We thank the reviewers and editor for their constructive comments on our paper and appreciate the helpful suggestions.

This paper is concerned with developing workflows to derive parameters and the <u>evaluation</u> of the resulting model performance; we have clarified this in the revision and improved the paper in the following aspects:

- Expansion of data sources analysed to the 38 sites from the now available FLUXNET2015 dataset.
- Expansion of surface conductance parameters evaluation to include the both the FLUXNET2015 site derived values but also the literature-based values of NOAH (a popular land surface model used in NWPs, e.g., WRF).
- Only vegetated areas are analysed in three groups: evergreen and deciduous trees, grass (including crops). Bare soil and water are no longer included in this work as less data are available to generalise findings compared to other land cover types in an appropriate way.
- Generalised workflows are provided to derive model parameters with variability/uncertainties (e.g. standard deviations and/or inter-quantile ranges).

Our responses below refer to the **new** Section/Figure/Table/Appendix in the revised manuscript unless otherwise indicated. Given we have made substantial changes to the paper some comments are no longer applicable, so we indicate as N/A. Some comments are applicable, elsewhere in the revised manuscript, and we have taken them into account.

Reviewer 1

General Comments:

1. *This work concerns use of the SUEWS model in non-urban areas. The manuscript includes some recent developments to the SUEWS model, some analysis of observed data from eddy covariance sites, estimation of SUEWS model parameters relevant to latent heat fluxes using the observed datasets and assessment of model performance for different time periods. This study includes many components, yet the overall purpose of the paper is unclear. Is it new model developments, new parameter values, a new method for parameter derivation or an assessment of model performance? It seems likely the authors would like to tackle several, if not all, of these aspects here, but lack of a clear structure makes the manuscript very hard to follow and, in my opinion, none of these aspects are covered in sufficient detail.*

2. *One of the main problems is that the manuscript is not well organised. It often reads more like notes than a journal article and it is very hard for the reader to follow what has been done and why. It is not necessary to have six appendices when the main text is only about 450 lines! The figures and tables should also be improved (many are not especially useful and there seems to be a lot of repetition).*

3. *This work would be of much greater use if the findings were analysed at a deeper level and set in context against the literature. In general, more evidence of awareness of the literature is necessary. As an example, the authors refer mainly to the body of work on the SUEWS model, in particular the recent model development papers of Järvi et al. (2011) and Ward et al. (2016). One question that arises is why the current study is needed at all, given that much of SUEWS was originally based on non-urban models and parameters. Here the authors take the recent 'urbanised' sub-models and parameters and seem to 'un-urbanise' them again.*

4. More detailed suggestions for improvement are given below, along with questions about various parts of the methodology. Providing these are given consideration, I believe the manuscript can be substantially improved. I therefore recommend publication after major revisions.

Our focus is the determination of parameters for use fully vegetated areas that commonly occur adjacent to cities.

We have restructured the paper and removed the non-vegetated land covers to simplify the storyline.

- The original parameters in SUEWS for conductance were all based on urban environments, so the analysis of the non-urban parameters is all new.
- More discussions on relevant literature have been added along with expanded analysis.

Major Comments:

5. Throughout the Methods and Results section, the various aspects of the study are mentioned interchangeably so it is not clear whether observations, model output, calibration or evaluation is being discussed. Currently the model description is spread throughout the Methods section.

a) I suggest first adding a section where the model is described along with the equations, including the new developments currently in Appendix A (these seem quite important, especially for a GMD paper, and I'm not sure why these are in the appendix whereas LAI is in the main text). For this to be a standalone publication, more general details about the SUEWS model also need to be given (e.g. what are the required inputs, what are the outputs, what scale does the model operate at). This section should simply describe what the model does, without including any methodological details about the parameterisation or evaluation approaches in this study.

The new paper structure is:

- Section 2: Vegetation related physics in SUEWS.
- Section 4: Model parameters derivation workflows

b) The readability of the section describing the observations should be improved so that the reader quickly gets an overview of the sites and starts to feel familiar with their different characteristics. For example, mentioning the site names in the text, not just the table, and giving a very brief description. It would also help if the abbreviations of the sites contained the land cover code instead of the country (which is relatively unimportant). Some details about quality control of observed data should be given. It would also be useful here to mention the representativeness of the site years (e.g. the low rainfall mentioned in L342). Then there should be the section on how each of the model parameters were derived. Seeing as this is a key part of the manuscript, more details are needed about how these parameters were fitted. It is currently not at all clear how this was done (using multiple SUEWS runs with various parameter values and minimising the MAE?). What range of parameter values was considered? Was more than one parameter allowed to vary at once to allow for interdependencies? Was any bootstrapping done?

The new paper structure includes:

- Section 3: Datasets used
- 3.1 FLUXNET2015 - by using one source for the flux data for simplicity.

    Also, we keep site ID in the format "country-site" to be consistent with the FLUXNET naming convention as it is widely adopted in FLUXNET-related studies.

- 3.2 MODIS for LAI
- 3.3 SoilGrids for soil properties
- Section 4: Model parameters derivation workflows

6. There also needs to be a section describing the approach used for the SUEWS runs, e.g. spinup, initial conditions, forcing variables used.

The new paper structure includes:

- Section 5.1: Model configurations.

7. *Having so many appendices for a short paper is not helpful. Much of the material in the appendices is not useful anyway and should be edited as strictly as in the main text. Here are some suggestions.*

a) *Appendix A seems to be important and should be in the main text (in the new model description section). If I understood correctly, SUEWS now calculates Q* based on Ts which is based on the sensible heat flux and MOST. This has the potential to cause large errors in Q*, especially in SUEWS since the sensible heat flux is calculated as the residual of the energy balance, and also because of the poor performance of MOST and all the uncertainties of the roughness length for heat, etc. The effect of this change needs to be shown here, and makes it more important that the model's ability to calculate Q* is dealt with.*

N/A: removed in the revised manuscript as this is an experimental development and not used in this work.

b) *Appendix B: this is probably appropriate as an appendix but needs to be rewritten as it is not at all clear what has been done here. Start with one or two sentences describing the purpose of this analysis. It needs to be made clear that this section concerns observed data (not SUEWS). How do the results obtained compare to rule-of-thumb values? What do the lines in Fig B1 represent? The discussion of fetch in L523-527 is very unclear and needs rewriting. Make the points smaller and axes ticks consistent in Fig B2.*

Remains as Appendix B but with the following updates:

- New analysis using the FLUXNET2015 dataset.

- Text rewritten and figures updated.

- New text (Appendix B, L796–L801)

