DHSVM-WSF

The Distributed Hydrology Soil Vegetation Model for Water Supply Forecasts

I. Abstract

Mountain Hydrology LLC introduces a first-in-class physically based operational water supply forecast (WSF) platform built on the Distributed Hydrology Soil Vegetation Model (DHSVM). The DHSVM-WSF framework supports real-time ensemble streamflow prediction to generate probabilistic water supply forecasts on weekly to seasonal timescales. With a generalized setup and calibration pipeline, DHSVM-WSF is ready for use on any watershed in the Western U.S. with area 100-10,000 km^2. The model setup uses machine learning to harmonize distributed land surface, subsurface, and meteorological data from numerous sources, including yearly dynamic vegetation updates to capture wildfires and other forest disturbances. The model calibration leverages multi-objective Bayesian optimization to constrain uncertainty in hydrological processes. DHSVM-WSF includes native support for snow data assimilation using the Snow Assimilation Water Accounting Method (SAWAM), an exclusive processing workflow that assimilates measured maps of snow water equivalent (SWE) while preserving the water mass balance and model dynamics (patent pending). Sub-seasonal ensemble weather forecasts are integrated into DHSVM-WSF to further constrain uncertainty. The historical performance of DHSVM-WSF typically obtains a daily Nash-Sutcliffe Efficiency (NSE) of 0.8 to 0.9 in snowy mountain watersheds, with seasonal water yield error typically on the order of 10% across wet and dry years using observed (backcast) weather.

II. Motivation

Reliable WSF predictions are crucial for water management planning and adaptation in the Western U.S., with forecasts informing decision-making to mitigate hazards and create economic value. Compared to statistical methods, which are mean-centralizing and rely on assumptions of stationarity, physically based forecast methods like DHSVM-WSF are better suited to predict extreme events and can operate under non-stationary conditions, such as climate change. Physical models are inherently interpretable, support new data sources, enable error detection and correction, and are more likely to lead to the right answers for the right reasons. Compared to other operational physically based forecast models, DHSVM-WSF parameterizes the landscape at much finer resolution (e.g., 90 m instead of 1,000 m in WRF-Hydro) and robustly handles critical hydrological processes such as vegetation-snow interactions. Since mountain hydrology is extremely nonlinear, high-resolution models exploiting the latest advances in remote sensing and machine learning are the natural endpoint as operational forecasts transition from lumped empirical relationships to fully distributed physical models.

DHSVM (Wigmosta et al. 1994, 2002) simulates the water mass and energy balance for each grid cell across a landscape (typically at 90 m resolution in DSHVM-WSF), with overland, subsurface, and channel flow routing. DHSVM was identified as the “preferred” hydrological model for use in mountainous terrain by Beckers et al. (2009) due to the model’s sophisticated treatment of hydrological processes and historical validation with observational data. However, DHSVM has seen limited operational application due to the model’s perceived complexity and computational burden. Beckers et al. (2009) estimate that setting up DHSVM for a single watershed typically requires a time commitment of 2-6 months, incurring a cost of $40,000 to >$100,000 per basin. By taking advantage of modern advances in automation and the burgeoning availability of large-scale, high-resolution datasets, Mountain Hydrology has developed the first-ever automated setup, calibration, and forecast pipeline for DHSVM, enabling the deployment of this powerful research-grade model to support the operational needs of water managers in an unlimited number of basins across the Western U.S.

Citation: Boardman, E. N. (2024). DHSVM-WSF: The Distributed Hydrology Soil Vegetation Model for Water Supply Forecasts. Mountain Hydrology LLC, White Paper 001v3. https://mountainhydrology.com/mh001_dhsvm-wsf/

© 2024 [Mountain Hydrology LLC](#)
Released under [CC BY-NC-SA 4.0](#)

