A. Atchley
aatchley@lanl.gov

Comment by Adam L. Atchley, Dylan R. Harp, Ethan T. Coon, Cathy J. Wilson

The authors present an impressive modeling effort investigating 3D water and energy simulations of polygonal tundra. Their research is of interest to the community as it undoubtedly yields insight into the thermal hydrology of polygonal tundra. Furthermore, a modeling effort informed by extensive field measurements provides a unique opportunity to validate and properly shape the process rich models currently being developed for terrestrial Arctic applications. It is for this particular reason, that we are first interested in this manuscript and second concerned with the message in section 5.1. In particular, we refer to line 2-3 on page 25, "At our study sites, while calibration may compensate for lack of data, it does not improve our understanding of the system."
In section 5.1 the authors provide reasons for not calibrating the model to the observed data available at the study site, specifically that process rich models have high degrees of freedom and therefore are plagued with non-uniqueness (equifinality). In other words, there are multiple combinations of parameters, or more generally model structures (Beven, 2006) that can produce optimized results which fit observed data equally well. While the authors do not quantitatively demonstrate the existence of equifinality here, non-unique parameter combinations certainty exist in this situation, as has been systematically identified for thermal hydrological models at the same site by Atchley et al. (2015) using multi-try calibration and rigorously quantified by Harp et al., (2016) using Null-Space Monte Carlo. However, it is our understanding that the literature addressing equifinalty does not argue for giving up calibration as a lost cause, but rather strongly suggests that additional efforts are required to account for a set or distribution of parameter combinations consistent with observations (e.g. Vrugt et al., 2009, Vrugt and Ter Braak, 2011; Bárdossy, 2007; Tonkin, 2009) and model structural error (Beven, 2005; Clark et al., 2008; Fenicia et al., 2011; Larson et al., 2014).

The research behind calibration and model optimization has long since evolved from simple parameter fits to more strategic calibration methods (Hill, 1998). Therefore, we believe that equifinality does not provide a justification to avoid calibration, especially if the objective of the modeling exercise is to improve understanding of system and model behavior. On the contrary, it has been our experience that, while difficult, time-consuming, and computationally expensive, extensive, systematic multi-try calibration can yield important system understanding and identify model capabilities and limitations. The work presented in Atchley et al., (2015) at the same site at the Barrow Environmental Observatory, shows that systematic multi-try calibration can be used as a tool to reduce model structural error and achieve system understanding. For example, calibration efforts led to the recognition of the importance of the representation of snow distribution and depth hoar formation in our models. These insights are not simply better

model parameters, but are physical representations of system components; this effort led to better system understanding. Furthermore, quantifying the equifinality of the combined model and represented system then allows for quantification of model uncertainty, where for example the projected ALT uncertainty attributed to parameter uncertainty can be measured and compared to meteorological and/or climate model uncertainty (Harp et al., 2016). Moreover, the parameter sensitivity quantified by such exercises has, in our opinion, provided valuable information for reducing model uncertainty. Porosity and material thermal conductivity measurements are shown to have the greatest potential to reduce projected ALT uncertainty (Harp et al., 2016), thereby directing which additional field data and process understanding are necessary to reduce uncertainty.

In the context of the model presented in this manuscript, we realize that exhaustive model calibration may be computationally infeasible, and we also do not over look the valuable contribution presented here as the 3D representation of energy and water fluxes in freeze-thaw polygonal tundra indeed pushes the boundaries of process-rich mechanistic modeling. Therefore, it is not our wish to force model calibration and parameter sensitivity analysis on the current manuscript. However, we strongly encourage the authors to reconsider the stated view of model calibration and to discuss how calibration and parameter sensitivity may provide insight into model performance as well as system understanding in polygonal tundra.

We thank Atchley et al. for their interest in our work and insightful comments. We completely agree with the comments above. Model calibration has since long been used a powerful and insightful tool in hydrology and it was certainly not our intention to overlook its power, usefulness and validity. Objective of our presented study was to apply the modeling framework to the study site and synthesize and use available observational data sets for the presented case studies. Work by Atchley et al. 2015 and Harp et al. 2016 provides methodology for calibration, insights in parameter sensitivity and calibrated parameters for three phase thermal hydrology model similar to ours and would certainly complement and guide our future modeling studies. We have updated our manuscript to add a reference to Atchley et al. 2015, Harp et al. 2016 and relevance of their work for our study (Page 24 Lines 16-18).

References
Atchley, A., S. Painter, D. Harp, E. Coon, C. Wilson, A. Liljedahl, and V. Romanovsky (2015a), Using field observations to inform thermal hydrology models of permafrost dynamics with ATS (v0.83), Geoscientific Model Development, 8, 2701-2722. Doi: 10.5194/gmd-8-2701-2015.
Bárdossy, A. "Calibration of hydrological model parameters for ungauged catchments." Hydrology and Earth System Sciences Discussions 11.2 (2007): 703-710.
Beven K. 2005. On the concept of model structural error. Water Science and Technology. 52(6): 167-175

Beven K. 2006. A manifesto for the equifinality thesis. Journal of Hydrology. 320: 18-36. doi:10.1016/j.jhydrol.2005.07.007.

Clark, M. P., Slater, A. G., Rupp, D. E., Woods, R. A., Vrugt, J. A., Gupta, H. V., ... & Hay, L. E. 2008. Framework for Understanding Structural Errors (FUSE): A modular framework to diagnose differences between hydrological models. Water Resources Research, 44(12) W00B02, doi:10.1029/2007WR006735,

Fenicia, F., Kavetski, D., & Savenije, H. H. 2011. Elements of a flexible approach for conceptual hydrological modeling: 1. Motivation and theoretical development. Water Resources Research, 47(11). W11510, doi:10.1029/2010WR010174

Harp, D., A. L. Atchley, S. L. Painter, E. Coon, C. Wilson, V. Romanovsky, and J. Rowland (2015), Effect of soil property uncertainties on permafrost thaw projections: a calibration-constrained analysis, The Cryosphere 10(3), 1-18. doi: 10.5194/tc-10-1-2016.

Hill, Mary C. Methods and guidelines for effective model calibration. Denver, CO, USA: US Geological Survey, 1998.

Kavetski, D., & Fenicia, F. 2011. Elements of a flexible approach for conceptual hydrological modeling: 2. Application and experimental insights. Water Resources Research, 47(11). W11511, doi:10.1029/2011WR010748,

Keating, Elizabeth H., et al. "Optimization and uncertainty assessment of strongly nonlinear groundwater models with high parameter dimensionality." Water Resources Research 46.10 (2010).
Printer-friendly version

Larsen L., Thomas C., Eppinga M. 2014. Exploratory modeling: extracting causality from complexity. EOS, 95(12) 285-292.

Vrugt, Jasper A., Cajo JF Ter Braak, Hoshin V. Gupta, and Bruce A. Robinson. "Equi-finality of formal (DREAM) and informal (GLUE) Bayesian approaches in hydrologic modeling?." Stochastic environmental research and risk assessment 23, no. 7 (2009): 1011-1026. TCD

Tonkin, Matthew, and John Doherty. "Calibration‐constrained Monte Carlo analysis of highly parameterized models using subspace techniques." Water Resources Research 45, no. 12 (2009). Interactive
comment

Vrugt, Jasper A., and C. J. Ter Braak. "DREAM (D): an adaptive Markov Chain Monte Carlo simulation algorithm to solve discrete, noncontinuous, and combinatorial posterior parameter estimation problems." Hydrology and Earth System Sciences 15, no. 12 (2011).