  Using the derived z_{0m} and z_d, f_0 and f_d parameters can be obtained (Eqn. 9 and 10). These is considerable intra-PFT variability of both f_0 and f_d (Fig. B1). There are also intra-site variations associated with varying H_c. Given the large variability in both f_0 and f_d, the rule-of-thumb approach would incur large bias in estimated aerodynamic and surface resistances and subsequently the modelled Q_E. To reduce such bias, in the evaluation of the other sub-models and parameter determinations in this paper, we use the derive z_{0m} and z_d determined for each vegetation stage and site.
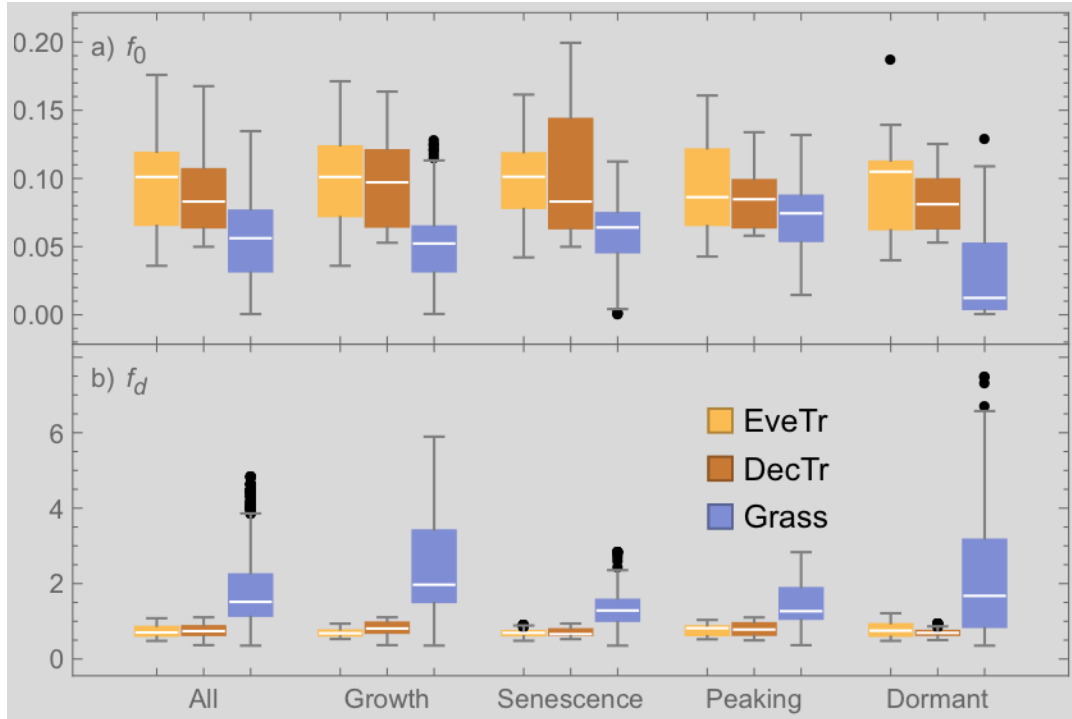
Figure B1. Relations between canopy height (H_c) and a) roughness length for momentum (z_{0m}, Eqn. B2) and b) displacement height (z_d, Eqn. B3) for different vegetation stages based on LAI (see Sect. 4.1 for classification details)

8. *Appendix C should be moved to the appropriate place in the main text. Perhaps adding boxplots for different temperature bins would help support your decision. I would suggest rephrasing as there does not seem to be a point where evaporation 'switches off' and three of the sites have very little data below the suggested cut-off. It should be made clear that this is observed data, not model output.*

9. *Appendix D: what do the authors want the reader to take away from these three tables? L564-565 makes no sense. Suggest deleting.*

10. *Appendix E: as explained before, this comparison does not make sense. The Ward et al. (2016) parameters were derived for bulk urban surfaces, and were not intended to be used for non-urban areas. Suggest deleting.*

11. *Appendix F: What does this plot add to what is already shown (and more) in Fig 10? Suggest deleting.*

New paper structure:

-   Old Appendices C-F removed.
-   New Appendix C added but with different material: now Matsumoto et al.'s (2008) upper-boundary-based method (Sect. 4.3) is adopted to determine the surface conductance related parameters; here we report detailed site-level values derived for SUEWS parameters.

12. *In addition, general readability could be improved by:*

   a) *using fewer cross-references: the reader has to work very hard to follow the text when we are constantly directed to Equation/Table/Figure/Appendix X. Use cross-references where necessary and helpful, but try to ensure the reader knows what variable/site you're talking about.*

   b) *avoiding vague language; instead specify what you mean (particularly with respect to this study versus previous studies and what is generally true/what is true in the model/what is done here).*

c) using more words so that the text flows more naturally and is therefore more easily understandable to the reader. This is particularly true for the table and figure captions, many of which don't really make sense.

d) don't include key methodological information in captions instead of the text (e.g. L381-382).

Done as suggested.

13. A lot of space (Figures and Tables) is given to presenting MAE, MBE and nMAE for the different sites at different times and different states of vegetation. However, there is little insight gained and very little discussion in the text. I therefore suggest removing these figures and, if necessary, compiling the information into a single figure or table

New Sect. 5.2–5.4 (which includes new Fig. 14–24) are undertaken to address this:

- Bias attribution: a new Sect. 5.2 is added in the revised manuscript to analytically attribute the bias in modelled Q_E to different parameter contributors using a sensitivity analysis framework by McCuen (1974), the results of which indicate surface conductance g_s is critical to the model performance in Q_E prediction.

- Impact of g_s parameters: given the importance of g_s suggested by the above analysis (Sect. 5.2), we have compared the model performance by simulations with two sources of g_s parameters – FLUXNET- and NOAH-based values – to examine their impacts (Sect. 5.3), which indicate the better model performance using FLUXNET-based values in particular at finer temporal scales (monthly and hourly) compared to the NOAH-based ones.

- Site-scale performance and key determinants: moreover, we have chosen sites of each PFT with best and poorest performance to understand the causes (Sect. 5.4) and found that correct prediction of LAI timing has a crucial influence on overall performance.

14. The background of this work seems to be the SUEWS model – i.e. an urban land surface scheme that has been developed by 'urbanising' sub-models developed over non-urban environments. The latent heat flux calculation is based on the Penman-Monteith equation with the Jarvis formulation of the surface conductance. This work seems to 'start' from SUEWS and then 'un-urbanise' the equations again by setting the anthropogenic heat flux to zero and fitting parameters for non-urban sites. In many places, the manuscript needs adjusting to reflect that these non-urban forms exist – and in fact existed long before SUEWS!

We did not intend to imply that non-urban parameters did not exist. Text has been changed (Sect. 1, L76–L86)

Central to the SUEWS biophysics, is the Penman-Monteith approach (Penman 1948; Monteith 1965) with a Jarvis-type (Jarvis 1976) surface moisture conductance (Grimmond and Oke 1991). Despite various parameters having been derived to account for different urban areas (e.g. land cover differences) and regions (e.g. high/mid-latitude) to allow for changing phenology, conductance and storage heat flux related parameters (e.g. Järvi et al. 2011, 2014; Ward et al. 2016), urban parameter estimates are lacking partly because of limited observations and lack of a standard workflow for deriving parameters. Other land surface schemes have parameters for a wide range of plant functional types (PFT) (e.g. NOAH within WRF, Chen et al. 1996, Chen and Dudhia 2001) but are often derived from a small number of observational sites and their widespread applicability is unexamined. For example, NOAH largely adopted values from the HAPEX-MOBILHY observational program (Andre et al. 1986) following Noilhan and Planton (1989).

15. Not only is this acknowledgement missing in the model description, but also in the Introduction, Results and Conclusion. In the Introduction, the motivation for this work needs to be set in the wider context of land surface modelling – i.e. at least a paragraph describing previous work that has been done on albedo, LAI and evaporation in forest, grassland, agricultural environments and for water

and bare soil surfaces too. In the Results section, the results obtained here should be compared to previous results obtained in some of these previous non-urban studies. In Table 6, how do the nonSUEWS specific parameters (albedo, LAI, roughness length and displacement height) obtained here compare to the body of literature over the last few decades and why should future users of SUEWS use the values presented here instead of those in the literature?

Section 1 motivation:

- Necessity of using the same modelling framework for urban-rural comparison (L66–L70):

  As SUEWS v2020a (Tang et al. 2021) can diagnose near surface meteorology in the roughness sub-layer and canopy layer (e.g., air temperature and humidity at 2 m agl (above ground level), wind speed at 10 m agl), it is essential to ensure that any urban-rural comparison in these diagnostics has the proper rural skill and parameters (i.e. values used in parameterisations).

- Requirement by WRF-SUEWS coupling (L72–L75):

  With plans to couple SUEWS to a meso-scale model (e.g. Weather Research and Forecasting (WRF), Skamarock and Klemp (2008)), most regions have extensive areas that have completely pervious grid cells. As these need to be simulated using a consistent surface scheme, it is essential to have appropriate parameters for these areas.

- Necessity of examining the widely used surface conductance parameters using more recent observations: Text modification as indicated in previous response to R1C14.

Comparison in model parameters between this and previous studies are now added:

- Albedo (Sect. 4.1, L357–L358):

  see Cescatti et al. 2012 for a detailed analysis of albedo dynamics at FLUXNET sites.

