Supplement of Hydrol. Earth Syst. Sci., 30, 5145–5171, 2026
https://doi.org/10.5194/hess-30-5145-2026-supplement

Hydrology and Earth System Sciences
Open Access

Supplement of

Groundwater hysteresis increasingly decouples flowing network length from streamflow as snow shifts to rain

Elijah N. Boardman et al.

Correspondence to: Elijah N. Boardman (eli.boardman@mountainhydrology.com)

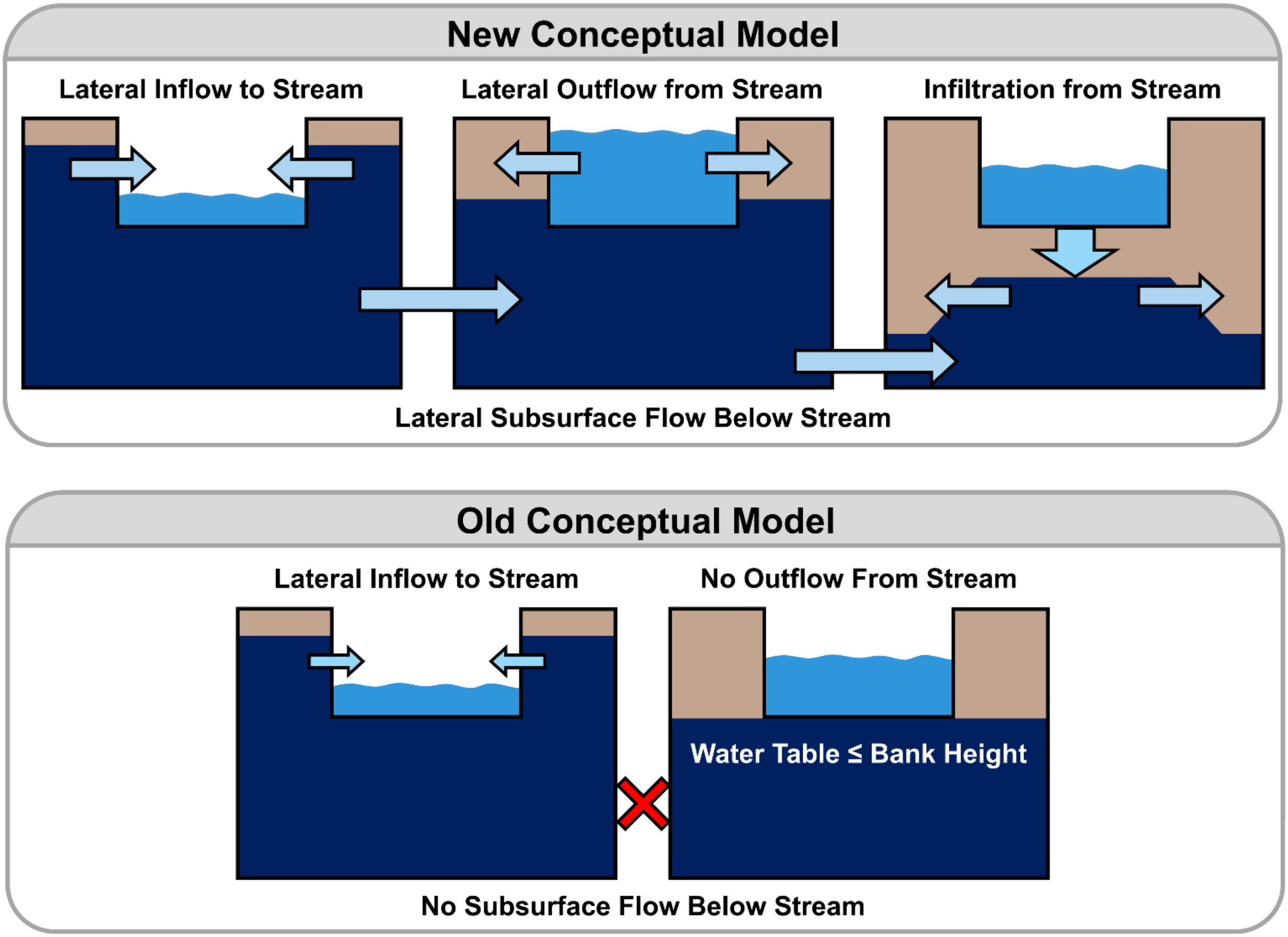


Figure S1. Cross-sectional view of the updated DHSVM grid-scale surface-groundwater coupling mechanism compared to the original DHSVM stream channel implementation.

