New methods for coupling climate-driven hydrology with hillslope and channel geomorphic processes at the watershed scale

Jeffrey Keck

A dissertation

submitted in partial fulfillment of the

requirements for the degree of

Doctor of Philosophy

University of Washington

2023

Reading Committee:

Erkan Istanbulluoglu, Chair

Jessica Lundquist

Alexander R. Horner-Devine

Program Authorized to Offer Degree:

Civil & Environmental Engineering

University of Washington

Abstract

New methods for coupling climate-driven hydrology with hillslope and channel geomorphic processes at the watershed scale

Jeffrey Keck

Chair of the Supervisory Committee:
Erkan Istanbulluoglu
Civil & Environmental Engineering

As human demand for natural resources grows, the historic hydrologic conditions that permitted certain slopes and channels to remain stable in the past are shifting. Consequently, relying on historic or observed data to inform management decisions, that may also affect slope and channel stability, is no longer reasonable and models that incorporate climate predictions are becoming increasingly necessary. Many numerical approaches for modeling watershed-scale sediment production and transport response to land use and climate already exist but they share similar shortcomings. This thesis improves hydrology-driven, watershed-scale sediment production and transport modeling methods and understanding. First, I examine hydrologic representation and its impact on modeled-network-scale sediment transport. Then, I develop a new landslide runout model, called MassWastingRunout, suitable for predicting probabilistic runout extent, sediment transport and topographic change. Finally, as part of a study on climate change impacts on

landslides, I develop a new method for coupling climate and hydrology to sediment production and transport models, called DistributedHydrologyGenerator. The new modeling techniques are coded in Python and implemented as components of the package Landlab. This thesis ends by synthesizing findings and tools from each section and briefly proposing a watershed-scale sediment production and transport modeling framework for future work.

TABLE OF CONTENTS

ACKNOWLEDGEMENTS

This is a Covid thesis. By that I mean, the Covid pandemic shaped much of my PhD experience and to a certain degree, the content of this thesis. I'm not thanking Covid, but I do want to acknowledge it. Of course, this thesis would not have been possible without the support of my family, especially Ichun and our two kids Max and Nolan, my parents, Jerry and Linda and my brothers, Daren and Mason. This thesis would not have been possible without support from my advisor Erkan Istanbulluoglu, who's flexibility allowed me to pursue grad school while working and training helped me improve.

The Covid pandemic began in late 2019, a little before the December, 2019 American Geophysical Union conference in San Francisco. A week after the conference and back home in Olympia, I developed chest pains and was diagnosed with pneumonia. I had planned field work but was told by a doctor to stay home and…."be careful, something's going around". I would later wonder, was AGU, 2019 the super spreader event that ignited the global pandemic?

When Covid was officially recognized in the US, federal and state governments responded by politicizing the virus. In some communities, people were told wearing masks wouldn't protect them. In other areas, outdoor parks were closed. I remember trying to take my family to the beach on the Olympic Peninsula in April, 2020. A patrol of four or five national park police, lights blazing, were hastily erecting barriers and signs along Highway 101 to prevent people from accessing the beach. If you've ever been to the beach on the Olympic Peninsula during non-summer periods, you know that you're often lucky to see a single person. I raced ahead of police to an area where they hadn't yet blocked off, parked and we ran down to the beach and hid. After about two hours of nervously trying to enjoy the beach, we slowly walked back to the car, checking to see if they were looking for us. They were gone, but our car had been encircled with flagging, signs and cones.

Back in Olympia, public schools were closed. Teachers tried to establish an online school system but it was largely ineffective. Kids sat idle. Brains turned off. Watching my kids attend online classes, you could see that many of their classmates were spending their days lying in bed. The teachers really tried their best to engage the kids, but it wasn't working. In Olympia, kids would ultimately be left that way for over a year. Ichun and I initially tried to homeschool during that period (for me, this was between my job and PhD research), but we were only marginally successful and eventually decided that for the 2020-2021 school year, we would move our kids to Taiwan. In Taiwan, schools were still open and Ichun had family, which allowed our kids to attend school there.

Taiwan very successfully managed Covid for about 15 months. Taiwan is an island, which perhaps made preventing mass infection more manageable, but basic science was never politicized like it was in the US and you could still go the park. They did close their beaches though, and in May, 2021, the Taiwan police caught us. This time, there was no flagging or signs, but they did warn that if we were caught again, we'd get a US $500 ticket. We found a new beach.

While the kids went to school, Ichun spent time with her family and I continued PhD research from Taipei. I still needed to return to the US to complete field work for my job but Taiwan required a strict 2-week police monitored quarantine for anyone entering the country. I ended up spending over a month in quarantine. I've never been able to focus like I did during those quarantines. I woke up and went to bed dreaming of my research, which at that time, was the initial conceptualization and coding of MassWastingRunout.

By late May, 2021, covid finally broke through Taiwan's defenses by a person the media dubbed the "Lion King". The Lion King was the head of a local division of the Lions Club. After visiting the airport and unknowingly contracting Covid, he attended a meeting and gave a speech in a small room packed with people. After the meeting, he then visited Taiwan's red-district in Taipei to have sexy tea. The pandemic spread from the members of the Lions Club and the red district. For the last 6 weeks of the 2020-2021 school year, public schools in Taipei were closed too and our kids were back to online school. At the same time, schools in Olympia were starting to open back up. That summer, we moved back to Olympia.

In Olympia, the kid's lives were returning to normal. I continued to work on my PhD, refining MassWastingRunout, surveying landslide sites in the field and developing the DistributedHydrologyGenerator. My first paper was also finally accepted after a year-long review. Later, during summer of 2022, my family returned to Taipei to visit Ichun's family and I went to Boulder, Colorado as an Open Earthscape visiting graduate student where I spent several months working around geomorphology and computational earth surface modeling pros and visiting researchers from all over the world. That was a great experience. Oddly, it was also the first time in over two years that I had worked around people.

Fast forward to May 2023, and our kids are both one foot taller than they were before the pandemic. We managed to (mostly) keep them in school. We weren't locked up for going to the beach and although slower than I would have liked, it appears I am finally going to graduate.

In addition to the people I have already thanked, I also need to thank the following: advisors Jessica Lundquist and Alex Horner-Devine. Much of my fluvial geomorphology know-how came from Brian Collins. Herve Capart at National Taiwan University helped with early conceptualization and evaluation of MassWastingRunout. Much of my DHSVM and model calibration skills were taught by Christina Bandaragoda. Climate change and hydrology support from Guillaum Mauger. LandslideProbability help from Ronda Strauch. Course work and DHSVM help from Claire Beveridge. DHSVM help and snow-hydrology instruction from Nicoleta Cristea. Coding inspiration and help from Jimmy Phuong, Sai Nudurupati and Amanda Manaster. Computational geomorphology inspiration and help from Greg Tucker, Benjamin Campforts, Albert Kettner, Marisa Repasch, Eric Hutton, Shelby Ahrendt and Mark Piper. Although the review process for my first paper was painful, I learned a great deal from John Buffington, Charlie Luce, David Gaeuman and Dieter Rickenmann. Github pushing and pulling with Allison Pfeiffer. Debris flow surveying with Eli Schwat. My neighbors Sam Merrill, Lee Antles and Alan Copsey who gave me an intense mock final exam. The Forest Resources Division of the WA Department of Natural Resources, in particular Casey Hanell, Allen Estep, Teddy Minkova, John Jenkins and Jenn Parker for supporting me while I juggled multiple workloads. Finally, I need to thank Kate Huntington and again thank Brian Collins and Allison Pfeiffer who also participated in my committee.

CHAPTER 1. INTRODUCTION

Growing demand for natural resources and a rapidly changing climate are making land management and infrastructure decisions increasingly dependent on modeled predictions (McKelvey et al., 2021). This is already observable in Washington State. For example, in many of the mountainous regions west of the Cascade Mountain range that have historically supported both timber production as well as ecosystems sensitive to changes in hydrology and sediment, model predictions suggest that in the near future many present-day timber production practices meant to protect habitat may no longer be adequate (Barik et al, 2017; Halofsky et al., 2011) and in some scenarios land managers are already actively updating their protocols (e.g., Wilhere et al., 2017).

Additionally, levees, which protect many urban centers from floods, are often designed based on historic channel flow and hydraulic conditions. The onset of rapid climate change may impact both the frequency of floods and sediment inputs near the levee and in turn increase the likelihood of levee failure (Vahedifard et al., 2020). In Washington, levees are now actively being upgraded to account for predicted changes in flood magnitude and frequency (Mauger et al., 2015).

Both the levee, which is typically located along a high order channel and the timber harvests, which often occur above the 1st order channels, are impacted by/or impact the movement of water and sediment. Both are intertwined in watershed-scale sediment production and transport processes, or sediment cascades (e.g., Burt and Allison, 2010). In a sediment cascade, climate drives surface runoff, saturated and unsaturated soil water flow and channelized flow. That movement of water in turn drives sediment production via surface erosion, landslides and channel erosion. Sediment production becomes the supply to the channels, where it is transported as bed, suspended and wash load at a rate determined by upslope hydrology and the transport capacity of the channel. Conceptual models like Lane's balance (Lane, 1955) help evaluate how channel

conditions might change given a likely change in sediment or water inputs, but in order to predict the magnitude and timing of those changes, numerical models are often needed.

Numerous numerical models exist for modeling sediment cascades. These models vary from (1) landscape evolution models suitable for exploratory studies (e.g., Murray 2007) to determine cause-and-effect impacts of changes in specific climate or threshold variables over large spatial and geologic times scales (Densmore et al., 1998; Campfort et al., 2020; Istanbulluoglu & Bras 2005; Tucker & Bras; 1998); (2) lumped, highly conceptualized models (Bennett et al. 2014; Beveridge et al. 2020) that couple a simplified representation of hillslope processes with a hydrology model or sediment transport model and (3) detailed, distributed approaches that explicitly represent both the sediment production and transport process across the watershed (e.g., Burton and Bathurst, 1998; DHSVM: Doten et al., 2006; tRIBS-Erosion: Francipane et al., 2012). This thesis improves on existing sediment-cascade modeling methods and understanding. In the following sections, I review key aspects and limitations of each modeling technique and conclude with a summary of the key contributions of this thesis.

1.1 LANDSCAPE EVOLUTION MODELS

Landscape evolution models are generally run over geomorphic time scales (100 to >10e6 years) and use a highly-simplified hydrology model, geomorphic transport rules or Geomrophic Transport Laws (sensu Dietrich et al., 2003; Tucker & Hancock, 2010) and mass continuity to drive sediment production and transport processes. Precipitation is often applied at a uniform rate across the landscape and flow at a grid cell is approximated as the precipitation rate times the upstream contributing area (Campforts et al., 2020; Densmore et al., 1998, Tucker & Slingerland, 1997). In some cases, a simple model for precipitation losses may be used (Tucker & Bras, 2000). Landslide processes are implicitly represented with detachment limited excess-shear stress GTLs

(Tucker & Bras, 2000) or explicitly represented with simple slope stability models (Istanbulluoglu and Bras, 2005; Istanbulluoglu et al., 2005). In some cases, both the initiation and runout of the landslide are represented (Campforts et al., 2020). The triggering event for the landslides is often based on a threshold slope (Campforts et al., 2020; Densmore et al., 1998; Istanbulluoglu & Bras, 2005) or critical geometry (Istanbulluoglu et al., 2005). Fluvial erosion and sediment transport are modeled as a function of excess shear stress or stream power. These models often include uplift and explore landscape topographic response to precipitation, vegetation and uplift rates. For example, Tucker and Slingerland (1997) used the landscape evolution model GOLEM to investigate climate change impacts on basin morphology (Figure 1a) and Istanbulluoglu and Bras (2005) used the landscape evolution model CHILD to explore vegetation impacts on watershed morphology (Figure 1b).

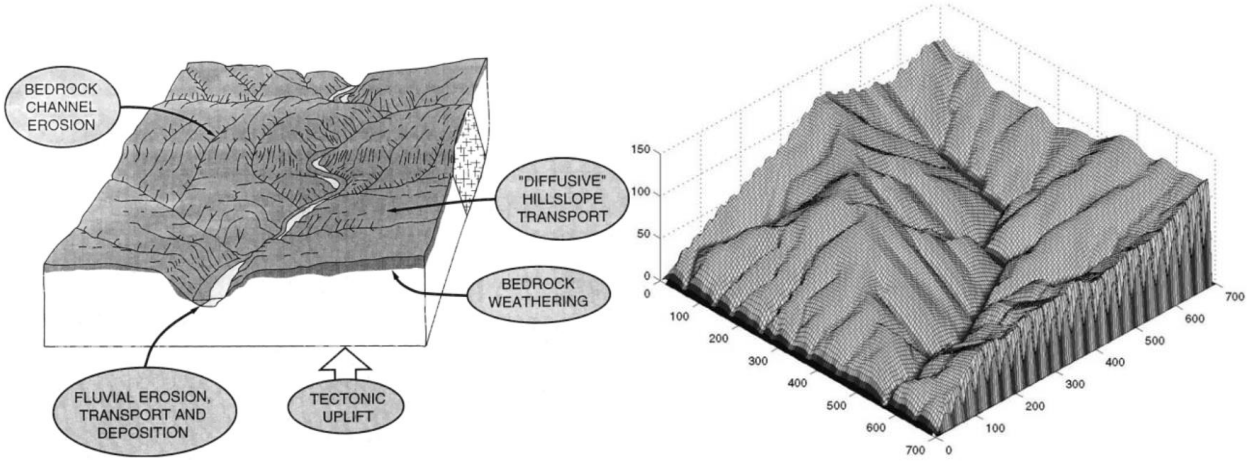

Figure 1. (a) Conceptual illustration of the model components included in the landscape evolution model GOLEM, from Tucker and Slingerland (1997), reproduction of their Figure 2; and (b) A modeled landscape that evolved in repsonse to vegetation-modulated landslide processes from Istanbulluoglu and Bras (2005), reproduction of their Figure 8.

1.2 LUMPED, HIGHLY CONCEPTUALIZED MODELS

Lumped and highly conceptualized models are intended for site specific, precipitation-event-scale to geomorphic-time-scale applications (days to thousands of years). A model by Bennet et al. (2014) consists of a chain of sediment reservoirs, each reservoir representing a link in a sediment cascade, fed by stochastic hillslope and water inputs (Figure 2). Other models consist of a detailed link-node representation of the channel network paired with a simple sediment production model. Often, the sediment sources are treated as black boxes, that release sediment as a function of a simple empirical equation (Murphy et al., 2019), randomly as a function of contributing area (Beveridge et al., 2020) or are simply a user input (Schmidtt et al., 2016), independent of climate and flow rate in the channel. Others explicitly represent a specific landslide process. For example in Benda and Dunn (1997a), sediment is derived from reservoirs at the tip of the channel network, that represent colluvium-filled topographic hollows that stochastically release debris as a function of sediment supply to the hollow and hydrologic conditions.

Transport capacity is generally determined using established transport formula (e.g., Wilcock & Crowe, 2003) or stochastically from an empirical PDF of annual maximum transport rates, scaled as a function of contributing area (Benda & Dunn, 1997b). In most models, actual transport rate varies as a function of transport capacity and sediment availability. Some models track both coarse and fine sediment transport using formula for bedload and suspended load (Benda & Dunne, 1997b; Bevridge et al. 2020; Schmidtt et al. 2016). Others are focused primarily on gravel bedded rivers and ignore suspended load (Czuba et al., 2018).

Despite the detailed representation of the channel network, many of the above models rely on crude estimates of channel hydraulics to force the sediment transport models. For example, in Murphy et al. (2019), they force their model with observed flow rates from the basin outlet, by

extrapolating the observed hydrograph to each upstream reach using a hydraulic geometry relationship. In Schmitt et al. (2016), they model sediment transport using specific quantile values of daily average flow recorded at flow gages and extrapolated to each upstream reach as a function of contributing area at the reach relative to the contributing area to the gage. In Beveridge et al. (2020), they use modeled daily flow rates from the Distributed Hydrology Soil and Vegetation model (DHSVM; Wigmosta et al. 1994), forced with daily average precipitation rates.

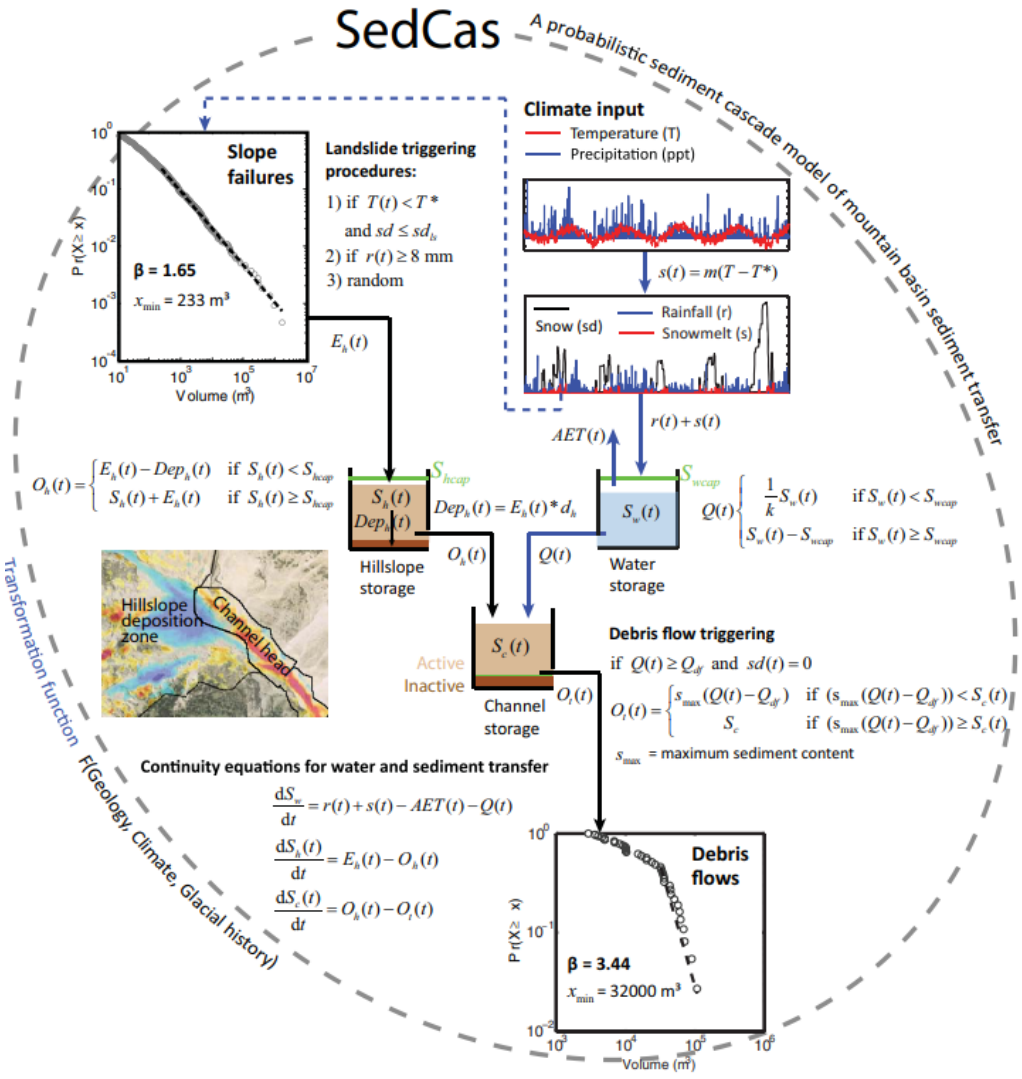

Figure 2. Model structure of SedCas, a sediment cascade model by Bennet et al. (2014) that conceptualizes a watershed as a series of interconnected, stochastically fed sediment and water reservoirs. Reproduction of their Figure 2.

1.3 DETAILED DISTRIBUTED APPROACHES

Detailed distributed approaches use physically based models that rely on mass, momentum and energy conservation and experimentally determined empirical equations to explicitly represent the movement of water and sediment at scales close to the real processes. They generally couple sediment production and transport models with a distributed hydrology model and attempt to represent hydrologic and sediment processes over the entire model domain. Many detailed distributed approaches can be parameterized from field measurements. Nonetheless, to a certain degree, calibration is still required because the real-world processes that actually control the movement of water (e.g. soil macropores, heterogenous soil types or channel hydraulics) and sediment are not fully represented in the model.

One of the first detailed distributed approaches was a model by Burton and Bathurst (1998). They added a landslide initiation and runout model to the sediment transport version of the distributed hydrology model SHE (Abbot et al., 1986a; 1986b) called SHETRAN (Evan et al., 1996). Hillslope and channel hydrology were modeled on a coarse grid and landslide initiation was modeled using a finer grid, with soil water hydrology interpolated from the coarser hydrology model to the finer grid using a topographic wetness index. Landslide initiation was modeled using the infinite slope model, landslide runout model consisted of a few slope-dependent rules and delivery of sediment to the channel was determined as a function of the distance from the beginning of deposition to the channel link.

Later, Doten et al. (2006) (Figure 3) added landslide, landslide runout, surface erosion and channelized sediment transport to the distributed hydrology model DHSVM using many of the methods from Burton and Bathurst (1998). Like Burton and Bathurst (1998), they computed

landslide initiation and runout on a finer grid scale than the distributed hydrology model and interpolate coarse-grid-scale hydrology to the fine-grid-scale model using a wetness index.

More recently, Francipane et al. (2012) coupled surface erosion and fluvial sediment transport model to the distributed hydrology model tRIBS. This model was then paired with a simple landslide and runout model by Arnone et al. (2011). Again, the runout model more or less consisted of the runout rules developed by Burton and Bathurst (1998).

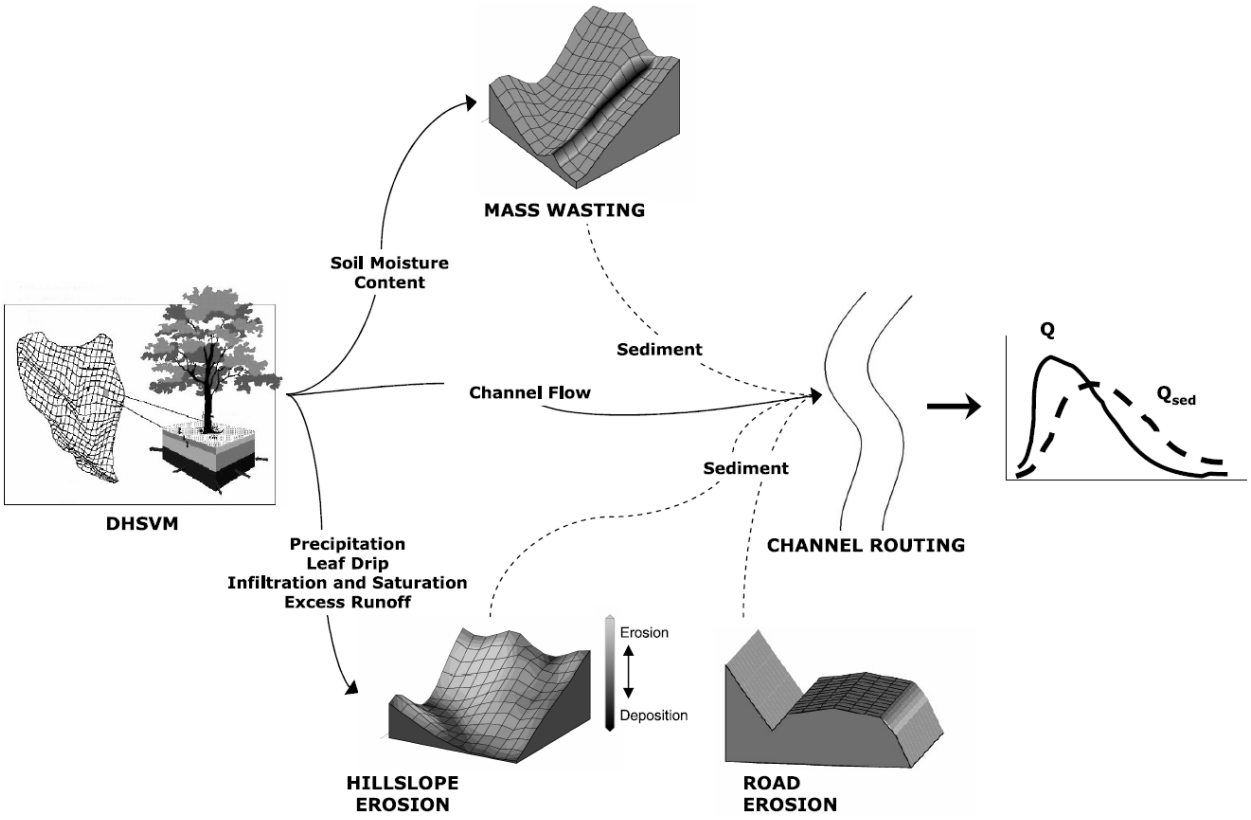

Figure 3. Conceptual diagram of the coupled sediment production, transport and hydrology model developed by Doten et al. (2006), reproduction of their Figure 1.

1.4 EXISTING MODEL LIMITATIONS AND THESIS CONTRIBUTION

Despite a diverse range of numerical methods for modeling sediment cascades, many of the methods share similar limitations. Many of the models rely on a simple representation of hydrology

or use a detailed hydrology model but force it with averaged meteorology data. Also, all of the models use a simple conceptualization of the landslide and sediment delivery process to the channel network. Finally, many of the approaches appear to be configured with a pre-determined hydrology model. For example, all of the detailed distributed models use a specific hydrology model.

In this thesis, I develop new methods for coupling climate-driven modeled hydrology with landslide and sediment transport models and investigate the importance of accurate hydrologic representation on modeled network-scale sediment transport. The new methods include a new model for predicting landslide runout, sediment transport and topographic change and a new program that ingests the raw output of a distributed hydrology model and converts it into a format that can be used to force any hydrologically driven model in the Landlab modeling framework (Barnhart et al., 2020). The overall contribution of this thesis is an advancement of climate-drive watershed-scale sediment production and transport modeling techniques as well as an improved understanding of data and methods needed for accurate sediment transport and landslide runout modeling.

This thesis is structured as follows: in the Chapter 2, I examine the sensitivity of a network-scale sediment transport model to the temporal representation of precipitation used to force the model as a function of location in the channel network and sediment availability. In Chapter 3, I develop the landslide runout model, called MassWastingRunout, which is a new method for routing landslides to the channel network that can be calibrated to observed runout deposits to accurately represent both the topographic change caused by the landslide as well as the redistribution of sediment. In Chapter 4, I develop the program for ingesting externally modeled hydrology to the grid scale of a landslide model, called DistributedHdyrologyGenerator, and use

it to couple DHSVM with a landslide model (LandslideProbability) to evaluate climate change impacts on landslide rates in the Skagit watershed. The final chapter ends with a brief summary of the main contributions of each chapter and a proposal for combining all of the new tools and insights into a new climate-driven, watershed-scale sediment production and transport modeling framework, that improves on existing methods for modeling the entire sediment cascade.

1.5 References

1. Abbott, M. B., Bathurst, J. C., Cunge, J., O'Connell, P., & Rasmussen, J. O. (1986a). An introduction to the European Hydrological System — Systeme Hydrologique Europeen, "SHE", 1: History and philosophy of a physically-based, distributed modelling system. *Journal of Hydrology*, *87*(1–2), 45–59. https://doi.org/10.1016/0022-1694(86)90114-9
2. Abbott, M. B., Bathurst, J. C., Cunge, J., O'Connell, P., & Rasmussen, J. O. (1986b). An introduction to the European Hydrological System — Systeme Hydrologique Europeen, "SHE", 2: Structure of a physically-based, distributed modelling system. *Journal of Hydrology*, *87*(1–2), 61–77. https://doi.org/10.1016/0022-1694(86)90115-0
3. Arnone, E., Noto, L., Lepore, C., & Bras, R. L. (2011). Physically-based and distributed approach to analyze rainfall-triggered landslides at watershed scale. *Geomorphology*, *133*(3–4), 121–131. https://doi.org/10.1016/j.geomorph.2011.03.019
4. Barik, M. G., Adam, J., Wein, A. J., & Muhunthan, B. (2017). Improved landslide susceptibility prediction for sustainable forest management in an altered climate. *Engineering Geology*, *230*, 104–117. https://doi.org/10.1016/j.enggeo.2017.09.026
5. Benda, L., & Dunne, T. (1997a). Stochastic forcing of sediment routing and storage in channel networks. *Water Resources Research*, *33*(12), 2865–2880. https://doi.org/10.1029/97wr02387
6. Benda, L., & Dunne, T. (1997b). Stochastic forcing of sediment supply to channel networks from landsliding and debris flow. *Water Resources Research*, *33*(12), 2849–2863. https://doi.org/10.1029/97wr02388
7. Bennett, G. L., Molnar, P., McArdell, B. W., & Burlando, P. (2014). A probabilistic sediment cascade model of sediment transfer in the Illgraben. *Water Resources Research*, *50*(2), 1225–1244. https://doi.org/10.1002/2013wr013806
8. Beveridge, C., Istanbulluoglu, E., Bandaragoda, C., & Pfeiffer, A. M. (2020). A Channel Network Model for Sediment Dynamics Over Watershed Management Time Scales. *Journal of Advances in Modeling Earth Systems*, *12*(6). https://doi.org/10.1029/2019ms001852
9. Burt, T., & Allison, R. R. (2010). Sediment cascades : an integrated approach. In *Wiley-Blackwell eBooks*. http://ci.nii.ac.jp/ncid/BB01752261
10. Burton, A., & Bathurst, J. C. (1998). Physically based modelling of shallow landslide sediment yield at a catchment scale. *Environmental Geology*, *35*(2–3), 89–99. https://doi.org/10.1007/s002540050296
11. Campforts, B., Shobe, C. M., Steer, P., Vanmaercke, M., Lague, D., & Braun, J. (2020). HyLands 1.0: a hybrid landscape evolution model to simulate the impact of landslides and

landslide-derived sediment on landscape evolution. *Geoscientific Model Development*, *13*(9), 3863–3886. https://doi.org/10.5194/gmd-13-3863-2020
12. Czuba, J. A. (2018). A Lagrangian framework for exploring complexities of mixed-size sediment transport in gravel-bedded river networks. *Geomorphology*, *321*, 146–152. https://doi.org/10.1016/j.geomorph.2018.08.031
13. Densmore, A. L., Ellis, M. A., & Anderson, R. H. (1998). Landsliding and the evolution of normal-fault-bounded mountains. *Journal of Geophysical Research*, *103*(B7), 15203–15219. https://doi.org/10.1029/98jb00510
14. Dietrich, W. E., Bellugi, D., Sklar, L. S., Stock, J. P. J., Heimsath, A. M., & Roering, J. J. (2013). Geomorphic Transport Laws for Predicting Landscape form and Dynamics. In *Geophysical monograph* (pp. 103–132). American Geophysical Union. https://doi.org/10.1029/135gm09
15. Doten, C. W., Bowling, L. C., Lanini, J., Maurer, E. P., & Lettenmaier, D. P. (2006). A spatially distributed model for the dynamic prediction of sediment erosion and transport in mountainous forested watersheds. *Water Resources Research*, *42*(4). https://doi.org/10.1029/2004wr00382
16. Ewen J. (1995). Contaminant transport component of the catchment modelling system (SHETRAN). In: Trudgill ST (ed) *Solute modelling in catchment systems*. Wiley, Chichester, UK, pp 417–441
17. Halofsky, J. E., Peterson, D. L., O'Halloran, K. D., & Hoffman, C. (2011). *Adapting to climate change at Olympic National Forest and Olympic National Park*. https://doi.org/10.2737/pnw-gtr-844
18. Istanbulluoglu, E. (2005a). Implications of bank failures and fluvial erosion for gully development: Field observations and modeling. *Journal of Geophysical Research*, *110*(F1). https://doi.org/10.1029/2004jf000145
19. Istanbulluoglu, E. (2005b). Vegetation-modulated landscape evolution: Effects of vegetation on landscape processes, drainage density, and topography. *Journal of Geophysical Research*, *110*(F2). https://doi.org/10.1029/2004jf000249
20. Lane, E. W. (1955). The Importance of Fluvial Morphology in Hydraulic Engineering. *Proceedings of the American Society of Civil Engineers*, *81*(7), 1–17. https://cedb.asce.org/CEDBsearch/record.jsp?dockey=0353595
21. Mauger, G. S., Casola, J. H., Morgan, H., Strauch, R. L., Jones, B. L., Curry, B., Isaksen, T. B., Binder, L. W., Krosby, M., & Snover, A. (2015). State of Knowledge: Climate Change in Puget Sound. *Climate Impacts Group, University of Washington*. https://digital.lib.washington.edu:443/researchworks/handle/1773/34347
22. McKelvey, K. S., Block, W. M., Jain, T. B., Luce, C. H., Page-Dumroese, D. S., Richardson, B. A., Saab, V. A., Schoettle, A. W., Sieg, C. H., & Williams, D. R. (2021). Adapting Research, Management, and Governance to Confront Socioecological Uncertainties in Novel Ecosystems. *Frontiers in Forests and Global Change*, *4*. https://doi.org/10.3389/ffgc.2021.644696
23. Murphy, B., Czuba, J. A., & Belmont, P. (2019). Post-wildfire sediment cascades: A modeling framework linking debris flow generation and network-scale sediment routing. *Earth Surface Processes and Landforms*, *44*(11), 2126–2140. https://doi.org/10.1002/esp.4635
24. Murray, A. B. (2007). Reducing model complexity for explanation and prediction. *Geomorphology*, *90*(3–4), 178–191. https://doi.org/10.1016/j.geomorph.2006.10.020
25. Schmitt, R., Bizzi, S., & Castelletti, A. (2016). Tracking multiple sediment cascades at the river network scale identifies controls and emerging patterns of sediment connectivity. *Water Resources Research*, *52*(5), 3941–3965. https://doi.org/10.1002/2015wr018097

26. Tucker, G. E., & Slingerland, R. (1997). Drainage basin responses to climate change. *Water Resources Research*, *33*(8), 2031–2047. https://doi.org/10.1029/97wr00409
27. Tucker, G. E., & Bras, R. L. (1998). Hillslope processes, drainage density, and landscape morphology. *Water Resources Research*, *34*(10), 2751–2764. https://doi.org/10.1029/98wr01474
28. Tucker, G. E., & Bras, R. L. (2000). A stochastic approach to modeling the role of rainfall variability in drainage basin evolution. *Water Resources Research*, *36*(7), 1953–1964. https://doi.org/10.1029/2000wr900065
29. Tucker, G. E., & Hancock, G. J. (2010). Modelling landscape evolution. *Earth Surface Processes and Landforms*, *35*(1), 28–50. https://doi.org/10.1002/esp.1952
30. Vahedifard, F., Jasim, F. H., Tracy, F. T., Abdollahi, M., Alborzi, A., & AghaKouchak, A. (2020). Levee Fragility Behavior under Projected Future Flooding in a Warming Climate. *Journal of Geotechnical and Geoenvironmental Engineering*, *146*(12). https://doi.org/10.1061/(asce)gt.1943-5606.0002399
31. Wigmosta, M. S., Vail, L. W., & Lettenmaier, D. P. (1994). A distributed hydrology-vegetation model for complex terrain. *Water Resources Research*, 30(6), 1665–1679. https://doi.org/10.1029/94wr00436
32. Wilcock, P. R., & Crowe, J. C. (2003). Surface-based Transport Model for Mixed-Size Sediment. *Journal of Hydraulic Engineering*, *129*(2), 120–128. https://doi.org/10.1061/(asce)0733-9429(2003)129:2(120
33. Wilhere, G. F., Atha, J. B., Quinn, T. J., Tohver, I., & Helbrecht, L. (2017). Incorporating climate change into culvert design in Washington State, USA. *Ecological Engineering*, *104*, 67–79. https://doi.org/10.1016/j.ecoleng.2017.04.009

CHAPTER 2. HOW DOES PRECIPITATION VARIABILITY CONTROL BEDLOAD RESPONSE ACROSS A MOUNTAINOUS CHANNEL NETWORK IN A MARITIME CLIMATE?

2.0 ABSTRACT

Modeled stream discharge is often used to drive sediment transport models across channel networks. Because sediment transport varies non-linearly with flow rates, discharge modeled from daily total precipitation distributed evenly over 24-hrs may significantly underestimate actual bedload transport capacity. In this study, we assume bedload transport capacity determined from a hydrograph resulting from the use of hourly (1-h) precipitation is a close approximation of actual transport capacity and quantify the error introduced into a network-scale bedload transport model driven by daily precipitation at channel network locations varying from lowland pool-riffle channels to upland colluvial channels in a watershed where snow accumulation and melt can affect runoff processes. Transport capacity is determined using effective stresses and the Wilcock and Crowe (2003) equations and expressed in terms of transport capacity normalized by the bankfull value. We find that, depending on channel network location, cumulative error can range from 10 - 20% to more than two orders of magnitude. Surprisingly, variation in flow rates due to differences in hillslope and channel runoff do not seem to dictate the network locations where the largest errors in predicted bedload transport capacity occur. Rather, spatial variability of the magnitude of the effective-bankfull-excess shear stress and changes in runoff due to snow accumulation and melt exert the greatest influence. These findings have implications for flood-hazard and aquatic habitat models that rely on modeled sediment transport driven by coarse-temporal-resolution climate data.

2.1 INTRODUCTION

Key to predicting how a river will respond to a given flow rate is predicting the channel conveyance at the time of the flood (Sturm, 2010). In mountainous watersheds, channel conveyance is maintained by the capacity of the channel to move bedload. If bedload transport capacity falls below the supply rate, the channel fills with sediment and channel conveyance decreases until consummate increases in channel slope balance transport and supply rates (Lane, 1955).

At gaged locations, an estimate of bedload transport capacity under historic hydrologic conditions can be determined from a survey of the reach geometry, measurement of bed surface grain size and the hydrograph (Wilcock et al., 2009). For ungauged locations or future climate scenarios, a hydrograph must be approximated. If the hydrograph is modeled from precipitation derived from observation networks like the US National Weather Service Cooperative Observer Network (NWS, 2020) and Livneh et al. (2013) dataset, or a future climate meteorology dataset such as the Salathé et al. (2014) dataset, the precipitation may largely be recorded as daily average values.

The frequency of precipitation observations necessary to accurately model floods was described decades ago (Bras, 1979; Eagleson & Shak, 1966; Singh, 1997) but time series of daily average precipitation are still commonly used to model hydrologic response to climate change (Dan et al., 2012; Shrestha et al., 2012), with precipitation assumed to fall at a constant rate over the day for hourly hydrologic simulations. As the understanding of hydrologic processes has improved, detailed physical models have confirmed the necessity of accurate temporal representation of precipitation for flood prediction at the basin outlet (Paschalis et al., 2014). Furthermore, over geologic timescales, precipitation variability has been shown to control erosion

rates and channel morphology of modeled landscapes (Istanbulluoglu et al., 2005; Solyom & Tucker, 2004; Tucker & Bras, 2000); not representing it therefore risks removing a key control on channel response to precipitation.

Recently, studies began to explore modeled differences in network-scale hydrologic response driven by temporal representations of precipitation data (e.g., Wehner et al., 2021); however network-scale differences in sediment transport remain unknown. At the reach scale, analyses of modeled bedload sensitivity to the temporal representation of measured flow rates have shown that differences can be as high as several orders of magnitude (Chen et al., 2011; Rosburg et al., 2016). In this study, we ask:

1. what is the magnitude of modeled bedload transport capacity error caused by using daily (24-h) rather than temporally accurate (1-h) precipitation data to drive streamflow and bedload?
2. how does that error vary across a channel network as a result of fluvial geomorphic, hydrologic, and storm hydrometeorologic conditions, and are there any associated topographic thresholds that amplify errors?

Errors in modeled bedload transport capacity can result from uncertainties in the critical or reference shear stress of the bed material (the shear stress at which small, but measurable transport of the entire bed mixture begins), one's choice of transport equation, and uncertainties in streamflow discharge (Wilcock et al., 2009; Barry et al., 2007; Yager et al., 2012). We expect the error caused by flow rates driven by a 24-h hyetograph would systematically vary across a channel network as a function of spatially varying watershed-scale runoff processes as well as antecedent wetness conditions. For example, runoff rates vary between the snowpack, hillslopes and channels (D'Odorico & Rigon, 2003; Lundquist et al., 2005; Penna et al., 2011; Rinaldo et al., 1991; Rinaldo

et al., 1995; Robinson et al., 1995) and with antecedent soil water and flow conditions (e.g., Asano & Uchida, 2018; Dunne & Black, 1970; Lundquist et al., 2005; McGlynn et al., 2004; McGuire & McDonnell, 2010). Saturated hillslope conditions permit rapid runoff and the contribution of runoff from hillslopes is higher in lower order channels (McGlynn et al., 2004; McGuire & McDonnell, 2010; Penna et al., 2011). At any given reach of a watershed, flow response to precipitation depends not only on the magnitude of the precipitation event but also on the antecedent hydrologic conditions, the proportion of the runoff path via hillslopes versus channels, as well as the state of snow in the watershed.

We anticipate that runoff response in low-order channels will more closely reflect the 24-h hyetograph during heavy precipitation events or events preceded by heavy antecedent precipitation causing larger error. Furthermore, analogous to the contributing area threshold representing the transition from channel-controlled to hillslope-controlled runoff response suggested by Robinson et al. (1995) and McGlynn et al. (2004), we suspect the channel order or location in the channel network at which bedload transport capacity error becomes large may correspond to some contributing area threshold.

To address our study questions, we conduct our study in the Sauk River watershed, a 1896 km^2 mountainous, alluvial watershed in Washington State, USA. We model snow and rainfall runoff processes using the Distributed Hydrology Soil Vegetation Model (DHSVM; Wigmosta et al., 1994) forced by an hourly, 6 km resolution, 35-year-long, modeled meteorology dataset produced by the Pacific Northwest National Laboratory (PNNL) using the Weather Research Forecast model (Chen et al., 2018; described below and herein referred to as PNNL WRF, see section 3.1. for details). PNNL WRF includes 55 pseudo-modeled weather stations distributed across the watershed (Figure 1).

We determine bedload transport capacity using the Wilcock and Crowe (2003) equations with effective shear stress (portion of the total shear stress exerted by the flow on the grains) following Schneider et al. (2015). To reduce the number of parameters needed to model bedload transport capacity across the channel network, we develop a nondimensional bedload transport equation as the ratio of bedload transport capacity of a given flow to capacity at bankfull flow (section 3.2). Assuming bedload transport capacity modeled from flow driven by a 1-h hyetograph closely approximates actual transport capacity (i.e., small error), we quantify the bedload transport capacity error caused by using a 24-h hyetograph to drive flows as the ratio of cumulative 1-h bedload transport capacity to cumulative 24-h bedload transport capacity. This ratio serves as an index of hydrometeorology-driven bedload transport capacity error, or simply hydrometeorologic bedload error (HBE), caused by using hydrographs resulting from daily precipitation data. We relate HBE to several other indices that represent the variabilities in precipitation intensity and peak streamflow discharge, as well as normalized 1-h streamflow to bankfull discharge to investigate the varying controls on HBE across the watershed (section 3.3).

Figure 2 illustrates how hydrologic processes across the watershed might impact HBE in an example precipitation event selected from the PNNL-WRF data, used to drive DHSVM. Streamflow hydrographs are plotted relative to the reference flow rate (Q_r) which is the flow rate that corresponds to the reference shear stress of the bed material. The cumulative bedload transport capacity of a flow event in gravel-bedded channels is a function of the shear stress exerted by the flow relative to the reference shear stress (Costa & O'Conner, 1995; Phillips et al., 2018) and can be inferred from flow rates above the reference flow. The storm consists of a single, intense (high precipitation rate) but short-duration (<1 day) burst of precipitation. At the outlet reach, runoff from the upstream channel is delayed and attenuated. In contrast, at the headwater reach, runoff

response is initially synchronous with precipitation and ends rapidly. A simple comparison of runoff response greater than Q_r reveals that at both the headwater and outlet reaches, the duration that the 1-h hydrograph exceeds Q_r is roughly 60% of the duration the 24-h hydrograph exceeds Q_r; however the difference between the 1-hr peak flow rate and Q_r is over four times larger than that of the 24-hr peak flow rate at the headwater reach and only two times larger than that of the 24-h peak flow rate at the outlet reach. If we assume similar bedload transport capacity response above the reference flow rate at the two reaches, during this particular storm, HBE was considerably higher at the headwater reach.

This paper is structured as followed: in the results (section 4), we first present hydrological and the nondimensional bedload model calibration. We then divide our analyses into three parts. First, we detail the hydrologic processes that drive the response to 1-h and 24-h hydrographs across the watershed and resultant HBEs of three hydrometeorologically unique precipitation events that are typical in the region: an extreme, rain-dominated atmospheric river event (Storm I), a rain-and-snow-accumulation event (Storm II), and a rain-plus-snowmelt event (Storm III). The first two events are commonly observed in late fall and early winter, the last event represents storms during the spring snowmelt season. Second, we calculate HBEs across the watershed using all storm events extracted from a 35-year-long distributed hydrologic model simulation and investigate the sensitivity of HBE to hydrometeorologic conditions to infer which storm types require high-frequency representation of precipitation intensity in space and time to accurately drive bedload transport. Third, we compare probability distributions of cumulative nondimensional bedload transport as a function of flow magnitude and location in the channel network to infer sensitivity of HBE to flow magnitude. The main findings from these analyses are further discussed in section

5, where we focus on the sensitivity of HBE to watershed hydrologic response and fluvial geomorphologic conditions used to characterize bed mobility.

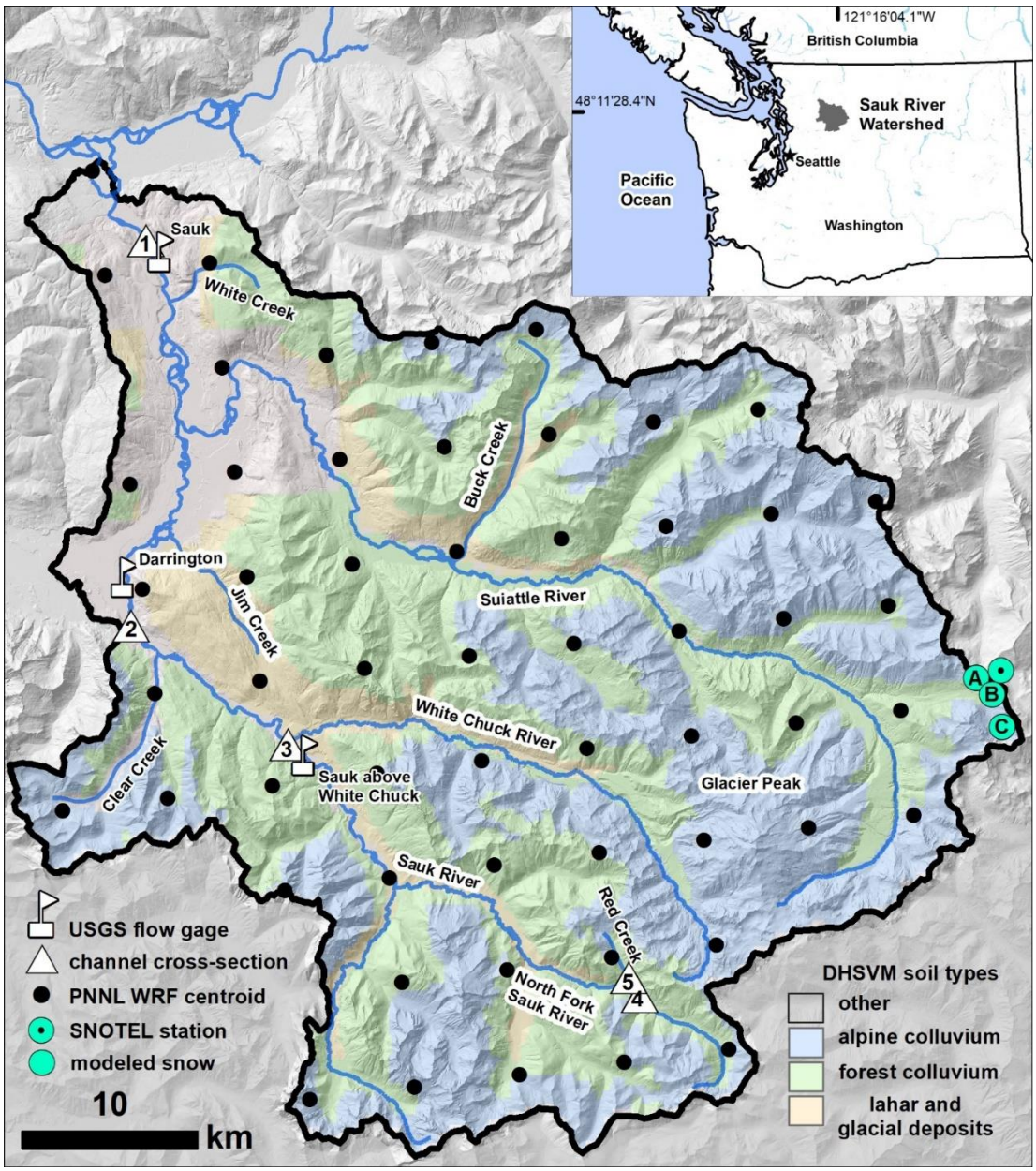

Figure 1. Sauk River Watershed. DHSVM was forced with PNNL WRF. DHSVM calibration was evaluated at the Sauk (Sauk at Sauk) and Sauk above White Chuck USGS gages and the SNOTEL station. Modeled flow response was adjusted by changing the hydrologic properties of the alpine colluvium, forest colluvium and lahar and glacial deposits soil layers. Observed effective-bankfull-excess shear stress (τ_b'/τ_r') was determined based on grain-size distributions at cross sections 1 through 5.

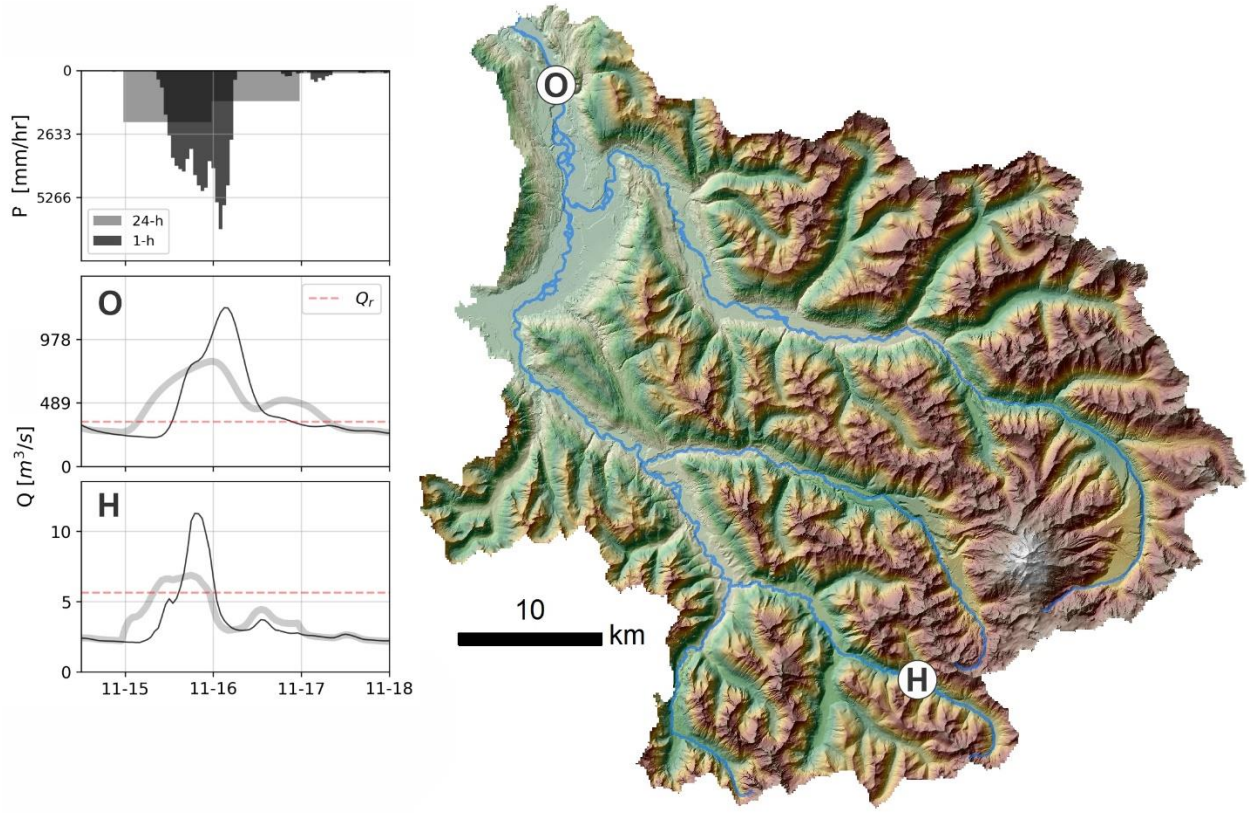

Figure 2. Illustration of 1-h and 24-h hyetographs of a rain-dominated precipitation event and their resultant modeled streamflow hydrographs at the basin outlet (O) and a headwater (H) reach in the Sauk River Basin, WA. Reference flow rate (Q_r) represents the discharge required to mobilize the channel bed at respective locations, estimated from field measurements of bed material.

2.2 Study Region

In the 1896 km2 Sauk River basin, elevation ranges from 70 m at the outlet to 3200 m at Glacier Peak, which is a small stratovolcano located in the headwaters of the Suiattle tributary to the Sauk river. The longest channel is roughly 90 km long. Average channel slopes range from < 0.5% in the valleys to > 40% in the colluvial headwater channels. In terms of the Montgomery and Buffington (1997) channel types, the lowland alluvial valleys primarily contain actively migrating pool-riffle channels and the upland channels consist of single-thread, step-pool, cascade and

colluvial channels. The bankfull flow, which in this region is equivalent to the 1.2-year return interval flow (Castro & Jackson, 2001), is 1060 m3 s-1 at the outlet. Suspended sediment transport occurs during 0.25-year and larger flow events (Jaeger et al., 2017).

The basin is predominately underlain by uplifted low- and high-grade metamorphic rock (Tabor et al., 2002). In areas of the basin underlain by low-grade metamorphic rocks, shallow landslides provide the primary source of sediment. Areas underlain by high-grade metamorphic bedrock have fewer landslides but receive sediment inputs from the proglacial zones on Glacier Peak (Jaeger et al., 2017). In the lowlands, bedrock is covered by continental glacial drift, lahar and alluvial deposits (Beechie et al., 2001). Clasts in the glacial drift and bedrock mass wasting deposits are relatively high strength and resistant to abrasion (Scott & Collins, 2021); over the entire channel network, grains consist almost entirely of coarse gravel. Cobbles and boulders are present in the upland channels.

The Sauk River basin has a maritime climate. Average annual precipitation ranges from approximately 1,500 mm/yr in the valleys and 3,300 mm/yr on the ridges (Livneh et al., 2013). The largest precipitation events are caused by atmospheric-rivers, which are long-duration (1 to 3 day), relatively constant-intensity events (up to 15-20 mm/hr) that result when long, narrow plumes of tropical water vapor are brought north from the equatorial region of the planet by extratropical cyclones (Neiman et al., 2011). Temperatures in the valleys are generally high enough that most precipitation falls as rain. On the ridges and Glacier Peak, precipitation often falls as snow and does not melt until warmer temperatures arrive in the spring. High flows are common during a heavy precipitation period between November to January, and the snowmelt period between May to July (Jaeger et al., 2017). About 2% of the landscape is perennially covered with snow. Less

than 1% of the basin is presently glaciated. In adjacent rivers, glacial melt on average accounts for 5 to 12% of summer low flows when snowmelt and rain inputs are minimal (Frans et al., 2018).

2.3 METHODS

In this section we present details on the distributed hydrology model (3.1); derivation of a new nondimensional bedload equation we used in an index of hydrometeorologic bedload error (HBE), and estimation of its parameters at five locations across the Sauk channel network (3.2); details of HBE and indices for precipitation intensity and streamflow variability which we relate to HBE to study its variability across the study watershed (3.3).

2.3.1. Distributed hydrology and flow routing models

We describe the distributed hydrology model by beginning with the PNNL WRF weather forcing. The centroids of the PNNL WRF grid cells are mapped on the Sauk River Watershed (Figure 1). PNNL WRF is a dynamically downscaled version of the North American Regional Reanalysis data (NARR, Mesinger et al., 2006) for the western part of the United States. NARR is a high-temporal and spatial resolution atmospheric and land surface hydrology dataset for North America. PNNL downscaled NARR using WRF version 3.1 (Skamarock et al., 2008). PNNL WRF extends from January, 1st 1981 to December, 31st, 2015 (35 years). We created the 24-h PNNL WRF hyetograph by replacing each 1-hour average precipitation depth with a daily average value (aggregated 24-h precipitation divided by 24). All other meteorologic variables in the 24-h PNNL WRF dataset (temperature, radiation, wind) are the same as the 1-h PNNL WRF dataset. We found that near the Sauk basin, PNNL WRF has a 3 to 5 degree C° cold bias during the spring and 1 to 2 C° warm bias during the summer relative to observed temperatures (Minder et al., 2010; NWS, 2020). Since snow accumulation and melt are sensitive to temperature and exert strong control on

spring and early summer flows in the upland channels of the Sauk watershed, we corrected the temperature field. Further details on PNNL WRF and the temperature bias correction can be found in the Supporting Information, sections 3 and 4.

A map of the soil types used in our DHSVM representation of the Sauk watershed are shown in Figure 1. DHSVM tracks snow pack, soil moisture and shallow water table states and models evapotranspiration, snowmelt, surface runoff and lateral subsurface fluxes on a gridded modeling scheme. Hillslope runoff processes that contribute to rapid response like return flow, infiltration-excess and saturation-excess flow (Dunne & Black, 1970) are included in DHSVM while other processes such as translatory and macropore flow (McGuire & McDonnell, 2010) are not (Wigmosta et al., 2002). Channel flow and individual reaches are represented by a network of inter-connected links (Wigmosta et al., 1994). Flow through each link (reach) is modeled as a linear reservoir. Incoming and outgoing flow rates are computed, including lateral flow from adjacent hillslopes (Wigmosta & Perkins, 2001). Many studies have successfully calibrated DHSVM to both small (Du et al., 2014; Surfleet et al., 2010; Waichler et al., 2005) and large, bedrock watersheds in the Western US (Cristea et al., 2013; Frans et al., 2018; Storck et al., 1998).

