Impacts of model initialization on an integrated surface water – groundwater model

Hoori Ajami^{1,2}, Matthew F. McCabe³, Jason P. Evans^{4,5}

¹ School of Civil and Environmental Engineering, University of New South Wales, Sydney, Australia

² Connected Waters Initiative Research Centre, University of New South Wales, Sydney, Australia

³ Water Desalination and Reuse Center, King Abdullah University of Science and Technology, Thuwal, Saudi Arabia

⁴ Climate Change Research Centre, University of New South Wales, Sydney, Australia

⁵ ARC Centre of Excellence for Climate System Science, University of New South Wales, Sydney, Australia

This article has been accepted for publication and undergone full peer review but has not been through the copyediting, typesetting, pagination and proofreading process, which may lead to differences between this version and the Version of Record. Please cite this article as doi: 10.1002/hyp.10478

Abstract:

Integrated hydrologic models characterize catchment responses by coupling the subsurface flow with land surface processes. One of the major areas of uncertainty in such models is the specification of the initial condition and its influence on subsequent simulations. A key challenge in model initialization is that it requires spatially distributed information on model states, groundwater levels and soil moisture, even when such data are not routinely available. Here, the impact of uncertainty in initial condition was explored across a 208 km² catchment in Denmark using the ParFlow.CLM model. The initialization impact was assessed under two meteorological conditions (wet versus dry) using five depth to water table and soil moisture distributions obtained from various equilibrium states (thermal, root zone, discharge, saturated and unsaturated zone equilibrium) during the model spin-up. Each of these equilibrium states correspond to varying computation times to achieve stability in a particular aspect of the system state.

Results identified particular sensitivity in modeled recharge and stream flow to the different initializations, but reduced sensitivity in modeled energy fluxes. Analysis also suggests that to simulate a year that is wetter than the spin-up period, an initialization based on discharge equilibrium is adequate to capture the direction and magnitude of surface water-groundwater exchanges. For a drier or hydrologically similar year to the spin-up period, an initialization based on groundwater equilibrium is required. Variability of monthly subsurface storage changes and discharge bias at the scale of a hydrological event, show that the initialization impacts do not diminish as the simulations progress, highlighting the importance of robust and accurate initialization in capturing surface water-groundwater dynamics.

KEY WORDS: surface water-groundwater interaction, model initialization, ParFlow.CLM, equilibrium state

INTRODUCTION

Catchment scale soil moisture, surface fluxes and groundwater level distributions often exhibit large spatial variability (Merz and Plate, 1997). Unfortunately, the spatial variability of a catchments physical response or state (i.e. soil moisture and depth to water table; DTWT) is typically unknown due to a scarcity of spatially distributed observations, making robust model initialization, calibration and assessment difficult (McCabe *et al.*, 2005; Stisen *et al.*, 2011). Although, recent advances in measuring soil moisture via remote sensing platforms (Liu *et al.*, 2012; Vereecken *et al.*, 2012) or developments in novel sensors such as cosmic ray probes (Zreda *et al.*, 2008; Zreda *et al.*, 2012; Franz *et al.*, 2013) provide promising tools to represent the spatial variability of soil moisture, remote measurement technique are unable to provide a fully three dimensional distribution of soil moisture in the subsurface (Kampf and Burges, 2007). This limitation in observational data presents a major challenge for initialization of distributed hydrologic models, where spatially and temporally distributed information on catchment state is required (Manfreda *et al.*, 2007).

It has long been recognized that the predictive uncertainty of hydrologic models is impacted by a catchments antecedent moisture condition, as it controls the timing and spatial variability of hydrologic responses (Nikolopoulos *et al.*, 2011). Previous investigations have shown that runoff sensitivity to a catchments initial condition is dependent on the basin size (Nikolopoulos *et al.*, 2011) and that it is controlled by precipitation types and patterns, catchment characteristics (topography, soil and vegetation types) and dominant runoff generation processes (Castillo *et al.*, 2003; Zehe *et al.*, 2005; Vivoni *et al.*, 2007).

In coupled/integrated hydrologic models, the surface and subsurface flow processes are combined to simulate hydrologic fluxes across the flow domains. In these models, the issue of model initialization is of particular importance, as both DTWT and soil moisture distributions must be specified at the start of a simulation. Various approaches have been implemented to initialize coupled/integrated hydrologic models, including: (1) initialization from a calibrated steady state model (Goderniaux *et al.*, 2009), (2) performing a drainage experiment by starting a simulation from a fully saturated condition and continuing the simulation until the simulated discharge matches the observed base flow (Vivoni *et al.*, 2007; Jones *et al.*, 2008), (3) determining the DTWT using a topographic-soil index approach (see Sivapalan *et al.* (1987) and Troch *et al.* (1993)), and (4) initialization from an equilibrium state by specifying water table elevation at a certain depth below the land surface and then repeatedly running the model over a single year or multiple years of forcing data until the system equilibrates according to one or multiple criteria (spin-up process) (Maxwell and Kollet, 2008; Rihani *et al.*, 2010; Ajami *et al.*, 2014a). In a recent multi-criteria assessment of the ParFlow.CLM integrated hydrologic model, the spin-up response illustrated the presence of multiple equilibrium states that were dependent on the convergence metrics used (Ajami *et al.*, 2014a). In that study, thermal equilibrium was reached the fastest, while the system presented a slower convergence to equilibrium based on changes in groundwater and unsaturated zone storages.