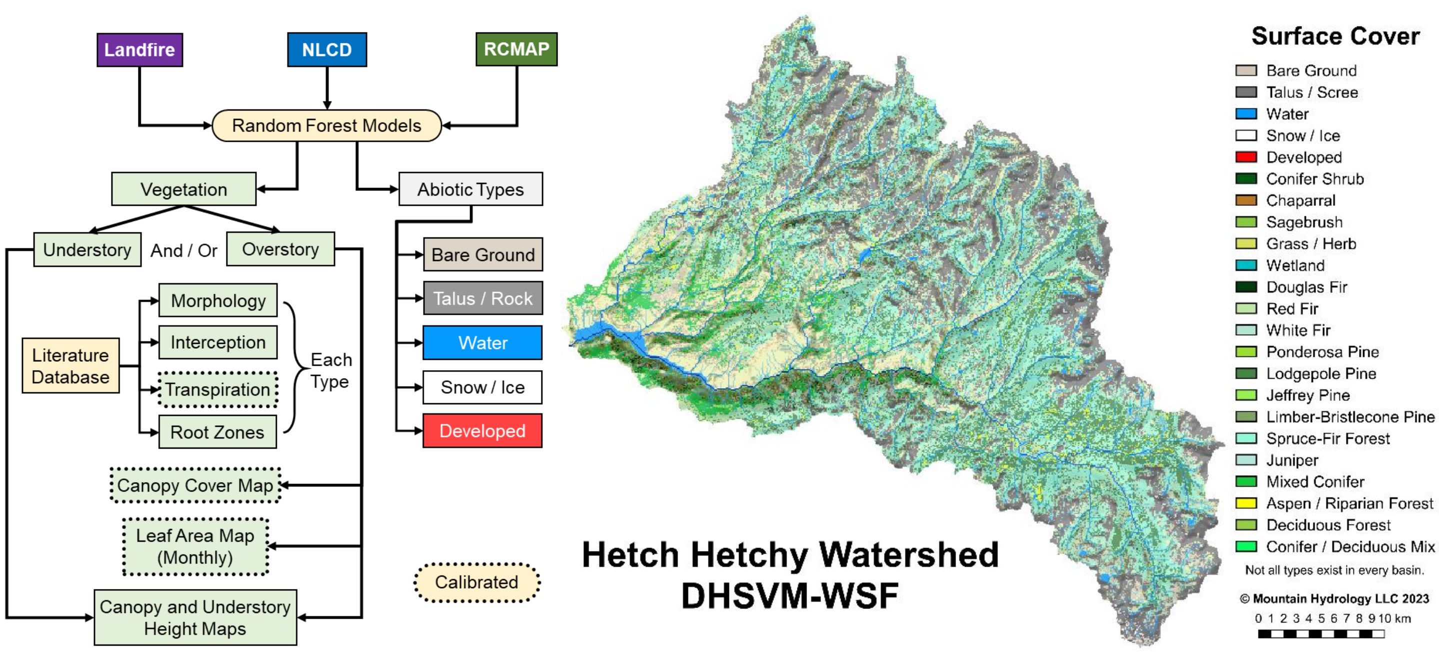
III. Model Setup

A. Terrain and Flow Routing

The Mountain Hydrology DHSVM setup pipeline begins by defining a watershed upstream of a user-supplied pour point, typically a stream gage, reservoir, or forecast point. SRTM digital elevation data (Farr et al. 2007) are hydrologically corrected with embankments or breaches to match natural flow paths. Channels are delineated using a minimum flow accumulation area that is determined by visual agreement with channel initiation points in local remote sensing imagery. The bankfull geometry of each channel segment is estimated from regional regression equations (Bieger et al. 2015), and variable Manning’s N roughness coefficients are estimated from a typical d84 particle size using Limerinos (1970).

