Letter to the Editor:

Dear Dr. Niko Wanders,

Thank you very much for handling our manuscript in HESSD and facilitating the fruitful open discussion process. We did our best to address reviewers' comments in the revised manuscript. The major changes are as follow:

1- We have reshuffled the structure of the manuscript so that the concept of vector-based configuration of land models is presented earlier and separately in the manuscript (currently in Section 2).
2- We have reworded the VIC-GRU to vector-based configuration of land models instead. We used the concept of the GRU for parameter allocation in this work [which can be different for other studies and parameter allocation].
3- We have re-calibrated the parameters with snow roughness as requested by the first reviewer.
4- We have merged two of the Figures and simplified the first Figure. The numbering of the Figures may have changed due to the change in the structure of the manuscript.
5- We fully removed Experiment-3 which was exploring the presence of the macropore water movement in the VIC model.

We hope the current changes in the manuscript satisfy the editor and the reviewers and hopefully warrant publication in the current format. Looking forward to hearing back from you.

With kind regards,

Shervan Gharari, on behalf of the co-authors

Answer to the comments by anonymous reviewer#1

We thank the reviewer for their constructive comments on our work. For convenience, the reviewer's comments are given in green and our response is in blue.

Gharari et al present an application of the VIC model, using Grouped Response Units to define computational units, rather than grids. It is acknowledged that this concept was already presented in 1993. I do think it is justified to re-introduce older concepts if these can serve the science of today, however, then the re-introduction should also deal with some of the challenges of today, and this is currently not the case.

As the reviewer rightly mentioned, this manuscript does not intend to introduce the concept of GRU, instead, it tries to use it as a base for implementation of the VIC model in a vector-based fashion which has been used worldwide and hopefully draw the attention for wider land model community to use the concept of vector-based setup (and based on GRUs parameterization).

Solving the "challenges of today" (which the reviewer describes in more depth in their later comments, and which we respond to in more depth later in this document) is not the main goal of this paper. In this manuscript, we try to point out the technical and scientific advantages of using a vector-based setup. To do so, here we reflect on one of often not very well explored challenges in land modeling community, that is trade-off between accuracy of the land models' spatial representation and their performance. We hope our paper sheds some light on the ongoing discussion. The vector-based implementation concept was very helpful in this respect as we could change the resolution of forcing without really affecting the parameter values at the computational units level as there is no upscaling to the grid resolution. It was also very helpful to use the GRU concept in the parameterization of the VIC model however other parameterization techniques could be explored within the vector-based implementation of land models. We think that is a major advantage that we highlighted in this manuscript. Also, following the suggestion by the second reviewer, we have moved the use of GRU concept in grouping the parameter values for a specific combination of geospatial data after introducing the vector-based implementation concept.

Firstly, the reader has to do quite some searching to fully capture the concept of GRU's, and its comparison to HRU's. Only when the investigated cases are presented it be- comes clear what a GRU exactly is and the choices it encompasses when defining GRU's. This seems to be the result of an overall quite weak structure in the manuscript; the introduction does not clearly present the aim or goal, probably because the structural test (case 3, presented in the intro in line 103-110) seems to be completely out of context. In the same fashion, 3.1.3 is not well embedded. Furthermore, sections are not logically structured, e.g. subsection 3.3 only consists of 2 sentences while some subsubsections are longer, and parameter are presented well after the calibration is discussed and the cases are introduced. I suggest restructuring the manuscript, clearly introducing the concepts with simple examples, and omitting parts that do not fit the aim or goal of the study.

We agree with the reviewer that the introduction structure can be improved (this is also mentioned by the second reviewer). We agree that the third experiment is somewhat out of the scope of this paper (although experiment three is related to the parameter/process uncertainty; will be mentioned in the following). We removed Experiment-3. We have reworked the structure so that the vector-based implementation of land models comes earlier as a separated section (Section 2). We hope the current structure is easier to read and follow.

One of the key questions in defining the spatial discretization of models is of course the calibration. Whereas the GRU's conceptually might make sense compared to grid cells, it introduces new questions on how to calibrate the parameters, and this is not well explained in the text. Does each GRU receive its own set of parameters? And is this then related in any way to the underlying data? As the authors rightly suggest, parameter ranges can be adapted based on soil type or land use, but it seems this was not done by the authors. Not surprisingly, the results demonstrate some of the already known flaws from calibrating on discharge outlet; the everlasting problem of equifinality and overparameterization. If the authors believe the GRU concept is valuable to reintroduce (and I can see it has potential), this value should be demonstrated in a more sophisticated calibration. If the same calibration is done as for usual grid-models, of course we know we can achieve good model performance because there are enough buttons to push, but what do we learn from it compared to a grid-based model and what does it add? 1000 evaluations in the calibration procedure seems rather limited given the dimensions of the problem; this is understandable from a computational point of view, but also a chance to demonstrate why GRU's make more sense than grids within these bounds, by making use of the opportunities that GRU's offer in comparison to grids.

We thank the reviewer for this comment. We fully agree with the reviewer's comments on the parameters' values and estimation. Everything boils down to how computational units are parameterized (which can be very well based on the GRU concept or other techniques). We tried to briefly explain this in the VIC parameter specification in Section 3.3.1. Just a brief explanation here:

1- The soil layers get the same set of parameters for bulk density, saturated hydraulic conductivity (no difference between the vertical soil layers). This is added to Section 3.3.1.
2- The conceptual soil parameters such as b_{inf} are unified across the scale (similar to most of VIC application).
3- The soil depths that conceptually define the storage of the system are defined based on land cover. The forested areas have deeper soil (or root zone) to allow for larger storage and transpiration. Just to mention that this is an advantage of the vector-based implementation and an illustration of more sophisticated calibration strategy. If more hydrological knowledge at the scale of interest is available that can be translated into the constraint.
4- K_{slow} is similar for the entire system (or a gauges) as it can be inferred/calibrated only from the recession analysis.

Of course, more intuitive and sophisticated parameter allocation can be explored but the above-mentioned parameter selection is aligned with what is often done for calibrating the VIC model. This is purposefully not to make the regionalisation so complex that the manuscript deviates from its own message (which is vector-based implementation and accuracy-efficiency trade-off implementation).

I know my colleagues who work with MESH model sometime do this distinction between parameter values of various GRUs in their applications/scientific explorations, for example, different soil with different land cover have different parameters. I personally do not move to that direction for few reasons: (1) the parameters of the spatially largest GRU will be the most sensitive ones when calibrating against the observed streamflow (or polishing of smaller GRUs that have very small contribution, <1%, may be needed), (2) expansion of parameter for calibration that we don't know how to tied actually will unnecessarily add to the dimension of the problem (no information tangible construct them). There is an ongoing effort to relate the parameters to physical characteristics but each of the decisions in itself is an assumption and cannot be inferred directly from the data (for example Mizukami et al., 2017 Table-3). We totally agree with the reviewer on "this value should be demonstrated in a more sophisticated calibration" but at the same time we have not much data for the sophisticated calibration especially the entire subsurface flow movement.
We should also emphasize that part of the motivation for vector-based configuration is computational efficiency with respect to optimization, for example in the MESH model, the underlying assumption is that grouping units from a parametrization perspective, since we expect them to behave in a physical similar way, allows us to characterize the variability with respect to the forcing and of course the subsequent hydrological response, while maintaining the degrees of freedom for parameter estimation reasonable. Our accuracy-performance trade-off is aligned with this mentality also.

Reflecting on reviewer comment, there might be two benefits of allocating the computational units parameter values:

1- Technical aspect; which is the ease of parameter allocation to a computational unit (as each computational unit has a specific land cover and soil type), or better implementation of regionalization rules if applicable. Easier coupling with vector-based routing.
2- The scientific values of implementing the models in a vector-based fashion. That is a grand challenge and an ongoing development. For example, how to effectively parameterize the model simulation at computational units. I personally think part of the reason that the GRU parameter allocation was overlooked or not implemented widely for land models, as the reviewer mentioned with sophisticated calibration, is the lack of data and proper understanding of how parameters behave at the scale of modeling. One of the scientific applications we had here is the trade-off between accuracy of spatial representation and model performance.

Perhaps, I agree that the current manuscript has more emphasis on the first than the latter but at the same time the vector-based implementation can be a vehicle to explore the more scientific questions. For sake of simplicity and to emphasize on the advantages of the vector-based configuration, we tried to put less emphasis on the sophisticated parameterization. As computational hydrology team at University of Saskatchewan we are moving to face this grand challenge in future.

The modeling set up, calibration and parameter perturbation for calibration are based on the objective of the modeling. Our objective here is accuracy-efficiency trade-off for land model and hydrograph simulation.

Reflecting on the computational costs of the setups:

1- 1000 simulations are selected as an arbitrary value based on the computational infrastructure (and time) we had available for the Case4-4km which has 6000 computational units (the most computationally expensive setup).
2- For context, 6000 computational units will be equivalent of approximately 1/3 of the CONUS domain with standard gridded simulation at 0.25 by 0.25 degree lat/lon.
3- Consequently, running 1000 simulations for the Case4-4km takes 5 days on 50 CPUs of ComputeCanada infrastructure (assuming nothing goes wrong with the job), or approximately 8 months on a single CPU.
4- We have tried the impact of more calibration runs for simpler cases (up to couple of thousands of simulations) but did not find a noticeable increase in NS scores.
5- Also, our choice for hourly forcing data increases the storage space needed by almost a factor of 40 compared to daily inputs that is used for VIC-4 and earlier.

An example: The results from Figure 4 are criticized in the text as: "The result indicates the two parameters that are often fixed or a priori allocated based on look up tables can exhibit significant uncertainty and non-identifiability". The Brooks-Corey coefficient is from such a high conceptual level that it might be challenging to find good values in lookup tables, but Ksat might be able to be estimated. The lookup tables can then provide an indication for a search range for the parameter and decrease the equifinality issues with these two parameters.

The soil data that need to inform this choice are themselves so uncertain that they may not guarantee more "realistic k_{sat} ranges". Also, sub-resolution heterogeneity is not account for, nor is k_{sat} very relevant if the dominant flowpath is macropores (which might be very well the case given the significant elevation differences in the region of study). Therefore, we didn't try to a priori limit the k_{sat} calibration ranges too much. This uncertainty in k_{sat} also has high implication for the future regionalization that might be built partly on k_{sat}.
We also wanted to indicate that parameters may be more uncertain that what is suggested in look up tables. It is often the case the land modeling community "kill" the potential uncertainty in parameters and processes either by assigning (hardcoded) parameter values from look up tables or using calibration techniques that yield single best solution (Mendoza et al., 2014).

Shortly, I can see why GRU's might have added value in land-surface modeling. How- ever, the re-introduction of this concept in this manuscript might not make a very good case to convince people of this fact, given that calibration is one of the main challenges and the potential for GRU's in this context is not well explored.

We thank the reviewer for his/her constructive suggestions. We try to improve the manuscript flow. Meanwhile, we would like to emphasize that the focus of this manuscript is not to come up with the parameterization scheme for the model but instead to provide an alternative representation of spatial data that can be beneficial for land modeling community. We demonstrated that how an existing model, such as VIC, can easily be implemented in a vector-based framework. Moreover, the vector-based implementation is not tied to any calibration strategy. A modeller may use the vector-based set up for a land model while avoiding any automatic calibration.

Other suggestions:

In section 2.3, it remains unclear why structural changes to the model were made. Some of the most sensitive parameters of the model (Ds, Dm) have been replaced by a linear reservoir coefficient. Furthermore, the description focusses on VIC4 while VIC5 was explored. Why is that?

It is a very good point for discussion.

1- VIC has a baseflow formulation that has 5 (or one can say 6) parameters, 4 for the baseflow formulation and 2 for the physical specification of the depth and porosity of the lower layer. These 6 parameters are impossible to be inferred from the recession analysis of a hydrograph.
2- The common structure of the land models does not allow for recession analysis based on master recession curve. The soil layers act like a cascade of reservoirs. The water movement from the second layer of VIC can already be damped enough that it may not even need a slow baseflow reservoir (Gharari et al., 2019).
3- One simple solution to that is to basically seize the micropore water movement to the baseflow layer, and only allow macropore water movement based on fraction of surface runoff for example (experiment three).
4- The recession coefficient of the Bow River at Banff is in scale of 0.01 day^{-1} or 100 days. This is not similar to what we get from the automatic calibration. K_{slow} is one identifiable and sensitive parameter given the regime of the Bow River but much higher value [this was wrongly stated in our first reply]. Similarly, the D_s and D_{smax} can show sensitivity but are they really identifiable if a simpler one-parameter baseflow does not hydrologically and logically does what is it expected to do [to simulate the baseflow and get close to the recession analysis]. These are the diagnostics checks that should be done even before any sensitivity analysis.

We mentioned VIC-4 to emphasize why VIC was used so widely. We will remove the explanation on VIC-4. We simplified the VIC description and focused only on VIC version 5.

It is not explained how the parameters in Table 2 were selected for calibration. It is for instance remarkable that no snow parameters, such as snow roughness, are included in the calibration – is this because GRU's focus on soil and land use? Furthermore, it is not clarified to which soil layer Eexp and Ksat refer, or is this kept constant over both soil layers?

We have indeed chosen to only calibrate those parameters that relate to aspects of computational units' configuration (and more importantly forcing resolution for accuracy-efficiency trade-off). We also set the K_{sat} and E_{exp} similar for both layers (clarified in the Section 3.3.1).

Additionally, and based on the reviewer's suggestion we included the snow roughness parameter in our calibration.

Minor for tables and figures:

Table 1 the unit of forcing resolution is missing (degree) Figure 3 the a,b,c labels are missing, the legend is not readable. Figure 4 not sure if this is very informative. More interesting to see a boxplot of every parameter to demonstrate the wide range. Figure 5 Caption says "deviation" but you demonstrate NSE compared to benchmark run, and not the deviation in NSE.

We fixed the forcing resolution in original Table-1 (which is now Table-2).

We fixed original Figure-3 (which is now Figure-2).

We replaced Figure-4 with a normalized illustration of the parameter values above E_{NS} of 0.7. We added extra interpretation of this Figure in the Results and Discussion Sections.

Please not that this is deviation from synthetic case and not standard deviation. We rephrased the caption.

We thank the reviewer for the constructive comments, and we hope to enrich our manuscript by addressing the reviewer's comment sufficiently and successfully.

With kind regards,

Shervan Gharari, on behalf of the co-authors

References:

Gharari, S., Clark, M., Mizukami, N., Wong, J.S., Pietroniro, A. and Wheater, H., 2019. Improving the representation of subsurface water movement in land models. *Journal of Hydrometeorology*, (2019).

Mendoza, P.A., Clark, M.P., Barlage, M., Rajagopalan, B., Samaniego, L., Abramowitz, G. and Gupta, H., 2015. Are we unnecessarily constraining the agility of complex process-based models?. *Water Resources Research*, *51*(1), pp.716-728.

