This manuscript describes an interesting modelling study of soil water balance components in lysimeters with grassland production, including the simulation of cutting/grazing, performed in two contrasting climate zones in Germany. Six lysimeters were collected in a wetter region (Rollesbroich), three of them were transported to a drier zone (Selhausen) to allow studying the same soil under different climate conditions. This practice, here called “space-for-time substitution”, could be helpful to mimic future climate conditions.

For the modelling, a Richards equation-based hydrological model is developed, using Van Genuchten-Mualem soil hydraulic parameters. Special attention is given to root growth and distribution, and a process-based function is used to predict root water uptake as a function of depth, and actual transpiration rates. Potential grassland growth (dry matter accumulation) is modelled using the concept of RUE. Actual growth rates are calculated from potential rates by reducing for water and temperature stresses. The effect of water stress on dry matter accumulation is included by assuming a linear proportionality between relative transpiration (predicted by process-based root water uptake modelling) and relative growth. Temperature stress is added by establishing a piecewise linear function with zones delimited by specific threshold temperatures. Simplifications, especially with respect to root growth and distribution, are unavoidable in this kind of modelling approach and are well presented and justified.

The manuscript is generally well written. There are some issues to be addressed by the authors, among others referring to the soil hydraulic parameterization and clarity about the lysimeter soil contents. See my specific comments below.

Besides these, in my opinion, the most important shortcoming (making me suggest a major revision) of the manuscript refers to the calibration and validation procedure. You used the GLUE method to identify the best parameterizations using 6 years of lysimeter observations. You then discuss the model performance based on the 30 best parameter combinations selected by some criteria out of 2000 original combinations (parameter realizations). Posterior results seem fairly good, in terms of soil water content and ET (Figs 4 and 5), as well as model efficiencies (Table 6). But isn’t that to be expected when selecting the 30 best performing parameter sets? I would challenge you to follow a more rigorous calibration-validation protocol, performing the GLUE method on three or four years of your data, selecting the best parameter combinations, and then testing them on
the remaining two or three years. This would reveal an unbiased and much more convincing model performance.

The conclusion section contains some discussion but does not mention the interesting fact that plants were able to mitigate water stress in the drier climate by enhancing below-ground (root) growth. If this is confirmed, soil organic matter contents could be expected to rise when climate change triggers more water stress. Also, in agricultural (grain) crops, aboveground yields might then be expected to diminish to favour belowground development. This is important in several ways, one of them being the expected yield of agricultural crops.

Scientific questions/issues:

118-126 This final part of the introduction appears to refer to Materials and methods. I suggest finalizing the introduction with a concluding statement (possibly containing the objectives), and start the M&M section with the content of these lines. Related to this, please state more clearly that the six lysimeters contain soil from Rollesbroich, but that three of them were transferred to Selhausen. Somehow contains this information, but it should be stated more clearly. The sentence in 179-181 is clear with this respect but should be located at an earlier position in the text.

194-196 This net upward flow would imply a depletion of the groundwater body in the long term? I.e., if plants invest more in the belowground biomass to avoid drought stress, large-scale hydrological cycles would be affected? A comment in this respect seems appropriate here.

210-213 Here you suggest that yield may have been affected by nutrient deficiencies. The performed simulations assume optimum plant nutrition (1356). Shortly discuss if this would affect the fairness of assessing model performance based on experimental data including nutrient issues.

284-294 No details are given about how you (numerically) solved the Richards equation (13). Was an implicit algorithm scheme used? How about the discretization? Time and space steps? Averaging K and h? Some more details would be important.

429 Regarding “the fraction of the total root length that is effective for water uptake” (epsilon), you use a fixed value of 0.05 (after Faria et al., 2010). A high model sensitivity to this parameter is to be expected, and it is also plausible to assume that epsilon will increase in periods of more intense root growth and decrease when the root system is shrinking due to stress. Please discuss shortly the effect of assuming a fixed static value for this parameter.

481-495 There are a lot of assumptions, guesses, and uncertainties in the soil hydraulic parameterization described here. It is especially peculiar that the saturated water content of the surface layer is much higher in the lysimeters at Rollesbroich, as “estimated from the data by eye”. Assuming all lysimeters contain a similar soil monolith (collected at Selhausen), can this be made plausible? In 468-469, the reader is informed that a common parameterization for the soil hydraulic properties in all six lysimeters will be assumed. But apparently, this does not include ThetaS? Please improve the description/justification of the soil hydraulic parameterization.

582 “No differences between the two sites were found for two of the parameters” – this is an interesting result and merits more attention. I think it is corroborated by common sense. Both parameters (the radiation extinction coefficient and the parameter controlling
dry matter allocation and leaf loss as a function of water stress) are expected to be mostly genotypical and not affected by environmental conditions. On the other hand, unstressed stomatal conductance (discussed in l591-602) is expected to be a function of sink (atmosphere) conditions, especially VPD, temperature, radiation; rooting depth will probably be affected by soil moisture and temperature distribution in the soil profile, both a function of weather conditions.

l627 Here one of the reasons for the high E values at Rollesbroich is assumed to be “the capillary nature of the soil”. This is a too vague description. Soils are very similar at both locations. I suggest you remove this reason.

l700-705 (“A major ... growth”) These statements are no conclusions from the presented research, they would fit better in the Discussion than in the Conclusion section.

Some other minor "technical corrections":

l62 SVAT stands for "Soil-Vegetation-Atmosphere Transfer". Please add "Transfer".

l79 Johnson et al. (2008) would fit well in this list of citations (doi:10.1071/EA07133)

l121 Delete “In this study”.

l235-240 For clarity, please add the units (dimensions) of the parameters of eqs. 1-7 in these lines.

l270 To be more precise, one would need pressures and conductances, both unavailable in bucket models.

l288 (here and on other occasions) The unit “day” is officially abbreviated as d. I suggest using d instead of day, days throughout the text.

l306 I would prefer to say that tau (just like alpha and n) is a (fitting) shape parameter. The way you wrote it here, it seems tau could be independently measured.

l700 after “below-ground biomass” you might add: “thus mitigating drought stress”.

Figs. 1 and 2, X-axis label and Figs. 6, 7, 11, and 12, Y-axis: replace yr-1 by y-1

Fig. 4 Symbols in this figure can hardly be distinguished, I suggest using colours for the three lysimeters. It would also be good to clearly identify both columns of figures by Selhausen and Rollesbroich (they are now identified only by the codes of the respective lysimeters).

Figs. 4, 5, 8, 9, and 10: X-axis label is unclear. I would prefer “Time from onset of simulation (01-Jan-2013) [d]”

Figure 7: Interpretation would be easier if Tp and Ta (and Ep and Ea) would appear in the same figure, side by side per location. In the current version of the figure, it is difficult to detect any difference between Tp and Ta.