- OHM coefficients (Sect. 4.2, L410–L412):

  In addition to the values derived here, we note that more detailed ΔQ_S observations are available for vegetated sites to derive such OHM coefficients (e.g. McCaughey (1985), Oliphant et al. (2004))

- Surface conductance related parameters (Sect. 4.3, L455–L458):

  The g_{max} results are consistent with Hoshika et al. (2018) in terms of inter-PFT magnitude (Grass > EveTr and DecTr). The grass and crop values are comparable (Table C3) to Hoshika et al. (2018). However, our derived deciduous trees values are smaller (22 cf. 31 mm s^{-1}) and EveTr values larger (20 cf. 12 mm s^{-1}).

16. *Are the roughness length and displacement height values given in Table 6 really useful? Wouldn't it be more reasonable to use the rule of thumb relating these parameters to vegetation height at the site?*

We now use our observation derived roughness length and displacement height by phenology state (dormant, growing, peak and senescence as detailed in Sect. 4.1) – instead of the rule-of-thumb approach – to remove the additional source of uncertainty in the Q_E simulations, because the relational ratios show large intra-PFT variability (even the same site with varying height; see Fig. B1, reproduced above under response to R1C7b) and may lead to considerable bias in modelled aerodynamic resistance if using the rule-of-thumb approach. This has been discussed in Appendix B (L798–L801):

  Given the large variability in both f_0 and f_d, the rule-of-thumb approach would incur large bias in estimated aerodynamic and surface resistances and subsequently the modelled Q_E. To reduce such bias, in the evaluation of the other sub-models and parameter determinations in this paper, we use the derive z_{0m} and z_d determined for each vegetation stage and site.

17. *For SUEWS applications in the urban environment, Järvi et al. (2011) and Ward et al. (2016) derived parameters for the surface conductance using datasets collected in urban areas with the aim of better capturing latent heat fluxes in urban environments. The comparison using the Ward et al. (2016) parameter values therefore does not make sense at these urban sites, as those parameter values were never intended to be used for non-urban surfaces! It would make more sense to compare the current results to those derived over non-urban surfaces, such as Ogink-Hendriks (1995) or Stewart (1988) over forest.*

Discussion added in Sect. 4.3 (L424–428):

> As the Jarvis-type formulation of stomatal/surface conductance is widely used for many land cover types, many parameter sets exist (e.g. Stewart 1988; Grimmond and Oke 1991; Ogink-Hendriks 1995; Wright et al. 1995; Bosveld and Bouten 2001; Järvi et al. 2011). Hoshika et al.'s (2018) comprehensive meta-analysis of published Jarvis-type stomatal conductance parameter values includes major woody and crop plants broadly similar to PFTs examined here.

Sect. 5.3: comparison changed to use the PFT-specific NOAH values (Appendix A) with the FLUXNET-based g_s parameters derived in the paper. Other comparison removed.

18. *A comparison of SUEWS model performance over these non-urban surfaces with previously published results of SUEWS model performance over urban surfaces could also be useful.*

See response to R1C17.

19. *The analysis/interpretation is generally superficial and needs to be developed substantially.*

Additional analysis includes:

- Variability in the derived parameters (Sect. 4).
- Bias attribution of modelled Q_E using an analytical framework (Sect. 5.2).
- Model performance in Q_E prediction at both intra-annual and sub-daily scales (Sect. 5.3 and 5.4).

20. *For example, in L353-355 the timing of the decrease in LAI for wheat is much worse than for rice but this is not discussed.*

N/A - as sites changed to a consistent dataset and crop-specific work removed.

21. *For a more complete paper and, crucially, to avoid drawing misinformed conclusions, the observed data must also be analysed. What is the explanation for the observed variation in the albedo of grassland (are the data even reliable)?*

Changed to FLUXNET2015 dataset to ensure better consistency and QC of all data used. Text added to discuss aspects that have been analysed of different processes (Sect. 1, L90–L96):

> Extensive analysis of FLUXNET datasets for the variety of terrestrial PFTs have considered various surface atmosphere controls (e.g., albedo: Cescatti et al. 2012; latent heat flux: Ershadi et al. 2014; spatiotemporal representativeness: Chu et al. 2017, Villarreal and Vargas 2021; energy balance closure: Franssen et al. 2010; landscape heterogeneity: Göckede et al. 2008, Stoy et al. 2013) to enhance understanding of land-atmosphere interactions. As such, this is an ideal data source for deriving widely applicable parameters and assessing performance of SUEWS over different land covers.

22. *How is the variation in albedo at US-AR1 related to the variation in LAI – more explanation is required, i.e. what is the mechanism proposed behind the low rainfall in 2011 mentioned in L341-343?*

New Appendix D is added to demonstrate the rationale for hydrological control of LAI dynamics.

A different site, US-SRG, with has more pronounced relational pattern between LAI and precipitation is chosen to demonstrate this (L838–L845):

b) Rainfall and thermal controls (US-SRG; Fig. D2): at this grassland site in Arizona the intra-annual precipitation has clear dry and wet seasons. The monsoon wet season after the peak air temperature in July through September (Fig. D2a), which has warmest air temperatures, Unlike US-MMS (Fig. D2b), the peak air temperature is more distinct (for a shorter period). A clear relation between the onset of rainfall and LAI enhancement can be seen but the GDD and SDD relation differs from US-MMS and it not captured by the current models in SUEWS. The rainfall and enhanced LAI and Q_E are associated with cooler daily air temperatures. Sites where the LAI dynamics are not captured are not explored further in this paper.
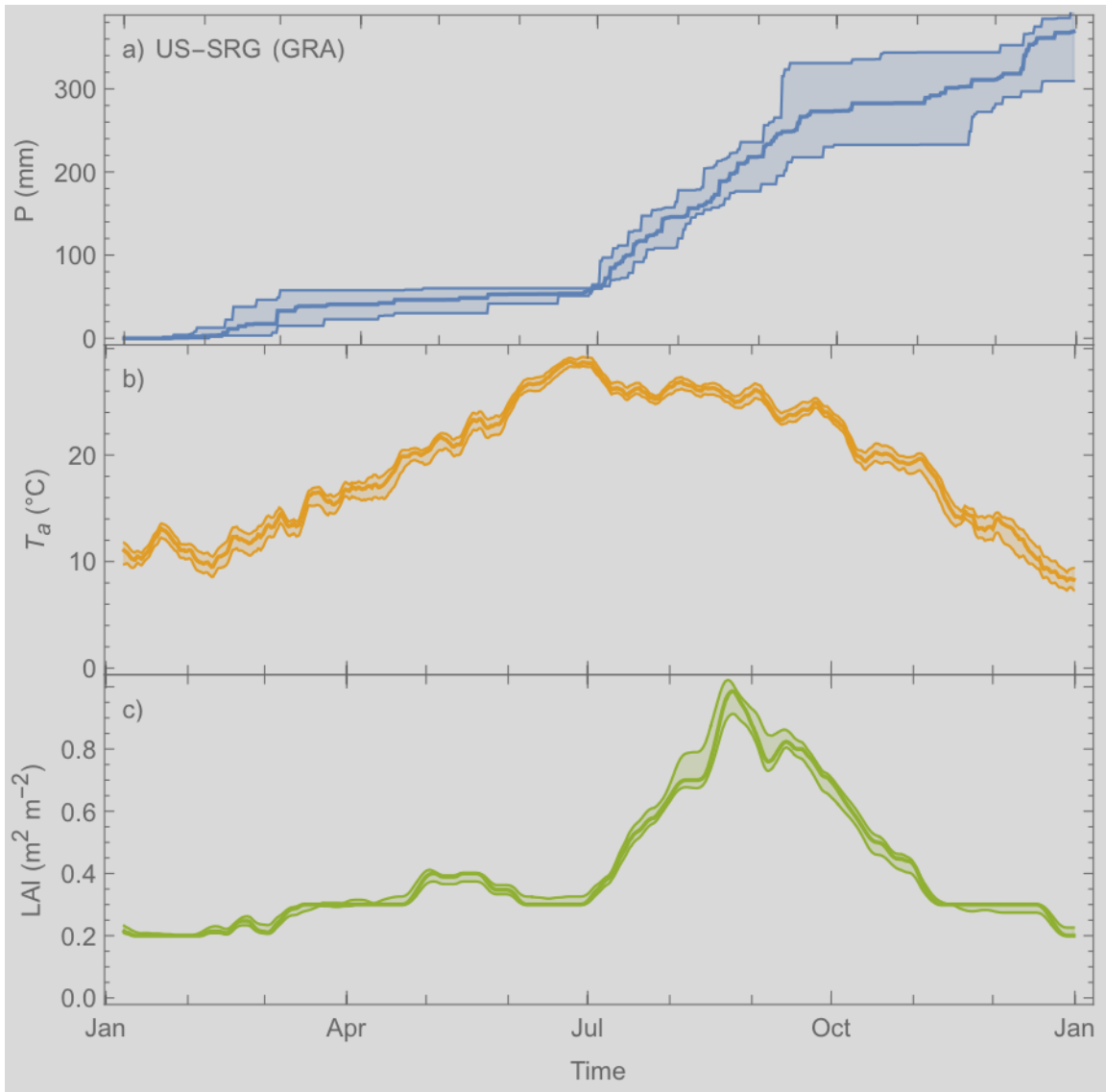