Despite the importance of initial conditions on simulated hydrologic fluxes and their overall contribution to predictive uncertainty, there is no clear consensus on the optimum approach to initialize coupled/integrated hydrologic models. Initial condition uncertainty has mostly been discussed in the context of land surface and conceptual hydrologic models and their use in applications such as flood forecasting (Berthet *et al.*, 2009) or in evaluating catchment runoff with respect to precipitation using physically based hydrologic models with

different levels of complexity (Castillo *et al.*, 2003; Noto *et al.*, 2008; Nikolopoulos *et al.*, 2011). In the case of coupled/integrated hydrologic models, assessing the impact of initial condition uncertainty on hydrologic predictions is made even more challenging due to the computational demand for catchment scale simulations. Indeed, it is typical that only a few scenarios are selected to assess the impact of initialization on catchment runoff at the hydrological event scale, rather than examining at monthly or annual time scales.

A number of past efforts have employed different approaches to examine the impact of model initialization in coupled hydrologic models. Shah *et al.* (1996) initialized the SHE model by setting up the phreatic surface uniformly at four different depths below the land surface for ten storms. Results showed larger error in peak discharge and total runoff volume under dry initial conditions compared to wet conditions. Vivoni *et al.* (2007) and Noto *et al.* (2008) used the groundwater rating curve (representing the relation between the base flow and mean DTWT) to choose different initial states for the water table distribution in the tRIBS model (Ivanov *et al.*, 2004). The groundwater rating curve was obtained by initializing the tRIBS model from a fully saturated condition and performing a drainage experiment in the absence of rainfall and evapotranspiration. In that approach, soil moisture was assumed to be in hydrostatic equilibrium with the position of the water table. Although runoff generation under the dry initialization was lower than the shallow water table initializations, simulated flood response to various storm size and wetness condition was not constant (Vivoni *et al.*, 2007). Noto *et al.* (2008) showed that the impact of the initial condition was controlled by rainfall intensity, and the sensitivity of the hydrologic response to initialization varied across the catchment, with less sensitive areas located near the stream network and ridgelines. Due to short simulation times, the role of climatic forcing on modulating the initialization impact has not been explored in previous studies.

Here we employ another approach to assess the sensitivity of surface water-ground water exchange to subsurface initialization. Instead of using arbitrary water table distributions or distributions obtained from a drainage experiment, we use initial conditions obtained from multiple equilibrium states that are identified by various spin-up measures through recursive simulations of the ParFlow.CLM model. The equilibrium states represent equilibrium in a particular aspect of the system state or response and correspond to thermal, root zone, discharge, saturated and unsaturated zone equilibrium. To obtain these equilibrium states, the ParFlow.CLM model was run recursively using forcing data for a particular year. Percent changes in ground surface temperature and subsurface storage through recursive simulations were then used to define the various equilibrium states. This form of initialization from an equilibrium state has an advantage of presenting a soil moisture distribution where its vertical and horizontal variability is controlled by the location of the water table and applied forcing through recursive simulations. Contrary to other approaches for initializations discussed earlier, soil moisture distribution in equilibrium based initialization is not in hydrostatic equilibrium with the water table.

Using water table distributions derived from the different equilibrium states noted above to initialize the ParFlow.CLM model, our objectives were to: (1) assess the impact of subsurface initialization on surface water-groundwater exchange using initial conditions from multiple equilibrium states obtained from ParFlow.CLM spin-up with spatially uniform forcing, (2) examine the role of meteorological condition (wet vs. dry year) on modulating the impact of initialization, and (3) examine convergence of various equilibrium based initializations in time by assessing the monthly water balances, and also changes in daily discharge in response to individual storms.

METHODS

Study site

The focus of this investigation is a sub-catchment of the Skjern River basin, located in western Denmark, which has an area of 208 km² and an elevation range from 32 to 126 m above sea level (Figure 1). Agricultural lands cover the majority of the catchment area (78%), with the remainder covered by evergreen needle leaf forest. The climate is temperate with mean annual precipitation of 984 mm over the period 1990-2010. Meteorological forcing from years 2003 to 2005 were used for the modelling experiment in this study. Monthly precipitation, air temperature and discharge at the gauge for the experimental period are presented in Figure 2.

Experimental set-up

The ParFlow.CLM integrated groundwater-surface water-land surface model (Ashby and Falgout, 1996; Jones and Woodward, 2001; Kollet and Maxwell, 2006; Kollet and Maxwell, 2008) was set up for the study catchment to assess the impact of subsurface initialization on surface water-groundwater interactions. ParFlow is a 3-D variably saturated flow model that solves the three dimensional Richards' equation for the subsurface, and has a fully integrated overland flow simulator (Kollet and Maxwell, 2006). ParFlow is coupled to the Common Land Model (Dai *et al.*, 2003) to simulate water and energy balances at the land surface (Maxwell and Miller, 2005; Kollet and Maxwell, 2008).