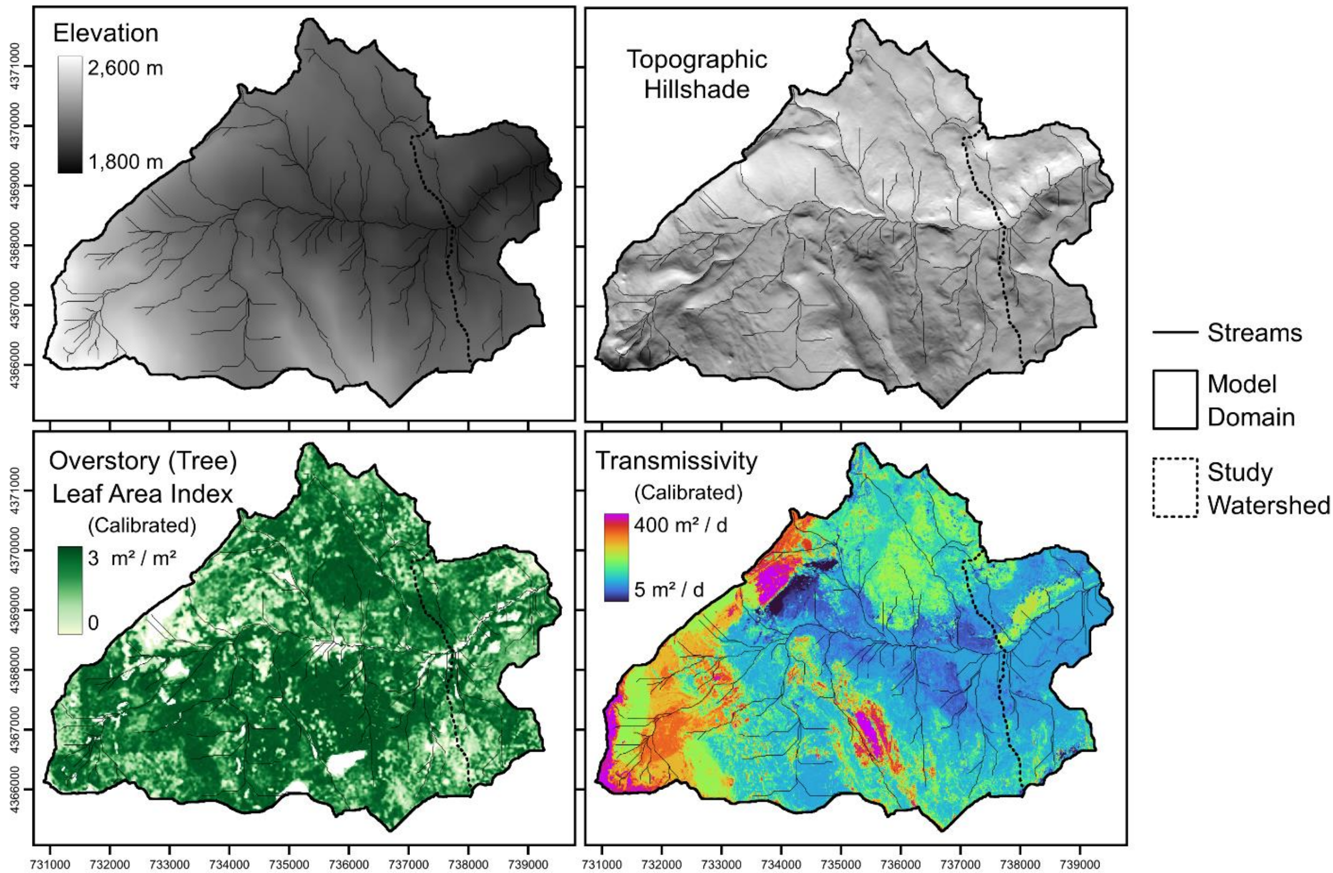


Figure S2. Maps of the Sagehen study watershed, showing (A) elevation, (B) hillshade, (C) area-averaged overstory leaf area index (LAI) within each grid cell (product of calibrated tree-scale LAI and grid-scale fractional cover), and (D) saturated hydraulic transmissivity (depth-integrated from calibrated surface conductivity, exponential decrease, and soil depth). Note that the model domain extends beyond the outlet of the study watershed to prevent no-flow boundary conditions from impacting surface-groundwater interactions at the stream gauge.

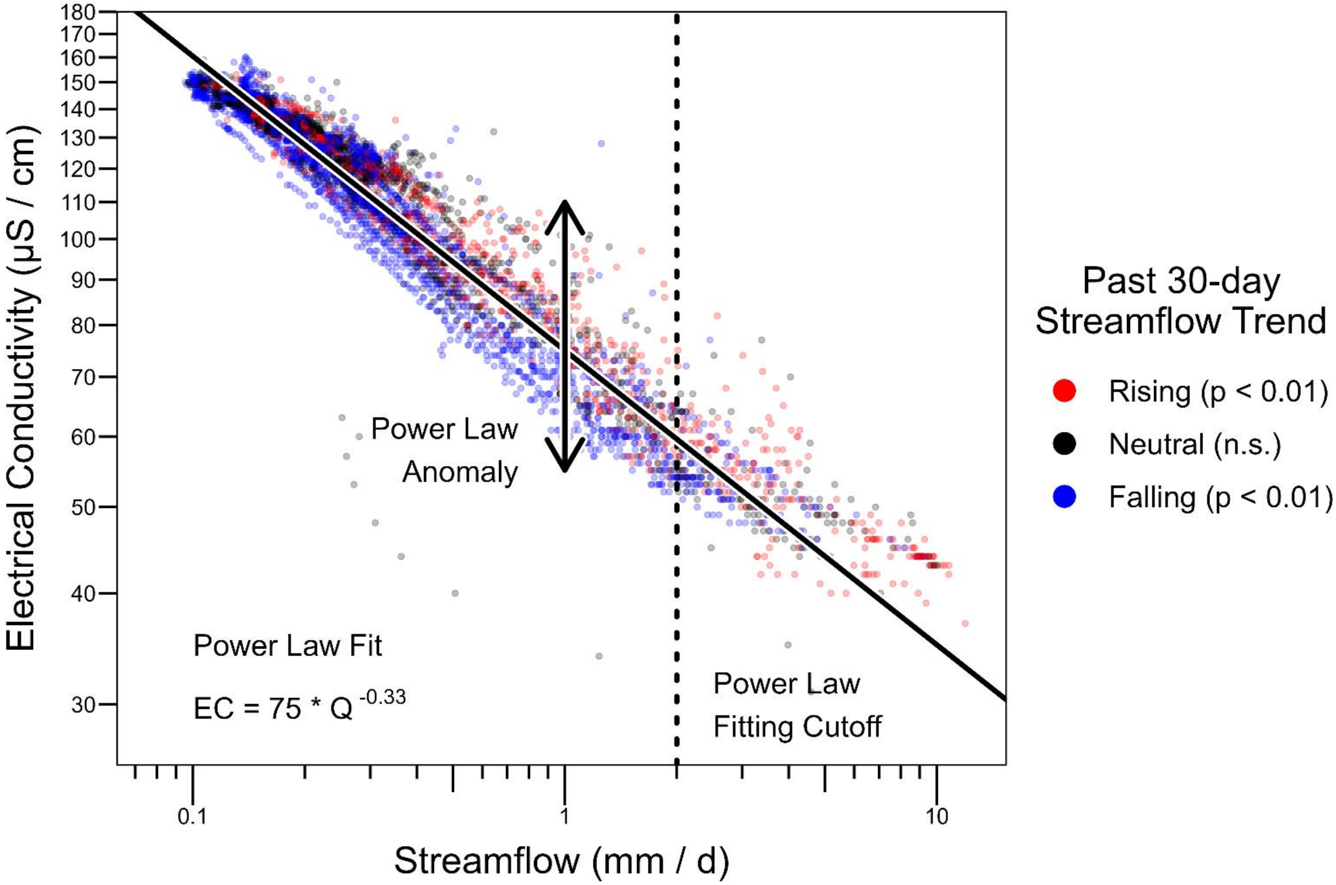


Figure S3. Log-scaled scatterplot between measured daily streamflow and measured daily mean electrical conductivity (EC). Each point represents one day. Points are colored identically to Figure 3 in the main manuscript (Mann-Kendall test). The power law anomaly is defined by the residual between measured EC and the best-fit regression line. Negative anomalies (lower-than-predicted EC) are generally associated with the falling limb of the hydrograph, suggesting an outsized contribution from fast groundwater flowpaths during the early recession season.