Figure D2 As Fig. D1 but for US-SRG (GRA according to IGBP; time span: 2008–2015; DOI: 10.18140/FLX/1440114).

23. *Why are the results for water and bare soil not shown?*

N/A: work related to these land covers are removed from paper as all data now from FLUXNET2015 dataset

24. *What is the reason for the very different annual cycles for evergreen trees seen in Fig 10?*

N/A: these sites are not included in the revision as the main data source has been changed to FLUXNET2015.

25. Why are the conductance parameters for the individual sites not shown (L438)?

Site-specific values now given in Table C3 and intra-PFT variability in Table 7.

26. Some physical interpretation of the G2-G6 values should be attempted (i.e. why did the fitting procedure result in these values and what does it tell us?).

An upper-boundary-based approach is now used to derive these parameters (Sect. 4.3, L435–L438):

> However, as the optimisation may not return values because of the complexity in Eqn. 14 and the challenge of interpreting the derived parameter values, we adopt Matsumoto et al.'s (2008) approach to derive these parameters. Rather than using all the data combinations for g_s, the upper boundary of each forcing variable component (e.g. $g(K_\downarrow)$) is considered as the response for unconstrained conditions.

Specific comments:

27. The Abstract is quite vague and suggests some topics will be analysed more deeply than they are. Some suggestions:

 a) add 'Here' at the start of the sentence on L34 to make clear this is what you did in this study and is not a general feature of SUEWS

 b) add 'from multiple sites' to the end of this sentence

 c) meaning of 'guidance to apply SUEWS are provided' (L38) is unclear

 d) L39: 'impacts' on what?

 e) 'The relation between LAI and albedo is explored' (L39-40) – I'm not sure this is really covered in the analysis

 f) Add 'in the model' after 'captured' in L44

 g) L44-45: the meaning is unclear and there is no discussion in the manuscript of how latent heat fluxes affects modelled canopy-layer air temperature.

We have incorporated the suggestion and rewritten the abstract as follows:

> To compare urban and rural areas, the fully vegetated areas (e.g. deciduous trees, evergreen trees and grass) commonly found adjacent to cities need to be modelled. Here we provide a general workflow to derive parameters for SUEWS (Surface Urban Energy and Water Balance Scheme), including those associated with vegetation phenology (via leaf area index, LAI), heat storage and surface conductance. As expected, attribution analysis of bias in SUEWS modelled Q_E finds the surface conductance (g_s) plays the dominant role, hence there is need for more estimates of surface conductance parameters. The workflow is applied at 38 FLUXNET sites. The derived parameters vary between sites with the same plant functional type (PFT), demonstrating the challenge of using a single set of parameters for a PFT. SUEWS skill at simulating monthly and hourly latent heat flux (Q_E) is examined using the site-specific derived parameters, with the default NOAH surface conductance parameters (Chen et al. 1996). Overall evaluation for two years has similar metrics for both configurations: median hit rate between 0.6 and 0.7, median mean absolute error less than 25 W m^{-2}, and median mean bias error ~5 W m^{-2}. Performance differences are more evident at monthly and hourly scales, with larger mean bias error (monthly: ~40 W m^{-2}; hourly ~30 W m^{-2}) results using the NOAH-surface conductance parameters, suggesting that they should be used with caution. Assessment of sites with contrasting Q_E performance demonstrates how critical capturing the LAI dynamics is to the SUEWS prediction skills of g_s and Q_E. Generally g_s is poorest in cooler periods (more pronounced at night, when underestimated by ~3 mm s^{-1}). Given the global LAI data availability and the workflow provided in this study, any site to be simulated should benefit.

28. *The treatment of snow cover is rather strange. Judging from Fig 10, either the detection of snow cover is not appropriate or the modelled albedo does not capture the true seasonal variability in vegetation characteristics. As snow cover is not considered in this work, perhaps additional sites with much less snow during leaf-off periods would be more valuable.*

Sect. 2.2 (L168–L170): We clarify that

"our focus is on snow-free conditions"

and indicate:

"evaluating the snow module is a large task in its own right"

Sect. 4.1: We illustrate how the snow-affected albedo values are filtered out (L349–L355):

- α_{min} / α_{max}: 10th/90th percentile of daily albedo values after the growth and before the senescence. A daily albedo is calculated from 30/60 min FLUXNET observations of incoming and outgoing shortwave radiation for the period 10:00 to 14:00 (local standard time). To remove outliers a clustering method is applied (`ClusterClassify` of *Mathematica* v12.3.1 Wolfram Research 2020). For example, at some high-latitude sites (e.g. CA-Oas) snow occurs, the winter values are based on data from shortly after senescence to shortly before growth (next spring) and the clustering approach removes the snow period albedo values.

As such, although we didn't explicit model snow-related physical processes in this work, we deem our treatment can effectively select albedo values under snow-free conditions for deriving the desired albedo-related parameters (i.e. α_{min} and α_{max}).

29. *Why are albedo and LAI parameters derived for each site but surface conductance parameters derived for each land cover type? Analysis of the variation in conductance parameters between sites would be informative and may help to inform about applicability to other sites.*

The parameters are derived and reported for all 38 sites (Table C3). The text has been updated (Sect. 4.3, L452–L458):

The derived surface conductance parameters for the 38 FLUXNET sites (Table 7 and C3) have different intra-PFT variability based on the IQR (dotted lines, Fig. 10) and demonstrates the benefit of the observations and of deriving site-values when possible. It may help in selecting appropriate PFT from other sources (e.g. NOAH values in Appendix A). The g_{max} results are consistent with Hoshika et al. (2018) in terms of inter-PFT ordering (Grass > EveTr and DecTr) and the grass and crop values are comparable (Table C3). However, our derived deciduous trees values are smaller (22 cf. 31 mm s^{-1}) and evergreen trees values larger (20 cf. 12 mm s^{-1}).
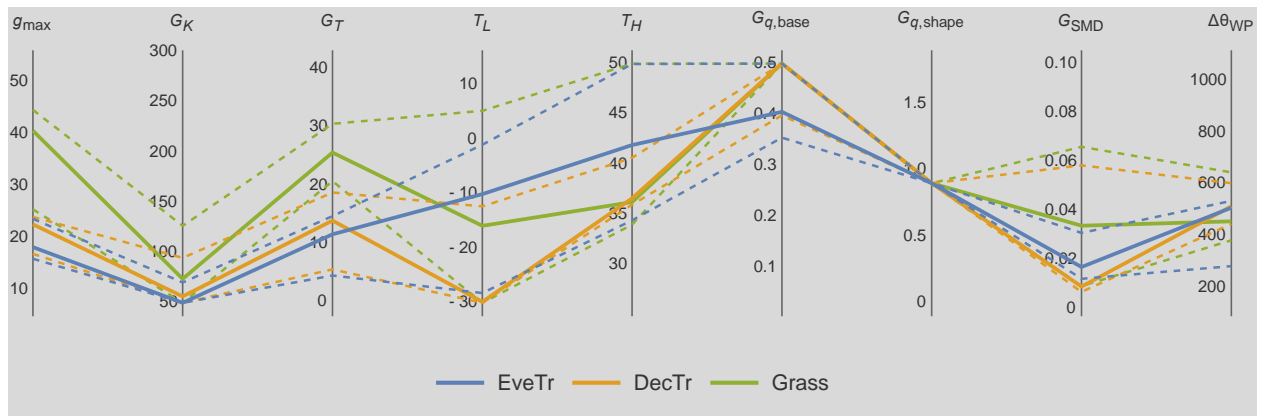