Mizukami, N., Clark, M.P., Newman, A.J., Wood, A.W., Gutmann, E.D., Nijssen, B., Rakovec, O. and Samaniego, L., 2017. Towards seamless large-domain parameter estimation for hydrologic models. *Water Resources Research*, *53*(9), pp.8020-8040.

Answer to the comments by anonymous reviewer#2

We thank the reviewer for their comments on our manuscript. For convenience, the reviewer's comments are given in black and our response is in blue.

The paper titled "Flexible vector-based spatial configurations in land models" uses a new spatial configuration approach with the VIC model that is based on the group response unit concept. The main goals in the paper are to first introduce a method to defining heterogeneity in VIC and then to assess the added value of multiple spatial configurations over the Bow River basin at Banff. The paper is a novel contribution and will be an excellent addition to the land surface/hydrologic modeling community. However, there are multiple issues that I describe below that should be addressed before publication.

The abstract is too long. I strongly suggest reducing the size of the abstract by 20%- 40%.

We have reduced the length of the abstract to ~75% of the original abstracts.

I don't really understand the difference between GRU and HRU in this study; from my understanding a GRU is composed of many HRUs. For example, a sub-basin (GRU) will have multiple HRUs. But based on what is being done here, these GRUs are just the classic GIS partitioning happening and thus very similar to the original definition of a HRU. Maybe I am misunderstanding something. In any case, please clarify the use of the GRU term here.

We thank the reviewer for their comment. The reviewer is correct on the original concept of GRUs and HRUs. As the reviewer correctly points out, GRUs and HRUs typically define a hierarchal spatial organization where HRUs are nested within GRUs (e.g., Clark et al., 2015). For example, in land models, a GRU could be a model grid box and HRUs could be the vegetation tiles within a model grid box; in hydrology models, a GRU could be a sub-basin and HRUs the hydrologically similar areas within a sub-basin. The forcing data could be either constant across the HRUs or distributed to each HRU (distributed forcing within GRUs is important in cases where there are strong climate gradients within GRUs, e.g., due to large elevation range). GRUs are also used to describe classifications of the landscape across the modelling domain (Kouwen et al., 1993, Pietroniro et al., 2007).

We agree with the reviewer that using the vector-based implementation there is little difference between the concept of GRU and HRU. To avoid confusion, we abandon the concept of GRU entirely and present the model as a vector-based implementation (that can benefit from the concepts of GRU and HRU). We only use the concept of GRU as a matter of parameter allocation in Section 3.3.1 and in the discussion to reflect on the validity of the assumption on parameter allocation.

Figures 1 and 2 - These two figures are not very informativeâATespecially Figure 1. I would remove them. Maybe some of these ideas could be merged into an improved (or split) Figure 3.

We have revised Figure 1 and made it more conceptual in the newer version. We have removed figure 2 to the discussion as an example the unrealistic combination of land cover and elevation zone in the VIC model for discussion purposes.

Figure 3 - You should have a, b, c, d coded on the panels themselves. Panel c is very unclear. I think this whole figure is critical to understand the implementation and thus should be improved (and perhaps split into two).

A, b, c, d are added to the panels. We have clarified the c as well to have a more readable legend. We believe the changes we made in Figure 1 now make the concept easier to understand. Therefore Figure 3 (now figure 2) is more directed to illustrates the basin of interest characteristics.

Figure 6 - The 3 dimensionality of this figure is unnecessary and frankly confusing. The 2d surface is more than enough to get the point across.

We have made it 2 D figure and merge it with original Figure 8.

Figure 8 - Again, the 2d surface would be much better here.

We have made it 2 D figure and merge it with original Figure 6.

Line 122 - Only using tmin, tmax, precipitation, and wind speed is only one option in the earlier VIC versions. One could also still use longwave in, shortwave in, among others.

We have revised the VIC description and now the description should be much clearer.

Line 153 - Although I am certainly a fan of “killing the grid”, it is not entirely true that "resolution loses its meaning" with the introduced approach. You still have an effective spatial resolution which is governed by the level of details that needs to exist in your polygons. Of course the advantage here is that you can have the size of those polygons vary as a function of space; however, you will still have the concept of an effective spatial resolution present. I’d suggest thinking more carefully of what moving to a polygon based approach really means and how it can be "upscaled" in more informative ways than the classic coarsening of the grid.

An excellent point raised by the reviewer. The reviewer is certainly correct that we still have the same upscaling challenges in vector-based implementations. We discussed this issue in more detail in the revised paper (Section 2).

One of the ideas behind the vector-based modeling is the flexibility of the input data. For example, for a larger basin, the forcing can be set to a higher resolution for the mountainous headwater. For our test case (limited spatial domain), this concept can be explored in detail, and our current work focuses on resampling and coarsening of the forcing grids. We will rework this section to include more discussion on the importance of “polygons” vs “grid” and the implications for large/continental scale modeling.

* Check for typos; there appear to be a few throughout the text (e.g., VIV-GRU on line 182)

We tried to remove the typos as much as possible in the manuscript.

We thank the reviewer for the constructive comments, and we hope to enrich our manuscript by addressing the reviewer's comment sufficiently and successfully.

With kind regards,

Shervan Gharari, on behalf of the co-authors

References:

Clark, M.P., Nijssen, B., Lundquist, J.D., Kavetski, D., Rupp, D.E., Woods, R.A., Freer, J.E., Gutmann, E.D., Wood, A.W., Brekke, L.D. and Arnold, J.R.: A unified approach for process-based hydrologic modeling: 1. Modeling concept. *Water Resources Research*, *51*(4), pp.2498-2514, 2015.

Pietroniro, A., Fortin, V., Kouwen, N., Neal, C., Turcotte, R., Davison, B., Verseghy, D., Soulis, E. D., Caldwell, R., Evora, N., and Pellerin, P.: Development of the MESH modelling system for hydrological ensemble forecasting of the Laurentian Great Lakes at the regional scale, Hydrol. Earth Syst. Sci., 11, 1279–1294, https://doi.org/10.5194/hess-11-1279-2007, 2007.

Kouwen, N., Soulis, E.D., Pietroniro, A., Donald, J. and Harrington, R.A.: Grouped response units for distributed hydrologic modeling. *Journal of Water Resources Planning and Management*, *119*(3), pp.289-305, 1993.

Flexible vector-based spatial configurations in land models

Shervan Gharari[1,*], Martyn P. Clark[1], Naoki Mizukami[2], Wouter J. M. Knoben[1], Jefferson S. Wong[3], Alain Pietroniro[4]

1- University of Saskatchewan Coldwater Laboratory, Canmore, Alberta, Canada.
2- National Center for Atmospheric Research, Boulder, Colorado, USA.
3- Global Institute for Water Security (GIWS), Saskatoon, Saskatchewan, Canada.
4- Environment and Climate Change Canada (ECCC), Saskatoon, Saskatchewan, Canada.
*Corresponding author Shervan Gharari, shervan.gharari@usask.ca

Abstract. Land models are increasingly used in terrestrial hydrology due to their process-oriented representation of water and energy fluxes. A priori specification of the grid size of the land models is typically defined based on the spatial resolution of forcing data, the modeling objectives, the available geo-spatial information, and computational resources. The variability of the inputs, soil types, vegetation covers, and forcing are masked or aggregated based on the *a priori* grid size. In this study, we propose an alternative vector-based implementation to directly configure a land model using unique combinations of land cover types, soil types, and other desired geographical features that has hydrological significance, such as elevation zone, slope, and aspect. The main contributions of this paper are to (1) implement the vector-based spatial configuration using the Variable Infiltration Capacity (VIC) model; (2) illustrate how the spatial configuration of the model affects simulations of basin-average quantities (i.e., streamflow) as well as the spatial variability of internal processes (SWE and ET); and (3) describe the work/challenges ahead to improve the spatial structure of land models. Our results show that a model configuration with a lower number of computational units, once calibrated, may have similar accuracy to model configurations with more computational units. However, the different calibrated parameter sets produce a range of, sometimes contradicting, internal states and fluxes. To better address the shortcomings of the current generation of land models, we encourage the land model community to adopt flexible spatial configurations to improve model representations of fluxes and states at the scale of interest.

Deleted:

Formatted: Font: +Headings (Calibri Light), 28 pt, Not Bold, Font color: Auto

Deleted: ¶

Deleted: ¶

Deleted: Land models can be set up at a range of spatial configurations, often ranging from grid sizes of 0.02 to 2 degrees (approximately 2 to 200 km) and applied at sub-daily temporal resolutions for simulation of energy fluxes.

Deleted: is derived from

Deleted: resolutions,

Deleted: data

Deleted: Typically, the choice of model configuration and grid size is based on modeling convenience and is rarely examined for adequate physical representation in the context of modeling.

Deleted: and parameters, forcings

Deleted: and

Formatted: Font: Italic

Deleted: chosen

Deleted: set up

Deleted: based on the concept of Group Response Unit ... [1]

Deleted: combination

Deleted: type

Deleted: Computational units

Deleted: defined as GRUs that are forced at a specific forcing ... [2]

Formatted: Font color: Auto, English (CAN)

Deleted: data and forcings. We set up

Formatted: Font color: Auto, English (CAN)

Deleted: , based on the GRU concept (VIC-GRU). Utilizing ... [3]

Formatted: Font color: Auto, English (CAN)

Deleted: following questions: (1) how well a model

Formatted: Font color: Auto, English (CAN)

Deleted: simulates an output variable, such as streamflow, for ... [4]

Formatted: Font color: Auto, English (CAN)

Deleted: structure and parameters are for the land model.

Formatted: Font color: Auto, English (CAN)

Deleted: , although case dependent,

Deleted: the models may similarly reproduce output with

Deleted: in the context of modeling (streamflow for ... [5]

Deleted: identification and model diagnostic suites, such as ... [6]

Deleted: approaches that will provide a better understanding

Deleted: accuracy-performance tradeoff in land models

1 Introduction

Land models have evolved considerably over the past few decades. Initially, land models (or land-surface models) were developed to provide the lower boundary conditions for atmospheric models (Manabe, 1969). Since then land models have increased in complexity, and they now include a variety of hydrological, biogeophysical, and biogeochemical processes (Pitman, 2003). Including this broad suite of terrestrial processes makes land models enables simulations of energy and water fluxes and carbon and nitrogen cycles.

Deleted: suitable to simulate

Despite the recent advancements in process representation in land models, there is currently limited understanding of the appropriate spatial complexity that is justified based on the available data and the purpose of the modelling exercise (Hrachowitz and Clark, 2017). The increase of computational power, along with the existence of more accurate digital elevation models and land cover maps, encourage modelers to configure their models at the finest spatial resolution possible. Such hyper-resolution implementation of land models (Wood et al., 2011) can provide detailed simulations at spatial scales as small as 1-km^2 grid over large geographical domains (e.g., Maxwell et al., 2015). However, the computational expense for hyper-resolution models could potentially be reduced using more creative spatial discretization strategies (Clark et al., 2017).

Deleted: modellers

It is common to adopt concepts of hydrological similarity to reduce computational costs. In this approach, spatial units are defined based on similarity in geospatial data, under the assumption that processes, and therefore parameters, are similar for areas within a spatial unit (e.g., Vivoni et al., 2004, Newman et al., 2014). Hydrological Response Units (HRUs) are perhaps the most well-known technique to group geospatial attributes in hydrological models. HRUs can be built based on various geospatial characteristics; for example, Kirkby and Weyman 1974, Knudsen and Refsgaard (1986), Flügel (1995), Winter (2001), and Savenije (2010) all have proposed to use geospatial indices to discretize a catchment into hydrological units with distinct hydrological behaviour. HRUs can be built based on soil type such as proposed by Kim and van de Giessen (2004). HRUs can also be built based on fieldwork and expert knowledge (Naef et al., 2002, Uhlenbrook 2001), although the spatial domain of such classification will be limited to the catchment of interest and the spatial extent of the field measurements. HRUs are often constructed by GIS-based overlaying of various maps of different characteristics and can have various shapes

Deleted:).

such as for non-regular (sub-basins), grid, hexagon, or triangulated irregular network also known as TIN (Beven 2001, Marsh et al., 2012, Oliviera et al., 2006). Land models are also beginning to adopt concepts of hydrological similarity (e.g., Newman et al., 2014; Chaney et al., 2018). Traditionally land models use the tiling scheme where a grid box is subdivided into several tiles of unique land cover, each described as a percentage of the grid (Koster and Suarez, 1992). Similarly, the concept of Grouped Response Units (GRUs, Kouwen et al., 1993), assumes similar hydrological property for areas with identical soil, vegetation, and topography. The GRU concept is utilized in the MESH land modeling framework (Pietroniro et al., 2007).

Moved (insertion) [1]

Deleted: , Pietroniro et al., 2007). Similar approaches are used in land models.

Moved up [1]: Land models are also beginning to adopt concepts of hydrological similarity (e.g., Newman et al., 2014; Chaney et al., 2018).

A long-standing challenge is understanding the impact of grid size on model simulations (Wood et al., 1988). The effect of model grid size can have a significant impact on model simulation across scale especially if the model parameters are linked to characteristics which are averaged out across scale (Bloschl et al., 1995). Shrestha et al. (2015) have investigated the performance of Community Land Model (CLM) v4.0 coupled with ParFlow across various grid sizes. They concluded the grid size changes of more than 100 meters can significantly affect the sensible heat and latent heat fluxes as well as soil moisture. Also using CLM, Singh et al. (2015) demonstrated that topography has a substantial impact on model simulations at the hillslope scale (~100 meters), as aggregating the topographical data changes the runoff generation mechanisms. This is understandable as the CLM is based on topographical wetness index (Beven and Kirkby 1979, Niu et al., 2005). However, Melsen et al. (2016) evaluated the transferability of parameters sets across the temporal and spatial resolutions for the Variable Infiltration Capacity (VIC) model implemented in an Alpine region. They concluded that parameter sets are more transferable across various grid sizes in comparison with parameter transferability across different temporal resolutions. Haddeland et al. (2002) showed that the transpiration from the VIC model highly depends on grid resolution. It remains debatable how model parameters and performance can vary across various grid resolutions (Liang et al., 2004; Troy et al., 2008; Samaniego et al., 2017).

The representation of spatial heterogeneity is an ongoing debate in the land modelling community (Clark et al., 2015). The key issue is to define which processes are represented explicitly and which processes are parameterized. The effect of spatial scale on emergent behavior has been studied for catchment scale models – the concepts of Representative Elementary Areas (REA), or Representative Elementary Watersheds (REW), were introduced to study the effect of spatial

Deleted: behaviour

aggregation on system-scale emergent behavior (Wood et al., 1995, Reggiani et al., 1999). The effect of scale on model simulations is not well explored for land models. More work is needed to understand the extent to which the heterogeneity of process representations is sufficient for the purpose of a given modelling application, and the extent to which the existing data can support the model configurations (Wood et al., 2011, Beven et al., 2015) and guarantee a *fidelius* model.