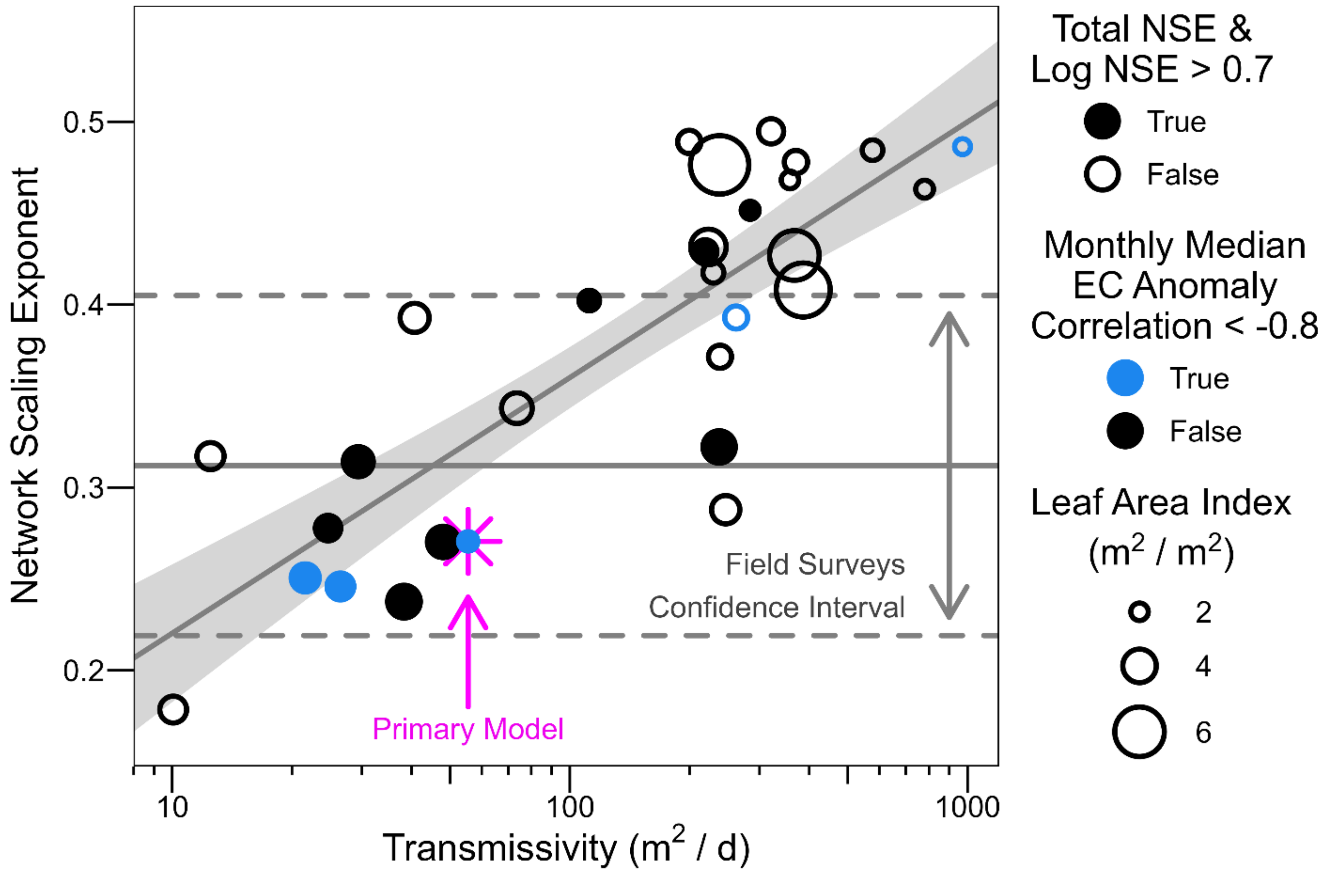


Figure S4. Parameter sensitivity and relationship between goodness-of-fit metrics for streamflow and the active network dynamics. Models with higher subsurface transmissivity tend to produce more dynamic stream networks (larger scaling exponent for the power law relationship between streamflow and network length). Models with higher transmissivity also have more variable leaf area index (LAI), indicated by circle size. Filled circles indicate models with daily NSE and log-scale NSE both > 0.7 over the full 2001 – 2024 simulation period. Blue circles indicate models with a strong correlation (< -0.8) between the simulated median monthly network anomaly and the measured median monthly electrical conductivity (EC) anomaly (Fig. 7). Gray lines indicate the confidence interval estimated from field surveys by Godsey and Kirchner (2014), as provided in their Table 1.

