Simulation and statistical modelling approaches to investigate hydrologic regime transformations following Eastern hemlock decline

Kanishka Singh | James Knighton | Mark Whitmore | M. Todd Walter | James P. Lassoie

Abstract

Ecohydrological processes occurring at or near the Earth’s surface are strongly influenced by Eastern hemlock [EH; *Tsuga canadensis* (L.) Carrière], a foundation tree species of eastern North American forests. EH populations are currently threatened by the invasive hemlock woolly adelgid (HWA; *Adelges tsugae* Annand). HWA HWA populations have been expanding rapidly throughout the EH’s range. Catchment-scale research examining the hydrological consequences of HWA infestation is lacking, and plot-scale studies remain conflicted in their findings. Given the complex relationships between canopy interception, unsaturated and saturated groundwater storage, and root water uptake, it is not immediately clear how EH loss will affect the hydrologic cycle. We investigated the impact of EH mortality on stream discharge characteristics across a regional sample of catchments utilizing both simulation and statistical modelling approaches. We first examined the relationship between various catchment characteristics, including EH health, and three hydrological variables through regression analysis. We then employed a non-parametric statistical test to evaluate differences in hydrologic regime trends between non-infested and infested catchments. Finally, we calibrated a physically based hydrologic model and considered differences in optimal model parameter values and simulated overland runoff between non-infested and infested catchments. HWA presence modified several ecohydrological characteristics and precipitation partitioning between groundwater flows and surface runoff, potentially driving higher stream flashiness and overland flow, lower baseflow contributions and catchment storage, shorter flow-path lengths, and variable source area dilation at infested sites. Our results suggest that EH decline will augment flooding potential associated with the increasing frequency and intensity of Atlantic Basin tropical cyclone events. Further, our physically based simulation provides more determinate results than regression analysis, indicating that a purely statistical methodology, commonly utilized in studying the relationship between landcover characteristics and hydrologic regime, neglects dynamic physical ecohydrologic relationships.
1 | INTRODUCTION

Ecological and hydrologic processes in forested ecosystems are often dependent on the functional and structural roles of foundation tree species (Brantley et al., 2017) – regionally abundant species that generally occupy low trophic levels and maintain stable conditions within their environment (Ellison et al., 2005). Such processes occurring at or near the Earth’s surface are strongly influenced by Eastern hemlock (EH; Tsuga canadensis (L.) Carrière), one such foundation tree species present throughout much of eastern North America. EH influences the hydrologic cycle by sustaining root water uptake and transpiration through all seasons, and by maintaining consistent spatial patterns of throughfall, thereby possibly modulating soil moisture, groundwater recharge, baseflow regime, stream discharge, and stable flow regimes into streams (Brantley, Ford, & Vose, 2013; Brantley, Miniát, Elliott, Laseter, & Vose, 2014; Daley, Phillips, Pettijohn, & Hadley, 2007; Ford & Vose, 2007; Guswa & Spence, 2012; Kim et al., 2017; Knighton, Coneelly, & Walter, 2019a; Knighton, Souter-Kline, et al., 2019b; Snyder, Young, Lemarié, & Smith, 2002; Stadler, Müller, Orwig, & Cobb, 2005).

Forest communities dependent on EH-regulated microclimatic conditions are threatened by the hemlock woolly adelgid (HWA; Adéges tsugae Annand), an invasive insect that feeds on starch reserves found in the parenchyma cells of EH xylem (Young, Shields, & Berlyn, 1995). HWA infestations progress rapidly both spatially and temporally; the insect was first reported in the southeastern US in the 1950s (Gouger, 1971 as in Trotter & Shields, 2009) and its range has since extended across the eastern coast of the US from Georgia to Maine, and West to Tennessee (Skinner, Parker, Gouli, & Ashikaga, 2003). EH feeding results in the death of developing EH buds, precluding the production of new foliage, and although the rate of EH decline following the onset of infestation varies depending on a variety of context-specific factors such as EH stand size and density, soil type, and mean annual temperature (MAT), past studies have recorded mortality in as little as 4 years (Ford, Elliott, Clinton, Kloeppel, & Vose, 2011; Young et al., 1995).