Figure 10 Median (thick), interquartile range (dashed) and site (thin lines) derived surface conductance related parameters for three land cover types (colour).

	g_{max} [mm s^{-1}]	G_K [W m^{-2}]	G_T [°C)	T_L [°C]	T_H [°C]	$G_{q,base}$ [-]	$G_{q,shape}$ [-]	G_θ [-]	$\Delta\theta_{WP}$ [mm]
EveTr	20.5±1.7	62±5	10.3±1.8	-13±4	41.4±2.0	0.391±0.033	0.9	0.033±0.009	511±75
DecTr	21.2±2.5	100±23	18.0±4.0	-18±5	38.0±1.5	0.439±0.024	0.9	0.029±0.010	521±58
Grass	38.6±2.8	87±13	26.1±1.9	-13±5	40.1±2.2	0.467±0.033	0.9	0.048±0.010	521±54

30. *The paper is missing a balanced consideration of shortcomings of this study. For example,*

   a) *Would these non-urban parameters be expected to be appropriate for e.g. deciduous trees in residential areas (L66) given the possibility of increased urban temperatures or advective effects?*

The starting premise is that the parameters are for large extensive vegetated areas (see Sect. 3.1 where fetch of the sites are discussed), these are unlikely to be found in many urban areas. However, the workflows provided can be used in urban areas when the required observations are available.

   b) *Considering the importance placed on seasonal variation the impact of assuming constant OHM coefficients should be addressed.*

The seasonally-varying OHM coefficients are given in Table C2 and discussed (Sect. 4.2, L398–L412):

> The derived OHM coefficients (Fig. 6) can be determined by season (Anandakumar 1999; Ward et al., 2016; Sun et al., 2017). We distinguish warm ("summer") and cold ("winter") seasons using months (summer: Northern Hemisphere JJA; Southern Hemisphere: DJF; winter: DJF (JJA), respectively). For simplicity, we omit periods when LAI may be changing rapidly. If the daily mean air temperature is warmer/cooler than the annual mean of daily median temperature, then summer/winter OHM coefficients are used in the simulations.

> The OHM coefficients derived for the 38 FLUXNET sites (Table 6, Fig. 7) vary between land cover types and seasons. For each land cover type, a_1 and a_3 are notably larger in winter than in summer while the seasonal difference in a_2 is relatively small. Thus the overall fraction of heat stored does not vary much but diurnal hysteresis effect is weaker in winter. These results are consistent with previous analytical results (Sun et al., 2017). Within each PFT, there is larger variability in a_2 and a_3 (cf. a_1), notably for evergreen and deciduous trees, suggesting using the most appropriate site values (e.g. medians) may improve predictions of the storage heat flux. In addition to the values derived here, we note that more detailed ΔQ_S observations are available for vegetated sites to derive such OHM coefficients (e.g. McCaughey (1985), Oliphant et al. (2004)).

   c) *By not including snow cover much of the leaf-off periods are not useable, and the significance of accurately capturing seasonal variation in LAI is reduced. If snow cover is not addressed here, these sites seem like a strange choice.*

See response to R1C28.

31. *Previous shortcomings of LSMs have highlighted that parameters derived for particular conditions can lead to bias in model results at other sites. Although this study aims to provide new, generalised parameter values the range of sites used for each land cover type is not very large and does not cover a large geographical area. The limitations of this should at least be mentioned.*

We have added the following recommendations in the concluding remarks (Sect. 6, L709–L713 and L744–L746):

- Where observations are available, we recommend determining local parameters, as derived parameters vary within PFT (Appendix C). The tools provided here are designed to facilitate this (Sect. 4).

- Given the global availability of MODIS LAI and reanalysis-based air temperature datasets (e.g., ERA5), it is feasible to derive site by site LAI parameters for SUEWS (Sect. 4.1).
- A potential source of parameters values for PFT beyond those studied here (i.e. values provided Appendix C, Sun et al. 2021) could be NOAH-based parameters (Appendix A) but these should be used with caution, as demonstrated (Section 5).

See also response to R1C14.

32. *Keep in mind the appropriateness of the statistics used when comparing different sites and different seasons, where the variables may have different magnitudes and different amounts of data available. Depending on the analyses that end up in the revised paper, this may not be an issue, but it is important to consider. Was any consideration given to other statistical measures (e.g. the correlation coefficient would indicate how well the model reproduces the variability, even if the magnitude is wrong)?*

We have added the non-dimensional metric, hit rate (HR) to examine the frequency the model performance is within an acceptable threshold (Sect. 5.1, L507–L513):

1) hit rate (HR):

$$\mathrm{HR} = \frac{\sum_{j=1}^{N} H(\delta_{Y,j} - |Y_{\mathrm{mod},j} - Y_{\mathrm{obs},j}|)}{N} \tag{31}$$

with Heaviside step function H defined by

$$H(x) = \begin{cases} 0, & x < 0 \\ 1, & x \geq 0 \end{cases} \tag{32}$$

and the threshold $\delta_{Y,j}$ being a value dependent on evaluation variable Y.

In particular, $\delta_{Y,j}$ for Q_E is determined as a function of net all-wave radiation $Q*$ following Hollinger and Richardson (2005) to be $\delta_{Y,j} = 0.1Q_j^* + 10$ (in W m^{-2}) based on measurement uncertainties.

33. *No uncertainty estimates are provided making it very difficult to judge the robustness and accuracy of these results. This should be rectified in the revised manuscript.*
- See response to R1C32 on hit rate.
- Also added uncertainty estimates in standard deviations into Tables 5–7 following the suggestions.

Minor comments

34. *L56-58: Why only mention anthropogenic heat and water here? Urban LSMs have many more additions than this.*

N/A (original text has been removed).

35. *L62: make it clear that these land covers can occur in a single grid. Somewhere the intended grid size for SUEWS needs to be mentioned.*

Rephrased as follows (Sect. 1, L55–L57):

SUEWS characterises the heterogeneity of urban surfaces allowing an integrated mix of seven land covers within a grid cell (neighbourhood scale: $O(0.1–10$ km)) of impervious (buildings, paved) and pervious (evergreen trees/shrubs, deciduous trees/shrubs, grass, soil, water) types.

36. *L64, 67: what is meant by 'integrated'?*

N/A: original text removed.

37. *L70: Here would be a good place to bring in some of the non-urban literature.*

Non-urban related references on LAI, heat storage and surface conductance have been added Sect. 4.1, 4.2 and 4.3, respectively.

38. *L71 'bridge this gap' – you have not really talked about a gap so this doesn't really make sense.*

N/A: original text removed.

39. *L76-79: This overview does not make sense (perhaps reflecting the confused structure)!*

Rephrased based on the new structure (Sect. 1, L99–L105):

> We briefly review the key vegetation biophysics schemes in SUEWS (Sect. 2), describe the FLUXNET2015 (Pastorello et al. 2020) and auxiliary datasets used (Sect. 3), and outline the workflows for deriving parameters (Sect. 4). To assess the quality of the derived parameters the SUEWS modelled latent heat flux is evaluated (Sect. 5). Model parameters related to surface conductance are derived for NOAH at the PFT level (Appendix A) as well as those related to surface roughness based on FLUXNET2015 dataset at the site level (Appendix B). Other model parameters derived following workflows (Sect. 4) are also provided (Appendix C).