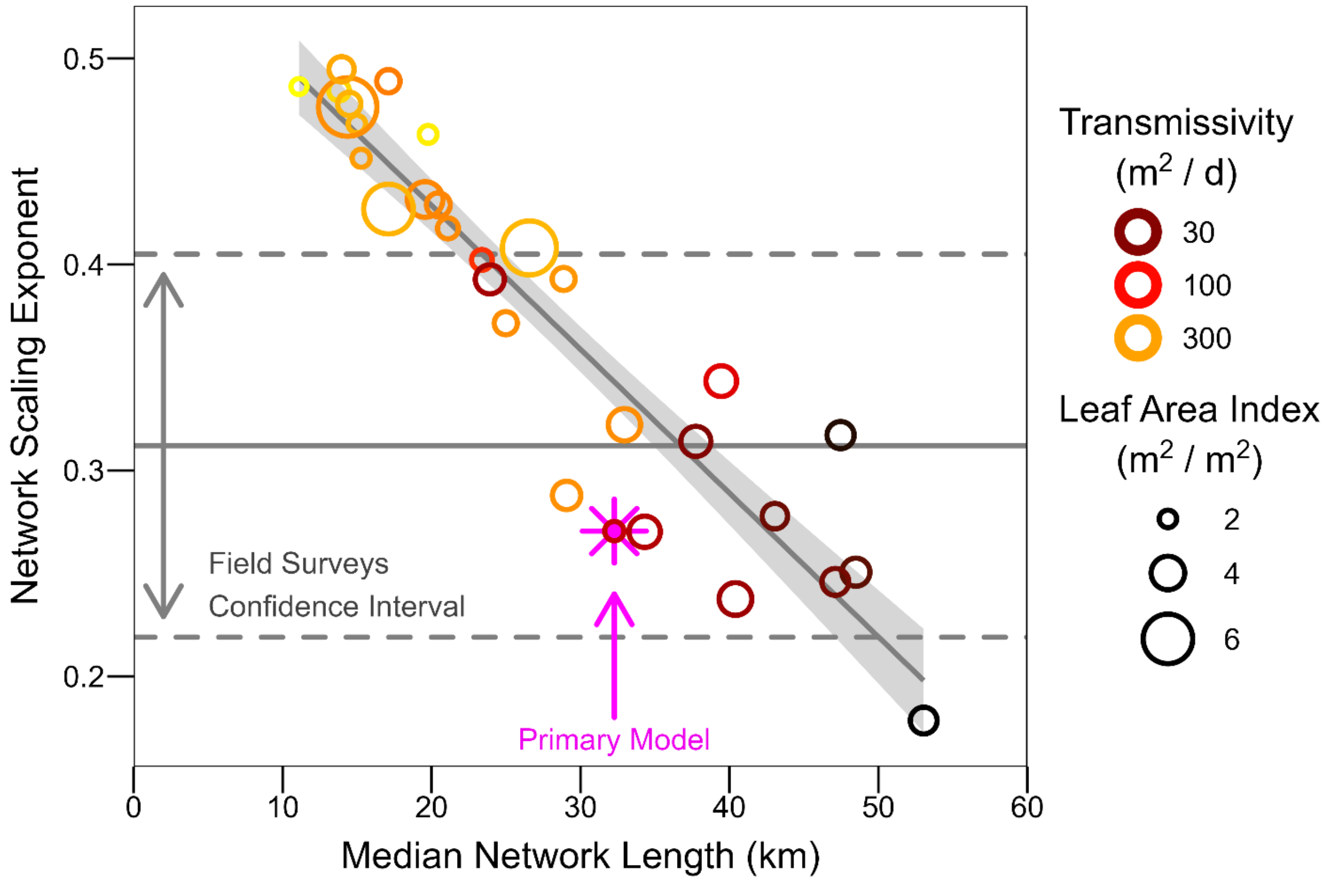


Figure S5. Parameter sensitivity and relationship between active stream network length and scaling dynamics. Models with higher transmissivity tend to produce smaller stream networks because a larger fraction of down-valley flow is partitioned to the subsurface. Models with larger networks also tend to predict a smaller scaling exponent for the power law relationship between streamflow and network length, which is the result of the network stabilizing on the surface with decreased transmissivity.

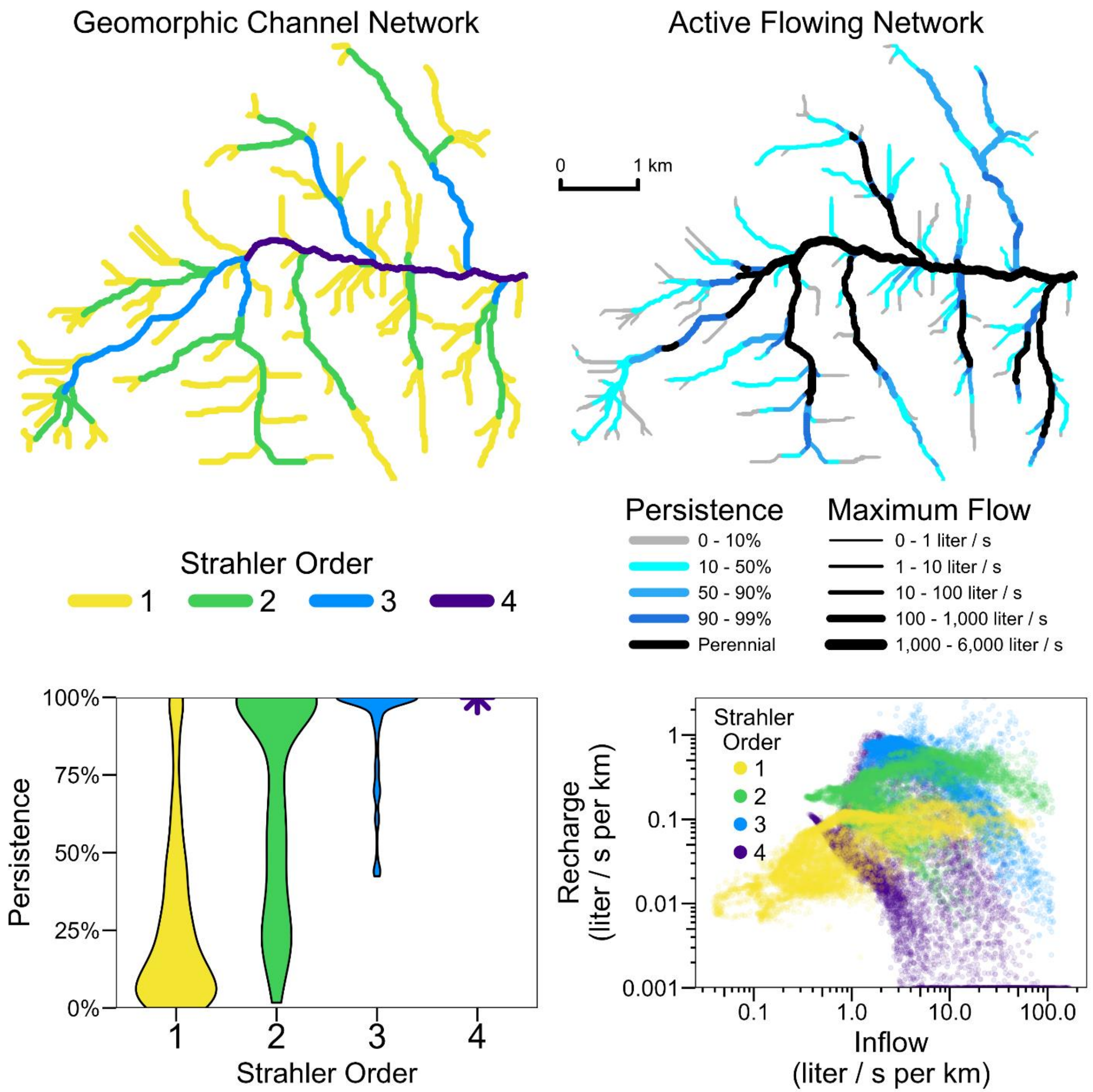


Figure S6. Comparison of geomorphic Strahler order and fractional streamflow persistence. Subplot (A) shows that higher-order streams are generally more persistent, but some first-order reaches are perennial, and some third-order reaches are ephemeral due to network disconnection. Subplot (B) shows the relationship between total inflow and recharge from all reaches of a given order on each day (normalized by length). First-order streams generally exhibit greater recharge at higher inflows as water is redistributed across relatively steep topography. In contrast, higher-order streams exhibit less recharge with increasing inflow as lower-gradient, higher-order valley bottoms reach saturation.