We used a 150-m grid resolution digital elevation model (DEM), aggregated from a 10- m DEM, to represent topography in DHSVM. A 150 m grid cell was chosen because it captured most topographic features and reduced model time to a practical duration. Regolith, the layer of all unconsolidated earth materials above the solid, impermeable bedrock surface; herein referred to as “soil” to be consistent with Wigmosta et al. (1994) and Beven, (1982), was approximated as a function of contributing area and slope (Westrick, 1999). A contributing area threshold of 10 ha was used to extract the channel network. This threshold was identified based on analysis of the 10-m DEM following concepts outlined by Tarboton et al. (1991) and Montgomery and Foufoula-

Georgiou (1993) and visual inspection of where concave profiles transition to convex, hillslope profiles in the 1-m LiDAR DEM of the watershed.

We parameterized the linear-reservoir representation of each channel reach using a constant flow width and depth value from hydraulic geometry relations for bankfull flow based on Castro and Jackson (2001) and a constant roughness value using an empirical formula for roughness determined from the three USGS gages shown in Figure 2 (detailed in the Supporting Information, section 5). Vegetation, snow, soil and glacier parameters were defined using sources listed in Table S2. Soil types were defined based on soil classification maps of the Sauk River (SSURGO, 2018). Detailed soil mapping was available for the north edge of the watershed. For the rest of the watershed, only coarse soil mapping was available. Since the coarse soil maps likely aggregate multiple soil types and may less accurately represent the spatial distribution of the soil type, parameters of the soils represented by the coarse soil maps were adjusted as part of DHSVM calibration.

We used the differential evolution adaptive Metropolis (DREAM) method for running multiple Markov chains in parallel (Vrugt & Ter Braak, 2011; Vrugt, 2016) to calibrate DHSVM. Calibration was quantitatively evaluated by comparing the modeled hydrograph to the observed hydrograph using the Nash-Sutcliffe efficiency metric (N-S; Nash & Sutcliffe, 1970) at the basin outlet and at a mid-basin location (cross sections 1 and 3; Figure 1). We used a single year due to the high computational overhead of DHSVM and the DREAM method which required running thousands of model iterations. Since PNNL WRF is a modeled meteorological dataset, it does not always match observed meteorological conditions. Water year 2009 was selected as the calibration year because the observed flow duration curve for water year 2009 best matched the flow duration curve of the entire observed flow record. Additionally, the PNNL daily accumulated precipitation

magnitude for water year 2009 was within 5% of that estimated from the observation-derived Livneh et al. (2013) gridded meteorology dataset. We manually evaluated DHSVM modeled Snow Water Equivalent (SWE) at locations A, B and C to SNOTEL observations for the best model runs (NRCS, 2020; Figure 2). Further details on calibration methods are included in the Supporting Information, section 6.

Our DHSVM representation of the Sauk River channel network included over 9,000 reaches. Modeled hydrographs from 1-h and 24-h precipitation hyetographs were computed for each reach. Due to memory and computation-time limitations, we saved flow data at only a subset of 582 reaches. The 582 reaches were selected at regular intervals along the main stem and larger tributaries of the Sauk River as well as randomly selected from low-order reaches located throughout the basin. The subset represents a range of channel types (colluvial to pool-riffle) and precipitation zones within the Sauk River watershed.

2.3.2. Nondimensional bedload transport equation and its parameter estimation

2.3.2.1. Nondimensional bedload equation

Throughout this paper, we express bedload transport capacity in terms of the following ratio:

$$\psi = \frac{q_s}{q_{sb}} \tag{1}$$

The variables q_s and q_{sb} are, respectively, the bedload transport capacity per unit channel width at some flow and the bankfull flow. Expressing bedload transport as this ratio has several advantages for modeling bedload transport across the channel network.

We define ψ from the high-shear-stress form of the Wilcock and Crowe (2003) equations for total transport rates as presented by Schneider et al. (2015). Below, we provide a brief summary

of our derivation of ψ. A detailed description is included in the Supporting Information, sections 1 and 2.

Total, volumetric bedload transport capacity per unit width is defined as:

$$q_s = \frac{W^*}{(s-1)g\rho^{1.5}} \tau^{1.5} \quad (2)$$

where s is specific gravity of the grains, g is gravity, ρ is the density of water, τ is total shear stress and W^* is the non-dimensional bedload transport capacity. Normalizing (2) by the bankfull value gives:

$$\psi = \frac{W^*}{W_b^*} \left(\frac{\tau}{\tau_b}\right)^{1.5} \quad (3)$$

Since no single equation describes both high- and low-shear-stress transport capacity, transport capacity functions like W^* are typically fit by multi-part equations (e.g., Parker et al., 1982; Wilcock & Crowe, 2003). We are primarily interested in tracking bedload transport capacity error during floods, when flow rates equal or exceed the reference flow rate. We therefore define (3) using the high-shear-stress form of W^*:

$$W^* = 14 \left(1 - \frac{0.894}{\phi^{0.5}}\right)^{4.5} \quad (4)$$

where $\phi = \tau/\tau_r$ (transport stage) and τ_r is the reference shear stress. Note that in the total transport rate form of the Wilcock and Crowe (2003), τ_r represents the reference shear stress of the median grain size. Since we are interested in mountainous channels, following Schneider et al. (2015), shear stresses herein use the effective shear stress and are denoted as τ'. Recasting τ' and τ_b' in terms of unit flow rate following Ferguson (2007) and Rickenmann and Recking (2011) and substituting (4) into (3), ψ can be expressed as (see Supporting Information, section 2, for full derivation):

$$\psi = \left(\frac{\left(\frac{\tau'_b}{\tau'_r} \left(\frac{Q}{Q_b} \right)^Y \right)^\alpha - B}{\left(\frac{\tau'_b}{\tau'_r} \right)^\alpha - B} \right)^\beta \left(\frac{Q}{Q_b} \right)^{Y(1.5-\alpha\beta)} \quad (5)$$

where Y is:

$$Y = \mathrm{C}(\mathrm{G} + 2.25) - 0.9(1 - m_w) \quad (6)$$

and C and G are power constants for discharge and velocity ratios for a given flow to bankfull flow. m_w accounts for changes in channel width with depth. In a rectangular channel, $m_w = 0$. Based on stage-discharge analysis at the first three gage locations shown in Figure 1, we obtained $m_w = 0.1$, for the Sauk River watershed. The parameters C and G are determined from the Ferguson (2007) flow velocity equations for high and low flow resistance domains which, following Rickenmann and Recking (2011), are defined as a function of dimensionless unit flow rate (q^{**}), giving the following for C and G (Supporting Information, section 2) for high ($q^{**} < 1$) and low ($q^{**} \geq 100$) resistance conditions:

$$C = \begin{cases} 0.6(1 - m_w) & , \ q^{**} < 1 \\ 0.4(1 - m_w) & , q^{**} \geq 100 \end{cases} \quad (7)$$

and:

$$G = \begin{cases} 0.67 \ , \ q^{**} < 1 \\ 1.5 \ , \ q^{**} \geq 100 \end{cases} \quad (8)$$

Where q^{**} is:

$$q^{**} = \frac{Q/_w}{\sqrt{gSD_{84}{}^3}} \quad (9)$$

D_{84} is the 84th percentile grain size and S is the energy slope, approximated by the channel slope. For the Sauk watershed ($m_w = 0.1$), Y is defined for the high and low-roughness domains as:

$$Y = \begin{cases} 0.77 & , q^{**} < 1 \\ 0.54 & , q^{**} \geq 100 \end{cases} \quad (10)$$

Appropriate application of the Ferguson (2007) flow resistance equations in (5) requires a definition of Y for intermediate flows ($1 \leq q^{**} < 100$) (personnel communication with Dr. Dieter Rickenmann). In practice, the Ferguson (2007) flow-depth based Variable Power Equation (VPE: described below) and Rickenmann and Recking (2011) unit-flow form of the VPE can be used to determine flow velocity over all flow resistance domains; however, neither of these are easily implemented in (5). For simplicity, we approximate Y for intermediate flows by linearly interpolating between the high and low roughness values of Y as a function of q^{**}.

The bedload transport capacity ratio ψ provides an index of relative bedload transport capacity, which can be effectively used to compare relative change in bedload transport capacity across channel morphologies. If flow rates fall primarily in the high- or low-roughness domains, then ψ can be implemented without q^{**} and defined with only two parameters, both of which relate to fluvial geomorphologic conditions of the channel: the bankfull flow (Q_b) and the ratio of the effective bankfull shear stress to the effective reference shear stress (τ_b'/τ_r').

In many gravel-bedded channels, Q_b is equivalent to the effective discharge, which transports the most sediment over time, and is therefore often considered equivalent to the channel-forming flow (e.g., Andrews and Nankervis, 1995; Wolman & Miller, 1960; but see exceptions noted by Hassan et al., 2014). Furthermore, in specific regions, Q_b often corresponds to a consistent return-interval flow (Barry et al., 2004; Castro & Jackson, 2001; Whiting et al., 1999) and can be determined objectively at any reach that has a sufficiently long hydrograph. In the Sauk River watershed, it is equivalent to the 1.2 year flow.

The ratio τ_b'/τ_r' is commonly written in terms of total Shields stress and referred to as the bankfull-excess Shields stress (τ_b^*/τ_r^*) . Herein, we refer to the ratio τ_b'/τ_r' as the effective-

bankfull-excess shear stress, which reflects the shear stress acting on the surface grains during the bankfull flow relative to effective stress needed to mobilize the surface grains. From the median grain size, channel slope and flow depth, τ_b'/τ_r' can be approximated from typical values of τ_b^*/τ_r^*. Values of τ_b^*/τ_r^* reflect bed mobility (Bunte et al., 2013) and sediment supply (Pfeiffer et al., 2017) and are reported for specific channel morphologies (Buffington & Montgomery, 2021). For a given reach morphology, τ_b'/τ_r' thus provides context for hypothetical sediment supply or bed mobility conditions of the reach. Since (5) is written as a function of τ_b'/τ_r', it is suitable for rapid sensitivity analysis to channel bed composition changes.

2.3.2.2 Estimation of effective-bankfull-excess shear stress (τ_b'/τ_r') at field sites

We used (5) to model nondimensional bedload transport capacity across the Sauk River watershed with flow from DHSVM. To obtain τ_b'/τ_r' across the channel network we first estimated τ_b' and τ_r' at five cross sections where relevant field data were gathered (Figure 1). Then τ_b'/τ_r' values were correlated to drainage area to scale τ_b'/τ_r' to each of the 582 reaches as a function of their respective drainage areas. Below, details for τ_b'/τ_r' estimation based on field information are discussed.

We calculated τ_b' using a reduced-energy-slope approximation for effective shear stress (Rickenmann 2012):

$$\tau' = \tau \left(\frac{U}{U_o}\right)^{1.5} \tag{11}$$

where τ is total shear stress, approximated as the depth-slope-product ($\tau = \rho g d S$), U_o is the virtual velocity corresponding to the base level resistance, U is the flow velocity and d is flow depth. We defined U using the Ferguson (2007) VPE, here in explicit form with shear velocity on the right-hand side:

$$U = \frac{a_1 a_2 \left(\frac{d}{D_{84}}\right)(gdS)^{0.5}}{\left[a_1^2 + a_2^2 \left(\frac{d}{D_{84}}\right)^{1.67}\right]^{0.5}} \quad (12)$$

Following Rickenmann and Recking (2011), we used a_1 = 6.5 and a_2 = 2.5 and U_o was determined as:

$$U_o = 6.5\sqrt{gdS}\left(\frac{d}{D_{84}}\right)^{0.167} \quad (13)$$

In both (12) and (13) d is bankfull flow depth. We solved for the bankfull flow depth by assuming bankfull flow channel hydraulics were reasonably approximated by a trapezoidal representation of the cross section and iteratively adjusting d until Q was equivalent to Q_b (1.2 year flow):

$$Q = A_x U \quad (14)$$

$$A_x = \frac{d}{2}[2w_x + d(m_l + m_r)] \quad (15)$$

where A_x is flow cross section area, w_x is the bottom width of the channel and m_l and m_r are the inverse of the slopes of the left and right channel wall (horizontal/vertical). Note that the exponent in (11) is 1.5 because we determine U and U_o as a function of depth rather than unit flow.

Typically, reference shear stress is determined from observations of shear stress and bedload transport by converting the bedload transport rates to W^* and defining the reference shear stress as the shear stress that corresponds to $W^* = 0.002$ (Parker et al., 1982; Wilcock, 2009). If the observed W^* values are less than 0.002, bedload transport formula are fit to the $\tau - W^*$ data set and used to extrapolate beyond the observations (Mueller et al., 2005). Alternatively, a theoretical reference shear stress can be determined from bedload transport formula (Wilcock, 2009). Not having bedload transport observations, we defined τ_r' by solving (4) for ϕ corresponding to $W^* = 0.002$ and converting ϕ to τ_r' as detailed below. Since (4) only includes the high-shear stress

formula, the ϕ at which $W^* = 0.002$ is slightly larger than 1 ($\phi \sim 1.1$). Transport stage, written in terms of the effective, Shields stresses acting on the median grain size is:

$$\phi = \frac{\tau^{*\prime}}{\tau_r^{*\prime}} \tag{16}$$

Setting $\phi = 1.1$ and $\tau_r^{*\prime} = 0.03$ (following Schneider et al. (2015) for effective stresses) we solve for $\tau^{*\prime}$. Finally, τ_r' can then determined with the Shields equation:

$$\tau_r' = \tau^{*\prime}(s-1)gD_{50} \tag{17}$$

We used the above method to determine τ_b'/τ_r' and parameterize ψ at each cross section. Rating curves of ψ, are compared with rating curves of explicitly computed q_s/q_{sb} (herein referred to as ψ_s) over $0.1Q_b < Q \leq 3Q_b$ in Figure 3. Explicitly computed q_s/q_{sb} are from the Wilcock and Crowe (2003) high-shear-stress bedload transport formula for total transport rates (see Appendix A). Figures 3A, 3B and 3C show ψ computed using the high-resistance ($Y = 0.77$), the low-resistance ($Y = 0.54$) and the variable (high-, intermediate- and low-resistance) exponents respectively. Figures 3D, 3E and 3F show error, or order of magnitude difference between ψ and ψ_s, expressed as $\log(\psi) - \log(\psi_s)$. A constant high-resistance Y causes ψ to overestimate ψ_s at high flow rates ($Q/Q_r > 5$) and underestimate ψ_s at low flow rates ($Q/Q_r < 1$) in the lowland channels (median error for all channels is +300 percent at high flows and -1.6 orders of magnitude at low flows). A constant low-resistance exponent Y causes ψ to underestimate ψ_s at high flow rates and to overestimate ψ_s at low flows in the upland channels (median error for all channels is -21 percent at high flows and +610% at low flows). A variable exponent Y doesn't systematically bias ψ in any part of the channel network but the median bias is -21% at high flow rates +59% at low flows. Values of q^{**} range from high as 3000 at cross section 1 to as low as 4 at cross section 5, therefore we use a variable exponent Y to determine ψ across the Sauk watershed; however, as

shown by Figure 3B and 3E, assuming a constant low-resistance exponent also does a reasonable job of predicting ψ_s at the cross sections, particularly when flow rates are greater than Q_r.

To determine Y across the watershed, we use D_{84} and drainage area values measured at the five cross sections to scale D_{84} to each of the 582 modeled reaches as a function of their respective drainage areas (Supporting Information, section 9). Surface grain size, reach geometry and water-surface slope used in the above computations were determined at each cross section as follows: Grain size distributions are measured from pebble counts collected at cross sections 4 (n=50) and 5 (n = 90), based on bank-to-bank random sampling of clasts, along the cross section. Grain size distributions for cross sections 1, 2, and 3 were obtained from county flood hazard assessment studies (DeVries, 2008: DeVries & Madsen, 2008). Cross sections 1, 2 and 3 are located in pool-riffle channels. Cross sections 4 and 5 are located in plane-bed and cascade morphologies respectively. The geometry of each cross section is described in the Supporting Information, section 8. In cross sections 1, 2 and 3, we estimated the cross-section topography and slope from a 1-m lidar DEM. Sub-aqueous topography was approximated from USGS cross sections recorded during flow measurements. At cross sections 4 and 5, cross-sectional topography and slope were determined from an auto-level survey we conducted during a 0.4 to 0.6-year flow (~5 times the annual average flow). Data are included in the Supporting Information, section 8.

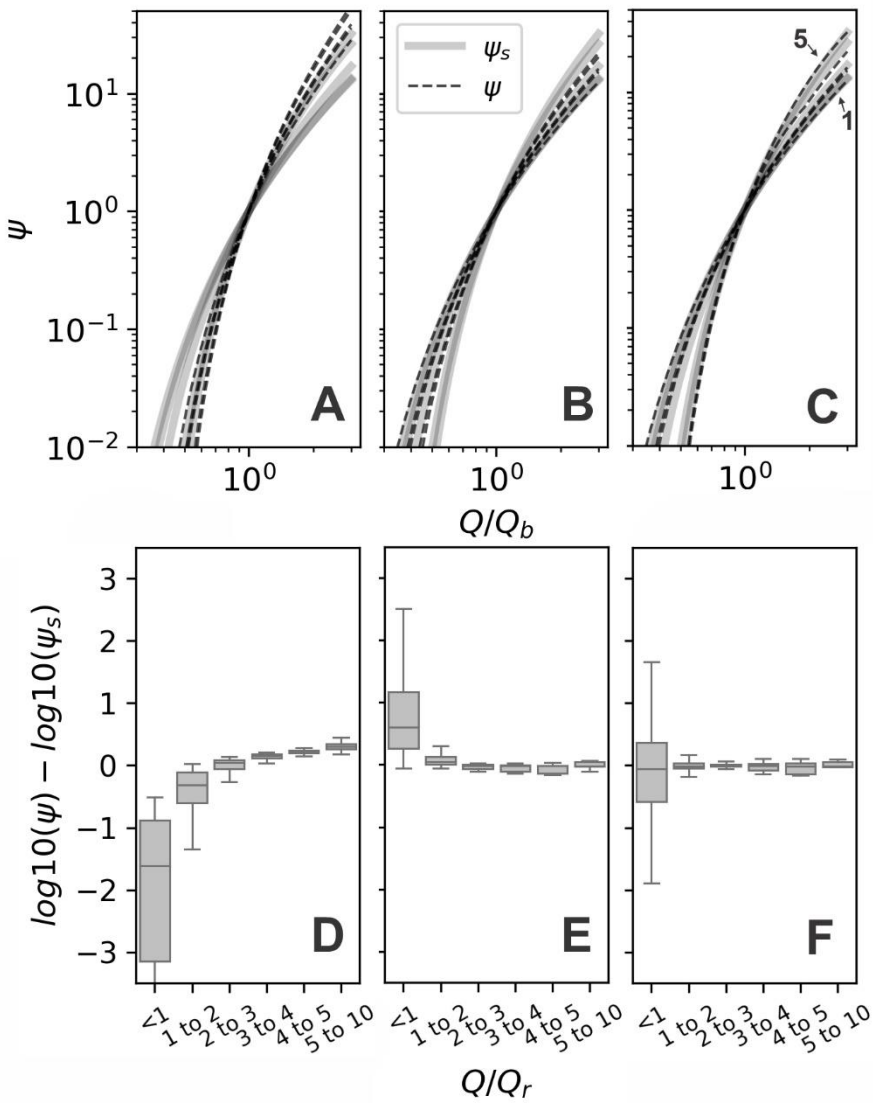

Figure 3. Rating curves of ψ relative to rating curves of explicitly calculated non-dimensional bedload (ψ_s) at each cross section. (A), (B) and (C) show ψ computed using the high-resistance ($Y = 0.77$) , low-resistance ($Y = 0.54$) and variable (high-, intermediate- and low-resistance) exponents respectively. (D), (E) and (F) show error, or order of magnitude difference between ψ and ψ_s expressed as $\log(\psi) - \log(\psi_s)$. Note that error is plotted relative to flow normalized by the reference flow rate (Q/Q_r).

2.3. Indices for hydrometeorological bedload error (HBE) and hydrologic variability

Flow events were extracted from the 1-h and 24-h hydrographs using a rules-based system similar to that applied in Jones and Grant (1996) and Tang and Carey (2017). Both the hydrograph and the time derivative of the hydrograph were used to identify the beginning, peak and end of each flow-event hydrograph. Precipitation events coinciding with each flow event were extracted from a reach-mean hyetograph using a modified form of algorithm number 6 in Jan et al. (2007). The reach-mean hyetograph was determined as the mean of all modeled hyetographs (PNNL WRF centroids) in the contributing area to the reach. Flow events and coincident precipitation events at each reach were assigned a basin-consistent time stamp and paired with their respective 24-h flow and precipitation events. Details on flow and precipitation event extraction and methods used to assign a consistent time stamp across the watershed are included in the Supporting Information, section 7.

For each flow event we estimated the cumulative non-dimensional bedload transport capacity (ψ_T) for 1-h and 24-h modeled hydrographs (ψ_{T1}, ψ_{T24}).

$$\psi_T = \sum_{i=1}^{n} \psi_i \tag{18}$$

The index of hydrometeorologic bedload error (HBE, denoted by ψ_T^*) is expressed as the ratio of the 1-h and 24-h cumulative non-dimensional bedload transport capacity as:

$$\psi_T^* = \frac{\psi_{T1}}{\psi_{T24}} \tag{19}$$

where n is the total number of hours in the flow event and i represents the i-th hour in the event.

The difference in the 1-h and 24-h hyetograph representation of a precipitation event is expressed using an index of precipitation variability based on the ratio of the peak precipitation intensity of the 1-h hyetograph ($P_{1\,max}$) to the 24-h hyetograph ($P_{24\,max}$):

$$P^*_{max} = \frac{P_{1\,max}}{P_{24\,max}} \quad (20)$$

The difference in the peak flow responses driven by the 1-h and 24-h hyetographs is expressed using an index of peak streamflow variability based on the ratio of the peak flow rate of the 1-h hydrograph ($Q_{1\,max}$) to the 24-h hydrograph ($Q_{24\,max}$):

$$Q^*_{max} = \frac{Q_{1\,max}}{Q_{24\,max}} \quad (21)$$

In addition to these two indices, to compare the flow magnitudes across low- and high-order channels, we express the peak flow rate (driven by 1-h hyetograph) as a proportion of the reach bankfull flow rate as $Q_{1\,max}/Q_b$.

In order to discuss our findings within the channel network continuum framework, we group the aforementioned indices with respect to the channel morphological classification of Montgomery and Buffington (1997) so that we can compare ψ^*_T and other indices for specific channel morphologies. We converted the slope-defined channel morphological classification of Montgomery and Buffington (1997) to a contributing area defined classification by measuring slope and contributing area from a sample of 100 different 33-m long segments and their respective contributing areas from the 1 meter Lidar DEM of the Sauk watershed (WA DNR Lidar Portal, 2018), fitting a power function to the resulting dataset and using the slope-contributing area relation to define slope as a function of contributing area. The sample of 100 different segments included lowland to upland channel reaches. Slope breaks listed in Montgomery and Buffington (1997) were used with slight modifications to match the observed slope ranges of the cascade and pool-riffle channels. Observed and predicted channel types matched the contributing area-defined channel morphology classification with some overlap at the colluvial to cascade and step-pool to plane-bed morphologies (See Supporting Information, sections 8 and 9).

2.4. RESULTS

2.4.1. Model Calibration - hydrology

Hydrologic calibration results relative to PNNL WRF precipitation, observed flow and observed SWE are shown in Figure 4. Through the DREAM calibration method (detailed in Supporting Information, section 6) we found hundreds of parameter sets that had nearly equivalent daily N-S values. Depending on the set of parameter values selected, modeled peak flow rates for water year 2009 varied by about 5 to 10% of the median value. Results herein use the model parameters that had minimal bias in the flow duration curve and cumulative modeled SWE error.

Box-and-whisker plots of basin-average hourly precipitation rates from PNNL WRF for each month are shown in Figure 4A. The precipitation rates are heaviest in the fall and winter. In the spring, precipitation rates drop to roughly 25 to 50% of the fall and winter precipitation rates, followed by a mostly dry summer period. There are two periods of elevated monthly average flow (box plots in Figures 4B and 4C) an October to January period driven by heavy precipitation events and an April to July period driven by spring snowmelt. Because bedload is transported during high flows, properly representing these two periods of elevated flow is important.

At the Sauk at Sauk gage, near the basin outlet, daily N-S efficiency was 0.71. At the Sauk above White Chuck gage, daily N-S efficiency was 0.66 (Figures 4E and 4F). These N-S values are considered satisfactory given that the PNNL data set may not accurately predict the magnitude and timing of specific events. In general, modeled flow captures the magnitude and timing of peak flows at both gages (Figure 4B and 4C); however, median peak flow rates during the month of November are overestimated by 30 to 40%. The modeled hydrographs at both gages capture flow frequency and magnitude of the observed hydrograph (Figures 4G and 4H). The three example flow events we used in this paper to illustrate the hydrometeorological controls on sediment

transport (see section 4.3. for details) are indicated by large circles relative to the bankfull flow (1.2 year flow) and the flow at which suspended sediment generally occurs in the Sauk Watershed (0.25 year flow; Jaeger et al., 2017). Event I is the largest event observed in the 35-year flow record. Events II and III have hourly peak flows that are slightly larger and smaller than the bankfull flow.

DHSVM-modeled SWE at points A, B and C (Figure 1), compared well with SWE at the Lyman Lake SNOTEL station (NRCS, 2020) in Figure 3I. None of the modeled snow locations are expected to match observations exactly; however, following correction of the cold-bias in the PNNL-WRF temperature time series (detailed in Supporting Information, section 4) and calibration, an average RMSE of 8.3 cm was achieved between years 2004 and 2011.

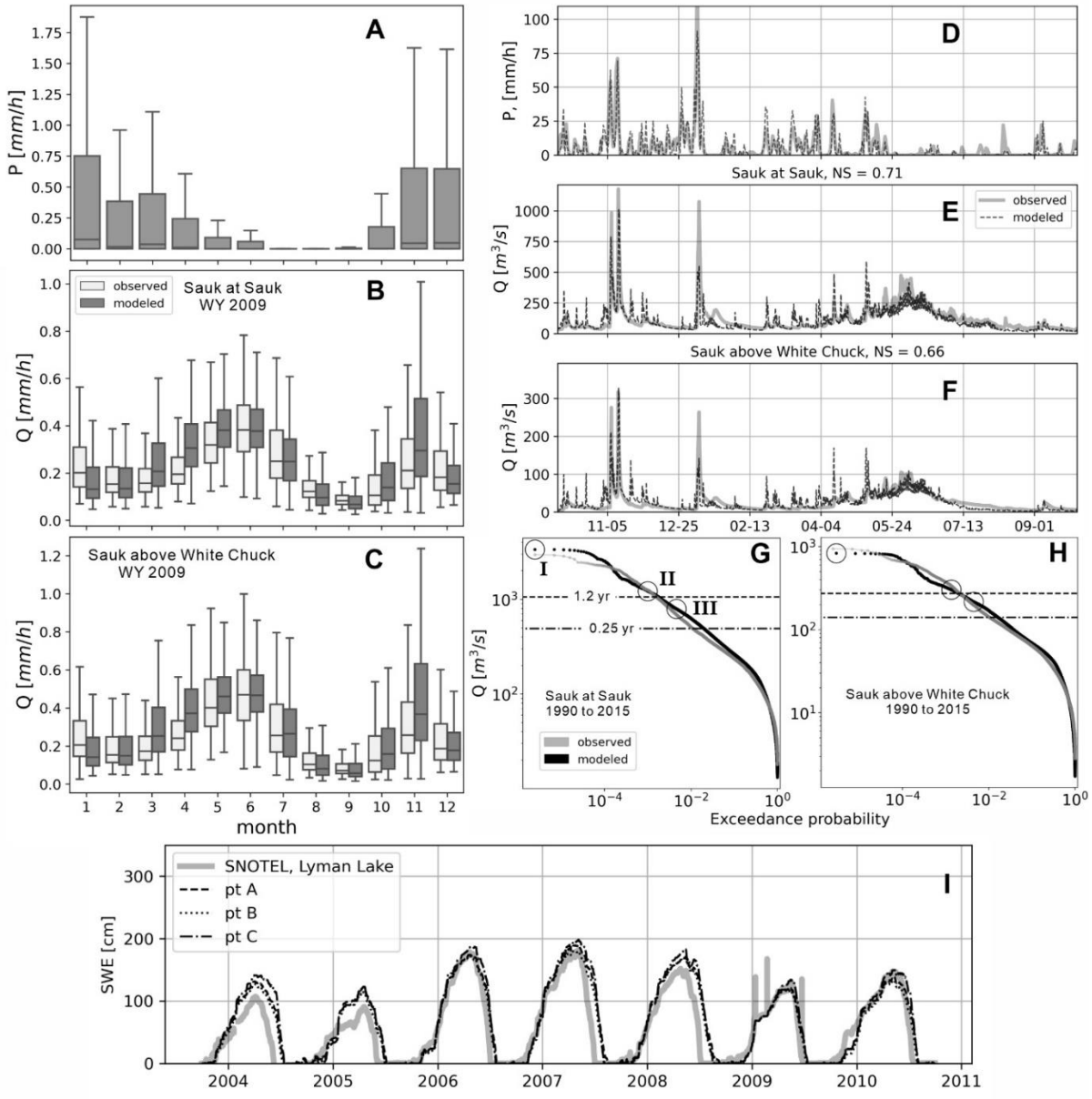

Figure 4. Box and whisker plots of 1990 to 2015, monthly (A) PNNL WRF basin-average precipitation rates and (B) observed and modeled flow at the Sauk at Sauk gage and (C) observed and modeled flow at the Sauk above White Chuck gage. (D) Water year 2009 basin-average hourly precipitation, and modeled and observed hourly flow at the (E) Sauk at Sauk gage and the (F) Sauk above White Chuck gage. Flow duration curves of hourly observed and modeled flow rates between 1990 to 2015 at the (G) Sauk at Sauk gage and the (H) Sauk above White Chuck gage. The peak hourly flow rate of example storms I, II and III (circles) relative to the 1.2 and 0.25 year flow (horizontal lines) are shown as figures in (G) and (H). (I) Modeled SWE at points A, B and C, which have similar elevation and are as close as possible to the SNOTEL station.

2.4.2. Scaling $\boldsymbol{\tau_b'/\tau_r'}$ across the channel network from limited data

Effective-bankfull-excess shear stress values (τ_b'/τ_r') calculated from observed channel conditions at the five sites, are plotted with respect to drainage area for each cross section (Figure 5A). Fitting a power function to the five contributing area and τ_b'/τ_r' points gives:

$$\tau_b'/\tau_r' \quad = 1.4A^{-0.043},\ r^2 = 0.88 \tag{22}$$

To infer the geomorphic implication of (22), the bankfull-excess Shields stress equivalent of τ_b'/τ_r' (τ_b^*/τ_r^*) is plotted as a function of slope in Figure 5B and compared to a constant value of 1.19 determined by Phillips and Jerolmack (2019) in channels that arguably met bankfull-threshold criteria. Values of τ_b^*/τ_r^* decrease from a value of 1.90 at cross section 1 (Figure 5B) to a value of 1.3 at cross section 5, suggesting that overall, the observed bed mobility (or sediment supply) is higher than expected for bankfull-threshold channel conditions in the lowlands but close to bankfull-threshold conditions at cross-sections 4 and 5. Values of τ_r and τ_r' are plotted as a function of channel slope in Figure 5C and show that in the lowland channels, where flow was generally in the low-resistance domain (high q^{**}), effective reference stress is close to total reference stress. In Figure 5D, we convert the observed values of τ_r to τ_r^* (using the median grain size) and compare to an empirical relationships for τ_r^* described by Mueller et al. (2005) for $S < 0.01$ and Pitlick et al. (2008) for $S \geq 0.01$ following Buffington and Montgomery (2021):

$$\tau_r^* = \begin{cases} 0.021 + 2.18S & ,S < 0.01 \\ 0.36S^{0.46} & ,S \geq 0.01 \end{cases} \tag{23}$$

Relative to τ_r^* predicted by (23) , observed τ_r^* is slightly low at cross sections 4 and 5 and slightly high at cross sections 1, 2 and 3 (Figure 5D), but overall in good agreement with the empirical model.

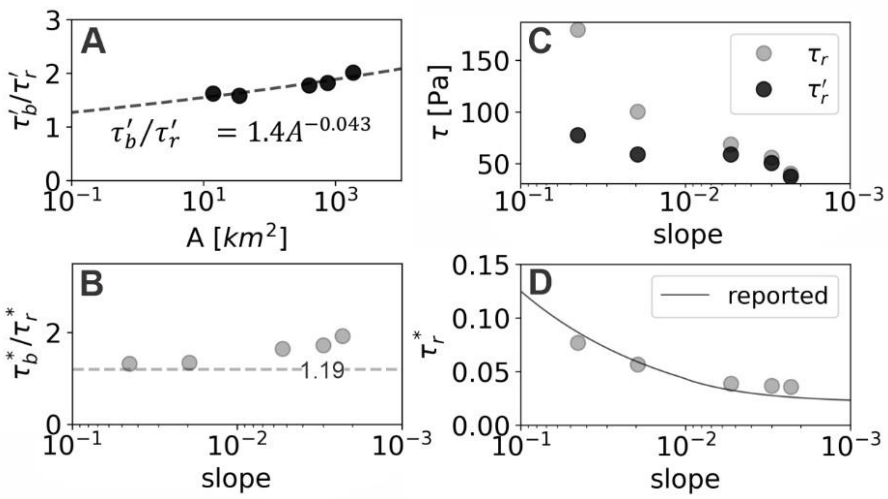

Figure 5. Observed conditions at each of the 5 cross sections plotted with respect to contributing area (A) and channel slope (B to D); note: slope is plotted in reverse to permit visual comparison with drainage area. (A) Observed effective-bankfull-excess shear stress ratio (τ_b'/τ_r') and the scaling relation used to estimate τ_b'/τ_r' in the watershed. (B) Bankfull excess Shields-stress (τ_b^*/τ_r^*). Dashed line is constant value of 1.19 indicative of bankfull-threshhold channels (Phillips and Jerolmack, 2019). (C) Total (τ_r) and effective reference shear stress (τ_r'). (D) Total reference Shields stress (τ_r^*), estimated from data plotted relative to empirical data as a function of channel slope reported in Buffington and Montgomery (2021), $\tau_r^* = 0.021 + 2.18S, S < 0.01$; $\tau_r^* = 0.36S^{0.46}, S \geq 0.01$.

2.4.3. Analysis 1: HBE of three example storm events

In the first part of our analysis, we examine watershed-scale runoff processes and their impact on HBE through case studies of three flow events. The three flow events are: (I) A long-duration, extreme (~ $3Q_b$), rain-dominated event (outlet hourly peak flow return interval, RIO, = 36 years; headwater hourly peak flow return interval, RIH, = 36 years; basin-average $P^*_{max} = 1.4$), (II) a short-duration, large (~ $1.2Q_b$), rain-and-snow-accumulation event (RIO = 1.8 years; headwater RIH = 2.5 years ; P^*_{max} = 3.1) and (III) a short-duration, moderate (~ $0.8Q_b$), rain-plus-snowmelt event (RIO = 0.6 years; RIH=0.9 years; P^*_{max}=4.5). Storm I, II and III total sediment transport capacity at the basin outlet, expressed relative to the average annual transport capacity is 322, 14.5 and 3.4% respectively. The peak flow rates of these storms are plotted relative to the flow duration curves at the Sauk at Sauk and Sauk above White Chuck USGS gages (cross sections 1 and 3 on watershed map, Figure 1; Figure 4G and 4H). For comparison, we list precipitation, discharge and bedload transport capacity statistics near the outlet (Sauk; cross section 1) and at a headwater location (Red Creek; cross section 5) in Table 1.

For each storm, reach-specific hydrographs from four reaches in the watershed and a basin-average hyetograph are shown in Figure 6. To infer whether the precipitation fell as snow or rain during each storm, Figure 6 also shows the average temperature of the PNNL WRF grid cells within the contributing area to each reach. The hydrographs are normalized by the reach bankfull flow rate (Q/Q_b). Similar to Figure 2, to infer whether the flow transported bedload, the reference flow rate is shown with the hydrograph (as Q_r/Q_b).

Spatial response in SWE are plotted over the watershed area (Figure 7) and watershed-scale indices for flow variability and HBE ($Q_{1\,max}/Q_b$, Q^*_{max} and ψ^*_T) are plotted relative to drainage area and channel type (Figure 7, right column). The plots are color coded according to sub-basin

corresponding to the color of each channel reach on the maps (labeled in lower left corner of Figure 7). Note that when $\psi_{T1} > 0$ but $\psi_{T24} = 0$, ψ_T^* is undefined. The range that $\psi_{T1} > 0$ in each channel is indicated as bar above the plot of ψ_T^*, again color coded corresponding to the color of each channel reach on the maps.

Table 1. Hydrologic and bedload transport statistics of each storm event at the Sauk outlet reach (cross section 1) and Red Creek reach (cross section 5). The bank-to-bank channel width and contributing area at the Sauk outlet is 120 m and 1865 km2 respectively. At Red Creek, it is 5 m and 13 km2 respectively.

Storm	I		II		III	
Location	Sauk	Red Creek	Sauk	Red Creek	Sauk	Red Creek
max hourly P rate [mm/hr]	13.0	14.7	6.6	8.8	6.2	4.9
P_{max}^*	1.4	1.5	3.1	3.1	4.5	3.4
P_{ant} [mm/11 days before storm]	132.9	200.2	541.8	649.1	29.8	24.9
cumulative P [mm/storm]	339.3	355.9	81.0	106.3	44.2	46.0
$Q_{1\,max}$ [mm/hr]	6.5	7.3	2.4	2.9	1.5	2.3
Q_{max}^*	1.2	1.1	1.5	1.6	1.5	1.5
$Q_{1\,max}/Q_b$	3.2	2.9	1.2	1.2	0.7	0.9
$Q_{1\,max}$ return interval	36	36	1.8	2.5	0.6	0.9
cumulative flow [mm/storm]	214.6	243.1	43.5	35.8	31.7	46.5
runoff ratio	0.63	0.68	0.54	0.34	0.72	1.0
ψ_{T1}	322	850	14.5	9.8	3.4	3.8
percent of annual average ψ_{T1}	2.4	11	0.11	0.13	0.025	0.049
ψ_T^*	1.2	1.4	2.1	10	3.0	26

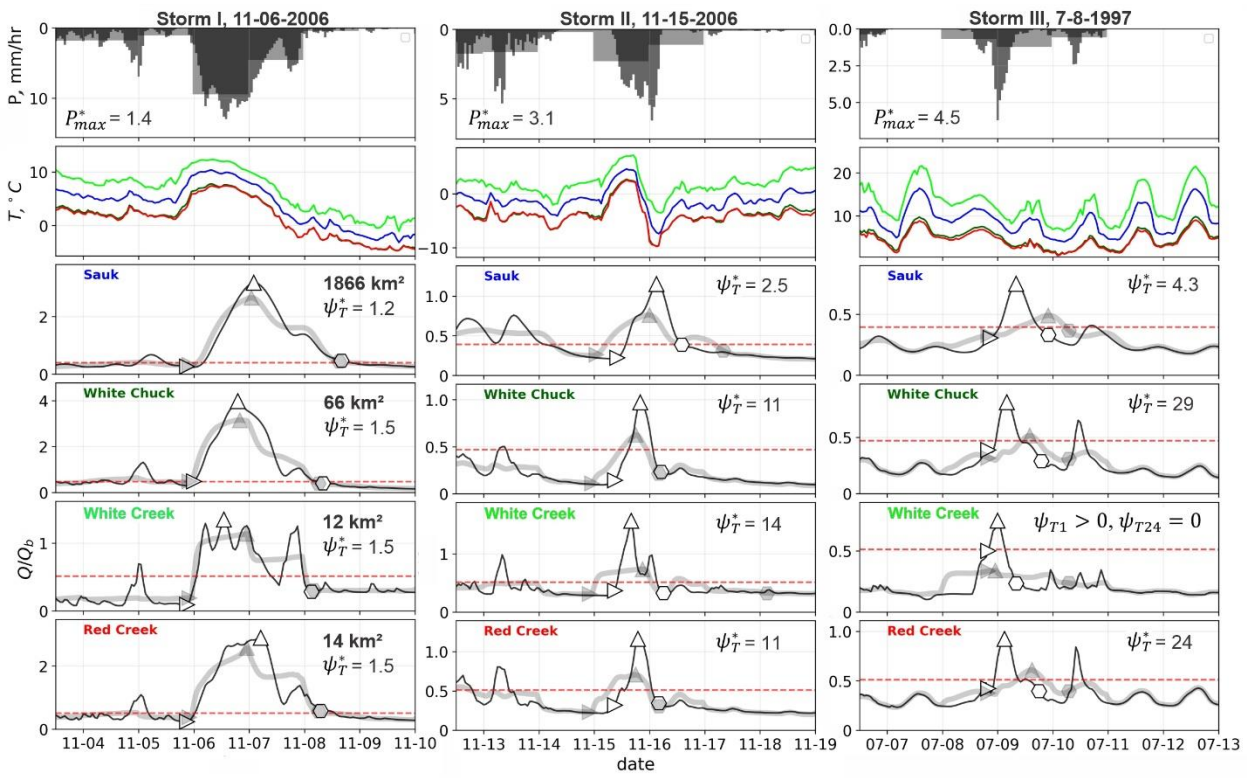

Figure 6. Hydrologic response in four representative reaches to the three example storms investigated in this paper. From top to bottom plots show: the 1-h and 24-h basin average hyetographs; max, min, and median air temperatures in the nodes upstream of each reach (color coded by reach name); normalized streamflow to bankfull flow (Q/Qb) plotted for models driven by 1h and 24h precipitation hyetographs, along with reference to bankfull flow ratio (Qr/Qb) (red dashed line) at four representative reaches. Triangles and octagon points indicate the beginning, peak and end of the flow event. The locations of the four reaches are shown as stars in Figure 7.

Storm I - The extreme, rain-dominated event

Storm I occurred between 6 and 9 November, 2006 and delivered a basin-average 340 mm of precipitation. Temperatures were warm which resulted in a cumulative reduction in SWE on the lower part of the Glacier Peak (note the ring of snowmelt) and small increase in SWE on the ridges and peak of Glacier Peak (Figure 7A). Basin average antecedent precipitation was 120 mm. Precipitation rates were relatively steady and the 24-h version of the hyetograph was nearly

identical to the 1-h version (Figure 6), resulting in relatively low (generally < 1.5) P^*_{max} and Q^*_{max} across the basin (Table 1).

Despite temporally steady 1-h precipitation rates, precipitation and hydrologic response showed some spatial variability, which lead to spatially variable $Q_{1\,max}/Q_b$, ranging between between 1 and 6, especially in cascade and colluvial channels. Channels downstream consistently showed $Q_{1\,max}/Q_b$ around 4 (Figure 7B). Maximum precipitation rates at White Creek (low elevation) and Jim Creek (intermediate elevation) were roughly 20 to 30% less than Red Creek (high elevation). Consequently, flow response at White Creek and Jim Creek ($Q_{1\,max}/Q_b \sim 1.5$) are much lower than Red Creek ($Q_{1\,max}/Q_b \sim 3$) (Figure 7B). Despite some spatial variability in precipitation, the generally long-duration and high-magnitude precipitation event caused muted differences in peak flow ratios (Q^*_{max}) and HBEs (ψ^*_T). Variability in ψ^*_T was confined to mostly colluvial and some cascade streams with a range from 1 to ~30. Downstream channels showed consistently low ψ^*_T between 1 and 2 (Figure 7D).

Storm II - The rain-and-snow-accumulation event

Storm II occurred between 15 and 17 November, 2006, one week after the end of Storm I. Basin-average precipitation was 81 mm. Basin average antecedent precipitation was 541 mm. Temperatures were below freezing before and near the end of the 1-h hyetograph. As a consequence of the temperature drop, many ridges received more than 3 cm of SWE (Figure 7E). Because the 1-h precipitation event spanned two calendar days, the 24-h hyetograph stretched the precipitation event duration from 20 hours to 48 hours and dropped the maximum precipitation rate by over 60%. Additionally, much of the event's 24-h precipitation event fell during the cold periods, mostly over the headwater elevations, that preceded and occurred at the tail end of the 1-h event, sending a larger portion of the 24-h hyetograph into storage as snow, as opposed to

contributing to runoff as in the case of 1-h hyetograph. The loss of runoff to snow-accumulation during the 24-h hyetograph dramatically reduced peak flow rates in high elevation, low-order channels like Red Creek (Figure 6), causing runoff ratios (runoff depth divided by precipitation depth) to be as low as 0.3 (Table 1). As a result, unlike Storm I, ψ_T^* was high in many of the higher order channels. It varied from 10 in the plane-bed channels to as high as several orders of magnitude in the cascade channels (Figure 7H).

Storm III - The early summer, rain-plus-snowmelt event

Storm III occurred between 8 and 9 July, 1997. It lasted only 10 hours. Basin average precipitation was 44 mm and basin average antecedent precipitation was only 30 mm. Like Storm II, the 24-h representation of the event spread precipitation evenly across a period of 48 hours and greatly reduced maximum precipitation rates; however, during Storm III, temperatures never fell below freezing and snowmelt contributed to runoff. The magnitude of snowmelt contribution varied across the channel network with snowpack distribution (Figure 7I) and in some parts of the basin, runoff ratios were as high as 1.0 (Table 1).

Values of ψ_T^* varied considerably at low and higher order channels but the magnitude of ψ_T^* appears to have been largely a function of whether or not snowmelt contributed to baseflow. In streams that received little to no snowmelt, like White Creek, baseflow before the storm was well below Q_r (Figure 6). Baseflow that was well below Q_r in combination with high P_{max}^* resulted in 1-h bedload transport capacity but no 24-h bedload transport capacity ($\psi_{T1} > 0$, $\psi_{T24} = 0$) in much of the basin. (Figure 7L). In contrast, in reaches that had snowmelt runoff, baseflow was closer to Q_r, and ψ_T^*, while still large, was closer to 10 (outlets of White Chuck and Buck Creek, Figure 7L).

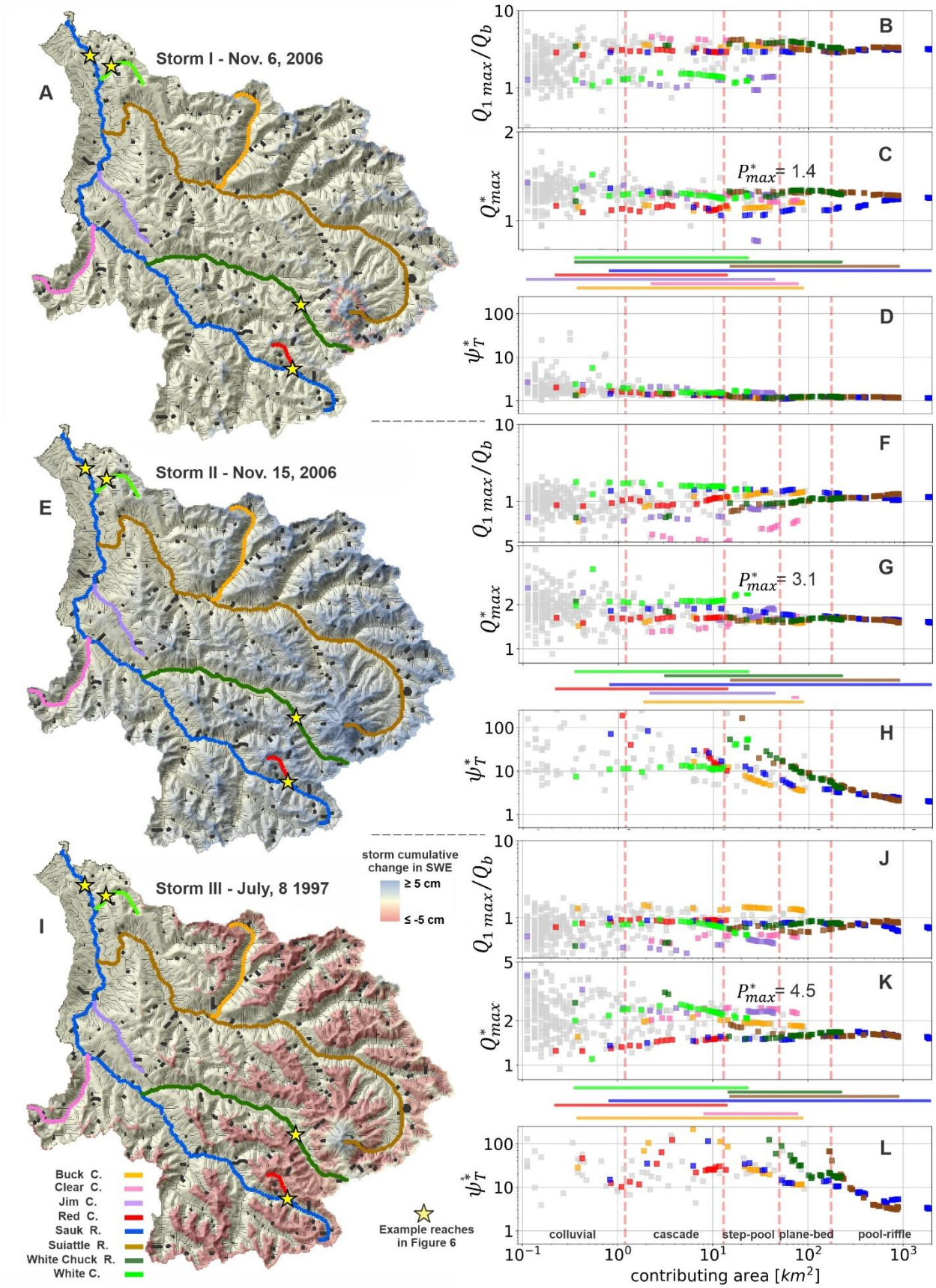
A
Storm I - Nov. 6, 2006
E
Storm II - Nov. 15, 2006
I
Storm III - July, 8 1997
storm cumulative change in SWE
≥ 5 cm
≤ -5 cm
Buck C.
Clear C.
Jim C.
Red C.
Sauk R.
Suiattle R.
White Chuck R.
White C.
Example reaches in Figure 6
B
C
P^*_{max}= 1.4
D
F
G
P^*_{max}= 3.1
H
J
K
P^*_{max}= 4.5
L
$Q_{1\,max}/Q_b$
Q^*_{max}
ψ^*_T
colluvial
cascade
step-pool
plane-bed
pool-riffle
contributing area [km^2]

Figure 7. Flow and bedload response relative to location in the channel network and channel morphology for three example storms. Maps of cumulative SWE during the 1-h representation of each storm are shown to the left in (A), (E) and (I). Colored tributaries and black reaches indicate location of the subset of 582 reaches used to track bedload response across the watershed. To the right, peak flow rate relative to the bankfull flow rate ($Q_{1\,max}/Q_b$) is plotted in (B), (F) and (J). In (C), (G) and (K), Q^*_{max} is plotted, which can be used to infer flow routing modulation of P^*_{max} for each storm. Reach HBE (ψ^*_T) is shown in (D), (H) and (L). Colored points on the right side of the figure correspond to the colored tributaries reaches shown in the SWE maps. Grey points are black reaches in the SWE maps. Note that when $\psi_{T1} > 0$ but $\psi_{T24} = 0$, ψ^*_T is undefined. The colored bars over the plots of reach ψ^*_T indicate the range that $\psi_{T1} > 0$. The vertical dashed red lines indicate channel morphology.

2.4.4. Analysis 2: Sensitivity of HBE to hydrologic variability

We identify regions of the basin where HBE (ψ^*_T) is correlated to P^*_{max}, and attempt the same for antecedent precipitation (P_{ant}) as well. To determine the relationship between HBE and precipitation metrics P^*_{max} and P_{ant}, we focus on two sub-samples of bedload transport events: large rain-dominated and, large rain-plus-snowmelt runoff events. Here "large" is any flow event greater than a 0.25-year flow that carries at least 5% of the annual average ψ_T. The 5% cutoff was selected because an empirical rainfall intensity-duration threshold for bedload transport (analogous to an intensity-duration threshold for landslides) in a cascade channel in the Rio Cordon basin in Italy (Badoux et al., 2012) fits the lower limit of the rain-dominated bedload events in cascade channels in the Sauk watershed if events that carry less than 5% of the annual average ψ_T are excluded from the sample group. The rain-dominated events are defined as all flow events that occurred during the month of November. The rain-plus-snowmelt events are all events that occur between April and July.

For the rain-dominated events, HBE expressed in orders of magnitude as log10(ψ_T^*), has a moderately positive, statistically significant ($p < 0.001$) trend with P_{max}^* (Figure 8A); however, the relation between log10(ψ_T^*) and P_{max}^* weakens and becomes increasingly heteroskedastic in the upstream direction. In the pool-riffle, plane-bed and step-pool channels, log10(ψ_T^*) is proportional to 0.3, 0.6, and 0.7 P_{max}^*, respectively (r2 = 0.4, 0.44 and 0.35). In the cascade and colluvial channels, log10(ψ_T^*) is proportional 0.7 and 0.9 P_{max}^* (r2 =0.33 and 0.21). Extreme values of ψ_T^* (i.e. ψ_{T24} is nearing 0 and ψ_{T1}>>0) occur at P_{max}^* > ~1.75 in the step-pool, cascade and colluvial channels. There is no correlation between antecedent rainfall (P_{ant}) and ψ_T^* (r2 ~ 0) (Figure 8B)

During rain-plus-snowmelt bedload events, values of log10(ψ_T^*) have a weak correlation to P_{max}^* (Figure 8C). This finding is expected because snowmelt-driven floods occur when temperatures exceed freezing regardless of whether or not there is precipitation. On average, log10(ψ_T^*) is proportional to 0.04 and 0.07 P_{max}^* in the pool-riffle and plane-bed morphologies (r2 = 0.03 and 0.11). In the step-pool and higher channel morphologies, the trend between ψ_T^* and P_{max}^* is slightly negative and r2 ≤ 0.13.

These results are highly variable, but demonstrate that, under the observed channel-substrate conditions, accurate temporal representation of rain-dominated precipitation events are necessary to accurately predict bedload transport capacity in step-pool, cascade and colluvial reaches. These findings also reinforce the concept that precipitation rates have a less predictable impact on bedload transport if flow response includes snowmelt. As such, we do not show the P_{ant} versus ψ_T^* plot for the rain-plus-snowmelt events. Note that some values of ψ_T^* plot below one during many of the snowmelt events and some of the rain-dominated events (Figures 8 A, C). This occurs when the 24-h hyetograph redistributes precipitation to a warmer period and it falls as rain rather

than snow. Additionally, some of the $\psi_T^* < 1$ storm events may be due to imprecision of the storm extraction algorithm (e.g., extraction results of Storm III at White Creek in Figure 6).

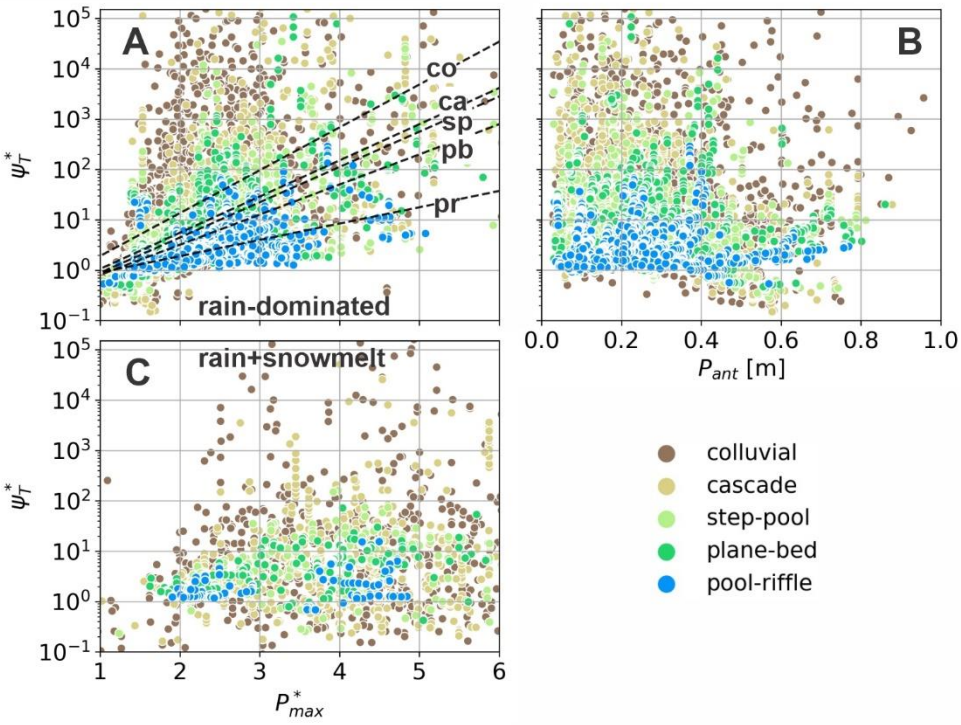

Figure 8. Precipitation and bedload characteristics of all flow events larger than a 0.25-year return interval flow that carry at least 5% of the annual average bedload (large events). (A) The index of hydrometeorological bedload error, HBE=$\psi_{T1}/\psi_{T24,} = \psi_T^*$ is plotted against the ratio of the 1-h to 24-h maximum hourly precipiation rate (P_{max}^*). Dashed lines are trend line for each channel morphology (pr: pool-riffle; pb: plain-bed; sp: step-pool; ca: cascade; co:colluvial) (B) ψ_T^* plotted against the 10-day antecedent cumulative precipitation. In (C), a plot the same as (A) is shown, but the sample group is changed to include only large, rain-plus-snowmelt bedload events. In Both B and C, trends are either insignificant or nearly flat. Details in trend lines in (A) are described in the text.

2.4.5. Analysis 3: Sensitivity of HBE to flow magnitude

We examine the role of flow magnitude (maximum daily average flow rate during the flow event) on ψ_T^*, by comparing the 1-h and 24-h empirical probability distributions of cumulative nondimensional bedload transport of the 35-year modeling record. We quantify the influence of

flow magnitude on equally probable (equal percentile) values of ψ_T for the following flow magnitude return interval (RI) bins: RI < 0.25, 0.25 < RI < 2.0 and 2.0 < RI. To minimize the impact of snow, we use only rain-dominated flow events. For each RI bin, empirical probability distributions of ψ_{T1} and ψ_{T24} are constructed. From the empirical probability distributions, specific percentile values of ψ_{T1} and ψ_{T24} (referred to as ψ_{T1_p} and ψ_{T24_p}, where p is the percentile value) are determined and compared. Results of this comparison are illustrated in Figure 9 and listed in Table 2. Note that the colluvial channels are excluded from Figure 9 (but included in Table 2) so that the distributions of ψ_T can be visualized for the other reaches.

