Interactive comment on "Diurnal, seasonal and long-term behaviour of high arctic tundra-heath ecosystem dynamics inferred from model ensembles constrained by time-integrated CO_2 fluxes" *by* Wenxin Zhang et al.

Anonymous Referee #2

Zhang et al. calibrated the Coup model using Zackenberg NEE eddy covariance measurements. In my view this manuscript needs a major revision of text and figures before submitting again.

1) The language seems to be inadequate. It is hard for me to understand all methods and results. One example is on lines 308-310 but the manuscript is full of sentences that I do not understand.

2) Figures show wrong units, miss axes titles and units. Figure captions are not always

understandable. Fig 4: Scaling ME to +/- 10 (units?) makes it hard to understand the figure.

3) It is hard to identify any research question / hypothesis in introduction, discussion and conclusion.

From the start of the discussion it seems that the research question of this study could be: Will tundra ecosystems move from a C sink towards a C source in future? It seems the authors want to calibrate the coup model using eddy covariance measurements from a specific site in order to address this question. Calibrating and running the model into the future using a climate scenario could give a first answer. Instead, the authors present a lot technical details about the calibration procedure which even does not show any technical advancement with respect to model calibration exercises.

4) Methods: A lot of text is written about the coup model which can be found also in the online documentation of the model. In contrast, methods important to the presented manuscript are described superficially, e.g. calibration procedure, wavelet analysis and model ensemble.

What is a behavior model ensemble? What are the three specific behavior model ensembles and how are they defined? This reviewer get a clue about it after reading the whole manuscript but it remains unclear from the methods section.

Model initialization: Unclear which CO2 concentration is used during the spin-up. Using 1996 climate can lead to extreme biases in state variables, and also is not expected to represent pre-industrial conditions.

Model evaluation: I suggest first two plots for model evaluation: 1) model vs. measurement time series of NEE, temperature, etc. and 2) scatter plots.