In this study, we configure the Variable Infiltration Capacity (VIC) model in a flexible vector-based framework to understand how model simulations depend on the spatial configuration. The remainder of this paper is organized as follows: In Section 2, we present the concept of vector-based configuration for land models. In Section 3 we describe the study area and the data sets used in this study as well as the design of the experiments, and elaborate the Variable Infiltration Capacity model (VIC) and mizuRoute as the vector-based routing model. In Section 4 we describe the results of the experiments. Section 5 discusses the implication of spatial discretization strategies on large-scale land model applications. The paper ends in Section 6 with conclusions of this study and implications for future work.

2 The vector-based configuration for land models

Land models are often applied at a regularly spaced grid. Land models are typically set up at a range of spatial configurations, ranging from grid sizes of 0.02° to 2° (approximately 2 to 200 km) and applied at sub-daily temporal resolutions for simulation of energy fluxes. A priori specification of the grid size of the land models is often derived from forcing resolutions, modeling objectives, available geo-spatial data and computational resources and is usually based on modeling convenience. Figure-1e-h illustrates the typical land model configuration – here the modeler selects a cell size, and then the soil, vegetation and forcing files are all aggregated or disaggregated to the target cell size. Original data resolution and spatial distribution of soil, land cover and forcing data are smeared while upscaled to the resolution of interest. Any change in the modeling resolution will require upscaling or downscaling of the geo-physical dataset once again.

In this study, we configure the land models using non-regular shapes. Figure-1a-d presents an example of non-regular shapes created through spatial intersections of the land covers and soil types shapes. These vector-based configuration of the geospatial data are then forced at the original meteorological forcing resolution, or its upscaled or downscaled values resulting in computational

Deleted: behaviour

Deleted:).

Deleted: In addition to the choice on model's spatial configurations, more work is needed to define the appropriate structure of land models. While many studies in hydrology have evaluated how model structure affects the smaller scale watershed response (Son and Sivapalan 2007, Clark et al., 2008, Fenicia et al., 2011, Shafii et al., 2017), this issue has received limited attention in the land modelling community (Desborough, 1999). Only recently, a few land models enable changing the process formulations within a limited range of model structural assumptions (Noah-MP, Niu et al., 2011, SUMMA, Clark et al., 2015) We explore effects of different choices of runoff generation process representation in the model. ¶

Deleted: VIC model, its

Deleted: implementation, and its coupling to the mizuRoute routing model.

Deleted: .

units. Therefore, each computational unit has unique geospatial data such as soil, vegetation, slope and aspect and is forced with a unique forcing. In this configuration changing of meteorological forcing resolution do not affect the decisions needed to upscale the geo-spatial data such as soil type and land cover to the grid resolution.

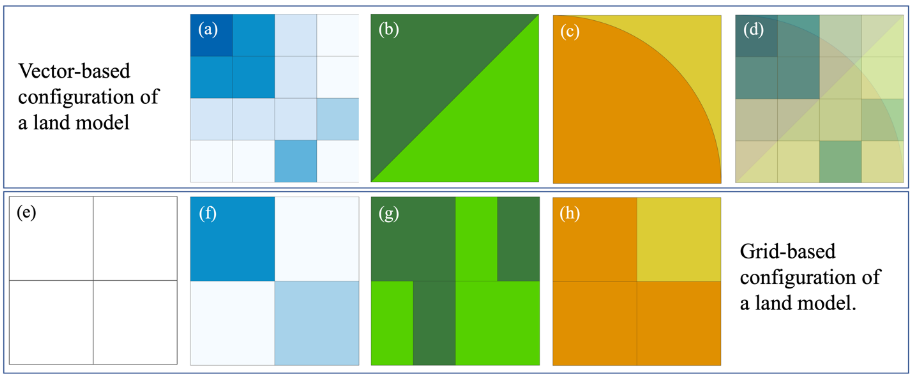


Figure-1- Top row indicates vector-based configuration of a land model; (a) meteorological forcing at its original resolution or upscaled and downscaled resolutions, (b) land covers, (c) soil types with their spatial extent, and (d) vector-based configuration with 28 computational units each with unique forcing, soil type and land cover type. The bottom row indicates typical grid-based configuration of a land model; (e) a priori resolution should be decided, (f) meteorological forcing should be upscaled or downscaled to the grid resolution, (g) land cover percentage should be calculated for each modeling grid; or a dominate landcover should be selected to represent that grid, and finally (h) soil characteristics for each modeling grid should be identified.

The benefits of vector-based configuration of land models can be summarized as follows:

1- No need for a priori assumption on modeling grid size. In traditional land model implementation, the modeler selects a grid resolution (which is often a regular latitude/longitude grid). The soil parameters and forcing data from any resolution must be aggregated, disaggregated, resampled or interpolated for every grid size. The land cover data often is only considered as a percentage for every grid and spatial location of the land cover is lost. However, in the vector-based setup these decisions are only based on the input and forcing data that are chosen to be used

Moved (insertion) [2]

in the modeling practice and no upscaling or downscaling to grid size is needed. Furthermore, the size of computational units can vary across modeling domain depending on the variability of the meteorological forcing and geospatial heterogeneity. For example, the spatial density of computational units can be higher in mountainous areas where temperature and precipitation gradients are larger while avoiding unnecessarily high number computational units in areas with lower gradient in meteorological forcing.

2- Reasonable relation between available meteorological forcing and geo-spatial data resolutions and number of computational units: computational units that are the result of available geophysical data sets forced with the original forcing data logically represent the maximum number of computational units that can be hydrologically unique. A higher number of computational units than the proposed setup will arguably provide an unnecessary computational burden due to identical forcing data and geospatial information.

3- Direct simplification of geospatial data. The vector-based implementation facilitates easier aggregation of computational units. It is easier to aggregate similar soil types or similar forested areas into a unified shapes with basic GIS function (dissolving for example) than this would be if all data had to be upscaled or downscaled into a different grid size.

4- Direct specification of physical parameters and avoiding unrealistic combinations of land cover, soil and other geo-physical information. As each computational unit has a specific type of land cover, soil type and other physical characteristics, it is straightforward to specify parameter values based on look up tables (i.e., no averaging, upscaling is needed). This is favorable because the modeler does not need to make decisions about methods used for upscaling of geophysical data at the grid level. Also, this might avoid the unrealistic combination of parameter sets that might be considered by the model at a grid scale, such as equiprobable combination of land cover on soil type which may not exist in reality which will be increasing the fidelity of the model representation of the processes (we will elaborate this further in the context of the VIC model in Discussion Section).

5- The ability to compare and constrain the parameter values for computational units and their simulations. The impact of land cover, soil type and elevation zone can be evaluated separately. For example, the vector-based implementation makes it easier to test if forested areas generate less surface runoff than grasslands. This might be more challenging at the gird-based configuration in which there are combination of different land cover types at grid scale. Similarly, the vector-based implementation may simplify regularization efforts across large geographical domains. This relative constrains can be utilized to translate often patchy expert knowledge into a sophisticated land model so that the model simulation will obey the modelers and hydrologists' expectations.

6- The possibility to incorporate additional data. If needed, additional data, such as slope and aspect for example, can be incorporated in building the computational units, accounting for changes in shortwave radiation or lapse rates for temperature. The changes can be implemented outside of the model in the forcing files. Computational units can be built also based on variation of leaf area index (LAI) giving an additional layer of information in addition to the land cover type. The additional information can be easily ingested into the model without extra effort in contrast to changing of the model parameter files at the grid scale.

7- Easier comparison of model simulations and in situ point-scale observation and visualization: The vector-based implementation of land models makes it easier to compare the point measurement to model simulation as the model simulations preserve extent of geospatial features.

8- Modular and controlled selection of models: The vector-based implementation identifies the characteristics and spatial boundary of geospatial domains. A model might not be suitable for processes of some of the geospatial domains. Alternatively, processes of a computational unit that is beyond the capacity of one model can be replaced with an alternative model. For example, computational units that are glaciered, can be replaced with more suitable models while the spatial configuration and forcings remain identical. Consequently, the effect of features such as glaciers can be better studied as more expert models can be applied to glacier while the rest of the computational units can be simulated with a model that includes general processes.

Moved (insertion) [6]

3 Data and methods

Moved (insertion) [7]

3.1 Study area:

Experiments are performed for the Bow River at Banff with a basin area of approximately 2210 km^2 located in province of Alberta, Canada. The Bow River is located in the Canadian Rockies in the headwaters of the Saskatchewan River Basin. Most of the Bow River streamflow is due to snow melt (Nivo-glacial regime). The average basin elevation is 2130 m ranging from 3420 m at the peak top to 1380 m above mean sea level at the outlet (town of Banff). The basin annual precipitation is approximately 1000 mm with range of 500 mm for the Bow Valley up to 2000 mm for the mountain peaks. The predominant land cover is conifer forest in the Bow Valley and rocks and gravels for mountain peaks above the tree line.

Moved (insertion) [8]

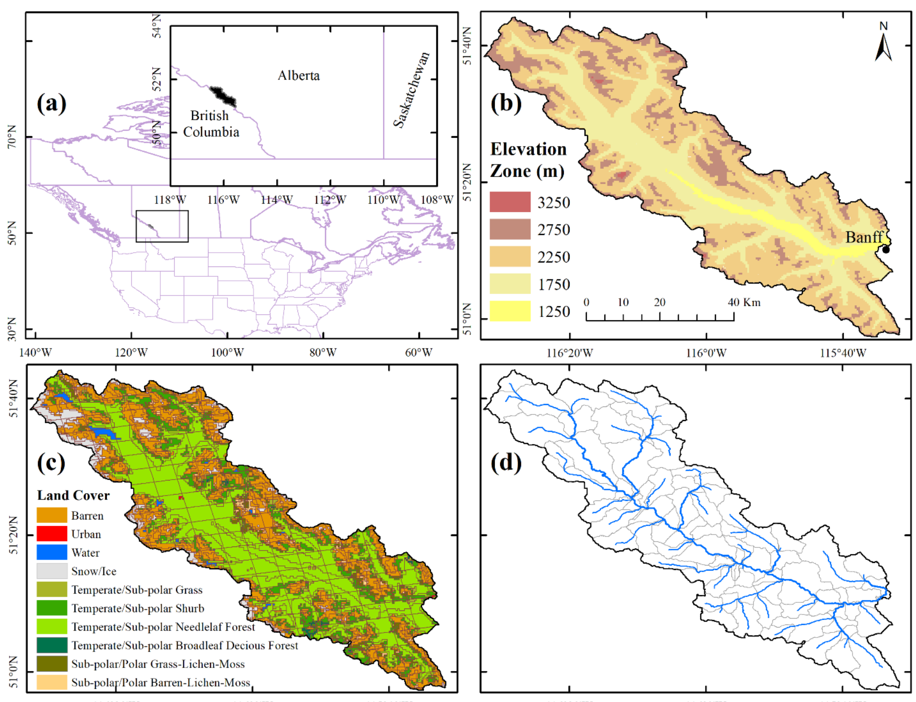


Figure – 2 (a) The location of the Bow River Basin at Banff (b) Bow River Basin elevation, (c) computational units for geospatial data of elevation zones, land cover and soil type forced at WRF original resolution at 4 km (Case-3-4km) and (d) river network topology and associated sub-basins that is used for the vector-based routing.

3.2 Geospatial data and meteorological forcing

Moved (insertion) [9]

3.2.1 Model input dataset and forcing:

The inputs and forcing we used to set up the model are as follows:

1- Land cover: We used the land cover map NALCMS-2005 v2 (North American Land Change Monitoring System, Latifovic et al., 2004) that is produced by CEC (Commission for Environmental Cooperation). NALCMS-2005 v2 includes 19 different classes. The land cover map is used to set up the vegetation file and vegetation library (look up table) for the VIC model (Nijssen et al., 2001).

Moved (insertion) [10]

2- Soil texture: We used the Harmonized World Soil Data, HWSD (Fischer et al., 2008). For each polygon of the world harmonized soil we use the highest proportion of soil type. The HWSD provide the information for two soil layers, in this study we base our analyses on the lower soil layer reported in HWSD to define the soil characteristics needed for the VIC soil file.

3- Digital Elevation Model: in this study we make use of existing hydrologically conditioned digital elevation models (DEM) to (1) derive the river network topology for the vector-based routing, mizuRoute and (2) to derive the slope, aspect and elevation zones which are used to estimate the forcing variables. For the first purpose we use hydrologically conditioned DEM of HydroSHED (Lehner et al., 2006) with resolution of 3 arc-second, approximately 90 meters; for the second purpose we use HydroSHED 15 arc-second DEM (approximately 500 meters).

4- Meteorological forcing: we used the weather research and forecasting (WRF) model simulation for continental United States with the temporal resolution of 1 hour and spatial resolution of 4 km (Rasmussen and Liu, 2017). For upscaling the WRF input forcing, we use the CANDEX package (DOI: 10.5281/zenodo.2628351) to map the 7 forcing variables to various

Moved (insertion) [11]

resolutions (1/16˚, 1/8˚, 1/4˚, 1/2˚, 1˚ and 2˚ from the original resolution of 4 km). We used the required variables from the WRF data set namely, total precipitation, temperature, short and long wave radiation at the ground surface, V, U components of wind speed and water vapor mixing ratio.

The shortwave radiation is rescaled based on the slope and aspect of the respective computational unit (refer to Appendix-A for more details). In this study we differentiated four aspects and five slope classes. The temperature at 2 meters are adjusted using the environmental lapse rate of -6.5°C for 1000 meters increase in elevation. The assumed lapse rate aligns with earlier findings from the region of study (Pigeon and Jiskoot, 2008).

3.2.2 Observed data for model calibration

The daily streamflow is extracted from the HYDAT (WSC, Water Survey Canada) for Bow at Banff with gauges ID of 05BB001. This data is used for parameter calibration/identification of the VIC parameters.

3.3 Land model and routing scheme:

3.3.1 The Variable Infiltration Capacity (VIC) model:

The VIC model was developed as a simple land surface/hydrological model (Liang et al. 1994) that has received applications worldwide (Melsen et al., 2016). In this study we use classic VIC version 5. The VIC model combines sub-grid probability distributions to simulate surface hydrology such as variable infiltration capacity formulation (Zhao, 1982). The VIC model uses three soil layers to represent the subsurface. While each soil layer can have various physical soil parameters (e.g., saturated hydraulic conductivity, bulk density), each layer is assumed to be uniform across the entire grid regardless of the vegetation type variability in that grid. The VIC model assumes a tile vegetation implementation within each grid similar to the mosaic approach of Koster and Suarez (1992) with bio-physical formulations for transpiration (Jarvis et al., 1976). To account for spatial variability in vegetation, the VIC model allows for root depths to be adjusted for every vegetation type. The vegetation parameters (e.g., stomatal resistance, LAI, albedo) are often identical for each land cover across the modeling domain. The VIC model can account for different elevation zones to account for temperature lapse rate given elevation difference in a grid cell, and also for the distribution of precipitation over the identified elevation zones.