The ParFlow.CLM model of the Skjern River sub-catchment was constructed over a 28 km by 20 km area encompassing the catchment, with the domain designed to reduce the impact of lateral boundary conditions on catchment scale fluxes. The horizontal grid

resolution was 500 m in both the X and Y directions, and the vertical discretization of the modelling grid was 0.5 m. Catchment topography was represented using a 500 m pre-processed DEM and the bottom elevation was set to a uniform -75 m, based on the bottom elevation of the extraction and observation wells in the domain. This modelling set-up resulted in a $56 \times 40 \times 406$ dimension grid. As the ParFlow.CLM version 605 does not support variable size grid discretization in the vertical dimension, the choice of horizontal and vertical discretization is influenced by the computational time in this study. For reference, one year of simulation for this modelling set-up required 1000 service units (a service unit is equivalent to 1 hour of time used by one processor) on a high performance parallel computing cluster: equivalent to 1.3 days of computation using 32 processors. The ParFlow overland flow boundary condition was specified at the land surface and no flow boundary conditions were assigned to all the lateral boundaries and bottom of the domain. The no-flow boundary condition of the bottom of the domain is supported by the geologic model of the catchment, which identifies a clay unit at this depth (Stisen *et al.*, 2011). The subsurface was assigned a uniform hydraulic conductivity of $0.3 \text{ (m.hr}^{-1}\text{)}$ using a trial and error procedure for a few initial simulations to match the observed base flow. The remaining subsurface hydraulic parameters were prescribed as follows: porosity (0.39), van Genuchten parameters ($\alpha=1.5 \text{ m}^{-1}$ and $n=2$), and relative residual saturation (0.1). In a prior investigation examining the spin-up process (Ajami *et al.*, 2014a), the ParFlow.CLM model of the same catchment was initialized with a 3 m uniform DTWT and 20 years of recursive simulations were required until equilibrium in subsurface storages were achieved using forcing data for the year 2003 (Ajami *et al.*, 2014a). Equilibrium of subsurface storages ensured equilibrium across a range of other variables including ground surface temperature, energy fluxes and discharge (Ajami *et al.*, 2014a).

To select multiple equilibrium states for initialization, data from 20 years of ParFlow.CLM recursive simulations were used. As noted in Ajami et al. (2014a), defining an equilibrium state depends very much upon the variable used to define spin-up, the time scales where percent changes of state and diagnostic variables are calculated through recursive simulations (daily, monthly, annual), and threshold levels that are used for defining system convergence based on percent change values (i.e. 1%, 0.1%, etc.) (Ajami *et al.*, 2014a).

Here, an equilibrium condition is generally achieved when the percentage changes of a particular variable during recursive simulations reach 0.1% or below for at least one spin-up criterion. For groundwater storage, a threshold level of 0.01% was used. Five different DTWT distributions were selected that corresponded to various equilibrium states in the catchment. These included thermal equilibrium, root zone equilibrium, discharge equilibrium, groundwater equilibrium and unsaturated zone equilibrium. The criteria used to define these equilibrium states and calculated spin-up times based on each criterion are summarized in Table 1. As presented in Table 1, thermal equilibrium, based on changes in ground and bottom layer soil temperature, is reached the fastest. The longest spin-up time corresponds to unsaturated zone equilibrium. Slower convergence of groundwater and unsaturated zone equilibrium is related to spatially distributed adjustment of the water table through recursive simulations, due to topography driven flow and in response to meteorological forcing.

Comparison of DTWT distributions from the above equilibrium states indicate deeper groundwater levels for the unsaturated zone equilibrium compared to other equilibrium states (Figure 3), particularly at high elevation regions in the catchment. Deeper groundwater levels from the unsaturated zone equilibrium state resulted in a larger unsaturated zone storage relative to the previous equilibrium states, because in ParFlow.CLM saturated and unsaturated zones are not defined explicitly and the extent of the unsaturated zone depends on the position of the water table.

To assess the impact of meteorological condition on surface water-groundwater exchange in subsequent simulations, two sets of ParFlow.CLM simulations at an hourly time step were performed using forcing data from the years 2004 and 2005 (Figure 2) along with the same five different DTWT distributions (Figure 3). For each ParFlow.CLM simulation, the subsurface pressure head distributions from the last day of each equilibrium simulation (Table 1) as well as the CLM restart files were used for initialization. Relative to the 2003 spin-up year with annual precipitation of 801.6 mm, total precipitation in years 2004 and 2005 were 1084.8 mm and 858.6 mm respectively. Although, both years are wetter than the spin-up year, the major difference in precipitation is in the second half of the year (August – December) where 2004 is much wetter than the spin-up year and 2005. These forcing data are referred to as wet and dry years in the remainder of this paper because they correspond to annual precipitation below or above the catchment's average precipitation. These particular years are not necessarily the driest and wettest years on record as our selection was limited by the availability of hourly meteorological forcing required for ParFlow.CLM. For our simulations, catchment averaged daily precipitation values were derived from 2 km krigged precipitation fields (Allerup *et al.*, 1997; Stisen *et al.*, 2012), and were evenly distributed to hourly intervals. Hourly air temperature, downward shortwave radiation, wind speed, and atmospheric pressure were obtained from the Danish meteorological institute. Hourly downward longwave radiation was downloaded from NASA's Modern Era Retrospective-Analysis for Research and Applications (MERRA) reanalyses product.