B. Land Surface Model

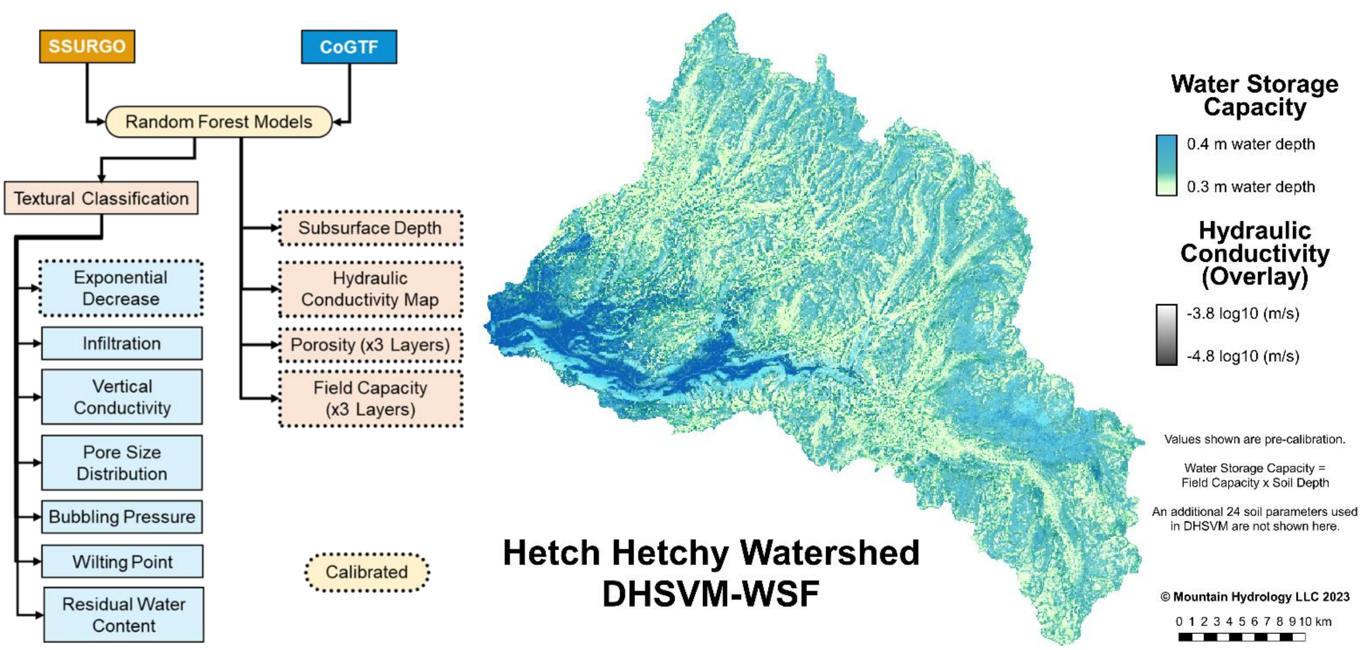
The DHSVM land surface model provides one of the most sophisticated treatments of grid-scale vegetation structure in any distributed hydrological model, requiring extensive data to parameterize. The Mountain Hydrology procedure involves harmonizing spatially distributed data from three main sources: RCMAP (Rigge et al. 2021), NLCD (Dewitz 2023), and Landfire (U.S. DOI 2022). Forest, shrub, and herb fractional cover maps are used to determine the presence and density of overstory and/or understory in every grid cell. Overstory leaf area (LAI) is estimated empirically from fractional cover (Pomeroy et al. 2002) and refined with calibration. Overstory and understory types are reclassified for each pixel to the species level where possible based on Landfire data, or otherwise reclassified to the functional type (e.g., DHSVM-WSF separately parameterizes distinct conifer classes like Douglas fir, red fir, lodgepole pine, etc., but a distinction is not always possible, so the model also has a Mixed Conifer class). Abiotic classes include bare ground (e.g., fire scars), talus/rock, water, ice, and developed areas, which can include impermeable areas and/or detention storage (Cuo et al. 2008). The attributes of each vegetation class (morphology, interception, transpiration, root zones, etc.) are determined at the species or community level based on an extensive review of more than 50 publications. Maps of overstory fractional cover, overstory monthly leaf area (LAI), overstory height, and understory height are directly input to the model and vary independently from the vegetation type, fully exploiting the distributed nature of DHSVM. Maps of vegetation type, cover, and other attributes are updated yearly based on RCMAP data derived from remote sensing, machine learning, and change-detection algorithms (Rigge et al. 2021). Dynamic vegetation maps are fully integrated into the DHSVM-WSF pipeline, reflecting fires and other disturbances during the calibration period and ensuring that the model provides the most up-to-date representation of the watershed during each forecast season.