Moved (insertion) [12]

Moved (insertion) [13]

Formatted: Font color: Accent 1

Formatted: Heading 3, Space After: 0 pt, Line spacing: single

Deleted: Land model and the routing model¶

Formatted: Font color: Accent 1

Formatted: Font color: Accent 1

Formatted: Heading 3, Space After: 0 pt, Line spacing: single

Deleted: (VIC-5, Hamman et al., 2018). The key features of VIC are: (i) traditionally, the VIC model (version 4 and earlier) simulates sub-daily energy variables with daily forcing of minimum and maximum temperature, precipitation and wind speed. This enables the VIC model to be easily forced with hydrological available data sets worldwide while being able to solve the energy fluxes over sub-daily time periods. (ii)

Deleted:) with bio-physical formulations for transpiration (Jarvis et al., 1976).¶

Deleted:).

Deleted: fixed

Deleted: every

Deleted: various

In the experiments for this study, we calibrate a subset of VIC parameters namely b_{inf}, E_{exp}, K_{sat}, $d_{2,forested}$, $d_{2,non\text{-}forested}$, K_{slow}, and $S_{roughness}$ (names are mentioned in Table-1). Following the concept of GRU, Kouwen et al., 1993, we assume the computational units with similar geophysical characteristics (soil and land cover) possess similar parameter values. We make sure that the $d_{2,forested}$ is larger than the $d_{2,non\text{-}forested}$ as the root depth are deeper for forested regions (constraining relative parameters). For the sake of simplicity, we limit the root zone to the upper soil layers and replace the 5-parameter VIC baseflow[1] with a linear reservoir (refer to Gharari et al., 2019 for further explanation). We also assume that the two top soil layers possess homogeneous soil characteristics.

3.3.2 mizuRoute, a vector-based routing scheme

In this study, we make use of the vector-based routing model mizuRoute (Mizukami et al., 2016). Vector-based routing models can be configured for different computational units than the land model uses (e.g., configuring routing models using sub-basins derived from existing hydrologically conditioned DEMs such as HydroSHEDS, Lehner et al., 2006, or MERIT Hydro, Yamazaki et al., 2019). This removes the dependency of the routing on the grid size or computational unit configurations and eliminates the decisions that are often made to represent routing-related parameters at grid scale. Therefore, we can ensure that two model configurations with different geospatial configurations are routed using the same routing configuration. The intersection between the computational units in the land model and the sub-basins in the routing model defines the contribution of each computational units from the land model to each river segment.

The Impulse Response Function (IRF) routing method (Mizukami et al., 2016) is used for this study. IRF, which is derived based on diffusive wave equation, includes two parameters – wave velocity and diffusivity. The diffusive wave parameters are set to 1 m/s and 1000 m^2/s respectively and remain identical for all the river segments. The river network topology, assuming approximately 25 km^2 starting threshold for the sub-basin size, is based on a 92-segment river network depicted in Figure-3d.

[1] The VIC baseflow parameters are: D_{smax}, maximum rate of baseflow; D_s, fraction of D_{smax} where non-linear baseflow begins; W_s, fraction of maximum soil moisture where non-linear baseflow occurs; c, exponent used for the non-linear part of the baseflow; and depth of the baseflow layer d_3.

Moved up [2]: <#> implementation, the modeler selects a grid resolution (which is often a regular latitude/longitude grid). The soil parameters and forcing data from any resolution must be aggregated, disaggregated, resampled or interpolated for every grid size.

Moved up [3]: <#> A higher number of computational units than the proposed setup will arguably provide an unnecessary computational burden due to identical forcing data and geospatial information.
<#>**Direct simplification of geospatial data.**

Moved up [4]: <#> This is favorable because the modeler does not need to make decisions about methods used for upscaling of geophysical data at the grid level.

Moved up [5]: <#> The impact of land cover, soil type and elevation zone can be evaluated separately.

Moved up [6]: <#> implementation identifies the characteristics and spatial boundary of geospatial domains. A model might not be suitable for processes of some of the geospatial domains.

Deleted: <#>*The VIC-GRU implementation: a vector-based configuration for land models*
<#>The VIC model is typically applied at regular grid. Figure-1a illustrates the typical VIC configuration – here the modeler selects a cell size, and then the soil, vegetation and forcing files are all aggregated or disaggregated to the target cell size. Original data resolution and spatial distribution of soil, land cover and forcing data are lost. In this study, we configure the VIC model using non-regular shapes, Grouped response Units (GRUs, Kouwen et al., 1993), depending on the soil, vegetation, and topography ... [7]

Deleted: <#>The land cover data is only considered as a ... [8]

Deleted: <#>The vector-based implementation makes the ... [9]

Deleted: <#> ... [10]

Deleted: <#>For example, the GRU concept makes it ... [11]

Moved (insertion) [14]

Deleted: <#>Alternatively, processes of a GRU that is ... [12]

Moved down [15]:

Formatted: Heading 3, Space After: 0 pt, Line spacing: single

Deleted: implemented several changes to the VIC process ... [13]

Formatted: Font color: Accent 1

Deleted: separate

Deleted: Hydrosheds

Deleted: Merit

Deleted: to

Deleted: GRUs

Deleted: ,

Deleted: in

Moved (insertion) [16]

Moved (insertion) [17]

Table-1 the VIC model parameters that are subjected to perturbation for model calibration for the designed experiments.

Parameter symbol	Parameter name	Minimum value	Maximum value	Unit	Explanation
b_{inf}	Variable infiltration parameter	0.01	0.50	[-]	
E_{exp}	The slope of water retention curve	3.00	12.00	[-]	
K_{sat}	Saturated hydraulic conductivity	5.00	1000.00	[mm/day]	
d_1	The depth of topsoil layer	0.2	0.2	[m]	Fixed at 20 cm for both forested and non-forested computational units
$d_{2,forested}$	The depth of the second soil layer for forest computational units	0.2	2	[m]	
$d_{2,non\text{-}forested}$	The depth of the second soil layer for non-forested computational units	0.2	$d_{2,forested}$	[m]	The maximum is bounded by the $d_{2,forested}$
D_{root}	The distribution of root in the two soil layers	0.5	0.5	[-]	Fixed at 50% for the top and lower soil layers.
K_{slow}	Slow reservoir coefficient	0.001	0.9	[1/day]	
$S_{roughness}$	Snow roughness	0.5	3	[mm]	

3.4 *Experimental design:*

In this study, we configure the VIC model in a flexible vector-based framework to understand how model simulations depend on the spatial configuration. We consider four different methods to discretize the landscape for seven different spatial forcing grids (see Table 2). The landscape discretization methods include (1) simplified land cover and soils; (2) full detail for land cover and soils; (3) full detail for land cover and soils, including elevation zones; and (4) full detail for land cover and soils, including elevation zones and slope and aspect. The different spatial forcing resolutions are 4-km, 0.0625°, 0.125°, 0.25°, 0.5°, 1°, and 2°. This design enables us to separate

Moved (insertion) [18]

Moved up [7]: <#>**Data and methods**¶

Deleted: 1

Deleted: grids

discretization of the landscape based on geo-spatial data from the spatial resolution of the forcing data.

Table – 2- The numbers of computational units for the Bow River at Banff, given different spatial discretization of land cover, soil type, elevation zones and slope and aspects forced with various forcing resolutions.

	Forcing resolution	**Case 4** 4 aspect groups; 5 slope groups; 19 classes of land cover; 500 meter elevation zones;	**Case 3** no aspect groups; no slope groups; 19 classes of land cover; 500 meter elevation zones;	**Case 2** no aspect groups; no slope groups; 19 classes of land cover; no elevation zones;	**Case 1** no aspect groups; no slope groups, 3 classes of land cover, one dominant soil type no elevation zones;
Number of unique combination of geo-spatial data (soil, land cover, elevation zones, slopes and	--	**582**	**65**	**56**	**3**
Number of Computational units	4 km	6631	1508	941	479
	0.0625° [~6.25 km]	5224	1098	663	290
	0.125° [~12.50 km]	3079	515	283	94
	0.25° [~25.00 km]	2013	306	154	39
	0.5° [~50.00 km]	1332	184	93	21
	1.0° [~100.00 km]	917	116	56	12
	2.0° [~200.00 km]	767	89	42	6

Deleted: our method to discretize

Deleted: Experiments are performed for the Bow River at Banff with a basin area of approximately 2210 km^2.

Moved (insertion) [19]

Formatted: Justified

Formatted Table

Formatted: Centered, Indent: Left: 0.2 cm, Right: 0.2 cm

Moved up [8]: The Bow River is located in the Canadian Rockies in the headwaters of the Saskatchewan River Basin. Most of the Bow River streamflow is due to snow melt (Nivo-glacial regime). The average basin elevation is 2130 m ranging from 3420 m at the peak top to 1380 m above mean sea level at the outlet (town of Banff). The basin annual precipitation is approximately 1000 mm with range of 500 mm for the Bow Valley up to 2000 mm for the mountain peaks. The predominant land cover is conifer forest in the Bow Valley and

Deleted: bare soil and rocks for mountain peaks above the tree line.

We design three experiments:

Formatted: Font: 12 pt

3.4.1 Experiment-1: How does the spatial configuration affect model performance?

As the first experiment, we focus on how well the various configurations simulate observed streamflow for Bow River at Banff. We calibrate the parameters for the different configurations in Table 2. Model calibration is accomplished using the Genetic Algorithm implemented in the OSTRCIH framework (Mattot, 2005; Yoon and Shoemaker, 2001), maximizing the Nash-Sutcliffe Efficiency (E_{NS}, Nash and Sutcliffe 1970) using a total budget of 1000 model evaluations given the available resources limited by the most computationally expensive model (Case-4-4km).

Deleted: at the

Deleted: 1

3.4.2 Experiment-2: How well do calibrated parameter sets transfer across different model configurations?

As the second experiment, we focus on how various configurations can reproduce the result from the configuration with highest computational units for a given parameter set. In other words, this experiment evaluates accuracy-efficiency tradeoffs – i.e., the extent to which spatial simplifications affect model performance under the assumption that similar computational units possess identical parameters across various configurations. This is important as it enables modelers to understand accuracy-efficiency tradeoffs, given the available data and the purpose of the modelling application. This experiment is based on perfect model experiments using the model with the highest computational unit as synthetic case (Case-4-4km). Synthetic streamflow for every river segment is generated using a calibrated parameter set for Case-4-4km. The models with lower number of computational units are then simulated using the exact same parameter set used for generating the synthetic streamflow. The differences in streamflow simulation, quantified using E_{NS}, provide an understanding of how the simulations deteriorate when the spatial and forcing heterogeneities are masked or upscaled. This also will bring an understanding on how sensitive the changes are along the river network and at the gauge location at which the models are calibrated against the observed streamflow data. Similarly, we compare the spatial patterns of snow water equivalent for the different spatial configurations.

Deleted:

Deleted: GRUs

Deleted: -accuracy

Deleted: -4km

Deleted: up-scaled

4 Results

4.1 Experiment-1

The various model configurations are compared with respect to the Nash-Sutcliffe performance metric (E_{NS}). Results show that all the models, including the ones that are configured with coarser resolution forcings, can simulate streamflow with E_{NS} as high as 0.70 (Table-3). It is noteworthy to mention that the configuration of Case-4-1° has higher E_{NS} value compared to the cases with highest computational units, Case-4-4km for example. This might be due to various reasons including: (1) compensation of forcing aggregation on possible forcing bias at finer resolution; (2) compensation of forcing aggregation on model states and fluxes and possible adjustment for model structural inadequacy and hence directing the optimization algorithm to different possible solutions across configurations.

Table-3 – The highest calibrated Nash-Sutcliffe performance metric (E_{NS}) for the different model configurations. Details on the geospatial cases are provided in Table 2.

Forcing resolution	**Case 4** 4 aspect groups; 5 slope groups; 19 classes of land cover; 500-meter elevation zones;	**Case 3** no aspect groups; no slope groups; 19 classes of land cover; 500-meter elevation zones;	**Case 2** no aspect groups; no slope groups; 19 classes of land cover; no elevation zones;	**Case 1** no aspect groups; no slope groups, 3 classes of land cover, one dominant soil type no elevation zones;
4 km	0.80	0.81	0.78	0.75
0.0625° [~6.25 km]	0.79	0.79	0.77	0.75
0.125° [~12.50 km]	0.82	0.81	0.75	0.75
0.25° [~25.00 km]	0.81	0.83	0.77	0.76
0.5° [~50.00 km]	0.79	0.82	0.76	0.76
1.0° [~100.00 km]	0.83	0.81	0.79	0.78
2.0° [~200.00 km]	0.77	0.77	0.77	0.80

We use a single objective calibration algorithm for model calibration, however and for investigating the parameter uncertainty, we check the behavioral parameter sets with E_{NS} higher than 0.7 (an arbitrary values). These parameter sets may have very different soil parameters

Moved up [9]: <#>*Geospatial data and meteorological forcing*¶

Moved up [10]: <#>-2005 v2 includes 19 different classes. The land cover map is used to set up the vegetation file and vegetation library (look up table) for the VIC model (Nijssen et al., 2001). ¶
<#>Soil texture: We used the Harmonized World Soil Data, HWSD (Fischer et al., 2008). For each polygon of the world harmonized soil we use the highest proportion of soil type. The HWSD provide the information for two soil layers, in this study we base our analyses on the lower soil layer reported in HWSD to define the soil characteristics needed for the VIC soil file. ¶

Moved up [11]: <#>For upscaling the WRF input forcing, we use the CANDEX package (DOI: 10.5281/zenodo.2628351) to map the 7 forcing variables to various resolutions (1/16°, 1/8°, 1/4°, 1/2°, 1° and 2° from the original resolution of 4 km). We used the required variables from the WRF data set namely, total precipitation, temperature, short and long wave radiation at the ground surface, V, U components of wind speed and water vapor mixing ratio. ¶
<#>The shortwave radiation is rescaled based on the slope and aspect of the respective

Moved up [12]: <#>In this study we differentiated four aspects and five slope classes. The temperature at 2 meters are adjusted using the environmental lapse rate

Moved up [13]: <#> The assumed lapse rate aligns with earlier findings from the region of study (Pigeon and

Moved up [14]: <#>We make sure that the $d_{2,\ forested}$ is

Moved up [17]: <#>The diffusive wave parameters are set

Deleted: <#>Experiment 3: How do different model ... [14]

Deleted: <#>The inputs and forcing we used to set up the ... [15]

Deleted: <#>Digital Elevation Model: in this study we ... [16]

Deleted: <#>VIC-GRU parameter values.¶ ... [17]

Deleted: <#>GRUs (refer to Appendix-A for more details).