Model evaluation and multi-model intercomparison

To assess the impact of initializations on surface water-groundwater exchanges, simulated discharges from each model run were compared with stream flow observations at the gauge, as shown in Figure 1. Percent bias in runoff is computed as follows:

$$bias = \frac{\sum_{t=1}^N (QS_t - QO_t)}{\sum_{t=1}^N QO_t} \times 100$$

where t is the time step, QS is simulated discharge and QO is observed discharge. As there is no evaluation data for energy fluxes, annual changes in saturated and unsaturated zone storages and DTWT, comparisons were made against model simulations initialized from the unsaturated zone equilibrium state: referred to herein as the baseline simulation. Since initialization from the unsaturated zone equilibrium state ensures equilibrium across a range of variables (including energy fluxes, discharge, root zone soil moisture, groundwater storage and ground surface temperature), simulations from this initialization were considered appropriate to use as the baseline case. To minimize the impact of boundary conditions on catchment scale fluxes, only data from the grid cells inside the catchment boundary are considered.

RESULTS

Impact of initialization on discharge and surface water-groundwater exchanges

The percent bias in the ParFlow.CLM simulated discharge relative to stream flow observations showed that initializations based on thermal equilibrium produced the largest bias in discharge. The magnitudes of these biases for the wet and dry years were 66% and 37%, respectively (Figure 4). Similar to discharge, percent change differences in annual

groundwater and unsaturated zone storage changes relative to the baseline simulations were largest for thermal equilibrium initializations in wet and dry years (Figure 4). Initialization from thermal equilibrium in the dry year even resulted in no groundwater recharge (negative changes in groundwater storage) (Figure 5a). Large differences in discharge and changes in storages indicate that initializations based on thermal equilibrium introduce large errors in both dry and wet year simulations.

Annual groundwater recharge, estimated from changes in groundwater storages in ParFlow.CLM during wet and dry years, showed increases in recharge for simulations initialized from groundwater equilibrium and unsaturated zone equilibrium. Higher groundwater recharge than the thermal equilibrium initializations are due to deeper water tables (Figure 3) in simulations initialized from groundwater and unsaturated zone equilibrium (Figure 5a).

Analysing the differences in mean annual DTWT obtained from four simulations initialized from various equilibrium states to that from the baseline simulation, illustrated that the differences in DTWT were larger in the simulations initialized from a thermal equilibrium state for both dry and wet year simulations (Figure 6). The largest differences in DTWT distributions were observed across the eastern side of the domain, corresponding to higher elevation regions. Ajami et al. (2014a) showed that it takes longer for these high elevation zones to equilibrate in the catchment as a result of uniform DTWT initialization. A comparison of mean annual DTWT distributions with the initial DTWT distributions (Figure 3) highlights the fact that the impact of DTWT initializations were sustained throughout the simulations in both wet and dry years.

Impact of initialization on energy fluxes

The variability in simulated annual evapotranspiration across multiple simulations initialized from different equilibrium states was small, with the exception of initialization from thermal equilibrium, which resulted in higher evapotranspiration due to shallower initial DTWT. Differences in latent heat and sensible heat fluxes relative to the baseline simulations (that initialized from an unsaturated zone equilibrium state) for dry and wet years showed small differences between simulations initialized from root zone and discharge equilibrium (Figure 7).

To further explore the simulated response, the extent of the groundwater coupling zone was calculated across the different simulations. To delineate the groundwater coupling zone, the relationship between mean annual DTWT and annual evapotranspiration for every pixel in the catchment was established based on the equilibrium simulation from year 2003 (Ajami *et al.*, 2014a). Based on this relationship, three distinct zones were identified that correspond to two DTWT thresholds related to changes in the slope of latent heat flux versus mean annual DTWT relationship: a temperature controlled zone ($DTWT < 0.6$ m), a precipitation controlled zone ($DTWT > 5$ m) where evapotranspiration was controlled by precipitation, and a groundwater controlled zone ($0.6 < DTWT < 5$) where evapotranspiration was controlled by the groundwater contribution. As can be seen in Figure 8, with the exception of simulations initialized from thermal equilibrium, differences in the extents of the groundwater coupling zone were small (about 1%) between simulations initialized from different equilibrium states. Therefore, no significant differences were observed in simulated evapotranspiration or the other energy fluxes amongst different simulations for the wet and dry year. For this catchment, the groundwater controlled zone constitutes the majority of the catchment area (more than 45%). As expected, correct initialization of the DTWT is important in simulating land surface fluxes.

Role of meteorological condition on modulating the initialization impacts

In dry year simulations, biases in simulated discharge from various equilibrium based initializations are smaller relative to the wet year simulations. However, despite the smaller percent bias in discharge for the dry year simulations, percent changes in annual groundwater and unsaturated zone storage changes relative to the baseline simulations were larger in the dry year compared to the wet year (Figure 4). Changes in groundwater storage in a model initialized from the root zone equilibrium were more than twice as large as the model initialized from discharge equilibrium, particularly in the dry year (Figure 4).

To investigate this further, we examined the spatial variability of annual changes in subsurface storages in both wet and dry years. As can be seen in Figure 5b, the percentage of recharge cells inside the catchment remained constant across simulations initialized from discharge, groundwater and unsaturated zone equilibriums in the wet year. This result indicated that in a wet year, initialization from discharge equilibrium was adequate to capture the direction of changes in groundwater storage. Small differences amongst simulations initialized from various equilibrium states during the wet year were present in the magnitudes of recharge (Figure 5a). However, in the dry year, both magnitude and directions of recharge changed across multiple simulations initialized from different equilibrium states (Figure 5b). In the dry year, initialization based on discharge equilibrium was not able to capture directions of surface water-ground exchange, particularly in high elevation areas in the catchment. As shown in Figure 5b, the impact of DTWT initializations was more pronounced on groundwater recharge in the dry year. In line with expectation, during the wet year, groundwater recharge was higher than in the dry year.