© 2024 Mountain Hydrology LLC
Released under CC BY-NC-SA 4.0

C. Subsurface Characterization

DHSVM implements a complex and spatially explicit representation of subsurface processes using a four-layer soil model with three root-zone layers and one deeper quasi-aquifer layer. The DHSVM representation is best suited to modeling subsurface hydrology in environments characterized by steep slopes and shallow bedrock, which is typical of the Western U.S. mountains. The Mountain Hydrology setup pipeline uses machine learning to reanalyze and downscale a combination of spatial and tabular data from SSURGO soil survey datasets (NRCS 2022) and CoGTF water retention parameter maps (Gupta et al. 2022) to derive spatially coherent multi-layer maps of numerous soil properties at the DHSVM grid scale, including textural classification, near-surface hydraulic conductivity, porosity, field capacity, etc. Vertical profiles of transmissivity and various water retention parameters are derived for each soil textural class by leveraging the three-dimensional structure of the SSURGO soil horizon database and the multi-layer CoGTF maps. The soil depth pattern is based on a machine learning reanalysis of the SSURGO soil depths and the 30 m terrain curvature (Patton et al. 2018), with root zone depths and fractions primarily based on Jackson et al. (1996). The depth of the deep (quasi-aquifer) soil layer is estimated through calibration relative to the soil depth pattern map. Additionally, other highly sensitive parameters affecting transmissivity and water storage capacity (i.e., conductivity, the exponential decrease in conductivity with water table depth, porosity, and field capacity) are also calibrated relative to the respective maps.



D. Meteorological Forcing

DHSVM requires precipitation, temperature, humidity, wind, and shortwave/longwave radiation forcing data to drive the temporal simulation. The gridMET meteorological dataset (Abatzoglou 2013) provides daily precipitation, wind, and minimum/maximum temperature estimates at the 4 km scale based on long-term spatial patterns, meteorological reanalysis, and gauge measurements. These data are disaggregated to a 3-hour timestep and complementary humidity and radiation data are simulated using MetSim (Bennett et al. 2020), which is based on MTCLIM (Hungerford et al. 1989). Precipitation and temperature are downscaled to the DHSVM resolution using machine learning and historical spatial patterns to redistribute precipitation and derive spatially variable monthly temperature lapse rates. If spatial snow data are available (usually from remote sensing), the redistribution of snow by wind and terrain is simulated using an exclusive kernel-based method. In forecast mode, a 40-year ensemble of historical gridMET meteorology is appended to the 48-member CFSv2 sub-seasonal (30-day) forecast ensemble (Saha et al. 2014) to propagate variability in the post-forecast weather into ensemble streamflow predictions while constraining uncertainty wherever possible.