Deleted: <#>for temperature of 6.5 km per 1000 meters.

Deleted: <#>The parameters for the routing scheme and ... [19]

Deleted: <#>¶

Formatted: Font color: Accent 1

Deleted: <#>¶ ... [18]

Moved up [16]: <#>Impulse Response Function (IRF)

Deleted: These results indicate that the coarse resolution ... [20]

Deleted: 0.5°

Moved (insertion) [15]

Formatted: Font: 12 pt

Formatted: Normal, Justified

Deleted: simulations,

Deleted: for example,

combinations. Figure-3a illustrates the possible combinations of behavioral parameter sets for Case-2-4km (E_{NS}> 0.7). As a specific example, saturated hydraulic conductivity, K_{sat}, and slope of water retention curve, E_{exp}, have very different combinations of values within the specified parameter ranges for calibration. The result indicates the two parameters that are often fixed or a priori allocated based on look up tables can exhibit significant uncertainty and non-identifiability. It is also noteworthy to mention that among the parameters, K_{slow} seems to be the most identifiable parameter while it is set to the upper limit range. There might be two explanation for this behavior: (1) this might be related to the Nivo-glacier regime of the basin of study that has strong yearly cycle due to snow accumulation and snow melt (2) and the lack of macropore water movement to the baseflow component which results in dampen input to this component and in return result in K_{slow} to be higher than expected for a baseflow reservoir (for further reading refer to Gharari et al., 2019). Overall, the results indicate that calibrating the VIC model parameters using a sum-of-squared objective function at the basin outlet does not constrain the VIC subsurface parameters. Additionally, we examine the difference between the fluxes, in this case transpiration, for all the parameter sets presented in Figure-3a. Figure 3-b illustrates differences between the yearly transpiration flux for the computational units of case-2-4km. This difference can be as high as 250 mm per year indicating the internal uncertainty of fluxes and related states in reproducing similar performance metric. This difference can be the basis of model diagnosis to understand which computational units are causing the internal uncertainty and perhaps the underlying reasons.

Deleted: configuration.

Deleted: an

Deleted: can

Deleted: the parameters. Figure-4 illustrates the possible combinations of K_{sat} and E_{exp} with performance higher than E_{NS} greater than 0.7 for Case-2-4km.

Deleted: Moreover,

Deleted: soil parameters.

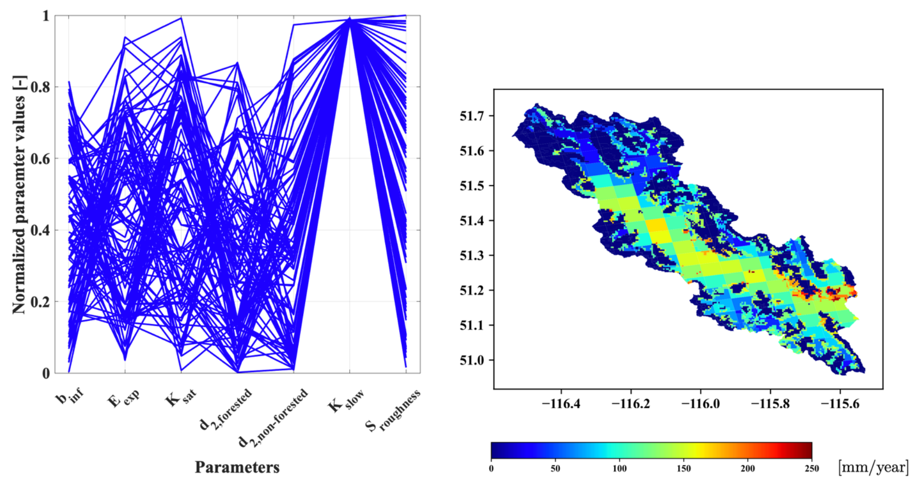


Figure-3 – (a) The normalized values for the parameters of Case-2-4km that have E_{NS}, Nash-Sutcliffe efficiency, values of higher than 0.7. (b) The difference of largest and smallest yearly simulated transpiration for parameter sets with E_{NS} above 0.7.

4.2 Experiment-2

The second experiment compares the performance of a parameter set from the Case-4-4km across the configurations with degraded geophysical information and aggregated spatial information. Here we choose a parameter set that has E_{NS} of above 0.7 (this can be any other parameter sets). Figure-4 shows the evaluation metric, E_{NS}, for the streamflow of every river segment across the domain in comparison with the synthetic case (Case-4-4km). From Figure-4, it is clear that the E_{NS} is less sensitive for river segments with larger upstream area (i.e. segments that are located more downstream). This result has two major interpretations (*i*) the parameter transferability across various configurations is dependent on the sensitivity of simulation at the scale of interest meaning that as long as good performance is achieved in the context of modeling, for example for the streamflow at the basin outlet, the parameters can be said to transferable for that scale and (*ii*) often inferred parameters at larger scale may not guarantee good performing parameters at the smaller

Deleted:

Deleted: with E_{NS} of above 0.7

Deleted: Figure-5

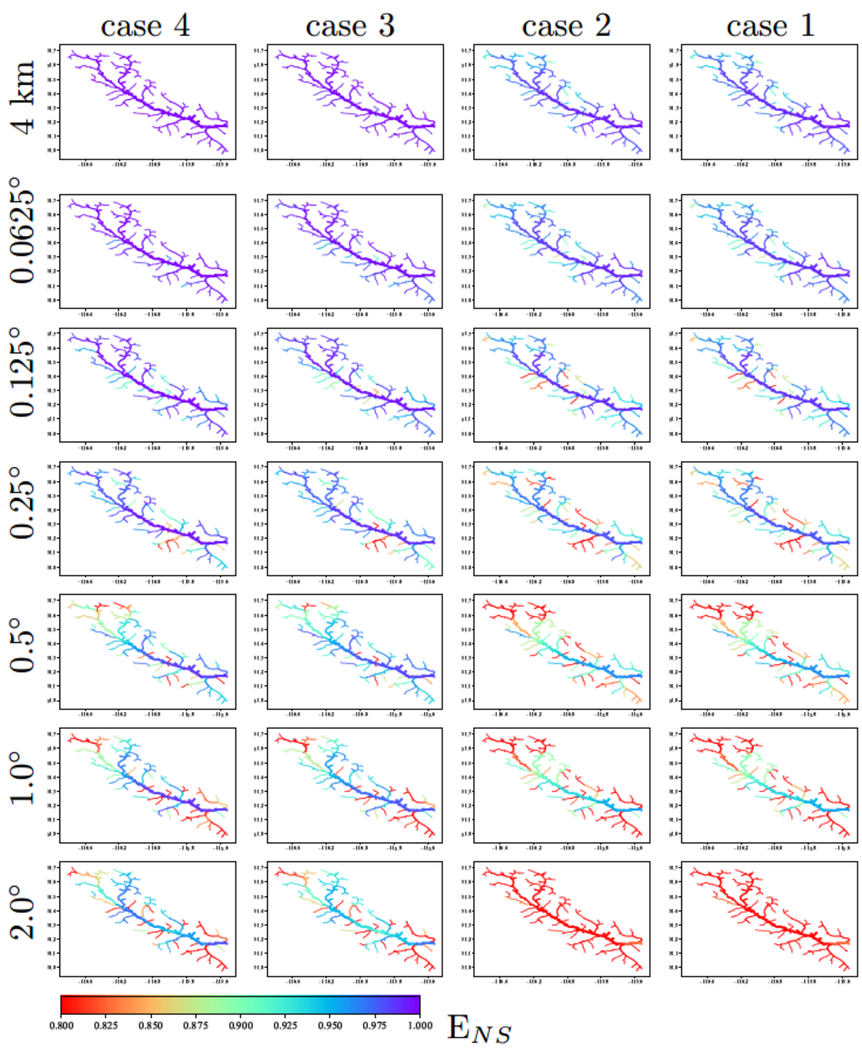
Deleted: 5

Deleted:

Deleted: read

Deleted: configuration

scales (read upstream areas) as the changed in the performance metric varies significantly across scale for the smaller modeling elements.



Deleted: .

Deleted: Figure-6

Figure 4 – Differences of the simulated streamflow at river segments in comparison with the synthetic case, Case-4-4km, expressed in performance metric, E_{NS}.

To understand the spatial patterns of model simulations for all the configurations, we evaluate the distribution of the snow water equivalent, SWE, for the computational units on 5th of May 2004 (Figure-5). In general, the SWE follows the forcing resolution and its aggregation. Although coarser forcing resolutions results in coarser SWE simulation, the geospatial details such as elevation zones and slope and aspects result in more realistic representation of SWE as the snow layer is thinner for south facing slopes where more melt can be expected to occur, and thicker for higher elevation zones (compare SWE simulations for Case-4-2° and Case-3-2° in Figure-5) which is consistent with higher precipitation volumes and slower melt at higher elevation. Another observation from Figure-5 is the unrealistic distribution of SWE for configurations without elevation zones (Case-2 and Case-1). The lack of elevation zones results in both valley bottom and mountain tops to be forced with the same temperature. Snow is more durable in the forested areas as the result of model formulation, which are at lower elevation, while SWE is less for higher mountains, which is unrealistic. We remind the reader that the various spatial pattern of SWE across different configurations are from the simulations that results in rather similar performance metric, E_{NS}, for the streamflow at the outlet of the basin.

Moved (insertion) [20]

Moved (insertion) [21]

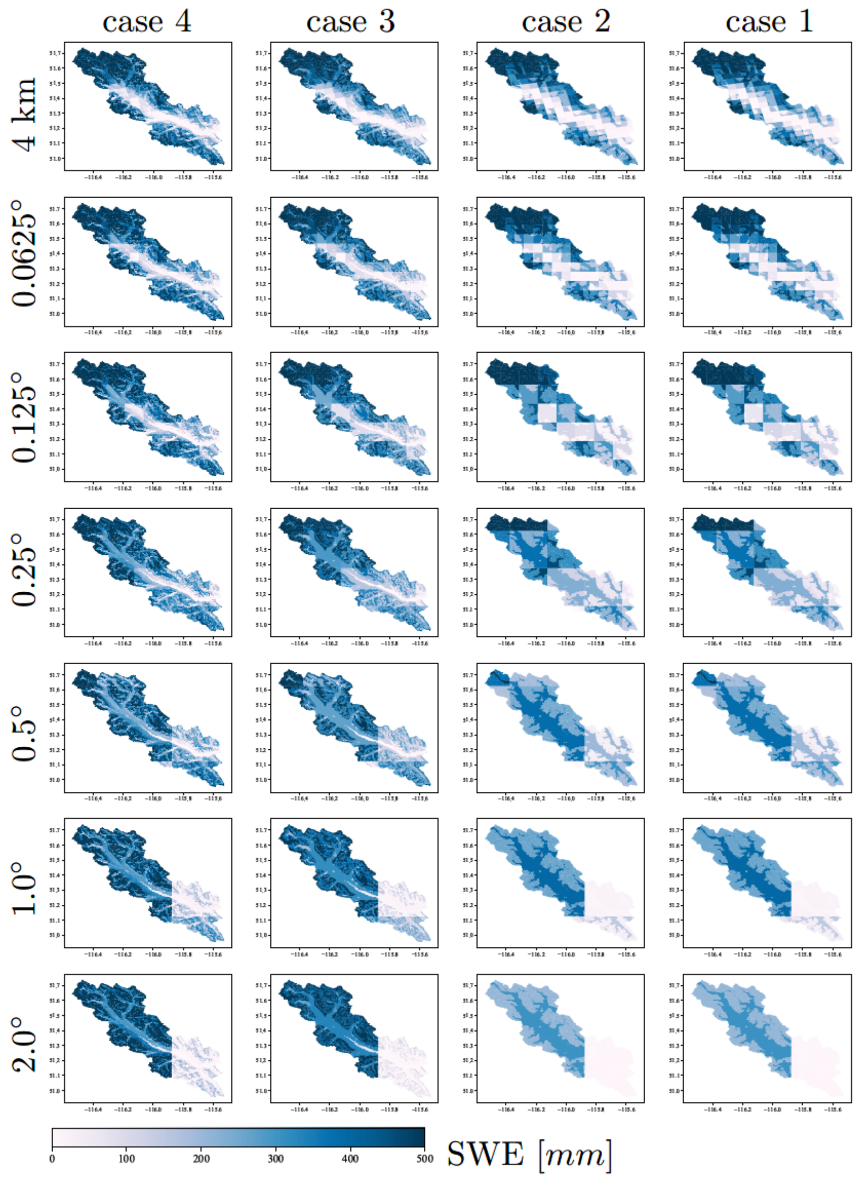


Figure 5- Comparison of the snow water equivalent for 5th of May 2004 for various configurations.

Figure-6a shows the performance of the streamflow across various configurations for the most downstream river segment (the gauged river segment which is used for parameter inference through calibration). Figure 6a illustrates that most of the configurations have similar scaled E_{NS} at the basin outlet. We compared the maximum snow water equivalent across different configurations for a computational unit located in the Bow Valley Bottom (an arbitrary location of -116.134˚W and 51.382˚E) for the year 2004 (Figure-6b). The result indicates that the SWE is higher for configurations with coarser forcing resolutions (almost triple). This is due to the reduced temperature as a result of masking warmer valley bottom by cooler and higher forcing grids over the Rockies. Such analyses can provide insights on the appropriate model configurations for different applications. Also and as an example, if model configurations of different complexity are known to show similar performance for a given parameter set, uncertainty and sensitivity analysis can be done initially on the models with fewer computational units and the results of the analysis can be applied to models with a higher number of computational units. This analysis can be repeated for different parameter sets, e.g., poorly performing parameter sets or randomly selected parameter sets, to better understand accuracy-efficiency tradeoffs of the model within its specified parameters ranges.

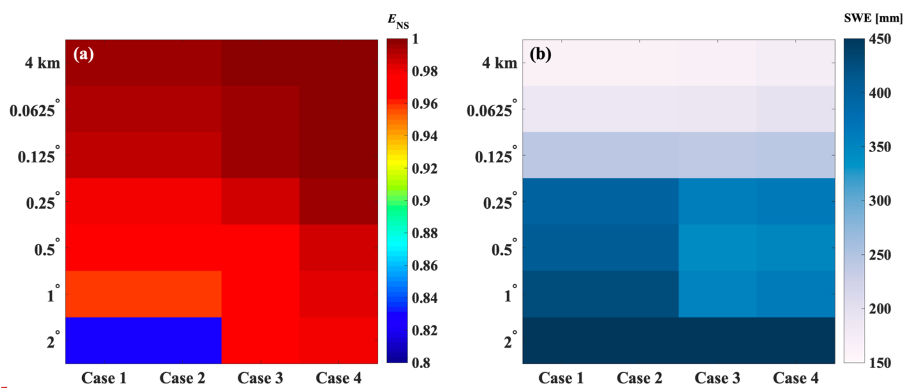


Figure -6 (a) The relative performance of model simulation across various configurations with a single parameter set. (b) Maximum of snow water equivalent for an arbitrary location of -116.134˚W and 51.382˚E located in Bow Valley Bottom across various model configurations for the year 2004.

