We appreciate the reviews from Referee#2 and believe that the comments and suggestions have significantly improved our manuscript. In the following, we address specific reviewer comments.

Anonymous Referee #2

The technical note presents a methodology to nudge the predicted groundwater table depth, thereby reducing the number of years required for spin-up of integrated surfacewater-groundwater model Parflow.CLM, based on subsurface storage spin-up criteria. The methodology however does not reduce the real computation time of the model itself, but only reduces the number of years of recursive runs required to initialize the model based on the spin-up criteria. Also, it does not distinguish between the computation time required for each year of spin-up on whether it decreases or it is constant. Although the problem size used in this study cannot be considered to be computationally intensive, which is also a relative term, the idea presented does show some potential to reduce the spin-up period to generate initial soil moisture data. In general, the manuscript is very well written, but at the same time, there are some short comings in the paper that needs to be addressed. There are several instances where the content of the paper is intangible, and inadequacy in experiment designs for the proposed methodology. In addition, the figure quality are very poor in terms of the size of figure, fonts and scale, rendering them unreadable.

We agree with the reviewer that the methodology does not decrease the computational time and it only reduces the spin-up period. As the reviewer suggested we changed spin-up time to spin-up period to clarify this point in the revised manuscript. We do not have the exact record of the computational time for each year of simulation during the spin-up, but our observations show no significant decrease in computational time as system equilibrates. We have significantly revised the manuscript to clearly describe the methodology and highlight the limitations of our approach.

We checked the quality of all the figures and improved them where possible.

Specific comments:

1.Benchmark the term "computationally intensive", which is loosely used throughout the manuscript. Eg., in comparisons to : : :.

We revised the manuscript as follows:

The challenge lies in designing methodologies to reduce spin-up period in computationally intensive integrated hydrologic models such as ParFlow.CLM (Ashby and Falgout, 1996; Jones and Woodward, 2001; Kollet and Maxwell, 2006) when initialization from equilibrium states is required for transient simulations. In integrated hydrologic models like ParFlow, numerical solution of the Richards equation in 3D increases computational time (Kim et al., 1997; Maxwell et al., 2014) in comparison to approaches that use a 1D Richards equation for the vadoze zone and a 2D groundwater flow formulation for simulating subsurface flow.

2.The title "Reducing the spin-up time" appears to be misleading in the sense, whether it is reducing the computation time itself or the number of iterative years required, it needs to be cleared. Eg. "Reducing the spin-up period"...

In the revised manuscript, we changed spin-up time to spin-up period throughout the text.

3.Model description is absent. Which version of the model is being used here, is it the terrain following co-ordinate system or the older version? Are the catchments delineated for the simulation or a box domain is used? This needs to be all clarified. If the terrain following co-ordinate is used, the number of vertical levels can be reduced. In addition, the real computation time can also be reduced using delineated catchments.

We included a brief description of the ParFlow.CLM as well as the version numbers in the revised manuscript. The terrain following coordinate system was not used in our simulations. We used the box domain in our simulations to reduce the impact of boundary conditions on catchment scale fluxes.

A brief description of ParFlow.CLM is added at the beginning of section 2.1 as follows:

ParFlow is a 3D variably saturated groundwater flow model that solves the mixed form of the three-dimensional Richards equation for the subsurface (Ashby and Falgout, 1996; Jones and Woodward, 2001; Maxwell et al., 2014). ParFlow has a fully integrated overland flow simulator (Kollet and Maxwell, 2006) and performs routing of the ponded water on the land surface via the kinematic wave equation. The Common Land Model (CLM 3.0) (Dai *et al.*, 2003) is integrated into ParFlow to simulate water and energy fluxes at the land surface (Maxwell and Miller, 2005; Kollet and Maxwell, 2008). ParFlow.CLM versions 605 and 653 were used for the Skjern River and Baldry simulations respectively, which are described below. The terrain following grid of Maxwell (2013) is not implemented in these modelling set-ups.

Section 2.1.1 is revised as follows:

To reduce the impact of boundary conditions on catchment scale fluxes, the computational domain is extended beyond the delineated catchment boundary. As such, the ParFlow.CLM model domain covered a 28 km by 20 km area that encompasses the Skjern River sub-catchment (Fig. 2).

Section 2.1.2 is revised as follows:

The ParFlow.CLM model of the site was set up over a 2.9 by 2.9 km area encompassing the Baldry sub-catchment (Fig. 2) in order to reduce the impact of boundary conditions on catchment scale fluxes.