40. *L93: 'phenology changes key model parameters' – changes what to what?*

N/A: original text removed

41. *L93-102: these paragraphs do not really make sense. Suggest rephrasing and incorporating at the appropriate stage in the model description section*

Restructured the model description part (Sect. 2.2) as suggested.

42. *Table 1: I personally find the last 3 columns unhelpful. There are details missing from the Definition column (GDD abbreviation, 'coefficients' alone is not informative, 'shortwave radiation', OHM) but possibly this table could be deleted as everything should be defined and explained in the text anyway.*

We respectfully keep the last three columns of Table 1 as we consider they include essential information. We have added the definitions in the caption.

43. *L163: what is t?*

t is time; this definition has been added in Sect. 2.2.2 (L180):

> … and t time.

44. *L165-169: This methodology should be separated from the model description and more details should be added, including what timestep and what type of regression. What is the justification for ignoring variation in these OHM coefficients, particularly for this paper on seasonal variation – was any observed? At least a couple of sentences with references should be added here.*

Separated the description of model physics (Sect. 2) and parameter derivation (Sect. 3).

See response to R1C30b for variations in OHM coefficients.

45. *L181: measurement height for wind speed appears to be Hu in L271*

Measurement height now z_m throughout the paper.

46. *L187: 'stability scale' → 'stability parameter'*

Corrected as suggested.

47. *L201: 'Phenological state is critical' – what is the justification for this statement? Presumably all functions are important if they are not correctly parameterised.*

N/A: original text removed.

48. *L120: It may be easier to follow if this new LAI equation was presented along with Fig 5 so the reader understands why a different parameterisation is needed. Some justification for this equation would be helpful.*

N/A: original text removed; see also response to R1C20.

49. *L135: In reality or in the model?*

N/A: original text removed.

50. *L141-142: Meaning unclear. Which 'model parameters'? How does this paragraph fit with L148-154?*

N/A: original text removed.

51. *L143: And also the longwave radiation components*

Added as suggested (Sect. 2.2.1, L171–L172):

> Within SUEWS the albedo is used with the observed incoming shortwave radiation and longwave radiation to obtain Q^*.

52. *Table 2: Without discussion in the text this table is not much use. The caption does not make sense. Why were OHM coefficients for some surfaces derived here and others from the literature?*

OHM coefficients are now given for each site (Table C2) and their features analysed in Sect. 4.2.

53. *L218: How was it ensured that the surface was dry?*

Only days with zero precipitation are used in this work as clarified in L433–L434:

> This requires the surface be dry (Section 2.2.4) which we define as being without recorded rainfall in 24 h.

54. *L253-263: Table 3 is vaguely referenced 4 times here. Why not provide the appropriate information in the text if necessary?*

N/A: original text removed.

55. *Table 3: Why are the DOIs given here? They should be included in the Reference list.*

As the websites linked by DOIs store all related information of these FLUXNET sites and related datasets, we provide them in a summary table to ease the access to these resources.

56. *Table 3: How were the study years chosen and then how was it decided which to use for calibration and which for testing? Without information, a cynical reader may suspect the combination was selected which gives the best results…*

We have added both data available and SUEWS simulation periods in Table 3. Also, the rationale for choice of these periods is clarified in the revised manuscript:

- Site selection (Sect. 3.1, L258–L259):

> 2)   data availability (56/206): require both MODIS LAI data (available from 2002, Sect. 3.2) and long-term continuity (defined here as ≥ 3 years for the multiple needs).

- Model configuration (Sect. 5.1, L496–L499):

> Simulations are conducted, with forcing data interpolated to a 5 min timestep (Ward et al. 2016), for three years (Table 3, Evaluation period) starting in mid winter. The first year is discarded to allow for model spin-up. The two subsequent years are evaluated when observed latent heat flux are available.

57. *Figure 2: This figure does not seem very useful. Suggest deleting unless it is better described in the text and provides more insight.*

New Fig. 1 gives overview with details presented in Figs. 3, 6 and 8.

58. *L334-334: Units of LAI missing*

Units added throughout revised manuscript.

59. *L356-359: Meaning unclear.*

N/A: original text removed.

60. *Figure 7, 12, and 14 take up a lot of space, are not very easy to read and are hardly discussed in the text. Perhaps a more useful way to summarise this information could be found, or only the most relevant results displayed. The same goes for Figure 9d and Fig 11d. Note it is frustrating for the reader to have to adjust to essentially reading the same information as Fig 7 and 12 but presented in a different way.*

N/A: original figures removed

61. *Panel labels are missing from Fig 9 and 11.*

N/A: original figures removed

62. *Would it make sense to merge Fig 6 and 9a-c and Fig 10 and 11a-c so that these similar results are presented in a more comparable way? I would suggest perhaps even making the x-axis all one year long so the difference for the crops can be seen more immediately. Please use smaller points.*

N/A: original figures removed

63. *Table 4: Tplant, GDDv and GDDLAImax would be better as separate columns.*

N/A: rice related work not in the revision

64. *Fig 12 and L403: the MAE and MBE seem small for the grassland sites considering the results shown in Fig 10. Please check.*

N/A: original figures removed

65. *L427-429: Where is the justification for this statement? Was the performance of other fluxes checked (e.g. net radiation, storage heat flux)? If these fluxes are not modelled adequately, wouldn't they result in inappropriate conductance parameters being derived (e.g. parameters which are tuned to give the right results for the wrong reasons)?*

N/A: work related to WAT and BSV is not used in the revision. See also Sect. 3.1 for site selection in this revision.

66. *L430: More explanation needed here. Also Q* needs to be reasonably accurate. The assumptions in the resistance approach and the uncertainties in the roughness parameters should be discussed too. The assumption of homogeneous fetch requires more explanation if it is included here.*

A more detailed analysis of bias attribution has been added in the revised manuscript (Sect. 5.2).

67. *L433-436: See above for explanation of why this comparison does not make sense.*

N/A: removed

68. *Fig 13: Difficult to read (make full-page?). Use smaller points. The annual diurnal pattern is of limited use considering the huge seasonal variation which makes the large interquartile ranges. Consider using daily or (if data availability is an issue) monthly evaporation totals instead, which may allow insight into when the latent heat flux is modelled well and when not.*

N/A: removed

69. *L456-457: This cannot be concluded here as it is not demonstrated in the paper. This hypothesis is suggested as 'a possible explanation' in L341-344 but no other possibilities are discussed and there is no further analysis to substantiate or contradict this suggestion.*

This has been clarified in Sect. 6 (L738–L741):

> None of the simple LAI schemes in SUEWS account for hydrological impacts on LAI. Vegetation with shallow roots (e.g. US-SRG in Arizona, US, categorised as grassland, Fig. D2) are not well modelled when air temperature if the only phenology forcing variable. Hydrological feedback should be considered in future development of the LAI scheme in SUEWS.

70. *L459: It should be stated somewhere that (presumably) to obtain fitted parameters for a specific site observations must be available.*

See response to R1C31.

71. *L558: 'Given this' – given what?*

N/A: text removed

72. *There are also a few typos that would need to be corrected at a later stage.*

Corrected throughout the paper.

Reviewer 2

General Comments:

1. *The manuscript aims to extent the SUEWS model to non-urban surfaces, with the overall goal to estimate the energy-balance fluxes in such areas. Therefore, specific parameters used to estimate the surface heat fluxes are inferred from observational data sampled at energy balance stations, which includes different vegetation types and different climate zones. The modelled surface fluxes with SUEWS were compared against observational data to evaluate the performance of the SUEWS model over rural surfaces. The topic of the paper itself fits well into the journal and is of interest to the research community, especially since reliable input of phenological data and surface data play a key role for reliable estimation of the surface-energy balance components in models. However, after extensive review, I cannot recommend the manuscript for publication until major revisions have been done and extended analysis is presented. My major concerns are outlined in the following.*

Major comments

2. *You added new methods and tuned parameters to model impervious surfaces in rural areas. However, the description of newly developed parameter estimation such as for LAI or albedo is mixed with parts of model description, so that it is hardly possible to extract what is new and what has been there already before. I would recommend to first described the state-of-the-art model and describe newly developed approaches separately. Also, the manuscript provides no condense model description of SUEWS but refers to previous papers. The manuscript itself should be readable as a standalone paper. Hence, even though not all details need to be brought-up, the manuscript needs to provide a proper overview of the model at one place. Further, please give all information concerning model description in the text, not within the appendix.*

See response to R1C5.