EH serves an important function with regards to the water cycle due to sustained transpiration during leaf-off months and potentially deeper root water uptake during the growing season; thus, loss of this species, or a vitiation of its hydraulic potential, is projected to cause significant seasonal shifts or declines in annual catchment actual evapotranspiration (ET) (e.g., Adams et al., 2012; Brantley et al., 2013; Daley et al., 2007; Ford & Vose, 2007; Kim et al., 2017; Knighton, Coneelly, & Walter, 2019a; Stadler et al., 2005). At a northeastern US catchment, Knighton, Coneelly, and Walter (2019a) demonstrated that riparian EH likely uptake more recently infiltrated (i.e. younger) water from larger soil pores relative to uptake by the deciduous American beech (Fagus grandifolia), possibly further explaining the mechanism by which near-stream EH influences catchment water yield and diurnal discharge variations. Once infested with HWA, tree stomatal conductance and root to stem conductance decrease (Domec et al., 2013; Gonda-King, Gómez, Martin, Orians, & Preisser, 2014). As a result, HWA infestation and subsequent EH mortality can reduce annual stand water use by 10–40% (Domec et al., 2013; Ford & Vose, 2007), depending on season and stand composition in southern Appalachian forests. In these forests where EH comprises 50–60% of total basal area, for example annual and winter ET decline by 22% and 74%, respectively (Brantley et al., 2013). Infested forests also intercept less precipitation, thus experiencing increased throughfall volume and a distortion of throughfall spatial distribution pattern (Guswa & Spence, 2012; Stadler et al., 2005).

Mortality of EH could indirectly affect soil moisture and streamflow timing and volume across different temporal and spatial scales. Given unaltered precipitation patterns and a reduction in ET, more water may infiltrate the soil and recharge groundwater, and create overland flow and shallow subsurface flow. The partitioning between these processes will, however, depend on soil infiltration capacity and water table depth (Adams et al., 2012), and therefore may not be consistent across space. There is some evidence that EH loss will precipitate increases in rooting zone soil moisture and annual catchment water yield, along with associated decreases in daily amplitude of streamflow and a dilation of variable source areas (VSAs) (Ford & Vose, 2007; Kim et al., 2017; Knighton, Coneelly, & Walter, 2019a; Lustenhouwer, Nicoll, & Ellison, 2012). Prior research also shows an overall increase in annual ET rates over the long-term, particularly in the summer, as tree species that replace EH following mortality exhibit greater transpiration rates (Daley et al., 2007; Hadley et al., 2008).

Studies investigating the ecohydrologic repercussions of tree decline and die-off, in general, are few (Adams et al., 2012), and research on the hydrologic consequences of EH mortality specifically is mostly limited to the tree- or plot-scale (Daley et al., 2007; Ford & Vose, 2007; Hadley et al., 2008; Lustenhouwer et al., 2012; Domec et al., 2013), although there are some studies at the catchment- or watershed-scale (Brantley et al., 2014; Kim et al., 2017). Given the complex relationships between canopy interception, unsaturated and saturated groundwater storage, and root water uptake, it is not immediately clear how impacts of EH loss will carry over into the hydrologic cycle. A decline in EH stand density could impart non-linear changes to stream discharge as observed by Brantley et al. (2014), although whether this is a consistent and general pattern in all EH forests is uncertain. For example, it is possible that observed hydrologic impacts

**KEYWORDS**

Eastern hemlock, ecohydrology, flooding, hemlock woolly adelgid, hydrologic modelling, plant hydraulic regulation, regression analysis, variable source area
of HWA infestation are more pronounced in northeastern forests, particularly those in New England and upstate New York, than in southern Appalachian forests, as the former often feature higher densities of EH compared to the latter (Albani, Moorcroft, Ellison, Orwig, & Foster, 2010).

It is necessary for forest and watershed management efforts to have an approximate understanding of the hydrologic changes that are most likely to result from HWA infestation, which may not be obvious from previous context- and scale-dependent research. Thus, our objective was to evaluate, at a regional scale, relationships between hydrologic fluxes and the condition and presence of EH stands across the tree’s range in the U.S. through an analysis of long-term hydrologic records. Specifically, we asked:

- Does EH infestation impart changes to stream discharge characteristics?
- Does EH infestation change how catchments partition precipitation into baseflow and surface runoff?
- Can we better understand the dominant controls on local hydrology under infestation with a semi-physically based hydrologic model?