Formatted: Space After: 0 pt, Don't adjust space between Latin and Asian text, Don't adjust space between Asian text and numbers

Deleted: often

Deleted: 6

Deleted: This analysis can be repeated for

Deleted: parameter sets, e.g., poorly performing parameter sets or randomly selected parameter sets,

Deleted: better understand accuracy-efficiency tradeoffs

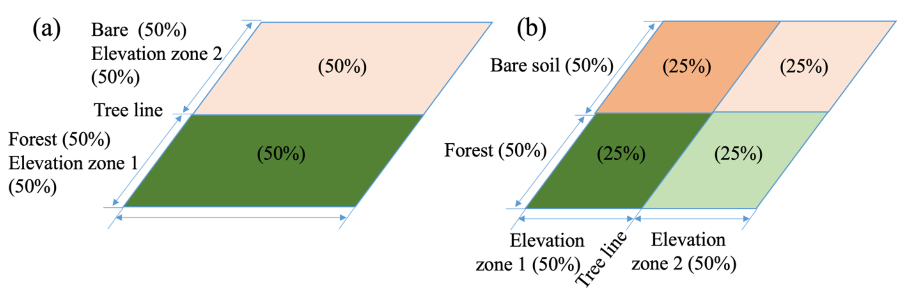
Deleted: As

Deleted: This is however under assumption that parameters are transferable based on the concept of GRU

Deleted: To understand the spatial patterns of model simulations for all the configurations, we evaluate the distribution of the snow water equivalent, SWE, for the computational units on 5th of May 2004 (Figure-7).

5 Discussion

In this study, we proposed a vector-based configuration for land models and applied this setup to the VIC model. We used a vector-based routing scheme, mizuRoute, which was forced using output from the land model (one-way coupling). Unlike the grid-based approach, there is no upscaling of land cover percentage or soil characteristics to a new grid size. This enables us to separate the effects of changes in forcing from changes in the spatial configurations. As mentioned earlier in Section 2, the vector-based configuration of land models may help avoiding unrealistic configuration of soil type, land cover or elevation zones that may happen in traditional grid-based implementation and hence increase the model fidelity. As an example, VIC configuration at grid scale assumes equal distribution of land cover over different elevation zones. Figure-7b illustrates how the traditional VIC configuration at grid-scale wrongly considers forested land cover above tree line. This issue is avoided in vector-based configuration as the set up will only include two computational unit of forested area below tree line and bare soil above the tree line (Figure 7a). The vector-based setup also provides more flexibility in comparing the model simulations across computational units (as an example, refer to Figure 5), and also comparing model simulations with point measurements, such as snow water equivalent. Moreover, the vector-based routing results in complete decoupling of the land model computational units' spatial extent from routing sub-basins. For the grid-based configuration of land models, it is often the case that in land model grid and routing grids are identical which result in further decision on upscaling of the routing direction to the land model grid scale.



Moved up [20]: In general, the SWE follows the forcing resolution and its aggregation. Although coarser forcing resolutions results in coarser SWE simulation, the geospatial details such as elevation zones and slope and aspects result in more realistic representation of SWE as the snow layer is thinner for south facing slopes where more melt can be expected to occur, and thicker for higher elevation zones (compare SWE simulations for Case-4-2˚ and Case-3-2˚ in Figure-

Moved up [21]: is the unrealistic distribution of SWE for configurations without elevation zones (Case-2 and Case-1). The lack of elevation zones results in both valley bottom and mountain tops to be forced with the same temperature. Snow is more durable in the forested areas as the result of model formulation, which are at lower elevation, while SWE is less for higher mountains, which is unrealistic.

Deleted: 7) which is consistent with higher precipitation volumes and slower melt at higher elevation. Another observation from Figure-7

Deleted: ¶
We compared the maximum snow water equivalent across different configurations for a computational unit located in the Bow Valley Bottom (an arbitrary location of -116.134˚W and 51.382˚E) for the year 2004. Figure-8 illustrated the maximum snow water equivalent for the period of simulation. The result indicates that the SWE is higher for configurations with coarser forcing resolutions (almost double). This is due to the reduced temperature as a result of masking warmer valley bottom by cooler and higher forcing grids over the Rockies.¶
Experiment-3¶
The calibration of model Case-2-4km and Case-2-4km-macro result in similar E_{NS} values of 0.78 and 0.75, indicating that both models are able to reproduce the observed streamflow to a similar extent although they are structurally different (the slow reservoir is recharged through only micropore and only macropore water movement in Case-2-4km and Case-2-4km-macro respectively). Figure-9 shows the streamflow hydrographs for the best performing parameter sets. . Observed Bow River at Banff has a minimum streamflow of approximately 15 cubic meter per second during snow accumulation months. This flow may be the result of regulation, return flow from human activities or unaccounted processes such as groundwater flow which are rather difficult for the linear reservoir of baseflow to simulate. Results (not shown) also indicate that both models structures generate 4 to 5 times more baseflow than surface runoff. This might be very intuitive as the model structure and parameters only have one processes, the slow reacting component, to simulate ... [21]

Formatted: Space After: 0 pt, Don't adjust space between Latin and Asian text, Don't adjust space between Asian text and numbers

Deleted: We term this new modelling approach VIC-GRU.

Deleted: The vector-based setup also provides us with more flexibility in comparing the model simulations across GRUs, and also comparing model simulations with point measurements, such as snow water equivalent

Formatted: Font: Times New Roman

Figure-7 – (a) The realistic configuration of a natural system with land cover consist of 50% Bare soil and 50% forest within a grid located in two different elevation zones above and below the tree line which is preserved with vector-based configurations and (b) the traditional VIC configurations for the given system at the grid for the two elevation zones and 2 land cover which results in unrealistic combination of forested land cover above the tree line and bare soil below the tree line.

Our results illustrate various vector-based spatial configuration of the VIC model generates similar large-scale simulations of streamflow when the setups are calibrated by maximizing the Nash-Sutcliffe score at the basin outlet. Similarly, we have shown that often behavioral parameter sets yield similar E_{NS} and can be significantly uncertain (Figure-3a) or have significant differences for their internal behavior which may be very well masked by aggregation of the result at the grid scale or basin scale (Figure-3b and Figure-5). Generally, both parameter and states and fluxes uncertainties are not often evaluated or reported for land models (Demaria et al., 2007) or is ignored by tying parameters, linking specific hydraulic conductivity to the slope of water retention curve, for example, so that the possible combination of parameters are reduced. Moreover, the behavior of K_{slow} parameter can be revealing of the VIC model structural deficiencies which are not often explored for land models. The recession coefficient obtained from recession analysis on the observed hydrograph is approximately 0.01 1/day while the calibrated K_{slow} has much higher values of around 0.90 1/day. This can be due to damped response from the two top soil layers and lack of macropore water movement to the baseflow component. Similarly, and due to lack of macropore water movement in the VIC model, and land models in general, it is impossible to infer the K_{slow} based on recession analysis on the observed hydrograph (for further reading on this and also recession analysis refer to Gharari et al., 2019). This finding can be generalized to the 5-parameter VIC baseflow, highlighting the need to properly evaluate the often not observable but calibrated baseflow parameters for the VIC model and if it is possible to identify 5 parameters based on the recession limps of a hydrograph.

Land models are often applied at large spatial scales. The results clearly show that the deviation of streamflow is much lower in river segments with larger upstream area (Figure 4 and 6a). It is often the case that the model parameters and associated processes are inferred thought calibration

Deleted: that

Deleted: -GRU approach

Deleted: across the various spatial configurations

Deleted: VIC-GRU is

Deleted: the VIC soil parameters can be very different when calibrated using different spatial configurations and that parameter transferability to different forcing resolutions and model setups is limited. This uncertainty is

Deleted: them are reduced

Formatted: Font: +Headings CS (Times New Roman)

Deleted: 5

Deleted: 6

Deleted: based on

on the streamflow at the basin outlet or over a large contributing area. We argue that this may not be a valid strategy for process understanding at the smaller scale (read computational units), given the large uncertainty exhibited by the parameters. Therefore, hyper-resolution modeling efforts, Wood et al. 2011, may suffer from poor process representation and parameter identification at the scale of interest (Beven et al., 2015). What is needed instead of efficiency metrics that aggregate model behavior across both space (e.g. at the outlet of the larger catchment) and time (e.g. expressing the mismatch between observations and simulations across the entire observation period as a single number), is diagnostic evaluation of the model's process fidelity at the scale at which simulations are generated in case of available observations (e.g. Gupta et al., 2008; Clark et al., 2016).

One might argue that the spatial discretization is important for realism of model fluxes and states. Moving to significantly high number of computational units may result in computational units that are similar in their forcing and geo-spatial fabric (such as soil and land cover types). Based on the result of this study for snow water equivalent (Figure-5), we can argue that the snow patterns are fairly similar for the configurations that have elevation zones and finer resolution of forcing (case3 and 4 and forcing resolution less than 0.125 degree). It can be further explored if the model simulation at finer resolutions can be approximated by interpolating result of a model with coarser resolution ($m(x|\theta) \sim \overline{m(x|\theta)}$), in which m is the model, x is forcing and $\boldsymbol{\theta}$ is the model parameter set).

The analysis on the accuracy-efficiency tradeoff presented in this study, Figure-6, can be used in model analysis such as sensitivity and uncertainty. One can assume a configuration with fewer computational units can be a surrogate for a model with more computational units, under the condition that both models are known to behave similarly for a given parameter set. The calibration can be done on the model configuration with less computational unit and the parameters can be transferred directly to the model with more computational units or can be used as an initial point for optimization algorithm to speed up the calibration process. Similarly, the sensitivity analyses can be done primarily on the model with less computational units.

In this study and following the concept hydrological similarity, we assume the parameters of computational units are identical for computational units with similar soil and land cover. The

Deleted: GRU

Deleted: ,

Moved (insertion) [22]

Deleted: We have shown that changes in model structure can result in identical performance for system-scale evaluation metrics (in this case, the Nash Sutcliffe Efficiency). We have changed the land model structure by replacing the micropore with macropore water movement to the slow reservoir. Similar to the parameter uncertainty, this indicates that lack or inclusion of macropore processes at the GRU scale does not have any notable effect on the efficiency score of the model simulation of streamflow at the outlet of the basin, even though the process simulations at the GRUs are different. Alternatively, this also shows that the micropore and macropore processes and their parameters may not be identifiable through calibration on the observed streamflow, which supports the argument against assuming that fine-scale parameters and processes can be inferred from large-scale observations. Although in this study we only focus on the processes and parameters that are often used to calibrate for the VIC model such as subsurface processes, it is possible to repeat the same analysis on wider range of processes such as snow processes or routing parameters. ¶
It is often computationally expensive to evaluate the uncertainty and sensitivity of land models. Following the results presented in Figure-6, one

Moved (insertion) [23]

Deleted: ,

Moved up [22]: One might argue that the spatial discretization is important for realism of model fluxes and states.

Deleted: Moving to significantly high number of GRUs may result in computational units that are similar in their forcing and spatial variability. Based on the result of this study for snow water equivalent (Figure-8

Moved up [23]:), we can argue that the snow patterns are fairly similar for the configurations that have elevation zones and finer resolution of forcing (case3 and 4 and forcing resolution less than 0.125 degree). It can be further explored if the model simulation at finer resolutions can be

Deleted: parameter set). ¶

Deleted: of GRUs, Grouped Response Units, we assumed the physical characteristics of soil and vegetation

Deleted: a given GRU across various model configurations.

Moved down [24]: Techniques such as multiscale parameter regionalization (MPR, Samaniego et al., 2010) can be used to scale parameter values for different model configurations.

Deleted: However, applying these techniques, such as in this case that has significant parameter and process uncertainty and significance accuracy-performance

Moved down [25]: tradeoff, should be put through rigorous tests (Merz et al., 2020, Liu et al., 2016

Deleted:).¶
Also, the

degree of validity of hydrological similarity concepts is debatable. For example, at the catchment scale, Oudin et al. (2010) have shown that the overlap between catchments with similar physiographic attributes and catchments with similar model performance for a given parameter set is only 60%. Physiographic similarity (in our case expressed through GRUs) does thus not necessarily imply similarity of hydrologic behavior, even though this is the critical assumption underlying GRUs. The VIC parameters can be linked to many more characteristics such as slope, height above nearest drainage (HAND, Renno et al., 2008), or Topographical Wetness Index (Beven and Kirkby, 1979) as has been done by Mizukami et al. (2017) and Chaney et al. (2018). Techniques such as multiscale parameter regionalization (MPR, Samaniego et al., 2010) can be used to scale parameter values for different model configurations. However, the functions that are used to link computational units and physical attributes to model parameters remains mostly based on inference, (i.e., calibration), and the reproducibility of those relationships are not very well explored. However, applying these techniques, such as in this case that has significant parameter and process uncertainty and significance accuracy-efficiency tradeoff, should be put through rigorous tests (Merz et al., 2020, Liu et al., 2016).

A key outstanding challenge is for models to provide the right results for the right reasons (Kirchner, 2006). Thoughtful strategies to formulate parameter and process constraints based on expert knowledge can reduce the plausible range of behavioral parameter sets. In this study, we imposed a simple parameter constraint that the root zone moisture storage of forested area should be larger than the non-forested area (Table-1). Additional process constraints, if available, can be increasingly difficult to satisfy. More rigorous parameter estimation methods that satisfy the fidelity constraints based on expert knowledge are required (e.g., Gharari et al., 2014).

6 Conclusions

The vector-based configuration of land models can provide modelers with more flexibility, e.g. representing the impact of various forcing resolution or geospatial data representation. The conclusions from this study can be summarized as follows:

1) The land model configuration with the highest number of computational units may not result in improved performance and better spatial simulation, in terms of obtained

Deleted: the concept of GRU,

Deleted: basd on physical attribute similarities,

Deleted: Although the GRUs in this study include slope and aspect, these characteristics were not translated into the model parameters and was only used for forcing manipulation.