4.In both studies, spatially uniform atmospheric forcing is used, could this be possibly one of the reason why the domain mean DTWT function performs well for the relatively flat topography used in this study. How will it effect the empirical DTWT functions, if spatially varying forcing is used? A case study with relatively larger extent, and spatially varying forcing should be presented to prove the presented methodology for its suitability in other applications.

We agree with the reviewer comments regarding examining the performance of DTWT functions across sites with steep topography, larger extent and spatially varying forcing. However, performing additional experiments requires huge computational demand and it is beyond the scope of this Technical Note. We revised the summary section to include this point:

In addition, the role of topography and spatially distributed forcing should be further examined.

5.Pg. 6977, Ln 1-17, This paragraph is very confusing, show the formulation of calculation *of MAE and RMSD in terms of the grid points, and then proceed to discussion, else the figure says otherwise. What does the mean DTWT in y-axis refer to, is it the domain mean or catchment mean? Fig. 3B is addressed before discussion about Fig. 3 itself.*

We included the formula for calculating the objective functions in the methodology section.

Root mean square difference (RMSD), mean absolute error (MAE) and bias were computed to find the best performing DTWT function. These objective functions are calculated as follows:

$$RMSD = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(B_i - M_i)^2} \tag{3}$$

$$MAE = \frac{1}{N}\sum_{i=1}^{N}\left|B_i - M_i\right| \tag{4}$$

$$\%bias = \frac{\sum_{i=1}^{N}(B_i - M_i)}{\sum_{i=1}^{N}(B_i)} \times 100 \tag{5}$$

where N is the number of grid cells in the domain, B is the mean annual DTWT from the baseline simulation of Ajami et al. (2014) for every grid cell, and M is the estimated mean annual DTWT at a grid cell obtained from a DTWT function.

We revised the Results section. Y-axis refers to the domain mean. We revised section 3.1 as follows:

Optimum parameter values for single and double exponential DTWT functions were obtained using nonlinear least squares method. Performance of the single and double exponential DTWT functions in predicting 14 years of DTWT were compared against ParFlow.CLM baseline spin-up simulations (years 7 through 20) of Ajami et al. (2014) to find the optimum empirical DTWT function for the Skjern River sub-catchment. Post-simulation analysis indicates that global DTWT functions based on domain or catchment averaged percentage change values are better predictors of DTWT response compared to local DTWT functions developed for every grid cell. Instability of local DTWT functions occurs in grid cells where percent changes in DTWT oscillate between positive and negative values through initial spin-up simulations. Spatial distribution of these grid cells are shown in Fig S1.

Calculated RMSD and percent bias relative to the baseline spin-up simulations indicate that global double exponential functions using ParFlow.CLM spin-up simulations 2 to 6 provide a better fit compared to various single exponential functions obtained from different spin-up simulation years (e.g. 2 to 3, 2 to 4, etc.). Because the first six cycles of ParFlow.CLM simulations were the same between the baseline spin-up simulations and DTWT distributions from DTWT functions presented in Fig 4, comparisons were made with simulations 7 to 20 of the baseline spin-up approach of Ajami et al. (2014).

As can be seen from Fig. 4a, the mean annual DTWT over the domain derived from the single exponential functions (fitted to percentage change data from simulations 2 to 6) under-predict the baseline spin-up simulations, due to their consistent small underestimates in comparison to double exponential functions fitted to the same data points. Only for the mean absolute error (MAE) calculated at each pixel do single exponential functions based on simulations 2 to 6 perform slightly better and produce smaller errors on average than the double exponential functions (Fig. 4b). It should be noted that the percent bias in mean annual DTWT for simulation cycle 20 is -1.6% for the domain based double exponential function and -6.2% for the single exponential function, with both functions derived from simulation cycles 2 to 6. Therefore, single exponential functions are not further examined in re-initializations of the DTWT. In terms of mean DTWT across the domain (Fig. 4a), the catchment delineated double exponential DTWT function provides a better prediction and the smallest mean bias when compared to the function based on the entire model domain. However, Figure 4b indicates that the mean absolute error values are slightly smaller for the domain based double exponential function. The higher MAE of the catchment based double exponential function is a result of slightly more regions with over and underestimated DTWT values that contribute to a good overall mean DTWT (Fig. 4a), but contains more errors spatially compared to the domain based double exponential function.

Minor Comments:

1.Ln 6970, Ln. 24 : Rephrase.

We rephrased the sentence in the revised manuscript:

Since information on the spatial pattern of water table and soil moisture distributions is generally unavailable, various approaches have been developed to determine the initial DTWT variation.

2.Pg. 6971, Ln 23: spin-up period

It is revised.

3.Pg. 6971, Ln 27: number of years of spin-up required for

The sentence is revised:

The objective of the current study is to develop a *hybrid* spin-up approach that significantly reduces the number of years of spin-up required for model state equilibrium.