At all flow magnitudes and channel morphologies, the median values of ψ_{T1}distributions are significantly greater than the median values of ψ_{T24} (Mann-Whitney U test, p<0.05), but differences in specific quantile values are largest during small-magnitude, upland-channel flow events, and smallest for the large, lowland channel flow events (Table 2). For example, comparing the 99th percentile value of ψ_{T1} to the 99th percentile value of ψ_{T24} value in the $RI < 0.25$ bin, ψ_{T1_99} is over 200 times greater than ψ_{T24_99} in cascade channels. In contrast, in pool-riffle channels, in the $2.0 < RI$ bin, ψ_{T1_99} is 1.2 times (20%) greater than ψ_{T24_99}. Generally speaking, under the observed channel bed conditions, the need to more accurately represent hydrologic variability grows from lowland to headwater channels. In the Sauk River watershed, the larger the flow return interval, the smaller the difference between equally likely ψ_{T1} and ψ_{T24} (and likewise, smaller HBE); however, considering that even the smallest differences in HBE are 10 to 40%, high-temporal variability precipitation data may be needed to accurately model bedload transport capacity. Furthermore, HBE in the 0.25 < RI < 2.0 flow range, which brackets the bankfull flow in the Sauk Watershed are 40% to over an order of magnitude.

Table 2. Ratio of equal percentile (p) values of ψ_{T1} and ψ_{T24} (ψ_{T1_p}/ψ_{T24_p}) for three flow return interval (RI) bins.

RI bin	$RI < 0.25$		$0.25 < RI \leq 2.0$		$2.0 < RI$	
p	75	99	75	99	75	99
colluvial	1-h = 0, 24-h = 0	1-h > 0, 24-h = 0	82	2.9	2.0	1.5
cascade	1-h > 0, 24-h = 0	222	11	1.4	1.5	1.4
step-pool	1-h > 0, 24-h = 0	52.0	6.5	1.7	1.2	1.3
plane-bed	1-h > 0, 24-h = 0	12.3	3.3	1.4	1.2	1.2
pool-riffle	9.0	3.1	1.8	1.7	1.1	1.2

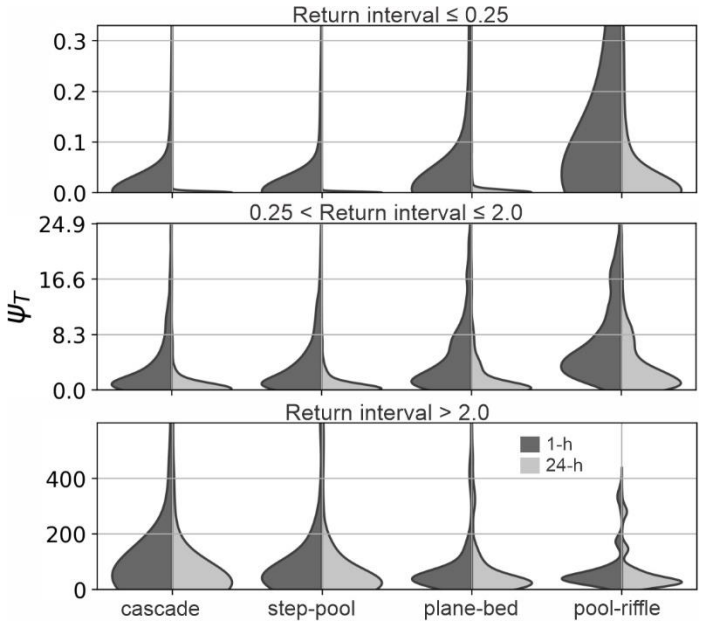

Figure 9. Empirical probability distributions of event ψ_T from the 1-h (dark-grey) and 24-h (light-grey) hydrographs for all rain-dominated (November) storm events. Values of ψ_T in the colluvial channels (not shown) are included in Table 2.

2.5. DISCUSSION

2.5.1 Key variables that control Hydrometeorological Bedload Error (HBE)

We have shown possible spatial patterns in modeled hydrometeorological bedload error (HBE or ψ_T^*) caused by using daily precipitation rates. HBE can be relatively small and uniform (Storm I) or high and variable (Storms II and III) depending on the hydrometeorology of the storm event and location in the channel network. Our results suggest that the following conditions that can lead to high HBE:

1. High precipitation variability (P_{max}^*): the peak precipitation intensity of the 1-h hyetograph is much larger than the 24-h event.
2. Runoff losses to snow storage: changing precipitation timing from a temporally accurate representation to a daily average, shifts precipitation from a period of above freezing temperatures to a period of below-freezing temperatures, and reduces rainfall runoff.
3. Baseflow before the precipitation event is much lower than Q_r.
4. A large and infrequent flow is required to mobilize the channel bed (high Q_r or low τ_b'/τ_r')

Condition 1 tends to have the most impact on HBE in upland channels where the hydrograph may more closely reflect the hyetograph. Condition 2 is likely to occur if a precipitation event is preceded and followed by periods of cold air. Condition 3 tends to occur when runoff lacks a snowmelt contribution or antecedent precipitation is low. Reaches that have Condition 4 are more likely to be affected by the other three conditions and because τ_b'/τ_r' increases in the upstream direction, Condition 4 generally, systematically increases in the upstream direction.

Here we demonstrate the effects of Conditions 1, 3 and 4 during Storms II and III, with references to specific tributaries. First, both Storms II and III had relatively high P^*_{max} values, at the watershed scale (3.1, and 4.5, respectively) and were thus impacted by Condition 1. Furthermore, during Storm III, antecedent precipitation was low and consequently, in streams that lacked a snowmelt contribution, Condition 3 enhanced HBE. For example, in White Creek, base flow was about $0.2Q_r$ and flow rates in the 24-h hydrograph never built up to Q_r. In contrast, during Storm II, baseflow in all reaches of the channel network were elevated as a consequence of Storm I. For example, baseflow at White Creek before Storm II was about 0.5 Q_r, which permitted runoff response in both the 1-h and 24-h hydrographs to exceeed Q_r (Figure 6). The absence of Condition 3 during Storm II thus limited HBE and a greater P^*_{max} combined with lower initial baseflows during Storm III led to generally higher and more undefined HBEs compared to those in Storm II (compare Figures 7H and 7L); however, during both storms, where τ'_b/τ'_r values are low, HBE was especially high (Condition 4).

The effect of Condition 2 was demonstrated during Storm II in most colluvial and cascade reaches of the watershed. As shown in Figure 6, the 24-h hyetograph redistributed precipitation from the period of warm air that coincided with the 1-h hyetograph to a cold period preceding and following the precipitation. This redistribution of precipitation to cold air caused the last 12 hours of the 24-h hyetograph to fall as snow, which in turn limited the magnitude of the 24-h hydrograph and ultimately prevented bedload transport during the later part of the 24-h flow event (Figure 6, Red Creek).

2.5.2 Sensitivity of HBE to watershed hydrologic response

We began this study expecting that HBE would increase in the upstream direction due to an increased proportion of direct runoff derived from saturated hillslopes, especially when hillslopes

were partially saturated due to wet antecedent conditions or during heavy precipitation events. We also anticipated that spatially varying watershed-scale runoff processes, such as snow hydrology, would influence network-scale trends in bedload transport capacity and sensitivity to the temporal variability of precipitation. Furthermore, we were expecting to detect a clear contributing area threshold, analogous to the transition from channel-controlled to hillslope-controlled runoff response suggested by Robinson et al. (1995) and McGlynn et al. (2004). Below this threshold, we expected HBE would become increasingly sensitive to the temporal variability of precipitation.

Rapid rainfall-runoff resulting from high antecedent precipitation may have enhanced differences in the 1-h and 24-h hydrographs during Storm II in channels like White Creek (Figure 6). But a simple regression between event antecedent precipitation and HBE from all rain-dominated, large events revealed no significant trend between the two (Figure 8B). One explanation for this is that high antecedent precipitation also increases baseflow, which removes Condition 3 and thus limits the impact of high P_{ant} on enhancing HBE. A regression between P^*_{max} and HBE shows that P^*_{max} (Figure 8A) causes higher HBE, but a clear connection to P^*_{max} is only detectible in the rain-dominated events (see lack of trend in Figure 8C). Under the observed channel bed conditions, HBE is most sensitive to P^*_{max} in upland channel reaches such as the step-pool, cascade and colluvial reaches, but as example Storm II and III showed, this result is likely due to the systematic decrease in τ'_b/τ'_r in the upland channels. While the driving mechanisms are somewhat different than expected, this finding supports our initial expectation for greater HBE in upland reaches.

Regarding the expectation that HBE would become increasingly sensitive to the temporal variability of precipitation below some channel contributing area threshold, we use the following logic to infer the location of such a threshold from Figures 7C, 7G and 7K. Assuming that hillslope

flow response is highly variable, but rapid and on some slopes may closely reflect precipitation inputs, especially for saturated conditions ($Q^*_{max} \cong P^*_{max}$), and assuming channel response is less variable than hillslope response but delayed and attenuated ($Q^*_{max} < P^*_{max}$), the location where values of Q^*_{max} increase in variability and begin to approach P^*_{max} represents the location in the channel network where direct hillslope runoff becomes the largest contribution to flow in a reach.

Applying this logic to Figure 7G (Storm II), we found that Q^*_{max} was relatively consistent at a contributing area greater than roughly 30 to 50 km2 (Fig. 7G). Upstream of 30 to 50 km2, Q^*_{max} increases but is still relatively uniform for specific sub-basin channels such as White and Clear Creek. At a contributing area of roughly 3 km2, Q^*_{max} becomes highly variable and the largest values approach P^*_{max}. Switching to Storm III (Figure 7K), Q^*_{max} was low and uniform in reaches that have a contributing area greater than roughly 100 km2. In reaches that have a contributing area of roughly 3 to 100 km2, Q^*_{max} was higher, but like Storm II, still relatively uniform across individual sub-basin channels. Below a contributing area of 3 km2, Q^*_{max} became highly variable and the largest values approach P^*_{max}. Finally, switching to Storm I (Figure 7C), Q^*_{max} was low and nearly uniform in all sub-basins and reaches, reflecting that the entire Sauk River watershed was in a consistent, highly saturated hydrologic state, but upstream of a contributing area greater than roughly 3 km2, Q^*_{max} became highly variable and maximum values approach P^*_{max}.

In all three storms, Q^*_{max} became highly variable and the largest values approached P^*_{max} at a contributing area of roughly 3 km2. We interpret this 3 km2 contributing area threshold to be hillslope-channel control contributing area threshold, above which using daily precipitation data in bedload modeling would lead to much greater errors in our DHSVM model of the Sauk River watershed. While a contributing area threshold of 3 km2 falls within the range of reported values

in the literature, sensitivity of this threshold to a different gridding scheme and resolution or model calibration should be investigated in the future.

Surprisingly, only during Storm I did the 3 km2 contributing area threshold coincide with the location of a consistent uptick in the magnitude and variability of ψ_T^*. During Storms II and III, the increase in the variability and magnitude of ψ_T^* occurred at a contributing area of roughly 50 to 200 km2, well downstream of the 3 km2 contributing area threshold.

Based on review of our distributed modeling results, we conclude that the reason the variability of HBE did not track with the variability of Q_{max}^* during Storms II and III is partially explained by spatially variable precipitation rates but mostly a consequence of spatially variable snow-hydrology. For example, during Storm III, all of the sub-basins that lack a clear snowmelt had elevated HBE or no transport at all (Clear, Jim and White Creeks; Figure 7I) while those with snow-melt generally had lower HBE. During Storm II, cumulative precipitation in White Creek and Red Creek were roughly 30% and 115% greater than precipitation amounts in Clear Creek respectively. Higher precipitation rates led to higher flow rates and as described in section 4.5, higher flow rates tend to result in smaller HBE because both the 1-h and 24-h hydrographs exceed the Q_r. However, in Red Creek, despite having higher precipitation than White Creek, flow rates were relatively low (lower Q/Q_b; Figure 7F) because runoff was lost to snow accumulation (compare White and Red Creek in Figure 7E). Reduced flow rates because of Condition 2 in combination with systematically increasing τ_b'/τ_r' in upland channels caused high HBE well downstream of the apparent 3 km2 contributing area (Figure 7H). Our expectation that spatially varying watershed-scale runoff processes would influence network-scale trends in HBE was correct but the degree to which snow hydrology would influence the location of high HBE during Storms II and Storm III was not anticipated.

2.5.3. Sensitivity of HBE to effective-bankfull-excess shear stress

As shown above, the spatial variability of HBE varies with channel bed composition, as represented by the ratio τ_b'/τ_r'. However, the observed trend in τ_b'/τ_r' represents just one possible instance in the transitory behavior of sediment in mountainous channels (i.e., Benda & Dunne, 1997). The sensitivity of HBE to channel bed composition is revealed when differences in the effective flow resulting from a 25 percent increase, a 25 percent decrease and the unchanged scaling relationship for observed τ_b'/τ_r' are compared (Figure 10). Here, effective flow is the flow rate that transports the largest proportion of all bedload transported through a reach (Wolman & Miller, 1960), determined following methods detailed in the Supporting Information, section 11.

For the observed bed composition (Figure 10B), the effective discharge is consistently about $1.0Q_b$ in all channel morphologies except the pool-riffle channels, which was $0.8Q_b$. If the τ_b'/τ_r' scaling relationship is increased by 25% (Figure 10A), the effective flow decreases to $0.4Q_b$ in the pool-riffle channels, $0.8Q_b$ in the plane-bed, step-pool and cascade channels and $1.0Q_b$ in the colluvial channels. If the τ_b'/τ_r' scaling relationship is decreased by 25% (Figure 10C), the effective flow increases to $1.3Q_b$ in the plane-bed, step-pool, cascade and colluvial channels and to 1.0 in the pool-riffle channels and there appears to be a clear, secondary effective discharge equal to $2.8Q_b$ to $3.2Q_b$ in all channel morphologies. This simple illustration demonstrates how the assumed spatial variability of τ_b'/τ_r' can change the flow magnitude that has the highest potential to transport bedload. As demonstrated in Figure 9 and Table 2, HBE may be 40% if the effective flow is closer to a 2 year event and over an order of magnitude if smaller ($0.25 < RI < 2.0$ flow bin, Figure 9 and Table 2).

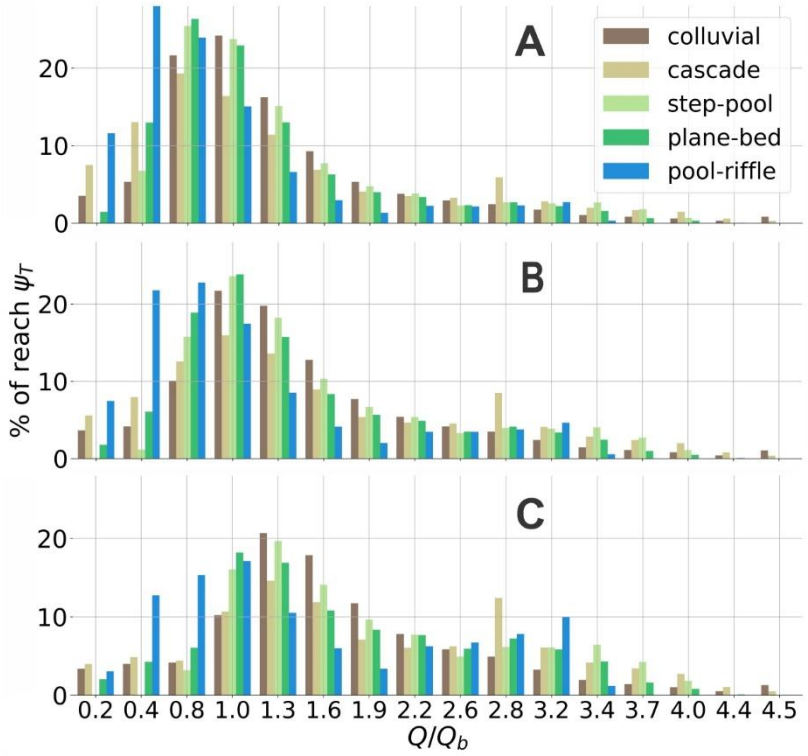

Figure 10. Effective flow, determined as the flow bin corresponding to the peak of the distribution of total non-dimensional bedload transport capacity (ψ_T) relative to normalized peak flow rate (Q/Q_b) for each channel morphology for the (A) observed τ_b'/τ_r'-drainage area scaling relationship increased by 25 percent and (B) observed τ_b'/τ_r'-drainage area scaling relationship and (C) observed τ_b'/τ_r'-drainage area scaling relationship reduced by 25 percent.

The impact of spatially variable τ_b'/τ_r' on HBE as a function of storm hydrometeorology can be visualized by re-computing HBE for Storms I, II and III using constant values of τ_b'/τ_r' between 0.8 and 100, which represent a range of bed compositions from low- to high-bed mobility. For comparison, the observed values ranged from 1.9 at the outlet to ~1.2 or less in the headwater channels (Figure 5 A). This analysis is conducted on the main stem of Sauk River (blue points in Figure 7). Note that ψ_T^* can only be plotted when both $\psi_{T1} > 0$ and $\psi_{T24} > 0$ (Figure 11). Also, to avoid the need to simultaneously scale D_{84} with τ_b'/τ_r', we use the low-resistance form of ψ (exponent Y is constant at 0.54), which as discussed in section 3.2, under estimates ψ in upland

channels, and therefore likely provides an estimate of minimum potential HBE in the upland channels but is representative in the low land channels. By using fixed values of τ_b'/τ_r' , differences in bedload transport capacity driven by the 1- and 24-hyetographs become evident.

For Storm I, the change in bed composition results in little change in bedload transport capacity sensitivity to HBE (Note scale of Y-axis). Regardless of the channel bed composition, ψ_T^* is relatively low (highest around 2) and constant across the channel network because as previously described, flow rates at all points in the network, in both the 1-h and 24-h hydrographs, are much higher than Q_r for roughly the same period of time. Both the 1- and 24-h hyeteographs result in indistinguishable hydrological variability in the watershed to trigger large differences in bedload transport (Figure 7C).

In contrast, a similar adjustment in τ_b'/τ_r' during Storm II results in orders of magnitude increase in ψ_T^* for a given drainage area as the channel bed coarsens (Figure 11). Furthermore, no transport conditions arise when the model is forced by 24h-precipitation for τ_b'/τ_r' >1.2. Spatial variability in the hydrologic response leads to a greater sensitivity on channels upstream of plane-bed channels. This is because a larger proportion of precipitation during the 24-h hyetographs falls in the form of snow (Condition 2), reducing the magnitude of the 24-h hydrograph. During the late spring rain-on-snow event (Storm III), the pattern observed in Storm II is reversed. Constant values of τ_b'/τ_r' result in smaller ψ_T^* in upland channels because snow-melt contribution in the upland channels maintain baseflow closer to Q_r and limit the effect of Condition 3.

The reach-scale assessments of bedload transport error caused by using daily flow rates by Chen et al. (2011) and Rosburg et al. (2016) suggested that differences in modeled bedload transport are higher in steep, flashy, coarse-grained channels. Tucker and Bras (2000) showed that over geologic timescales, hydrologic and transport thresholds control topographic evolution

response to precipitation variability. Here we have shown that, at time scales equivalent to a single flood event, spatially (and temporally) variable transport thresholds (τ_b'/τ_r') and hydrologic conditions, particularly storm intensity and snow-hydrology, determine the location in the watershed where accurate representation of precipitation variability is necessary to accurately model bedload transport capacity.

Finally, while the level of error reported here is generally within the range of uncertainty due simply to selection of the bedload transport formulae (Gomez & Church, 1989; Barry et al., 2007; Schneider et al., 2015), the error introduced by using daily precipitation data to model flows is in addition to that uncertainty. Therefore, the index of HBE we proposed may need to be considered when interpreting network-scale trends in bedload predictions driven by daily-hydrologic models. When high HBE is suspected (when condition 1 and any of the others occur), it may be necessary to either obtain hourly precipitation data and model hourly flows or correct the daily hydrograph. Several methods are available for estimating an hourly hydrograph from the daily hydrograph such as using the precipitation and response hydrograph durations to increase the peak and shorten the duration of the hydrograph (Solyom & Tucker, 2004) or using the slope of the rising limb to disaggregate a daily hydrograph into an equivalent hourly hydrograph (Tan et al., 2007).

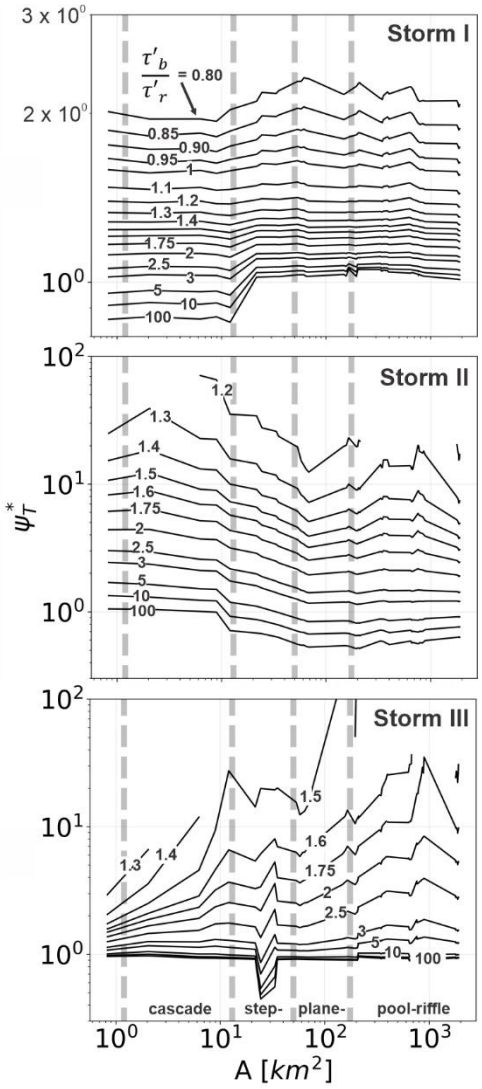

Figure 11 Values of HBE (ψ_T^*), along the main stem Sauk River (blue channel in Figure 7) recomputed for constant values of τ_b'/τ_r' for $0.8 \leq \tau_b'/\tau_r' \leq 100$ for example Storms I, II and III. The resulting τ_b'/τ_r' isolines can be used to infer the effect of different spatial configurations of τ_b'/τ_r' on HBE. The vertical dashed red lines indicate channel morphology. Note different y-axis scale for Storm I.

2.6. CONCLUSION

In this study, we developed a new non-dimensional bedload transport capacity metric (ψ), which is simply the ratio q_s/q_{sb}, but expressed in terms of Q and two parameters, Q_b and τ_b'/τ_r'. We characterized modeled bedload transport capacity error, expressed as the ratio of the 1- and 24-h

cumulative ψ (hydrometeorologic bedload error, HBE or ψ_T^*). We assessed HBE with respect to storm hydrometeorologic conditions and location in the channel network. We showed how ψ can be effectively used to compare relative change in bedload transport capacity across channel morphologies under changing bed-mobility/sediment supply conditions. Findings from this study include the following:

- The conditions that increase HBE are: high precipitation variability (P_{max}^*); loss of rainfall-runoff to snow accumulation caused by averaging of precipitation throughout the day and redistributing it to a period of colder air; low baseflow prior to a storm event; and high thresholds required for bedload transport (low τ_b'/τ_r') especially in upland reaches.
- HBE tends to be low in channel networks characterized by highly mobile (high τ_b'/τ_r') channel beds and/or during high intensity (high return-interval), uniform storm events that last several days.
- Regardless of the storm type and magnitude, we found generally high HBEs in upland channel reaches (in particular colluvial and cascade) where τ_b'/τ_r' values are low relative to downstream reaches, but the range of HBE in uplands and the rate at which it declines downstream are largely controlled by hydrologic and hydroclimatologic factors.
- Precipitation variability contributes to HBE. A positive response of HBE to P_{max}^* was detectible in the rain-dominated events, however not equally in all channel types. HBE was found more sensitive to P_{max}^* in upland channel reaches. In contrast, we found generally low HBEs for long-duration (several-day) uniform-rate atmospheric river events.
- A comparison of event antecedent precipitation and HBE from all rain-dominated, large events revealed no correlation between the two. High antecedent precipitation may

accelerate hillslope runoff response leading to higher P^*_{max}, but it also raises baseflow, which leads to lower HBE.

- We interpreted a contributing area of 3 km2 to be the transition from hillslope-controlled runoff to channel-controlled runoff in our DHSVM representation of the Sauk Watershed. Only during Storm I does this 3 km2 contributing area threshold coincide with the location of a consistent uptick in the magnitude and variability of HBE. Sub-basin hydrologic conditions, particularly snowmelt processes, can overwhelm hillslope-channel impacts on the timing of runoff. Snowmelt contributions to runoff reduce HBE while runoff-losses to snow-accumulation increase HBE.

2.7 Supporting Information

2.7.1 Justification for selected form of $\boldsymbol{\psi}$

We express bedload transport capacity in terms of the following ratio:

$$\psi = \frac{q_s}{q_{sb}} \tag{S1}$$

The variables q_s and q_{sb} are the bedload transport capacity per unit channel width at some flow and the bankfull flow. When ψ is written in terms of the Barry et al. (2007) bedload equation we can obtain:

$$\psi = \left(\frac{Q}{Q_b}\right)^{-2.45q^*+3.56} \tag{S2}$$

where Q and Q_b are some flow and the bankfull flow and the exponent of the bedload transport power function is parameterized as a function of the Dietrich et al. (1989) armoring metric q^*. The metric q^* can be interpreted as a proxy for the ratio of supply to transport capacity in a channel and is defined as:

$$q^* = \left(\frac{\tau_b - \tau_{cs}}{\tau_b - \tau_{cl}}\right)^{1.5} \quad \text{(S3)}$$

where τ_b, τ_{cs} and τ_{cl} are, respectively, the total stress at bankfull flow, the critical shear stress of the median surface grain-size and the critical shear stress of the median subsurface grain-size. A q^* value near zero implies supply limited and armored bed conditions. A value near 1 implies transport-limited and unarmored bed conditions. Most of the Sauk watershed basin is coarse-grained, and transport in coarse-grained rivers tends to be curved in log-log space (i.e., transport is threshold-like in behavior). Bedload transport modeled as a simple power function can greatly overestimate the upper or lower part of the bedload transport to flow relation depending on the range of bedload-and flow observations used to develop or fit the model. (Barry et al., 2008; Gaeuman et al., 2018; Nash, 1994).

Alternatively, ψ can be defined as a power function of the difference between shear stress and the reference shear stress:

$$\psi = \left(\frac{\tau - \tau_r}{\tau_b - \tau_r}\right)^n \quad \text{(S4)}$$

where τ and τ_r are, respectively, the total shear stress at a given flow and the total reference shear stress of the bed surface. In this form, ψ resembles the armoring metric q*. As observed by Barry e al. (2004), threshold-based equations result in no bedload bedload transport when $\tau < \tau_r$. This limitation is particularly problematic in armored channels where the reference shear stress of surface material may be greater than the bankfull shear stress and cause ψ to be undefined (note: q^* also has this limitation). Additionally, when $\tau < \tau_r < \tau_b$, (S4) results in a complex number. Upland channels, such as cascade and step-pool channels, make up a large portion of mountainous channel networks and often have channel beds that are mobilized during only the most extreme floods (Grant et al., 1990; Grant et al., 2008).

Expressing ψ in terms of the general form of a non-dimensional bedload transport equation as developed by Parker (1978), Wilcock and Kenworthy (2002) and Wilcock and Crowe (2003):

$$q = \frac{W^*}{(s-1)g\rho^{1.5}} \tau^{1.5} \quad \text{(S5)}$$

where s is specific gravity of the grains, g is gravity, ρ is the density of water and W^* is the non-dimensional bedload transport rate that normalizes transport rates across different grain-sizes (Wilcock et al., 2009) gives:

$$\psi = \frac{W^*}{W_b^*} \left(\frac{\tau}{\tau_b}\right)^{1.5} \quad \text{(S6)}$$

This form of ψ has the benefit of being threshold-like in behavior while still allowing small amounts of bedload at flow rates below the reference flow rate. Also, it can be applied to channels in which the reference flow rate of the surface grains is greater than the bank full flow rate. Finally, when used in mountainous streams with effective shear stresses, transport capacity equations of this form have been shown to more closely match observed transport rates (e.g., Schneider et al., 2015).

2.7.2 Derivation of $\boldsymbol{\psi}$ as a function of bankfull-excess-shear-stress and discharge

Nondimensional bedload transport rate functions like W^* are typically fit by multi-part equations (e.g., Parker et al., 1982; Wilcock & Crowe, 2003) because no single equation describes the behavior of both the high- and low-shear-stress transport capacity. Since we are primarily interested in modeling changes in transport capacity during floods, or flow rates near or exceeding the reference flow rate, we define (S6) using the high shear-stress form of W^*:

$$W^* = A\left(1 - \frac{B}{\phi^\alpha}\right)^\beta \quad \text{(S7)}$$

where $\phi = \tau/\tau_r$, and is commonly referred to as the transport stage (Wilcock et al., 2009) and A, B, and α are empirical parameters determined from regressing the equation against observed transport. Note that in the total transport rate form of the Wilcock and Crowe (2003), τ_r represents the reference shear stress of the median grain size. Also, since we are interested in mountainous channels, shear stresses herein use the effective shear stress and are denoted as τ'. For Wilcock and Crowe (2003) equation: A=14, B=0.894, α =0.5, β = 4.5. Substituting Equation (S7) into Equation (S6), for both W^* and bankfull conditions (W_b^*), ψ can be rewritten as:

$$\psi = \left(\frac{\frac{\phi^{\alpha-B}}{\phi^{\alpha}}}{\frac{\phi_b^{\alpha-B}}{\phi_b^{\alpha}}} \right)^{\beta} \left(\frac{\tau'}{\tau'_b} \right)^{1.5} = \left(\frac{\phi_b^{\alpha}}{\phi^{\alpha}} \right)^{\beta} \left(\frac{\phi^{\alpha-B}}{\phi_b^{\alpha-B}} \right)^{\beta} \left(\frac{\tau'}{\tau'_b} \right)^{1.5} \tag{S8}$$

Writing ϕ in terms of effective shear stress gives:

$$\psi = \left(\frac{\left(\frac{\tau'}{\tau'_r}\right)^{\alpha} - B}{\left(\frac{\tau'_b}{\tau'_r}\right)^{\alpha} - B} \right)^{\beta} \left(\frac{\tau'}{\tau'_b} \right)^{1.5-\alpha\beta} \tag{S9}$$

Now we will express this equation as a function of bankfull excess shear stress τ'_b/τ'_r and discharge to bankfull discharge ratio Q/Q_b. Following Rickenmann (2012) τ', can be determined using the depth-slope product with a reduced slope, S_o :

$$\tau' = \rho g d S_o \, . \tag{S10}$$

where d is flow depth. We use S_o from Rickenmann and Recking (2011):

$$S_o = S \left(\sqrt{\frac{f_o}{f_t}} \right)^{1.5} \tag{S11}$$

where S is energy slope approximated from channel slope, f_o is base level (skin) resistance, and f_t is the total resistance. When flow velocity is determined from unit flow rate (q), the friction term $\sqrt{f_o/f_t}$ is defined as (Rickenmann and Recking (2011):

$$\sqrt{\frac{f_o}{f_t}} = \left(\frac{U(q)}{U_o(q)}\right)^{1.5} \tag{S12}$$

where U is flow velocity and U_o is the virtual velocity corresponding to base level grain resistance with friction factor, f_o. Substituting (S12) into (S11) results in:

$$S_o = S\left(\left(\frac{U(q)}{U_o(q)}\right)^{1.5}\right)^{1.5} = S\left(\frac{U(q)}{U_o(q)}\right)^{2.25} \tag{S13}$$

And substituting (S13) into (S10) gives:

$$\tau' = \rho \mathrm{gdS}\left(\frac{U}{U_o}\right)^{2.25} \tag{S14}$$

Defining the effective shear stress ratio $\tau'/{\tau'}_b$ using (S14) and assuming water surface slope is constant gives:

$$\frac{\tau'}{{\tau'}_b} = \frac{d}{d_b}\left(\frac{U_o}{U_{ob}}\right)^{-2.25}\left(\frac{U}{U_b}\right)^{2.25} \tag{S15}$$

The ratio U/U_b can be defined in terms of unit flow rate following Ferguson (2007) for shallow (high resistance) flow as:

$$U = 1.443(gS)^{0.2}{D_{84}}^{-0.4}\left(\frac{Q}{w}\right)^{0.6} \tag{S16}$$

and deep (low resistance) flow as:

$$U = 3.074(gS)^{0.3}{D_{84}}^{-0.1}\left(\frac{Q}{w}\right)^{0.4} \tag{S17}$$

where w is flow width, D_{84} is the 84th percentile grain size. Rickenmann and Recking (2011) used (S17) to define the virtual velocity (U_o), a velocity representation of the grain roughness. They also showed that between the low and high roughness flow domains, the exponent over the Q/w term (S16) and (S17) ranges between 0.4 and 0.6. They defined three roughness domains (high, intermediate and low) using the non-dimensional metric:

$$q^{**} = \frac{{}^{Q}/_{w}}{\sqrt{g S {D_{84}}^3}} \tag{S18}$$

Here we derived a nondimensional bedload equation for both high ($q^{**} < 1$) and low ($q^{**} \geq 100$) flow resistance domains and interpolate between high and low flow resistance domains to define the intermediate flow resistance domain. The ratio U/U_b rewritten using (S16) and (S17) is:

$$\frac{U}{U_b} = \left(\frac{Q}{Q_b}\right)^{\mathrm{C}} \tag{S19}$$

where C is defined as:

$$C = \begin{cases} 0.6(1 - m_w) & , \ q^{**} < 1 \\ 0.4(1 - m_w) & , q^{**} \geq 100 \end{cases} \tag{S20}$$

The parameter m_w is the exponent of an at-a-station hydraulic geometry relation for width:

$$\frac{w}{w_b} = \left(\frac{Q}{Q_b}\right)^{m_w} \tag{S21}$$

If width does not change significantly with flow depth, it can be excluded. For the Sauk River, we obtained $m_w = 0.1$, based on stage-discharge analysis at the three gage locations shown in Figure 1. Using (s17) the ratio of U_o/U_{ob} in (S15) can be defined in terms of flow as:

$$\frac{U_o}{U_{ob}} = \left(\frac{Q}{Q_b}\right)^{0.4-0.4m_w} \tag{S22}$$

The ratio d/d_b in (S15) can also be defined in terms of flow for the high and low flow resistance domains. Following Ferguson (2007), U is written as a function of flow depth for shallow flows (high resistance) as:

$$U = 2.5(gdS)^{0.5}\frac{d}{D_{84}} \tag{S23}$$

Rearranging to solve for d gives:

$$d = \left(\frac{D_{84}U}{2.5(gS)^{0.5}}\right)^{0.67} \tag{S24}$$

For deep (low resistance) flows, U is written as a function of flow depth as:

$$U = 6.5(gdS)^{0.5}\left(\frac{d}{D_{84}}\right)^{0.167} \tag{S25}$$

Rearranging to solve for d gives:

$$d = \left(\frac{U{D_{84}}^{0.167}}{6.5(gS)^{0.5}}\right)^{1.5} \tag{S26}$$

The ratio d/d_b rewritten using (S24) or (S26) is:

$$\frac{d}{d_b} = \left(\frac{U}{U_b}\right)^{G} \tag{S27}$$

$$G = \begin{cases} 0.67 \ , \ q^{**} < 1 \\ 1.5 \ \ , \ q^{**} \geq 100 \end{cases} \tag{S28}$$

Substituting (S19) into (S27) results in:

$$\frac{d}{d_b} = \left(\frac{Q}{Q_b}\right)^{\mathrm{CG}} \tag{S29}$$

Now $\tau'/{\tau'}_b$ can be defined as a function of only Q and Q_b by combining (S19), (S22) and (S29) as:

$$\frac{\tau'}{{\tau'}_b} = \left(\frac{Q}{Q_b}\right)^{\mathrm{CG}}\left(\frac{Q}{Q_b}\right)^{-2.25(0.4-0.4m_w)}\left(\frac{Q}{Q_b}\right)^{2.25\mathrm{C}} = \left(\frac{Q}{Q_b}\right)^{\mathrm{C(G+2.25)}-2.25(0.4-0.4m_w)} \tag{S30}$$

We can write $\tau'/{\tau'}_r$ in (S9) as a function of ${\tau'}_b$ and Q_b by multiplying both sides of (S30) by ${\tau'_b}/{\tau'}_r$:

$$\frac{\tau'}{{\tau'}_r} = \frac{{\tau'}_b}{{\tau'}_r}\left(\frac{Q}{Q_b}\right)^{\mathrm{C(G+2.25)}-2.25(0.4-0.4m_w)} \tag{S31}$$

Substituting (S30) and (S31) into (S9), we arrive at a form of ψ that can be parameterized with only two parameters, bankfull excess shear stress ${\tau'}_b/{\tau'}_r$ and discharge to bankfull discharge ratio Q/Q_b.

$$\psi = \left(\frac{\left(\frac{{\tau'}_b}{{\tau'}_r}\left(\frac{Q}{Q_b}\right)^{Y}\right)^{\alpha} - B}{\left(\frac{{\tau'}_b}{{\tau'}_r}\right)^{\alpha} - B}\right)^{\beta}\left(\frac{Q}{Q_b}\right)^{Y(1.5-\alpha\beta)} \tag{S32}$$

where Y is:

$$Y = \mathrm{C}(\mathrm{G} + 2.25) - 0.9(1 - m_w) \tag{S33}$$

Including $m_w = 0.1$ and the values of C and G stated above, Y is defined for the low and high flow depths as:

$$Y = \begin{cases} 0.77 & , q^{**} < 1 \\ 0.54 & , q^{**} \geq 100 \end{cases} \tag{S34}$$

For the intermediate roughness domain ($1 \leq q^{**} < 100$), Y is linearly interpolated as a function of q^{**} and the values for the low and high flow depths:

$$Y = 0.77 - (q^{**} - 1)\left(\frac{0.77 - 0.54}{99}\right) \tag{S35}$$

Incorporating all exponent values (high, intermediate and low) gives ψ for all flow conditions.

2.7.3 PNNL WRF

PNNL WRF variables include hourly mean air temperature [°K], mean wind at 10 m elevation [m s-1], mean mixing ratio at 2 m elevation [kg kg-1], mean surface air pressure [Pa], mean downward shortwave flux at the ground surface [W m-2], mean downward long wave flux at the ground surface [W m-2] and accumulated precipitation [mm]. We convert the mixing ratio and surface air pressure to relative humidity following methods described in the appendix of Feld et al., (2013).

Comparison with hourly networks near the Sauk watershed and flow response in the Sauk revealed that the PNNL WRF dataset contains numerous artificial precipitation events and the timing of some events can be 5 to 10 hours different than observed. However, because of known issues with observing precipitation in complex terrain, particularly at sub-daily timesteps (see Lundquist et al. 2019 for a review), we use the PNNL WRF precipitation as is. This enables

consistent comparison between hourly and daily precipitation input, although offsets from actual precipitation timing results in some mismatches with observed streamflow timing.

2.7.4 Temperature bias correction

We compared PNNL WRF temperature to the Minder et al. (2010) temperature dataset and to low-elevation temperature records from the Cooperative Observer Network stations (NWS, 2020) for water years 1981 to 2010. The Minder et al. (2010) dataset consists of 21 sensors that recorded hourly at elevations ranging from 360 to 2120 m during water year 2008. The COOP dataset includes three stations near the outlet of the Sauk watershed. Results showed that PNNL WRF monthly maximum temperature have a 3 to 5 °C cold bias during the spring and 1 to 2 °C warm bias during the summer. The bias varies by time and elevation as illustrated in Figure S1.

Because the temperature bias varies by time and location, simply adding or subtracting a constant temperature difference to correct the monthly maximum temperature bias causes temperatures during all other parts of the day to differ from the temperature observations. Therefore, to bias correct PNNL WRF temperature and account for changes in the bias correction due to time and location, we apply a diurnal, monthly bias correction unique to each node (centroid) as follows:

$$T_c = T_u + \frac{(T_{dx} - T_{dn})}{2}\left(2\pi\left[\frac{T_{hr} - hr_x - 12}{24}\right]\right) + \frac{(T_{dx} + T_{dn})}{2} \quad \text{(S36)}$$

Where T_c is the bias corrected hourly temperature value, T_u is the uncorrected PNNL WRF hourly temperature value, T_{dx} is the monthly mean difference between PNNL WRF and sensor maximum temperature, T_{dn} is the monthly mean difference between PNNL WRF and sensor minimum temperature, $T_{hr}$is the hour of the PNNL WRF temperature, hr_x is the number of hours

from noon at which the daily maximum temperature is reached, which is assumed to be at 3 pm based on the timing of the temperature high in the sensor data.

To determine T_{dx} and T_{dn}, for each PNNL WRF node, we compare the node monthly average minimum and maximum temperature to temperature sensors in the elevation range of the node. Three elevation ranges are used: 0 to 900 m, 900 to 1500 m, and greater than 1500 m. These elevation ranges roughly correspond to the rain dominant/rain-on-snow, snow dominant, and highland precipitation zones in the region (Washington State Department of Natural Resources, 2020). The temperature observations in the corresponding elevation range are extrapolated to the elevation of the node. Temperature observations come from Minder et al., 2010 dataset because it spans a range of elevations and temperature during water year 2008 are a reasonable approximation of the long term average based on comparison with the 1981 to 2010 COOP (2020) dataset. Equation S36 is applied to each node. Each node has a unique set of T_{dx} and T_{dn} values for the month and node.

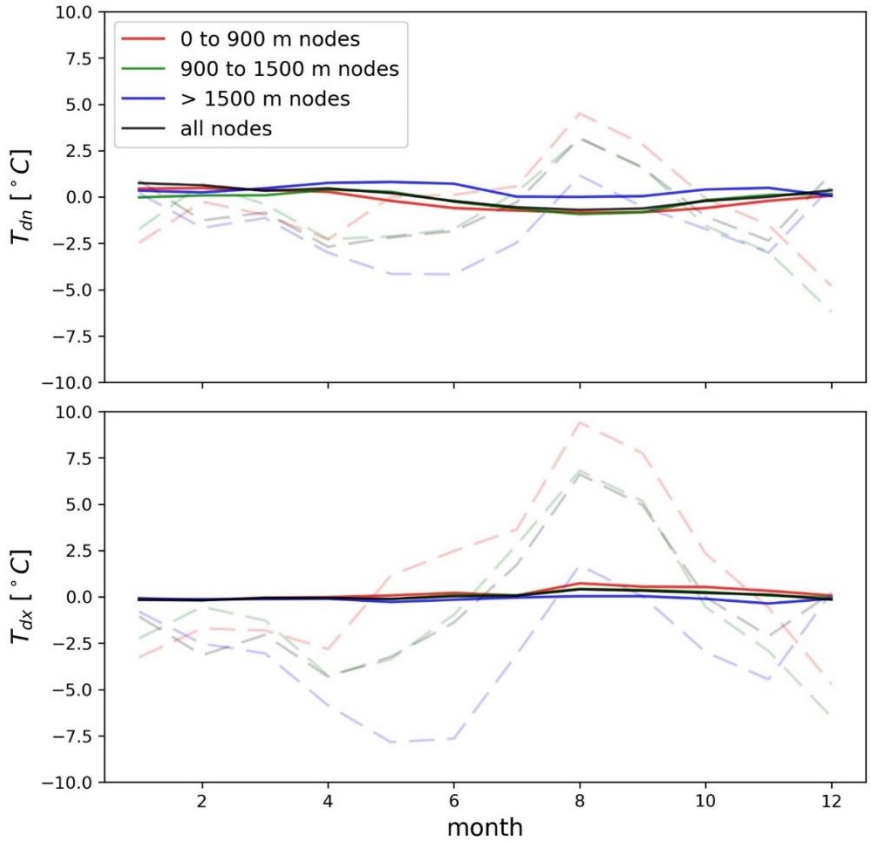

Figure S1. Difference between observed and PNNL WRF monthly minimum temperature (T_{dn}) and maximum temperature (T_{dx}) for different elevation ranges. Dashed, seim-transparent lines are before bias correction. Solid lines are after bias correction.

2.7.5 Longwave and shortwave radiation:

Longwave and shortwave radiation were recalculated using the bias corrected temperature time series and the PNNL WRF vapor pressure and relative humidity. There are many different methods for approximating longwave radiation; however, the Swinbank (1963) clear-sky long wave radiation method paired with the Unsworth and Monteith (1975) all-sky longwave radiation method were selected because these methods increased longwave radiation in the spring without changing the overall average longwave radiation value. Shortwave radiation was computed using

the Thornton and Running (1999) method as implemented in the python module MetSim (Bennett et al., 2020).

2.7.6 DHSVM channel network

Channel hydraulic properties used determine an initial parameterization of the linear-reservoir representation of each link of the DHSVM model of the Sauk Watershed channel network were assigned based on link contributing area using 500 different contributing-area classes. Bankfull flow width and depth were assigned to each class as a function of contributing area (A) using hydraulic geometry equations determined for the Pacific Northwest region of the United States (Castro and Jackson (2001). Roughness was assigned using a hydraulic geometry relation we developed between bankfull channel roughness (nb) and contributing area, from roughness back-computed from all 1- to 2-year return interval flows at the Sauk, Sauk at Darrington and Sauk above White Chuck USGS gages:

$$n_b = 1.6642A^{-0.177}, r^2 = 0.99 \quad \text{(S37)}$$

Note that in Equation S37, A is in square meters. Also note that these hydraulic geometry relations were not used for parameterizing ψ and that the parameterization of the channel network was adjusted as part of the DHSVM calibration procedure (detailed in next section).

2.7.7 DHSVM calibration and the DREAM algorithm

We calibrated DHSVM to observed hourly average flow at two locations: the Sauk at Sauk gage and Sauk above White Chuck River (United States Geological Society, 2020; Figure 1). To address parameter equifinality (Beven, 2006), which increases with basin area (Surfleet et al., 2010), we adjusted parameters using the differential evolution adaptive Metropolis (DREAM) method for

running multiple Markov chains in parallel (Vrugt and Ter Braak,2011; Vrugt, 2016). We implemented DREAM using the SPOTpy Python package by Houska et al., (2015).

The DREAM calibration algorithm was configured to adjust five DHSVM parameters: the exponential decay coefficient of saturated hydraulic conductivity with respect to soil depth (α), lateral saturated conductivity ($ksat_L$), the exponent of the channel roughness hydraulic geometry relation (Equation s19), the temperature below which all precipitation falls as snow and the temperature lapse rate used to extrapolate temperature to locations between PNNL WRF grid nodes (centroids). The final calibration parameter values were: 2.8, 0.00073 m s-1, -0.2, -0.41 °C and -0.0085 °C km-1, respectively. The lapse rate is relatively high but was identified by DREAM as the best parameter because it reduced the amplitude of diurnal fluctuations in flow caused by limitations of the current DHSVM snow model.

Initial parameter ranges for the DREAM method were selected from sources in Table S1 and then refined by repeat runs of the DREAM algorithm. Since the timing of PNNL WRF precipitation can be incorrect, we were able to improve calibration by adjusted the timing of precipitation (add and subtract a fixed number of hours to the DHSVM precipitation timestamp). An additional 10 soil parameters for soil types in the coarse soil map layer were related to the two soil parameters based on the SSURSGO soil classification and values in Rawls et al., 1982 and Das, 2005 (listed in Table S2). The hillslope parameters were adjusted in three of the coarse soil map (shown as the Forest Colluvium, Alpine Colluvium and Lahar and Glacial deposit layers in Figure 2).

Table S1. Data types and sources used for distributed hydrologic modeling with DHSVM

Spatial data	Source	Parameter value
vegetation	National Land Cover Database, 2016	Cuo et al, 2011, Jassal et al., 2009
soil	SSURGO, 2018	Das 2005, Rawls et al., 1982, Meyer et al., 1997, Kelleher et al., 2015
glacier, snow	Bandaragoda et al., 2019	Bandaragoda et al., 2019, Frans et al., 2018

Table S2. Relation between α and $ksat_L$ in the sandy loam layer to 11 additional soil parameters based on Rawls et al., 1982 and Das, 2005.

Soil type	parameter	relation
forest colluvium	exponential decrease	α
	lateral conductivity	$ksat_L$
	maximum infiltration	$ksat_L * 2$
	vertical conductivity, 3 layers	$ksat_L$, $ksat_L$, $ksat_L$
lahar and glacial deposits	exponential decrease	$\frac{\alpha}{5}$
	lateral conductivity	$\frac{ksat_L}{5}$
	maximum infiltration	$\frac{ksat_L * 2}{5}$
	vertical conductivity, 3 layers	$\frac{ksat_L}{5}$, $\frac{ksat_L}{5}$, $\frac{ksat_L}{5}$
alpine colluvium	exponential decrease	$\alpha * 2$
	lateral conductivity	$ksat_L * 3$
	maximum infiltration	$ksat_L * 3$
	vertical conductivity, 3 layers	$ksat_L * 3$, $ksat_L * 3$, $ksat_L * 3$

2.7.8 Flow and precipitation event extraction and pairing

Individual flow events were defined from the hydrograph using a rules based system similar to that applied in Jones and Grant (1996) and Tang and Carey (2017). First, both the hydrograph and

the time derivative of the hydrograph were used to identify the begin, peak and end of each flow event in the hydrograph. Then, for each flow event, characteristics of the coinciding precipitation event were recorded. A "coinciding" precipitation event was defined as any event that started between the end time and 24 hours before the begin time of the flow event. Precipitation events were identified using a modified form of the storm extraction algorithm from Jan and Lee (2004). In the modified form, a precipitation event begins when the hourly precipitation exceeds 2 mm. The precipitation event ends at the first hour in which hourly precipitation falls below 1 mm and the maximum hourly precipitation rate in the preceding 12 hours is less than 2 mm. If the paired precipitation event lasted longer than the flow event, it was truncated so that it ended no later than the flow event. If multiple precipitation events coincided with the flow event, they were combined into one precipitation event. Some flow events during the snow-melt season did not have coincident precipitation events. Because both the hydrograph and hyetograph were hourly (rather than 15-minute), a preliminary smoothing as applied in Tang and Carey (2017) was not needed.

Flow events were ranked and assigned an annual return interval based on flow magnitude of the 1-h version of the flow event. Flow magnitude was defined as the maximum, 24 hour flow rate during the duration of the flow event. If the flow event was less than 24 hours long, the maximum, 24 hour rate was computed as the mean flow rate during the 24 hour period that began 12 hours before and ended 12 hours after the peak flow rate of the storm. The flow event begin time was designated as the time stamp of the event.

Flow events at a mid-basin location in the watershed were paired with events in all other reaches in the basin so that all events had a basin-consistent time stamp. A mid-basin reach was selected for pairing with all other reaches because differences in runoff response between the mid-basin reach and other reaches in the watershed were small relative to differences in runoff response

between the outlet and headwater reaches. Occasionally, a flow event that appeared as a single flow event at the mid-basin reach appeared as a series of independent flow events in the headwater reaches. If more than one flow event in a reach coincided with a flow event in the mid-basin reach, flow event statistics (including ψ) of the multiple events were combined into a single event. If a coincident flood was not found to exist in a reach (was not identified using automated method described above), a flow event was identified that had a matching begin, peak and end time as the mid-basin reach flow event. For each reach, once flow events in the 1-h hydrograph were successfully paired and assigned a basin-consistent date stamp, flow events in the reach 24-h hydrograph were then paired with the 1-h hydrograph flow events.

2.7.9 Channel geometry and grain size tables

Table S3. Fraction of surface grains in each grain-size [m] bin

Cross section	1	2	3*	4	5
0.008	0	0	0	0	0
0.016	0.020	0.025	0	0	0
0.032	0.060	0.095	0.078	0.078	0.040
0.064	0.180	0.210	0.122	0.122	0.120
0.128	0.440	0.370	0.433	0.433	0.300
0.256	0.280	0.260	0.333	0.333	0.440
0.512	0.020	0.040	0.033	0.033	0.100
1.024	0	0	0	0	0

* Cross section 3 grain-size was unavailable and assumed equal to measured grain-size at cross section 4

Table S4 Cross section geometry and grain size [m]. b is channel width. Slopes m1 and m2 are expressed as ratio of horizontal to vertical distance.

Cross section	1	2	3*	4	5
Contributing area [km^2]	1865	763	402	35	14
Water surface slope **	0.0023	0.0040	0.0053	0.0195	0.0450
b	120	45	40	15	5
m1	0.25	5	0.75	0.5	5
m2	0.25	0.25	0.6	0.5	2
D16	0.045	0.040	0.050	0.050	0.064
D50	0.070	0.094	0.110	0.110	0.145
D65	0.120	0.125	0.130	0.130	0.160
D84	0.142	0.190	0.170	0.170	0.200

* Cross section 3 grain-size was unavailable and assumed equal to grain-size measured at cross section 4

** Water surface slope (Sf) estimated from 1 meter lidar dem at cross sections 1, 2 and 3 and an autolevel survey at cross sections 4 and 5.

2.7.10 Assigning channel network morphology, grain size and bankfull width

We approximated channel morphology, the 84th percentile grainsize (D84) and the bankfull width (w_b) at each link in the channel network as a function of link-contributing area or the bankfull flow. These bankfull flow width and D84 values were used in transport capacity calculations. Bankfull flow width was determined as a function of bankfull flow using an empirical relation for Pacific Maritime mountain streams in the Pacific Northwest (Castro & Jackson 2001):

$$w_b = 2.37{Q_b}^{0.50} \tag{S38}$$

The 84th percentile grain size was approximated as a power function of contributing area using the values in Table S4:

$$D_{84} = 0.21A^{-0.039},\ r^2 = 0.39 \tag{S39}$$

Finally, channel morphology was assigned based on reach slope according slope-defined channel morphological classification of Montgomery and Buffington (1997). Since bedload transport values are highly sensitive to channel slope, slope was inferred from reach contributing area and slope measured form a 1 m Lidar DEM of the Sauk watershed, rather than the 150 m DEM used for DHSVM. We fit a power function to a sample of slope measurements and their respective contributing areas from 100 reaches distributed across a range of channel morphologies (Figure S2):

$$S = 0.28A^{-0.57},\ r^2 = 0.58 \quad \text{(S40)}$$

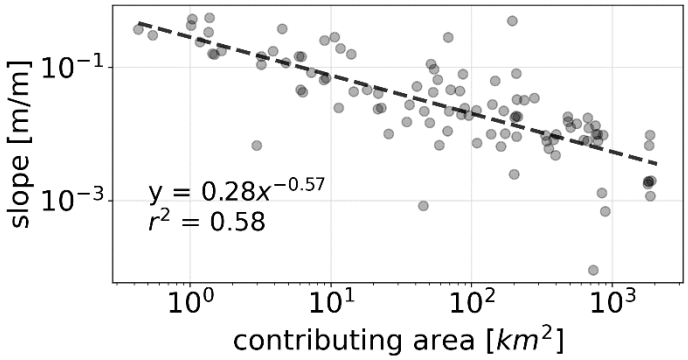

Figure S2 Relation between contributing area and channel slope used to define channel morphology as a function of contributing area

Table S5 Slope and contributing area bins used to categorize channel morphology

Channel morphology	Slope	Equivalent contributing area [km²]
pool-riffle	0.015 – 0.004	>175
plane-bed	0.03 – 0.015	50 – 175
step-pool	0.065 – 0.03	13 – 50
cascade	0.25 – 0.065	1.2 – 13
colluvial	>0.25	0 – 1.2

2.7.11 Example reaches and channel morphology classification

Figure S3 Sauk River near the cross section 1 during a snow-melt event

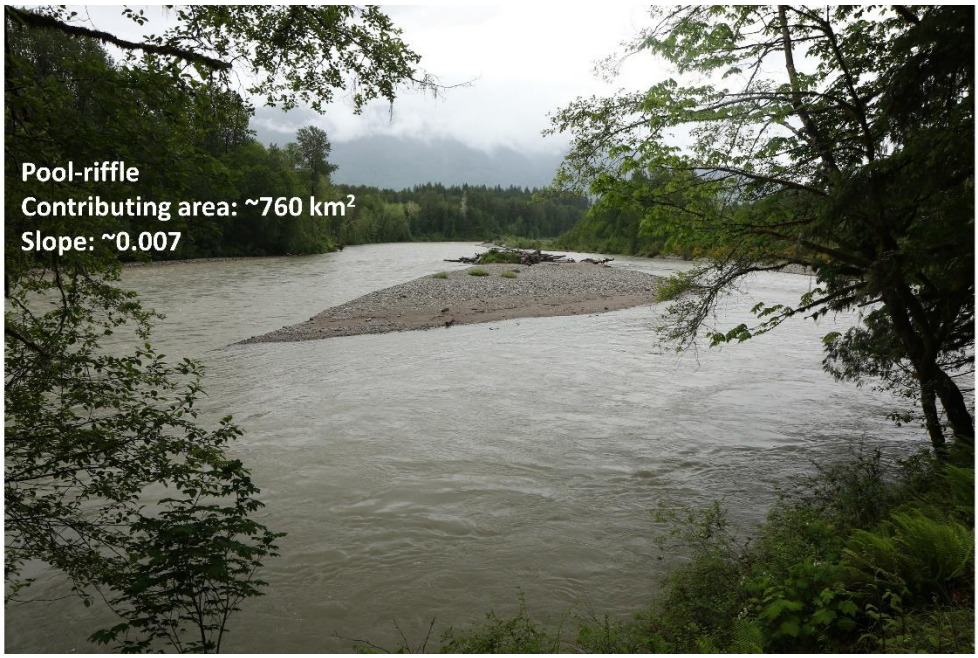

Figure S4 Sauk River, 1 km upstream of the Suiattle River confluence, during a snow-melt event

Figure S5 Cross section 4 in the main stem of the Sauk River

S6 Near the outlet of Red Creek at Cross section 5

S7 Example cascade channel that drains to the main stem Sauk River.

2.7.12 Effective flow

Effective flow, the flow rate that transports the largest proportion of all bedload that passes through a reach (Wolman & Miller, 1960), was computed for each of the 582 channel reaches using the 1-h ψ in the "mean" approach for calculating effective flow described by Lenzi et al. (2006). The "mean" approach involves creating a histogram of cumulative sediment transport relative to flow classes or bins. The central value of the flow bin denotes the value of the bin. Note that the "mean" approach is typically applied to bedload transport observations whereas we applied it to modeled bedload transport capacity. In order to simultaneously compare histograms of cumulative ψ from different channels, we express flow as the ratio to bankfull flow and cumulative ψ in each flow bin as a proportion of total ψ at the reach over the 35 year modeling record. Reaches are combined into channel morphology bins. The cumulative ψ in each channel morphology and flow bin is the average cumulative ψ of all reaches in the bin. The effective flow is determined as the flow bin that contains the highest cumulative ψ value.

