Response to Referee #1

We'd like to thank Referee #1 for the positive and helpful review. The comments of Referee #1 concern valid points that will help to improve out manuscript. Below, we will address the comments of Referee #1, with the referee comments written in italics.

However, I felt that the goals/sequence of the analysis were not explicit enough and the presentation of the methods rather unorganized, that is, it was difficult for me to clearly understand why things were done from the beginning, and identify/distinguish what parameterizations are part of the 2015 model versus the new version. It seems important for the goal of this paper that these aspects be as clear as possible.

We will revise the Methods section, also based on the specific comments, and will describe the model more generally, and the modifications of the model only in Section 2.2.8, to avoid confusion. More specifically, we believe most confusion came from the descriptions about the water balance and the model input data, where both model versions were described. We will generalize these sections and add the specific details to Section 2.2.8. We will also describe specific goals of each modification in the point-wise description in Section 2.2.8.

The results are clear but I was left thinking at the end that something was missing in the conclusions to broaden the relevance of this paper to the community and comment on the utility and value of doing such a systematic analysis of individual model changes; discuss the robustness of VOM overall; the implications of the findings; put the information in a broader context versus replicating the Whitney 2016 boundary conditions.

This is an important point raised by the referee and we will add more context to the conclusions. We will clarify that model development very often suffers from lack of transparency about the effects of various improvements when applied in isolation or in combination. Therefore, our one-step-at-a-time benchmark approach is likely of use to the broader modelling community. In addition, our findings show that the common assumption of freely draining conditions in TBMs can have large effects on the simulated fluxes and therefore the potential for groundwater influence needs to be carefully assessed when interpreting the results.

I have mostly made comments on the presentation, which I hope improve the structure of the paper for an easier read and increase the relevance of the information.

Specific comments
The word 'step-wise' in the title is poorly chosen. I expected the modification to be done one at a time in a sequence building on each other – but this was not exactly the case – they were just done one-by-one.

We agree that the word step-wise leads to confusion and decided to abandon the use of it completely.

The introduction is not specific enough and difficult to follow to take away the important information. What are the shortcomings that are meant to be addressed? What is proposed to address them, why and how? what is the specific outline of boundary conditions that need to be changed? Referencing the companion and it's goals could be helpful to provide more context and understand the relevance of this paper.

This is also an important point raised by the reviewer. We elaborated in the introduction on the accompanying paper and its goals, and have re-written the mentioned paragraphs, currently starting from line 50. In addition, we have added a paragraph explaining the importance of tracing back the effects of model modifications to previous applications of the same model, in order to maintain generality of a model.

In section 2.2, it is helpful to mention in the first paragraph that detailed descriptions are in Schymanski et al. 2009, 2015 (and perhaps even mention the few other earlier ones referenced later about specific processes). But then it doesn't seem necessary to constantly repeat "according to Schymanski XX", "after Schymanski XX"; defined as "Schymanski XX" in the rest of the section. To lighten up the rest of the model description, I suggest clearly stating at the beginning of the section that all parameterizations and processes are the same as in the original references, except those explicitly mentioned. This would better highlight what is different and needs to be remembered and relevant here.

We agree that for readability, it will be better to mention the references once. We will follow the suggestion of the referee, and add one statement in the first part of section 2.2 and remove the redundant references.

It may be even worth completely separating the description of the original model structure versus the parameterizations (and their rational) relevant in this paper, that is, having a completely separate description/list of individual modifications versus having multiple changes mixed into eachother and described together as they seem to be here. A clearer structure may need more thought as these are the key aspect in the paper. Maybe even a table of the information in section 2.2.8 would be an effective summary.

We believe that most confusion came in the section about the water balance (2.2.5) and the model input data (2.2.7), where both set-ups were described. We will generalize these descriptions and move the specific descriptions to the section about the modifications (2.2.8, now 2.2.9). We will follow a similar approach about the model input data, with a generalized description in section 2.2.7, and a description of the changes in section 2.2.8 (now 2.2.9). Hence, we will separate the model description from the modifications, as suggested by the referee.

It would be helpful to name or number the multiple different model variants in a more systematic/tractable way versus referring to the 'new' model and "Schymanski 2015" or "previous application" versus "here" or "current". And be able to more easily reference the text while looking at the figures for the step modifications.

We will refer to the previous model version as VOM-AoB2015, and the new model VOM-v0.5 for clarity, and remove the referring as described by the referee.

It would be helpful if results in 3.1 follow the same sequence as how the cases are presented in the methods and in the figures. Is there a rational for this sequence (can you explain it?) and can it be consistent throughout?

Originally, this sequence was chosen based on the importance of the change, with the changes that had strong effects at the end. However, we feel that this is not obvious anymore, and will change the sequence in the section as the referee suggests.

Line-specific comments

L26 -28: These are a very general statement maybe be more specific about what are "novel modelling approaches" ; "fluxes"; "vegetation dynamics". What are specific shortcoming that are relevant to the VOM?

We will change "vegetation dynamics" to "vegetation dynamics, such as vegetation cover or root surfaces" , "fluxes" to "carbon and water fluxes", and replaced "novel model approaches are needed, especially related to vegetation dynamics" with "models with explicit vegetation dynamics are needed" We will also add a paragraph about the accompanying paper (also based on the specific comments), where we also elaborated on the specific shortcomings in TBMs that are relevant for the VOM.

L28 "therefore, we use here" seems misplaced relative to the broad explanations in the next 2 paragraphs. Maybe just simply state "Optimality theory predicts …"

Will be changed accordingly.

L40-42 this is repeating what was stated in the previous paragraph

We will remove the sentence.

L50-63 What shortcomings versus what modifications? In the next few paragraphs there seems to be a mix of information that should be in the model descriptions / site description. I suggest structuring more in parallel and in a more explicit outline of what shortcomings or what boundary conditions are addressed and what modifications were required and tested here to address them.

We will re-write this part of the introduction, also with respect to the general and specific comments of the referee.

L65 Could be helpful to mention how many steps were taken.

We will add the number of changes here, and in section 2.2.8 (now 2.2.9) the number of model cases (9 changes, with a reproduction of the VOM-Aob2015, re-optimization of the VOM,-AoB2015 and the final VOM-v0.5).

L84 "found in" ?

Will be changed accordingly.

L91: Here and elsewhere use Net Carbon Profit or NCP consistently rather than defining it multiple times

We will remove the multiple definitions of NCP and refer to it only as NCP throughout the manuscript.

Table 1: maybe write out all words like precipitation. potential evaporation; radiation ; delete '.' After aridity ; correct units of net raditation MJ m-2 year-1

Will be changed accordingly.

L103: components – plural?

Will be changed accordingly.

L114: “essentially” is informal writing

We will remove “essentially”.

L142: space between “cost” and “factor”

Will be changed accordingly.

L173: Maybe separate model input data from evaluation data in different sections?

We will separate the paragraph now into “Meteorological data” and “Model evaluation data”