4.Pg. 6972, Ln 4: Confusing statement, Fig. 1 mentions 3 stages, but the paragraph begins with two stages.

The sentence is modified:

The hybrid approach consists of three main stages: a two-stage model simulation step and an intermediate state-updating step using the DTWT function.

5.Pg. 6972, Ln 7: "against the equilibrated initial condition for the subcatchment of

: : :.., using the ParFlow.CLM model."
It is revised.
The utility of the proposed scheme is compared against the equilibrated initial condition for a sub-catchment of the Skjern River basin in Denmark, using the ParFlow.CLM model as developed by Ajami et al. (2014) that employed a traditional baseline spin-up approach.

6.Pg. 6972, Ln 20: Mention grid point numbers. Also mention the annual precipitation received and min-max annual temperatures in the text.
This section is revised as follows:
The modelling grid had a horizontal resolution of 500 m and a vertical discretization of 0.5 m. Catchment topography was determined via a 500 m digital elevation model (DEM) and the bottom elevation of the domain was a uniform -75 m, resulting in a 56 × 40 × 406 dimension grid…

In 2003, annual precipitation was 801.6 mm and minimum and maximum daily air temperature were 261.2 K and 295.2 K respectively.

7.Pg. 6973, Ln 24: Also mention the annual precipitation received and min-max annual temperatures in the text. Why 400m deep layer here?
The annual precipitation and min-max annual temperature is added for the Baldry sub-catchment in the revised manuscript:
In 2004, annual precipitation was 674.8 mm, and minmum and maximum daily air temeprature were 277 K and 305.5 K respectively.

The subsurface thickness was not 400 m. The bottom elevation of the domain was 400m. The manuscript is revised as follows:
The bottom elevation of the modelling grid was a uniform 400 m resulting in a subsurface thickness of 43 to 101 m across the computational domain.

8.Pg. 6974, Ln 4: Does these function depend on the initial condition of prescribed groundwater table depth ?

The coefficients and shape of these functions depends on the initial condition of prescribed groundwater table depth. The sentence is revised as follows:

Analysis of ParFlow.CLM spin-up behavior using the baseline spin-up approach for the sub-catchment of the Skjern River identified that percentage changes in subsurface storages and DTWT had the form of an exponential decay for a model initialized from a uniform 3 m DTWT (Ajami et al., 2014). Due to spatial adjustment of the water table during the spin-up, groundwater levels declined near the catchment divide and reached the land surface along the channel network causing an overall decline in mean annual DTWT relative to the initial condition.

9.Pg. 6975, Ln7-14: Paragraph not comprehensible. Rephrase.
The paragraph is revised.

The empirical DTWT functions calculated above estimate percentage changes in mean annual DTWT as a function of simulation year. To predict spatially distributed mean annual DTWT from a global DTWT function, the mean annual DTWT from the final cycle of the ParFlow.CLM spin-up simulation for every grid cell is used as the initial value to successively estimate DTWT distributions as a function of simulation year. These DTWT distributions are based on the predicted percent change values from the global DTWT function. Sensitivity of DTWT functions to the number of ParFlow.CLM cycles was also examined by developing a number of DTWT functions using data from 2 to 6 cycles of ParFlow.CLM. To assess the performance of these DTWT functions, estimated mean annual DTWT from the DTWT functions were compared against mean annual DTWT from the ParFlow.CLM model of Ajami et al. (2014) that had been spun-up for 20 years.

10.Pg.6975, Ln 19: predicted for : : :...
It is revised.

11.Pg. 6976, Ln 20: It has be to discussed clearly, whether the computation domain consists of delineated catchment or a rectangular domain in the experiment description itself.
This point is clarified in the revised manuscript. The computational domain consists of the rectangular domain which includes the delineated catchment.

12.Pg. 6976. Ln 24: A plot showing these oscillations will be illustrative.
A Figure is added to the supplementary information that illustrates regions in the Skjern River sub-catchment were changes in DTWT oscillate.