2.8 APPENDICES

2.8.1 Appendix A: Wilcock and Crowe (2003) Fractional bedload equations with effective stress

We compute dimensional rating curves of q_s/q_{sb} using effective stresses and the total bedload transport form of the Wilcock and Crowe (2003) bedload transport capacity formula following Schneider et al. (2015):

$$W^*_{tot} = \begin{cases} 0.002\phi^{7.8} & \phi < 1.33 \; and \; D > 4mm \\ 14\left(1 - \frac{0.894}{\phi^{0.5}}\right)^{4.5} & \phi \geq 1.33 \; and \; D > 4mm \end{cases} \quad \text{(a1)}$$

Grain size in the Sauk river is larger than 4mm (Supporting Information, Table S3). Since ψ is defined using only high-shear stress from of (a1), we only use the high-shear stress form of (a1) to compute W^*_{tot}. Transport stage (ϕ) is defined using effective Shields stresses acting on the median grain size as:

$$\phi = \frac{\tau^{*\prime}}{\tau^{*\prime}_r} \quad \text{(a2)}$$

where $\tau^{*\prime}_r$ = 0.03, $\tau^{*\prime}$ is determined from $\tau^{\prime}$ using the Shields equation and the median grain size and $\tau^{\prime}$ is determined for a given flow rate following the same steps used to determine $\tau^{\prime}_b$ (Equations 11-15, section 3.2.2). Volumetric, total bedload transport capacity per unit width (q_s), is:

$$q_s = \frac{W^*_{tot} u^{*3}}{(s-1)g} \quad \text{(a3)}$$

where $u^* = (\tau^{\prime}/\rho)^{0.5}$.

2.8.2. Appendix B: Notation List

A contributing area to reach [L2]
A_x flow cross-sectional area [L2]
d flow depth [L]
D_{50} median grain size [L]
g acceleration due to gravity [L T-2]
m_w exponent of hydraulic geometry relation for channel width
m_l inverse of the slope of the left channel wall [L L-1]
m_r inverse of the slope of the right channel wall [L L-1]
$P_{1\,max}$ maximum 1-h precipitation rate [L T-1]
$P_{24\,max}$ maximum 24-h precipitation rate [L T-1]
P^*_{max} ratio of the maximum 1-h and 24-h precipitation rates
q_s volumetric bedload transport capacity per unit channel width [L3 T-1 L-1]
q_{sb} volumetric bedload transport capacity per unit channel width at the bankfull flow [L3 T-1 L-1]
q^{**} dimensionless unit flow rate
Q flow [L3 T-1]
Q_b bankfull flow [L3 T-1]
Q_r reference flow [L3 T-1]
$Q_{1\,max}$ peak flow rate of the 1-h hydrograph [L3 T-1]
$Q_{24\,max}$ peak flow rate of the 24-h hydrograph [L3 T-1]
Q^*_{max} ratio of the maximum 1-h and 24-h peak flow rate
S slope [L L-1]
s specific gravity
U flow velocity for a given flow [L T-1]
U_o virtual velocity corresponding to the base level resistance
U_b flow velocity at bankfull flow [L T-1]
w channel width [L]
w_x bottom width of cross-sectional area
W^* dimensionless bedload transport rate: the ratio of the power required to transport bedload scaled by the power available
W^*_b dimensionless bedload transport rate at bankfull flow
ρ density of water [M L-3]
ϕ transport stage – ratio of effective shear stress to reference shear stress
ϕ_b transport stage at bankfull flow - ratio of effective shear stress at bankfull flow to reference shear stress
τ total shear stress - shear stress exerted by flow on the channel [ML-1 T-2]
τ_r total reference shear stress - [M L-1 T-2]
τ_b total shear stress at bankfull flow [M L-1 T-2]
τ' effective stress – shear stress exerted by the flow on the grains [M L-1 T-2]
τ'_r effective reference stress [M L-1 T-2]
τ'_b effective shear stress at bankfull flow [M L-1 T-2]

τ^*	total Shields stress
$\tau^{*\prime}$	effective Shields stress
τ_r^*	total reference Shields stress
τ_b^*	total bankfull Shields stress
τ_b^*/τ_r^*	bankfull-excess Shields stress
τ_b'/τ_r'	effective-bankfull-excess shear stress
ψ	non-dimensional bedload transport capacity as defined by Equation 5
ψ_s	explicitly determined non-dimensional bedload transport capacity using the Wilcock and Crowe (2003) bedload formula and effective stress to determine q_s and q_{sb}
ψ_{T1}	total non-dimensional bedload transport in the 1-h sediograph
ψ_{T24}	total non-dimensional bedload transport in the 24-h sediograph
ψ_T^*	hydrometeorologic bedload error (HBE): ratio of the 1-h and 24-h total non-dimensional bedload transport

2.9 REFERENCES

1. Asano, Y., & Uchida, T. (2018). The roles of channels and hillslopes in rainfall/run-off lag times during intense storms in a steep catchment. *Hydrological Processes*, 32(6), 713–728. https://doi.org/10.1002/hyp.11443
2. Badoux, A., Turowski, J. M., Mao, L., Mathys, N., & Rickenmann, D. (2012). Rainfall intensity–duration thresholds for bedload transport initiation in small alpine watersheds. *Natural Hazards and Earth System Sciences*, *12*(10), 3091–3108. https://doi.org/10.5194/nhess-12-3091-2012
3. Bandaragoda, C., Lee S., Istanbulluoglu E., & Hamlet A. *Hydrology, Stream Temperature and Sediment Impacts of Climate Change in the Sauk River Basin. Report prepared for Sauk-Suiattle Indian Tribe, Darrington, WA and the Skagit Climate Consortium, Mt. Vernon, WA.* (2019). Retrieved from https://www.hydroshare.org/resource/e5ad2935979647d6af5f1a9f6bdecdea/
4. Barry, J. J., Buffington, J. M., & King, J. G. (2004). A general power equation for predicting bed load transport rates in gravel bed rivers. *Water Resources Research*, 40(10). https://doi.org/10.1029/2004wr003190
5. Barry, J. J., Buffington, J. M., & King, J. G. (2007). Correction to "A general power equation for predicting bed load transport rates in gravel bed rivers." *Water Resources Research*, 43(8). https://doi.org/10.1029/2007wr006103
6. Barry, J. J., Buffington, J. M., Goodwin, P., King, J. G., & Emmett, W. W. (2008). Performance of bed-load transport equations relative to geomorphic significance: predicting effective discharge and its transport rate. *Journal of Hydraulic Engineering*, 134(5), 601–615. https://doi.org/10.1061/(asce)0733-9429(2008)134:5(601)
7. Beechie, T.J., Collins, B.D. & Pess, G.R. 2001. Holocene and recent geomorphic processes, land use, and salmonid habitat in two north Puget Sound river basins. In J.M. Dorava, D.R. Montgomery, B.B. Palcsak and F.A. Fitzpatrick (Eds.) *Geomorphic processes and riverine habitat* (pp.37-54). DOI:10.1029/WS004
8. Benda, L., & Dunne, T. (1997). Stochastic forcing of sediment routing and storage in channel networks. *Water Resources Research*, *33*(12), 2865–2880. https://doi.org/10.1029/97wr02387
9. Bennett, A., Hamman, J., & Nijssen, B. (2020). MetSim: A Python package for estimation and disaggregation of meteorological data. *Journal of Open Source Software*, 5(47), 2042. https://doi.org/10.21105/joss.02042
10. Beven, K. (1982). On subsurface stormflow: an analysis of response times. *Hydrological* Sciences Journal, 27(4), 505–521. https://doi.org/10.1080/02626668209491129
11. Beven, K. (2006). A manifesto for the equifinality thesis. Journal of Hydrology, 320(1–2), 18–36. https://doi.org/10.1016/j.jhydrol.2005.07.007
12. Bras, R. L. (1979). Sampling of interrelated random fields: The rainfall-runoff case. *Water Resources Research*, 15(6), 1767–1780. https://doi.org/10.1029/wr015i006p01767
13. Buffington, J. M., & Montgomery, D. R. (1999a). Effects of hydraulic roughness on surface textures of gravel-bed rivers. *Water Resources Research*, *35*(11), 3507–3521. https://doi.org/10.1029/1999wr900138
14. Buffington, J. M., & Montgomery, D. R. (1999b). Effects of sediment supply on surface textures of gravel-bed rivers. *Water Resources Research*, *35*(11), 3523–3530. https://doi.org/10.1029/1999wr900232

15. Buffington JM. 2012. Changes in channel morphology over human time scales. In *Gravel-bed rivers: processes, tools, environments*, Church M, Biron PM, Roy AG (eds). Wiley: Chichester; 181–213.
16. Buffington, J.M., Montgomery, D.R., (2021) Geomorphic classification of rivers: An updated review. *Treatise on Geomorpholo*gy. (2)1-47.
17. Bunte, K., Abt, S. R., Swingle, K. W., Cenderelli, D. A., & Schneider, J. M. (2013). Critical Shields values in coarse-bedded steep streams. *Water Resources Research*, 49(11), 7427–7447. https://doi.org/10.1002/2012wr012672
18. Castro, J. M., & Jackson, P. L. (2001). Bankfull discharge recurrence intervals and regional hydraulic geometry relationships: patterns in the Pacific Northwest, USA. *Journal of the American Water Resources Association,* 37(5), 1249–1262. https://doi.org/10.1111/j.1752-1688.2001.tb03636.x
19. Chen, L., Schumer, R., Knust, A., & Forsee, W. (2011). Impact of temporal resolution of flow-duration curve on sediment load estimation. *Journal of the American Water Resources Association*, 48(1), 145–155. https://doi.org/10.1111/j.1752-1688.2011.00602.x
20. Chen, X., Leung, L. R., Gao, Y., Liu, Y., Wigmosta, M., & Richmond, M. (2018). Predictability of extreme precipitation in western U.S. watersheds based on atmospheric river occurrence, intensity, and duration. *Geophysical Research Letters*, 45(21), 11693-11701 https://doi.org/10.1029/2018gl079831
21. Costa, J.E., O'Connor, J.E., 1995. Geomorphically effective floods. In: Costa, J.E., Miller, A.J., Potter, K.P., Wilcock, P.R. (Eds.), *Natural and anthropogenic influences in fluvial geomorphology* (The Wolman Volume) AGU Geophysical Monograph. American Geophysical Union, Washington, D.C., pp. 45–56.
22. Cristea, N. C., Lundquist, J. D., Loheide, S. P., Lowry, C. S., & Moore, C. E. (2013). Modelling how vegetation cover affects climate change impacts on streamflow timing and magnitude in the snowmelt-dominated upper Tuolumne Basin, Sierra Nevada. *Hydrological Processes,* 28(12), 3896–3918. https://doi.org/10.1002/hyp.9909
23. Cuo, L., Giambelluca, T. W., & Ziegler, A. D. (2011). Lumped parameter sensitivity analysis of a distributed hydrological model within tropical and temperate catchments. *Hydrological Processes, 25*(15), 2405–2421. https://doi.org/10.1002/hyp.8017
24. Dan, L., Ji, J., Xie, Z., Chen, F., Wen, G., & Richey, J. E. (2012). Hydrological projections of climate change scenarios over the 3H region of China: A VIC model assessment. *Journal of Geophysical Research: Atmospheres*, 117(D11), https://doi.org/10.1029/2011jd017131
25. Das, B. M. (2005). Principles of geotechnical engineering (5th Edition). Pacific Grove, CA: Brooks/Cole.
26. Dietrich, W. E., Kirchner, J. W., Ikeda, H., & Iseya, F. (1989). Sediment supply and the development of the coarse surface layer in gravel-bedded rivers. *Nature*, 340(6230), 215–217. https://doi.org/10.1038/340215a0
27. DeVries, P., 2008, *Results of geomorphic analysis of hydrologic, sediment, and habitat conditions in the upper Sauk River between Clear Creek and the Suiattle River: Technical Memorandum prepared for Snohomish County Surface Water Management,* R2 Resource Consultants, Inc.
28. DeVries, P., and Madsen, S., 2008, *Avulsion risk assessment for Sauk River—Clear Creek to Suiattle River: Technical Memorandum prepared for Snohomish County Surface Water Management,* R2 Resource Consultants, Inc.

29. D'Odorico, P., & Rigon, R. (2003). Hillslope and channel contributions to the hydrologic response. *Water Resources Research*, 39(5), 1113. https://doi.org/10.1029/2002wr001708
30. Du, E., Link, T. E., Gravelle, J. A., & Hubbart, J. A. (2014). Validation and sensitivity test of the distributed hydrology soil-vegetation model (DHSVM) in a forested mountain watershed. *Hydrological Processes*, 28(26), 6196–6210. https://doi.org/10.1002/hyp.10110
31. Dunne, T., & Black, R. D. (1970). Partial area contributions to storm runoff in a small New England watershed. *Water Resources Research*, 6(5), 1296–1311. https://doi.org/10.1029/wr006i005p01296
32. Eagleson, P. S., & Shack, W. J. (1966). Some criteria for the measurement of rainfall and runoff. *Water Resources Research*, 2(3), 427–436. https://doi.org/10.1029/wr002i003p00427
33. Feld, S. I., Cristea, N. C., & Lundquist, J. D. (2013). Representing atmospheric moisture content along mountain slopes: Examination using distributed sensors in the Sierra Nevada, California. *Water Resources Research*, 49(7), 4424–4441. https://doi.org/10.1002/wrcr.20318
34. Frans, C., Istanbulluoglu, E., Lettenmaier, D. P., Fountain, A. G., & Riedel, J. (2018). Glacier recession and the response of summer streamflow in the Pacific Northwest United States, 1960–2099. *Water Resources Research*, 54(9), 6202–6225. https://doi.org/10.1029/2017wr021764
35. Gaeuman, D., Stewart, R. L., & Pittman, S. (2018). Toward the prediction of bed load rating curve parameter values: the influence of scale, particle size, and entrainment threshold. *Water Resources Research*, 54(5), 3313–3334. https://doi.org/10.1002/2017wr021627
36. Grant, G. E., Lewis, S. L., Swanson, F. J., Cissel, J. H., & McDonnell, J. J. (2008). *Effects of Forest Practices on Peak Flows and Consequent Channel Response: A State-of- Science Report for Western Oregon and Washington* (PNW-GTR-760). United States Department of Agriculture, Forest Service, Pacific Northwest Research Station. Retrieved from https://www.fs.usda.gov/treesearch/pubs/30179
37. Grant, G. E., Swanson, F. J., & Wolman, M. G. (1990). Pattern and origin of stepped-bed morphology in high-gradient streams, Western Cascades, Oregon. *Geological Society of America Bulletin*, *102*(3), 340–352.
38. Hassan, M. A., Brayshaw, D., Alila, Y., & Andrews, E. (2014). Effective discharge in small formerly glaciated mountain streams of British Columbia: Limitations and implications. *Water Resources Research*, 50(5), 4440–4458. https://doi.org/10.1002/2013wr014529
39. Houska, T., Kraft, P., Chamorro-Chavez, A., & Breuer, L. (2015). SPOTting Model Parameters Using a Ready-Made Python Package. *PLOS ONE*, 10(12), e0145180. https://doi.org/10.1371/journal.pone.0145180
40. Istanbulluoglu, E., Bras, R.L., Flores-Cervantes, H., Tucker, G.E. (2005). Implications of bank failures and fluvial erosion for gully development: Field observations and modeling. *Journal of Geophysical Research*, 110(F1). https://doi.org/10.1029/2004jf000145
41. Jaeger, K.L., Curran, C.A., Anderson, S.W., Morris, S.T., Moran, P.W., and Reams, K.A.,(2017) *Suspended sediment, turbidity, and stream water temperature in the Sauk River Basin, Washington, water years 2012–16*: *U.S. Geological Survey Scientific Investigations Report 2017–5113*, 47 p., https://doi.org/10.3133/sir20175113
42. Jan C-D, Lee M-H (2004) A debris-flow rainfall-based warning model. *Journal of Chinese Soil and Water Conservation* 35(3):275–285
43. Jassal, R. S., Black, T. A., Spittlehouse, D. L., Brümmer, C., & Nesic, Z. (2009). Evapotranspiration and water use efficiency in different-aged Pacific Northwest Douglas-fir

stands. *Agricultural and Forest Meteorology*, 149(6–7), 1168–1178. https://doi.org/10.1016/j.agrformet.2009.02.004
44. Jones, J. A., & Grant, G. E. (1996). Peak Flow Responses to Clear-Cutting and Roads in Small and Large Basins, Western Cascades, Oregon. *Water Resources Research*, *32*(4), 959–974. https://doi.org/10.1029/95wr03493
45. Kelleher, C., Wagener, T., & McGlynn, B. (2015). Model-based analysis of the influence of catchment properties on hydrologic partitioning across five mountain headwater subcatchments. *Water Resources Research*, *51*(6), 4109–4136. https://doi.org/10.1002/2014wr016147
46. Lane, E.W. (1955). The importance of fluvial morphology in hydraulic engineering. *Proceedings of the American Society of Civil Engineers* 81 (art. 745)
47. Lenzi, M., Mao, L., & Comiti, F. (2006). Effective discharge for sediment transport in a mountain river: Computational approaches and geomorphic effectiveness. *Journal of Hydrology*, 326(1–4), 257–276. https://doi.org/10.1016/j.jhydrol.2005.10.031
48. Livneh, B., Rosenberg, E. A., Lin, C., Nijssen, B., Mishra, V., Andreadis, K. M., Lettenmaier, D. P. (2013). A long-term hydrologically based dataset of land surface fluxes and states for the conterminous United States: update and extensions. *Journal of Climate*, 26(23), 9384–9392. https://doi.org/10.1175/jcli-d-12-00508.1
49. Lundquist, J. D., Dettinger, M. D., & Cayan, D. R. (2005). Snow-fed streamflow timing at different basin scales: Case study of the Tuolumne River above Hetch Hetchy, Yosemite, California. *Water Resources Research*, 41(7). https://doi.org/10.1029/2004wr003933
50. Lundquist, J., Hughes, M., Gutmann, E., & Kapnick, S. (2019). Our skill in modeling mountain rain and snow is bypassing the skill of our observational networks. *Bulletin of the American Meteorological Society*, *100*(12), 2473–2490. https://doi.org/10.1175/bams-d-19-0001.1
51. McGlynn, B. L., McDonnell, J. J., Seibert, J., & Kendall, C. (2004). Scale effects on headwater catchment runoff timing, flow sources, and groundwater-streamflow relations. *Water Resources Research*, 40(7). W07504, https://doi.org/10.1029/2003wr002494
52. McGuire, K. J., & McDonnell, J. J. (2010). Hydrological connectivity of hillslopes and streams: characteristic time scales and nonlinearities. *Water Resources Research*, 46(10). W10543, https://doi.org/10.1029/2010wr009341
53. Mesinger, F., DiMego, G., Kalnay, E., Mitchell, K., Shafran, P. C., Ebisuzaki, W., Shi, W. (2006). North American regional reanalysis. *Bulletin of the American Meteorological Society*, 87(3), 343–360. https://doi.org/10.1175/bams-87-3-343
54. Meyer, P. D., Rockhold, M. L., & Gee, G. W. (1997). *Uncertainty analyses of infiltration and subsurface flow and transport for SDMP Sites* (NUREG/CR-6565 PNNL-11705). Richland, WA: Pacific Northwest National Laboratory.
55. Minder, J. R., Mote, P. W., & Lundquist, J. D. (2010). Surface temperature lapse rates over complex terrain: Lessons from the Cascade Mountains. *Journal of Geophysical Research*, 115, D14122. https://doi.org/10.1029/2009jd013493
56. Montgomery, D. R., & Foufoula-Georgiou, E. (1993). Channel network source representation using digital elevation models. *Water Resources Research*, 29(12), 3925–3934. https://doi.org/10.1029/93wr02463
57. Montgomery, D. R., & Buffington, J. M. (1997). Channel-reach morphology in mountain drainage basins. *Geological Society of America Bulletin*, 109(5), 596–611.

58. Nash, D. B. (1994). Effective sediment-transporting discharge from magnitude-frequency analysis. *The Journal of Geology*, 102(1), 79–95. https://doi.org/10.1086/629649
59. Nash, J., & Sutcliffe, J. (1970). River flow forecasting through conceptual models part I — A discussion of principles. *Journal of Hydrology*, *10*(3), 282–290. https://doi.org/10.1016/0022-1694(70)90255-6
60. National Weather Service. (2020). COOP. Retrieved 2020, from https://www.weather.gov/coop/
61. Natural Resource Conservation Service, United States Dept. of Agriculture. (2020). Retrieved 2020, from https://www.wcc.nrcs.usda.gov/snow/
62. Neiman, P. J., Schick, L. J., Ralph, F. M., Hughes, M., & Wick, G. A. (2011). Flooding in western Washington: the connection to atmospheric rivers. *Journal of Hydrometeorology*, 12(6), 1337–1358. https://doi.org/10.1175/2011jhm1358.1
63. Palucis, M. C., & Lamb, M. P. (2017). What controls channel form in steep mountain streams? *Geophysical Research Letters*, 44(14), 7245–7255. https://doi.org/10.1002/2017gl074198
64. Parker, G. (1978). Self-formed straight rivers with equilibrium banks and mobile bed. Part 2. The gravel river. *Journal of Fluid Mechanics*, 89(1), 127–146. https://doi.org/10.1017/s0022112078002505
65. Parker, G., Klingeman, P. C., & McLean, D. G. (1982). Bedload and size distribution in paved gravel-bed streams. *Journal of the Hydraulics*, 108(4), 544–571. https://doi.org/10.1061/jyceaj.0005854
66. Penna, D., Tromp-van Meerveld, H. J., Gobbi, A., Borga, M., & dalla Fontana, G. (2011). The influence of soil moisture on threshold runoff generation processes in an alpine headwater catchment. *Hydrology and Earth System Sciences*, 15(3), 689–702. https://doi.org/10.5194/hess-15-689-2011
67. Pfeiffer, A. M., Finnegan, N. J., & Willenbring, J. K. (2017). Sediment supply controls equilibrium channel geometry in gravel rivers. *Proceedings of the National Academy of Sciences*, 114(13), 3346–3351. https://doi.org/10.1073/pnas.1612907114
68. Phillips, C. B., Hill, K. M., Paola, C., Singer, M. B., & Jerolmack, D. J. (2018). Effect of flood hydrograph duration, magnitude, and shape on bed load transport dynamics. *Geophysical Research Letters*, 45(16), 8264–8271. https://doi.org/10.1029/2018gl078976
69. Phillips, C. B., & Jerolmack, D. J. (2019). Bankfull transport capacity and the threshold of motion in coarse-grained rivers. *Water Resources Research*, 55(12), 11316–11330. https://doi.org/10.1029/2019wr025455
70. Rawls W. J., Brakensiek D. L., & Saxtonn K. E.. (1982). Estimation of soil water properties. *Transactions of the ASAE*, 25(5), 1316–1320. https://doi.org/10.13031/2013.33720
71. Rickenmann, D., & Recking, A. (2011). Evaluation of flow resistance in gravel-bed rivers through a large field data set. *Water Resources Research*, 47(7). https://doi.org/10.1029/2010wr009793
72. Rinaldo, A., Marani, A., & Rigon, R. (1991). Geomorphological dispersion. *Water Resources Research*, 27(4), 513–525. https://doi.org/10.1029/90wr02501
73. Rinaldo, A., Vogel, G. K., Rigon, R., & Rodriguez-Iturbe, I. (1995). Can one gauge the shape of a basin? *Water Resources Research*, 31(4), 1119–1127. https://doi.org/10.1029/94wr03290
74. Robinson, J. S., Sivapalan, M., & Snell, J. D. (1995). On the relative roles of hillslope processes, channel routing, and network geomorphology in the hydrologic response of natural

catchments. *Water Resources Research*, 31(12), 3089–3101. https://doi.org/10.1029/95wr01948
75. Rosburg, T. T., Nelson, P. A., Sholtes, J. S., & Bledsoe, B. P. (2016). The effect of flow data resolution on sediment yield estimation and channel design. *Journal of Hydrology*, 538, 429–439. https://doi.org/10.1016/j.jhydrol.2016.04.040
76. Salathé, E. P., Hamlet, A. F., Mass, C. F., Lee, S. Y., Stumbaugh, M., & Steed, R. (2014). Estimates of twenty-first-century flood risk in the Pacific Northwest based on regional climate model simulations. *Journal of Hydrometeorology*, 15(5), 1881–1899. https://doi.org/10.1175/jhm-d-13-0137.1
77. Schneider, J. M., Rickenmann, D., Turowski, J. M., Bunte, K., & Kirchner, J. W. (2015). Applicability of bed load transport models for mixed-size sediments in steep streams considering macro-roughness. *Water Resources Research*, 51(7), 5260–5283. https://doi.org/10.1002/2014wr016417
78. Schneider, J. M., Rickenmann, D., Turowski, J. M., Schmid, B., & Kirchner, J. W. (2016). Bed load transport in a very steep mountain stream (Riedbach, Switzerland): measurement and prediction. *Water Resources Research*, 52(12), 9522–9541. https://doi.org/10.1002/2016wr019308
79. Scott, D. N., & Collins, B. D. (2021). Frequent Mass Movements From Glacial and Lahar Terraces, Controlled by Both Hillslope Characteristics and Fluvial Erosion, are an Important Sediment Source to Puget Sound Rivers. *Water Resources Research*, 57(4). https://doi.org/10.1029/2020wr028389
80. Shrestha, R. R., Schnorbus, M. A., Werner, A. T., & Berland, A. J. (2012). Modelling spatial and temporal variability of hydrologic impacts of climate change in the Fraser River basin, British Columbia, Canada. *Hydrological Processes*, *26*(12), 1840–1860. https://doi.org/10.1002/hyp.9283
81. Singh, V.P. (1997) Effect of spatial and temporal variability in rainfall and watershed characteristics on stream flow hydrograph. *Hydrological Processes*, 11(12), 1649-1669.
82. Skamarock, W. C., Klemp, J. B., Dudhia, J., Gill, D. O., Barker, D. M., Wang, W., & Powers, J. G. (2008). *A description of the Advanced Research WRF version 3*. (NCAR Technical note-475+STR). Retrieved from http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.484.3656
83. Sólyom, P. B. and Tucker, G. E (2004). Effect of limited storm duration on landscape evolution, drainage basin geometry, and hydrograph shapes. *Journal of Geophysical Research*, 109(F3). https://doi.org/10.1029/2003jf000032
84. SSURGO Database | Natural Resource Conservation Service , United States Dept. of Agriculture. (2018). Retrieved 2018, from https://www.nrcs.usda.gov/wps/portal/nrcs/detail/soils/survey/?cid=nrcs142p2_053627
85. Storck, P., Bowling, L., Wetherbee, P., & Lettenmaier, D. (1998). Application of a GIS-based distributed hydrology model for prediction of forest harvest effects on peak stream flow in the Pacific Northwest. *Hydrological Processes*, 12(6), 889–904.
86. Sturm, T. W. (2010). *Open Channel Hydraulics* (2nd edition). Singapore: McGraw-Hill Medical Publishing.
87. Surfleet, C. G., Skaugset, A. E., & McDonnell, J. J. (2010). Uncertainty assessment of forest road modeling with the Distributed Hydrology Soil Vegetation Model (DHSVM). *Canadian Journal of Forest Research*, 40(7), 1397–1409. https://doi.org/10.1139/x10-079

88. Swinbank, W. C. (1963). Long-wave radiation from clear skies. *Quarterly Journal of the Royal Meteorological Society*, 89(381), 339–348. https://doi.org/10.1002/qj.49708938105
89. Tabor, R.W., Booth, D.B., Vance, J.A., & Ford, A.B (2002) *Geologic map of the Sauk River 30- by 60-minute quadrangle*, Washington. Geologic Investigations Series I-2592, USGS
90. Tan, K. S., Chiew, F. H. S., & Grayson, R. B. (2007). A steepness index unit volume flood hydrograph approach for sub-daily flow disaggregation. *Hydrological Processes*, 21(20), 2807–2816. https://doi.org/10.1002/hyp.6501
91. Tang, W., & Carey, S. K. (2017). HydRun: A MATLAB toolbox for rainfall-runoff analysis. *Hydrological Processes*, 31(15), 2670–2682. https://doi.org/10.1002/hyp.11185
92. Tarboton, D. G., Bras, R. L., & Rodriguez-Iturbe, I. (1991). On the extraction of channel networks from digital elevation data. *Hydrological Processes*, 5(1), 81–100. https://doi.org/10.1002/hyp.3360050107
93. Thornton, P. E., & Running, S. W. (1999). An improved algorithm for estimating incident daily solar radiation from measurements of temperature, humidity, and precipitation. Agricultural and Forest Meteorology, 93(4), 211–228. https://doi.org/10.1016/s0168-1923(98)00126-9
94. Tucker, G. E., & Bras, R. L. (2000). A stochastic approach to modeling the role of rainfall variability in drainage basin evolution. *Water Resources Research*, 36(7), 1953–1964. https://doi.org/10.1029/2000wr900065
95. United States Geological Survey (2020) Current water data for the nation. Retrieved 2020, from https://waterdata.usgs.gov/nwis/rt
96. Unsworth, M. H., & Monteith, J. L. (1975). Long-wave radiation at the ground I. Angular distribution of incoming radiation. *Quarterly Journal of the Royal Meteorological Society*, 101(427), 13–24. https://doi.org/10.1002/qj.49710142703
97. Vrugt, J. A., & Ter Braak, C. J. F. (2011). DREAM: an adaptive Markov Chain Monte Carlo simulation algorithm to solve discrete, noncontinuous, and combinatorial posterior parameter estimation problems. *Hydrology and Earth System Sciences*, 15(12), 3701–3713. https://doi.org/10.5194/hess-15-3701-2011
98. Vrugt, J. A. (2016). Markov chain Monte Carlo simulation using the DREAM software package: theory, concepts, and MATLAB implementation. *Environmental Modelling & Software*, 75, 273–316. https://doi.org/10.1016/j.envsoft.2015.08.013
99. Washington Department of Natural Resources (2018) Lidar Portal. Retrieved 2018 from https://lidarportal.dnr.wa.gov/
100. Washington Department of Natural Resources (2020) Rain on Snow ArcGIS layer. Retrieved 2020, from https://gis.dnr.wa.gov/site2/rest/services/Public_Forest_Practices/WADNR_PUBLIC_FP_Watershed/MapServer/6
101. Waichler, S. R., Wemple, B. C., & Wigmosta, M. S. (2005). Simulation of water balance and forest treatment effects at the H.J. Andrews Experimental Forest. *Hydrological Processes*, 19(16), 3177–3199. https://doi.org/10.1002/hyp.5841
102. Wehner, M., Lee, J., Risser, M., Ullrich, P., Gleckler, P., & Collins, W. D. (2021). Evaluation of extreme sub-daily precipitation in high-resolution global climate model simulations. *Philosophical Transactions of the Royal Society* A, 379(2195), 20190545. https://doi.org/10.1098/rsta.2019.054
103. Westrick, K. (1999). Soil depth calculation script implemented in python. Available from https://github.com/pnnl/DHSVM-PNNL (accessed January 2019)

104. Whiting, P. J., Stamm, J. F., Moog, D. B., & Orndorff, R. L. (1999). Sediment-transporting flows in headwater streams. *Geological Society of America Bulletin*, 111(3), 0450–0466.
105. Wigmosta, M. S., Vail, L. W., & Lettenmaier, D. P. (1994). A distributed hydrology-vegetation model for complex terrain. *Water Resources Research*, 30(6), 1665–1679. https://doi.org/10.1029/94wr00436
106. Wigmosta M.S. & Perkins W.A (2001) Simulating the Effects of Forest Roads on Watershed Hydrology Land Use in Wigmosta, M. S., & Burges, S. J. (Eds.), *Watersheds: Human Influence on Hydrology and Geomorphology in Urban and Forest Areas*. Water Science and Application 2. Washington DC : American Geophysical Union. https://doi.org/10.1029/ws002
107. Wigmosta MS, Nijssen B, Storck P. 2002. The distributed hydrology soil vegetation model. In Singh VP, Frevert DK (Eds.), *Mathematical models of dmall eatershed hydrology applications*, Water Resources Publications LCC, Highlands Ranch, Colorado.
108. Wilcock, P. R. (2001). Toward a practical method for estimating sediment-transport rates in gravel-bed rivers. *Earth Surface Processes and Landforms*, 26(13), 1395–1408. https://doi.org/10.1002/esp.301
109. Wilcock, P. R., & Crowe, J. C. (2003). Surface-based transport model for mixed-size sediment. *Journal of Hydraulic Engineering*, 129(2), 120–128. https://doi.org/10.1061/(asce)0733-9429(2003)129:2(120)
110. Wilcock, P., Pitlick, J; Cui, Y (2009*) Sediment transport primer: estimating bed-material transport in gravel-bed rivers*. (Gen. Tech. Rep. RMRS-GTR-226). Fort Collins, CO: U.S. Department of Agriculture, Forest Service, Rocky Mountain Research Station.
111. Wolman, M. G., & Miller, J. P. (1960). Magnitude and Frequency of Forces in Geomorphic Processes. *The Journal of Geology*, 68(1), 54–74. https://doi.org/10.1086/626637
112. Yager, E. M., Dietrich, W. E., Kirchner, J. W., & McArdell, B. W. (2012). Prediction of sediment transport in step-pool channels. *Water Resources Research*, 48(1). https://doi.org/10.1029/2011wr010829

CHAPTER 3. PROBABILISTIC LANDSLIDE RUNOUT, SEDIMENT DELIVERY AND TOPOGRAPHIC CHANGE USING MASSWASTINGRUNOUT

3.0 ABSTRACT

We developed a new rule-based cellular-automaton debris flow algorithm for predicting the hazard extent, sediment transport and topographic change associated with the runout of a potential landslide. This algorithm, which we call MassWastingRunout, is coded in Python and implemented as a component for the package Landlab. Given the location and geometry of an initial landslide body (i.e., landslide polygon), MWR models the downslope progression of the runout process and evolves the underlying terrain. Runout behavior is controlled by mass continuity, underlying-topographic slope and field-data-informed rules for erosion and deposition. MWR includes a calibration utility that uses a Markov Chain Monte Carlo algorithm to sample model parameter space and tune the model to match observed landslide runout extent, deposition and erosion. Output from the calibration utility informs probabilistic implementation of MWR. Here we demonstrate calibrated model performance relative to a range of observed runout phenomena and terrain, including debris flows in channelized, low-energy-dissipation terrains and debris avalanches on open-slope, moderate, to high-energy-dissipation terrains. We test model ability to predict runout probability at a case study site using parameters that were determined through calibration to a different site. Finally, we show how to use a calibrated MWR model to determine runout-probability from an expert-defined, potentially unstable slope and a landslide hazard map.

3.1 INTRODUCTION

Over geologic timescales, landslides (e.g., Hungr et al. 2014) and their runout shape the topographic expression of mountain ranges and channel networks (Campforts et al., 2022; Korup, 2006; Larsen and Montgomery, 2012; Montgomery and Deitrich, 1988). Over more pragmatic engineering and environmental risk management timescales, landslides and their runout pose a primary risk to humans and infrastructure but also support numerous ecosystem benefits, including carbon and nutrient exports from hillslopes and the creation of ecological habitats in the receiving stream (Benda et al., 2003; Bigelow et al., 2007; Goode et al., 2012). Therefore, explicit representation of the landslide runout should be an integral component of (1) landscape evolution models, with emphasis on topographic change prediction; (2) landslide risk assessment tools, with emphasis on probabilistic impact and; (3) sediment budget models, with emphasis on the mobilization and deposition of sediment carried by the landslide.

Landslide runout processes can be generalized into three phases: initiation, erosion and deposition. After a landslide initiates, it may break apart and become a granular flow such as a debris flow or dry debris slide. Field observations show that erosion rates are sensitive to the moisture content of the channel bed (McCoy et al. 2012) and, although highly variable, generally increase with flow depth (Schürch et al. 2011). At the same time, laboratory experiments show that entrainment rates vary with granular stress (Capart et al, 2015) and increase with flow grainsize (Egashira et al., 2001). Deposition occurs by layered accretion rather than emplacement of a single, massive deposit (Major, 1997) and the length of the deposit is a function of resistance to the flow, the initial landslide volume and the volume of material entrained during the runout (Iverson, 1997). The grain to water ratio and friction angle of the slide material control the slope at which deposition begins (Takahashi, 2014, Major and Iverson, 1999; Zhou et al., 2019) but the

friction angle of the material varies as a function of the grains in the flow and fluidization (fluctuation of interstitial fluid and the solid particles) of the flow material (Hutter et al. 1996). If the water content of the runout material is high enough, as the solid fraction of the debris flow compresses, the water is squeezed out of the debris flow and may continue to transport sediment as an immature debris flow (sensu Takahashi, 2014) or intense bedload (sensu Capart & Fraccarolo, 2011), which also causes the deposition angle to decrease and extent of the runout to increase.

Existing runout models attempt to represent the above processes with a variety of different numerical approaches and level of detail. Most models are parameterized from measurements of frictional characteristics of the runout material and model performance is validated against historic observations (e.g., measurements of an actual landslide and runout). These models can broadly be grouped into four categories: (1) site-specific-empirical/statistical models (Benda and Cundy 1990, Fannin and Wise, 2001; Miller and Bennet, 2008; Reid et al. 2016); (2) detailed, continuum-based mechanistic models that predict flow physics and granular dynamics (Takahashi, 1978; Iverson and Denlinger, 2001); (3) top-down, simplified abstractions (i.e. Murray, 2003) of the key physical processes that control the geomorphic phenomenon (e.g., Clerici and Perego, 2000; Guthrie et al., 2008); note that this group might include simple flow-routing models or reduced complexity models and (4) hybrid modeling approaches that combine detailed, mechanistic approaches with empirical models and top-down, rule-based abstractions (D'Ambrosio et al., 2003; Iovine et al., 2005; Lancaster et al., 2003; McDougall and Hungr 2004; Medina et al., 2008; Frank et al, 2015; Han et al., 2017, 2021). While any of the above models can be implemented in a probabilistic manner for risk assessment purposes, simply by repeatedly running the model for different initial landslide volumes or parameter values (e.g. Hurlimann et al., 2008), these models

generally either lack a built-in probabilistic implementation or are too computationally intensive for Monte Carlo simulations. In addition, only a few models are designed for use in watershed sediment budget and landscape evolution applications, which involves evolving both the topography and regolith during the runout event.

Of the landslide runout models employed in landscape evolution applications, they generally include the removal and redistribution of hillslope material via the initial landslide and subsequent runout but ignore erosion and entrainment of new material during the landslide runout. Deposition of eroded landslide material is often calculated by redistributing it downstream either based on a depositional slope threshold, or through a calculated transport distance formulation (Carretier et al., 2016; Campforts et al., 2020). Landslide hazard or risk maps typically use a hillslope stability model to delineate regional or watershed-scale landslide risk at the initiation zone (e.g., Wu et al., 1995; Lee and Park, 2015; Lee et al., 2018; Strauch et al., 2017) but exclude the subsequent runout hazard. While fully mechanistic models can predict runout, their computational expense and heavily parameterized entrainment rules can limit their application at the scale required for regional mapping and rapid response assessments (Hurlimann et al., 2008; Reid et al., 2016; Gorr et al., 2022). Recently developed hybrid, top-down and empirical-statistical methods provide opportunities for mapping downslope landslide hazard (Gorr et al., 2022; Guthrie & Befus, 2021; Horton et al., 2013; Liu et al., 2022, Wallace et al., 2022), but the models are still heavily parameterized and may require extensive regional observations to calibrate.

Here we describe a new rule-based cellular-automaton algorithm for predicting the hazard extent, sediment transport and topographic change caused by the runout of a landslide. This algorithm is coded in Python and implemented as a set of components for the Landlab earth surface modeling toolkit (Hobley et al., 2017; Barnhart et al., 2020) and is called MassWastingRunout

(MWR). MWR is designed to (1) model topographic change caused by the erosion and deposition that occurs during the landslide runout and; (2) map runout impact in space probabilistically as a function of uncertainties in the initial landslide volume and model parameters. Model behavior is controlled by several parameters that can be inferred from field observations and calibration. For the purposes of rapidly calibrating MWR to the observed runout extent and depositional patterns of a specific site, we developed a Markov Chain Monte Carlo (MCMC) calibration utility. Additionally, MWR tracks attributes of the regolith such as organic content and sediment size, represented as a numerical value on a model grid and transferred as a conserving and well-mixed mass with the runout material.

In this paper, we present the MWR model (Section 2), describe the calibration utility and probabilistic implementation of MWR (Section 3) and demonstrate calibrated model performance at four topographically and geologically unique field sites (Section 4). The field sites include six different landslides and encompass the low-, moderate- and high-energy-dissipation settings (Figure 1) described by Nicoletti and Sorriso-Valo (1991). The sites also span a range of landslide-runout erosion rates. After demonstrating calibrated model performance, we then test the model using the parameterization of one site to predict runout probability at a neighboring site. Finally, we show how to couple MWR with a probabilistically modeled landslide hazard map or a field-mapped, potentially unstable slope to determine runout probability (Section 5).

Figure 1. Runout at the four validation sites: (a) Cascade Mountains, WA: a large debris avalanche over steep, broadly convergent terrain (photo credit: Stephen Slaughter). (b) Black Hills, WA: a large debris flow over a broadly convergent, gently sloped valley (photo credit: Stephen Slaughter). (c) Rocky Mountains, CO: a moderate sized debris avalanche over steep, unconfined to divergent hillslope. (d) Olympic Mountains, WA: small debris flows in steep, highly convergent channels.

3.2 Description of the Mass Wasting Runout model

3.2.1 Overview of the cellular-automaton modeling approach

MWR is coded as a discrete cellular automaton (CA) model. CA models can track the movement of sediment over rugged boundary conditions (i.e., natural topography) and depending on implementation, can be numerically efficient (Tucker and Hancock, 2010; Tucker et al., 2016). In the context of predicting the runout from a landslide, the CA approach in MWR includes only the key processes that control the runout at the DEM grid scale (See Section 4) and therefore may be faster than an equivalent detailed, continuum-based mechanistic model that attempts to include grain-scale level processes. CA models are not without limitations though, in particular, the use of simple topographic routing algorithms to approximate flow results in a grid-size dependence in flow depth and velocity (Tucker and Hancock, 2010).

CA models iteratively apply a set of equations or rules to individual cells of a grid (Codd, 1968) and are typically employed in geomorphology models in which gravitationally directed erosion and depositional processes modify a digital elevation model (DEM) representation of a landscape (i.e. Chase, 1992; Crave & Davy, 2001; Murray & Paola, 1994; Tucker et al., 2018). Other cellular-automaton based landslide runout models include DebrisFlow Predictor (Guthrie and Befus, 2021), SCIDDICA (D'Ambrosio et al., 2003) and a model by Han et al (2021). In those models and MWR, computations only occur at the location of moving debris rather than the entire model grid, in a manner analogous to the "mobile" cellular automaton implementation of Chase (1992). Chase (1992) modeled precipitation-driven fluvial erosion by randomly placing single packets of precipitation on a DEM, which then moved downslope, eroding and transporting sediment as a function of slope. The individual packets of precipitation were referred to as precipitons. After one iteration, both the precipiton and the landscape evolved and the location of

the precipiton for the next iteration was determined by the slope of the landscape and the state of the precipiton. In our implementation, we route debris released from a specified landslide source area and we refer to the packets of debris as “debritons”. The debritons represent debris flux. Each debriton occupies a single cell. Multiple debritons can traverse the landscape at once. As they move, the debritons change volume, split to form new, smaller debritons or combine to form larger debritons. In this model, debris flux is defined as the volume of debris per unit area of the terrain transferred downslope in one model iteration [L3/L2/iteration]. Erosion and transport capacity of the debritons are a function of slope as well as properties of the debritons.

The present implementation of the MWR algorithm uses the Landlab raster model grid. The raster grid consists of a lattice of rectangular cells. Properties of each cell, such as topographic elevation, slope, and other spatially varying properties such as regolith depth or grain size, are recorded at nodes in the center of each cell (see Figure 5 of Hobley et al. 2017) and evolve as the model runs. In the next section, we describe the methodology used to determine flow direction and how each debriton interacts with the terrain and regolith as it moves down slope.

3.2.2 Mobilization of the initial landslide (Algorithm 1):

Algorithm 1 operates on all nodes within a user-specified initial landslide polygon. For each node n within the landslide polygon (Figure 2, iteration 0), Algorithm 1 sets the topographic elevation to the elevation of the landslide slip-surface (surface along which the landslide body initially fails/moves), designates the material above the slip surface as mobile debris (a debriton) and determines which nodes will receive debris (receiver nodes: R_n), the flux of debris sent to each of the node n’s i-th receiver nodes (qs_{n,R_n}) and attributes of the sent debris (ξ_n) based on topographic slope and attributes at node n (Figure 3). Movement is initiated at the lowest debriton in the landslide body, where topographic slope is computed from the upper surface of the debriton. At

all other debritons, and for all other stages of modeled runout (Algorithm 2), slope is determined from the underlying terrain (slip surface during Algorithm 1).

The depth of the slip surface is specified by the user, which can be inferred from field evidence for a specific landslide (e.g. shallow failure along the fresh-bedrock to regolith contact) or determined from an external landslide model (See Section 5). All nodes that send material are designated as donor nodes and appended to a list of donor nodes (D). All R_n are appended to receiver node list R. Flux and attributes that will be passed from the donor nodes to the receiver nodes are appended to lists qs and ξ. During Algorithm 1, no debris actually moves; it is simply staged for movement by preparing the lists D, R, qs and ξ. Actual movement occurs in Algorithm 2. Lists D, R, qs and ξ evolve with each iteration of Algorithm 2.

Algorithm 1 Determinin the initial D, R, qs and ξ lists

sort initial landslide nodes from losest elevation to highest
for each node n in the initial landslide nodes, beginning from the lowest node
    determine R_n, qs_{n,R_n} and ξ_n using (1)
    subtract landslide thickness from topographic elevation
    update the topographic slope using new topographic elevation
    append n, R_n, qs_{n,R_n} and ξ_n to lists D, R, qs and ξ
end for

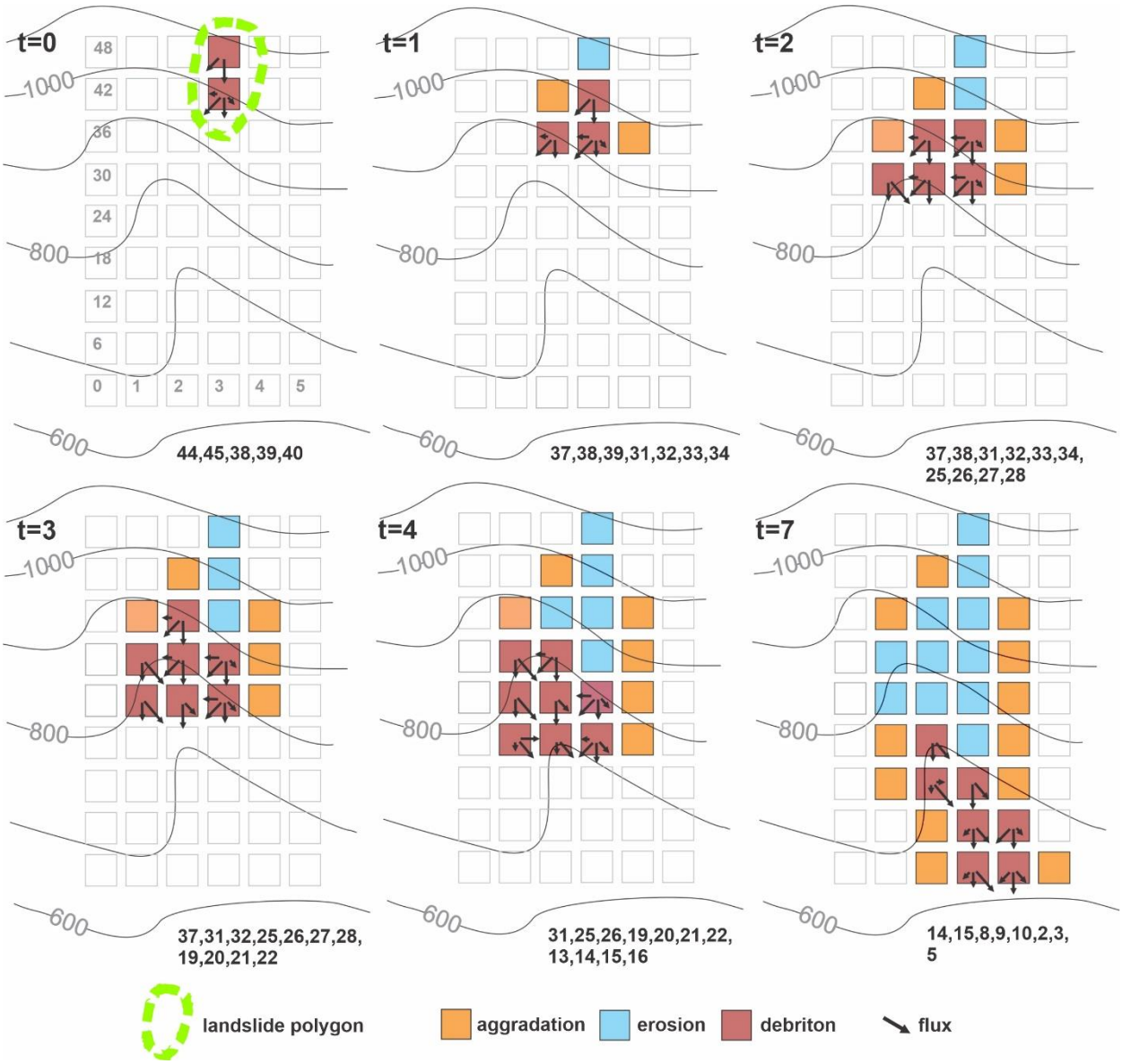

Figure 2. Illustration of initial mass wasting release (t = 0) and runout. Notice how the receiver node list R, shown at the base of each iteration, changes with each iteration. The flow elongates and widens as the number of nodes in R increase and stops when the number of nodes in R is zero. If the incoming flux (sum of all incoming debritons) to a node is less than qs_c, the material stops, causing aggradation.

3.2.3 Rules for debris flow erosion and deposition (Algorithm 2)

Algorithm 2 is essentially the runout model. An illustration of several iterations of Algorithm 2 is shown in Figure 2 and in the Validation section (Section 4, Figure 7). Algorithm 2 determines how material traverses and interacts with the terrain, beginning with the initial D, R, qs and ξ lists passed from Algorithm 1. While Algorithm 2 processes each debriton, it appends n (node index of the debriton), R_n, $qs_{n,R_{ni}}$ and ξ_n to temporary lists D', R', qs' and ξ'. At the end of Algorithm 2, D', R', qs' and ξ' become D, R, qs and ξ. Algorithm 2 repeatedly runs while the length of R is greater than zero (while there are receiver nodes for the next model iteration).

Debriton movement is controlled by mass continuity and a simple threshold flux constraint. At a node n, outgoing sediment flux, qs_n^O, is determined as (Figure 3c):

$$qs_n^O = \begin{cases} qs_n^I - A_n + E_n, & qs_n^I \geq qs_c \\ 0 \quad , & qs_n^I < qs_c \end{cases} \tag{1}$$

where, qs_n^I is total incoming sediment flux to the node (described in Equation 2); A_n and E_n are local deposition (aggradation) and erosion [L/iteration], both of which are calculated prior to (1); qs_c is the threshold flux for deposition. When $qs_n^I < qs_c$, qs_n^I deposits and qs_n^O becomes zero. Conceptually, qs_c represents the flow depth below which flow resistance is large enough to cease the forward momentum of the flow, whether in the form of frictional resistance along the base of the flow or debris and vegetation in the path of the flow.

Summing the outgoing flux from each donor node towards node n, gives the total incoming flux, qs_n^I [L/iteration]:

$$qs_n^I = \sum_{j=1}^{Nd_n} qs_{D_{nj},n} \tag{2}$$

where Nd_n is the number of donors to node n, and $qs_{D_{nj},n}$ is the flux from node D_{nj} (the j-th donor to n) to node n. The flux from node n to each of the node n's i-th receiver nodes (qs_{n,Rn_i}) is determined using the Freeman (1991) multiflow direction algorithm:

$$qs_{n,R_{ni}} = qs_n^O \frac{S_i^a}{\sum_{i=1}^{Nr_n} S_i^a} \tag{3}$$

where Nr_n is the number of receiving nodes of node n, and S_i is the underlying topographic slope ($\tan\theta$) to each of node n's i -th receiver and qso_n.

The multiflow direction algorithm is common to many models, and in this implementation it is handled by a pre-existing Landlab flow-routing component. The exponent a controls how material is distributed to downslope nodes as a function of slope. In a braided river cellular-automaton model, Murray and Paola (1994) used an approximation for turbulent shallow water flow to justify a = 0.5 (which is the exponent on the slope factor in channel friction laws). For our application, we found MWR was easier to calibrate to observed mass wasting runout if a = 1.

Local aggradation (A_n) and erosion (E_n) in (1) are determined using a series of field-data-informed rules (Figure 3c). These rules are equivalent to the "transition functions" applied in SCIDDICA (D'Ambrosio et al. 2003) and the Han et al. (2021) models. If the local slope (S_n) is greater than a critical slope (S_c), then no material deposits, whereas if S_n is less than S_c, then a portion or all of qs_n^I deposits:

$$A_n = \begin{cases} 0 & , \quad S_n \geq S_c \Delta x \\ \min\left(A_{p,n|N_a}, qs_n^I\right), & S_n < S_c \Delta x \end{cases} \tag{4}$$

where, $A_{p,n|N_a}$ is the potential deposition depth at node n, which is the first node where the debriton begins to deposit because the deposition criterion is met (described below, see Figure 4). The amount of deposition at node n and each subsequent downslope node is governed by the

topographic position of each node. To determine $A_{p,n|N_a}$, we initially used the nonlinear, nonlocal deposition scheme described by Carretier et al. (2016) and Campforts et al. (2020); however, when implemented with the flow routing method described above, unreasonably tall deposits resulted when qs_n^I was large and $S_n << S_c$. Alternatively, we experimented with limiting the deposition depth to $A_n \leq S_c \Delta x$, but when qs_n^I was large, this constraint resulted in long deposits that paralleled the underlying slope. Therefore, a different constraint was needed. We ultimately developed a new deposition rule that determines deposition at a node by back-computing the deposition amount necessary to form a hypothetical deposit that spreads over multiple nodes, whose geometry is controlled by S_c.

The hypothetical deposit is defined using three assumptions: (1) the volume of the deposit is equal to $qs_n^I \Delta x^2$ and spreads over N_a nodes; (2) the surface slope of the deposit is equal to the critical slope S_c for deposition of the debritons material and; (3) the underlying topographic slope of the N_a nodes is uniform and equal to S_n. Assumption 3 is necessary because the terrain evolves with each model iteration and the future downslope terrain is unknown. From the hypothetical deposit, we can analytically define $A_{p,n|N_a}$ and N_a as a function of qs_n^I, S_c and S_n as follows (see Figure 4):

From assumption 1, qs_n^I can be expressed as the sum the N_a deposits that make up the overall deposit as:

$$qs_n^I = \sum_{i=1}^{N_a} A_{p,i} \quad (5)$$

where $A_{p,i}$ is the i-th deposition amount in the hypothetical deposit. Since we assume the hypothetical deposit slope and underlying topographic slope are uniform, the deposition amount at any of the N_a nodes can be determined from the last deposition amount ($A_{p,1}$) as:

$$A_{p,i} = A_{p,1} + (i-1)\Delta x(\,S_c - \mathrm{S}_n) \tag{6}$$

From (6) we can re-write (5) as a function of $A_{p,1}$ (write each of N_a deposits in (6) in terms of $A_{p,1}$) and rearrange to define $A_{p,1}$ as a function of qsi_n:

$$A_{p,1} = \frac{1}{N_a} qs_n^I - \frac{N_a - 1}{2}\Delta x(\,S_c - \mathrm{S}_n) \tag{7}$$

Substituting (7) into (6) and solving for $i = N_a$ (i.e., A_{p,N_a}), we get an expression for $A_{p,n|N_a}$:

$$A_{p,n|N_a} = \frac{1}{N_a} qs_n^I + \frac{N_a - 1}{2}\Delta x(S_c - \mathrm{S}_n\,) \tag{8}$$

Equation (7) can be rearranged into a quadratic equation and solved for N_a as:

$$N_a = \frac{-A_{p,1} + \frac{1}{2}\Delta x(S_c - S_n\,) \pm \sqrt{\left(A_{p,1} - \frac{1}{2}\Delta x(S_c - S_n\,)\right)^2 + 2\Delta x(S_c - S_n\,)qs_n^I}}{\Delta x(S_c - S_n\,)} \tag{9}$$

We use (8) to solve for $A_{p,n|N_a}$ and (9) to solve for N_a assuming $A_{p,1} = 1/2\,\Delta x S_c$ and rounding the positive solution to the nearest integer. When implemented using a single debriton released on a 2-dimentional hillslope like illustrated in Figure 4 (i.e., a single row of nodes), beginning at the slope break, the debriton deposits over N_a nodes at a uniform slope equal to S_c. When implemented on an actual terrain, the interaction between multiple debritons in multiple directions creates a complex deposit whose slope changes with S_c, as will be shown in the results.

The third variable in Equation 1 is debriton erosion rate, E_n. We constrain $E_n$to the lesser of a potential erosion rate, he, and local regolith depth, h_r:

$$E_n = \min\,(h_r, h_e) \tag{10}$$

where h_e [L/iteration] is computed as a function of the basal shear stress of the granular flow (τ, described below in Equations 12 and 13) and the critical shear stress (τ_c) of the regolith at the node:

$$h_e = k(\tau - \tau_c)^m \quad (11)$$

Shen et al. (2020) used τ_c = 5 Pa to represent a boulder-colluvium substrate. Stock and Deitrich (2006) used normal stresses rather than shear and exclude τ_c ($\tau_c = 0$). The exponent m controls the non-linearity of h_e. Many authors (Chen & Zhang 2015; Frank et al., 2015; Shen et al., 2020) use a value of 1 for m but field measurements by Schurch et al. (2011) (see their Figure 3) suggest that m may be less than 1 if τ is assumed to vary linearly with flow depth, particularly at high flow depths ($h > 3$m). We experimented with various values of m, ranging from 1 to 0.2 and found that as long as k is determined as a function of m (Section 4.3, Equation 36), the impact of the value of m on model behavior is relatively minor.

The coefficient k in (11) is an erodibility parameter. Stock and Dietrich (2006) showed that k encapsulates rock properties. If h_e is used to represent erosion over geomorphic time scales, with repeated debris flow occurrences in a single model iteration, k becomes associated with debris flow length and frequency (Perron, 2017). In our application we are modeling the erosion associated with a single runout event, as represented by the debritons. The coefficient k therefore needs to scale h_e so that it is on the order of the total erosion caused by a single debriton.