Deleted: Cheney

Moved (insertion) [24]

Deleted: linked the attribute

Deleted: characteristics

Deleted: assumptions rather than

Deleted: them

Moved (insertion) [25]

Deleted: (if possible

Formatted: Font: Times New Roman

Deleted: In this study, the vector-based routing configuration does not include lakes and reservoirs. This is often a neglected element of land modeling efforts and has only attracted limited attention compared to the its impact on terrestrial water cycle (Haddeland et al., 2006, Yassin et al., 2018). The presence of lakes and reservoirs and their interconnections reduces the, already limited, ability of inference of land model parameter based on calibration on the observed streamflow because streamflow variability is reduced. ¶
Although not primary the result of this study, however, the Nivo-glacial regime of the Bow River Basins is mostly dominated by snow melt that contributes to streamflow through baseflow (slow component of the hydrograph). The high Nash-Sutcliffe Efficiency, E_{NS}, is partly due to the fact that it is rather easy for the land model to capture the yearly cycle of the streamflow only with snow processes while rapid subsurface water movement, such as macropore, are largely missing in the land models but do not lead to notably increased efficiency scores when they are included in the model structure. More caution is needed for use of the land model for flood forecasting (Vionnet et al., 2019) for this region and all the Nivo-glacial river systems in western Canada, McKenzie, Yukon and Colombia River Basins.¶

Deleted: setup for the

Deleted: model

Deleted: , while avoiding decisions that are often taken for model configuration at grid level.

Deleted: conclusion and messages

Deleted: <#>Regardless of observations at the scale of modeling, a model configuration with lower computational units, coarser resolution and less geospatial information, can produce model simulations with similar efficiency scores as configurations with higher computational units. The choice of model set up should be tested within the context and purpose of modeling for every different case.¶
<#>The model

efficiency scores, while the internal model state and fluxes can show significant uncertainty.

2) There is significant parameter and structural uncertainty associated with the land model (in this case, the VIC model). This uncertainty poses challenges for the process and parameter inference when the model is calibrated by minimizing the sum-of-squared differences between simulated and observed streamflow. Any parameter regionalization efforts should take these uncertainties into account. Our results emphasize that more attention is needed on the topic of parameter and process inference at finer modelling scales.

3) A model configuration with lower computational units, coarser resolution and less geospatial information, may reproduce model simulations with similar efficiency scores as configurations with higher computational units. Less computationally expensive configurations can be used instead for primary uncertainty and sensitivity analysis.

A key scientific challenge is hydrological scaling. i.e., how do small-scale heterogeneities shape large-scale fluxes. Addressing this challenge requires a mix of both explicit representations of spatial heterogeneity (enabled through spatial discretization of the landscape) and implicit representations of heterogeneity (enabled through sub-grid parameterizations). The contribution in this paper is to advance flexible spatial configurations for land models – our approach improves the explicit representation of spatial heterogeneities, at least compared to traditional approaches that simply drape a grid over the landscape. Much more work is required across all spatial scales to carefully evaluate how a mix of implicit and implicit representations of spatial heterogeneity can improve process representations.We encourage the community to develop tools which can enable easier and more flexible configuration of land models that can be used to explore the above-mentioned research questions.

Acknowledgment. This research was undertaken thanks in part to funding from the Canada First Research Excellence Fund.

Data availability. All the data used in this study are available publicly (refer to references).

Deleted: <#>parameters

Deleted: <#>be transferred without substantial performance changes

Deleted: certain model setups

Deleted: <#>There is significant parameter and structural uncertainty associated with the VIC-GRU model. This uncertainty creates challenges for the process and parameter inference using calibration on streamflow. Any regionalization for parameters of the model should take into account these significant uncertainties. Our results recommend caution and more attention to the topic of parameter and process inference at finer modelling scales.¶
<#>We also encourage the need for tools which can facilitate easier and more flexible set up of land models that in turn can facilitate the above mentioned research questions.¶

7 Appendix

7.1 Appendix – A

This appendix reflect on the method and equations that have been used to calculate the ratio of the solar radiation on a surface with slope and aspect to a flat surface. Please note that the angles in the equations are in radian but for better communication we express angles in degree in the text.

Declination angle: declination angle can be calculated for each day of year and is the same for the entire Earth (Ioan Sarbu, Calin Sebarchievici, in Solar Heating and Cooling Systems, 2017):

$$\delta = 23.45 \frac{\pi}{180} \sin \left[\frac{2\pi}{360} \frac{360}{365} (284 + d) \right] \quad \text{(A-1)}$$

in which δ is declination angle in radian and d is the number of day in a year starting from 1st of January.

Hour angle: hour angle is the angle expressed the solar hour. The reference of solar hour angle is solar noon (hour angle is set to zero) when the sun is passing the meridian of the observer or when the solar azimuth is 180° (north direction with azimuth of 0°). The hour angle can be calculated based on the:

$$\sin \omega = \frac{\sin \alpha - \sin \delta \sin \emptyset}{\cos \delta \cos \emptyset} \quad \text{(A-2)}$$

In which α, ϕ and δ are the altitude angle, latitude of the observer and declination angle. The sunset and sunrise hour can be calculated as (when sun is at horizon and solar altitude angle is zero):

$$\cos \omega_s = -\tan \emptyset \tan \delta \quad \text{(A-3)}$$

More caution is needed using equation A-3 for latitude above and below 66.55° north and south respectively where it can be always day or night with no sunrise or sunset during part of the year. The number of daylight hours that can be split before and after the solar noon equally can be calculated based on (assuming 15° for every 1 hour):

$$n = \frac{2\omega_s}{15} \frac{180}{\pi} \quad \text{(A-4)}$$

Deleted: methods
Deleted: ration
Deleted: to flat surface and
Deleted: based on
Deleted: *N*
Deleted: In
Deleted: N
Deleted: beginning
Deleted: .
Deleted: inclination angle.
Deleted: The solar noon is not exactly coinciding with12 am of the local time zone. However in this study we assume the two property are coinciding. The sunset and sunrise hour can be calculated from:
Deleted: For beyond 66.55 degree if the value of the right hand side is above 1 then there is 24 hour of daylight and if the right hand side is less than 1 the will be 24 hour of darkness.
Deleted: degree

And therefore, hour angle can be easily calculated for time before and after solar noon the (relationship between the 15° equals to an hour). Hour angle is negative for the time before solar noon and positive for the time after solar noon. Note the solar noon does not often coincide with 12 pm of the local time zone. There are relationships to find the local time of solar noon.

Solar altitude angle: Solar altitude angle is the angle of sun rays with the horizontal plane of an observer. This angle is maximum at solar noon and 0° for subset and sunrise. The altitude angle can be calculated based on the:

$$\sin\alpha = \sin\delta\sin\emptyset + \cos\delta\cos\omega\cos\emptyset \qquad \text{(A-5)}$$

For the solar noon when ω, hour angle, is zero the question simplifies to:

$$\sin\alpha = \sin\delta\sin\emptyset + \cos\delta\cos\emptyset = \cos(\emptyset - \delta) = \sin\left(\frac{\pi}{2} - \emptyset + \delta\right) \qquad \text{(A-6)}$$

Solar Azimuth: The solar azimuth angle, A_{Sun} reflect on the angle of the sun on the sky from the north with clockwise rule. The azimuth angle can be calculated as:

$$\sin A_{Sun} = \frac{\sin\omega\cos\delta}{\cos\alpha} \qquad \text{(A-7)}$$

The solar azimuth angle for the solar noon is set to be 180°.

The azimuth at the sunset and sunrise can be calculated:

$$\sin A_{Sun,rise} = -\sin\omega_s\cos\delta \qquad \text{(A-8)}$$

$$\sin A_{Sun,set} = \sin\omega_s\cos\delta \qquad \text{(A-9)}$$

Surface Azimuth (a.k.a. aspect): The surface azimuth angle, $A_{Surface}$ reflect the direction of the any tilted surface to the north direction. This azimuth is fixed for any point while the solar azimuth changes over hours and seasons.

Angle of incidence θ: this angle represents the angle between a sloped surface and the sun rays. The model angle of the incidence for a slope surface β, and aspect of $A_{Surface}$ over latitude of ∅

Deleted: Altitude
Deleted: :
Formatted: Font: Not Bold
Deleted: This result the altitude angle for the solar noon to be:
$\alpha = \frac{\pi}{2} - \emptyset + \delta$ (A-7)
Deleted: θ_{Sun}
Deleted: North
Deleted: $\sin\theta_{Sun}$
Deleted: 8
Deleted: degree (calculated clockwise from north direction).
Deleted: using
Deleted: $\sin\theta_{Sun,rise}$
Deleted: 9
Deleted: $\sin\theta_{Sun,set}$
Deleted: 10
Deleted: $\theta_{Surface}$
Deleted: represent
Deleted: that reaches this sloped surface.
Deleted: $\theta_{Surface}$

can be calculated as (Kalogirou, in Solar Energy Engineering, 2009, in the reference formulation the Azimuth is from south which is corrected here for North):

$$\cos\theta = \sin\delta\sin\emptyset\cos\beta + \sin\delta\cos\emptyset\sin\beta\cos A_{Surface} + \cos\delta\cos\emptyset\cos\beta\cos\omega - \cos\delta\sin\emptyset\sin\beta\cos A_{Surface}\cos\omega - \cos\delta\sin\beta\sin A_{Surface}\sin\omega \qquad \text{(A-10)}$$

For the flat surface, both $A_{Surface}$ and β, is set to 0°, the incident angle can be calculated for the flat surface as

$$\cos\theta_{flat} = \sin\delta\sin\emptyset + \cos\delta\cos\emptyset\cos\omega \qquad \text{(A-11)}$$

In case where the angle of incident is larger than 90° the surface shades itself.

Correction of short-wave radiation based on slope and aspect. In this study we correct the WRF short wave radiation based on the surface slope and aspect. We first back calculated the incoming short-wave radiation by dividing the provided short wave radiation by the cosine of the incident angle of the flat surface. Then we can calculate the solar radiation of the sloped surface multiplying this value to the cosine of the incident angle of the slope surface. Basically, this ratio is:

$$R = \frac{\cos\theta}{\cos\theta_{flat}} \qquad \text{(A-12)}$$

The effect of the atmosphere is considered in the WRF product itself. However, and for incident level close to 90 degrees the ratio, *R*, might be very high values which result in the surface receiving unrealistically high value of radiation even higher than the solar constant, 1366 W/m2, at the top of the atmosphere. For cases with cosine values of incident angle lower than 0.05 we set the ratio to 0 to avoid this unrealistic condition.

Deleted: $\cos\theta_{Surface}$
Deleted: $\cos\theta_{Surface}$
Deleted: $\sin\theta_{Surface}$
Deleted: 11
Deleted: $\theta_{Surface}$
Deleted: zero,
Deleted: 12
Deleted: degrees
Deleted: Amendment
Deleted:
Deleted:
Deleted: 13
Deleted: cos

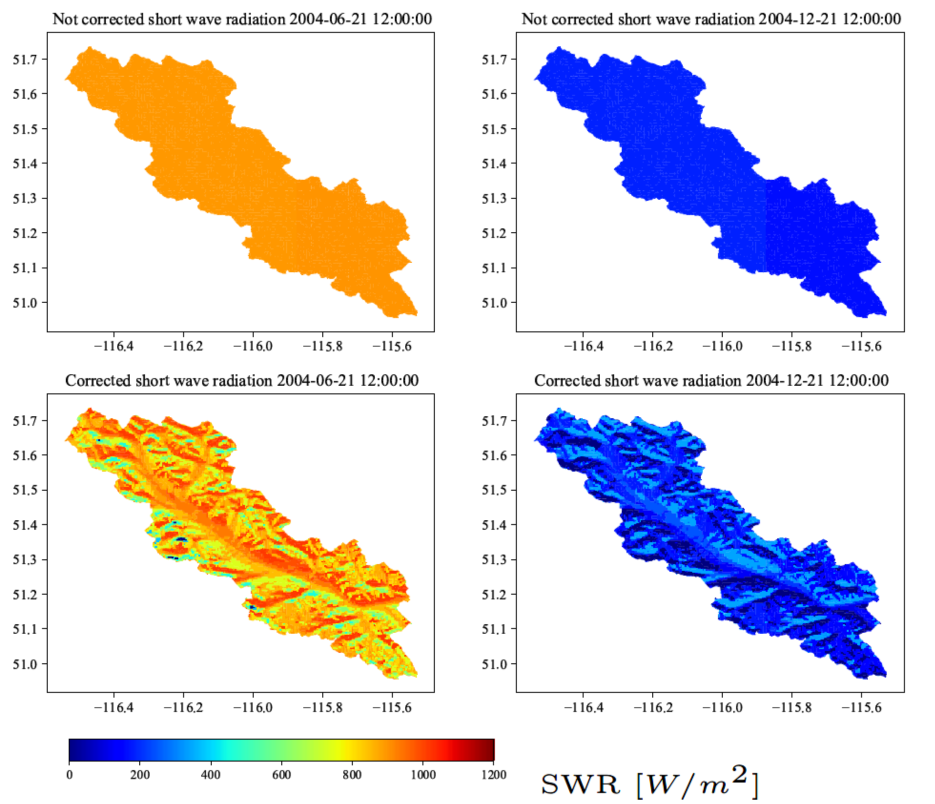


Figure A-1 Short wave radiation for (top left) not corrected for slope and aspect and (bottom left) corrected for slope and aspect for 21st June 2020 and (top right) not corrected for slope and aspect and (bottom right) corrected for slope and aspect for 21st December 2020.

8 References:

Annear, R.L. and Wells, S.A.: A comparison of five models for estimating clear-sky solar radiation. Water resources research, 43(10), 2007.

Beven, K., Cloke, H., Pappenberger, F., Lamb, R. and Hunter, N.: Hyperresolution information and hyperresolution ignorance in modelling the hydrology of the land surface. Science China Earth Sciences, 58(1), pp.25-35, 2015.

Beven, K.J. and Kirkby, M.J.: A physically based, variable contributing area model of basin hydrology/Un modèle à base physique de zone d'appel variable de l'hydrologie du bassin versant. Hydrological Sciences Journal, 24(1), pp.43-69, 1979.

Beven, K.J.: Rainfall-runoff modelling: the primer. John Wiley & Sons, 2011.

Blöschl, Günter, Rodger B. Grayson, and Murugesu Sivapalan: On the representative elementary area (REA) concept and its utility for distributed rainfall-runoff modelling. Hydrological Processes 9, no. 3-4 313-330, 1995.

Chaney, N.W., Van Huijgevoort, M.H., Shevliakova, E., Malyshev, S., Milly, P.C., Gauthier, P.P. and Sulman, B.N.: Harnessing big data to rethink land heterogeneity in Earth system models. Hydrology and Earth System Sciences, 22(6), pp.3311-3330, 2018.

Clark, M.P., Nijssen, B., Lundquist, J.D., Kavetski, D., Rupp, D.E., Woods, R.A., Freer, J.E., Gutmann, E.D., Wood, A.W., Brekke, L.D. and Arnold, J.R.: The structure for unifying multiple modeling alternatives (SUMMA), Version 1.0: Technical description. NCAR Tech. Note NCAR/TN-5141STR, 2015.

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

Clark P. M., Marc FP Bierkens, Luis Samaniego, Ross A. Woods, Remko Uijlenhoet, Katrina E. Bennett, Valentijn Pauwels, Xitian Cai, Andrew W. Wood, and Christa D. Peters-Lidard. "The evolution of process-based hydrologic models: historical challenges and the collective quest for physical realism." Hydrology and Earth System Sciences (Online) 21, no. LA-UR-17-27603, 2016.