© 2024 Mountain Hydrology LLC
Released under CC BY-NC-SA 4.0

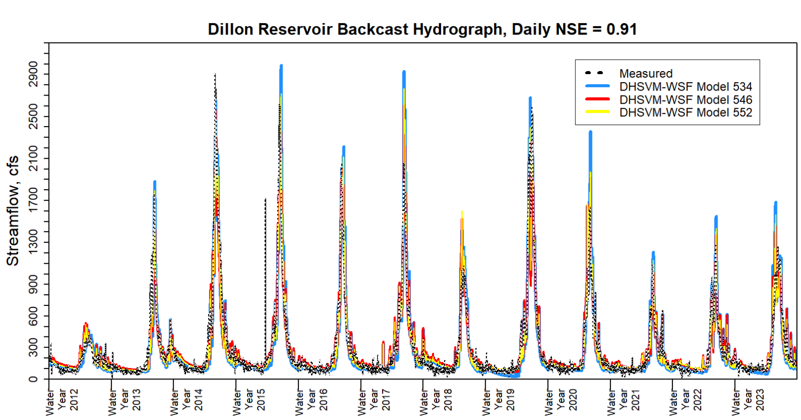
IV. Calibration

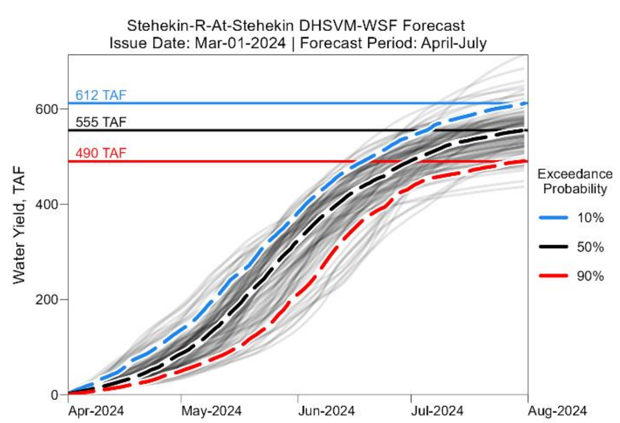
Although most of the hundreds of parameters and maps in DHSVM can be constrained *a priori* using field data and other observations, several key processes like subsurface flow and snowpack accumulation/ablation are emergent phenomena that do not permit a “correct” parameterization. Emergent hydrological processes can be constrained in physically based distributed hydrological models even when variables like soil depth or hydraulic conductivity remain uncertain (Boardman et al., in prep.). This is accomplished by using multi-objective Bayesian optimization to learn the Pareto frontier for objective functions targeting different hydrological signatures like baseflow, peak flows, snowmelt, and water yield. Mountain Hydrology implements these procedures using Gaussian Process surrogate models (Roustant et al. 2012) and parallel Particle Swarm Optimization (Zambrano-Bigiarini & Rojas 2013) to optimize the Expected Hypervolume Improvement (Binois & Picheny 2019) in a generational calibration framework. Pareto-optimal models are filtered to identify endmembers balancing tradeoffs between different objectives. The resulting DHSVM ensemble constrains residual uncertainty in modeled hydrological processes and provides the first fully Bayesian framework for hydrological prediction using a distributed physical model. Since DHSVM is a very computationally expensive model, Mountain Hydrology has invested in an on-premise high performance computing (HPC) cluster to enable robust calibration and ensure that real-time forecast results are available in a timely manner. As a point of interest, the Mountain Hydrology HPC is approximately 25x as powerful as the world’s fastest supercomputer in 1994, when DHSVM was released.

V. Forecasting

Mountain Hydrology has operationalized DHSVM for real-time water supply forecasting by developing an efficient pipeline to handle meteorological data and generate ensemble streamflow predictions. Forecasts are made using a 48-member ensemble sub-seasonal weather forecast from CFSv2 (Saha et al. 2014) downscaled and bias-corrected using historical gridMET data (Abatzoglou et al. 2023) combined with long-term weather from the prior 40 years of historical meteorology data propagated through the ensemble of calibrated DHSVM models. A stochastic Bayesian framework also propagates residual uncertainty in the calibrated hydrological models. In watersheds with SWE maps acquired by remote sensing, DHSVM-WSF can assimilate these SWE data while preserving the water mass balance and model dynamics. The Snow Assimilation Water Accounting Method (SAWAM) uses a spatially explicit local linearization of the pre-assimilation modeled water mass balance fluxes to construct a time-reversible surrogate model, from which it is possible to infer and correct errors in precipitation, streamflow, subsurface storage, and channel storage based on a single SWE map (patent pending).

The DHSVM-WSF deliverable is a CSV file containing daily ensemble streamflow traces for a given forecast period and a written report containing a summary of current conditions, with tabular water supply forecast guidance at the 10%, 50%, and 90% exceedance levels. Mountain Hydrology targets forecast delivery by close of business on the forecast issue date.





© 2024 Mountain Hydrology LLC
Released under CC BY-NC-SA 4.0

Get DHSVM-WSF in Your Watershed!