To define τ, MWR includes two options: (1) a quasi-static basal shear stress approximation or (2) a grain-size-based shear stress approximation. The quasi-static basal shear stress approximation (e.g., Takahashi, 2014) is defined as:

$$\tau = \rho g h \sin\theta \quad (12)$$

where ρ is the density of mass wasting material (grain and water mixture), g is gravity and h is flow depth, which in MWR is equivalent to qs_n^I. The grain-size-based shear stress approximation is defined as follows: Bagnold (1954) defined τ as a function of grain-collision dependent normal stress (σ).

$$\tau = \sigma \tan \varphi \tag{13}$$

where φ is the collision angle between grains, measured from the vertical axis (See Bagnold 1954), with a value of $\tan \varphi$ typically equal to 0.32. Following Bagnold (1954) and Iverson (1997), Stock and Dietrich (2006) defined σ as:

$$\sigma = \cos \theta v_s \rho_s {D_s}^2 \left(\frac{du}{dz}\right)^2 \tag{14}$$

where θ is the topographic slope measured in degrees, v_s is the volumetric solids concentration, ρ_s is density of the solids, u is flow velocity, z is depth below the flow surface, du/dz is strain and D_s is the representative grain size. Stock and Dietrich (2006) suggested that D_s corresponds to a small percentile of the coarsest fraction of the runout material (D_{88} to D_{96}) and they approximated du/dz as:

$$\frac{du}{dz} = \frac{u}{h} \tag{15}$$

Solely for the purpose of computing du/dz, we approximate velocity at a node following Julien and Paris (2010) as:

$$u = 5.75 u^* \log \left(\frac{h}{D_s}\right) \tag{16}$$

where $u^*$is shear velocity $\left(\sqrt{gh \tan \theta}\right)$. Since bedrock type determines the size and type of grains transported by a landslide (e.g., Roda-Boluda et al. 2018), defining an erosion rule as a function of grain size permits landslide-driven erosion rates to vary by lithologic region, as represented by spatial variation in D_s. Furthermore, D_s can be quickly approximated in the field by measuring the coarser grain sizes in existing runout-deposits, road-cuts and tree-throw pits.

Once A_n and E_n are determined at a node, we assume uniform material density and determine the change in elevation at the node ($\Delta\eta_n$) as:

$$\Delta\eta_n = A_n - E_n \tag{17}$$

Finally, for each regolith attribute being tracked by the model (e.g., grain size), the attribute value at the node (ξ_n) and the attribute value sent from the node to its receiver nodes (ξ_{Rn}) is determined using a volumetric-weighted average as:

$$\xi_n = \frac{\xi_n(h_r - E_n) + \xi_{Dn}(A_n)}{D + h_r - E_n} \tag{18}$$

$$\xi_{Rn} = \frac{\xi_E E_n + \xi_{Dn}(qsi_n - A_n)}{qso_n} \tag{19}$$

Where ξ_{Dn} is the weighted average attribute value delivered to the node from its donor nodes:

$$\xi_{Dn} = \frac{\boldsymbol{\xi_D \cdot qs_{D_n}}}{qsi_n} \tag{20}$$

and $\boldsymbol{qs_{D_n}}$ is a vector containing all $qs_{D_{nj},n}$ sent to the node and $\boldsymbol{\xi_i}$ is a vector containing the incoming attribute values for each $qs_{D_{nj},n}$.

For each model iteration t, as Algorithm 2 processes each unique node n in the receiver node list (R), it appends $\Delta\eta_n$ to a list $\Delta\eta$ and n to D', R_n to R', $qs_{n,Rn}$ to qs' and ξ_{Rn} to ξ'. After Algorithm 2 has cycled through all receiver nodes, D', R', qs' and ξ' become D, R, qs and ξ and maps of regolith thickness ($h_{r\,t}$), topographic elevation (η_t) and the elevation of the runout flow surface (F_t) are updated for the entire model domain using vector operations as:

$$\eta_t = \eta_{t-1} + \Delta\eta \tag{21}$$

$$h_{r\,t} = h_{r\,t-1} + \Delta\eta \tag{22}$$

$$F_t = \eta_t + qs \tag{23}$$

After η_t, $h_{r\,t}$ and F_t are updated, the multi-direction slope of the DEM, which is used for routing the debritons the next model iteration, is recomputed from η_t.

Since MWR uses the underlying topographic slope to route the debritons, debritons are obstructed if they encounter a topographic pit or flat topography. To allow a debriton to pass an obstruction, we rely on a simple work-around: upon encountering the obstruction, the debriton is

directed to itself ($R_n = n$) and some portion of the debris is deposited based on (4). At the end of the model iteration, the node elevation and slope are updated, and then during the next iteration, if the remaining mobile debris is no longer obstructed, it moves to its downslope node(s). If the node is still obstructed, it is again sent to itself until either all material has deposited or the elevation of the node exceeds that of its neighbor nodes, allowing the debriton to move downslope.

Finally, as noted at the beginning of this section, flow routing algorithms used in CA models, such as (3), can cause flow-depth dependence on the grid-size used to represent the terrain. Since aggradation and erosion rates determined by (8) and (11) are sensitive to flow depth, we thus constrain the flow depth used to determine A and E to no more than the maximum observed flow depth as:

$$h' = \min(h_{max}, h) \tag{24}$$

where h' is the flow depth used in (8) and (11) and h_{max} is the maximum observed flow depth. This constraint does not impact mass continuity in (1) and in practice, h_{max} can be inferred from field indicators (See Section 4).

Algorithm 2 Runout algorithm

while length(R) >0
  for node n in R
    determine flux and attributes of incoming material from lists qs and ξ
    compute erosion, deposition and outflow as a function of underlying topographic slope
    determine the new receiver nodes (R_n), receiving flux ($qs_{n,Rn}$), receiving properites (ξ_{Rn})
    append n to D', R_n to R', $qs_{n,Rn}$to qs and ξ_{Rn} to ξ'.
  end for
  update the topographic elevation and slope for erosion and deposition
  lists D', R', qs' and ξ' become D, R, qs and ξ
end while

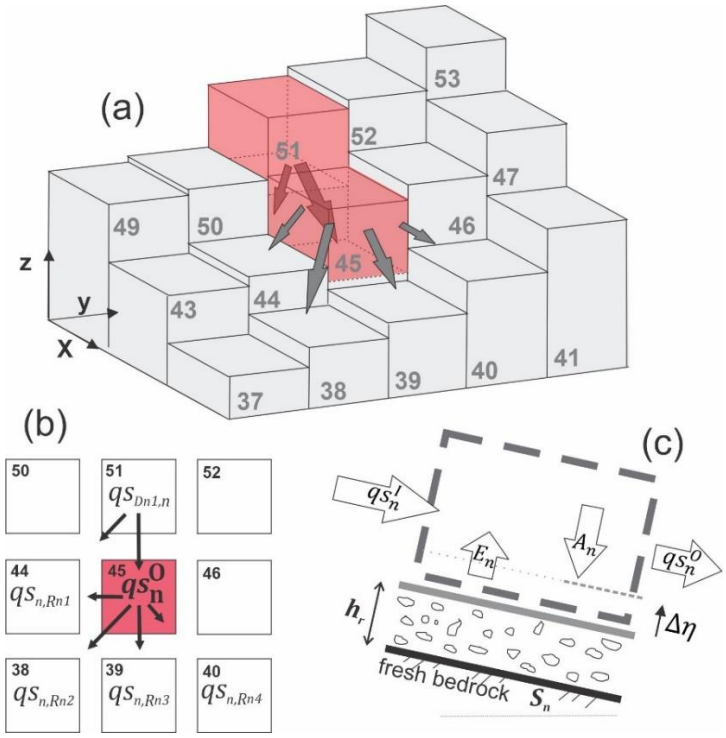

Figure 3. (a) Three-dimensional illustration of iteration 0 in Figure 1, showing initial landslide nodes (initial debritons) and flux towards downslope nodes. Except for the first released debriton (lowest elevation node in the landslide), all debritons are directed downslope based on the underlying topographic slope; (b) Distribution of qs_n^O to downslope nodes and incoming material qs_n^I (here equal to $qs_{D_{n1},n}$); (c) Aggradation (A_n) and erosion (E_n) are determined from field informed rules and qs_n^O is determined from mass continuity.

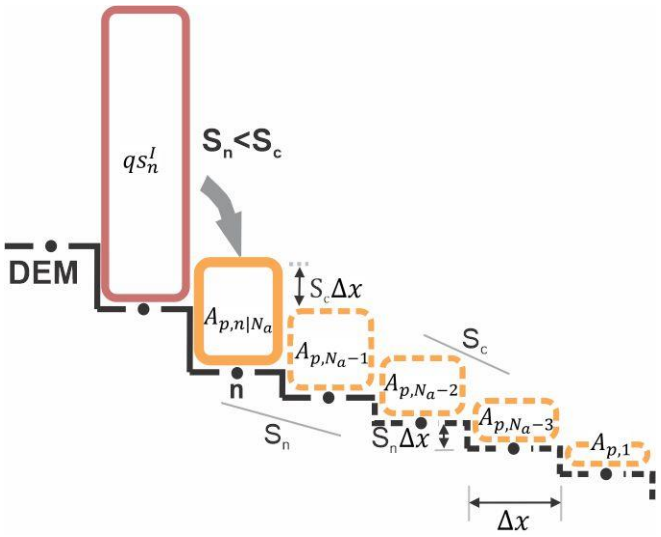

Figure 4. Conceptual model used to define deposition rule, here illustrated using $N_a = 5$. The rule relies on three assumptions about the depositional process: (1) deposition of the debriton, which has a flux of qs_n^I at node n, spreads over N_a nodes to form a hypothetical deposit as illustrated by yellow boxes above; (2) the underlying topographic slope of the N_a nodes is uniform and equal to S_n and; (3) the surface slope of the hypothetical deposit will be equal to S_c. Solid yellow box indicates deposition at node n. Dashed yellow boxes and lines indicate hypothetical deposition and underlying topography. Dots along DEM surface are nodes.

3.3 Calibration and MWR probability

3.3.1 Calibration utility

We developed a calibration program, written in Python as a utility for the package Landlab, to calibrate MWR. This utility uses an adaptive Markov Chain Monte Carlo (MCMC) sampling algorithm to automatically sample model parameter space, evaluate model performance relative to observed runout, and determine the most likely set of parameter values that calibrate the model.

A minimalistic calibration of MWR requires a shape file, representing the boundaries of a landslide and its runout. For a more elaborate calibration, the utility uses a DEM of difference (DoD) obtained by subtracting the DEMs obtained before and after the landslide and subsequent runout. In case of the former, the parameter likelihood function is limited to an evaluation of

planimetric extent (Ω_T, described below), and the model tends to have high parameter equifinality (e.g., Beven 2006); multiple parameter sets result in an equally calibrated model as evaluated by Ω_T. The calibration utility is presently coded to modify S_c and qs_c, k, landslide thickness, and mean soil depth; however, any parameter of the MWR can potentially be calibrated using this approach.

As will be detailed in Section 4, most parameters in MWR can be inferred from field observations and a DoD map, but parameters S_c and qs_c are best determined through calibration. A range of S_c values can be inferred from deposition slopes measured at channel-filling deposits of the observed runout but given the intermittency and variability of most debris flow deposits, identifying a single representative S_c value in the field can be difficult. Conceptually, qs_c represents a minimum flow thickness below which resistance due to friction or debris, such as vegetation or boulders in the path for the flow, causes the flow to cease. While it's possible that qs_c may be inferred from the minimum observed flow thicknesses along the margins and termini of the runout, this evidence, especially at the termini of debris flows, can be erased rapidly as streams erode deposits and may also vary along the runout path. Through model calibration to observed runout, a single value of S_c and qs_c that best represents the heterogeneity of observed depositional slope and flow thickness can be determined.

The adaptive MCMC sampling algorithm is described by Coz et al. (2014), who used it to calibrate stage-discharge relationships to observed river flow in France. The algorithm works as follows: First, the user identifies an initial (prior) PDF for each parameter to be calibrated, currently only limited to normal and uniform distributions, and provides initial estimate of the distribution parameters. The algorithm then randomly selects a set of parameter values (θ) from the prior PDFs, evaluates the prior likelihood of the parameter set ($p(\theta)$), and runs MWR using θ.

Once the model has completed the run, the algorithm evaluates the posterior likelihood of the parameter set ($L(\theta)$, described below, based on how well modeled runout matches the planimetric extent, topographic change and cumulative flow along the observed runout path.

After the first $L(\theta)$ has been determined, the algorithm selects a new set of parameters (θ_{t+1}) by jumping some distance from each parameter in θ space. The jump distance is determined by randomly sampling a normal distribution of possible jump distances, whose standard deviation is initially specified by the user, but changes as the algorithm iteratively runs and evaluates model performance. The next iteration, $t+1$, MWR is run using θ_{t+1} and once the model has completed the run, L_{t+1} is determined. If $L_{t+1}$is larger than L, θ is updated to the value of θ_{t+1} and L is updated to the value of L_{t+1}. If $L_{t+1} < L$, then θ and L have a probability of remaining unchanged equal to 1- L_{t+1}/L where the fraction L_{i+1}/L is referred to as the acceptance ratio. This Markov process is repeated a user-specified N times. The algorithm is adaptive because the jump distance changes depending on the average acceptance ratio, which is evaluated every user-specified number of iterations ($Ncycles$). Two additional parameters control the jump size: a_{min} and a_{max}. If the average acceptance ratio is below a_{min}, the variance of the jump-distance distribution is decreased by 10 percent and if the average acceptance ratio is greater than a_{max}, the variance is increased by 10 percent. We initially tested the adaptive MCMC algorithm using values for $Ncycles$, a_{max}, a_{min} and the initial jump size ($jump\_size$) from Coz et al. (2014) but through a series of tests, found that values listed in Algorithm 3 worked best for our application (see Algorithm 3).

For each iteration of the calibration utility, $L(\theta)$ is determined as the product of $p(\theta)$ and three other metrics as:

$$L(\theta) = p(\theta) * \Omega_T * \frac{1}{{\Delta\eta_E}^2} * \frac{1}{{Q_{s_E}}^2} \quad (25)$$

where Ω_T is the Lee-Salle index (Heiser et al. 2017) and indicates modeled planimetric fit and ${\Delta\eta_E}^2$ and ${Q_{s_E}}^2$ are new dimensionless indices, proposed for this study, that represent modelled topographic and flow error. Larger values of Ω_T indicate a the modeled planimetric more closely fits observed. Smaller values of ${\Delta\eta_E}^2$ and ${Q_{s_E}}^2$ indicate better fit. Note that the reciprocal values of ${\Delta\eta_E}^2$ and ${Q_{s_E}}^2$ are used in (25) so that improved fit causes $L(\theta)$ to increase. As mentioned above, if a DoD is not included, then (25) is determined as the product of $p(\theta)$ and Ω_T only. To ensure that all values of $L(\theta)$ are greater than or equal to zero, here we add a value of 1 to Ω_T so that it scales from 0 to 2 as:

$$\Omega_T = \frac{\alpha-\beta-\gamma}{\alpha+\beta+\gamma} + 1 \quad (26)$$

where α, β and γ are the areas of matching, overestimated and underestimated runout extent, respectively. The root values of ${\Delta\eta_E}^2$ and ${Q_{s_E}}^2$ ($\Delta\eta_E$ and Q_{s_E}) reflect the volumetric error of the modeled topographic change relative to the observed total mobilized volume (observed landslide + erosion) and the mean modeled cumulative flow error along the runout path relative to the observed mean cumulative flow. The index ${\Delta\eta_E}^2$ is determined by dividing the squared, cumulative, volumetric square error (SE) of modeled topographic change by the squared observed-total-mobilized volume (V):

$${\Delta\eta_E}^2 = \frac{\alpha_E+\beta_E+\gamma_E}{V^2}. \quad (27)$$

Here values α_E, β_E and γ_E are the SE of $\Delta\eta_M$ in areas α, β and γ respectively, individually calculated as the sum of the squared differences of debris flow between the observed and modeled elevation change at each cell in the model:

$$SE = \sum_{i=0}^{p}[(\Delta\eta_{Oi} - \Delta\eta_{Mi})dx^2]^2 \quad (28)$$

where, p is the number of nodes within each of the α, β, and γ areas, and $(\Delta\eta_{Mi})$ and $\Delta\eta_{Oi}$ are the observed topographic change [L] at the i-th node (cell) in that erosion area.

Cumulative flow volume (Q_s) is determined along the channel profile, as the integrated, upslope-volumetric topographic change relative to a point along the profile in a manner similar to Fannin and Wise (2001), who computed cumulative flow at individual reaches along a channel profile (see their Figure 4). We compute cumulative flow at each node (grid cell) j along the channel profile as:

$$Q_s = -dx^2 \sum_{i=1}^{q_j} \Delta\eta_{i,j} \quad (29)$$

where $\Delta\eta_{ij}$ is the topographic change [L] at the i-th node located upstream of node j, q_j is the number of all nodes in the DEM located upstream of node j, and node j is a node located along the channel profile. The negative sign converts upslope erosion and deposition to positive and negative change in cumulative flow at the downslope point in the runout path. For each node j along the profile, Q_s is computed from both the observed and modeled runout ($Q_{sO\ j}$ and $Q_{sM\ j}$) and ${Q_s}_E{}^2$ is the mean square error of the modeled-cumulative flow at all nodes along the profile, normalized by the observed-mean-squared-cumulative flow along the profile:

$${Q_s}_E{}^2 = \frac{\frac{1}{r}\sum_{j=1}^{r}(Q_{sO\ j} - Q_{sM\ j})^2}{\frac{1}{r}\sum_{i=0}^{r} {Q_{sO\ j}}^2} \quad (30)$$

where r is the number of upstream nodes for each j node used in the calculation of this statistic and ${Q_s}_E{}^2$ is simply the mean square error of modeled Q_s normalized by mean observed Q_s. Values of ${Q_s}_E{}^2$ and ${\Delta\eta}_E{}^2$ (or ${Q_s}_E$ and $\Delta\eta_E$) close to zero indicate a good fit and large values (>>0) indicate

a poor fit. As noted above, in the final objective function, $L(\theta)$, we use the reciprocals of ${\Delta\eta_E}^2$and ${Q_{s_E}}^2$, so that smaller values of ${Q_{s_E}}^2$ and ${\Delta\eta_E}^2$result in larger $L(\theta)$ values.

An advantage of defining $L(\theta)$ as the product rather than summation of these terms is that a change in any one metric has an equal impact on the relative value of $L(\theta)$ (e.g, a 20% reduction of any of the metrics causes a 20% reduction in $L(\theta)$). Furthermore, because we use the product of the reciprocals of the squared topographic and flow error (e.g., error-based weighting, Foglia et al., 2009), this formulation of $L(\theta)$ is very sensitive to changes in calibration performance and results in values that vary over several orders of magnitude. Since the MCMC algorithm searches parameter space based on changes in the relative value of $L(\theta)$, at the validation sites included in this study, it worked well for guiding the algorithm.

Finally, one of the major advantages of using an MCMC approach for model calibration is that after an adequate number of model runs have been performed, the resulting histogram of parameter indicates the posterior PDF of the parameter values (Gelman et al., 2021; Renard et al., 2006), which can then be used in Monte Carlo modeling runs. Typically the initial series of jumps are considered "warm up" and depending on context, may be discarded from the histogram (Gelman et al., 2021). Also, the minimum number of iterations needed for convergence on best-fit parameters (maximum $L(\theta)$) can be determined as a function of the ratio of between sequence variance (e.g., variance between two sequences of $Ncycles$ parameters) and variance within each sequence.

In the Model Validation section, we run the calibration utility using a single chain of 2000 repetitions to estimate the posterior PDF for each parameter. At most sites, the model converges relatively quickly on a solution and we therefore don't consider burn-in, or evaluate convergence. Future implementations of the calibration utility may include multiple chains, burn-in and a check

for convergence. We use the parameter set with the highest $L(\theta)$ value (best-fit parameter set) to compare model performance relative to observed runout. Then, in Section 5, we iteratively run MWR with parameter values sampled from the posterior PDFs to illustrate the calculation of runout probability as a function of parameter uncertainty.

Algorithm 3 Adaptive Markov Chain Monte Carlo calibration algorithm

N: user defined number of times the while loop runs
θ_{t-1}: initial set of model parameter from which the first jump in parameter space is made
$jump\_size$: ratio of standard deviation of the jump size expressed as the ratio of the range between the
minimum and maximum parameter values
$L(\theta_{t-1}) = 1$
t = 0
while t < N
from θ_{t-1}, jump random distance to new parameter set, (θ_t)
compute liklihood of parameter values using prior distribution $(p(\theta_t))$
run MWR with with θ_t
evaluate model calibration with ${\Delta\eta_E}^2$, ${Q_{s_E}}^2$ and Ω_T
compute posterior likihood of parameter values ($L(\theta_t)$) using (22)
acceptance_ratio = min(1, $L(\theta_t)$/ $L(\theta_{t-1})$)
append acceptance_ratio to acceptance_ratio_list
rv = random value between 0 and 1

if rv < acceptance_ratio
$\theta_{t-1} = \theta_t$

if c = $Ncycles$
if mean acceptance_ratio > a_{max}
increase variance of jump size distribution by 10%
if mean accepatnce_ratio < a_{min}
decrease variance of jump size distribution by 10%

t = t+1

Where $Ncycles$ = 100, a_{max} = 0.7, $a_{min} = 0.3$, $jump\_size$ = 0.09

3.3.2 MWR probability

MWR includes an additional utility called MWR Probability that relies on a Monte Carlo implementation of MWR using the parameter PDF's produced by the calibration utility. This utility is developed to produce debris flow hazard maps, and indicates the probability of runout below a potential landslide. MWR Probability repeatedly runs MWR for a user specified N times,

each repetition with a different S_c, qs_c pair, randomly sampled from the parameter PDFs. After the N model runs, it determines probability of runout impact at each model node as:

$$P(\Delta\eta) = \frac{n(|\Delta\eta|>0)}{N} \quad (31)$$

where $n(|\Delta\eta| > 0)$ is the number of times topographic elevation at the node changed as a result of erosion or deposition from the N model runs.

MWR Probability includes three options for specifying the initial landslide: (1) a single, fixed-size landslide, represented as a polygon with a fixed regolith depth, (2) a single, variable-size landslide within a user defined potentially unstable slope with fixed regolith depth (3) landslide(s)s that change location, area and depth each iteration. If option 2 is used, for each model repetition, MWR Probability randomly selects MWR parameter values from the parameter PDFs and a landslide area within the potentially unstable slope polygon. The randomly selected landslide area is at least as large as a user defined minimum size but no larger than the potentially unstable slope and can form anywhere within the potentially unstable slope. If option (3) is used, for each model repetition, MWR reads externally determined regolith depth and landslide locations from a total of N regolith depth and landslide maps.

3.4 Model Validation:

3.4.1 Overview

In this section, we demonstrate calibrated model performance relative to observed runout at field sites located in the western USA. Simply calibrating a model to match field data does not constitute a satisfactory test of model predictive ability (Iverson, 2003). Strategic testing, which involves calibrating the model to one site or period of time and then running the calibrated model at a separate site or period of time (Murray 2013), is a better indicator. Two of our validation sites,

the Cascade Mountain and Olympic Mountain sites, include two separate landslides and subsequent runout and we explore model predictive ability at these sites in Section 5.

Calibrated model performance is demonstrated at the following field sites: (1) a large debris avalanche and a moderately-sized debris flow in the Cascade Mountains (Washington state [WA], USA) that inundated and flowed within a first-order channel until perpendicularly intersecting a narrow river valley (Figure 1a); (2) channelized debris flows sourced from a small failure along the toe of a deep-seated landslide and a moderately-sized debris avalanche sourced from a large road fill in the Black Hills (WA) that flowed several kilometers along a flume-like channel before stopping (Figure 1b); (3) a single, moderately-sized debris avalanche in the Rocky Mountains (Colorado state) that flowed topographically unobstructed, onto a broadly convergent to divergent hillslope (Figure 1c); and (4) a 30-year chronology of small-landslides and subsequent debris flows in the Olympic Mountains (WA) that coalesced into a single runout deposit in a dendritic, channelized watershed (Figure 1d).

These validation sites encapsulate the three runout morphologies categorized in terms of runout-energy-dissipation by Nicoletti and Sorriso-Valvo (1991): channelized runout (low-energy-dissipation; Figure 1d), unobstructed spreading of the runout (moderate-energy-dissipation; Figure 1c) and runout that perpendicularly intersects a channel and collides with the opposite channel wall (high-energy-dissipation; Figure 1a). The validation sites also span different climates and land use. The Cascade Mountain, Black Hills and Olympic Mountain sites are located in a wet, maritime climate, in landscapes actively managed as tree plantations. The Rocky Mountain site is located in a semi-arid climate, in a landscape managed as a national park. Despite differences in runout morphology, land use and climate, all landslides initiated during heavy rainfall or rain-plus-snowmelt storm events (WRCC, 2022; NRCS, 2022; Table 1).

In the next sections, we summarize the remote sensing and field methods used to parameterize MWR and the calibration utility. Then, for each site, we describe modeled performance relative to observed runout. At all locations, we use (13) to approximate shear stress as a function of grain size. Model performance is discussed relative to the pre-event topography of the runout path in terms of the overall, planimetric, topographic and total volumetric flow.

Table 3.1 Landslide and runout characteristics

site	Cascade Mountain, 09	Cascade Mountain, 22	Black Hills, south	Black Hills, north	Rocky Mountain	Olympic Mountain
landslide length, ℓ [m]	185	55	80	75	40	45
landslide width [m]	80	50	15	65	35	15
landslide volume [m^3]	110,000	22,000	1,500	18,500	4,600	400 - 2,200
2-day cumulative precipitation + snowmelt [mm]	120+85	140+75	205+50	205+50	193+0	100 - 220 + ?
maximum grain size [m]	0.316	0.316	0.48	0.206	0.984	0.8
slope at positive-net deposition [%]	1 - 15	1 - 15	<1 - 10	<1 - 8	16 - 25	5 - 15
average flow depth in scour zone [m] *	4	2	2	3	3	3
average channel slope in scour zone [m/m]	0.25	0.25	0.15	0.15	0.4	0.3
average channel width in scour zone [m]	45	20	25	35	55	10
length of erosion, [m]	600	340	1210	1345	360	2550
erosion area, $\mathbb{A}$ [m2]	28,400	6,600	22,800	52,400	20,800	28,900
total erosion volume, $\sum E\Delta x^2$ [m^3]**	44,547	5,125	12,332	26,815	34,275	33,725
average erosion per unit length of landslide, $\bar{E}/\ell$, [m/m]	0.0085	0.014	0.0068	0.0068	0.041	0.026
k	0.020	0.034	0.017	0.020	0.076	0.051
growth factor, [m^3/m]	74.2	15.1	10.2	19.9	95.2	13.2
average observed $\lvert\Delta\eta\rvert$	2.4	2.2	0.53	0.63	0.89	1.4
total erosion volume / total mobilized volume***	0.29	0.19	0.89	0.59	0.88	0.97

* rough approximation based on landslide volume, channel width and height of scour marks in the scour zone of the channel

** excludes landslide volume

***total moblized volume = total erosion volume+ landslide volume

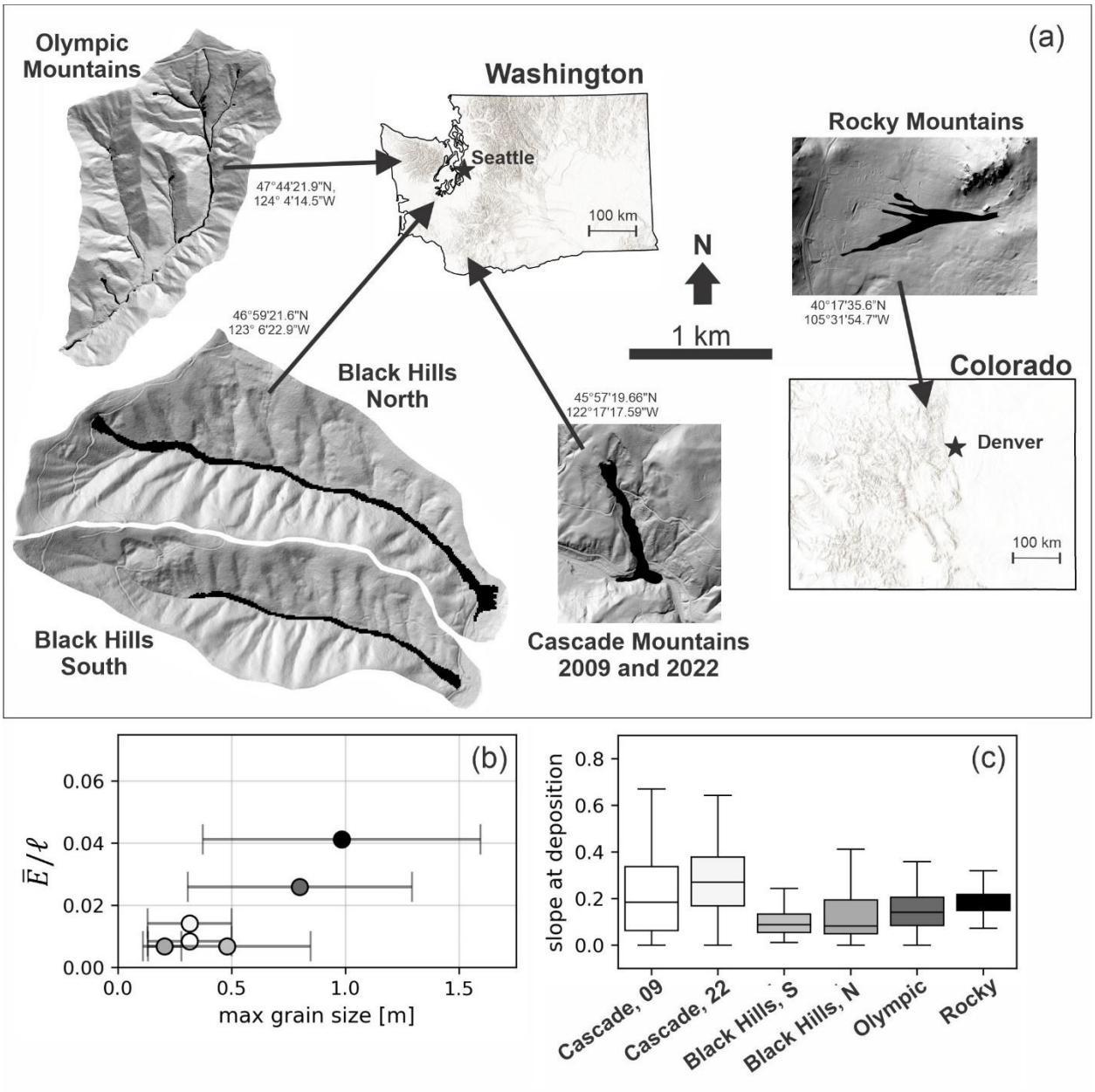

Figure 5. (a) Landslide locations in Washington and Colorado states. Coordinates next to each site are WGS84. Shaded DEMs of each site are sized to the same scale. (b) Observed average erosion rate per unit landslide length ($\bar{E}/\ell$) relative to the observed average-maximum grainsize. Error bars indicate standard deviation. (c) Underlying topographic slope of observed deposition locations.

3.4.2 Model setup, data acquisition and field reconnaissance

For each validation site, we used a combination of field and remote sensing observations to set up and parameterize MWR. Each model was set up on a 10-meter grid, with the following gridded inputs: (1) the pre-event DEM; (2) a map of regolith depth and (3) landslide boundary polygon as a shape file. If the landslide slip plane depth differed from the regolith depth, a list of all nodes in the landslide(s) and their respective slip plane depths were used to adjust the regolith depth to the slip plane depth. For implementing the calibration utility, we also prepared a map of elevation change, $\Delta\eta$, obtained from DEM differencing (a DoD) or a field survey.

In the field, we walked the runout path from the head of the landslide to the terminus of the runout deposit noting the location and thickness of deposition and erosion, deposit composition, the slope of channel filling deposits (reaches where deposition exceeds erosion and results in positive net deposition; Guthrie et al., 2008) and maximum grain size in the deposit. Geometric measurements in the runout path were measured using a combination of a clinometer, steel tape and range finder. Location was measured using a handheld GPS with maximum accuracy = +/- 1.8 meters. Maximum grain size (which at all sites was cobble to boulder sized) was visually approximated in the field or from photos that included a scale object, to the nearest 5 cm at eight to sixteen points along the runout path.

Remote observations included air photos (DNR, 2022; Google Earth, 2022) and lidar DEMs (DNR, 2022; Open Topography, 2022) of the landslide and runout path. The lidar DEMs were used to define the model domain. The air photos were used in combination with field observations and the lidar DEM to interpret the lateral extent of the observed runout. Where pre- and post-mass wasting event lidar DEMs were available, we subtracted the pre-event DEM from the post-event DEM to determine maps of observed $\Delta\eta$ (DoD). We found that not all DEMs or regions of the

DEM were of equal quality and therefore, where possible, DoD maps derived from lidar differencing were checked in the field. At sites with only one post-event lidar DEM (Rocky Mountain and Olympic Mountain), we created the map of $\Delta\eta$ entirely from field observations of deposition and erosion.

In the field, we estimated deposition depth from the thickness of fluvially eroded terraces and burial depth of stumps or trees. In channels that were eroded to the fresh bedrock surface, we estimated erosion depth using the regolith thickness visible along the edges of scoured channel walls (Olympic Mountain site). At other locations, we used vegetative indicators such as the remnants of in-situ tree roots (Rocky Mountain site). Key site reconnaissance results that informed model parameterization and model behavior following parameterization at each site are summarized in Table 1. Model setup details specific to each site are included in the Supplementary Information.

3.4.3 Calibration at each validation site

To calibrate MWR to the observed-runout datasets, we estimated all parameter values from field observations of the observed runout except S_c and qs_c. For S_c and qs_c, we defined uniform prior distributions based on the field observations and then used the calibration utility to find the best-fit parameter values (parameter values corresponding to the highest $L(\theta)$). Minimum and maximum values of S_c were initially estimated from the range of observed slope of areas of positive net deposition (Table 1). Minimum and maximum values of qs_c were set as 0.01 to 1.75, which roughly represents the range of minimum observed thickness of debris flow termini in the field at all of the sites. The range of S_c and qs_c were later adjusted if we found the calibration utility tended to sample parameter values near the edge of the initial parameter range. We approximated k from the observed runout path as follows:

The average erosion depth caused by the observed runout ($\bar{E}$) is equal to the total erosion volume ($\sum E\Delta x^2$) divided by the erosion area ($\mathbb{A}$):

$$\bar{E} = \frac{\sum E\Delta x^2}{\mathbb{A}} \tag{32}$$

where $\sum E\Delta x^2$ and $\mathbb{A}$ exclude the initial landslide volume and area. Since $\mathbb{A}$ is erosion area, it excludes areas of deposition ($+\Delta\eta$) and areas with no change in elevation ($\Delta\eta = 0$). In terms of the debriton conceptualization used in MWR, $\bar{E}$ can also be written as a function of the mean number of times a debriton passes over a grid cell ($\bar{n}$) multiplied by an average erosion depth per debriton ($\bar{h}_e$) as:

$$\bar{E} = \bar{n}\bar{h}_e \tag{33}$$

An estimate for $\bar{n}$ can be determined from the average length of the runout material, which we estimate simply as the mapped landslide length (ℓ) divided by the cell width:

$$\bar{n} = \frac{\ell}{\Delta x} \tag{34}$$

Note that if the observed runout formed as a result of multiple landslides (as was the case at the Olympic Mountain site), then ℓ was determined as the sum of the initial landslide lengths. Also, as the debritons move down slopes in excess of S_c, they entrain material, split, and spread, and the runout material tends to lengthen. Using the initial landslide length to represent the runout length thus represents a minimum value for $\bar{n}$ and if needed, (34) can be multiplied by a coefficient to scale ℓ into a more representative runout length. Combining (33) and (34), $\bar{h}_e$ can be defined as the average erosion rate per unit length of landslide ($\bar{E}/\ell$) times the cell width.

$$\bar{h}_e = \frac{\bar{E}\,\Delta x}{\ell} \tag{35}$$

Rewriting (11) using an average shear stress in the erosion-dominated reaches of the runout path ($\bar{\tau}$) and assuming $\tau_c \cong 0$, debris flow erodibility parameter k can be estimated as:

$$k = \frac{\bar{h}_e}{\bar{\tau}^m} \quad (36)$$

To solve for k, we estimated $\bar{\tau}$ from field-approximated debris flow depth and channel slope measurements in the erosion-dominated reaches of the runout path. To estimate flow depth, we used the height of scour marks on the channel wall or tree trunks, above the channel bed (Table 1). We used Equation 13 to define $\bar{\tau}$. For D_s, we used the average maximum grain size observed over the whole runout path. As noted above, when k is determined as a function of m, as in (36), the impact of m on model behavior is relatively small; however, the calibration utility did tend to more quickly converge on a set of best-fit parameter values using a lower value for m and results herein use m = 0.2.

3.4.4 Observations and model calibration results

This section summarizes field observations used for model parameterization and calibration at the four field sites. Detailed descriptions, including a narrative of the landslide failure and runout process, are included in the Supporting Information. Specific field measurements used to parameterize MWR to each landslide are listed in Table 1.

3.4.4.1 Observed runout characteristics

The observed landslides (e.g., the initial landslide body) ranged in volume from 400 to 110,000 m^3. At all sites, erosion and subsequent entrainment added to the total mobilized volume (total erosion volume plus the initial landslide volume, also referred to as the total flow volume or total debris flow volume; e.g. Barnhart et al. 2021; Reid et al; 2016), but the contribution was highly variable. The erosion volume divided by the total mobilized volume was as low as 0.19 at the Cascade Mountain, 2022 landslide to as high as 0.97 at the Olympic Mountain landslides (Table 1).

The average maximum grain size (see Section 4.2 for field methods) varied from 0.2 m at the Black hills sites to nearly 1 meter at the Rocky Mountain Site (Figure 5b). Values of $\bar{E}/\ell$ are relatively consistent between landslides that occur in the same area (see Cascade Mountain and Black Hills sites) and range from 0.007 to 0.041 [m/m] with the highest rate occurring at the Rocky Mountain landslide and the lowest at the Black Hills sitess. For comparison with reported values the average volumetric erosion per unit length of the erosion-dominated region of the runout path (e.g., yield rate, sensu Hungr et al. 1984 or growth factor, sensu Reid et al., 2016), dividing $\sum E\Delta x^2$ by the erosion-dominated channel length (Table 1), shows that growth factors ranged from 10 m^3/m at the Black Hills South site to 95 m^3/m during the Rocky Mountain landslide.

The topographic slope at which observed deposition occurred, measured as the slope of all cells of the 10 meter DEM representation of the terrain in which observed $\Delta\eta > 0$, was as high as 65% at some sites (Figure 5c); however positive net-deposition tended to occur at lower slopes (Table 1). The slope of areas with positive net deposition tended to be lowest at the Black Hills site (<1% to 10%) and highest at Rocky Mountain site (16% to 25%).

3.4.4.2 Calibration results

Markov chains, colored according to model calibration score, $L(\theta)$ are plotted in the S_c- qs_c domain, along with histograms of sampled S_c and qs_c values for each landslide in Figure 6. The chains show a wide array of sampling patterns and parameter ranges. Broadly speaking, at all sites, the chains converged, with growing $L(\theta)$, towards either a small bullseye (b, c, d, f), or wider aerial patterns (a, e). In the former, the model shows high performance with only a single or a very narrow range of few S_c- qs_c pairs while in the latter, similar model performance (nearly equivalent $L(\theta)$ values) can be achieved or a wider range of sample pairs. The convergence patterns also vary.

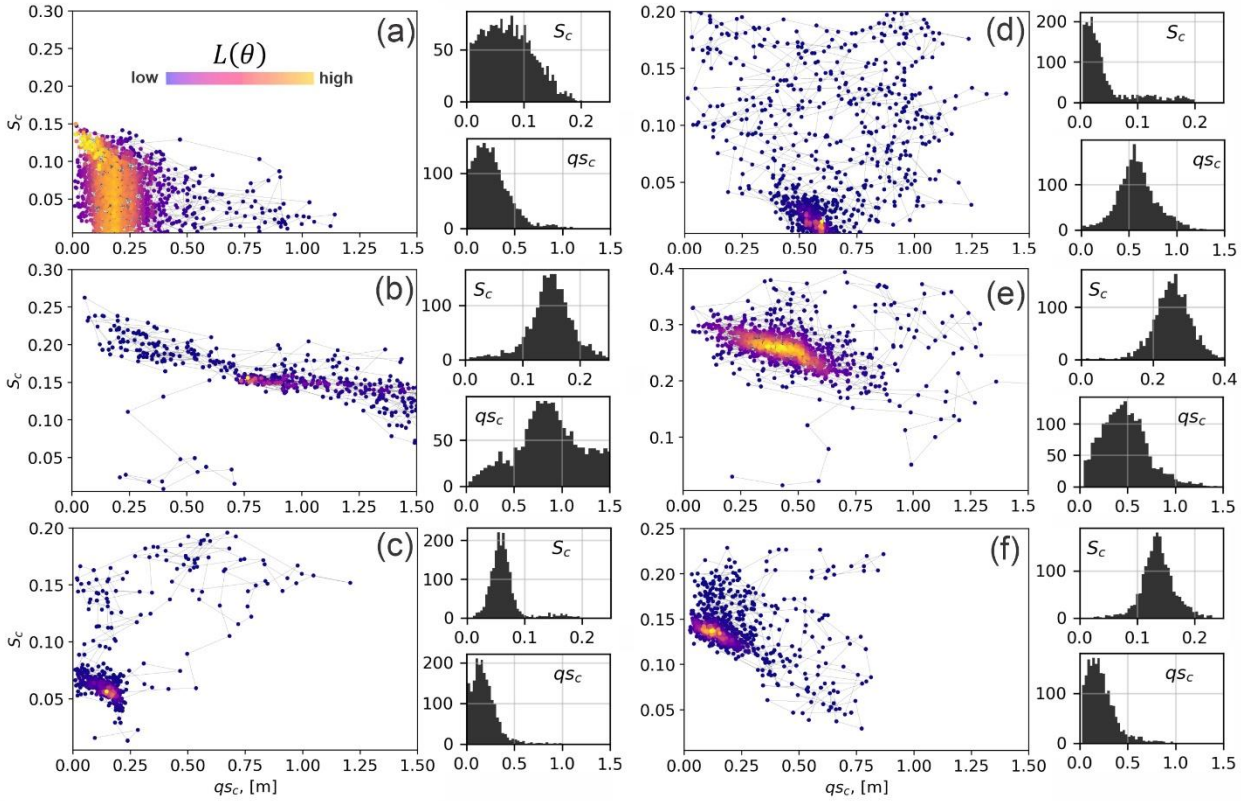

Figure 6. MWR calibration results to landslides described in Section 4 for (a) Cascade Mountain, 2009; (b) Cascade Mountain, 2022, (c) Black Hills, South; (d) Black Hills, North; (e) Rocky Mountain and; (f) Olympic Mountain. Each result shows a scatter plot of the sampled S_c and qs_c values, colored by their relative $L(\theta)$ value. To the right of each scatter plot are histograms of the iterated S_c and qs_c parameters that improved the model results, which represent an empirical PDF of the possible S_c and qs_c values that calibrate MWR to each site.

The Cascade Mountain, 2009 and Cascade Mountain, 2022 landslides, which flowed over the same hillslope, reveal that best-fit model parameters can vary at the same location. The calibration utility converged to smaller qs_c and S_c values for the Cascade Mountain, 2009 event, which permitted thinner flow over lower slopes and effectively made the 2009 modeled runout more mobile relative to the 2022 modelled runout. Even though MWR does not have explicit representations of complex debris flow rheology and momentum, it appears it can be parameterized in relation to the initial volume and runout behavior of the slide. On the other hand,

the Cascade Mountain, 2009 and 2022 models used different underlying initial DEMs and differences in parameterization could reflect the degree to which those DEMs represented actual terrain (See Supporting Information).

Additionally, modeled flow of the Cascade Mountain, 2009 slide show how underlying topographic slope controlled runout direction (Figure 7). In Figure 8, profile plots of modeled Q_s (Equation 29) and maps of the modeled planimetric runout extent, colored to indicate where the runout matched (α), overestimated (β) or underestimated (γ) the observed runout (Equation 26) illustrate observed and modeled flow behavior. At the Cascade Mountain site, because a large portion of the transported sediment was lost to fluvial erosion in the valley, we limited the comparison of observed and modeled Q_s to the upper 750 m of the hillslope (see Supporting Information). The profile plots of Q_s at the Cascade Mountain site (Figure 8a and 8b) show that during the 2009 landslide, all of the runout material flowed past the first 750 meters of the runout path (Figure 8a) while during the 2022 landslide, material began to deposit just down slope of the initial landslide scar (Figure 8b).

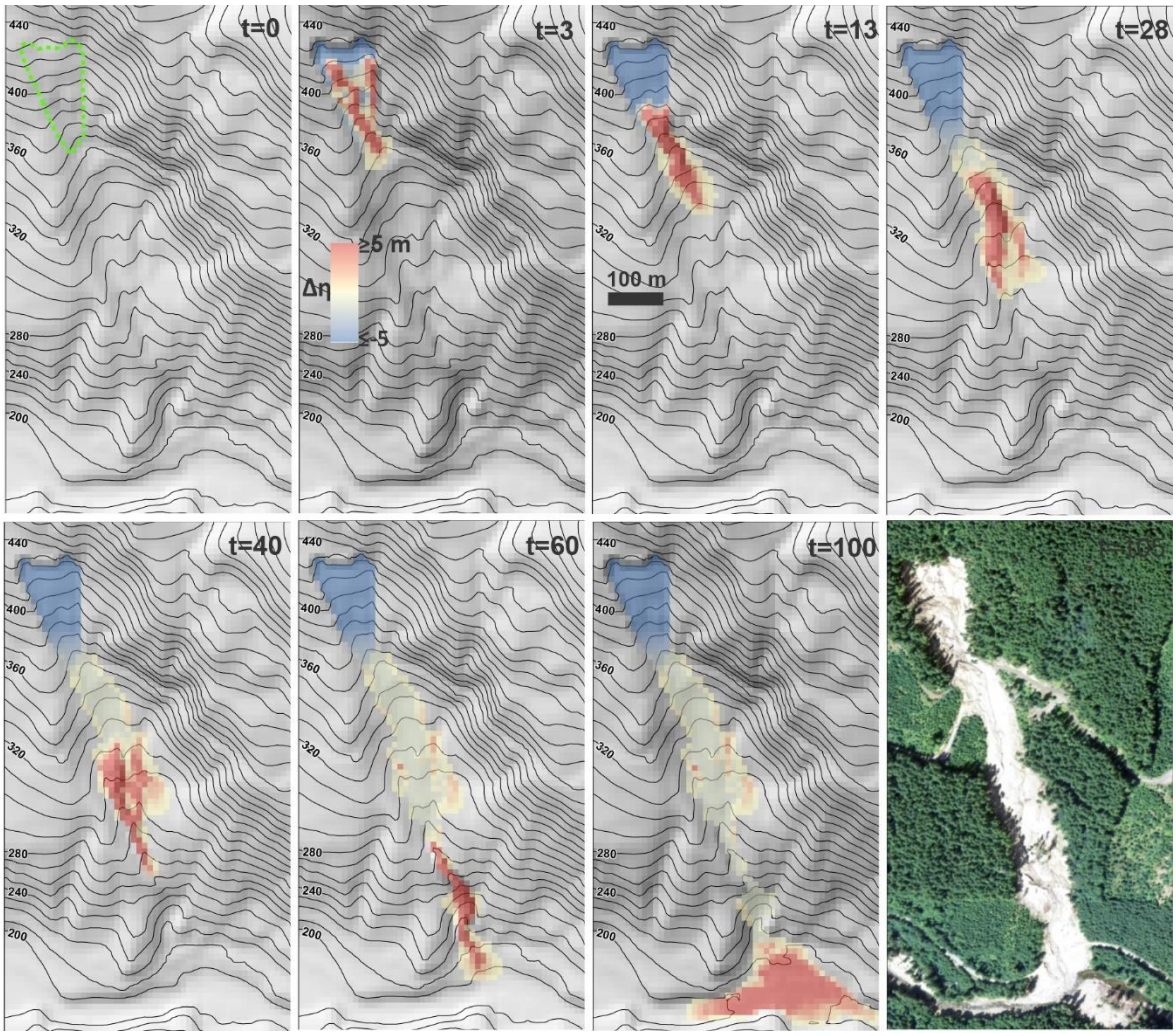

Figure 7. Illustration of modeled runout at the Cascade Mountain, 2009 landslide. At iteration $t = 0$, Algorithm 1 determines the direction and flux of the initial debritons (all nodes located in the landslide green-dashed polygon). In later iterations, Algorithm 2 routes the debritons down slope, entrains material, forms new debritons and updates the terrain. By the end of the modeled runout, a colluvial fan forms at the base of the slope. Topography lines reflect the underlying terrain, which is updated after each iteration.

Using the best-fit parameters from the calibration utility, MWR successfully captures the spatial variation in Q_s and runout extent at most sites (Figure 8). Values of Ω_T are comparable or higher than reported values of Ω_T for landslides in California that were modeled using a variety of detailed, mechanistic and top-down models (Gorr et al., 2022; Barnhart et al., 2022; Note, to

compare Ω_T values, subtract 1 from values reported in this study). Calibration was best at the Cascade Mountain, 2009 landslide (values of Ω_T are highest and value of $\Delta\eta_E$ and Q_{s_E} are lowest) and calibration was poorest at the Rocky Mountain and Olympic Mountain sites (Values of Ω_T are lowest Q_{s_E} and $\Delta\eta_E$ are highest). As noted in Section 4.2, at both the Rocky Mountain and Olympic Mountain sites, we lacked repeat lidar and instead created the DoD from a map of field estimated erosion and deposition depths and estimated the pre-vent DEM. The lower calibration scores may indicate that field estimated DoDs and pre-event DEMS were not as accurate as those determined from repeat lidar DEMs.

MWR tended to underestimate the runout width of large landslides where the apparent momentum of the observed the runout allowed it to flow upslope or slope-parallel (Figure 8a and 8c). For example, after the modeled Cascade Mountain, 2009 landslide initially failed in the uplands, the debritons split and then rejoined to flow around a broad ridge in the middle of the runout path and then were funneled into a narrow ravine (iteration t=40 in Figure 7), rather than flowing over and inundating the surrounding topography as was observed (Large red zone in Figure 8a; see Supporting Information). On the contrary, for small to moderately sized landslides that flowed in highly convergent sections or channels, (Olympic Mountain and Cascade Mountain, 2022 sites; Figures 8f and 8b), or for the moderately-sized Rocky Mountain landslide that flowed on planner and slightly diverging topography (Figure 8e) the model had a tendency to overestimate the flow width. Overestimation of width is particularly pronounced at the Olympic Mountain and Cascade Mountain, 2022 sites, across most locations of the channel network.

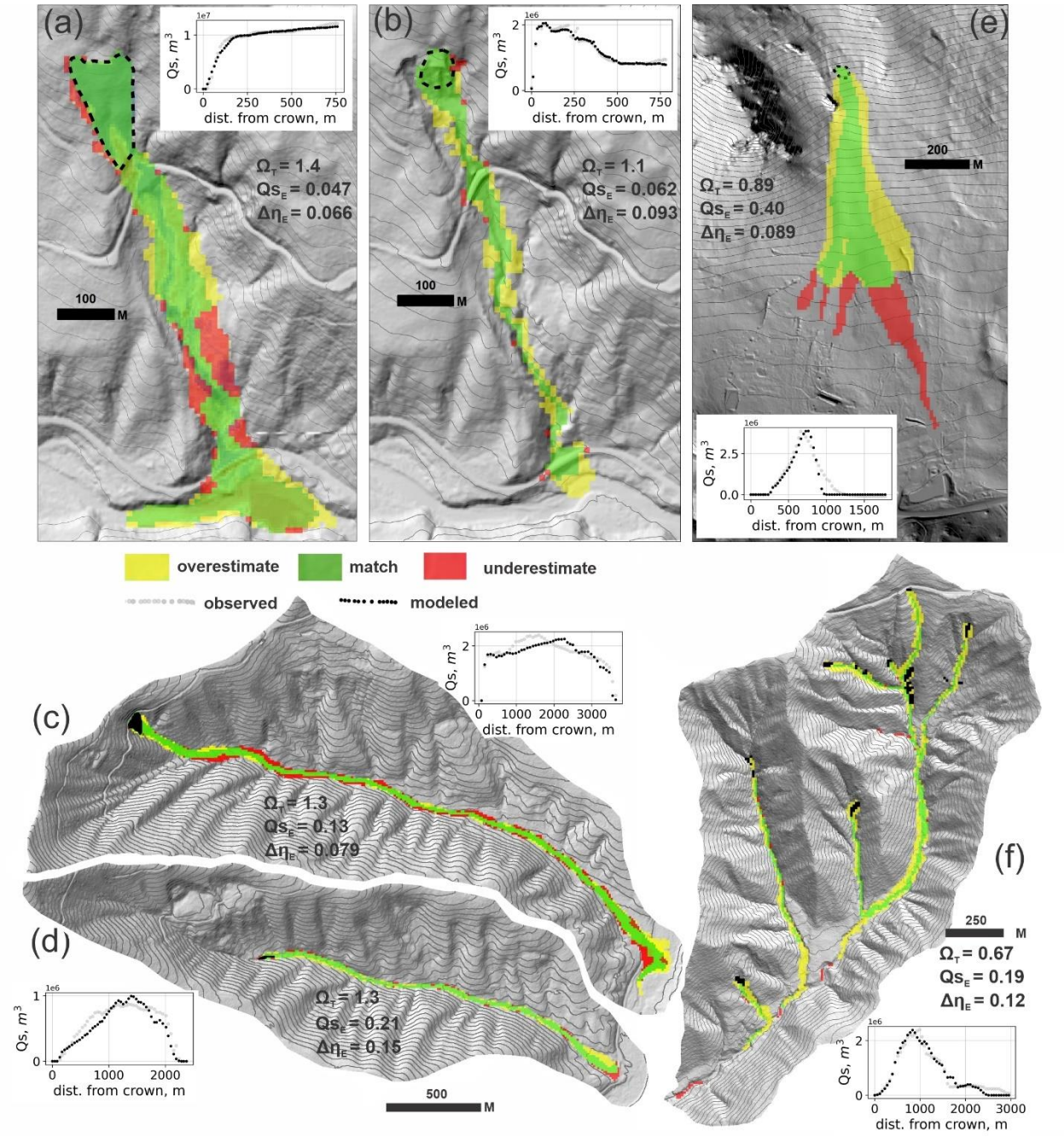

Figure 8. Calibrated model performance as indicated by modeled runout extent, profile plots of Q_s, and reported values of Ω_T, $\Delta\eta_E$ and Q_{s_E}. In all maps, up is north except in (e), north is towards the left. (a) Cascade Mountain, 2009; (b) Cascade Mountains, 2022; (c) Black Hills, South; (d) Black Hills, North; e) Rocky Mountain; (f) Olympic Mountain.

To understand whether the ability to calibrate MWR systematically varies with topography, we compared model performance with three topographic indices computed from the terrain in the observed runout extent. The indices included the relief ratio (H/L), mean total curvature (κ) and the mean specific stream power index (SPI, Chen & Yu, 2011) (Figure 9). The index H/L equals the average slope of the runout path (or relative relief), determined as the total topographic relief of the runout (measured from the top of the initial landslide to the end of the runout path) divided by the horizontal length of the runout and indicates the internal and external resistance to the observed mass wasting runout (Iverson, 1997). Index κ represents topographic curvature, which is the second derivative of the terrain surface, with increasingly positive values of index κ reflecting increasing topographic convergence (e.g., Istanbulluoglu et al., 2008). The index SPI is determined as the natural log of the product of the contributing area and slope. We computed indices κ and SPI at each node in the observed runout extent and the mean was determined from all nodes in observed runout extent.

Evaluation of model performance with respect to the topographic indices shows: slightly improved model performance on less convergent terrain (lower SPI and κ) and slightly improved performance on steeper terrain (higher H/L). The former finding may have to do with how well topographic features that controlled the observed runout were represented by the model DEM. In all models, we used a 10-meter DEM (see Section 4.2), which may have lacked sufficient topographic detail to represent the relatively small first and second order channels that controlled observed runout extent of small to moderate sized landslides like the Olympic Mountain and the Cascade Mountain, 2022 sites (Supporting Information). Without a sufficiently accurate topographic representation of the topographic features that controlled observed flow, modeled flow tended to flow over a wider area of the hillslope than observed. The latter finding appears to

be mostly a result of how well modeled sediment transport and topographic change (Q_{s_E} and $\Delta\eta_E$) replicated observed, as there does not appear to be a trend in Ω_T with H/L. At the high H/L sites, like the Cascade Mountain landslides, a distinct slope break largely determined zones of deposition and erosion. Slope controls on erosion and deposition rates are explicitly represented in MWR. At the lower H/L sites (Black Hills and Olympic Mountain sites), other factors, such as woody debris, also impacted the distribution of deposition and scour (detailed in Supporting Information). Although MWR lacks an explicit representation of woody debris impact on runout, through the calibration processes, values of qs_c and S_c that allowed MWR to more closely replicate the observed flow were identified, but modeled runout didn't replicate observed quite as well as sites where slope appeared to be the primary control on observed runout behavior.

In summary, using a variety of methods to prepare the DoD and model terrain, the calibration utility was able to identify clear best-fit S_c and qs_c values at a range of different landslide types. Calibration likelihood ($L(\theta)$) tended to vary with topographic indices of the runout path (H/L, κ and SPI). How well the DEM represented topographic features that controlled observed runout, the importance of runout processes that are not explicitly represented in MWR on the observed runout behavior (such as woody debris at the Black Hills and Olympic Mountain sites) and the accuracy of the initial DEM and the DoD datasets (field mapped vs lidar DEM determined) impacted model performance. To a certain degree, the calibration utility identified parameter values that minimized differences between modeled and observed runout, and thus allowed MWR to mimic the impacts of processes that are not explicitly represented in the model, such as momentum (Cascade Mountain, 2009 site) and the impact of woody debris (Black Hills and Olympic Mountain sites).