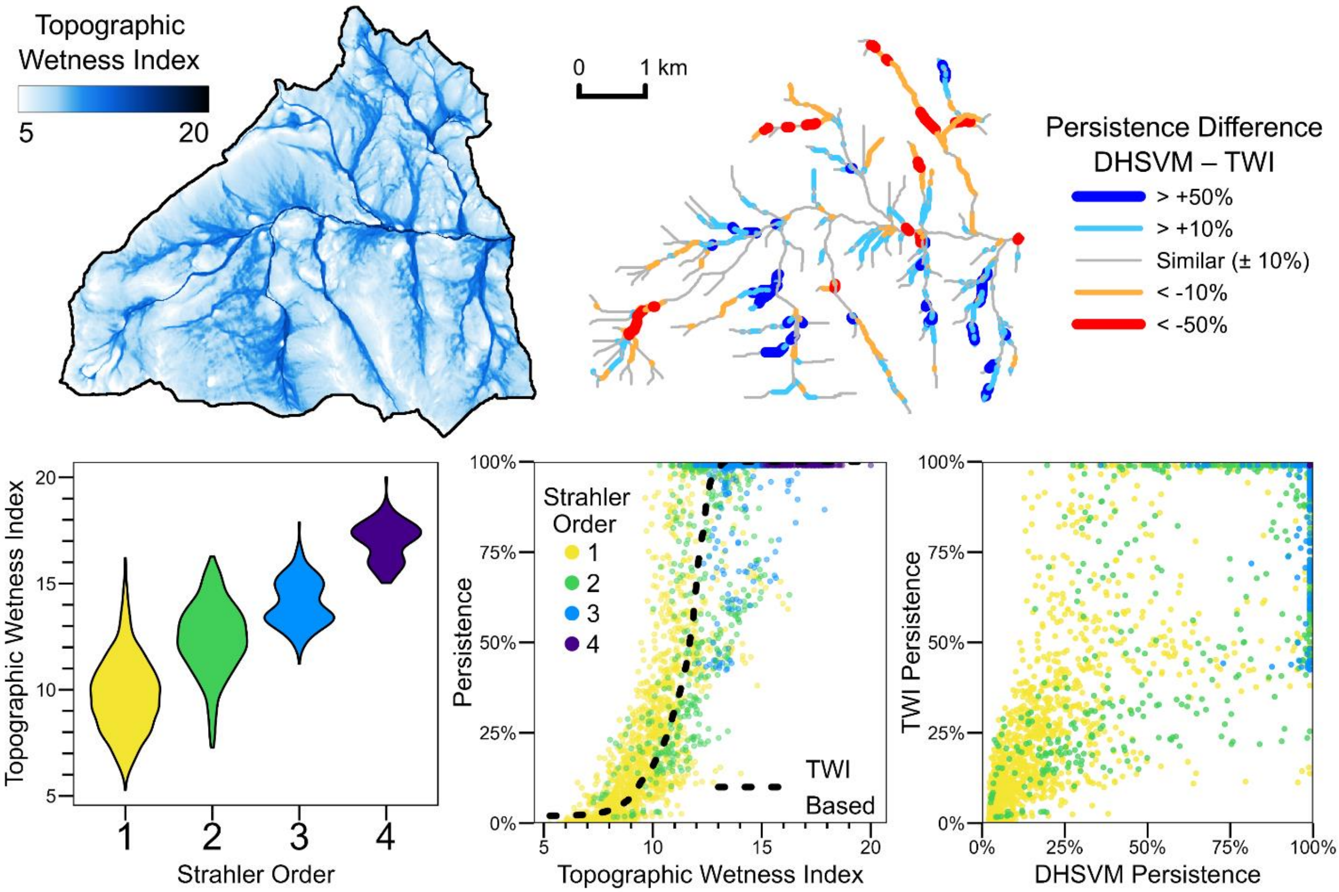


Figure S7. Comparison of stream persistence inferred using either DHSVM or a simplified scheme based on the power law with stream reaches activated in order of decreasing topographic wetness index (TWI). Subplot (A) shows the variability of TWI across Strahler orders. Subplot (B) shows the relationship between TWI and streamflow persistence, which is scattered for DHSVM and monotonically increasing for the TWI-based scheme. Subplot (C) compares the persistence of each stream reach simulated by DHSVM or estimated using the power law and TWI. All subplots are colored according to Strahler order.