Demaria, E.M., Nijssen, B. and Wagener, T.: Monte Carlo sensitivity analysis of land surface parameters using the Variable Infiltration Capacity model. Journal of Geophysical Research: Atmospheres, 112(D11), 2007.

Desborough, C.E.: Surface energy balance complexity in GCM land surface models. Climate Dynamics, 15(5), pp.389-403, 1999.

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

Fischer, G., F. Nachtergaele, S. Prieler, H.T. van Velthuizen, L. Verelst, D. Wiberg: Global Agro-ecological Zones Assessment for Agriculture (GAEZ 2008). IIASA, Laxenburg, Austria and FAO, Rome, Italy, 2008.

Flügel, W.A.: Delineating hydrological response units by geographical information system analyses for regional hydrological modelling using PRMS/MMS in the drainage basin of the River Bröl, Germany. Hydrological Processes, 9(3-4), pp.423-436, 1995.

Gharari, S., Clark, M., Mizukami, N., Wong, J.S., Pietroniro, A. and Wheater, H., 2019. Improving the representation of subsurface water movement in land models. Journal of Hydrometeorology, 2019.

Haddeland, I., Lettenmaier, D.P. and Skaugen, T.: Effects of irrigation on the water and energy balances of the Colorado and Mekong river basins. Journal of Hydrology, 324(1-4), pp.210-223, 2006.

Haddeland, I., Matheussen, B.V. and Lettenmaier, D.P., 2002. Influence of spatial resolution on simulated streamflow in a macroscale hydrologic model. Water Resources Research, 38(7), pp.29-1, 2002.

Hamman, J.J., Nijssen, B., Bohn, T.J., Gergel, D.R. and Mao, Y.: The Variable Infiltration Capacity model version 5 (VIC-5): infrastructure improvements for new applications and reproducibility. *Geoscientific Model Development (Online)*, *11*(8), 2018.

Formatted: English (US)

Formatted: English (US)

Formatted: English (US)

Hrachowitz, M. and Clark, M.P.: HESS Opinions: The complementary merits of competing modelling philosophies in hydrology. Hydrology and Earth System Sciences, 21(8), p.3953, 2017.

Jarvis, P.G.: The interpretation of the variations in leaf water potential and stomatal conductance found in canopies in the field. Philosophical Transactions of the Royal Society of London. B, Biological Sciences, 273(927), pp.593-610, 1976.

Kalogirou, S.: Solar Energy Engineering, edited by Soteris A. Kalogirou, 2009.

Kirkby, M. J. and Weyman, D. R: `Measurements of contributing area in very small drainage basins', Seminar Series B, No. 3. Department of Geography, University of Bristol, Bristol, 1974.

Knoben, W. J. M., Freer, J. E., Peel, M. C., Fowler, K. J. A., & Woods, R. A.: A brief analysis of conceptual model structure uncertainty using 36 models and 559 catchments. Water Resources Research, 56, e2019WR025975. https://doi.org/10.1029/2019WR025975, 2020.

Knudsen, J., Thomsen, A. and Refsgaard, J.C.: WATBALA Semi-Distributed, Physically Based Hydrological Modelling System. Hydrology Research, 17(4-5), pp.347-362, 1986.

Koster, Randal D., and Max J. Suarez: Modeling the land surface boundary in climate models as a composite of independent vegetation stands, Journal of Geophysical Research: Atmospheres 97, no. D3, 2697-2715, 1992.

Kouwen, N., Soulis, E.D., Pietroniro, A., Donald, J. and Harrington, R.A.: Grouped response units for distributed hydrologic modeling. Journal of Water Resources Planning and Management, 119(3), pp.289-305, 1993.

Latifovic, R., Zhu, Z., Cihlar, J., Giri, C., & Olthof, I.: Land cover mapping of North and Central America - Global Land Cover 2000. Remote Sensing of Environment, 89:116-127.

Lehner, B., Verdin, K. and Jarvis, A.: HydroSHEDS technical documentation, version 1.0. World Wildlife Fund US, Washington, DC, pp.1-27, 2006

Liang, X., Guo, J. and Leung, L.R.: Assessment of the effects of spatial resolutions on daily water flux simulations. Journal of Hydrology, 298(1-4), pp.287-310, 2004.

Liang, X., Lettenmaier, D.P., Wood, E.F. and Burges, S.J.: A simple hydrologically based model of land surface water and energy fluxes for general circulation models. Journal of Geophysical Research: Atmospheres, 99(D7), pp.14415-14428, 1994.

Liu, H., Tolson, B.A., Craig, J.R. and Shafii, M.: A priori discretization error metrics for distributed hydrologic modeling applications. *Journal of hydrology*, *543*, pp.873-891, 2016.

Manabe, Syukuro: Climate and the ocean circulation: I. The atmospheric circulation and the hydrology of the earth's surface. Monthly Weather Review 97, no. 11, 739-774, 1969.

Marsh, C.B., Pomeroy, J.W. and Spiteri, R.J.: Implications of mountain shading on calculating energy for snowmelt using unstructured triangular meshes. Hydrological Processes, 26(12), pp.1767-1778, 2012.

Matott, L.S.: OSTRICH: An optimization software tool: Documentation and users guide. University at Buffalo, Buffalo, NY, 2005.

Maxwell, R. M., L. E. Condon, and S. J. Kollet: A high-resolution simulation of groundwater and surface water over most of the continental US with the integrated hydrologic model ParFlow v3, Geoscientific model development 8, no. 3, 923, 2015.

Melsen, L., Teuling, A., Torfs, P., Zappa, M., Mizukami, N., Clark, M. and Uijlenhoet, R.: Representation of spatial and temporal variability in large-domain hydrological models: case study for a mesoscale pre-Alpine basin. Hydrology and Earth System Sciences, 20(6), pp.2207-2226, 2016.

Merz, R., Tarasova, L. and Basso, S.,: Parameter's controls of distributed catchment models–How much information is in conventional catchment descriptors?. *Water Resources Research*, p.e2019WR026008, 2020.

Mizukami, N., Clark, M.P., Sampson, K., Nijssen, B., Mao, Y., McMillan, H., Viger, R.J., Markstrom, S.L., Hay, L.E., Woods, R. and Arnold, J.R.: mizuRoute version 1: a river network routing tool for a continental domain water resources applications. Geoscientific Model Development, 9(6), pp.2223-2238, 2016.

Naef, F., Scherrer, S. and Weiler, M.: A process based assessment of the potential to reduce flood runoff by land use change. Journal of hydrology, 267(1-2), pp.74-79, 2002.

Newman, A.J., Clark, M.P., Winstral, A., Marks, D. and Seyfried, M.: The use of similarity concepts to represent subgrid variability in land surface models: Case study in a snowmelt-dominated watershed. Journal of Hydrometeorology, 15(5), pp.1717-1738, 2014.

Nijssen, B., O'Donnell, G.M., Hamlet, A.F. and Lettenmaier, D.P.: Hydrologic sensitivity of global rivers to climate change. Climatic change, 50(1-2), pp.143-175, 2001.

Niu, G.Y., Yang, Z.L., Dickinson, R.E. and Gulden, L.E.: A simple TOPMODEL-based runoff parameterization (SIMTOP) for use in global climate models. Journal of Geophysical Research: Atmospheres, 110(D21), 2005.

Niu, G.Y., Yang, Z.L., Mitchell, K.E., Chen, F., Ek, M.B., Barlage, M., Kumar, A., Manning, K., Niyogi, D., Rosero, E. and Tewari, M.: The community Noah land surface model with multiparameterization options (Noah-MP): 1. Model description and evaluation with local-scale measurements. Journal of Geophysical Research: Atmospheres, 116(D12), 2011.

Olivera, F., Valenzuela, M., Srinivasan, R., Choi, J., Cho, H., Koka, S. and Agrawal, A.: ARCGIS-SWAT: A GEODATA MODEL AND GIS INTERFACE FOR SWAT 1. JAWRA Journal of the American Water Resources Association, 42(2), pp.295-309, 2006.

Oudin, L., Kay, A., Andréassian, V. and Perrin, C.: Are seemingly physically similar catchments truly hydrologically similar?. Water Resources Research, 46(11), 2010.

Park, S.J. and Van De Giesen, N.: Soil–landscape delineation to define spatial sampling domains for hillslope hydrology. Journal of Hydrology, 295(1-4), pp.28-46, 2004.

Pietroniro, A., Fortin, V., Kouwen, N., Neal, C., Turcotte, R., Davison, B., Verseghy, D., Soulis, E.D., Caldwell, R., Evora, N. and Pellerin, P.: Development of the MESH modelling system for hydrological ensemble forecasting of the Laurentian Great Lakes at the regional scale. Hydrology and Earth System Sciences Discussions, 11(4), pp.1279-1294, 2007.

Pigeon, K.E. and Jiskoot, H.: Meteorological controls on snowpack formation and dynamics in the southern Canadian Rocky Mountains. Arctic, antarctic, and alpine research, 40(4), pp.716-730, 2008.

Pitman, A.J.: The evolution of, and revolution in, land surface schemes designed for climate models. International Journal of Climatology: A Journal of the Royal Meteorological Society, 23(5), pp.479-510, 2003.

Rasmussen, R., and C. Liu.: High Resolution WRF Simulations of the Current and Future Climate of North America. Research Data Archive at the National Center for Atmospheric Research, Computational and Information Systems Laboratory. https://doi.org/10.5065/D6V40SXP, 2017.

Reggiani, P., Hassanizadeh, S.M., Sivapalan, M. and Gray, W.G.: A unifying framework for watershed thermodynamics: constitutive relationships. Advances in Water Resources, 23(1), pp.15-39, 1999.

Rennó, C.D., Nobre, A.D., Cuartas, L.A., Soares, J.V., Hodnett, M.G., Tomasella, J. and Waterloo, M.J.: HAND, a new terrain descriptor using SRTM-DEM: Mapping terra-firme rainforest environments in Amazonia. Remote Sensing of Environment, 112(9), pp.3469-3481, 2008.

Samaniego, L., Kumar, R. and Attinger, S.: Multiscale parameter regionalization of a grid-based hydrologic model at the mesoscale. Water Resources Research, 46(5), 2010.

Samaniego, L., Kumar, R., Thober, S., Rakovec, O., Zink, M., Wanders, N., Eisner, S., Müller Schmied, H., Sutanudjaja, E., Warrach-Sagi, K. and Attinger, S.: Toward seamless hydrologic predictions across spatial scales. Hydrology and Earth System Sciences, 21(9), pp.4323-4346, 2017.

Sarbu, I. and Sebarchievici, C.: Thermal Energy Storage. Solar Heating and Cooling Systems, pp.99-138, 2017.

Savenije, H.H.G.: HESS Opinions" Topography driven conceptual modelling (FLEX-Topo)". Hydrology and Earth System Sciences, 14(12), pp.2681-2692, 2010.

Shafii, M., Basu, N., Craig, J.R., Schiff, S.L. and Van Cappellen, P.: A diagnostic approach to constraining flow partitioning in hydrologic models using a multiobjective optimization framework. Water Resources Research, 53(4), pp.3279-3301, 2017.

Shrestha, P., Sulis, M., Simmer, C. and Kollet, S.: Impacts of grid resolution on surface energy fluxes simulated with an integrated surface-groundwater flow model. Hydrology and Earth System Sciences, 19(10), pp.4317-4326, 2015.

Singh, R.S., Reager, J.T., Miller, N.L. and Famiglietti, J.S.: Toward hyper-resolution land-surface modeling: The effects of fine-scale topography and soil texture on CLM 4.0 simulations over the S outhwestern US. Water Resources Research, 51(4), pp.2648-2667, 2015.

Son, K. and Sivapalan, M.: Improving model structure and reducing parameter uncertainty in conceptual water balance models through the use of auxiliary data. Water resources research, 43(1), 2007.

Troy, T.J., Wood, E.F. and Sheffield, J.: An efficient calibration method for continental-scale land surface modeling. Water Resources Research, 44(9), 2008.

Uhlenbrook, S., Roser, S. and Tilch, N.: Hydrological process representation at the meso-scale: the potential of a distributed, conceptual catchment model. Journal of Hydrology, 291(3-4), pp.278-296, 2004.

Vionnet, V., Fortin, V., Gaborit, E., Roy, G., Abrahamowicz, M., Gasset, N. and Pomeroy, J.W.: High-resolution hydrometeorological modelling of the June 2013 flood in southern Alberta, Canada. Hydrology and Earth System Sciences Discussions, pp.1-36, 2019.

Vivoni, Enrique R., Valeri Y. Ivanov, Rafael L. Bras, and Dara Entekhabi: Generation of triangulated irregular networks based on hydrological similarity, Journal of hydrologic engineering 9, no. 4, 288-302, 2004.

Winter, T.C.: The concept of hydrologic landscapes 1. JAWRA Journal of the American Water Resources Association, 37(2), pp.335-349, 2001.

Wood, E.F., Roundy, J.K., Troy, T.J., Van Beek, L.P.H., Bierkens, M.F., Blyth, E., de Roo, A., Döll, P., Ek, M., Famiglietti, J. and Gochis, D.: Hyperresolution global land surface modeling: Meeting a grand challenge for monitoring Earth's terrestrial water. Water Resources Research, 47(5), 2011.

Wood, E.F., Sivapalan, M., Beven, K. and Band, L.: Effects of spatial variability and scale with implications to hydrologic modeling. Journal of hydrology, 102(1-4), pp.29-47, 1988.

Yamazaki, D., Ikeshima, D., Sosa, J., Bates, P.D., Allen, G. and Pavelsky, T.: MERIT Hydro: A high-resolution global hydrography map based on latest topography datasets. Water Resources Research, 2019.

Yassin, F., Razavi, S., Elshamy, M., Davison, B., Sapriza-Azuri, G. and Wheater, H.: Representation and improved parameterization of reservoir operation in hydrological and land-surface models. Hydrology and Earth System Sciences, 23(9), pp.3735-3764, 2019.

Yoon, J.-H.,Shoemaker, C. A.: Improved real-coded GA for groundwater bioremediation. Journal of Computing in Civil Engineering 15, 224-231, 2001.

Zhao, R.J.:. The xinanjiang model. In Proceedings of the Oxford Symposium, 1980.

Deleted: ¶
Tables¶
Table – 1 the number

Formatted: Indent: Left: 0 cm, First line: 0 cm

Moved up [19]: of computational units for the Bow River at Banff, given different spatial discretization of land cover, soil type, elevation zones and slope and aspects forced with various forcing resolutions. ¶

Deleted: Number of GRUs ... [22]

Moved up [18]: model parameters that are subjected to perturbation for model calibration for the designed experiments.¶
Parameter¶
symbol

Deleted: K_{sat} ... [23]