Impact of initialization across monthly and daily time scales

The monthly percent bias in discharge and the variability of monthly saturated and unsaturated storage changes were explored to assess whether the impact of initializations dissipate with time for a given year. Monthly percent bias in modelled discharge (relative to observations) decreased for the simulations initialized from groundwater and unsaturated zone equilibrium states for both dry and wet years (Figure 9), similar to the annual scale. However, for any given year, the magnitudes of bias did not decrease by the end of the year. The largest percent bias in discharge for the wet and dry years were in September and November across all the initializations respectively. Although precipitation was higher in these months relative to the rest of the year (Figure 2), increases in bias were related to the distribution of large precipitation events in a given month (uniform vs. skewed distribution).

To explore the sensitivity of monthly saturated and unsaturated zone storage changes to various initializations, monthly groundwater and unsaturated zone storage changes were calculated for every simulation. Overall, monthly groundwater storage changes constitute a larger component of the water balance compared to unsaturated zone storage changes. As can be seen from Figure 10, differences in monthly groundwater storage changes amongst the simulations are within the magnitude of recharge or discharge. For the unsaturated zone storage, the direction and magnitude of monthly storage changes across simulations initialized from different equilibrium states are impacted. These differences are most pronounced in spring and summer. This analysis illustrates that while simulations from different equilibrium states tend to agree in predicting storage changes in certain months, differences across model predictions do not diminish with time as the simulations progress.

At the individual storm scale, percent changes in daily stream flows for a given storm decreased for simulations initialized from groundwater and unsaturated zone equilibrium. However, the impact of initialization did not diminish based on the timing of the storm during

the year, indicating the importance of proper initialization of coupled/integrated hydrologic models even when continuous simulations are used to estimate daily discharge. It should be noted that for this particular catchment, with a base flow index of greater than 0.7, groundwater has a major influence on discharge. Therefore, DTWT initialization is especially important in capturing discharge dynamics.

DISCUSSION

Impacts of subsurface initializations on surface water-groundwater exchange

While the impacts of subsurface initializations have previously been assessed for simulating flood response in coupled hydrologic models, their role on surface water-groundwater exchanges at longer time scales (monthly and annual) has not been fully explored. Our results indicate that subsurface initializations have large impacts on stream discharge and changes in subsurface storages even at longer time scales such as monthly and annual. Although the impact of initialization can be partly modulated depending on the meteorological condition (as has been observed in a wet year here), improper initialization can have large impacts on both the magnitude and direction of surface water-groundwater exchanges.

Amongst the simulated fluxes, energy balance variables are the least sensitive to the equilibrium based initializations used here. Our results indicate that initialization based on root zone equilibrium is adequate to capture the dynamics of energy balances in subsequent simulations. In this particular catchment, where groundwater constitutes a major component of stream flow, equilibrium of subsurface storages is of particular importance to capture discharge and changes in surface water-groundwater storages. It should be noted that equilibrium of snow storage is also important in catchments where snow melt represents a major component of discharge during the spring and summer.

Implications for model spin-up

The spin-up approach has been commonly used for initializing ParFlow.CLM models (e.g. Kollet and Maxwell, 2008). However, the spin-up process is computationally expensive, especially when the model is implemented at the catchment and regional scales. Here, our results illustrate that for a year similar to the spin-up year, an initialization based on groundwater storage is required to accurately capture the surface water-groundwater dynamics. This means that for this particular catchment, at least 16 years of recursive simulations are required until the groundwater system is equilibrated. The computational time to perform these simulations was 16,000 service units, which is equivalent to 21 days of computation time using 32 processors. For the wet year, only 11 years of spin-up simulations were required. Therefore, one approach for reducing spin-up time could be to perform two sets of model simulations. The first set of simulations include performing model spin-up using data immediately preceding a wet year and continue simulations until discharge is equilibrated. In the next stage, one year of simulation would be performed using the wet year forcing. Using this approach, the number of spin-up simulations could be reduced to 12 years instead of 16 years. However, to understand the consistency of reductions in spin-up time, additional experiments need to be performed across catchments with a range of climate, topography, subsurface heterogeneity and vegetation features. Additional investigation should be performed to assess the efficiency of this approach against the newly developed hybrid spin-up approach based on the empirical DTWT functions (Ajami *et al.*, 2014b). A combination of these approaches may further reduce the computational time of the spin-up process.

Limitations and future work

Due to computational demand, only a few scenarios were selected for assessing initialization impacts on an integrated hydrologic model. While the chosen scenarios were selected to represent the catchments equilibrium in relation to a particular aspect of the system state or response, more realizations are required to comprehensively assess the impacts of initialization on catchment response. Subsurface heterogeneity is not represented in this modelling experiment due to the conceptual uncertainty of geological models and their impact on model predictions (Refsgaard *et al.*, 2006). In addition, the impact of groundwater pumping in conjunction with the role of initializations on surface water-groundwater exchanges in the catchment was not explicitly assessed: an element that could be quite important in basins with considerable groundwater pumping and irrigation. It should be noted that this basic model set-up focuses on an examination of the role of surface and subsurface flow interactions on the spin-up times and on assessing the impact of initial conditions on subsequent simulations. Further investigation of the impacts on the equilibration time due to variability in subsurface heterogeneity and hydraulic conductivity, aquifer geometry and boundary conditions are required.