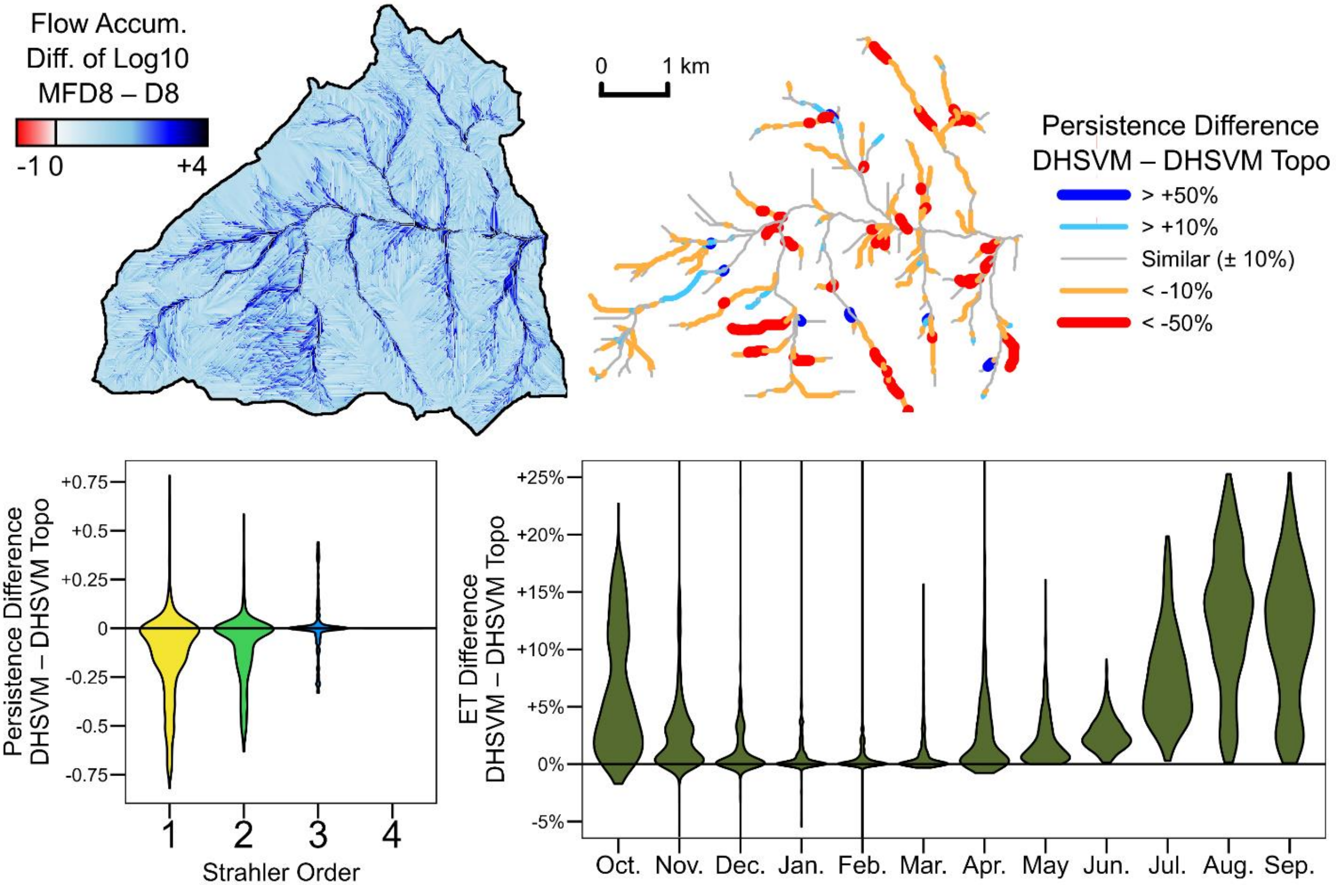
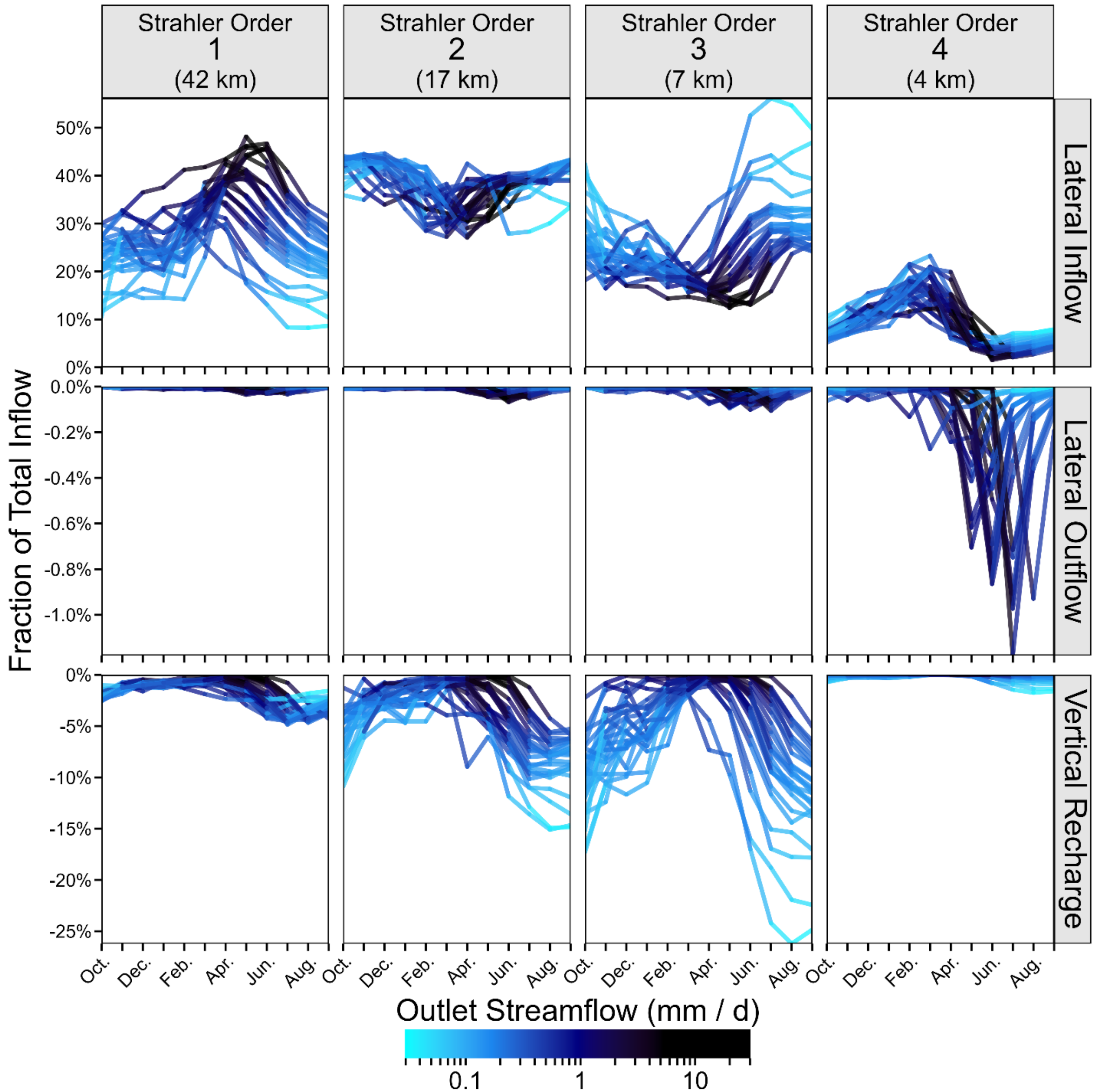


Figure S8. Comparison of stream persistence using the primary version of DHSVM (flow routed based on dynamic water table, multiple flow directions) relative to an alternate version where flow is routed using steepest-descent based on topographic gradients. The flow accumulation difference map (top left) shows areas where a MFD8 (multiple flow direction) algorithm implies a larger flow accumulation relative to a D8 (steepest-descent) algorithm. The topography-based steepest-descent routing scheme produces unphysical water table gradients (water "piled up" under stream) in valley bottoms that are greater than one grid cell in width (10 m), leading to an overestimation of streamflow persistence, particularly for first- and second-order streams. Additionally, the steepest-descent routing scheme frequently underestimates daily evapotranspiration (ET) by >20% during the warm season (July-October) because it neglects the contribution of dispersive groundwater flow away from losing stream reaches, which increases soil moisture and hence ET in the primary model (dynamic water table routing).