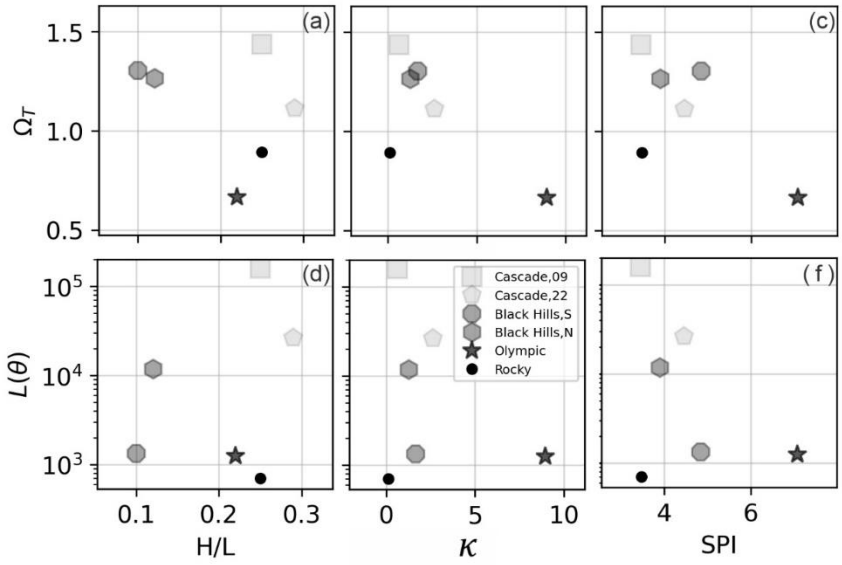

Figure 9. Illustration of model calibration, as reflected by the posterior parameter likelihood $L(\theta)$ and planimetric fit (Ω_T) relative to topographic indices.

3.5 Discussion

3.5.1 Strategic testing of MWR for hazard mapping applications

Having demonstrated that MWR can be calibrated to a variety of landslides and runout terrains, we now strategically test MWR using the Cascade Mountain and Black Hills sites. Since both of these sites consisted of two separate landslides, we can thus test model performance by swapping best-fit model parameters, rerunning the models and comparing results with the original, calibrated results. At the Cascade Mountain site, the 2009 and 2022 landslides originated on the same hillslope (Figure 8a and 8b). At Black Hills site, the two landslides occurred on different hillslopes but in adjacent east-west oriented watersheds (Figure 8c and 8d).

As shown in Figure 10, at both the Black Hills and Cascade Mountain sites, when the best-fit parameters from the other landslide are used to predict runout, the accuracy of modeled runout planimetric extent drops but resultant Ω_T values are still as high or higher than values reported in other studies (compare to equivalent Ω_T values in Gorr et al., 2022 and Barnhart et al., 2022). In

terms of modeled sediment transport and topographic change, swapping best-fit parameters has a larger effect. At the Cascade Mountain, 2009 landslide, using the 2022 best-fit parameter values causes about half of the modeled runout material to prematurely deposit on the hillslope, reducing the amount of sediment that reaches the valley floor (Q_{s_E} increases by a factor of nine; Figure 10). Using the 2009 parameter values for 2022 landslide (Figure 10b) increases modeled runout extent and results in nearly four times the entrainment and transport of sediment to the valley floor, causing Q_{s_E} to increase by a factor of 20 and $\Delta\eta_E$ by 83%. At the Black Hills site, using the South basin best-fit model parameters at the North basin causes Q_{s_E} and $\Delta\eta_E$ increase by 83% and 39% respectively (Figure 10c). Unlike the other three landslides, swapping best-fit parameters at the Black Hills, South basin results in both large sediment transport and runout extent error because the North basin best-fit parameters cause modeled landslide to entrain too little and stop only a few hundred meters from the initial source area (Figure 10d).

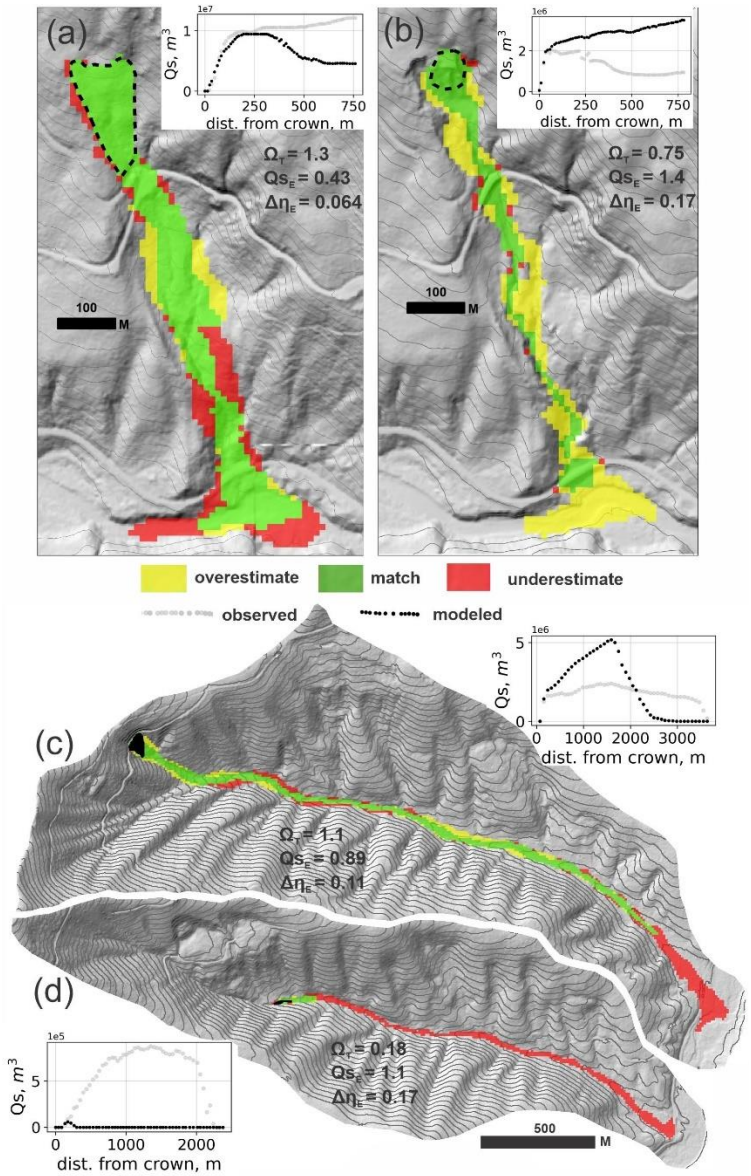

Figure 10. Model performance using the neighboring landslide parameter values, as indicated by modeled runout extent, profile plots of Q_s, reported values of Ω_T, $\Delta\eta_E$ and Q_{s_E} and profile plots of Q_s. Compare with Figure 8. (a) Cascade Mountain, 2009; (b) Cascade Mountain, 2022; (c) Black Hills, South; (d) Black Hills, North

An alternative test to simply swapping the best-fit parameters is to swap parameter PDFs determined from the calibration utility and compare probabilistic runout. For this test, we run MWR Probability with option (1), using the field mapped landslide polygons. As shown in Figure 11, at three of the landslides, using the parameter distribution associated with the neighboring

landslide results in relatively minor changes in the probability of runout (probability of ≥50%; Figures 11a, 11b and 11d) and mostly changes the terminal extent of the possible landslide runout area. Similar to the results of swapping the best-fit parameter values, at the Black Hills South landslide, swapping parameter PDFs causes a large change in runout probability (Figure 11c).

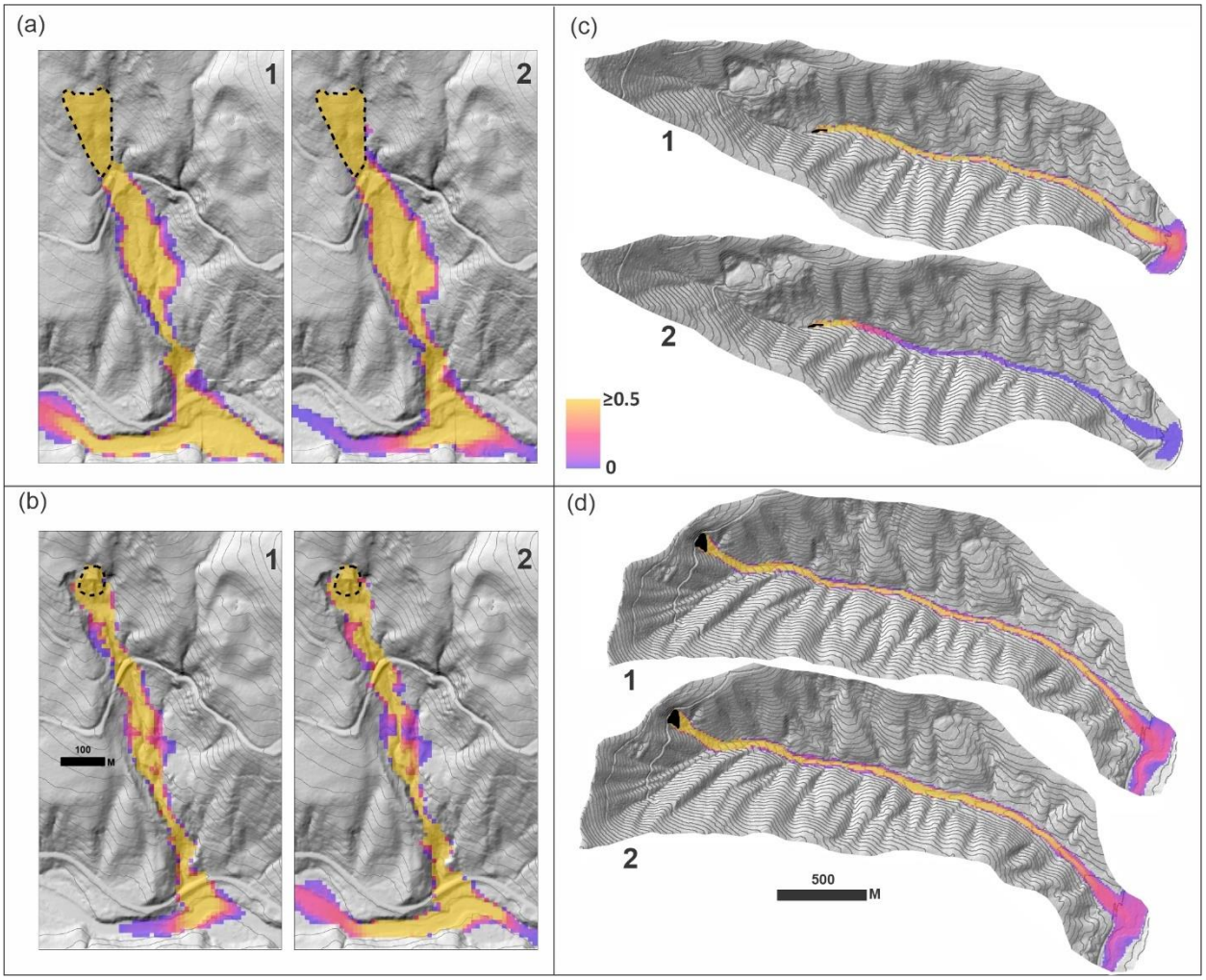

Figure 11. Model tests by swapping parameter PDFs and comparing runout probability at the (a) Cascade Mountain, 2009; (b) Cascade Mountain, 2022; (c) Black Hills, South and; (d) Black Hills, North sites. (1) the parameter distributions from calibration to the site and (2) the parameter distributions from calibration to the neighboring site.

These findings suggests that in most cases, once best-fit parameters or parameter PDFs have been established for a landslide, those values can then be used to determine runout extent at nearby sites. At the same time, sediment transport and topographic change may be more sensitive to model parameterization. However, we suspect that had landslide processes at the Cascade Mountain and Black Hills sites been more uniform, swapping best-fit parameters would have resulted in similar modeled sediment transport and topographic change as well. For example, to a certain degree, a single parameterization seemed to capture both the runout extent and depositional patterns of all slides at the Olympic Mountain site. There, most landslides tended to be stream adjacent and relatively uniform in size. In contrast, at the Cascade Mountain site, the 2009 landslide was five times the volume of the 2022 landslide. Both the volume and apparent momentum of the 2009 slide allowed it to flow topographically unconstrained while the much smaller 2022 landslide was topographically confined to a narrow channel ravine. The calibration utility found that relatively low S_c and qs_c values best replicated the Cascade Mountian, 2009 flow behavior while higher S_c and qs_c values best replicated the Cascade Mountain, 2022 behavior. Similarly, at the Black Hills site, the north watershed landslide was over ten times the volume of the southern watershed landslide, and initiated 150 hundred meters above the channel, on a steep hillslide (detailed in Supporting Information) while the much smaller south watershed landslide initiated at the elevation of the channel bed. Again, differences in landslide behavior equated to different parameterizations of MWR. Had the Cascade Mountain, 2009 and 2022 landslides or the Black Hills North and South landslides been similar in size and runout behavior, best-fit S_c and qs_c values and PDFs may have been similar and test results may have been better.

In summary, the Cascade Mountain and Black Hills sites each contained two landslides that we could use to independently calibrate MWR and thereby evaluate model predictive ability by

swapping parameter values and comparing model results with calibrated model results. At three of the landslides, the tests revealed a modest decrease in the accuracy of modeled runout extent (though accuracy of the reduced model extent was still comparable to values reported in the literature) but a substantial decrease in sediment transport and modeled topographic change. We suspect that the poor sediment transport and topographic change test results may have been caused differences in landslide processes. In regions where landslide processes are relatively uniform (like the Olympic Mountain site), calibration to one landslide might be sufficient to predict the depositional patterns of another without multiple landslide parameterizations. At sites like the Cascade Mountain and Black Hills sites, which consisted of two very different landslide types and runout processes, MWR may need to be calibrated to each type of landslide (e.g., small-low-slope vs. large-steep-slope) and predictive application of would involve applying the appropriate parameter set based on landslide type.

3.5.2 Mass wasting runout probability applications

In this section we demonstrate how to determine runout probability from a probabilistically determined landslide hazard map or a specific, potentially unstable slope using MWR. The first application may be appropriate for watershed- to regional-scale runout hazard assessments. The second application is an example site-scale hazard assessment. Both applications are demonstrated at the Olympic Mountain site where landslide size and type tended to be relatively uniform and parameter PDFs determined through calibration may therefore represent typical runout processes in the basin.

3.5.2.1 Runout probability from a landslide hazard map

To determine runout probability from a landslide hazard map, we ran MWR Probability using option (3) at Olympic Mountain site. We use LandslideProbability (Strauch et al., 2017), an existing component in Landlab that computes landslide probability by repeatedly determining the hillslope Factor-of-Safety (FS: ratio of the resisting to the driving forces) at each node on the raster model grid from stochastically selected soil (regolith) hydrology properties, (including soil depth, saturated hydraulic conductivity) soil strength (friction angle, cohesion) and recharge rates (precipitation input rate minus evapotranspiration and soil storage). We setup LandslideProbability using soil strength and hydrologic parameters based on the soil class (SSURGO, 2020) and vegetation type at Olympic Mountain site (Strauch et al., 2017). We defined soil depth as a function of contributing area, slope gradient and minimum and maximum depths following Westrick (1999). Minimum and maximum soil depths were estimated from field observations at channel banks, tree-throw pits and road cuts. For recharge, we assumed evapotranspiration and soil storage losses were small and used a uniform daily precipitation depth equivalent to the 50 year event (determined from rainfall records, WRCC, 2017). We ran $N = 1000$ model runs. For each model run, FS and soil depth maps were saved and after all N runs were complete, a single map of landslide probability was produced. We then ran MWR Probability using each of the N FS and soil depth maps and treated all nodes with $FS < 1$ as a landslide source. Each repetition began with the same initial topography but different landslide locations and regolith thickness. Landslide probability and landslide runout probability are shown in Figures 12a and 12b. If changes in the elevation of the channel were a concern, MWR probability could also be used to determine the probability of scour or deposition. For example, the probability of deposition greater than 1 meter is shown in Figure 12c.

Interestingly, while landslide probability is relatively low over most of the basin, runout probability is high in many of the second order channels but then low again at the basin outlet. In a field study of the spatial and temporal patterns of debris-flows, May and Gresswell (2004) observed that debris flow return interval tends to be low (debris flow probability tends to be high) in channels which are fed by a greater number of upslope landslide source areas; however, our model results illustrate how debris flow (landslide runout) probability is both a function of the number of upslope landslide source areas as well as the typical runout distance of those source areas. For example, despite having the largest number of potential landslide source areas, runout probability at the basin outlet is low because it is located beyond the typical runout distance of landslides at the Olympic Mountain site.

3.5.2.2 Runout probability for a specific, potentially unstable slope

When field evidence or other data indicate that a specific slope may be potentially unstable, but the exact area of a potential slide on that slope is unknown, MWR can be used to determine potential runout hazard of the slope. For this application MWR Probability is run using option (2) which requires a polygon representing the maximum extent of the potentially unstable slope. As noted above, for each model repetition, a landslide area is randomly selected within the potentially unstable slope, it can form anywhere within the potentially unstable slope and is least as large as a user defined minimum size but no larger than the potentially unstable slope. After a user defined N repetitions, runout probability of the potentially unstable slope is determined at each downslope node from (31).

As an example application of using MWR Probability option (2), we designated a 0.6 ha, zero-order basin in the headwaters of the Olympic Mountain site as a potentially unstable slope and modeled runout probability using $N = 1000$ (Figure 12d). This example shows that if a

landslide were to initiate at the potentially unstable slope, the probability of that material running out to the basin outlet as a mass wasting runout process is less than 5%. This result is again a function of the typical runout distance of landslides at the Olympic Mountain site.

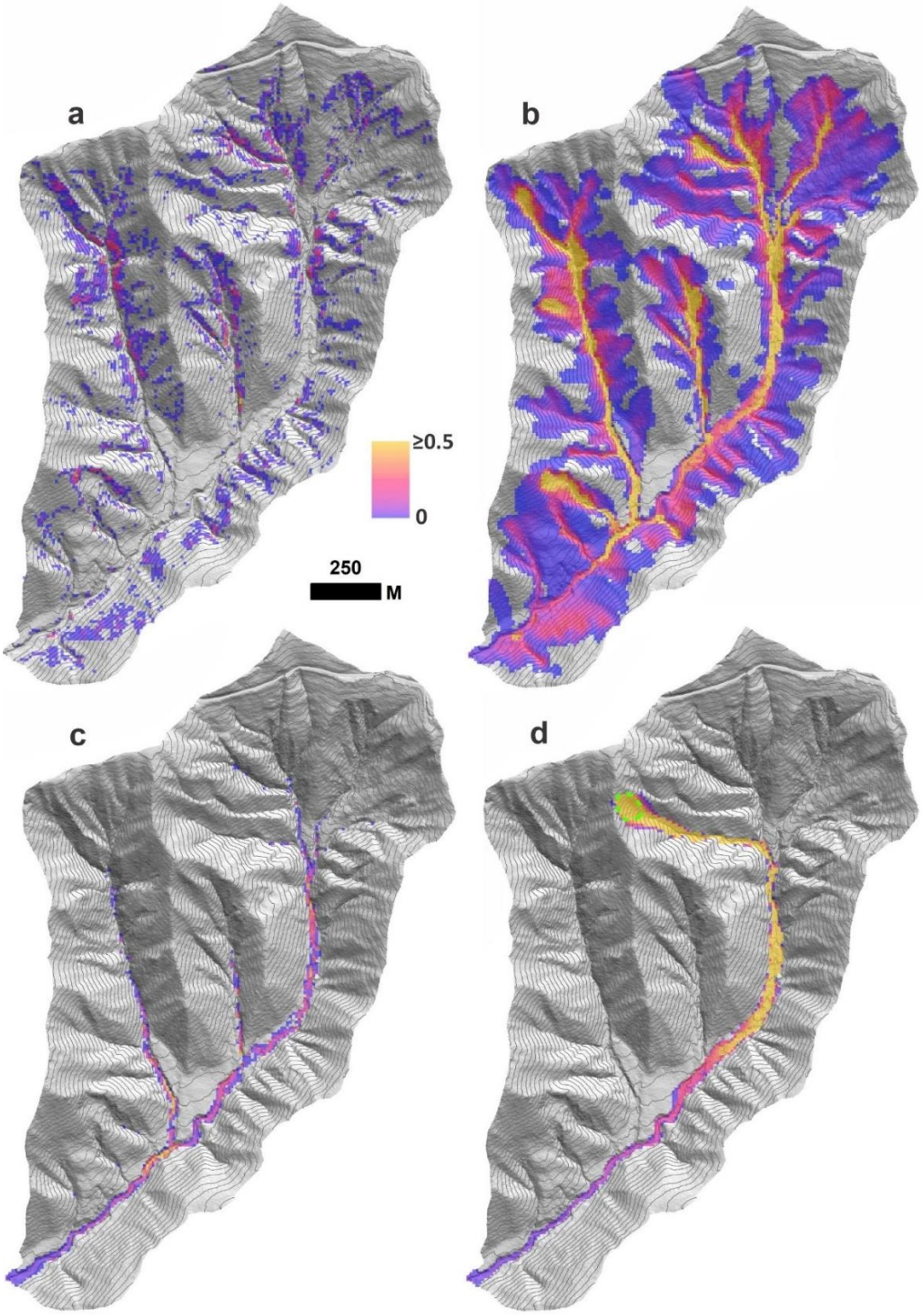

Figure 12. Olympic Mountains site: (a) Landslide probability, $P(FS \leq 1)$. (b) Corresponding runout probability, $P(\Delta\eta)$. (c) Probability of deposition greater than 1 m and (d) Runout probability for the potentially unstable slope (green-dashed polygon).

3.5.3 Model limitations and strengths

In the field, we observed a number of runout phenomenon that are not explicitly represented by MWR, such as the apparent affect of momentum, debris fluidization, sorted deposition and the impact of standing vegetation and woody debris on the runout process (detailed in Supporting Information). For example, observed runout at the Olympic Mountain site appeared to be restricted by log jams which in turn lead to large, valley-filling sediment deposits. At the Rocky Mountain site, the lateral extent of the runout path appeared to have been restricted by standing trees. Also, the momentum or perhaps fluidization of the initial landslide appears to have impacted small areas of the runout extent at Black Hills North landslide and the Cascade Mountain, 2009 landslide.

Additionally, grid-size appears to have affected the lateral extent of modeled runout. At all sites, we used a 10-m grid size. Topographic detail of the channels that constrained observed flow to a narrow width at the Olympic Mountain and Cascade Mountain, 2022 sites were likely lost in the 10-m DEM. On the other hand, for landslides like the Cascade Mountain, 2009 slide and the Rocky Mountain slide, the 10-m grid size more than adequately captured the key topographic features that controlled runout width and a lower resolution (larger grid size) DEM may have resulted in equivalent model performance. This last point is important because the number of debritons and computations needed to model the initial landslide is equal to the initial landslide area divided by the square of the grid size (e.g. a 100 m^2 landslide on a 10 m grid is represented by 1 debriton but on a 1-m grid is represented by 100 debritons). Using an excessively small grid size may result in an unnecessary number of computations. We used a 10-m grid size because it was the coarsest grid size that didn't exceed the minimum flow width of the four validation sites and it appeared to capture most of the topographic detail that controlled observed runout. In future

work, we hope to provide a more comprehensive assessment of grid size impact on model performance.

Nonetheless, MWR is a parameterized model, and to a certain degree, it appears MRW can be calibrated to compensate for DEM resolution or the absence of explicit representation of all processes that controlled the observed flow. At the Rocky Mountain site, the qs_c parameter could be used to mimic the resistance to lateral spreading caused by the trees. At the Olympic Mountain and Cascade Mountain, 2022 sites, even though modeled flows tended to be wider than observed, the calibration utility identified qs_c and S_c values that allowed MWR to replicate the longitudinal sediment transport behavior of the observed runout.

Finally, through strategic testing, we showed that once MWR is calibrated to a landslide it may reasonably evaluate the extent of runout or runout probability at neighboring slopes. This may be particularly true in areas that have relatively uniform landslide processes (Olympic Mountain site). At sites like the Black Hills and Cascade Mountain sites, where landslide processes are highly variable, MWR may have limited predictive ability for sediment transport and topographic change modeling and in some cases even runout probability. In a region that has a diverse range of landslide phenomenon, accurate landslide runout predictions may require calibration to each type of landslide.

The utility of a model is ultimately judged by its ability to explain or predict natural geomorphic phenomena, even if the interactions within the model greatly simplify the physical processes controlling the phenomena (Murray 2007; Murray 2013). Furthermore, the model should be rejected if it fails to react to changes in forcing in a reasonable way or cannot replicate the intended phenomena. From this standard, MWR appears to pass as an acceptable model, because it was able to replicate a range of landslide behavior and because it responds in an expected way

to changes in terrain or parameters like S_c, qs_c and grain size. For example, at the Olympic Mountain site, MWR was able to replicate the depositional patterns caused by a 30-year series of landslides, because like real mass wasting events, modeled mass wasting events in MWR modify both the underlying terrain and sediment availability and thus the occurrence of one runout event impacts the runout behavior of a later event. MWR is a parameterized model and while the predictive application of MWR may not be as restricted as implied by Iverson (2003), we suspect that calibration may be required for most applications. For the purpose of facilitating rapid calibration to different sites/watersheds and developing parameter PDFs, we made MWR parameterization as field-relatable as possible and developed a calibration utility.

3.6 CONCLUSION

In this study, we described, calibrated and tested MassWastingRouter (MWR), a new cellular-automata landslide runout model designed for probabilistic landslide hazard assessments, sediment transport and topographic change applications. MWR is implemented in Python as a component for the Landlab earth surface modeling toolkit. MWR includes a Markov Chain Monte Carlo calibration utility that determines the best-fit parameter values for a site as well as empirical Probability Density Functions (PDF) of the parameter values. MWR also includes a utility called MWR Probability that takes the PDF output from the calibration utility to model runout probability. It has three options for defining the initial landslide: (1) a single, fixed-size landslide, (2) a single, variable-size landslide within a user defined potentially unstable slope or (3) a single or multiple landslides that change in location and size each model iteration.

We used the calibration utility to parameterize the model to six different landslides. The following results were obtained:

- MWR can be calibrated to a range of landslide types. The ability to calibrate MWR is slightly better on a steep (high H/L) and broadly convergent (lower κ and SPI) runout path. But this finding may partly be a consequence of the 10-m grid-size used to model runout at each of the six landslides, and needs further research. The six landslides varied from small to large and the ability to calibrate MWR appears to be better at the sites in which the terrain features that controlled runout extent were larger than the grid size.
- The two model parameters, critical slope (S_c), and a threshold flux for deposition (qs_c) are found to effectively control model performance over a range of landslide and landscape conditions and, to a certain degree, through calibration, MWR can be parameterized to mimic the effect of processes or terrain not explicitly represented in the model.
- Once MWR is calibrated to an observed-runout dataset, it may work as a predictive tool for assessing the runout extent of neighboring potential landslides. To predict depositional and topographic change, depending on the diversity of landslide types in the region, a more rigorous calibration may be required such as calibrating the model to different landslide types.
- We demonstrated how to use MWR to determine mass wasting runout probability from an expert-defined potentially unstable slope polygon and a landslide hazard map produced by the model LandslideProbability.

3.7 Supporting Information: Observed runout description and model setup

3.7.1 Cascade Mountain landslide

3.7.1.1 Site and mass wasting runout description

The Cascade Mountain site includes two landslides that occurred in 2009 and 2022, one in each year, on a hillslope located roughly 30 km south of Mount St. Helens, USA, near the south-west edge of Washington State. The hillslope is moderately-steep (average slope ~40 to 60%), broadly convergent and dissected by a small, first-to-second order channel. The 2022 landslide was a reactivation of the 2009 landslide headscarp. Two logging roads cross the hillslope and impacted flow behavior, particularly during 2022 event. The underlying regional geology is Miocene-Oligocene volcaniscalistic deposits or rocks (DNR, 2022), which in the field consisted of hydrothermally altered basaltic andesite and epilcastic sandstone with gravel to cobble intrusions. In some areas, the rock was weathered to clay and rippable with a hand shovel. During both events, the runout path was covered in trees, though the size and age of the trees differed.

Both landslides were triggered during heavy rain-on-snow precipitation events (Table 1). For both landslides, overall runout extended 700 to 800 meters downslope from the base of the landslide, ending abruptly when the runout perpendicularly intersected a narrow river valley. The first event was a large, catastrophic debris avalanche (initial landslide volume = 110,000 m^3) that flowed relatively unconfined over a wide area of the downslope hillslope, completely inundating the small channel. Figure 7 shows calibrated modeled reconstruction of initial failure and subsequent runout. At the time of the 2009 event, 0.6 to 0.9 meter diameter conifer trees covered the hillslope. The 2009 event obliterated the logging roads and all trees in the runout path. Roughly 29% of the total mobilized volume was derived from erosion during the runout.

The second event was smaller than the 2009 event, but was still moderate in size (initial landslide volume = 22,000 m^3). By 2022, the roads were reconstructed and the 2009 scar was covered with small (0.1 to 0.2 diameter) hardwood trees. Runout of the second event was largely confined to the small first order channel in the center of the scar of the 2009 slide, forming a channelized debris flow and destroying most of the small hardwood trees in its path. Unlike the 2009 event, except for a ~3m wide area of erosion in the upper road, both the upper and lower road remained intact and acted as a check-dam. Roughly 19% of the total mobilized volume was derived from erosion during the runout.

In the field (March, 2022), deposits consisted of a mixture of 70 to 80% clayey sand matrix with 20 to 30% gravels and cobbles and trace boulders. The slope of areas of positive net deposition ranged from 5 to 15% but slopes as low as 1% may have formed along the lower river valley (Table 1). Along the margins of the runout path, hummocky, cohesive levees formed on a range of slopes (See Figure S1b for example).

3.7.1.2 Observed-runout dataset and model setup

At this site, we have high-resolution DEMs of the terrain before and after each runout event: a 2-meter Lidar DEM recorded in 2006, a 1-meter Lidar DEM recorded in 2019, and a structure-from-motion DEM of the runout path created from unmanned aerial vehicle images, recorded in 2022. The 2019 DEM was recorded before reconstruction of the upper logging road and thus does not accurately represent the topography at the time of failure in 2022. We added the upper road to the 2019 Lidar DEM using AutoCAD Civil3D based on the Engineer's design drawings.

We used the 2006 DEM as the pre-event terrain for the 2009 event and the modified 2019 DEM as the pre-event terrain for the 2022 event. We created the map of $\Delta\eta$ (DoD) for the 2009 event by subtracting the 2006 DEM from the unmodified 2019 DEM and the observed-runout

dataset for the 2022 event by subtracting the modified 2019 DEM from the 2022 DEM. Because the 2019 DEM was recorded nearly 10 years after the 2009 event, by the time the 2019 DEM was recorded, much of the deposit near the channel in the center of the runout path and most of the deposit in the river valley appears to have been eroded by fluvial erosion. Also, the lower road was reconstructed. For the 2022 event, most of the runout material was still in-place when the 2022 DEM was recorded.

For the model setup, we used a 10-m grid size. We assumed an average erodible regolith of 1.2 meters at all locations on the channel based on the depth of regolith visible in road cuts and the average observed erosion depth. Because much of the 2009 deposit in the river valley was eroded after the event and no longer represented in the 2019 DEM, we limited the comparison of Q_s to the hillslope above the lower river valley, which corresponded to roughly the first 750 meters of the runout path (measured from the landslide crown).

3.7.2 Black Hills site

3.7.2.1 Site and mass wasting runout description

The Black Hills landslides occurred in 2007, in adjacent, east-west oriented basins, in response to an extreme precipitation and snowmelt event. The southern basin has an area of 2 km2 and is drained by a second order channel. The northern basin has an area of 2.8 km2 and is drained by a third order channel. Logging roads cross both the headwater and the outlet of the basins. Underlying geology is weathered basalt covered by a thick regolith that includes a deep (2 to 5 meters) saprolite layer of clay-rich soil with occasional cornerstones of basalt and exotic, glacially derived clasts (DNR, 2022).

In the southern basin, a deep-seated bedrock landslide partially mobilized on a moderately steep slope (40%) but came to rest upon intersecting the channel that runs along the center of the

basin. A small portion of the toe of the landslide broke off (initial landslide volume ~1,500 m^3) in the channel and continued flowing until reaching the logging road at the mouth of the basin. In the northern basin, a large landslide (initial landslide volume ~ 18,500 m^3) catastrophically failed in the through-fill section of a logging road, located 250 meters upslope of the channel on a 60% slope (vertically 150 meters higher than the channel). Based on super elevation of the flow, evident in the elevation on the left and right sides of the runout path, flow velocity was high. Despite differences in initial landslide velocity near the top of the runout path, over most of the runout path, both landslides progressed down relatively low-gradient channels as debris flows, eroding and entraining debris and large trees (0.2 to 0.7 meters) until eventually depositing near the junction of a larger channel. The total mobilized volume derived from erosion during the runout was 89% and 59% at the southern basin and northern basin respectively.

In the field (December, 2021), the runout deposit in the southern basin consisted of 65% sandy silt matrix with 35% gravels. In the northern basin, the runout deposit consisted of 45% sandy matrix with 55% gravels. The slope of positive net deposition deposits measured in the field ranged from <1 to 10%. Most of the entrained woody debris was transported to the end of the runout path where it deposited as large, 2 to 3 meter high stacks of logs. Some woody debris did deposit at higher locations in the channel, which in turn caused sediment to deposit on relatively steep reaches of the channel. At this site, we assumed an average erodible regolith depth of 1.5 meters.

3.7.2.2 Observed-runout dataset and model setup

Here we have a 2-meter DEM recorded in 2005 and a 1-meter DEM recorded in 2011. We had to translate the 2005 DEM vertically roughly 0.5 m to get rid of a systematic offset with the 2011 dataset. We created the observed-runout dataset by subtracting the translated 2005 DEM from a 2-meter resampled version of the 2011 dataset. Much of the 2005 DEM appears to have been

impacted by vegetation interference but was generally higher quality over the valley bottoms in the mixed hardwood-conifer stands. Nonetheless, in the field, we verified areas of scour or deposition signals with the Lidar difference. Similar to 2019 DEM at the Cascade Mountain site, areas of deposition appear to have been fluvially eroded between between the pre-and post-event DEM.

We use the 2005 translated Lidar, resampled to a 10-meter grid and smoothed using a low-pass filter (ArcGIS, 2022), as the pre-event terrain. Smoothing was required because of a few artificial bumps in the channel surface that were not present in the field.

3.7.3 Rocky Mountain site

3.7.3.1 Site and mass wasting runout description

The Rocky Mountain landslide occurred in September, 2013, following an intense period of rainfall (Patton et al. 2016) on a steep, broadly convergent to divergent hillslope. It was a moderately sized debris avalanche (initial volume ~4600 m^3) sourced from thick (>6 m) colluvium. The area is underlain by gneiss, schist, migmatite and Proterozoic granites (USGS, 2022). In the field (August, 2022), bedrock was generally covered by thick colluvium, which consisted of a sandy gravel matrix mixed with cobbles and very large boulders (diameter > 2 meters). Evidence of macropore-flow and infiltration-excess surface runoff were visible in the headscarp. The only water available to mix with the landslide debris appears to have been limited to these runoff sources.

The hillslope below the landslide is planar and moderately steep (40 to 50 percent) for about 350 meters and then transitions to a 20 to 30 percent, broadly divergent slope. At the time of the failure, the hillslope was densely forested with 0.2 to 0.4-meter diameter conifer trees with rooting

depths of about 0.5 to 1 meters. Most erosion occurred on the planar hillslope and removed the upper 0.5 to 1 m of colluvium (inferred from the tips of tree roots visible in areas of erosion) but channelized erosion as deep as 4 to 5 meters also occurred.

Unlike the other validation sites, as the debris flowed from the landslide to the lower extent of the runout path, it spread laterally. Also, here, vegetation, rather than convergent topography, appears to have largely controlled the lateral extent of the flow. Along the edge of the runout path, a nearly continuous berm of logs developed as the mobilized material flowed down slope (see Figure S1c).

Positive net deposition began at gradient of 16 to 25%, near the slope break between the planar slope and the divergent slope. Once on the divergent hillslope, the runout split, with a small lobe of material flowing northwest, and the majority of the material flowing southwest. Relative to the other sites, deposited material near the slope break was unusually well sorted and thick (2-3 meters). Deposited material on the northwest lobe was mostly fine gravel to sand sized and deposition depths were relatively shallow (roughly 10 to 30 cm). Material that flowed to southwest included many large boulders, which deposited just downslope of the thick gravel deposit. A thin (0.2 to 0.5 meter) sand and fine gravel deposit continued past the boulders. An eyewitness account of the deposit shortly after the runout described the terminal end of the runout as a mudflow. The average composition of observed deposited material was 40% sandy gravel matrix with 60% gravels, cobbles and boulders. Here, 88% of the total mobilized volume was derived from erosion during the runout but the runout was not long, which resulted in an extreme growth factor of 95 m^3/m of channel length.

3.7.3.2 Observed-runout dataset and model setup

Here we only have one, post-event Lidar dataset, recorded in 2015. We rely on field observations of erosion and deposition depths to approximate the DoD. We used in-place tree roots in the runout-scar and adjacent, intact topography to estimate changes in elevation within the runout path. Scour depths and deposition depths were generally shallow (<1.5 m). Where deep scour occurred, it was localized to narrow channels, that may have formed after the runout event. At this site, we assumed an average erodible regolith depth of 1.2 meters. We approximated a pre-event DEM by subtracting the field estimated map of $\Delta\eta$ from the post-event DEM, resampling to a 10-meter grid cells and smoothing the DEM using the same low pass filter we applied to the Black Hills 2005 DEM.

3.7.4 Olympic Mountain site

3.7.4.1 Site and mass wasting runout description

The Olympic Mountain site is a small (2.1 km^2), dendritic, fourth order basin. Inspection of the air photo record revealed that over a 30 year period (1980 and 2010), at least twelve small (initial landslide volume ~ 400 to 2200 m^3) landslides initiated along the edges or heads of a first or second order channels and subsequently formed channelized debris flows. Between 1980 and 2010, at least six different precipitation events with 2-day precipitation totals between 100 and 220 mm occurred (Table 1).

The basin is underlain by Miocene-Eocene marine sedimentary rocks (WA DNR, 2022), which in the field appear as steeply bedded densely fractured 1 to 2-meter thick beds of siltstone with occasional beds of 0.3 to 0.5-meter beds of coarse grained sandstone. These beds generally break up into 10 to 30 cm sized clasts, though boulders larger than half a meter in diameter are

also common, with most of the larger clasts derived from the sandstone. Many of the channels were filled with thick colluvial deposits that predated the debris flows triggered after 1980.

Approximately 60% of basin area was logged and replanted between 1978 and 1990. At the time of most debris flows, valley bottoms were not vegetated but logs left from the tree harvests cluttered the channel. Here, few roads were constructed in the watershed and, except for a small landslide (~100 m^3) that initiated from the fillslope of a road, roads had little impact on mass-wasting processes.

In the field (March, 2021), debris flow deposits were visible throughout the channel network and consisted of 40% sandy-gravel matrix with 60% coarse gravels, cobbles and boulders. The surface slope of reaches of positive net deposition generally systematically increase in the upstream direction from ~5 percent near the basin outlet to 15 percent at the upper extent of deposition (Table 1). Several deposits formed behind thick log jams (Figure S1a), which restricted the runout distance of some flows. At this basin, the total mobilized material consisted almost entirely of eroded and entrained material (~97%) but because runout occurred over long distance in relatively narrow channels, the growth factors is relatively low (Table 1). Soil depth visible in tree-throw pits within the basin varied from 1 to 1.5 m but is as deep as 3 m on lower-gradient alluvial and colluvial deposits. Here, we defined soil depth as a function of contributing area, slope gradient and minimum and maximum depths following Westrick (1999).

3.7.4.2 Observed-runout dataset and model setup

Here we have a 2005, Lidar DEM, recorded after most of the landslides occurred (post-event). Like the Rocky Mountain site, we relied on field observations of erosion and deposition depths to approximate the map of $\Delta\eta$. Many of the deposits were fluvially eroded but the original extent and elevation of the deposit could be identified by terraces of material left along the channel wall. At

this site, the landslides were relatively small and the resulting runout deposits appeared to only add a new layer to pre-existing valley-filling debris flow deposits. We therefore used the 2005 DEM resampled to a 10-meter grid, as the pre-event DEM.

Figure S1. Example woody debris impact on observed runout behavior. (a) Wood jams at the Olympic Mountain site that caused valley-filling deposits. (b) Standing trees resisted thinner parts of the flow at the Cascade Mountain site but were knocked down by the main part of the flow (photo credit: John Jenkins). (c) Woody debris piled along the edge of the flow and controlled lateral spreading at the Rocky Mountain site.

3.8 REFERENCES

1. ArcGIS, https://desktop.arcgis.com/en/arcmap/latest/tools/spatial-analyst-toolbox/filter.htm, accessed 2022
2. Bagnold, R. A. (1954). Experiments on a gravity-free dispersion of large solid spheres in a Newtonian fluid under shear. *Proceedings of the Royal Society of London*, *225*(1160), 49–63. https://doi.org/10.1098/rspa.1954.0186
3. Barnhart, K. R., Hutton, E. W. H., Tucker, G. E., Gasparini, N. M., Istanbulluoglu, E., Hobley, D. E. J., Lyons, N. J., Mouchene, M., Nudurupati, S. S., Adams, J. M., & Bandaragoda, C. (2020). Short communication: Landlab v2.0: a software package for Earth surface dynamics. *Earth Surface Dynamics*, *8*(2), 379–397. https://doi.org/10.5194/esurf-8-379-2020
4. Barnhart, K. R., Jones, R., George, D. J., McArdell, B. W., Rengers, F. K., Staley, D. M., & Kean, J. W. (2021). Multi-Model Comparison of Computed Debris Flow Runout for the 9 January 2018 Montecito, California Post-Wildfire Event. *Journal of Geophysical Research: Earth Surface*, *126*(12). https://doi.org/10.1029/2021jf006245
5. Benda, L., & Cundy, T. W. (1990). Predicting deposition of debris flows in mountain channels. *Canadian Geotechnical Journal*, *27*(4), 409–417. https://doi.org/10.1139/t90-057
6. Benda, L., Veldhuisen, C. P., & Black, J. (2003). Debris flows as agents of morphological heterogeneity at low-order confluences, Olympic Mountains, Washington. *Geological Society of America Bulletin*, *115*(9), 1110. https://doi.org/10.1130/b25265.1
7. Beven, K. (2006). A manifesto for the equifinality thesis. *Journal of Hydrology*, *320*(1–2), 18–36. https://doi.org/10.1016/j.jhydrol.2005.07.007
8. Bigelow, P., Benda, L., Miller, D., & Burnett, K. M. (2007). On Debris Flows, River Networks, and the Spatial Structure of Channel Morphology. *Forest Science*, *53*(2), 220–238. https://doi.org/10.1093/forestscience/53.2.220
9. Campforts, B., Shobe, C. M., Overeem, I., & Tucker, G. E. (2022). The Art of Landslides: How Stochastic Mass Wasting Shapes Topography and Influences Landscape Dynamics. *Journal of Geophysical Research: Earth Surface*, *127*(8). https://doi.org/10.1029/2022jf006745
10. Campforts, B., Shobe, C. M., Steer, P., Vanmaercke, M., Lague, D., & Braun, J. (2020). HyLands 1.0: a hybrid landscape evolution model to simulate the impact of landslides and landslide-derived sediment on landscape evolution. *Geoscientific Model Development*, *13*(9), 3863–3886. https://doi.org/10.5194/gmd-13-3863-2020
11. Capart, H., & Fraccarollo, L. (2011). Transport layer structure in intense bed-load. *Geophysical Research Letters*, *38*(20), n/a. https://doi.org/10.1029/2011gl049408
12. Capart, H., Hung, C., & Stark, C. R. (2015). Depth-integrated equations for entraining granular flows in narrow channels. *Journal of Fluid Mechanics*, *765*. https://doi.org/10.1017/jfm.2014.713
13. Carretier, S., Martinod, P., Reich, M., & Goddéris, Y. (2016). Modelling sediment clasts transport during landscape evolution. *Earth Surface Dynamics*, *4*(1), 237–251. https://doi.org/10.5194/esurf-4-237-2016
14. Chase, C. G. (1992). Fluvial landsculpting and the fractal dimension of topography. *Geomorphology*, *5*(1–2), 39–57. https://doi.org/10.1016/0169-555x(92)90057-u
15. Chen, C., & Yu, F. (2011). Morphometric analysis of debris flows and their source areas using GIS. *Geomorphology*, *129*(3–4), 387–397. https://doi.org/10.1016/j.geomorph.2011.03.002

16. Chen, H., & Zhang, L. (2015). EDDA 1.0: integrated simulation of debris flow erosion, deposition and property changes. *Geoscientific Model Development*, 8(3), 829–844. https://doi.org/10.5194/gmd-8-829-2015
17. Clerici, A., & Perego, S. (2000). Simulation of the Parma River blockage by the Corniglio landslide (Northern Italy). *Geomorphology*, 33(1–2), 1–23. https://doi.org/10.1016/s0169-555x(99)00095-1
18. Codd, E. F. (1968). *Cellular Automata* (1st ed.). New York, Academic Press.
19. Coz, J. L., Renard, B., Bonnifait, L., Branger, F., & Boursicaud, R. L. (2014). Combining hydraulic knowledge and uncertain gaugings in the estimation of hydrometric rating curves: A Bayesian approach. *Journal of Hydrology*, *509*, 573–587. https://doi.org/10.1016/j.jhydrol.2013.11.016
20. Crave, A., & Davy, P. (2001). A stochastic "precipiton" model for simulating erosion/sedimentation dynamics. *Computers & Geosciences*, 27(7), 815–827. https://doi.org/10.1016/s0098-3004(00)00167-9
21. D'Ambrosio, D., Di Gregorio, S., Iovine, G., Lupiano, V., Rongo, R., & Spataro, W. (2003). First simulations of the Sarno debris flows through Cellular Automata modelling. *Geomorphology*, *54*(1–2), 91–117. https://doi.org/10.1016/s0169-555x(03)00058-8
22. Department of Natural Resources, https://www.dnr.wa.gov/programs-and-services/buy-maps-aerial-photos-or-survey-data
23. Egashira, S., Honda, N., & Itoh, T. (2001). Experimental study on the entrainment of bed material into debris flow. *Physics and Chemistry of the Earth, Parts a/B/C*, 26(9), 645–650. https://doi.org/10.1016/s1464-1917(01)00062-9
24. Foglia, L., Hill, M. C., Mehl, S. W., and Burlando, P. (2009), Sensitivity analysis, calibration, and testing of a distributed hydrological model using error-based weighting and one objective function, *Water Resources Research*, 45, W06427, doi:10.1029/2008WR007255.
25. Fannin, R. J., & Wise, M. P. (2001). An empirical-statistical model for debris flow travel distance. *Canadian Geotechnical Journal*, *38*(5), 982–994. https://doi.org/10.1139/t01-030
26. Frank, F., McArdell, B. W., Huggel, C., & Vieli, A. (2015). The importance of entrainment and bulking on debris flow runout modeling: examples from the Swiss Alps. *Natural Hazards and Earth System Sciences*, *15*(11), 2569–2583. https://doi.org/10.5194/nhess-15-2569-2015
27. Gelman, A., Carlin, J. B., Stern, H. S., Dunson, D. B., Vehtari, A., & Rubin, D. B. (2021). *Bayesian Data Analysis* (3rd ed.). Electronic Edition.
28. Goode, J. R., Luce, C. H., & Buffington, J. M. (2012). Enhanced sediment delivery in a changing climate in semi-arid mountain basins: Implications for water resource management and aquatic habitat in the northern Rocky Mountains. *Geomorphology*, 139–140, 1–15. https://doi.org/10.1016/j.geomorph.2011.06.021
29. Google Earth, https://earth.google.com/, accessed 2022
30. Gorr, A., McGuire, L. A., Youberg, A., & Rengers, F. K. (2022). A progressive flow-routing model for rapid assessment of debris-flow inundation. *Landslides*, 19(9), 2055–2073. https://doi.org/10.1007/s10346-022-01890-y
31. Guthrie, R. H., Deadman, P., Cabrera, A. R., & Evans, S. D. (2008). Exploring the magnitude–frequency distribution: a cellular automata model for landslides. *Landslides*, *5*(1), 151–159. https://doi.org/10.1007/s10346-007-0104-1
32. Guthrie, R. H., & Befus, A. D. (2021). DebrisFlow Predictor: an agent-based runout program for shallow landslides. *Natural Hazards and Earth System Sciences*, 21(3), 1029–1049. https://doi.org/10.5194/nhess-21-1029-2021

33. Han, Z., Li, Y., Huang, J., Chen, G., Xu, L., Tang, C. Y., Zhang, H., & Shang, Y. (2017). Numerical simulation for run-out extent of debris flows using an improved cellular automaton model. *Bulletin of Engineering Geology and the Environment*, 76(3), 961–974. https://doi.org/10.1007/s10064-016-0902-6
34. Han, Z., Ma, Y., Li, Y., Zhang, H., Chen, N., Hu, G., & Chen, G. (2021). Hydrodynamic and topography based cellular automaton model for simulating debris flow run-out extent and entrainment behavior. *Water Research*, 193, 116872. https://doi.org/10.1016/j.watres.2021.116872
35. Hobley, D. E. J., Adams, J. M., Nudurupati, S. S., Hutton, E. W. H., Gasparini, N. M., Istanbulluoglu, E., & Tucker, G. E. (2017). Creative computing with Landlab: an open-source toolkit for building, coupling, and exploring two-dimensional numerical models of Earth-surface dynamics. *Earth Surface Dynamics*, *5*(1), 21–46. https://doi.org/10.5194/esurf-5-21-2017
36. Horton, P., Jaboyedoff, M., Rudaz, B., & Zimmermann, M. N. (2013). Flow-R, a model for susceptibility mapping of debris flows and other gravitational hazards at a regional scale. *Natural Hazards and Earth System Sciences*, 13(4), 869–885. https://doi.org/10.5194/nhess-13-869-2013
37. Hungr, O., Morgan, G. J., & Kellerhals, R. (1984). Quantitative analysis of debris torrent hazards for design of remedial measures. *Canadian Geotechnical Journal*, 21(4), 663–677. https://doi.org/10.1139/t84-073
38. Hungr, O., Leroueil, S., & Picarelli, L. (2013). The Varnes classification of landslide types, an update. *Landslides*, 11(2), 167–194. https://doi.org/10.1007/s10346-013-0436-y
39. Hürlimann, M., Rickenmann, D., Medina, V., & Bateman, A. (2008). Evaluation of approaches to calculate debris-flow parameters for hazard assessment. *Engineering Geology*, *102*(3–4), 152–163. https://doi.org/10.1016/j.enggeo.2008.03.012
40. Hutter, K., Svendsen, B., & Rickenmann, D. (1996). Debris flow modeling: A review. *Continuum Mechanics and Thermodynamics*, *8*(1), 1–35. https://doi.org/10.1007/bf01175749
41. Iovine, G., D'Ambrosio, D., & Di Gregorio, S. (2005). Applying genetic algorithms for calibrating a hexagonal cellular automata model for the simulation of debris flows characterised by strong inertial effects. *Geomorphology*, 66(1–4), 287–303. https://doi.org/10.1016/j.geomorph.2004.09.017
42. Istanbulluoglu, E., O. Yetemen, E. R. Vivoni, H. A. Gutie´rrez-Jurado, and R. L. Bras (2008), Eco-geomorphic implications of hillslope aspect: Inferences from analysis of landscape morphology in central New Mexico, *Geophysical. Research Letters*, 35, L14403, 10.1029/ 2008GL034477.
43. Iverson, R. M. (1997). The physics of debris flows. *Reviews of Geophysics*, *35*(3), 245–296. https://doi.org/10.1029/97rg00426
44. Iverson, R. M., & Denlinger, R. P. (2001). Flow of variably fluidized granular masses across three-dimensional terrain: 1. Coulomb mixture theory. *Journal of Geophysical Research*, *106*(B1), 537–552. https://doi.org/10.1029/2000jb900329
45. Iverson, R.M. (2003) How should mathematical models of geomorphic processes be judged?. In Wilcock, P., & Iverson, R. (Eds.), *Prediction in Geomorphology*. American Geophysical Union.
46. Julien, P. Y., & Paris, A. (2010). Mean Velocity of Mudflows and Debris Flows. *Journal of Hydraulic Engineering*, *136*(9), 676–679. https://doi.org/10.1061/(asce)hy.1943-7900.0000224
47. Korup, O. (2006). Effects of large deep-seated landslides on hillslope morphology, western Southern Alps, New Zealand. *Journal of Geophysical Research*, 111(F1). https://doi.org/10.1029/2004jf000242

48. Lancaster, S. T., Hayes, S. K., & Grant, G. E. (2003). Effects of wood on debris flow runout in small mountain watersheds. *Water Resources Research*, 39(6). https://doi.org/10.1029/2001wr001227
49. Larsen, I. J., & Montgomery, D. R. (2012). Landslide erosion coupled to tectonics and river incision. *Nature Geoscience*, 5(7), 468–473. https://doi.org/10.1038/ngeo1479
50. Lee, C., Huang, W., Chang, Y., Chi, S., & Liao, W. (2018). Regional landslide susceptibility assessment using multi-stage remote sensing data along the coastal range highway in northeastern Taiwan. *Geomorphology*, 300, 113–127. https://doi.org/10.1016/j.geomorph.2017.10.019
51. Lee, J. P., & Park, H. (2015). Assessment of shallow landslide susceptibility using the transient infiltration flow model and GIS-based probabilistic approach. *Landslides*, *13*(5), 885–903. https://doi.org/10.1007/s10346-015-0646-6
52. Liu, J., Wu, Y., Gao, X., & Zhang, X. (2022). A Simple Method of Mapping Landslides Runout Zones Considering Kinematic Uncertainties. *Remote Sensing*, 14(3), 668. https://doi.org/10.3390/rs14030668
53. Major, J. J. (1997). Depositional Processes in Large-Scale Debris-Flow Experiments. *The Journal of Geology*, *105*(3), 345–366. https://doi.org/10.1086/515930
54. Major, J. J., & Iverson, R. M. (1999). Debris-flow deposition: Effects of pore-fluid pressure and friction concentrated at flow margins. *Geological Society of America Bulletin*, *111*(10), 1424–1434. https://doi.org/10.1130/0016-7606(1999)111
55. May, C. N., & Gresswell, R. E. (2004). Spatial and temporal patterns of debris-flow deposition in the Oregon Coast Range, USA. *Geomorphology*, 57(3–4), 135–149. https://doi.org/10.1016/s0169-555x(03)00086-2
56. McCoy, S. W., Kean, J. W., Coe, J. A., Tucker, G. S., Staley, D. M., & Wasklewicz, T. A. (2012). Sediment entrainment by debris flows: In situ measurements from the headwaters of a steep catchment. *Journal of Geophysical Research*, *117*(F3), n/a. https://doi.org/10.1029/2011jf002278
57. McDougall, S., & Hungr, O. (2004). A model for the analysis of rapid landslide motion across three-dimensional terrain. *Canadian Geotechnical Journal*, 41(6), 1084–1097. https://doi.org/10.1139/t04-052
58. Medina, V., Hürlimann, M., & Bateman, A. (2008). Application of FLATModel, a 2D finite volume code, to debris flows in the northeastern part of the Iberian Peninsula. *Landslides*, 5(1), 127–142. https://doi.org/10.1007/s10346-007-0102-3
59. Miller, D., & Burnett, K. M. (2008). A probabilistic model of debris-flow delivery to stream channels, demonstrated for the Coast Range of Oregon, USA. *Geomorphology*, *94*(1–2), 184–205. https://doi.org/10.1016/j.geomorph.2007.05.009
60. Montgomery, D. R., & Dietrich, W. E. (1988). Where do channels begin? *Nature*, *336*(6196), 232–234. https://doi.org/10.1038/336232a0
61. Murray, A. S., & Paola, C. (1994). A cellular model of braided rivers. *Nature*, *371*(6492), 54–57. https://doi.org/10.1038/371054a0
62. Murray, B.A. (2003) Contrasting the goals, strategies and predictions associated with simplified numerical models and detailed simulations. In Wilcock, P., & Iverson, R. (Eds.), *Prediction in Geomorphology*. American Geophysical Union.
63. Murray, A. B. (2007). Reducing model complexity for explanation and prediction. *Geomorphology*, *90*(3–4), 178–191. https://doi.org/10.1016/j.geomorph.2006.10.020

64. Murray A.B. (2013) Which Models Are Good (Enough), and When?. In: John F. Shroder (ed.) *Treatise on Geomorphology*, Volume 2, pp. 50-58. San Diego: Academic Press.
65. Natural Resources Conservation Service | Snow and Water Interactive Map (n.d.). Natural Resources Conservation Service. Retrieved April 1, 2022, from
66. Nicoletti, P. G., & Sorriso-Valvo, M. (1991). Geomorphic controls of the shape and mobility of rock avalanches. *Geological Society of America Bulletin*, 103(10), 1365–1373. https://doi.org/10.1130/0016-7606(1991)103
67. Open topography, https://opentopography.org/, accessed 2022
68. Patton, A. I., Rathburn, S. L., Bilderback, E. L., & Lukens, C. E. (2018). Patterns of debris flow initiation and periglacial sediment sourcing in the Colorado Front Range. *Earth Surface Processes and Landforms*, 43(15), 2998–3008. https://doi.org/10.1002/esp.4463
69. Perron, J. T. (2017). Climate and the Pace of Erosional Landscape Evolution. *Annual Review of Earth and Planetary Sciences*, *45*(1), 561–591. https://doi.org/10.1146/annurev-earth-060614-105405
70. Reid, M. J., Coe, J. A., & Brien, D. L. (2016). Forecasting inundation from debris flows that grow volumetrically during travel, with application to the Oregon Coast Range, USA. *Geomorphology*, *273*, 396–411. https://doi.org/10.1016/j.geomorph.2016.07.039
71. Renard, B., Garreta, V., & Lang, M. J. (2006). An application of Bayesian analysis and Markov chain Monte Carlo methods to the estimation of a regional trend in annual maxima. *Water Resources Research*, *42*(12). https://doi.org/10.1029/2005wr004591
72. Roda-Boluda, D. C., D'Arcy, M., McDonald, J., & Whittaker, A. C. (2018). Lithological controls on hillslope sediment supply: insights from landslide activity and grain size distributions. *Earth Surface Processes and Landforms*, 5), 956–977. https://doi.org/10.1002/esp.4281
73. Schürch, P., Densmore, A. L., Rosser, N., & McArdell, B. W. (2011). Dynamic controls on erosion and deposition on debris-flow fans. *Geology*, *39*(9), 827–830. https://doi.org/10.1130/g32103.1
74. Shen, P., Zhang, L. M., Wong, H., Peng, D., Zhou, S., Zhang, S., & Chen, C. (2020). Debris flow enlargement from entrainment: A case study for comparison of three entrainment models. *Engineering Geology*, *270*, 105581. https://doi.org/10.1016/j.enggeo.2020.105581
75. https://www.nrcs.usda.gov/resources/data-and-reports/snow-and-water-interactive-map
76. Soil Survey Geographic Database (SSURGO) *Ag Data Commons*. (n.d.). https://data.nal.usda.gov/dataset/soil-survey-geographic-database-ssurgo, accessed 2020
77. Stock, J. P. J., & Dietrich, W. E. (2006). Erosion of steepland valleys by debris flows. *Geological Society of America Bulletin*, 118(9–10), 1125–1148. https://doi.org/10.1130/b25902.1
78. Strauch, R. L., Istanbulluoglu, E., Nudurupati, S. S., Bandaragoda, C., Gasparini, N. M., & Tucker, G. E. (2017). A hydroclimatological approach to predicting regional landslide probability using Landlab. *Earth Surface Dynamics*, 6(1), 49–75. https://doi.org/10.5194/esurf-6-49-2018
79. Takahashi, T. (1978). Mechanical Characteristics of Debris Flow. *Journal of the Hydraulics Division*, *104*(8), 1153–1169. https://doi.org/10.1061/jyceaj.0005046
80. Takahashi, T. (2014). *Debris Flow* (2nd ed.). CRC Press, Taylor & Francis Group.
81. Tucker, G. E., McCoy, S., & Hobley, D. E. J. (2018). A lattice grain model of hillslope evolution. *Earth Surface Dynamics*, 6(3), 563–582. https://doi.org/10.5194/esurf-6-563-2018

82. USGS Colorado geologic map data, https://mrdata.usgs.gov/geology/state/state.php?state=CO, accessed September, 2022
83. Wallace, C. A., Santi, P. M., & Walton, G. (2022). Scoring system to predict landslide runout in the Pacific Northwest, USA. *Landslides*, 19(6), 1449–1461. https://doi.org/10.1007/s10346-021-01839-7
84. WA DNR geology portal, https://www.dnr.wa.gov/geologyportal, accessed June, 2022
85. Western Regional Climate Center. (n.d.). Retrieved 2017, from https://wrcc.dri.edu/
86. Westrick, K. (1999). Soil depth calculation script implemented in python. Retrieved from https://github.com/pnnl/DHSVM-PNNL
87. Wu, W., & Sidle, R. C. (1995). A Distributed Slope Stability Model for Steep Forested Basins. *Water Resources Research*, 31(8), 2097–2110. https://doi.org/10.1029/95wr01136
88. Zhou, G. G. D., Li, S., Song, D., Choi, C. E., & Chen, X. (2019). Depositional mechanisms and morphology of debris flow: physical modelling. *Landslides*, 16(2), 315–332. https://doi.org/10.1007/s10346-018-1095-9

CHAPTER 4. LANDSLIDE RESPONSE TO CLIMATE CHANGE AND A RETREATING SNOW ZONE IN A MOUNTAINOUS WATERSHED

4.0 ABSTRACT

In this study, we model the hydrologic and slope stability impacts of climate change in the North Cascades, a mountainous region where local climatology, ecosystems and terrain are highly variable. To evaluate the potential range of climate change impacts on hydrology and slope stability, we coupled an existing hydrology model (DHSVM) with an existing landslide model (LandslideProbability). We forced the coupled models with three climate scenarios that capture the variability of predicted precipitation and temperature from 20 global climate models selected for the Pacific Northwest. The three models represent a hot-dry, warm-dry and hot-wet scenario relative to 17 other possible future climate scenarios. We assessed results in terms of three historic precipitation zones: rain-dominated, transient and snow-dominated. To couple the landslide and the hydrology models, we developed the DistributedHydrologyGenerator (DHG), written in Python and implemented as a component for the package Landlab. DHG takes the raw flow and gridded hydrology outputs from a distributed hydrology model, and downscales and converts it to the Landlab grid. Using this coupled modeling approach, we found that in all three climate scenarios, by the end of the 21st century, the potential landslide area in the rain-dominated and transient snow zones increases and in the snow-zone decreases. Relative to the warm-dry scenario, the hot-dry scenario, which has nearly equivalent precipitation as the warm-dry scenario but is hotter, causes more potential landslides in the rain-dominated and transient zones because the warmer temperatures result in a larger shift from spring to winter peak runoff in the snow zone. Finally, as droughts become more prevalent, stand replacing fires may destroy tree roots and

reduce total soil cohesion. This is expected to elevate landslide probability in all regions of the North Cascades; however, the increase will be largest for the lower regions of the historically snow-dominated zones where slopes are likely most dependent on root cohesion for stability.