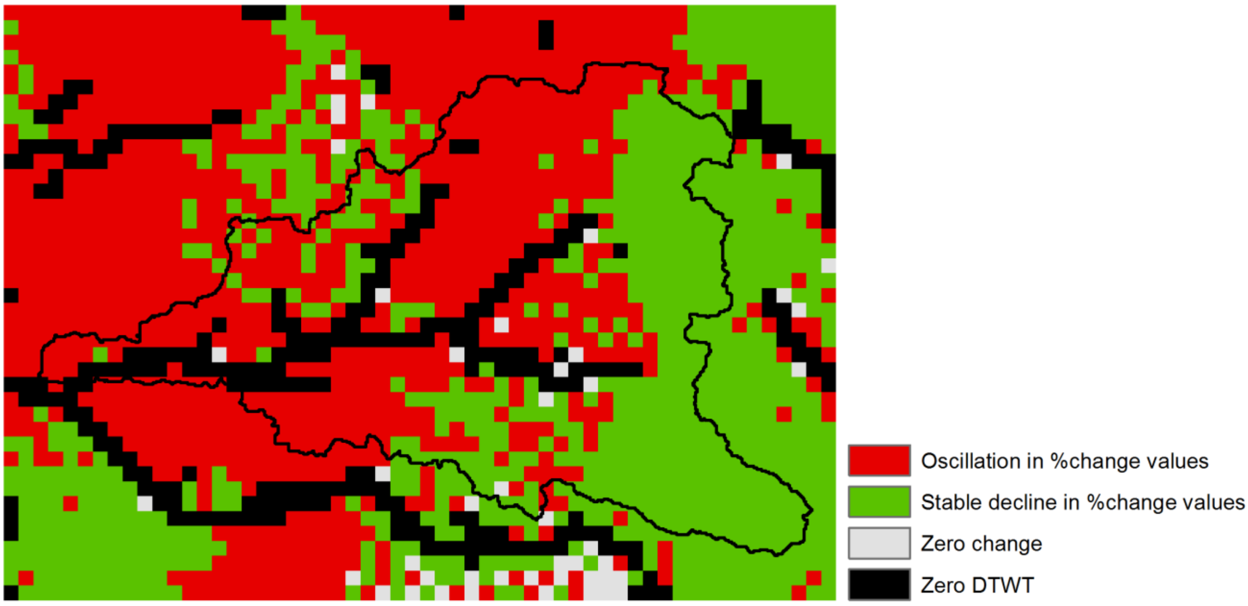

Fig. S1. Delineating four regions in the modelling domain according to percent changes in mean annual DTWT values from six cycles of ParFlow.CLM simulations. Red region represents grid cells where percent changes in DTWT values oscillate between negative and positive values, while the green region corresponds to grid cells with stable decline in DTWT. In grey regions, percent changes in DTWT have reached zero and black region corresponds to the channel network, where DTWT is zero.

13.Pg. 6977, Ln 19: Show the formulation of semi-variograms calculations.
The formulation of semi-variograms are included in the revised manuscript.

To investigate this result further, three empirical semi-variograms were generated. As the impact of an east-west spatial trend in the mean annual DTWT values was evident in the semi-variograms, the trend should first be removed from the mean annual DTWT values. To remove the trend, a plane was fitted to the observed mean annual DTWT values, with an equation of the form:

$$z = ax + by + c \qquad (6)$$

where a, b and c are fitted coefficients, x and y are the coordinates of every grid cell, and z is the mean annual DTWT. Residuals are computed by subtracting the estimated mean annual DTWT from Equation 6 from the observed mean annual DTWT values. Finally, the semi-variogram of the residuals as a function of distance is calculated.

14. Pg. 6979, Ln 2: Is this the result from the simulation using the initial condition from the different methods ?

yes. To clarify this point, the manuscript is revised as follows:

At equilibrium, differences in simulated DTWT from the last day of the ParFlow.CLM simulations after re-initializations (hydrostatic equilibrium and adjusted pressure head distribution) and the baseline spin-up approach varied by up to 2 m inside the catchment boundary (Fig. 6), although most areas were within 0.5m.

15.Pg. 6980, Ln 11: Is it the case in reality?
In our simulations, we assumed that the Baldry sub-catchment is covered by the plantation forest as indicated in the methodology section. In reality half of the catchment is covered by pasture and stage recordings at the catchment outlet indicate 335 days of no flow in 2004.

16.Pg. 6980, Ln 17: "simulation year". Is the computation time same for each year of simulation. Does is also exhibit some pattern?
We do not have the exact computation time for each year of simulation. In general, the computational time slightly decreases as the system equilibrates but the reduction in computational time is not significant.

17.Pg. 6981: Ln 1-7: Far-fetching conclusions. Please remove it.
We removed this statement and revised the manuscript as follows:
Previous efforts in calibrating coupled or integrated hydrologic models required a spin-up process after every parameter update (Stisen et al., 2011; Weill et al., 2013). Development of a computationally efficient spin-up approach will enable this type of systematic calibration of integrated or coupled hydrologic models.

18.Pg. 6981: Ln 9: "reducing number of years to …"
It is revised.

19.Pg. 6981: Ln 10: "spin-up years"

It is revised.