75 **Figure S9. Seasonal variation in the fractional inflow and recharge from stream reaches organized according to Strahler order (analogous to Fig. 3 in them main manuscript). Here, the collective "recharge" category from Fig. 3 is separated into lateral outflow and vertical recharge components (Fig. 1). Vertical recharge is the dominant outflow process in most stream reaches, but a small amount of lateral outflow may also occur in fourth-order reaches where channels are relatively deep. Note that each row of panels has a different vertical axis scale, unlike Fig. 3, which has uniform axes.**

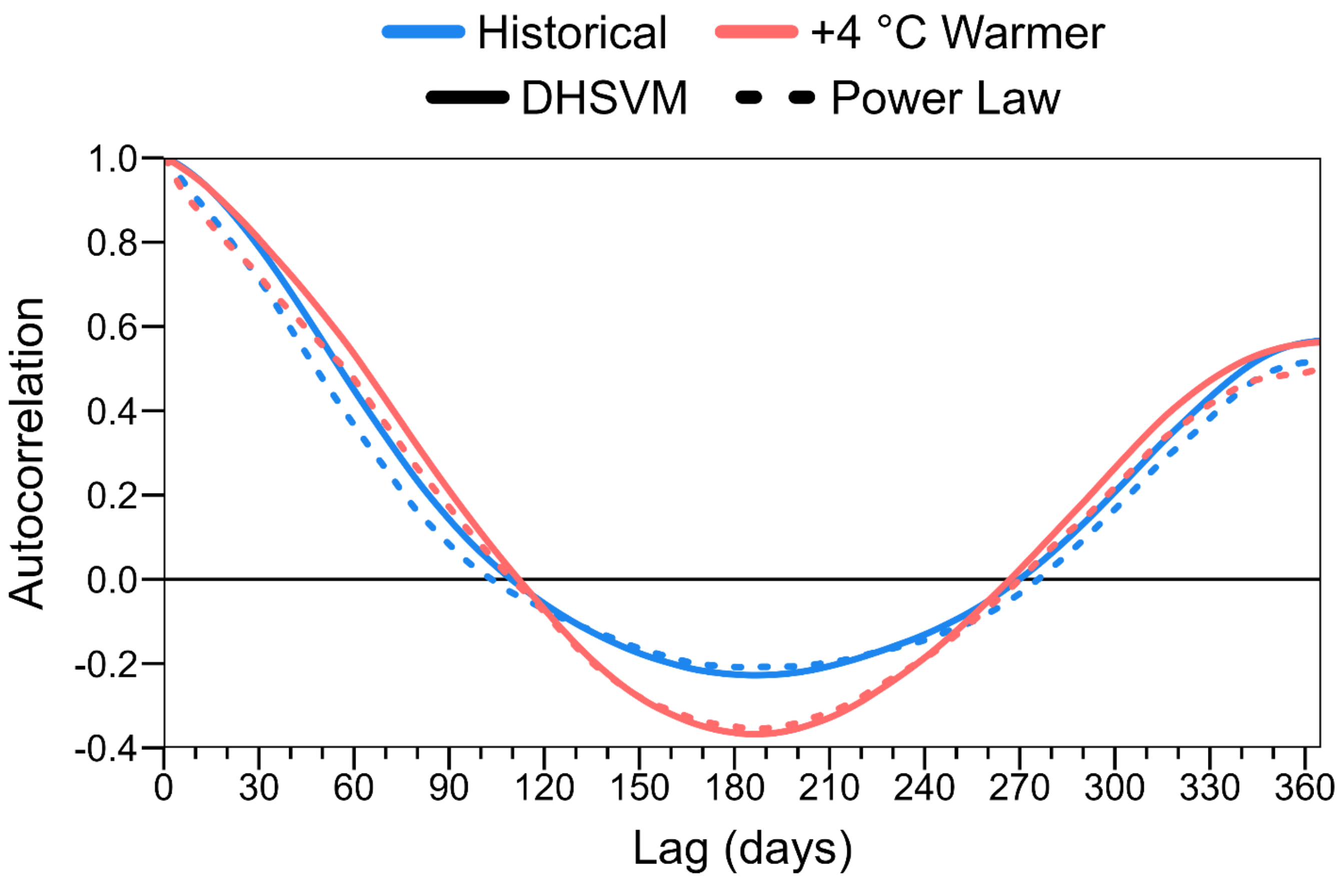


Figure S10. Autocorrelation for the timeseries of active network length simulated by DHSVM or estimated by a power law regression on streamflow (calculated for the full 2001-2024 period). The difference in autocorrelation at different lags (shown here) is the basis of Fig. 9 in the main manuscript. For lags of 1-90 days, DHSVM generally estimates a stronger autocorrelation than the power law, which indicates hysteresis in the relationship between streamflow and network length. For lags around a half year (183 days), the negative autocorrelation is primarily driven by seasonal variability in precipitation and snowmelt, i.e., the network length generally follows a seasonal cycle. Due to this seasonal cycle, the autocorrelation becomes positive again for lags around 1 year. We focus on shorter lags (< 90 days) in Fig. 9 to minimize the impact of the seasonal cycle, which is driven by meteorological dynamics instead of channel scaling relationships.