In addition to assessing the impact of initialization approach on catchment response across a range of different climate and watershed characteristics, the impact of spatial resolution and other model set-up and parameterization including distributed atmospheric forcing on modulating the initialization impact should also be examined. Performing additional experiments across multiple sites and with a range of coupled/integrated hydrologic models, will ultimately lead to the development of efficient methods for initialization of coupled/integrated hydrologic models.

SUMMARY AND CONCLUSIONS

An integrated hydrologic model was used to assess the impact of subsurface initialization on simulating water and energy fluxes when the model was initialized from various equilibrium states. Results showed that initializations from thermal equilibrium were not adequate for simulating catchment responses. While initialization based on discharge equilibrium was sufficient to simulate a wetter than average year, in order to simulate a drier than average year (similar to the spin-up year), initialization based on groundwater storage was required to capture surface water-groundwater dynamics. This means that 11 and 16 years of spin-up simulations were sufficient for model initialization prior to a wet and a dry year respectively. The analysis also illustrated that the impact of initialization on daily discharge or monthly storage changes did not diminish as the simulations progressed in any given year. For groundwater storage changes, the largest variability between simulations was observed for the driest months.

Results from these experimental simulations in wet and dry years illustrated the impact of meteorological condition on modulating the effect of initialization in an integrated hydrologic model. If possible, it is recommended to perform model spin-up using data immediately preceding a wet year in order to reduce the number of years of spin-up required to reach an adequate initial state, and to minimize the impact of the initial state on the subsequent simulation.

ACKNOWLEDGEMENTS

This research was funded by the Australian Research Council and the National Water Commission through support of the National Centre for Groundwater Research and Training (NCGRT). We acknowledge the support provided by the National Computational Infrastructure at the Australian National University through the National Computational Merit Allocation Scheme and Intersect partner share for high performance computing. The use of high performance computing facilities at the University of New South Wales is also appreciated. We acknowledge the HOBE Hydrological Observatory funded by the Villum Kann Rasmussen Foundation for providing model data. We would also like to thank the two anonymous reviewers for their valuable comments.

REFERENCES

- Ajami H, McCabe MF, Evans JP, Stisen S. 2014a. Assessing the impact of model spin-up on surface water-groundwater interactions using an integrated hydrologic model. *Water Resources Research*, **50**: 2636-2656. DOI: 10.1002/2013WR014258.
- Ajami H, Evans JP, McCabe MF, Stisen S. 2014b. Technical Note: Reducing the spin-up time of integrated surface water-groundwater models. *Hydrol. Earth Syst. Sci.*, **18**: 5169-5179. DOI: 10.5194/hess-18-5169-2014.
- Allerup P, Madsen H, Vejen F. 1997. A comprehensive model for correcting point precipitation. *Nordic Hydrology*, **28**: 1-20.
- Ashby SF, Falgout RD. 1996. A parallel multigrid preconditioned conjugate gradient algorithm for groundwater flow simulations. *Nuclear Science and Engineering*, **124**: 145-159.
- Berthet L, Andréassian V, Perrin C, Javelle P. 2009. How crucial is it to account for the antecedent moisture conditions in flood forecasting? Comparison of event-based and continuous approaches on 178 catchments. *Hydrology and Earth System Sciences*, **13**: 819-831. DOI: 10.5194/hess-13-819-2009.
- Castillo VM, Gómez-Plaza A, Martínez-Mena M. 2003. The role of antecedent soil water content in the runoff response of semiarid catchments: a simulation approach. *Journal of Hydrology*, **284**: 114-130. DOI: [http://dx.doi.org/10.1016/S0022-1694\(03\)00264-6](http://dx.doi.org/10.1016/S0022-1694(03)00264-6).
- Dai Y, Zeng X, Dickinson RE, Baker I, Bonan GB, Bosilovich MG, Denning AS, Dirmeyer PA, Houser PR, Niu G, Oleson KW, Schlosser CA, Yang ZL. 2003. The Common Land Model. *Bulletin of the American Meteorological Society*, **84**: 1013-1023. DOI: 10.1175/bams-84-8-1013.
- Franz TE, Zreda M, Ferre TPA, Rosolem R. 2013. An assessment of the effect of horizontal soil moisture heterogeneity on the area-average measurement of cosmic-ray neutrons. *Water Resources Research*, **49**: 6450-6458. DOI: 10.1002/wrcr.20530.