Availability

DHSVM-WSF historical simulations and real-time water supply forecasts are available for any watershed entirely contained in the Western U.S. with total area between 100 and 10,000 km^2. Larger basins and basins extending into Canada may be available on a case-by-case basis, as they require additional custom development. The model is best suited for use in mountainous environments such as the Sierra Nevada, Cascades, and Rocky Mountains.

Lead Time

DHSVM-WSF setup and calibration requires a minimum one to two month lead-time before the first forecast issue date.

Water Supply Forecast Deliverables

The primary deliverables from DHSVM-WSF, provided by close of business on the forecast issue date, are: (1) a CSV of daily ensemble streamflow traces for the period from the start of the current water year through the end of the customer-defined seasonal forecast period, and (2) a written report containing tabular water supply forecast guidance at the 10%, 50%, and 90% exceedance levels with accompanying qualitative and quantitative analysis of the current basin conditions and water supply outlook compared to other years. Additionally, a model setup and calibration report is prepared for each watershed each year detailing the land surface and subsurface characteristics and validating the model's performance.

Cost

The cost of operating DHSVM-WSF depends on the watershed size, number of forecast deliveries required, availability of SWE maps for assimilation, and other factors. The forecast cost includes yearly updates to the model setup to incorporate land surface disturbances and development costs to implement improved methods each year in addition to yearly re-calibration to take advantage of new SWE data and account for climate nonstationarity or land surface change. A detailed quote and complementary suitability analysis will be provided on request for any qualifying watershed (c.f. Availability). Please contact eli.boardman@mountainhydrology.com to request your customized estimate.

Research Applications

DHSVM is a powerful tool to investigate numerous hydrological research questions beyond water supply forecasting. For academic or government customers with technical expertise who wish to use DHSVM for *non-forecast* purposes, Mountain Hydrology can set up and calibrate DHSVM using state-of-the-art, reproducible procedures in a fraction of the time required to manually set up a distributed physical hydrology model from scratch. In support of public-benefit water science, Mountain Hydrology can offer the research community a considerable discount on model setup and calibration only. Deliverables will be negotiated on a case-by-case basis depending on specific project requirements but will include at minimum all materials necessary to run an ensemble of calibrated DHSVM models in the customer's watershed and a methodological appendix that can be included as supplementary material in scientific or technical publications.

© 2024 Mountain Hydrology LLC
Released under CC BY-NC-SA 4.0

Frequently Asked Questions

Isn't DHSVM a free model?
The DHSVM source code is distributed by the Pacific Northwest National Lab (PNNL) and is Public Domain (c.f. Perkins et al. 2019 "Software Availability"). However, one look at the configuration file template is usually enough to deter most would-be users. Mountain Hydrology is not selling DHSVM, but rather offering an economical service to make DHSVM operationally useful to the water resources community without having to dedicate years to learn the model's intricacies.

Are any other organizations affiliated with the DHSVM-WSF platform?
The DHSVM-WSF platform is exclusively developed and made available by Mountain Hydrology LLC, which is not affiliated with any other private or public entity. However, several of the original and ongoing contributors to DHSVM are aware of the project and enthusiastic about the potential for more widespread application of the model.

Can you share the scripts that help set up and calibrate the model?
Regrettably, no. The DHSVM-WSF processing pipeline uses several software packages with licenses that make it impossible to share the setup scripts with anyone, even just for personal investigation, without also releasing them under an overly permissive license. However, Mountain Hydrology is committed to furthering public benefit science on a good will basis by offering a large discount on model setup and calibration to qualifying academic or government researchers.

Is DHSVM-WSF better than XYZ other model?
DHSVM-WSF merges a comprehensive, process-based physical hydrological model with state-of-the-art data reanalysis, machine learning techniques, and physically based snow data assimilation to generate a first-in-class physical water supply forecast platform. For seasonal-scale water supply forecasting in mid-sized mountain watersheds, DHSVM-WSF could outperform many other forecast methods. However, past performance does not guarantee future results, and there is an element of unpredictability in any natural system. Sometimes, any given model might have lower error in a particular year just by random chance. Unlike most other physically based forecast platforms, DHSVM-WSF propagates uncertainty in the underlying model parameterization all the way to the final forecast exceedance levels, so the DHSVM-WSF framework is uniquely positioned to quantify uncertainty, a key element of any forecast-informed decision-making.