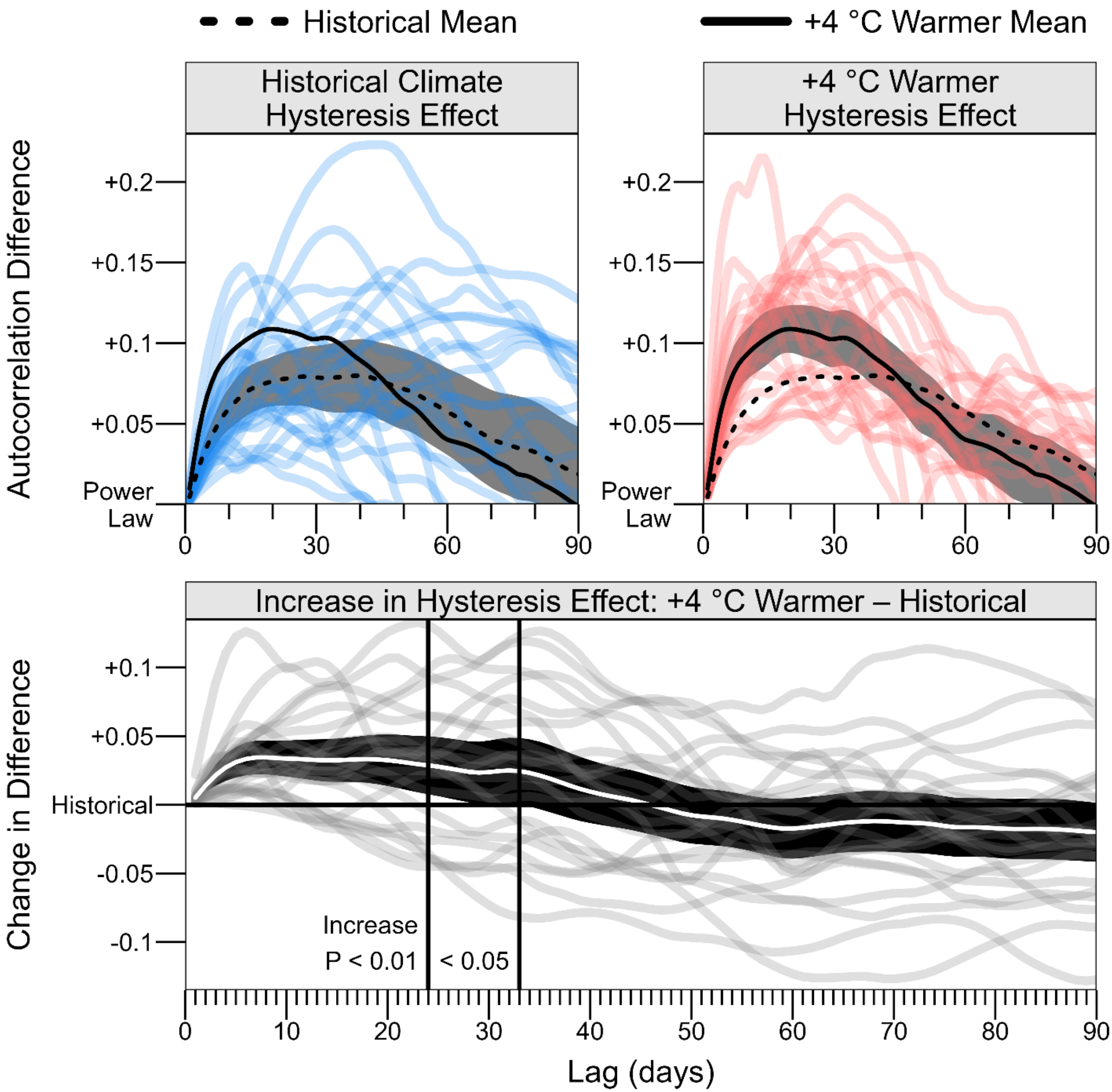


Figure S11. Difference in autocorrelation for the timeseries of active network length simulated by DHSVM or estimated by a power law regression on streamflow. Each line shows the autocorrelation at different lags within a single water year (2001-2024 baseline climatology). Black lines show the mean across years, and gray shading shows the 95% confidence interval of the sample mean. On average, DHSVM predicts a larger autocorrelation compared to the power law for lags of 1-80 days. In a warmer climate (more sporadic rain instead of consistent snowmelt), the autocorrelation underestimation by the power law becomes even greater for lags around 10-40 days.

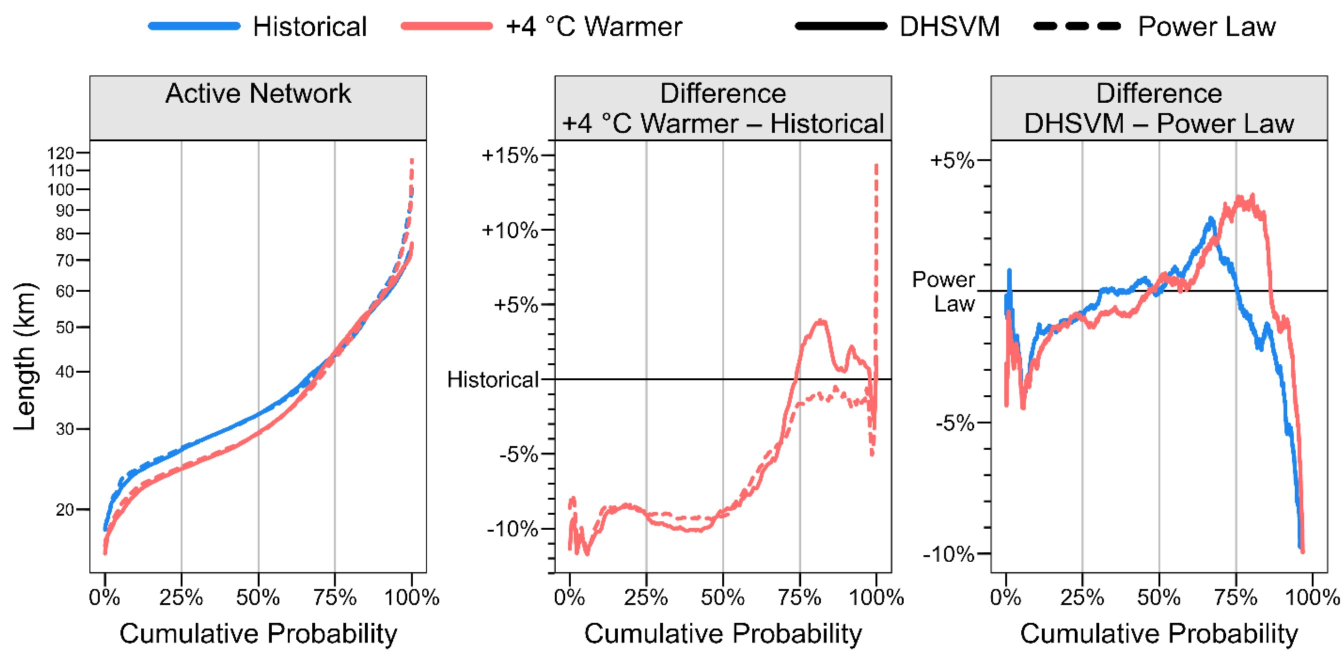


Figure S12. Cumulative distribution functions (CDF) for the daily active stream network length under historical and +4 °C warmer scenarios (2001-2024 baseline meteorology). The left panel shows the raw CDF for network length simulated using DHSVM (solid lines) or estimated from the power law regression (dashed lines). The center panel shows the difference in CDF between warmer and historical scenarios ($CDF_{warm} - CDF_{historical}$). The right panel shows the difference in CDF between DHSVM and the power law assumption ($CDF_{DHSVM} - CDF_{power\ law}$). In a warmer climate, the network length is further reduced by about 10% on below-median days, but the network length increases by up to ~5% on upper-quartile days. The power law assumption weakens in a warmer climate, as the power law predicts a decrease in upper-quartile network length, contrary to the DHSVM-predicted increase. Moreover, the power law predicts a large (>10%) increase in the absolute maximum network extent, which may not be possible depending on whether the definition of "active network" is restricted to the geomorphic channel network.

Movie S1. The supplemental video shows daily maps of simulated groundwater and stream networks in Sagehen Creek Basin for sample water years 2021-2024. The map is re-scaled based on the "dynamic storage" range of each grid cell, i.e., a value of 100% indicates that the cell is currently storing the maximum amount of groundwater simulated between 2021-2024, and a value of 0% indicates that the cell is storing the least amount of groundwater simulated on the same period. Note that each grid cell is re-scaled separately, so the 0-100% range represents a larger amount of water for some cells. Thus, the map is not intended to quantify total groundwater storage, water table level, or gaining/losing reaches; rather, it illustrates qualitative patterns of relative groundwater activation and flow. The magenta lines indicate streams that are actively flowing on each date.

Video link: https://doi.org/10.5446/72156