3. *The manuscript is sometimes hard to follow due to missing logical order between sentences. In several sections, sentences appear to be disconnected from each other rather than indicating a logical order. As a consequence the text reads more like a collection of notes.*

See responses to R1C5–12 for notable structural changes in this revision.

4. *The discussion of the results is not sufficient and lacks important aspects. For example, why is the bias error positive for some sites but negative for others. The authors provide the errors for all sites, but do not try to put these within the context of site-specific information. Also, one of the main problems of eddy-covariance measurements is the non-closure of the energy balance. Especially for the comparison of surface latent heat fluxes this needs to be discussed. In this context, the manuscript needs to provide also more information about the specific EC sites. At EC stations located in heterogeneous landscapes the measured fluxes are a mixture of signals emerging at different land-surface types rather than only one type (as assumed in this study), i.e. the footprint of the stations covers several land surfaces with different properties (LAI, roughness). As a consequence the value f_i (which is assumed to be 1 in this study) is not necessarily one. To be able to evaluate the validity of the inferred parameters in this study, site specific information should be provided, e.g. the degree of surface heterogeneity, which in turn need to be correlated to the overall error in the surface latent heat flux for the individual sites.*

- See response to R1C13 for our improved analysis of bias error.

- Site representativeness:
  First, we need to clarify a detailed observational analysis of flux measurements is out of scope of this work. Meanwhile, we fully agree with the reviewer that a better understanding of the measurement contexts may help interpret the results presented here. Given FLUXNET2015 has been extensively analysed in many studies, instead of repeating similar analysis with respect to surface heterogeneity, we provide related references and discussions as follows (Sect.3.1, L272–L277):

The landscape heterogeneity of many FLUXNET EC flux measurements sites have been systematically examined by Stoy et al. (2013) using satellite imagery. Of the sites they examined, they found them to be located within homogeneous parts of the targeted PFT, but the larger landscape (~20 km) may have considerable variability. As a FLUXNET site is typically assigned to one PFT for land surface model development/evaluation (e.g. Stöckli et al. 2008, Zhang et al. 2017), we configure each as a homogeneous grid cell and assume f_i =1.

Minor Comments

5. *L55-56: You mention that there is a number of LSM's, but you cite only one.*

N/A: original text has been removed.

6. *L54-58: In my opinion this leads the reader on a wrong track, the manuscript focuses on nonurban rural sites.*

We have restructured the introduction part to make the storytelling more rapidly reaching the nonurban topic in the first paragraph of revised manuscript (Sect. 1, L58–L61):

Although SUEWS has been evaluated in cities around the globe (e.g. Karsisto et al., 2016, Ward et al., 2016, Ao et al., 2018, Kokkonen et al., 2018, Harshan et al., 2018) with varying mixes of integrated impervious-pervious land covers, its performance has not been comprehensively examined in fully vegetated areas that commonly occur adjacent to cities.

7. *L63: I guess you mean "around the globe".*

Modified as suggested.

8. *L66: The word parameters is unclear at that point and need to be specified. Do you mean certain (bio)physical quantities such as leaf-are densities, surface or material properties, or do you mean certain values used in parametrizations?*

Clarified in Sect. 1 (L68–L70):

it is essential to ensure that any urban-rural comparison in these diagnostics has the proper rural skill and parameters (i.e. values used in parameterisations).

9. *L71: Which gap does the authors mean? Please be more specific.*

NA (original text has been removed).

10. *L96: It is unclear to what does "The former" refer to.*

Rephased as follows for clarification (Sect. 4.1, L324–L328):

LAI changes also modify both aerodynamic roughness parameters (roughness length z_0, zero plane displacement height z_d) (e.g. Kent et al. 2017) impacting aerodynamic resistance (r_a) and surface resistance (r_s). LAI directly moderates Q_E and canopy interception capacity, which modifies when potential evaporation occurs and aspects of the water balance.

11. *L98-99: "Model parameters ...": As a stand alone sentence this makes sense, though it becomes not directly apparent to the reader what is exactly meant. However, from this there is not obvious connection to the following sentence. With changes of the key parameters you may describe any type of vegetation, but how is this related to the statement that parameters need to be consistent?*

NA (original text has been removed).

12. *107-108: How are GDD and SDD defined? Are these vegetation-type specific?*

We have added the definitions of GDD and SDD in Sect. 2.2.1 in the revised manuscript, which are vegetation-type specific and given along with related symbols as follows (Sect. 2.2.1, L152–L157):

In SUEWS, leaf growth is tiggered by reaching a critical growing degree days (GDD) threshold ($T_{base,GDD,i}$), and similarly for leaf fall by senescence degree days (SDD, $T_{base,SDD,i}$) using daily (d) mean air temperatures (T_d) based on the previous day (d - 1) for each vegetation type i (one of evergreen trees, deciduous trees and grass). For forests and grass we use (Järvi et al., 2011):

$$LAI_{d,i} = \begin{cases} min(LAI_{max,i}, LAI_{d-1,i}^{\omega_{1,GDD,i}} GDD_{d,i}\, \omega_{2,GDD,i} + LAI_{d-1,i}), & T_{base,SDD,i} < T_d < T_{base,GDD,i} \\ max(LAI_{min,i}, LAI_{d-1,i}\, \omega_{1,SDD,i}(1-SDD_{d,i})\, \omega_{2,SDD,i} + LAI_{d-1,i}), & T_{BaseGDD,i} < T_d < T_{base,SDD,i} \end{cases} \quad (1)$$

with $GDD_{d,i} = GDD_{d-1,i} + (T_d - T_{base,GDD,i})$, $SDD_{d,i} = SDD_{d-1,i} + (T_d - T_{base,SDD,i})$, and $\omega_{1/2,GDD/SDD,i}$ curve factors needing to be derived.

13. *Eq. 3: Does the index i includes all vegetation types including or excluding crops?*

Work related to the crop-specific LAI model has been removed in this revision.

Also the meaning of *i* has been clarified (Sect. 2.2.1, L154–L155):

> … for each vegetation type *i* (one of evergreen trees, deciduous trees and grass).

14. *Eq. 3: Is LAI_max/min a function of the time of the year? If this is the case, please indicate this somehow within the equation or text.*

$LAI_{max/min}$ is *not* a function of time of year but an adjustable parameter.

In the revised manuscript, we have clarified its meaning as follows (Sect. 2.2.1, L160–L161):

> For each site and vegetation type *i*, the maximum and minimum *LAI* values ($LAI_{max,i}$, $LAI_{min,i}$) and $T_{base,GDD}$ and $T_{base,SDD}$ are determined for each site (Sect. 4.1).

15. *115-116: Where does these max/min values come from? Here, a reference is required in the text.*

Determination of these LAI related parameters have now been detailed in Sect. 4.1.

16. *121: The note within the parenthesis is unclear to me, how are shorter / longer LAI_max times are reflected in Eq. 3?*

N/A: original text removed; see also response to R1C20.

17. *175/189: I guess you mean water vapor.*

Yes; this has been clarified throughout the revised manuscript.

18. *199/200: The authors should elaborate why removing G1 from the first term is a valid approach. According to the text it sounds to be an arbitrary decision, though I assume there is a specific reason for this?*

The G1 was introduced in SUEWS as an adjusting parameter for grid with a mixture of different vegetated land covers (Jarvi et al. 2011, Ward et al. 2016) that rescales the contributions to total surface conductance from different vegetated land covers with respect to their LAI values.

In this work, given the focus on homogeneous land covers, we removed the adjusting parameter G1 for formulation simplicity (we also note mathematically G1 and g_{max} are interchangeable in the formulation for a fully homogeneous land cover as in this work).

19. *200-201 and following: This is not really a sentence but more a note. Also, the following sentences sound more like a note.*

Reworded.