But isn't machine learning the answer to everything? Are physically based models still relevant?
Artificial intelligence and machine learning methods may prove useful for hydrological prediction in certain situations, particularly on short timescales. However, since machine learning methods are based on statistical relationships, they cannot predict out-of-sample disturbances, such as the watershed response to wildfires and climate change. Furthermore, the sources of uncertainty in machine learning models are difficult to interpret, leaving managers without answers when the models go wrong. Physically based distributed hydrological models are the only tool that fuses the "why" and "how" of hydrological science with the "what" of big data to generate rigorously justified, interpretable hydrological predictions.

I heard DHSVM is a "research model." Is such a complicated model really justified for water supply forecasting?
Yes! In places where water supply forecasts affect water management, the skill and reliability of those forecasts (including accurate uncertainty quantification) is of paramount importance. When water supply forecasts are misleading, unforeseen droughts or floods can threaten the livelihood and safety of entire populations. Conversely, reliable water supply forecasts can increase decision-making confidence, leading to more efficient strategies for optimizing competing water uses. Forecast-informed water management decisions can contribute hundreds of thousands or millions of dollars in economic benefit through increased hydropower generation, agricultural drought mitigation, flood prevention, and more. Best practices for modern water management demand the deployment of research-grade tools like DHSVM-WSF to ensure the security and resilience of the West's water resources.

© 2024 Mountain Hydrology LLC
Released under CC BY-NC-SA 4.0

References

Abatzoglou, J. T. (2013). Development of gridded surface meteorological data for ecological applications and modelling. *International Journal of Climatology*, 33(1), 121–131. https://doi.org/10.1002/joc.3413

Abatzoglou, J. T., McEvoy, D. J., Nauslar, N. J., Hegewisch, K. C., & Huntington, J. L. (2023). Downscaled subseasonal fire danger forecast skill across the contiguous United States. *Atmospheric Science Letters*, 24(8), e1165. https://doi.org/10.1002/asl.1165

Beckers, J., Smerdon, B., Redding, T., Anderson, A., Pike, R., & Werner, A. T. (2009). Hydrologic models for forest management applications. Part 1: Model selection. *Streamline Watershed Management Bulletin*, 13(1). https://www.pacificclimate.org/sites/default/files/publications/Beckers.StreamlineModelsPartI.Dec2009.pdf

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

Bieger, K., Rathjens, H., Allen, P. M., & Arnold, J. G. (2015). Development and Evaluation of Bankfull Hydraulic Geometry Relationships for the Physiographic Regions of the United States. *JAWRA Journal of the American Water Resources Association*, 51(3), 842–858. https://doi.org/10.1111/jawr.12282

Binois, M., & Picheny, V. (2019). GPareto: An R Package for Gaussian-Process-Based Multi-Objective Optimization and Analysis. *Journal of Statistical Software*, 89(8). https://doi.org/10.18637/jss.v089.i08

Cuo, L., Lettenmaier, D. P., Mattheussen, B. V., Storck, P., & Wiley, M. (2008). Hydrologic prediction for urban watersheds with the Distributed Hydrology-Soil-Vegetation Model. *Hydrological Processes*, 22(21), 4205–4213. https://doi.org/10.1002/hyp.7023

Dewitz, J. (2023). National Land Cover Database (NLCD) 2021 Products [dataset]. U.S. Geological Survey. https://doi.org/10.5066/P9JZ7AO3

Farr, T. G., Rosen, P. A., Caro, E., Crippen, R., Duren, R., Hensley, S., Kobrick, M., Paller, M., Rodriguez, E., Roth, L., Seal, D., Shaffer, S., Shimada, J., Umland, J., Werner, M., Oskin, M., Burbank, D., & Alsdorf, D. (2007). The Shuttle Radar Topography Mission. *Reviews of Geophysics*, 45(2). https://doi.org/10.1029/2005RG000183