- Goderniaux P, Brouyère S, Fowler HJ, Blenkinsop S, Therrien R, Orban P, Dassargues A. 2009. Large scale surface–subsurface hydrological model to assess climate change impacts on groundwater reserves. *Journal of Hydrology*, **373**: 122-138.
- Ivanov VY, Vivoni ER, Bras RL, Entekhabi D. 2004. Catchment hydrologic response with a fully distributed triangulated irregular network model. *Water Resources Research*, **40**: W11102. DOI: 10.1029/2004wr003218.
- Jones JE, Woodward CS. 2001. Newton–Krylov-multigrid solvers for large-scale, highly heterogeneous, variably saturated flow problems. *Advances in Water Resources*, **24**: 763-774. DOI: 10.1016/s0309-1708(00)00075-0.
- Jones JP, Sudicky EA, McLaren RG. 2008. Application of a fully-integrated surface–subsurface flow model at the watershed-scale: A case study. *Water Resources Research*, **44**: W03407. DOI: 10.1029/2006wr005603.
- Kampf SK, Burges SJ. 2007. A framework for classifying and comparing distributed hillslope and catchment hydrologic models. *Water Resources Research*, **43**: W05423. DOI: 10.1029/2006WR005370.
- Kollet SJ, Maxwell RM. 2006. Integrated surface–groundwater flow modeling: A free-surface overland flow boundary condition in a parallel groundwater flow model. *Advances in Water Resources*, **29**: 945-958. DOI: 10.1016/j.advwatres.2005.08.006.
- Kollet SJ, Maxwell RM. 2008. Capturing the influence of groundwater dynamics on land surface processes using an integrated, distributed watershed model. *Water Resources Research*, **44**: W02402. DOI: 10.1029/2007wr006004.
- Liu YY, Dorigo WA, Parinussa RM, de Jeu RAM, Wagner W, McCabe MF, Evans JP, van Dijk AIJM. 2012. Trend-preserving blending of passive and active microwave soil moisture retrievals. *Remote Sensing of Environment*, **123**: 280-297. DOI: <http://dx.doi.org/10.1016/j.rse.2012.03.014>.
- Manfreda S, McCabe MF, Fiorentino M, Rodríguez-Iturbe I, Wood EF. 2007. Scaling characteristics of spatial patterns of soil moisture from distributed modelling. *Adv. Water Resour.*, **30**: 2145-2150. DOI: 10.1016/j.advwatres.2006.07.009.
- Maxwell RM, Miller NL. 2005. Development of a Coupled Land Surface and Groundwater Model. *Journal of Hydrometeorology*, **6**: 233-247. DOI:10.1175/JHM422.1.
- McCabe MF, Kalma JD, Franks SW. 2005. Spatial and temporal patterns of land surface fluxes from remotely sensed surface temperatures within an uncertainty modelling framework. *Hydrol. Earth Syst. Sci.*, **9**: 467-480. DOI: 10.5194/hess-9-467-2005.
- Merz B, Plate EJ. 1997. An analysis of the effects of spatial variability of soil and soil moisture on runoff. *Water Resources Research*, **33**: 2909-2922. DOI: 10.1029/97WR02204.
- Nikolopoulos EI, Anagnostou EN, Borga M, Vivoni ER, Papadopoulos A. 2011. Sensitivity of a mountain basin flash flood to initial wetness condition and rainfall variability. *Journal of Hydrology*, **402**: 165-178. DOI: <http://dx.doi.org/10.1016/j.jhydrol.2010.12.020>.
- Noto LV, Ivanov VY, Bras RL, Vivoni ER. 2008. Effects of initialization on response of a fully-distributed hydrologic model. *Journal of Hydrology*, **352**: 107-125. DOI: <http://dx.doi.org/10.1016/j.jhydrol.2007.12.031>.
- Refsgaard JC, van der Sluijs JP, Brown J, van der Keur P. 2006. A framework for dealing with uncertainty due to model structure error. *Adv. Water Resour.*, **29**: 1586-1597. DOI: <http://dx.doi.org/10.1016/j.advwatres.2005.11.013>.
- Rihani JF, Maxwell RM, Chow FK. 2010. Coupling groundwater and land surface processes: idealized simulations to identify effects of terrain and subsurface heterogeneity on land surface energy fluxes. *Water Resources Research*, **46**: W12523. DOI: 10.1029/2010wr009111.

- Shah SMS, O'Connell PE, Hosking JRM. 1996. Modelling the effects of spatial variability in rainfall on catchment response. 2. Experiments with distributed and lumped models. *Journal of Hydrology*, **175**: 89-111. DOI: [http://dx.doi.org/10.1016/S0022-1694\(96\)80007-2](http://dx.doi.org/10.1016/S0022-1694(96)80007-2).
- Sivapalan M, Beven K, Wood EF. 1987. On hydrologic similarity: 2. A scaled model of storm runoff production. *Water Resources Research*, **23**: 2266-2278. DOI: [10.1029/WR023i012p02266](http://dx.doi.org/10.1029/WR023i012p02266).
- Stisen S, McCabe MF, Refsgaard JC, Lerer S, Butts MB. 2011. Model parameter analysis using remotely sensed pattern information in a multi-constraint framework. *Journal of Hydrology*, **409**: 337-349. DOI: <http://dx.doi.org/10.1016/j.jhydrol.2011.08.030>.
- Stisen S, Højberg AL, Troldborg L, Refsgaard JC, Christensen BSB, Olsen M, Henriksen HJ. 2012. On the importance of appropriate precipitation gauge catch correction for hydrological modelling at mid to high latitudes. *Hydrology and Earth System Sciences*, **16**: 4157-4176. DOI: [10.5194/hess-16-4157-2012](http://dx.doi.org/10.5194/hess-16-4157-2012).
- Troch PA, De Troch FP, Brutsaert W. 1993. Effective water table depth to describe initial conditions prior to storm rainfall in humid regions. *Water Resources Research*, **29**: 427-434. DOI: [10.1029/92wr02087](http://dx.doi.org/10.1029/92wr02087).
- Vereecken H, Huisman JA, Pachepsky Y, Montzka C, van der Kruk J, Bogaen H, Weihermüller L, Herbst M, Martinez G, Vanderborght J. 2012. On the spatio-temporal dynamics of soil moisture at the field scale. *Journal of Hydrology*, **516**: 76-96. DOI: <http://dx.doi.org/10.1016/j.jhydrol.2013.11.061>.
- Vivoni ER, Entekhabi D, Bras RL, Ivanov VY. 2007. Controls on runoff generation and scale-dependence in a distributed hydrologic model. *Hydrology and Earth System Sciences*, **11**: 1683-1701. DOI: [10.5194/hess-11-1683-2007](http://dx.doi.org/10.5194/hess-11-1683-2007).
- Zehe E, Becker R, Bárdossy A, Plate E. 2005. Uncertainty of simulated catchment runoff response in the presence of threshold processes: role of initial soil moisture and precipitation. *Journal of Hydrology*, **315**: 183-202. DOI: <http://dx.doi.org/10.1016/j.jhydrol.2005.03.038>.
- Zreda M, Desilets D, Ferré TPA, Scott RL. 2008. Measuring soil moisture content non-invasively at intermediate spatial scale using cosmic-ray neutrons. *Geophysical Research Letters*, **35**: L21402. DOI: [10.1029/2008GL035655](http://dx.doi.org/10.1029/2008GL035655).
- Zreda M, Shuttleworth WJ, Zeng X, Zweck C, Desilets D, Franz T, Rosolem R. 2012. COSMOS: the COsmic-ray Soil Moisture Observing System. *Hydrol. Earth Syst. Sci.*, **16**: 4079-4099. DOI: [10.5194/hess-16-4079-2012](http://dx.doi.org/10.5194/hess-16-4079-2012).