4.1 INTRODUCTION

Climate change is expected to impact the entire planet but the localized effects will vary (IPCC, 2022). In many steep mountainous regions such as the Taiwan Alps or the Olympic Mountains, USA, where local climatology, ecosystems, and landscapes vary tremendously across an elevation gradient, the impacts of climate change could have far reaching consequences on biotic and abiotic processes, that co-evolve and are regulated by hydrology (Kuo et al. 2022; Wershow and DeChaine, 2018). In the North Cascade mountains of the United States, winter and spring snow pack are expected to decrease (Frans et al., 2018) and the frequency and intensity of heavy rainfall events are expected to increase. Extreme 24-h rainfall totals are predicted to increase by 19% on average under high greenhouse gas scenarios. As a consequence of higher rainfall totals, winter soil water content and landslide frequency are expected to increase (Mauger et al., 2020).

Landslide are often triggered by elevated soil pore water pressure, which reduces the frictional resistance of the soil (Kuriakose et al. 2008). However, multiple hydrologic processes control pore water pressures and several studies indicate that the hydrologic trigger for increased landslides may be more complex than a simple increase in rainfall rates. For example, the change in precipitation type from snow to rain might expose slopes that were previously conditioned to snow-melt runoff to precipitation runoff, both of which could have different consequences on subsurface flows and the resulting soil pore-water pressure. Multiple studies in the northwest have recorded a high number of landslide events occurring during rain-on-snow events (Harr 1981; Guthrie et al. 2010). Guthrie et al (2010) also reported the contribution of strong winds to landslide

occurrence as a consequence of wind-driven snow melt and wind-concentrated rainfall disturbance. Rain-on-snow floods are predicted to increase in western north America (Musselman et al (2018). At the same time, a warming climate can increase evapotranspirtation rates, which may reduce soil antecedent wetness (Gariano and Guzetti, 2016). These findings raise the following competing questions: Could an increase in temperature raise evapotranspiration rates such that the net change in soil water content is negligible? Or, could an increase in temperature without an increase in precipitation cause an increase in landslide rates simply because it causes precipitation to fall as rain on an increasing and over-steepened area or because it falls as a rain-on-snow event instead of as a snow event?

To decipher the potential impact of these hydrological processes on landslides in the North Cascades, we investigate landslide response to climate change using a coupled, distributed hydrology and landslide model representation of the North Cascades forced with future climate data (Figure 1). In much of the North Cascade mountains, precipitation falls as both rain and snow and rates vary as a consequence of westerly flowing atmospheric rivers and consequent orographic effects as they pass the mountains. The Cascade Mountains consist of a wide range of forest types as well as variable underlying geology. To model landslide response to climate change in this environment, we couple the Distributed Hydrology Soil and Vegetation Model (DHSVM; Wigmosta et al. 1994), forced with gridded meteorology data that capture the spatial variability of precipitation in the region, to a probabilistic slope stability model in Landlab, an earth surface modeling toolkit (Barnhart et al. 2020; Hobley et al. 2017). The coupled models are forced with four climate scenarios obtained from the University of Washington Climate Impacts Group (CIG, 2020): (1) historic climate between 1960 and 2010; (2) a hot-dry scenario that is considerably warmer than historic conditions but only moderately wetter than historic conditions; (3) a warm-

dry scenario that is only moderately warmer and wetter than historic conditions and (4) a hot-wet scenario that is both much warmer and wetter than historic conditions. Under these four climate scenarios, the two questions we posed can be formulated into three expected outcomes.

1. Precipitation control: The hot-wet scenario will trigger more landslides than the hot-dry scenario, as precipitation totals will be higher in the hot-wet scenario while temperature will be comparable in both scenarios.

2. Water balance control: Compared to the warm-dry scenario, the amount of increase in landslides in the hot-dry scenario will depend on the offsetting effects of higher evapotranspiration. We also expect the hot-wet scenario to have more landslides than warm-dry scenario, because the hot-wet scenario precipitation is higher than the warm-dry precipitation; however, high evapotranspiration rates driven by high temperature in the hot-wet scenario may offset some of that difference.

3 . Temperature control: We expect the warm-dry scenario to have more landslides than the hot-dry scenario because the warm-dry and hot-dry precipitation are comparable but hot-dry temperature is much higher, which might increase evapotranspiration rates and reduce soil water relative to the warm-dry scenario. On the other hand, rapid snowmelt can trigger more landslides, so it's possible enhanced snow melt in the hot-dry scenario will cause more landslides than the warm-dry scenario.

Additionally, forest fire intensity and frequency varies as a function of drought and fuel availability (Littell et al., 2016). In the future, summers are likely to become warmer and drier (Mauger et al., 2015) and fuels are presently at a historic high due to past forest management and fire suppression. As a consequence of warmer, dryer summers and increased fuels, large, intense stand-replacing fires may occur. Stand-replacing fires kill most or all trees which in turn reduces

the total cohesion (roots+soil cohesion) of the soil. Therefore, in addition to climate impacts, we also examine the relative impact of fire on landslide area by adjusting total cohesion to reflect the loss of roots following a stand replacing fire similar to Istanbulluoglu et al. (2005). We expect the impact of fire to be large given that cohesion values can drop by as much as 70% once root strength is removed.

In the following sections, we first present the study region, spatial data used to represent the study region and mapped landslides used to evaluate landslide model results. Second, we present DHSVM and LandslideProbability, a landslide model in Landlab, and describe the DistributedHydrologyGenerator, which is a new tool coded in Python for the package Landlab, designed to couple DSHVM modeled hydrology outputs with existing models in Landlab, like LandslideProbability. Following the description of the models, we then present landslide sensitivity to changes in temperature and precipitation as predicted by the three future climate scenarios.

4.2 STUDY REGION

The model domain spans over 3,300 km^2 of the upper Skagit River watershed in the North Cascade mountains within Washington State (Figure 1). The higher elevation, eastern region of the model domain consists of high-grade metamorphic rocks (Tabor et al. 2002). The western region of the mountains are lower grade metamorphic and sedimentary rocks (Beechie et al. 2001). The climate is maritime and the largest precipitation events occur during atmospheric river events mostly in the fall and some winter months (Neiman et al., 2011). In the model domain, annual accumulated precipitation varies from as high as 450 cm near glacier peak to 80 cm in the western lowlands (Beechie et al. 2001). The model domain contains a range of elevations and precipitation zones including rain-dominated, transient and snow-dominated zones (Elsner et al., 2009). Peak flow

rates often coincide with the largest precipitation events but also occur during the spring snow melt season.

Much of the surface morphology and regolith have been sculpted by past alpine and continental glaciations (Beechie et al., 2001), leaving behind a complex landscape of variable regolith thickness that is often composed of glacial drift and local colluvium. Treeline begins at an elevation of roughly 1,500 meters, above which, vegetation is limited to isolated pockets of conifers and regolith largely consists of a heterogeneous layer of thick glacial deposits, colluvial deposits of cobbles and boulders (debris fans and talus) or the thin weathered surface of bare rock (NRCS, 2023). Snowflow grooves and erosional processes (e.g., Blikra and Nemec, 1998) in the glacial and colluvial deposits are common. Below treeline, the landscape is covered by a continuous conifer dominated forest and regolith includes a wide range of soils but is predominantly colluvium derived stony, sandy loams intermixed with stony outcrops and thick valley fills. Soil depths are generally at least 1.5 meters in depth but there are numerous rock outcrops, where there is no soil (NRCS, 2023). Despite spatially variable precipitation rates, regolith cover and vegetation zones, much of the landscape becomes unstable during heavy rainfall. Heavy rainfall activated a multiple deep seated landslides in 2003 (Strouth et al., 2006) and active landslide scars are visible along channel walls and zero-order hollows throughout the landscape.

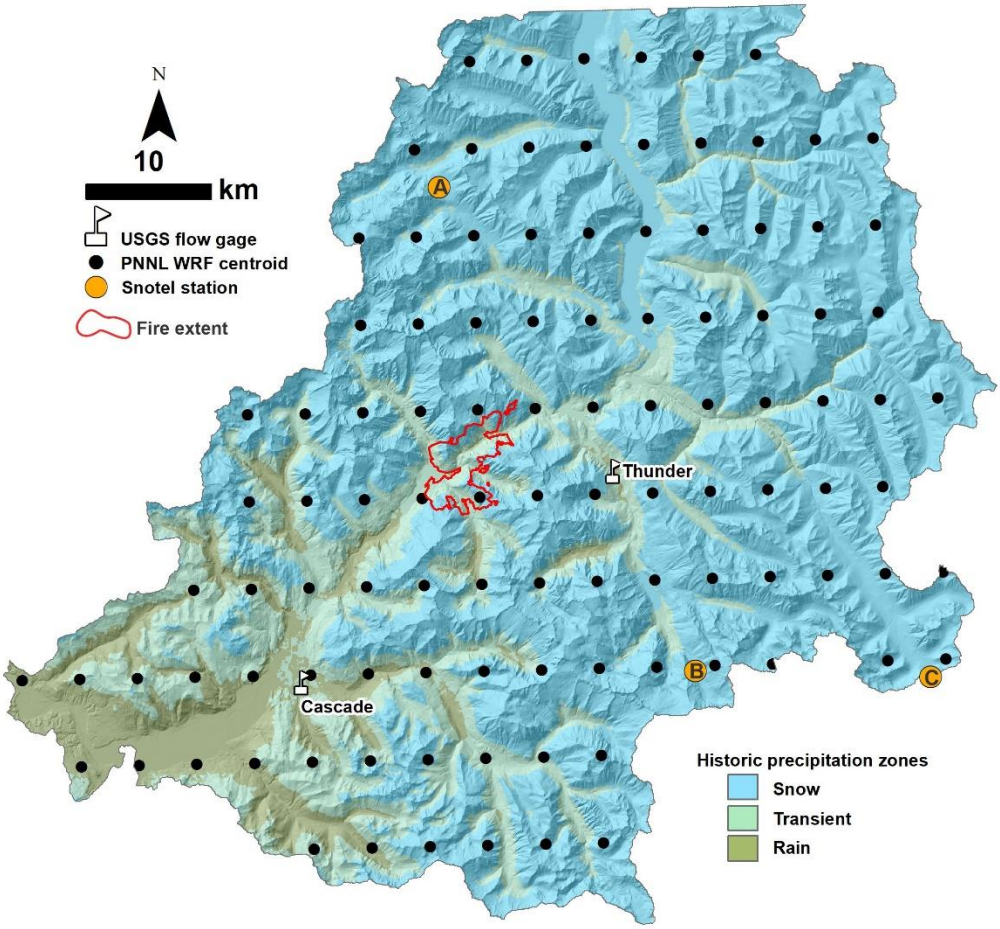

Figure 1. Shaded DEM of the model domain, which extends over most of the Skagit River watershed. DEM is colored according to historic precipitation zones, determined as a function of annual maximum SWE and cumulative cold season precipitation. Black dots indicate centroids of historic and projected climate used to force DHSVM. DHSVM model calibration is compared to the two flow gages and three snow gages (A - Beaver Pass; B - Thunder Creek; C - Rainy pass).

4.3 METHODS

4.3.1 Landslide mapping

To characterize landslide processes in the model domain and create a dataset that could be used to evaluate model performance, we manually mapped all active landslides (see definition of "active" below) visible in imagery of the model domain. Numerous landslide mapping standards have been

developed for the Pacific Northwest region (e.g. Burns et al., 2012; Slaughter et al. 2017). For the purpose of model evaluation, only the area and location of the landslide are needed but we collected basic data about each slide including the landslide type, confidence of the mapping and year of movement as inferred from the imagery date. For simplicity, landslide types were divided into three types Rock Fall (RF), Rock Avalanche (RA) and Earth Avalanche (EA). Landslides above tree line often consisted of rock topple from bluffs or exposed bedrock surfaces and were generally classified as being sourced in rock. Landslides below tree line often occurred in colluvium and were generally classified as sourced in earth. Any landslide that created an actual colluvial fan was assigned as an avalanche. Landslides that resulted in smaller deposits were categorized as falls. No distinction in speed or movement type (e.g., translation vs rotational) was made. Relict deep-seated slides and inactive topographic hollows were not mapped.

Landslide mapping was completed using repeat satellite imagery draped over a 30 meter DEM in Google Earth (Google Earth, 2023). The image years that were reliably clear enough to make out details like boulders, vegetation and scarp edges included 2013 and 2016 images. Other images years clear enough to make out changes in vegetation and bare areas included the 1998 and 2006 images. We toggled between the different image years and view orientations to interpret vegetation and colluvial/alluvial deposits. Since we are only modeling the initiation of the landslide, we only mapped the source area of the landslide (scar or inferred scar location). Features of the landslide runout path, such as areas of scour or deposition were excluded. Active landslides were interpreted as:

1. A landslide scar that formed or enlarged between photos
2. A scar that could be traced to an active colluvial fan, alluvial fan or talus deposit

Active colluvial fans and alluvial fans were identified as un-vegetated debris deposits at the base of a clear runout path in the vegetation or a low order channel. Talus deposits were identified as deposits below hillslopes or bluffs, unassociated with channels or clear runout paths. Many of the talus deposits were located above treeline. Active talus deposits were identified by fresh deposits, which often had a different color and arrangement of boulders, appeared only in later imagery or appeared as rockfall on top of the snow covering the talus deposit.

To identify the source location of the material on the active fan or talus deposit, we followed the clearest, de-vegetated path along the upslope channel or hillslope, as inferred from a continuous, channelized break in vegetation to the most active looking (fresh bedrock, denuded surface) landslide scar or topographic hollow. If the runout path bifurcated, we followed each to its source. If the source area of a runout path could not be confidently identified from the patterns in vegetation or the rock surface, the most likely area was selected based on topography and upslope position relative to the deposit (convergent areas directly upslope of the most active part of the deposits.

4.3.2 Hydrology model

We model surface and subsurface hydrology in a DHSVM representation of the north Cascade Mountains described by Bandaragoda et al. (2019). DHSVM models saturated and unsaturated soil water movement, and hillslope and channelized surface runoff and importantly, snow dynamics and vegetation interception and evapotranspiration (Wigmosta et al., 1994; Wigmosta & Perkins, 2001). Bandaragoda et al. (2019) used the model to evaluate climate impacts on floods and sediment dynamics in the Skagit River watershed. That same model was later used by the Climate Impacts Group to evaluate future water resources in the Skagit basin (CIG, 2020).

The Bandaragoda et al. (2019) DHSVM setup uses a 150-m DEM grid. Hydrologic properties of the soils were assigned based on soil classes mapped in the Soil Survey Geographic Database (SSURGO, 2019). Soil depth was defined as a function of observed minimum and maximum soil depths, contributing area and slope (Westerick, 1999). Vegetation cover was defined based on National Land Cover Dataset 2011 vegetation classes (NLCD, 2019). Additional details on the model setup and calibration can be found in Bandaragoda et al. (2019).

Calibration is an important component of any hydrologic model and evaluation of model calibration is necessary before inferences from model results can be made. For the North Cascade DHSVM model, Bandaragoda et al. (2019) reported calibration relative to observed flow at four basin locations, in terms of Daily Nash-Sutcliffe Efficiency values (Nash & Sutcliffe, 1970).

Since our study focus is the Skagit watershed, we further examined model calibration, specifically at two unregulated (no upstream dams) streamflow gages in the Cascade River and Thunder Creek watersheds (USGS, 2022) and three SNOTEL sites in the Skagit Watershed (see Figure 1). At all three SNOTEL sites, the first period of the SWE record, ranging from 5 to 10 years, was systematically lower than modeled SWE. We excluded these initial periods of systematically offset SWE from model evaluation.

4.3.3 Climate data

This study uses historical (Livneh et al., 2013) and future climate scenarios (Figure 2) from the Skagit Water Supply and Demand project (CIG, 2020) to force DHSVM. The CIG (2020) climate scenarios include 10 global climate models (GCM), each with two different Representative Concentration Pathways (RCP), a high RCP scenario (RCP 8.5) and a low RCP scenario (RCP 4.5). The 10 models represent the top-10 most representative models of the Pacific Northwest in terms of ability to replicate observed climate in the Pacific Northwest based on a sample of 41

models evaluated in Rupp et al. (2013). Each model was downscaled from the GCM model grid resolution to the finer grid resolution used in DHSVM using a multivariate adaptive constructed analog (MACA; see CIG, 2020 and Abatsoglou & Brown, 2012).

Jakob et al. (2012) found that landslide-sourced debris flows were correlated to 4-week antecedent and 2-day cumulative precipitation. At the same time, warming climate may increase ET rates, and offset precipitation impacts on slope stability (Goriano and Guzetti, 2016). Furthermore, Guthrie et al. (2010) reported strong winds contributing to landslide occurrence as a consequence of wind-driven snow melt and wind-concentrated rainfall. From the 20 future climate scenarios (two RCP versions of each of the 10 models), we refined the number of models used to model landslides to just three models that capture the variability of predicted precipitation and temperature from the 20 global climate models selected for the Pacific Northwest. We characterize each climate scenario in terms of the water-year- (Oct. 1 to Sept. 30) maximum 2-day cumulative precipitation, the water-year-maximum, monthly-cumulative precipitation and the water-year-maximum 2-day average temperature and wind. Together, the three climate scenarios represent: (1) an extreme increase in temperature and modest increase in precipitation scenario (hot, dry); (2) a modest increase in precipitation and temperature scenario (warm, dry) and; (3) an extreme increase in both temperature and precipitation scenario (hot, wet). In terms of wind, almost all models vary around the historic mean wind and are comparable (2c) and we do not distinguish models by wind. By refining the models to these three extreme scenarios, we can directly compare and contrast the most extreme predicted temperature and precipitation impacts on future landslide rates.

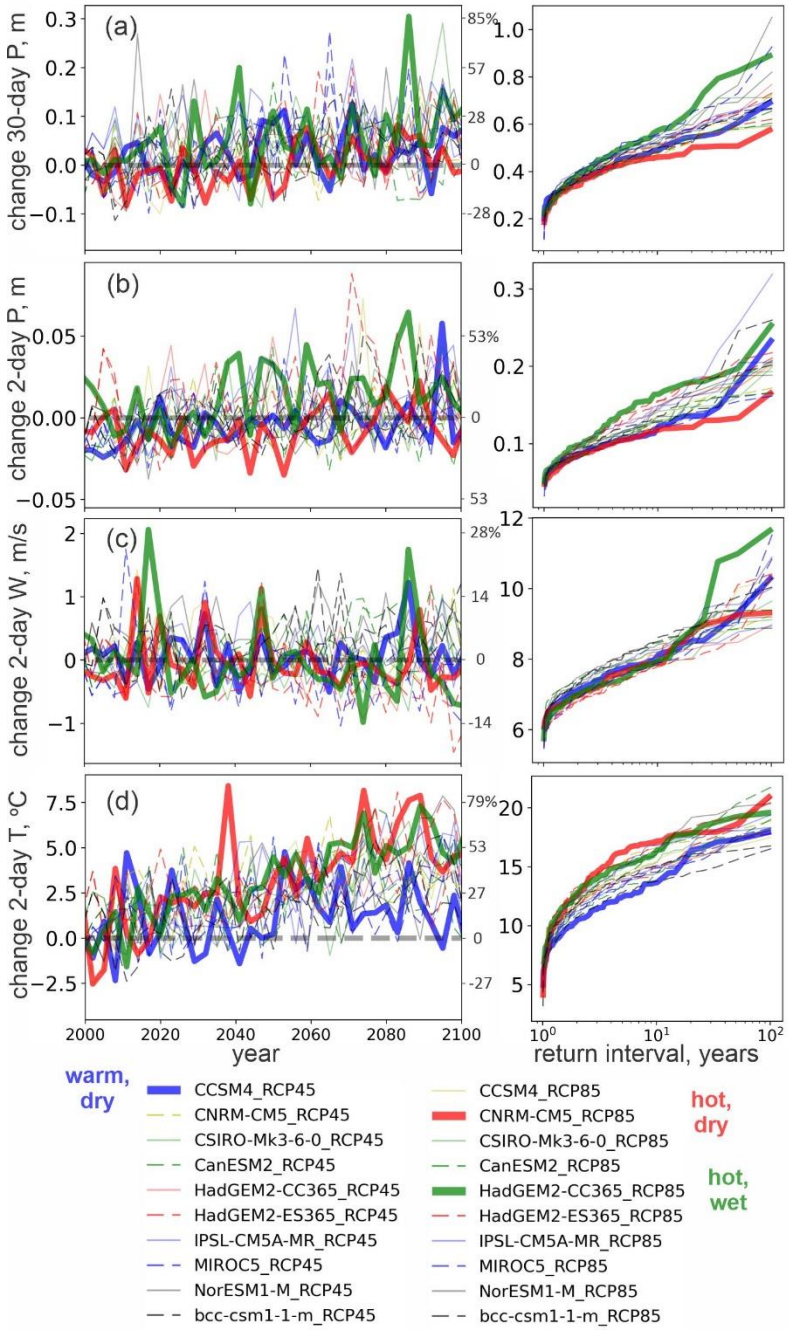

Figure 2. Time series of water-year maximum (left panel) and probability of exceedance, expressed as return interval (right panel) for (a) maximum 30-day accumulated precipitation and (b) 2-accumulated precipitation, (c) 2-day mean wind, and (d) 2-day mean temperature for the 20 climate scenarios considered in this study. The three scenarios selected for this study are highlighted in green, blue and red (CCSM4_RCP45, CNRM-CM5_RCP85 and HadGEM2-CC365_RCP85 respectively). Note, plots (a), (b), (c) and (d) show the differences from historic average (dashed gray line), using a 3-year average, to aid visualization of trend in data

4.3.4 Precipitation zones.

Most of the findings in this study are presented in terms of historic precipitation zones that represent the predominance of precipitation occurring as rain, snow or mixed-rain-and-snow. We determined the historic precipitation zones similar to the method described by Elsner et al. (2009). At each grid cell in the landslide model, we computed the ratio of the peak SWE (SWE at grid cell when basin average SWE was highest, SWE_{max}) to the cumulative cold-season (October to March) precipitation (P_c) at the grid cell:

$$SP = \frac{SWE_{max}}{P_c} \quad (1)$$

Where SP is the ratio of SWE_{max} to P_c, both of which were determined for each 150 m grid cell in DHSVM and then linearly interpolated to the 30-m landslide model domain. We defined the rain-dominated zone (R) as all cells as $SP < 0.1$, the transient zone (T) as $0.1 \leq SP < 0.5$ and the snow zone (S) as $SP \geq 0.5$ using the historic climate. The precipitation zones represent historic conditions (we do not evolve them with time) and are shown in Figure 1.

4.3.5 Landslide model

This study uses the landslide model LandslideProbability (Strauch et al., 2018) to predict changes in hillslope stability in response to climate-driven-changes in soil hydrology. LandslideProbability is a probabilistic implementation of the infinite slope model described by Montgomery et al (1998, 2000), Pack et al. (1998) and Wu and Sidle (1994). The Factor-of-safety (FS), or the ratio of the resisting forces to the driving forces is expressed as:

$$FS = \frac{C/h_s\rho_s g}{\sin\theta} + \frac{\cos\theta \tan\phi(1-R_w\rho_w/\rho_s)}{\sin\theta} \quad (2)$$

Where C is soil cohesion, h_s is soil depth [m], ρ_s is soil density [kg/m3], θ is slope of the slip-plane (assumed equal to the surface slope) ϕ is the soil friction angle [degrees], R_w is the relative wetness (ratio of saturated thickness to the soil thickness), ρ_w is the density of water. LandslideProbability determines relative wetness from a user provided recharge rate (Precipitation – losses) as:

$$R_w = \frac{h_w}{h_s} = min\left(\frac{Ra}{T \sin\theta}, 1\right) \quad (3)$$

Where R is recharge rate, a is contributing area and T is transmissivity (soil depth multiplied by saturated hydraulic conductivity). For each grid cell in the model domain, LandslideProbability samples from a user defined probability distribution function (PDF) of each parameter in (2), specific to the grid cell, to determine FS. After N model iterations, the probability of failure or probability of a factor-of-safety less than 1 ($P(FS \leq 1)$) is defined as:

$$P(FS \leq 1) = \frac{n(FS \leq 1)}{N} \quad (4)$$

Where $n(FS \leq 1)$ is the number of times modeled factor-of-safety is less than one.

To define PDFs of C at each node, we used land cover types in the National Land Cover Database (NLCD, 2011), grouped into eight broad classes: forest, shrubs, herbaceous, wetland, developed, barren, ice/snow, and water. We defined triangular PDFs of C for the forest, shrub, herbaceous and developed categories using the minimum, mode and maximum cohesion values (Figure 3b) from Strauch et al. (2018). We excluded the wetland, water and ice/snow categories from the model domain. Since many of the barren areas in the model domain consist of exposed, unweathered bedrock and mass wasting processes in these areas often consisted of topple like failures (see Section 3.1), rather than rock or debris slides, we exclude the barren areas from the model domain as well.

We defined triangular PDFs of ϕ at each node based on the SSURGO soil class (described in Section 3.2) assuming a minimum and maximum value equal ~ +/-30% of the median value. We defined triangular PDFs of soil depth at each node using the soil depth map for DHSVM (see Section 3.2) to represent the mode and minimum and maximum values equal to 70% and 110% of the mode, again following Strauch et al. (2018). Soil depths vary from 1 meter on the hillslopes to as deep as 2.8 meters in the valleys (Figure 3d). Note that while not as prevalent as the barren cover type, large, bedrock outcrops can be found within the shrub, forest and herbaceous land cover categories as well. Therefore, computed landslide probabilities within the Skagit watershed should be interpreted as a potential probability, analogous to sediment transport capacity. As in sediment transport, landslide potential depends on material availability. The potential landslide probability of a grid cell in which soil depth is assumed to be 1 meter may be high, but in reality, the landslide probability may be very low if there is no soil.

LandslideProbability includes a number of options for defining the probability distribution of R, however in our application, we force LandslideProbability with saturated thickness (thickness from impermeable bedrock surface to the phreatic surface), modeled from DHSVM. To run LandslideProbability using saturated thickness, we modified LandslideProbability. This updated version of LandslideProbability can be forced with saturated thickness using two methods: (1) a time series of saturated thickness values at each node or; (2) randomly sampled values from a lognormal PDF of saturated thickness, whose shape is parameterized from the mean and standard deviation of a time series of saturated thickness at each node. All results in this study use option 2. We used DHSVM to create a water-year maximum time series of saturated thickness at each node. To convert the coarse-scale DHSVM maps of saturated thickness to the landslide model grid

resolution, we developed and used a new utility for Landlab, called DistributedHydrologyGenerator, described below.

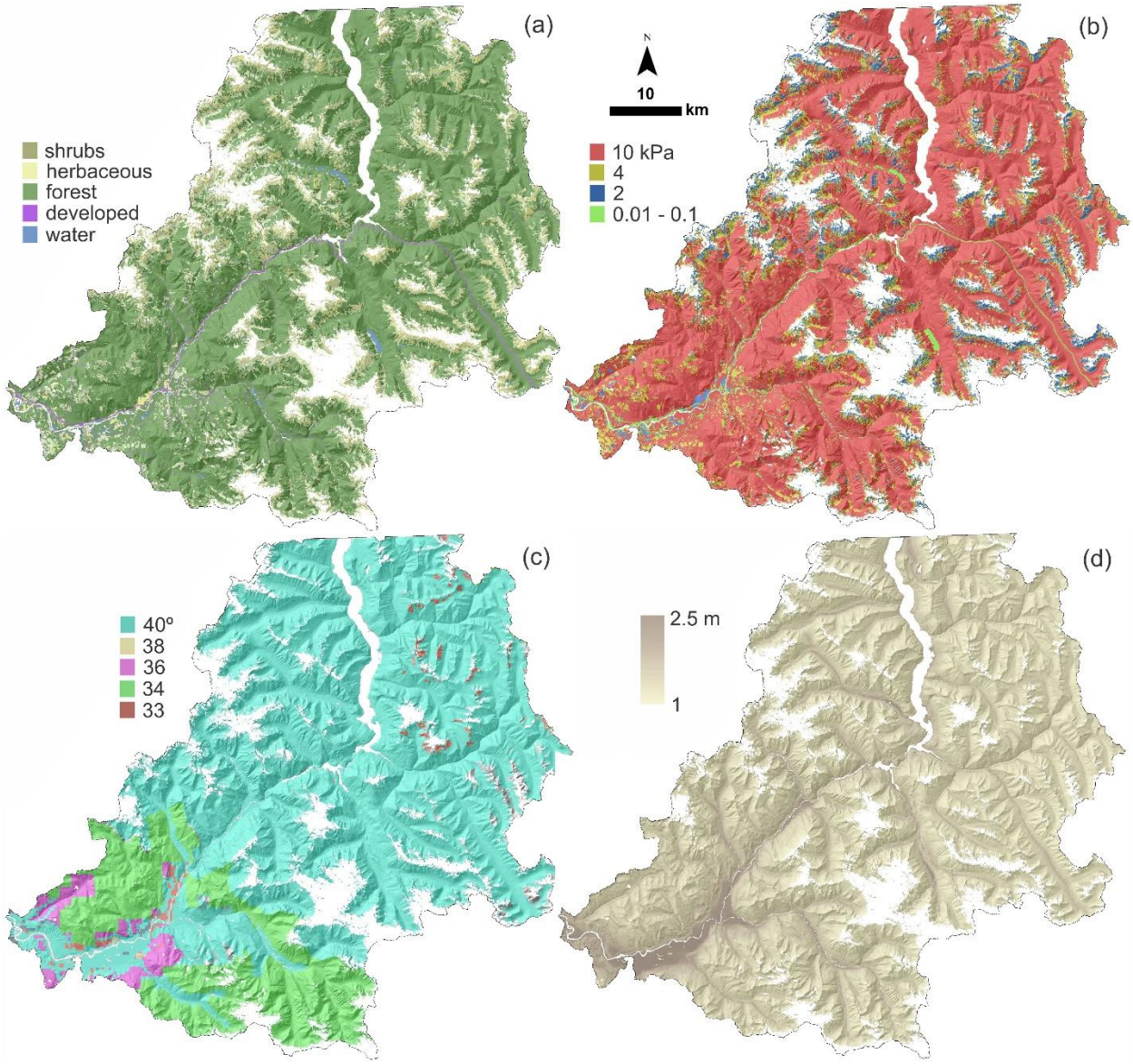

Figure 3 (a) Landuse and cover type map; (b) total (soil+roots) cohesion; (c) soil friction angle; (d) soil depth. Sources for each map listed in text. The landuse and covertype map was used to estimate soil cohesion values in (b). Soil friction angle values in (c) are estimated from soil type. Soil depth in (d) was determined as a function of contributing area and slope.

4.3.6 Coupling modeled hydrology and landslides: DistributedHydrologyGenerator

Water controls the topographic evolution of many landscapes. Therefore, prediction of landscape geomorphic response to climate, land use and land cover change requires realistic representations of landscape hydrologic response. While LandslideProbability includes its own hydrology model, it excludes processes like unsaturated flow, exfiltration and surface runoff that may impact the timing and spatial variability of actual soil water conditions. In fact, many models that evaluate watershed-scale geomorphic response to climate or land surface changes often lack relevant representations of hydrology (e.g., Murphy et al. 2019; Schmitt et al, 2016; Tucker and Slingerland, 1997). Of the models that do accurately represent hydrology, the landslide model is often are hard-coded with the hydrology models (e.g., Burton and Bathurst, 1998; Doten et al., 2006; Arnone et al., 2011). While such models are excellent research tools to study a known landslide event and its hydrologic drivers (e.g., Lepore et al., 2013), because of their rigid design, they are less suitable for rapid regional landslide risk assessment applications or broader studies that might include geologic factors other than hydrology.

Instead of recreating hydrologic routines inside of geomorphic models, we propose that existing hydrologic model outputs be ingested into the geomorphic models. For that purpose, we developd DHG, a new tool to link distributed hydrology models, such as DHSVM (Wigmosta et al., 1994) with the Landlab earth surface modeling toolkit (Hobley et al., 2017; Barnhart et al., 2020).

DHG is written in Python, implemented as a component of the package Landlab (i.e., a class in the Landlab python package) and is structured to match the standard Landlab initialization, computation and run-one step format. In order to bridge the typically coarse-gridded output of a hydrology model with finer-gridded geomorphic models, we developed an external utility that

downscales coarse scale modeled inputs to the finer-scale Landlab grid. The downscaling utility is run before running DHG (described in next section). DHG can be operated in three modes: Stochastic Storm Generator mode, Predefined Time Series mode or Basic Statistics mode.

The structure and data inputs for each mode are described in Figure 4. In Predefined Time Series mode, the Landlab models are forced with the raw hydrologic time series produced by distributed hydrology model. In Stochastic Storm Generator mode, a Landlab model can be forced with a time series of soil recharge and flow conditions that is much longer than, but informed (parameterized) by the raw distributed hydrologic time series. In the basic statistics mode, DHG simply populates the raster model grid with the mean and standard deviation saturated thickness and flow conditions from the raw hydrologic time series.

DHG can be run simultaneously on both the Landlab raster model and Landlab network model grids for coupled hillslope and channel process models or it can be run individually on only one of the grids. DHG is instantiated by providing a time series of modeled flow and/or a time series of modeled depth to water table outputs from a distributed hydrology model, Landlab model grids for the model domain and user selected parameters that control how the model is run (Figure 4). The time series of modeled flow depth can be provided in ascii form or as an Xarray dataset (Hoyer & Hamman, 2017).

4.3.7 Downscaling utility

The downscaling utility serves the sole purpose of downscaling the gridded hydrologic outputs (e.g., depth to water table or SWE) from a distributed hydrology model to the Landlab model grid scale. For large watersheds, like ours, distributed hydrology models are often implemented at 100 to 150 m grid cell widths to save computation time and memory. For stream flow modeling, these coarse DEMs may have little impact on modeled flow, especially for flow at the outlet of large

watersheds; however, modeled soil water and hillslope stability modeling are very sensitive to DEM grid size and quality. Following Burton and Bathurst (1998) we resample the coarse gridded phreatic surface to the finer scale Landlab DEM using a topographic index:

$$z_i - \bar{z} = \frac{\bar{I} - I_i}{f} \quad (5)$$

Where z_i is the depth to the water table for the finer DEM and $\bar{z}$ is the mean depth to the water table for the coarse DEM. The variables $\bar{I}$ and I_i are the topographic index for the coarse and fine resolution DEMs respectively. Topographic index, defined using the TOPMODEL wetness index (Beven and Kirby 1979) is:

$$I = \ln\left(\frac{a}{s}\right) \quad (6)$$

Where a is the specific contributing area (contributing area per unit contour width) and s is the local tangent slope [m/m]. The downscaling utility uses (5) to interpolate DHSVM modeled saturated thickness to the Landlab grid scale.

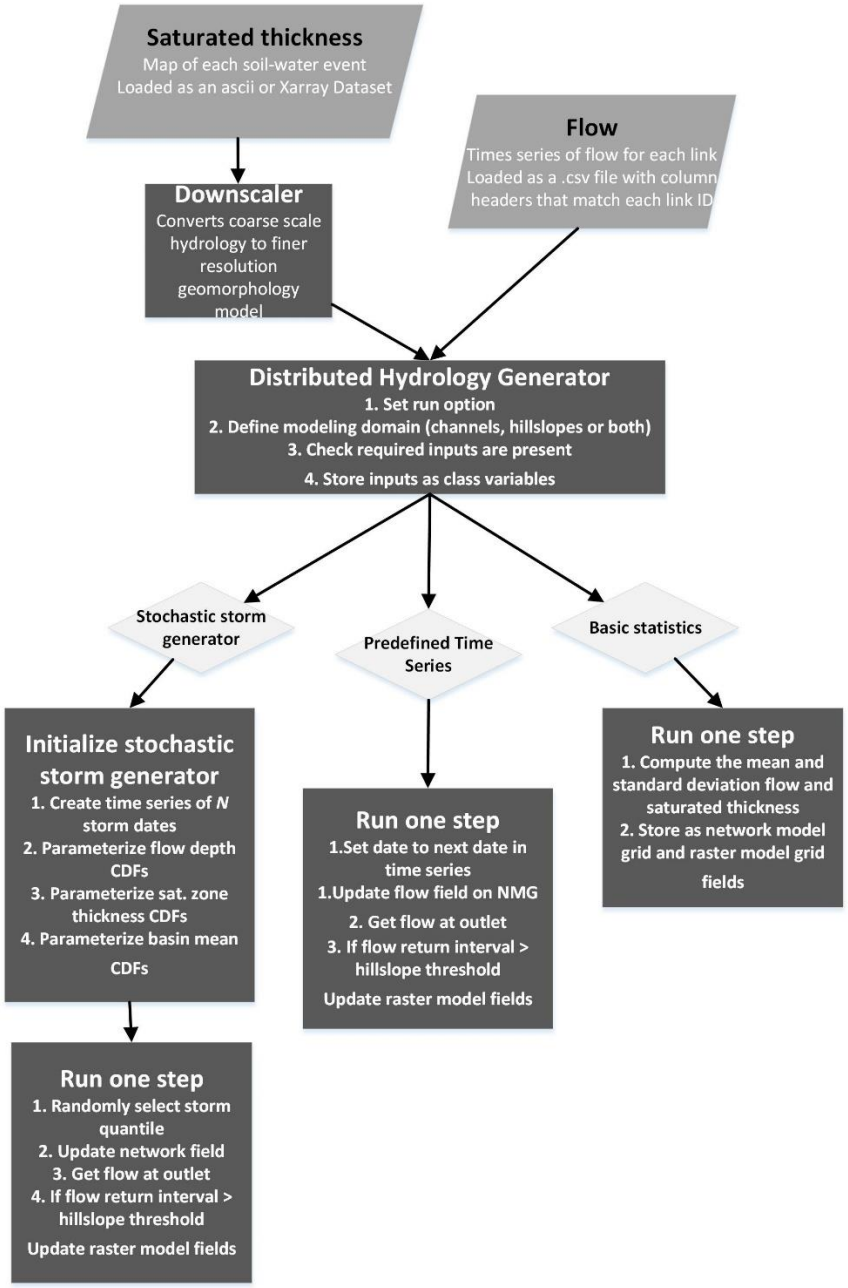

Figure 4. DistributedHydrologyGenerator (DHG) code structure. DHG has three run options. Saturated thickness is first downscaled and clipped to the Landlab model domain using the Downscaling utility.

4.3.8 DHG and LandslideProbability Implementation

Future uncertainty in R_w and FS is represented by variability in the climate forcings and their impact on soil water. We run DHG using the Basic Statistics option by providing a time series of water-year-maximum saturated thickness maps. These maps are then used to parameterize a lognormal distribution of saturated thickness at each grid cell in the watershed area. Because the Skagit watershed consists of both rain-dominated and snow-dominated precipitation zones, but the largest precipitation events occur as atmospheric river events during the winter and fall, the timing of maximum saturation varies with location and no single time (map) captures the maximum saturation at every cell. Historically, the maximum wetness in the rain-dominated zone tends to coincide with the atmospheric rivers, while the maximum wetness in the snow-zone tends to occur in the late spring to early summer months when precipitation from the atmospheric rivers melt. To create the water-year-maximum saturated thickness map, we produced seasonal maximum saturated thickness maps for the winter and spring seasons of each water year and then for each grid cell, took the maximum of the two values as the water year maximum.

$$S_{max} = \max(S_w, S_S) \quad (7)$$

Where S_{max} is the water-year maximum value, S_w is the winter maximum value and S_S is the spring season maximum value. We produced a water-year maximum saturated thickness map in this way for each water year in the DHSVM model run.

4.4 RESULTS

4.4.1 Hydrology calibration

Modeled flow and SWE relative to observations in the Skagit watershed are shown in Figure 5. At the three SNOTEL sites, modeled snow generally captures the timing of the observed begin and

end of snow accumulation (Figures 5a, 5b and 5c). The root mean square error (RMSE) of the three gages ranges from 8.4 to 18.3 cm. Some of the model ability to match observed SWE may be the results of a slight cold bias in the meteorology data used to force DHSVM. Comparison of the forcing dataset to observed mean daily average minimum and maximum temperature shows that they were generally 1 to 2 $°C$ cooler than the observed temperatures.

Daily mean flow NS values (NS_d) at Thunder Creek is 0.49 and monthly NS (NS_m) is 0.62 (Figure 5d) and At Cascade River, NS_d is 0.62 and NS_m is 0.74 (Figure 5e). At both the Thunder Creek and Cascade River gages, the modeled timing and magnitude of most daily average flow rates, as represented by box plots of all flow values in each month, generally matches the observed; however, the model generally underpredicts extreme high flow rates, as shown by flow duration curves inset into Figures 5d and 5e. We consider this level of calibration acceptable for landslide modeling.

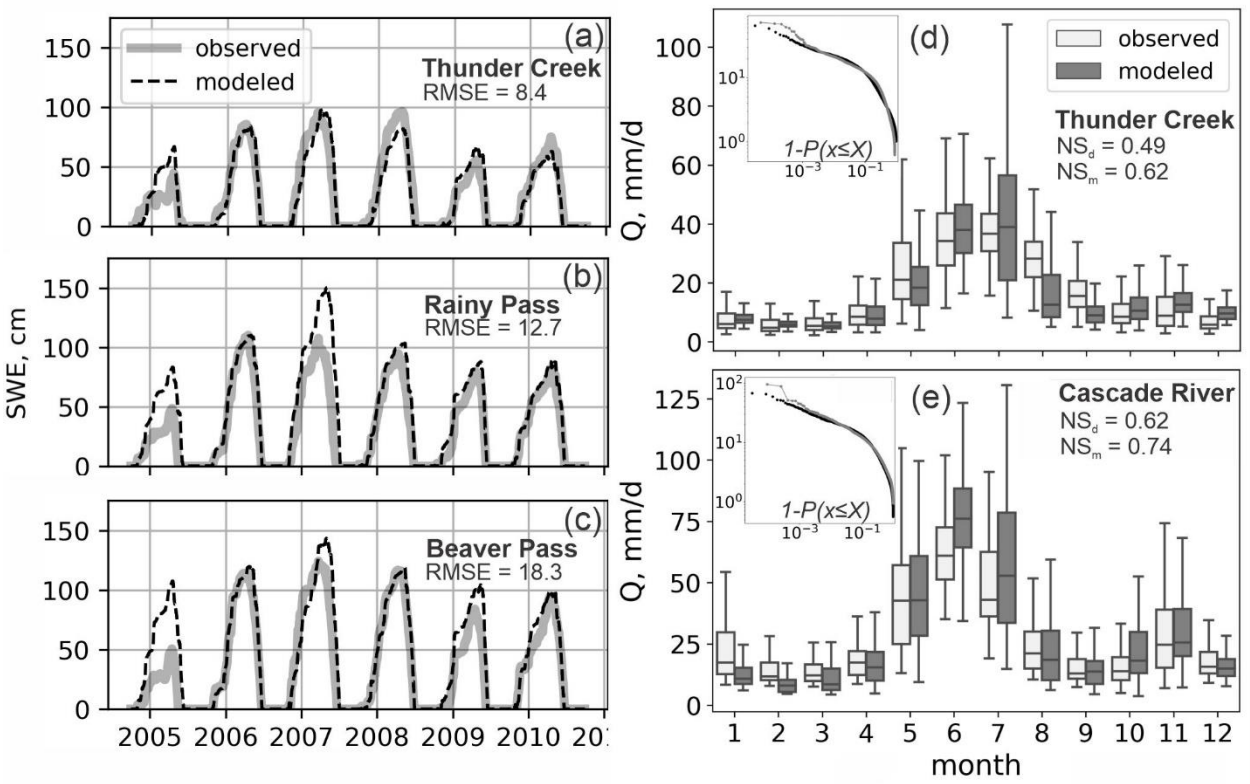

Figure 5. Hydrology calibration results. (a), (b) and (c) show daily SWE at the Thunder Creek, Rainy Pass and Beaver Pass SNOTEL stations and modeled RMSE. (d) and (e) show box plots of modeled and observed daily cumulative flow, normalized by basin area [mm/d], for each month and the daily and monthly NS values. Inset plots in (d) and (e) are flow-duration curves of daily cumulative flow.

4.4.2 Landslide Calibration

An example of the mapped landslides is shown in Figure 6a. Figure 6b shows empirically determined probability density of landslide size for the mapped landslides, determined as the count in each bin divided by the area of the bin (i.e., count times bin width). The number of bins used to create the empirical probability density function was selected based on sample size (n) as log(n)/log(2), as described following Iman and Conover (1983). The mapped landslides areas range between 25 to 120,000 m^2 (Figure 6b). In the model representation of the terrain, each landslide becomes a single cell and because the grid size is 30m, the smallest landslide area is 900

m^2. Figure 6c shows the slope-area characteristics of the terrain and the landslide cells based on this 30-m grid size representation of the terrain. The upslope contributing area at which landslides occur is generally smaller for steeper slides and the average slope of landslide cells is higher than the landscape average. The binned-average slope of landslide cells with the smallest contributing areas is greater than 100% while the average slope of the landscape at the smallest contributing area is closer to 70%. The transition from alluvial to colluvial slopes (e.g., Campforts et al., 2022) appears to take place at a contributing area of roughly 400,000 m^2, as manifested by a change to a steeper slope-area trend of the binned-average slope data.

LandslideProbability model results were evaluated against the mapped landslides using the area under the curve (AUC) value of the receiver operating characteristics (ROC) method (e.g., Anagnostopouls et al., 2015). The AUC value of the modeled landslide locations is 0.76 (Figure 6d), which suggests that there is a 76% chance that as parameterized, LandslideProbability can correctly identify a landslide cell out of two randomly sampled grid cells. Note, to assess model calibration, we excluded comparison with mapped landslides in the wetland, water, ice/snow and barren cover types. Considering that we implemented LandslideProbability without any calibration, and relied solely on the parameter ranges described in Strauch et al. (2018), we consider the model's AUC score adequate.

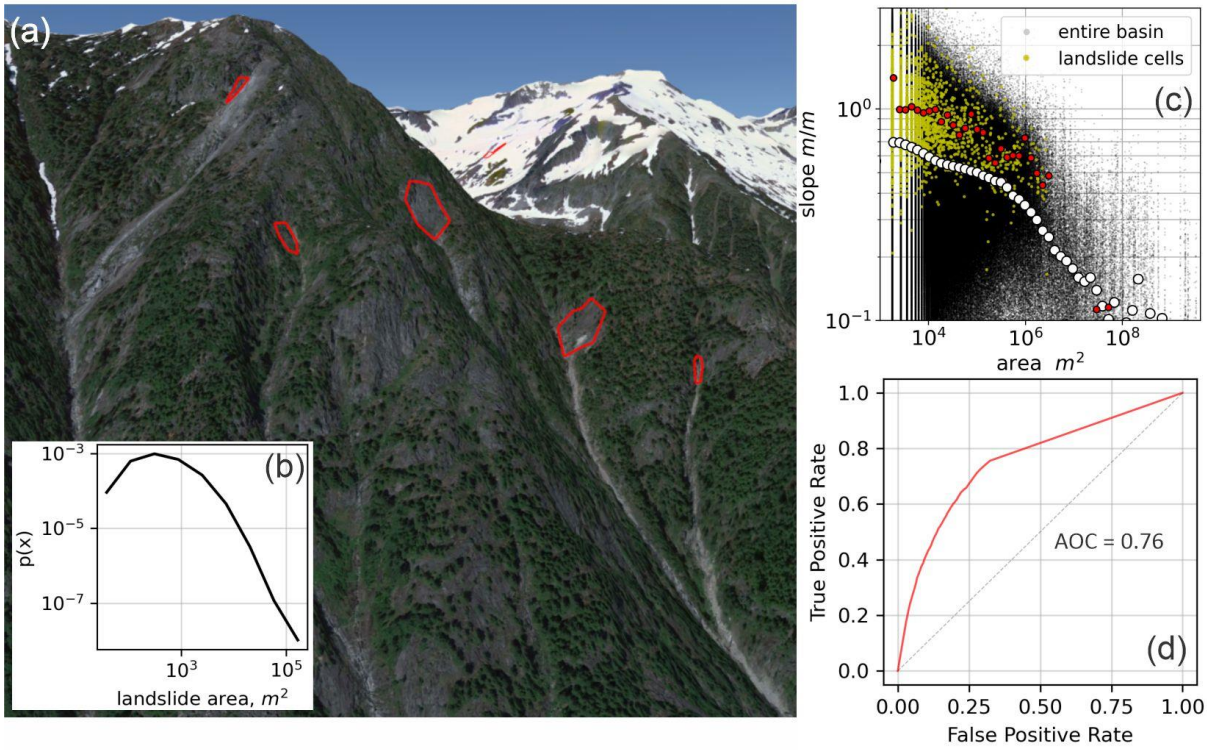

Figure 6. (a) Example of active landslides mapped from Google Earth imagery. Active landslides include both topographic evidence of landslide processes as well a clear runout path. (b) probability density plot of landslide area determined from the active landslides. (c) Slope-Area plot of the basin and the mapped landslide locations, red and white circles show binned and average slope with respect to area of landslide locations, and all landscape data, respectively. (d) LandslideProbability calibration to mapped landslides, as indicated by a ROC plot.

4.4.3 Landslide probability based on historic conditions

Landslide probability over the model domain for historic climatology-driven, modeled hydrology is shown in Figure 7a. Over 10% of the snow-dominated region (excluding barren, snow/ice zones) of the Skagit watershed has a 50% probability of failing any given year (Figure 7b). In contrast, only 4 and 2% of the transient and rain dominated regions have a 50% probability of failure any given year. The cause of the spatial variability in FS values can be inferred from histograms of the key model parameters that control slope stability (Figures 7c, 7d, 7e, 7f). Each precipitation zone consists of a comparable range of soil friction angles though a higher number of relatively low

cohesion values are found in the snow zone. Soil thickness is also thinnest in the snow zone (Figure 7f) but slope is highest: the median slope value is close to 70% (35°) in the snow zone, 60% (31°) in the transient zone and 50% (27°) in the rain zone (Figure 7c). FS (Equation 1) positively varies with cohesion and inversely varies with slope. The relatively steep and low-cohesion soils in the snow zone cause the relatively high landslide probabilities.

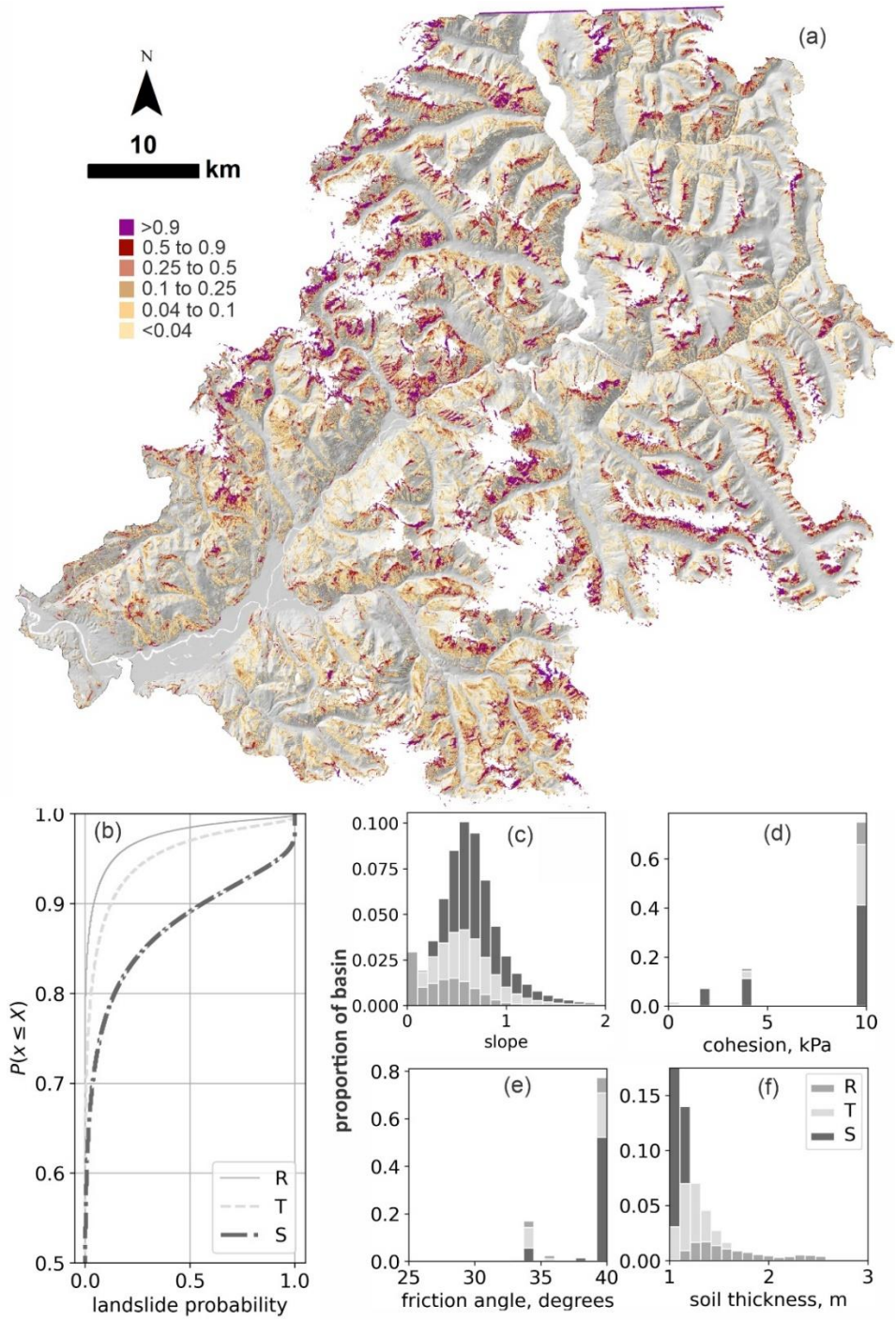

Figure 7. (a) Historic annual (water year) probability of failure and properties within the three precipitation zones: Rain (R), Transition (T), and Snow (S). (b) cumulative distribution function of probability of failure. (c), (d), (e) and (f) show histograms of the area in each slope, cohesion, friction angle respectively, normalized by the basin area. Note that (d) and (e) are stacked histograms and in (f), the y-axis is scaled to show the distribution in each precipitation zone and the first bin of the snow dominated soil thickness extends to 0.3.