20. *204: To be specific, soil moisture deficit is not really a meteorological quantity.*

Corrected.

21. *Eq. 13,14,15: G_2, G_3, G_4, G_5 are not defined in the text.*

Defined now in the revised manuscript in Sect 2.2.4; please also note we modified the notation with more explicit names for their physical meanings:

G2 → G_K: solar radiation (K) related parameter.

G3 → $G_{q,base}$: specific humidity (q) related parameter for the "base" value.

G4 → $G_{q,shape}$: specific humidity (q) related parameter for the curve shape.

G5 → G_θ: soil moisture (θ) related parameter depending on soil type.

22. *L240: Parameters itself cannot have a performance. What you mean is the performance of SUEWS using parameters for non-urban surfaces.*

We meant the model performance when configured with a specific set of parameters. Related text has been clarified throughout the revised manuscript.

23. *L246: What do the authors mean with surface state?*

"surface state (C_i)" refers to the water content on canopy, which has been clarified as follows (Sect. 2.2.4, L227–L228):

> The amount of water on the canopy of each surface (C_i)

24. *L285: What do the authors mean with "are not completely independent": among each other?*

N/A: original text removed.

25. *L347-349: In Fig. 6 the authors show the LAI distribution over the year. It does not become clear how this indicates that a constant LAI would lead to poor radiation and surface fluxes. If the authors see a link between these two things it should be given there.*

N/A: original text removed.

26. *Fig. 6: The LAI variation for the evergreen-tree sites is surprisingly high. The minimum LAI values for the respective Canadian sites are similar compared to the deciduous-tree sites. For evergreen trees I would expect a rather time-constant value, while here also the MODIS values indicate almost zero LAI. Could the maybe connected to snow cover on trees?*

We thank the reviewer for bringing up this concern, which led us to a more thorough investigation of intra-annual LAI dynamics of various land covers, including evergreen trees.

By looking into intra-annual LAI dynamics of evergreen trees using a long-term (1981–2015) MODIS LAI climatology dataset (Mao and Yan 2019), we find (Fig. R1):

- evergreen green broadleaf forest (EBF) keeps quasi-constant LAI values throughout a year (Fig. R1a).
- evergreen needleleaf forests (ENF) does show apparent intra-annual variability (Fig. R1b), which is consistent with Fig. 6 in our last submission.
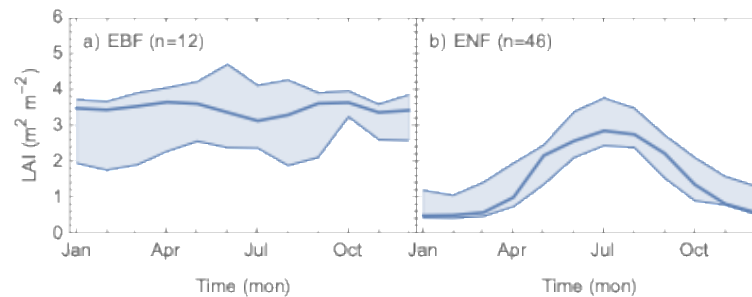

Figure R1 Ensemble intra-annual LAI dynamics of a) evergreen green broadleaf forest (EBF) and b) evergreen green needleleaf forest (ENF). Medians are in bold lines while shadings for inter-quartile ranges. "n" denotes number of FLUXNET sites used in plots.

As for the low LAI values found at ENF sites in winter, it is a known issue in MODIS LAI product that seasonal variation can be exaggerated by unrealistically low LAI retrievals over high latitude ENF in winter (Garrigues et al. 2008, Heiskanen et al. 2012). Related discussions have been added in Sect. 4.1 (L365–L368):

> For EveTr sites, the large contrast between LAI_{max} and LAI_{min} in the ENF sites analysed here is consistent with MODIS derived LAI for ENF having larger seasonal variability than EBF (Heiskanen et al. 2012), but some of this is caused by a known issue of particularly low winter values (Garrigues et al. 2008).

27. Fig 7b: MBE indicates that the modelled LAI values are biased towards smaller values (not for all sites, but for many), especially during the leaf-on period. However, I miss some discussion about this in the text (line 338-345).

Please see response to R1C13.

28. Fig 12: Please provide a full description what is shown in the figure. To switch between the figures to find out what is shown makes the figure hardly readable.

N/A: original figure removed.

29. Most of the equations: Punctuation is missing.

Punctuation is added wherever appropriate.

References

Cescatti, A., Marcolla, B., Vannan, S. K. S., Pan, J. Y., Román, M. O., Yang, X., Ciais, P., Cook, R. B., Law, B. E., Matteucci, G., Migliavacca, M., Moors, E., Richardson, A. D., Seufert, G. and Schaaf, C. B.: Intercomparison of MODIS albedo retrievals and in situ measurements across the global FLUXNET network, Remote Sens Environ, 121, 323–334, doi:10.1016/j.rse.2012.02.019, 2012.

Chu, H., Baldocchi, D. D., John, R., Wolf, S. and Reichstein, M.: Fluxes all of the time? A primer on the temporal representativeness of FLUXNET, J Geophys Res Biogeosciences, 122(2), 289 307, doi:10.1002/2016jg003576, 2017.

Ershadi, A., McCabe, M. F., Evans, J. P., Chaney, N. W. and Wood, E. F.: Multi-site evaluation of terrestrial evaporation models using FLUXNET data, Agr Forest Meteorol, 187, 46–61, doi:10.1016/j.agrformet.2013.11.008, 2014.

Franssen, H. J. H., Stöckli, R., Lehner, I., Rotenberg, E. and Seneviratne, S. I.: Energy balance closure of eddy-covariance data: A multisite analysis for European FLUXNET stations, Agr Forest Meteorol, 150(12), 1553–1567, doi:10.1016/j.agrformet.2010.08.005, 2010.

Heiskanen, J., Rautiainen, M., Stenberg, P., Mõttus, M., Vesanto, V.-H., Korhonen, L. and Majasalmi, T.: Seasonal variation in MODIS LAI for a boreal forest area in Finland, Remote Sens Environ, 126, 104–115, doi:10.1016/j.rse.2012.08.001, 2012.

Mao, J., and B. Yan: Global Monthly Mean Leaf Area Index Climatology, 1981-2015. ORNL DAAC, Oak Ridge, Tennessee, USA, doi:10.3334/ORNLDAAC/1653, 2019.

Stöckli, R., Lawrence, D. M., Niu, G.-Y., Oleson, K. W., Thornton, P. E., Yang, Z.-L., Bonan, G. B., Denning, A. S., and Running, S. W.: Use of FLUXNET in the Community Land Model development, 113, doi:10.1029/2007JG000562, 2008.

Villarreal, S. and Vargas, R.: Representativeness of FLUXNET Sites Across Latin America, J Geophys Res Biogeosciences, 126(3), doi:10.1029/2020jg006090, 2021.

Wang, Y., Broxton, P., Fang, Y., Behrangi, A., Barlage, M., Zeng, X., and Niu, G.: A Wet-Bulb Temperature-Based Rain-Snow Partitioning Scheme Improves Snowpack Prediction Over the Drier Western United States, Geophys Res Lett, 46, 13825–13835, https://doi.org/10.1029/2019gl085722, 2019.

Garrigues, S., Lacaze, R., Baret, F., Morisette, J. T., Weiss, M., Nickeson, J. E., Fernandes, R., Plummer, S., Shabanov, N. V., Myneni, R. B., Knyazikhin, Y. and Yang, W.: Validation and intercomparison of global Leaf Area Index products derived from remote sensing data, J Geophys Res Biogeosciences 2005 2012, 113(G2), doi:10.1029/2007jg000635, 2008.

Zhang, X., Dai, Y., Cui, H., Dickinson, R. E., Zhu, S., Wei, N., Yan, B., Yuan, H., Shangguan, W., Wang, L. and Fu, W.: Evaluating common land model energy fluxes using FLUXNET data, Adv Atmos Sci, 34(9), 1035–1046, doi:10.1007/s00376-017-6251-y, 2017.