), v (m s^{-1}), C_d (-), q_{surf} (kg kg^{-1}), and q_{air}(kg kg^{-1}) denote the air density, wind speed, drag coefficient, saturated surface air moisture, and specific humidity, respectively. **The bulk formulation yields evaporation in kg m^{-2} s^{-1}, which is integrated over the model time step (dt) to obtain values in kg m^{-2}. This quantity is then expressed in mm day^{-1} using the equivalence 1 mm = 1 kg m^{-2}."**

L210: Please improve formatting of the equation (use bigger parentheses). The same stands for the other equations

[Response] We have revised the formatting of the Eqs (6), (7), (14), (15) and (17) by using appropriately sized parentheses to improve readability.

L215: I believe that a more standard notation for canopy height could be h_c (or H_c if the authors want to be more consistent with the rest of the notation)

[Response] We have replaced 'Height' with H_c in Sect. 2.2.4.

L222: The meaning of "weighted height of vegetation with the maximum fcrop for winter wheat in each grid" is unclear

[Response] We have revised the description in Sect. 2.2.4 for clarification.
Line 245-247: "...and the vegetation **height weighted by** the maximum f_{crop} ($f_{crop,max}$) **of** winter wheat **within** each grid"

L225: H_ave appears on both left- and right-hand side of the equation (which I assume it's an error)

[Response] We agree that the original formulation was incorrect. The equation was intended to represent an area-weighted average of vegetation heights within each grid cell. In general, this equation is:

$$H_{ave} = \sum_i \ f_i H_i$$

where f_i is the fractional coverage and H_i is the corresponding crop height.

However, in this study we only consider winter wheat within each grid. Therefore, the formulation simplifies to:

$$H_{ave} = f_{crop,max} \times H_c$$

We have corrected the equation 10 in Sect. 2.2.4 and clarified this point in Sect.2.2: **In the present study, only winter wheat is considered**.

L274: meaning of k3 and k4 should go after eq. (14)

[Response] We explained k_3 and k_4 after Eq. (14).

Line 302-303: "Where α_{max} is the simulated α_{crop} on the last day of the albedo increase phase. (1 - f_{res}(t)) represent the daily variation of bare soil fraction. k_3 **and** k_4 **represent the fitting parameters that need to be calibrated at the site scale (see below).**"

L328: "Data from the remaining five sites with five winter wheat years" – why five sites? Isn't it just the remaining five years?

[Response] We deleted "five sites with" in Sect. 2.4.2 to be specific.

Line 362-363: "Data from the remaining five winter wheat years (n=179 daily α_{surf} observation) were used for model testing."

L332: "Simulations were performed for the 10 sites"

[Response] We have revised this sentence to be specific.

Line 366: "Simulations were conducted for 10 sites **encompassing 14 farms** (Table S2)"

L345: Fix Table 1 caption

[Response] We have revised the caption of Table 1 to improve clarity and make it more informative.

Line 380: "Table 1 **Overview of** simulation experiments **and model configurations.**"

L371: "of the old and new models and observation". Do the authors mean "the comparative performances of the old and new models against observations"?

[Response] Yes, we intended to refer to the comparative performance of the old and new models against observations. We have revised the wording accordingly for clarity.

Line 411-413: "With the calibration of the simulation of harvesting date for all sites, we further compared the **simulated** surface temperature (T_{surf}), *LE*, *H* from the old and new models **against** observation during the foliar yellowing and residue covering periods."

L372: "to assess the statistical significance of the results" is unclear. Please state clearly the null hypothesis that is being tested.

[Response] We have revised the text to explicitly state the null hypothesis being tested. In the paired t-test, the null hypothesis is that the mean paired difference between the simulations from the two model configurations is equal to zero. Statistical significance is then assessed based on this hypothesis.

Line 413-415: "**For each variable, a** paired t-test was also used to **assess the null hypothesis that the mean paired difference between the simulations from the old and new models differed significantly from zero.**"

L435: "ORC-AE indicates" instead of "ORC-AE suggests"

[Response] We have revised the wording accordingly in the legend of Figure 4 by replacing "ORC-AE suggests" with "ORC-AE indicates".

L505-506: Why would both positive and negative perturbations of k1 lead to enhancements of surface cooling?

[Response] In our model, k_1 controls the linear increase of residue albedo during the first 15 days, and also determines the albedo value at the transition to the decay phase (Sect. 2.3.2). This value influences the subsequent exponential decay of albedo. At the same time, the residue fraction gradually decreases over the 90-day period. As a result, changes in k_1 affect the full temporal evolution of the product of residue albedo and residue fraction, which determines the cumulative reflected radiation. This leads to a nonlinear response in surface cooling, where both increases and decreases in k_1 can enhance cooling, although the effect is stronger for positive perturbations.

We have clarified this point in the revised manuscript.

Line 566-568: "**This reflects a nonlinear response, as k_1 controls not only the rate of albedo increase but also the initial value of the subsequent decay phase (Sect. 2.3.2). Combined with the gradual decline in residue fraction over time, changes in k_1 modify the cumulative contribution of residue albedo to surface radiation, leading to enhanced cooling in both cases.**"

L527: I wouldn't say that the control is "even stronger" – to me it looks like the responses are usually below the 2°C cooling

[Response] We have removed "even stronger" in Line 590 in the revised manuscript.

L531: missing space between "increased" and "Z0"

[Response] We have corrected this error in Line 594.

L548: comma instead of period before "while"

[Response] We have revised this sentence in Sect. 4.2.

Line 611-613: "Our simulations revealed that the surface-atmosphere interactions (ORC-E) dominate the reduction in *LE* in both periods (Fig. 4e and Fig. S6b), with this effect jointly mediated by changes in β_4 and Z_0 (Fig. S11d)**, while** α_{surf}-driven radiative impact is less important.**"**

L568: "both and their"?

[Response] We have deleted "their" in Line 633 to correct this error.

290
291 *L621: "while removing the residue cover, temporarily boosting $Esoil$ and LE after the end of residue*
292 *covering" seems an anacoluthon. Please check wording.*

293 **[Response]** We have revised the wording in Sect.4.3 to improve clarity and readability.
294 **Line 689-691**: "Early tillage, as observed at the FR-Gri site in 2018 (three weeks after harvest), **retains**
295 more *SWC* in the topsoil **and removes** the residue cover, **thereby** temporarily **enhancing** E_{soil} and *LE* after
296 the end of residue covering (Fig. S12)."
297
298 *L627-628: Check wording.*

299 **[Response]** We have revised the wording in Sect.4.5 to improve clarity and readability.
300 **Line 696-697**: "Despite improvements **to** the ORCHIDEE-CROP model **in simulating** crop albedo
301 changes **associated with foliar** yellowing and **crop** residues, **as well as** their biophysical climate impacts,
302 uncertainties and limitations **remain**."