4.4.4 Climate change and change in landslide potential for the three climate scenarios

By the end of the 21st century, water year maximum 30-day cumulative precipitation is higher than historic conditions for all model runs in all precipitation zones (Figure 8a). In terms of the 2-day cumulative precipitation, by the end of the 21st century, the hot-wet scenario experiences 20% higher precipitation totals than historic conditions, while the warm-dry and hot-dry scenarios are only 1 to 2 percent above historic conditions in the snow zone and close to the historic average in the rain-dominated and transient zones (Figure 8b). There is little change in wind speed (Figure 8c). The change in temperature is largest for the hot-wet and hot-dry scenarios while the warm-dry scenario is only 2 degrees warmer than average over the whole 21st century (Figure 8d). Based on differences in predicted precipitation amounts, we anticipate the hot-wet scenario to result in the biggest increase in landslide probability.

In our landslide model, relative wetness is the only dynamic variable that responds to climate change. These climate signals equate to modest increases in relative wetness in the rain-dominated and transient zones and despite the elevated precipitation totals, a decrease in relative wetness in the snow zone for all three climate scenarios (Figure 8e). However, the timing and extent of saturated area does change. We explore when the maximum basin saturation is observed and how it changes in the future with respect to historical conditions (Figure 9). Historically, peak saturation in the basin occurs in the spring, around day 275 of the water year, which can be attributed to snowmelt (Figure 9a). Water-year maximum saturation area tends to decrease by the end of the century for the hot-dry and warm-dry runs (Figure 9b and 9c; see color bar in Figure 9a for color definition), while for the hot-wet runs it tends to increase (Figure 9d). However, in all three climate runs, there is a clear trend in early snowmelt-driven saturation, and emergence of late fall and early

winter peak saturation patterns (days 20-150 of the water year), roughly by the middle of the 21st century (Figure 9b, 9c and 9d).

In order to compare modeled landslide area between climate scenarios, we define all cells where $P(FS < 1) > 0.5$ as a potential landslide. The change in relative wetness causes the area of potential landslides to increase in the rain-dominated and transient precipitation zones and decrease in the snow-dominated zone (Figure 8f). Summed over the entire model domain, the total potential landslide area actually drops for all three climate runs because the decrease in potential landslide area in the snow-zone (see Figure 1) exceeds the increase in area in the rain and transient zones (Figure 8g).

Spatially, all three climate scenarios cause a similar change in landslide probability. Landslide probability drops in the high elevation regions of the model domain, particularly in the eastern peaks and ridges, and increases along the river valley walls, as shown for the hot-wet scenario in Figure 10a. As expected, the spatial response of landslide probability to the climate scenarios mirrors the spatial change in relative wetness (Figure 10b).

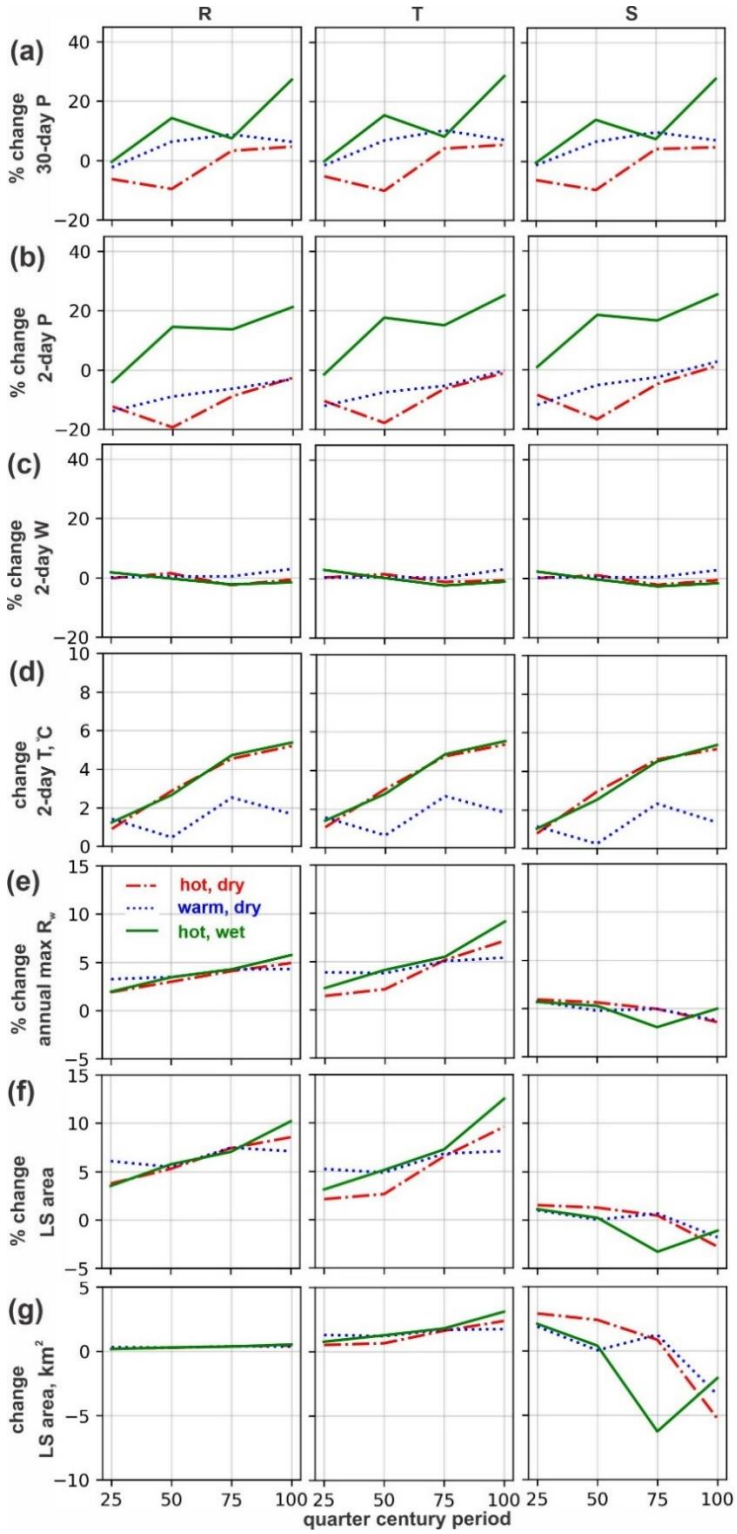

Figure 8. Quarterly mean average changes of water year maximums, for each of the climate scenarios, relative to historic averages in (a) and (b) cumulative 3-day and 2-day precipitation depths; (c) 2-day wind speed; (d) 2-day temperature; (e) relative wetness and (f) potential landslide area. (g) shows change in potential landslide area in km2.

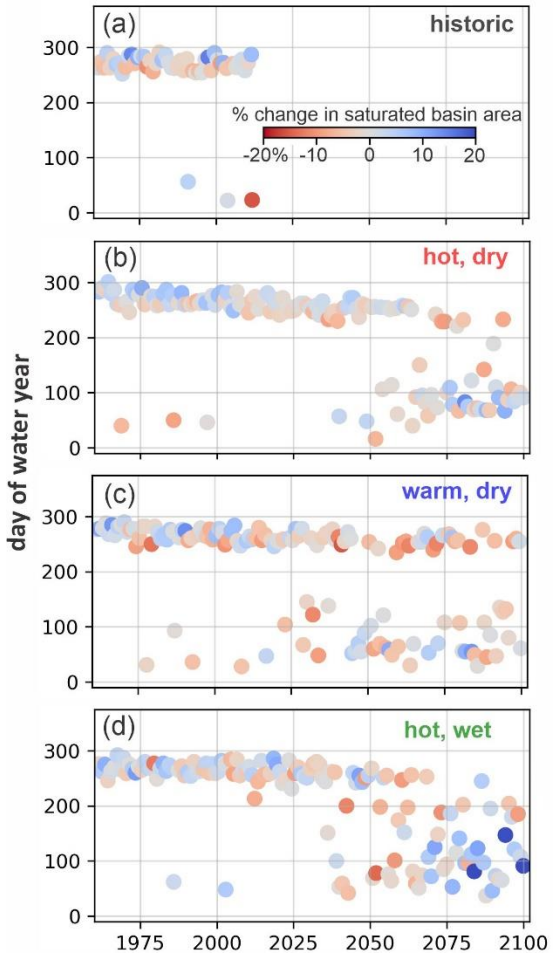

Figure 9. Peak saturation day of the water year, colored according to percent change in saturated basin area relative to the historic average for (a) historic; (b) hot-dry; (c) warm-dryand; (d) hot-wet model runs.

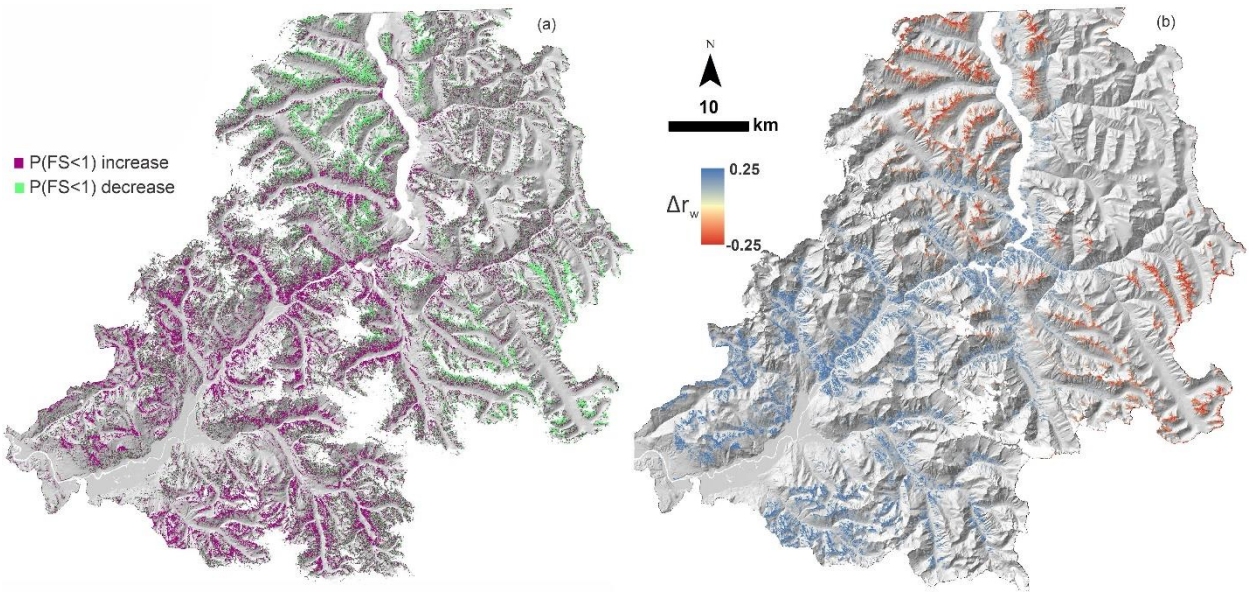

Figure 10. (a) Change direction (increase vs decrease) of Landslide probability ($P(FS < 1)$), relative to historic conditions, for the hot, wet climate scenario during the last quarter of the 21st century. (b) Difference between the historic and predicted average annual maximum relative wetness (R_w) for the last quarter of the 21st century (2076 to 2100) for the hot-wetter model run.

4.5 Discussion

4.5.1 What controls the change in landslide probability

At the beginning of this study, we anticipated that landslide rates would be highest for the hot-wet scenario because the average, water year cumulative 30-day accumulated and 2-day cumulative precipitation of this scenario is considerably higher than the other model runs, especially by the end of the 21st century (Figure 2 and Figure 8a, 8b). Not surprisingly, this expectation was correct (Figure 8g); however, as shown in Figure 10, most of the increase in landslide probability is in the rain-dominated and transient zones (see Figure 1 for precipitation zone location); landslide probability in most of the snow-dominated region of the basin decreases.

We also expected the warm-dry scenario to have more landslides than the hot-dry scenario because relative to the warm-dry scenario, the higher temperatures associated with the hot-dry scenario would drive up evapotranspiration rates, lower peak SWE (and resulting snowmelt rate),

altogether lower the relative wetness which in turn would lead to fewer potential landslides. This expectation was only partly correct. First, by the end of the 21st century, as noted in the results and shown in Figure 8g, for all climate scenarios, summed over the entire model domain, landslide area decreases relative to historic conditions. Second, the warm-dry scenario does have more landslides than the hot-dry scenario, but that's only because the reduction in total landslide area is small relative to the hot-dry scenario. Third, in the rain-dominated and transient zones, the hot-dry scenario in fact has more landslides, rather than less, relative to the warm-dry scenario. However, the mechanism leading to different potential landslide area in the lowland precipitation zones for the hot-dry scenario relative to the warm-dry scenario is not caused by differences in lowland evapotranspiration rates but rather changes in runoff processes in the snow-zone as described below.

In DHSVM, evapotranspiration (ET) positively varies as a function of temperature, leaf area index (LAI) and inversely varies with vapor pressure (Wigmosta et al., 1994). If as we anticipated, the hot-dry scenario caused large ET relative to the warm-dry scenario, at the very least one might expect to see lower relative wetness values compared to historic in the forested areas of the watershed (see Figure 3s for map of forest extent) where LAI is high; however, as shown in a map of the difference between the hot-dry (HD) and warm-dry (WD) average water year maximum relative wetness at the end of the century ($HD\ R_w$ - $WD\ R_w$; Figure 11a), the forested areas are actually wetter in the hot-dry scenario, not drier, but only for part of the model domain.

Plotting basin average water year cumulative ET for each climate scenario (Figure 11b) shows that by the end of the 21st century, the hot-dry scenario ET rates are at most 3 to 4 % greater than historic and the warm-dry scenario cumulative ET are roughly equal to historic. It turns out the difference in ET between the hot-dry and warm-dry scenarios are small. In contrast, changes in

water year cumulative snow water equivalent (SWE) are large (Figure 11c). By the end of the 21st century, basin average water year maximum SWE is roughly 25% lower than historic conditions for the hot-wet and hot-dry scenarios and roughly 16% lower for the warm-dry scenarios. By the end of the 21st century, the hot-dry scenario has considerably less SWE than the warm-dry scenario.

Lower SWE levels equate to a thinner and less extensive snow-pack. It's important to note that the low SWE levels in the hot-dry run relative to the warm-dry run are not due to less precipitation; water year cumulative precipitation in the hot-dry is actually slightly higher than the wet-dry scenario (Figure 11d) and by other metrics, they are nearly equal (Figures 2a, 2b). Rather, the difference in SWE is due to differences in temperature (Figure 2d, 8d) and a change in precipitation from snow to rain.

Precipitation that falls as rain is more likely to enter the soil at the time of the precipitation event and the timing of runoff from the snow-dominated zone changes and coincides with the rainfall triggered peak saturated thickness in the downslope transient and rain dominated zones. This in fact happens for all three future climate runs, but the shift in timing is most significant for the hot-wet and hot-dry scenarios (Figure 12b and 12d). Therefore, the increase in potential landslide area in the rain-dominated and transient zones for the hot-dry scenario, relative to the warm-dry scenario, is not due to differences in evapotranspiration rates, but rather is due to a change in the timing of peak upslope runoff. By the end of the 21st century, for the hot-dry scenario, a large portion of runoff from the snow-zone region, that historically occurred in spring, shifts to the winter and coincides with peak relative wetness in the rain-dominated and transient zones, which drives up relative wetness in those zones (Figure 11a) and causes more landslides relative to the warm-dry scenario (Figure 8g).

Finally, comparing the season at which the peak saturation occurs during the last quarter of the 21st century for the hot-wet scenario reveals that in the snow zone, where the peak saturation date shifts from spring to winter (earlier in the water year, Figure 12d), landslide rates are generally higher (compare to Figure 10). And where the peak saturation date remains in the spring, landslide rates drop.

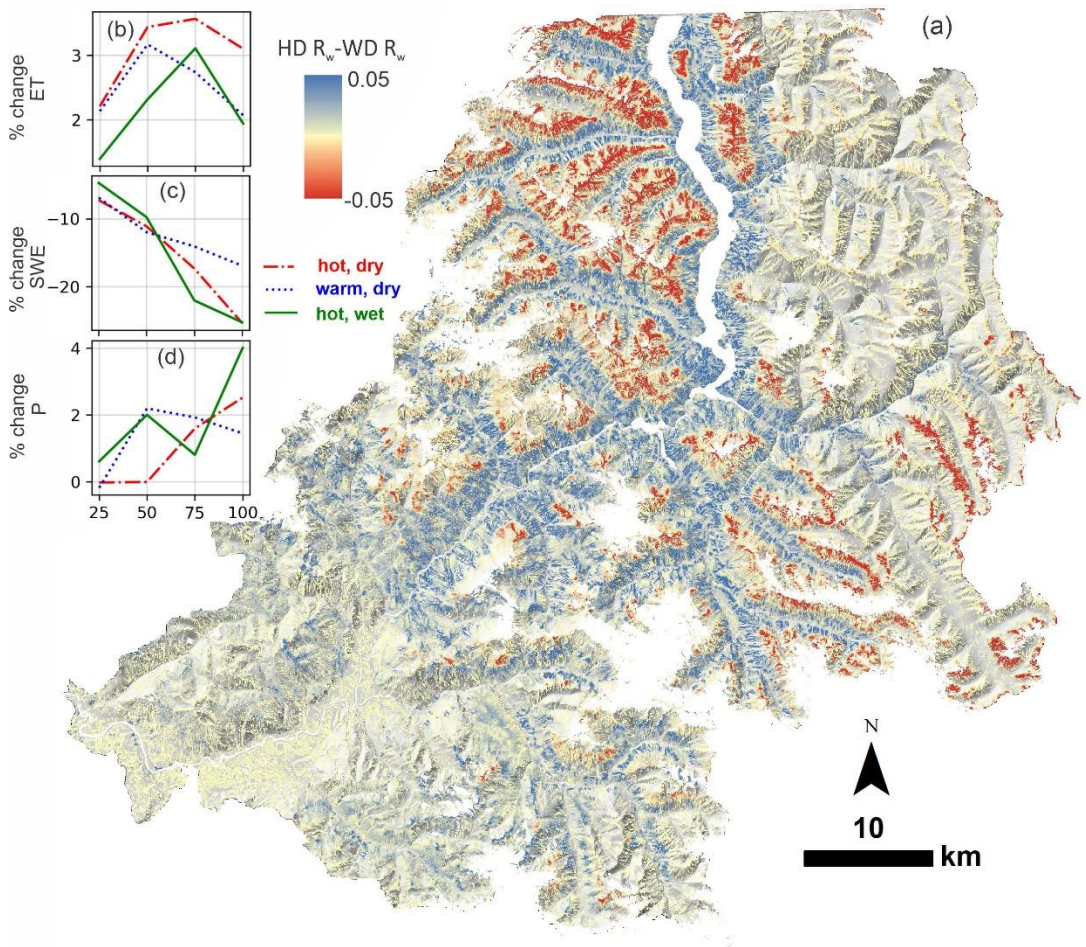

Figure 11 (a) Difference in average-water-year-maximum relative wetness (R_w) between the hot-dry (HD) scenario and the warm-dry (WD) scenario during the last quarter of the 21st century. The hot-dry scenario is generally drier than the warm-dry scenario at the top of the mountains and wetter at foot of the mountains on hillslopes that extend into the snow zone. (b) Percent difference in basin average water year maximum ET rate relative to historic conditions, (c) percent difference in water year maximum SWE relative to historic conditions and (d) the percent change in water year maximum cumulative precipitation.

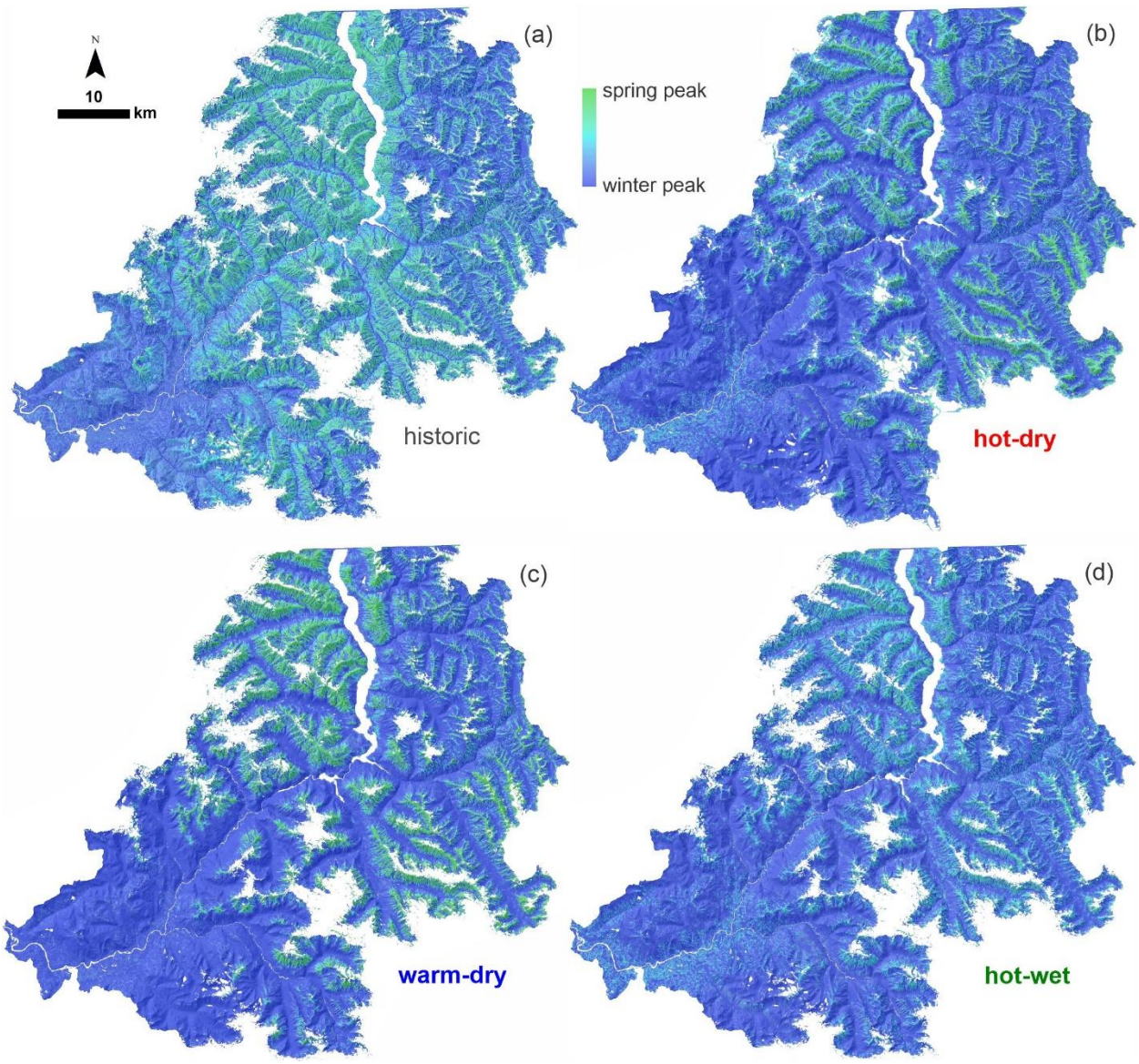

Figure 12. Average season that peak saturation at each grid cell occurs for (a) historic conditions and (b) the hot-dry scenario; (c) the warm-dry scenario and (d) the hot-wet scenario during the last quarter of the 21st century. Green indicates that peak season saturation occurs mostly in the spring. Blue indicates that peak season saturation occurs in the winter. Note how the season changes in the historically snow-dominated region of the watershed.

4.5.2 How do landslide rates change if a large fire occurs during the 21st century

As noted in the introduction, future fire frequency and intensity may increase as a consequence of drier, warmer summers and high fuel loads. Stand replacing fires kill both the trees and their roots which in turn causes the total cohesion of the soil to drop (e.g, Istanbulluoglu & Bras, 2005). To compare the relative impact of fire on potential landslide area we plotted the percent change in potential landslide area as a function of reduced total cohesion assuming the hot-wet climate scenario at the end of the 21st century (Figure 13b). This combination of both end-of-century hot-wet climate and a reduction in cohesion by stand replacing fire represents a worst-case scenario.

Following a wildfire, root cohesion can decline up to 70% from the initial value (e.g., Istanbulluoglu and Bras, 2005). These changes in cohesion are large relative to the change in relative wetness caused by climate change (5 to 12 %, Figure 8e) and consequently, the change in landslide probability is dramatic, particularly in the Cascade River watershed (upstream of Cascade gage shown in Figure 1) in the south region of the model domain (Figure 13a). Interestingly, the spatial patterns of landslide probability change differ from the spatial patterns caused by increased relative wetness alone (Figure 10). The largest change in landslide probability caused by fire occurs in the lower regions of the snow-dominated zone. The snow dominated zone is generally steeper than the other zones (see Figure 7c) and the lower regions of the snow-dominated zone are forested (compare Figures 1 and 4). The stability of the low-elevation snow-zone slopes is likely dependent on cohesion (e.g. the slope exceeds the friction angle of the soil). Therefore, once cohesion is lost in these areas, landslide probability rapidly increases.

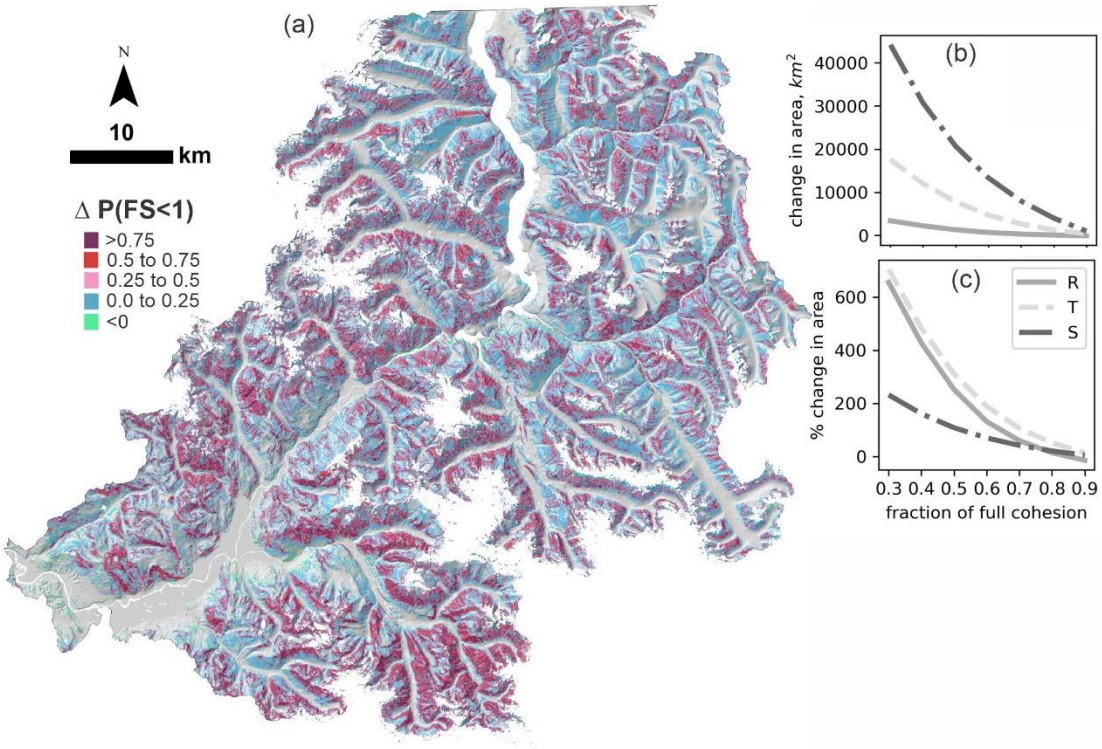

Figure 13. (a) Probability of FS<1 for the hot, wet climate, during the 2075 to 2100, assuming 70% decline from pre-fire total (roots+soil) cohesion; (b) change in potential landslide area versus cohesion, expressed as a fraction of full cohesion and; (c) percent change in potential landslide area versus cohesion.

4.5.3 Limitations of current findings

The findings reported above are preliminary as there are several important, additional factors that have not been accounted for. This analysis revealed that as a consequence of climate change, potential landslide rates in the historically, snow-dominated zone will decrease; however, this result was determined from a landslide model that does not include snow-pack effects on slope stability. Snow adds weight to the soil and can store water during snow-melt and rain-on-snow precipitation events. It's possible that the weight of the snow pack and/or pore water pressure from the snow pack could translate to the soil and lead to more landslides. There are multiple documented cases of snow-melt triggered landslides including a landslide that began as a snow-avalanche near Snoqualmie Pass, Washington (i.e., a glide avalanche; Stimberis and Rubin, 2011)

and a large, low angle landslide that was initiated following heavy snow-melt in volcanic debris near Mt. St. Helens, Washington (King5, 2023).

Additionally, the hydrologic mechanism leading to reduced water-year maximum relative wetness in the snow-zone needs to be further examined. In this assessment, we assumed that a thinner snow-pack causes lower relative wetness. However, the release of meltwater may be such that snowpack thickness has little effect on peak saturation and instead, melt rates might have a larger effect.

Finally, results have been presented in terms of the rain-dominated, transient and snow-dominated zones, but examination of the spatial trends in changes in relative wetness shows that there are different soil-water responses to the changes in temperature and precipitation within the same precipitation zone. For example, in 11a, the noted decrease in relative wetness in the snow-zone really only affects the western part of the snow-zone while the eastern half remains relatively unchanged. Similarly, the increase in relative wetness in the transient and rain dominated zones is most pronounced in areas with an upslope snow-zone. Areas where the rain-dominated and transient zones lack an upslope snow-zone, such as the west half of the model domain, appear to have little increase in relative wetness.

4.6 CONCLUSION

In this study we modeled the hydrologic and slope stability impacts of climate change in the North Cascades, a mountainous region where local climatology, ecosystems and terrain are highly variable. To evaluate the potential range of climate change impacts on hydrology and slope stability, we forced the model using three climate scenarios that capture the variability of predicted precipitation and temperature from 20 global climate models selected for the Pacific Northwest.

The three models represent a hot-dry, warm-dry and hot-wet scenario relative to the possible future climate scenarios. Relative to historic conditions, all three models have as much or more annual precipitation and all three are warmer.

To characterize climate change impacts across the landscape, we assessed results in terms of three historic precipitation zones: rain-dominated, transient and snow-dominated. We modeled hydrology using the Distributed Hydrology Soil and Vegetation Model (DHSVM) and landslide probability using LandslideProbability. To couple modeled hydrology from DHSVM with LandslideProbability, we developed the DistributedHydrologyGenerator (DHG), written in Python and implemented as a component for the package Landlab. DHG takes the raw flow and gridded hydrology outputs from a distributed hydrology model, like DHSVM, and downscales and converts it to the Landlab grid. It can be run using three options: (1) Stochastic Storm Generator mode; (2) Predefined Time Series mode or (3) Basic Statistics mode. In this study, we implemented DHG using option 3. Using this coupled modeling approach, we found the following:

- For all three climate scenarios, by the end of the 21st century, the potential landslide area (area of grid cells having a probability of failure > 0.5) increases in the rain-dominated and transient snow zones decreases and in the snow-zone relative to historic potential landslide area.
- For all three climate scenarios, the reduction in potential landslide area in in the snow-zone exceeds the increases in potential landslide area in the rain-dominated and transient zones and the total change in potential landslide area (sum of all three precipitation zones) goes down.
- Relative to the warm-dry scenario, the hot-dry scenario, which has nearly equivalent precipitation as the warm-dry scenario but is hotter, causes more potential landslides in the rain-dominated and transient zones because warmer temperatures cause a larger proportion of the snow

zone to shift from spring to winter peak runoff, which coincides with peak saturation conditions in the downslope transient and rain-dominated zones, and thus causes more landslides.

- Stand replacing fires that destroy tree roots and reduce total soil cohesion will elevate landslide probability in all regions of the North Cascades; however, the increase in probability will be largest for the lower regions of the historically snow-dominated zones where slopes are likely most dependent on root cohesion for stability.

4.7 REFERENCES

1. Abatzoglou, J. T., & Brown, T. M. (2012). A comparison of statistical downscaling methods suited for wildfire applications. *International Journal of Climatology*, 32(5), 772–780. https://doi.org/10.1002/joc.2312
2. Anagnostopoulos, G., Fatichi, S., & Burlando, P. (2015). An advanced process-based distributed model for the investigation of rainfall-induced landslides: The effect of process representation and boundary conditions. *Water Resources Research*, 51(9), 7501–7523. https://doi.org/10.1002/2015wr016909
3. Arnone, E., Noto, L., Lepore, C., & Bras, R. L. (2011). Physically-based and distributed approach to analyze rainfall-triggered landslides at watershed scale. *Geomorphology*, 133(3–4), 121–131. https://doi.org/10.1016/j.geomorph.2011.03.019
4. Bandaragoda, C., S. Lee, E. Istanbulluoglu, and A. Hamlet. 2019. Hydrology, Stream Temperature and Sediment Impacts of Climate Change in the Sauk River Basin. *Report prepared for Sauk-Suiattle Indian Tribe, Darrington, WA and the Skagit Climate Consortium*, Mt. Vernon, WA, Available at https://www.hydroshare.org/resource/e5ad2935979647d6af5f1a9f6bdecdea/
5. Barnhart, K. R., Hutton, E. W. H., Tucker, G. E., Gasparini, N. M., Istanbulluoglu, E., Hobley, D. E. J., Lyons, N. J., Mouchene, M., Nudurupati, S. S., Adams, J. M., & Bandaragoda, C. (2020). Short communication: Landlab v2.0: a software package for Earth surface dynamics. *Earth Surface Dynamics*, 8(2), 379–397. https://doi.org/10.5194/esurf-8-379-2020
6. Beechie, T. J., Collins, B. T., & Pess, G. R. (2013). Holocene and recent geomorphic processes, land use, and salmonid habitat in two north Puget Sound river basins. In *Water science and application* (pp. 37–54). American Geophysical Union. https://doi.org/10.1029/ws004p0037
7. Beven, K., & Kirkby, M. (1979). A physically based, variable contributing area model of basin hydrology, *Hydrological Sciences Bulletin*, 24(1), 43–69. https://doi.org/10.1080/02626667909491834
8. Blikra, & Nemec. (1998). Postglacial colluvium in western Norway: depositional processes, facies and palaeoclimatic record. *Sedimentology*, 45(5), 909–959. https://doi.org/10.1046/j.1365-3091.1998.00200.x
9. Burns, W. J.; Madin, I. P.; Mickelson, K. A., 2012, Protocol for shallow-landslide susceptibility mapping: *Oregon Department of Geology and Mineral Industries* Special Paper 45, 32 p. [http://www.oregongeology.org/pubs/sp/p-SP-45.htm]
10. Burton, A., & Bathurst, J. C. (1998). Physically based modelling of shallow landslide sediment yield at a catchment scale. *Environmental Geology*, 35(2–3), 89–99. https://doi.org/10.1007/s002540050296
11. Campforts, B., Shobe, C. M., Overeem, I., & Tucker, G. E. (2022). The Art of Landslides: How Stochastic Mass Wasting Shapes Topography and Influences Landscape Dynamics. *Journal of Geophysical Research: Earth Surface*, 127(8). https://doi.org/10.1029/2022jf006745
12. Climate Impacts Group, 2021, Skagit water supply and demand synthesis, https://cig.uw.edu/projects/skagit-water-supply-and-demand-synthesis/, visited 11/2022
13. Doten, C. W., Bowling, L. C., Lanini, J., Maurer, E. P., & Lettenmaier, D. P. (2006). A spatially distributed model for the dynamic prediction of sediment erosion and transport in mountainous forested watersheds. *Water Resources Research*, 42(4). https://doi.org/10.1029/2004wr00382
14. Elsner, M. M., Cuo, L., Voisin, N., Deems, J. S., Hamlet, A. F., Vano, J. A., Mickelson, K. E. B., Lee, S., & Lettenmaier, D. P. (2010). Implications of 21st century climate change for the

hydrology of Washington State. *Climatic Change*, 102(1–2), 225–260. https://doi.org/10.1007/s10584-010-9855-0
15. Frans, C. D., Istanbulluoglu, E., Lettenmaier, D. P., Fountain, A. G., & Riedel, J. L. (2018). Glacier Recession and the Response of Summer Streamflow in the Pacific Northwest United States, 1960–2099. *Water Resources Research*, 54(9), 6202–6225. https://doi.org/10.1029/2017wr021764
16. Gariano, S. L., & Guzzetti, F. (2016). Landslides in a changing climate. *Earth-Science Reviews*, 162, 227–252. https://doi.org/10.1016/j.earscirev.2016.08.011
17. Google Earth V 7.3.6., North Cascades, USA, http://www.earth.google.com [September, 2022].
18. Guthrie, R. H., Mitchell, S., Lanquaye-Opoku, N., & Evans, S. D. (2010). Extreme weather and landslide initiation in coastal British Columbia. Quarterly Journal of Engineering Geology and *Hydrogeology*, 43(4), 417–428. https://doi.org/10.1144/1470-9236/08-119
19. Harr, R. (1981). Some characteristics and consequences of snowmelt during rainfall in western Oregon. *Journal of Hydrology*, 53(3–4), 277–304. https://doi.org/10.1016/0022-1694(81)90006-8
20. Hobley, D. E. J., Adams, J. M., Nudurupati, S. S., Hutton, E. W. H., Gasparini, N. M., Istanbulluoglu, E., & Tucker, G. E. (2016). Creative computing with Landlab: an open-source toolkit for building, coupling, and exploring two-dimensional numerical models of Earth-surface dynamics. *Earth Surface Dynamics*, 5(1), 21–46. https://doi.org/10.5194/esurf-5-21-2017
21. Hoyer, S. & Hamman, J., (2017). xarray: N-D labeled Arrays and Datasets in Python. *Journal of Open Research Software*. 5(1), p.10. DOI: https://doi.org/10.5334/jors.148
22. IPCC, 2022: Summary for Policymakers [H.-O. Pörtner, D.C. Roberts, E.S. Poloczanska, K. Mintenbeck, M. Tignor, A. Alegría, M. Craig, S. Langsdorf, S. Löschke, V. Möller, A. Okem (eds.)]. In: Climate Change 2022: Impacts, Adaptation, and Vulnerability. *Contribution of Working Group II to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change* [H.-O. Pörtner, D.C. Roberts, M. Tignor, E.S. Poloczanska, K. Mintenbeck, A. Alegría, M. Craig, S. Langsdorf, S. Löschke, V. Möller, A. Okem, B. Rama (eds.)]. Cambridge University Press, Cambridge, UK and New York, NY, USA, pp. 3-33, doi:10.1017/9781009325844.001.
23. Istanbulluoglu, E. (2005). Vegetation-modulated landscape evolution: Effects of vegetation on landscape processes, drainage density, and topography. *Journal of Geophysical Research*, 110(F2). https://doi.org/10.1029/2004jf000249
24. Iman R L, Conover W J, 1983. A Modern Approach to Statistics, John Wiley and Sons, New York in Maidment, D. R. (1993). *Handbook of Hydrology*.
25. Jakob, M., Owen, T., & Simpson, T. (2012). A regional real-time debris-flow warning system for the District of North Vancouver, Canada. *Landslides*, 9(2), 165–178. https://doi.org/10.1007/s10346-011-0282-8
26. King5 News (2023), https://www.king5.com/article/news/local/sr-504-coldwater-lake-mount-st-helens-closed-debris-slide/281-6d452110-62d0-4f42-aa97-006444eae332, accessed 5/2023
27. Kuo, CC., Liu, YC., Su, Y. et al. Responses of alpine summit vegetation under climate change in the transition zone between subtropical and tropical humid environment. *Scientific Reports* 12, 13352 (2022). https://doi.org/10.1038/s41598-022-17682-2
28. Littell, J. S., Peterson, D. L., Riley, K. L., Liu, Y., & Luce, C. H. (2016). A review of the relationships between drought and forest fire in the United States. *Global Change Biology*, 22(7), 2353–2369. https://doi.org/10.1111/gcb.13275

29. Livneh, B., Rosenberg, E. S., Lin, C., Nijssen, B., Mishra, V., Andreadis, K., Maurer, E. P., & Lettenmaier, D. P. (2013). A Long-Term Hydrologically Based Dataset of Land Surface Fluxes and States for the Conterminous United States: Update and Extensions. *Journal of Climate*, 26(23), 9384–9392. https://doi.org/10.1175/jcli-d-12-00508.1
30. Mauger, G. S., Casola, J. H., Morgan, H., Strauch, R. L., Jones, B. L., Curry, B., Isaksen, T. B., Binder, L. W., Krosby, M., & Snover, A. (2018). State of Knowledge: Climate Change in Puget Sound. *Climate Impacts Group*, University of Washington. https://digital.lib.washington.edu:443/researchworks/handle/1773/34347
31. Montgomery, D. R., Schmidt, K., Greenberg, H. J., & Dietrich, W. E. (2000). Forest clearing and regional landsliding. *Geology*, 28(4), 311. https://doi.org/10.1130/0091-7613(2000)28
32. Montgomery, D. R., Sullivan, K. E., & Greenberg, H. J. (1998). Regional test of a model for shallow landsliding. *Hydrological Processes*, 12(6), 943–955. https://doi.org/10.1002/(sici)1099-1085(199805)12:6
33. Murphy, B., Czuba, J. A., & Belmont, P. (2019). Post-wildfire sediment cascades: A modeling framework linking debris flow generation and network-scale sediment routing. *Earth Surface Processes and Landforms*, 44(11), 2126–2140. https://doi.org/10.1002/esp.4635
34. Musselman, K. N., Lehner, F., Ikeda, K., Clark, M. P., Prein, A. F., Liu, C., Barlage, M., & Rasmussen, R. (2018). Projected increases and shifts in rain-on-snow flood risk over western North America. *Nature Climate Change*, 8(9), 808–812. https://doi.org/10.1038/s41558-018-0236-4
35. Natural Resource Conservation Service, https://www.nrcs.usda.gov/wps/portal/wcc/home/quicklinks/imap, accessed 2022
36. Natural Resource Conservation Service, SoilWeb, https://casoilresource.lawr.ucdavis.edu/gmap/, accessed 4/2023
37. Nash, J., & Sutcliffe, J. (1970). River flow forecasting through conceptual models part I — A discussion of principles. *Journal of Hydrology*, 10(3), 282–290. https://doi.org/10.1016/0022-1694(70)90255-6
38. Neiman, P. J., Schick, L. J., Ralph, F. M., Hughes, M., & Wick, G. A. (2011). Flooding in Western Washington: The Connection to Atmospheric Rivers*. *Journal of Hydrometeorology*, 12(6), 1337–1358. https://doi.org/10.1175/2011jhm1358.1
39. Rupp, D. E., J. T. Abatzoglou, K. C. Hegewisch, and P. W. Mote (2013), Evaluation of CMIP5 20th century climate simulations for the Pacific Northwest USA, *Journal of Geophysical Research: Atmospheres*, 118, 10,884–10,906, doi:10.1002/jgrd.50843.
40. Sekhar, L., Jetten, V., Van Westen, C., Sankar, G., & Van Beek, L. P. H. (2008). Pore Water Pressure as a Trigger of Shallow Landslides in the Western Ghats of Kerala, India: Some Preliminary Observations from an Experimental Catchment. *Physical Geography*, 29(4), 374–386. https://doi.org/10.2747/0272-3646.29.4.374
41. Slaughter, S., Burns, W., Mickelson, K., Jacobacci, K., Biel, A., & Contreras, T. (2017, April). Protocol for landslide inventory mapping from lidar data in Washington State. Washington Geologic Survey.
42. SSURGO Database | Natural Resource Conservation Service , United States Dept. of Agriculture. (2018). Retrieved 2018, from https://www.nrcs.usda.gov/wps/portal/nrcs/detail/soils/survey/?cid=nrcs142p2_053627
43. Stimberis and Rubin, (2011) Glide avalanche response to an extreme rain-on-snow event, Snoqualmie Pass, Washington, USA. *Journal of Glaciology*, 57(203)

44. Strouth, A., Burk, R. D., & Eberhardt, E. (2006). The Afternoon Creek rockslide near Newhalem, Washington. *Landslides*, 3(2), 175–179. https://doi.org/10.1007/s10346-005-0030-z
45. Tabor, R., Haugerud, R. A., Hildreth, W., & Brown, E. H. (2003). Geologic map of the Mount Baker 30- by 60 minute quadrangle, Washington. https://doi.org/10.3133/i2660
46. Tucker, G. E., & Slingerland, R. (1997). Drainage basin responses to climate change. *Water Resources Research*, 33(8), 2031–2047. https://doi.org/10.1029/97wr00409
47. United States Geological Survey, https://waterdata.usgs.gov/wa/nwis/rt, accessed 2022
48. United States Geological Survey: National Land Cover Data (NLCD), https://www.mrlc.gov/data?f%5B0%5D=category%3ALand%20Cover&f%5B1%5D=year%3A2011, accessed 2023.
49. Wershow, S. T. and E. G. DeChaine. 2018. Retreat to refu-gia: Severe habitat contraction projected for endemic alpine plants of the Olympic Peninsula. *American Journal of Botany* 105(4): 760–778.doi:10.1002/ajb2.1042
50. Westrick, K. (1999). Soil depth calculation script implemented in python. Available from https://github.com/pnnl/DHSVM-PNNL (accessed January 2019)
51. Wigmosta, M. S., Vail, L. W., & Lettenmaier, D. P. (1994). A distributed hydrology-vegetation model for complex terrain. *Water Resources Research*, 30(6), 1665–1679. https://doi.org/10.1029/94wr00436
52. Wigmosta M.S. & Perkins W.A (2001) Simulating the Effects of Forest Roads on Watershed Hydrology Land Use in Wigmosta, M. S., & Burges, S. J. (Eds.), Watersheds: Human Influence on Hydrology and Geomorphology in Urban and Forest Areas. *Water Science and Application 2*. Washington DC : American Geophysical Union. https://doi.org/10.1029/ws002
53. Wu, W., & Sidle, R. C. (1995). A Distributed Slope Stability Model for Steep Forested Basins. *Water Resources Research*, 31(8), 2097–2110. https://doi.org/10.1029/95wr01136

CHAPTER 5. CONCLUSION

5.1 SUMMARY OF FINDINGS AND NEW METHODS

This thesis began with a review of existing numerical approaches for modeling climate and land use impacts on watershed-scale sediment production and transport processes, or sediment cascades. Existing modeling approaches were grouped into three categories: landscape evolution models, lumped, highly conceptualized models and detailed, distributed approaches. From these models, three potential shortcoming of the existing models were identified:

1. Many of the models rely on a simple representation of hydrology or use a detailed hydrology model but force it with averaged meteorology data.
2. All of the models use a simple conceptualization of the landslide sediment delivery process to the channel network.
3. Many of the approaches appear to be configured with a pre-determined hydrology model.

In Chapter 2, I attempted to clarify when and where in a network-scale sediment transport model that impacts of Shortcoming 1 should be considered. Specifically, I modeled bedload transport capacity determined from two sets of hydrographs: (1) a high-accuracy hydrograph, modeled in DHSVM from temporally accurate precipitation (1-h) and (2) a low-accuracy hydrograph, modeled in DHSVM from daily (24-h) average precipitation rates. I compared modeled sediment transport capacity relative to location in the channel network and sediment availability. One of the by-products of this study was a new non-dimensional bedload transport capacity metric (ψ), that accounts for effective stresses in steep, mountain channels and requires

only one variable (a time series of flow) and two parameters, the bankfull flow rate (Q_b) and sediment availability (i.e., effective-bankfull-excess shear stress, τ_b'/τ_r'), to implement.

The major findings of Chapter 2 were that depending on channel network location, cumulative error can range from relatively small (10 - 20%), to large (more than two orders of magnitude) and that surprisingly, flow routing differences between hillslope and channels do not seem to dictate the network locations where modeled bedload transport capacity is sensitive to the accuracy of the modeled hydrology. Rather, spatial variability of the magnitude of τ_b'/τ_r' and changes in runoff due to snow accumulation and melt largely determined where accurate hydrology was needed. Error, which we quantified in terms of the ratio of the 1-h to 24-h bedload transport capacity, tended to be low in channel networks characterized by highly mobile (high τ_b'/τ_r') channel beds and/or during high intensity (high return-interval), uniform storm events that lasted several days. Error tended to be high in upland reaches, where τ_b'/τ_r' values are typically low relative to downstream reaches and snow accumulation and melt affect runoff rates. The key take-away from Chapter 2 is that if the focus of a network-scale sediment transport model is lowland, high τ_b'/τ_r' channels or extreme flows that greatly exceed the reference flow rate (flow rate corresponding to τ_r'), then a crude estimate of runoff hydrology is likely to result in relatively small error. Conversely, if the focus is upland, low τ_b'/τ_r' channels where snow-accumulation and melt processes affect runoff, a temporally accurate hydrograph from a distributed hydrology model or observations, that include snow-hydrology effects may be needed.

In Chapter 3, I addressed Shortcoming 2. This shortcoming is specifically referring to the rule-based landslide routing algorithm in Burton and Bathurst (1998), which shows up in slightly modified forms in later studies such as Doten et al. (2006) and Arenone et al. (2011). My goal was to develop a method for routing landslide sediment to the channel network that could more

accurately mimic sediment transport of real landslide runout processes. I developed, calibrated and tested Mass Wasting Router (MWR), a new cellular-automata landslide runout model capable of predicting the runout-extent, sediment transport and topographic change caused by a landslide. I found that once MWR is calibrated to an observed-runout dataset, it may work as a predictive tool for assessing the runout extent of neighboring potential landslides and a by-product of Chapter 3 was a new method for landslide runout hazard mapping.

Importantly, I demonstrated how to parameterize MWR from basic field observations of an actual runout path and developed a Markov Chain Monte Carlo calibration utility that determines the best-fit parameter values for a site as well as empirical Probability Density Functions (PDF) of the parameter values. Unlike many hydrology models, landslide runout models still largely rely on manual calibration and the use of an automated calibration technique is new. Also, since MWR evolves the underlying terrain, it may be applicable to long-time-scale model runs in which changes in sediment availability are an important aspect of the study.

In Chapter 4, I addressed Shortcoming 3. The detailed distributed approaches noted in the introduction are often hard-coded with a specific hydrology model (e.g., Burton and Bathurst ,1998; Doten et al., 2006; Arnone et al., 2011). As part of a study on the hydrologic and slope stability impacts of climate change in the North Cascades, I developed the DistributedHydrologyGenerator (DHG), a new tool that couples modeled distributed hydrology from models like DHSVM into the Landlab modeling framework. DHG can be run using three options: Stochastic Storm Generator mode, Predefined Time Series mode or Basic Statistics mode.

In our application of DHG, we used the output of DHSVM forced with three climate scenarios, to force the landslide model LandslideProbability. The three models represent a hot-dry, warm-dry and hot-wet scenario relative to the range of possible future climate scenarios typically

used to evaluate climate change in the Pacific Northwest (e.g., Bandaragoda et al. 2019). Relative to historic conditions, all three scenarios have as much or more annual precipitation and all three are warmer. To characterize climate change impacts across the landscape, we assessed results in terms of three historic precipitation zones: rain-dominated, transient and snow-dominated. In that study we expected evapotranspiration rates to modulate where landslide rates increased as a consequence of climate change. However, we found that changes in landslide rates were mostly due to changes in upland runoff processes. The transition from a spring-snow-melt driven peak runoff to winter-rainfall-runoff driven peak runoff caused potential landslide areas in the uplands to decrease but also caused lowland landslide area to increase. The increase appeared to be caused by the coincident arrival of upslope runoff with peak runoff rates in the lowlands.

5.2 A PROPOSED FRAMEWORK FOR MODELING SEDIMENT CASCADES

Given the tools and findings from Chapters 2, 3 and 4, a new framework for modeling sediment cascades is proposed in Figure 1. This proposed framework is implemented in Landlab (Barnhart et al. 2020) as a mixture of the new components developed in this thesis, existing components in Landlab and several new components that are still in development. There are three loops in the model. Loop 1 simply transports sediment available in the channel network. Like many of the other watershed-scale sediment production and transport models, the channel network is represented using a vector-node conceptualization. If channel-adjacent terraces exist, then Loop 2 is activated, which erodes the terraces and transports sediment in the channel network. If hydrologic conditions exceed a user defined threshold, Loop 3 is activated, which implements landslide runout, terrace erosion and sediment transport models. The hydrologic driver of the modeling framework is the DistributedHydrologyGenerator. In this case, operated in stochastic storm generator or predefined time series modes. In all loops, as long as sediment is present in the channel network, sediment

transport occurs. Sediment transport is handled by the existing component NetworkSedimentTransporter (Pfeiffer et al, 2019), which following Czuba (2018), conceptualizes and tracks sediment as “parcels”. Terraces are eroded only if they are present. Unstable hillslope locations are modeled using LandslideProbability. MassWastingRunout distributes landslide material to the channel network.

The new, in-development components shown in the proposed framework (Figure 1) include LandslideMapper, MassWastingEroder and the Pulser. LandslideMapper interprets a map of unstable model cells produced by a model like LandslideProbability into a map of discrete landslides. The shapes and sizes of the landslides interpreted by LandslideMapper can be calibrated to observed landslide sizes. The landslide maps produced by LanddslideMapper become the initial landslide polygons input into MassWastingRunout, which then redistributes the landslide material into the channel network via mass wasting runout processes. The runout material forms terraces and MassWastingEroder controls the rate that the terraces are eroded and enters the channel network as bedload. The output of MassWastingEroder is a table of sediment volumes, attributes and location on the vector-node representation of the channel network. The Pulser, which is already a part of Landlab, takes that output and actually inserts those volumes onto the channel network as parcels for transport via NetworkSedimentTransporter.

The result of this proposed framework is an improved spatial-temporal representation of watershed-scale sediment production and transport processes and the ultimate product of the proposed modeling framework is a time series of sediment transport at the watershed outlet that varies as a function of climate driven hydrology and landslide processes. The proposed framework is more flexible than past modeling approaches because it can be run with any hydrology model using the DistributedHydrologyGenerator. Finally, since the proposed framework is implemented

in the Python package Landlab, adding to or improving the proposed modeling framework is relatively easy.

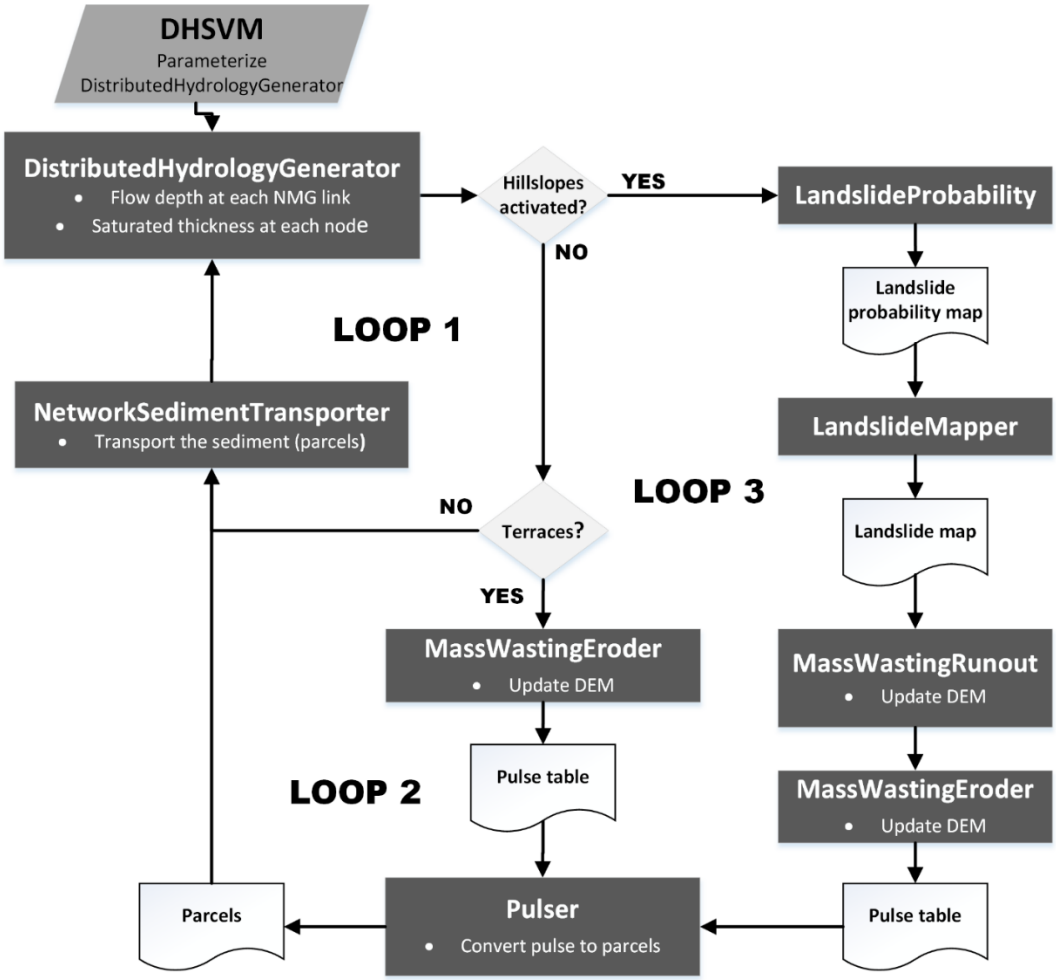

Figure 1. Proposed modeling framework for a new network-scale sediment production and transport model implemented in Landlab.

5.3 REFERENCES

1. Arnone, E., Noto, L., Lepore, C., & Bras, R. L. (2011). Physically-based and distributed approach to analyze rainfall-triggered landslides at watershed scale. *Geomorphology*, 133(3–4), 121–131. https://doi.org/10.1016/j.geomorph.2011.03.019
2. Bandaragoda, C., S. Lee, E. Istanbulluoglu, and A. Hamlet. 2019. Hydrology, Stream Temperature and Sediment Impacts of Climate Change in the Sauk River Basin. Report prepared for Sauk-Suiattle Indian Tribe, Darrington, *WA and the Skagit Climate Consortium*, Mt. Vernon, WA, Available at https://www.hydroshare.org/resource/e5ad2935979647d6af5f1a9f6bdecdea/
3. Burton, A., & Bathurst, J. C. (1998). Physically based modelling of shallow landslide sediment yield at a catchment scale. *Environmental Geology*, 35(2–3), 89–99. https://doi.org/10.1007/s002540050296
4. Czuba, J. A. (2018). A Lagrangian framework for exploring complexities of mixed-size sediment transport in gravel-bedded river networks. *Geomorphology*, 321, 146–152.
5. Doten, C. W., Bowling, L. C., Lanini, J., Maurer, E. P., & Lettenmaier, D. P. (2006). A spatially distributed model for the dynamic prediction of sediment erosion and transport in mountainous forested watersheds. *Water Resources Research*, 42(4). https://doi.org/10.1029/2004wr00382
6. Pfeiffer, A. M., Barnhart, K. R., Czuba, J. A., & Hutton, E. W. H. (2020). NetworkSedimentTransporter: A Landlab component for bed material transport through river networks. *Journal of Open Source Software*, 5(53), 2341. https://doi.org/10.21105/joss.02341