Gupta, S., Papritz, A., Lehmann, P., Hengl, T., Bonetti, S., & Or, D. (2022). Global Mapping of Soil Water Characteristics Parameters—Fusing Curated Data with Machine Learning and Environmental Covariates. *Remote Sensing*, 14(8), 1947. https://doi.org/10.3390/rs14081947

Hungerford, R. D., Nemani, R. R., Running, S. W., & Coughlan, J. C. (1989). MTCLIM: A mountain microclimate simulation model (INT-RP-414; p. INT-RP-414). U.S. Department of Agriculture, Forest Service, Intermountain Forest and Range Experiment Station. https://doi.org/10.2737/INT-RP-414

Jackson, R. B., Canadell, J., Ehleringer, J. R., Mooney, H. A., Sala, O. E., & Schulze, E. D. (1996). A global analysis of root distributions for terrestrial biomes. *Oecologia*, 108(3), 389–411. https://doi.org/10.1007/BF00333714

Limerinos, J. T. (1970). Determination of the manning coefficient from measured bed roughness in natural channels (Water-Supply Paper 1898-B). U.S. Geological Survey. https://doi.org/10.3133/wsp1898B

Patton, N. R., Lohse, K. A., Godsey, S. E., Crosby, B. T., & Seyfried, M. S. (2018). Predicting soil thickness on soil mantled hillslopes. *Nature Communications*, 9(1), Article 1. https://doi.org/10.1038/s41467-018-05743-y

Perkins, W. A., Duan, Z., Sun, N., Wigmosta, M. S., Richmond, M. C., Chen, X., & Leung, L. R. (2019). Parallel Distributed Hydrology Soil Vegetation Model (DHSVM) using global arrays. *Environmental Modelling & Software*, 122, 104533. https://doi.org/10.1016/j.envsoft.2019.104533

Pomeroy, J. W., Gray, D. M., Hedstrom, N. R., & Janowicz, J. R. (2002). Physically Based Estimation of Seasonal Snow Accumulation in the Boreal Forest. Proceedings of the 59th Eastern Snow Conference, 93–108.

Rigge, M. B., Bunde, B., Shi, H., & Postma, K. (2021). Rangeland Condition Monitoring Assessment and Projection (RCMAP) Fractional Component Time-Series Across the Western U.S. 1985-2020 [dataset]. U.S. Geological Survey. https://doi.org/10.5066/P95IQ4BT

Roustant, O., Ginsbourger, D., & Deville, Y. (2012). DiceKriging, DiceOptim: Two R Packages for the Analysis of Computer Experiments by Kriging-Based Metamodeling and Optimization. *Journal of Statistical Software*, 51, 1–55. https://doi.org/10.18637/jss.v051.i01

Saha, S., Moorthi, S., Wu, X., Wang, J., Nadiga, S., Tripp, P., Behringer, D., Hou, Y.-T., Chuang, H., Iredell, M., Ek, M., Meng, J., Yang, R., Mendez, M. P., Dool, H. van den, Zhang, Q., Wang, W., Chen, M., & Becker, E. (2014). The NCEP Climate Forecast System Version 2. *Journal of Climate*, 27(6), 2185–2208. https://doi.org/10.1175/JCLI-D-12-00823.1

Soil Survey Staff, Natural Resources Conservation Service, United States Department of Agriculture. (2022). Soil Survey Geographic (SSURGO) Database [dataset]. https://sdmdataaccess.sc.egov.usda.gov

U.S. Department of the Interior, Geological Survey, and U.S. Department of Agriculture. (2022). LANDFIRE 2.3.0 [dataset]. http://www.landfire/viewer

Wigmosta, M. S., Nijssen, B., & Storck, P. (2002). The Distributed Hydrology Soil Vegetation Model. In V. P. Singh & D. K. Frevert (Eds.), *Mathematical Models of Small Watershed Hydrology and Applications* (pp. 7–42). Water Resources Publications, LLC. ISBN 1-887201-35-1.

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

© 2024 Mountain Hydrology LLC
Released under CC BY-NC-SA 4.0