Table 1. Multiple equilibrium states defined based on ParFlow.CLM spin-up simulations using 2003 forcing.

Equilibrium state	Number of spin-up simulations	Criteria & threshold level	DTWT* μ (σ)
Thermal (Th)	2	Annual and daily criteria, 0.1%	3.8 (4.8)
Root zone (RZ)	8	Daily criteria, 0.1%	5.0 (6.4)
Discharge (Q)	11	Daily criteria, 0.1%	5.2 (6.6)
Groundwater storage (GW)	16	Daily criteria, 0.01%	5.3 (6.8)
Unsaturated storage (UZ)	20	Daily criteria, 0.1%	5.4 (6.9)

*Mean (μ) and standard deviation (σ) of DTWT across the domain obtained from various equilibrium states. Daily criteria are based on one or a few of the following criteria: percent change of a given variable between the first days of recursive simulation, last days of recursive simulation, beginning and end of a given simulation year, and differences in daily minimum and maximum values through recursive simulations. Annual criterion is based on changes in mean annual model output in recursive simulations (Ajami *et al.*, 2014a).

Figure 1. The studied sub-catchment of the Skjern river basin located in western Denmark. The rectangular blue area represents the computational domain of the ParFlow.CLM simulation.

Figure 2. Monthly precipitation (mm), monthly mean air temperature (K) and monthly mean discharge for the spin-up year (2003), wet year (2004) and dry year (2005).

Figure 3. Differences in initial DTWT distributions of various equilibrium states for a) thermal equilibrium, b) root zone equilibrium, c) discharge equilibrium, and d) groundwater equilibrium from the unsaturated zone equilibrium state. Larger differences in DTWT distributions were observed in high elevation regions at the eastern side of the catchment.

Figure 4. Percent bias in discharge relative to the stream flow observations at the gauge (Figure 1) and percent changes in simulated annual changes in groundwater and unsaturated zone storages relative to the baseline simulation (initialization from unsaturated zone equilibrium) for dry and wet years. See Table 1 for the description of equilibrium states (Th, RZ, Q, GW, UZ).

Figure 5. a) Annual groundwater recharge and b) percentage of recharge cells inside the catchment from models initialized using different equilibrium states (Table 1) in wet and dry years.

Figure 6. Differences in mean annual DTWT obtained from simulations initialized from a) thermal equilibrium, b) root zone equilibrium, c) discharge and d) groundwater equilibrium states compared to the baseline simulation for the wet year. No significant differences were observed between the dry and wet years. Differences in DTWT were larger in a simulation initialized from thermal equilibrium state.

Figure 7. Differences in annual latent heat and sensible heat fluxes relative to the baseline simulations for dry and wet years for simulations initialized from different equilibrium states (Table 1).

Figure 8. Variability in the extents of temperature, groundwater and precipitation controlled zones in the catchment relative to the equilibrium simulation from 2003. In both wet and dry years, the extents of groundwater coupling zones were adjusted depending on the climatic conditions in the catchment. The extents of temperature and groundwater controlled zones were greater in the wet year across all equilibrium states.

Figure 9. Monthly bias in runoff relative to observations in a) wet and b) dry years in the catchment. Overall, monthly bias was higher in the wet year compared to the dry year. The magnitudes of bias decreased for simulations initialized from groundwater and unsaturated zone equilibrium states.

Figure 10. Monthly groundwater (top) and unsaturated zone (bottom) storage changes in (mm) across simulations initialized from different equilibrium states (root zone (RZ), discharge (Q), groundwater (GW) and unsaturated zone (UZ) equilibrium) in wet and dry years. Results of initializations from thermal equilibrium are not shown.