Previous studies examining the role of forest cover as a control on flooding frequency that rely on purely statistical methodologies have produced inconsistent outcomes of varying statistical significance (Bradshaw, Sodhi, PEH, & Brook, 2007; Brogna et al., 2017; Li, Chai, Yang, & Li, 2016; Tan-So, Adnan, Ahmad, Pattanayak, & Vincent, 2016). One possible physical reason for this is that the relationship between runoff response and vegetation varies by spatial scale and catchment features, such as the prevalence of snow or tree species composition and distribution (Zhang et al., 2017). Another methodological reason may be that statistical models of natural processes rely on oversimplifications of dynamic physical phenomena (Van Dijk et al., 2009), and neglect known ecohydrologic relationships between variables of interest. Process-based models offer a viable alternative, integrating physical detail with greater fidelity, but may themselves be limited by model uncertainty (Beven, 1989; Singh & Woolhiser, 2002). In an effort to circumvent limitations associated with each of these approaches, we pursued a methodology that combined statistical and simulation-based analyses; we addressed our research questions within the context of three models, which varied in their assumptions, and in the extent of knowledge of catchment hydrology carried forward to help explain variance in the observations. We first performed a parametric statistical analysis (i.e. a multivariate linear regression) utilizing a suite of common physical catchment characteristics and EH infestation condition (non-infested or infested) to determine if HWA outbreak can be identified as a significant influence on the discharge regime. We then employed a non-parametric statistical test to evaluate differences in hydrologic regime trends between non-infested and infested catchments. Finally, we used a physically based hydrologic model to determine the influence of EH condition on optimal calibration parameter values.

2  |  METHODOLOGY

2.1  |  Description of regional hydroclimate and land-cover datasets

We collected 81 daily streamflow records from the U.S. Geological Survey (USGS, 2018) across 17 states (Figure 1) based on the availability of at least one full year of flow data; 42 sites recorded measurements from January 1, 1980 to December 31, 2016. Hence, the most temporally expansive datasets spanned 37 years of daily measurements, whereas the shortest, a subset of four sites, spanned 2–3 years. The mean and median record length for all gauging locations was 27 and 37 years, respectively. Streamflow data were collected via the EcoHydrology package for the R scripting language (Fuka, Walter, Archibald, Steenhus, & Easton, 2018).

We organized sites into two broad categories: within native EH range with no reported HWA presence (non-infested, n = 37), and within native EH range with reported HWA presence (infested, n = 44) based on the spatial classification of HWA range (USDA Forest Service, 2018). Importantly, our binary classification did not capture the extent or duration of infestation, only that HWA presence had been observed.

We gathered soil textures and physical properties for the most dominant soil type at each stream discharge gauging location from the Natural Resource Conservation Service Soil Survey Geographic Database (SSURGO) (Soil Survey Staff, 2018), specifically: available water capacity (AWC), and saturated hydraulic conductivity (K_{SAT}). We used daily precipitation totals and mean annual precipitation (MAP) for each site for the period 1976–2017 from the CPC Unified Gauge-Based Analysis of Daily Precipitation over CONUS, a 0.25° precipitation product derived from a dense network of precipitation gauges (e.g., Chen et al., 2008). Daily minimum and maximum temperatures and MAT for the period 1979–2017 were sourced from the CPC Global Daily Temperature dataset, a 0.50° product. All CPC data were provided by NOAA/OAR/ESLR PSD Boulder, Colorado, U.S. (2018).

We determined land-cover composition from the 2011 National Land Cover Database (NLCD), a land-cover classification product at 30-m spatial resolution that organizes cover into 16 distinct categories (Homer et al., 2015). We determined watershed boundaries for each site using an HUC12 map from the national Watershed Boundary Database (USGS, 2018). We clipped the landcover map to the hydrologic unit boundaries of each gauge site, calculated the percentage of total area that each landcover type contributed for each catchment, and recorded the percentages of evergreen, deciduous, and mixed forest cover (Anderson, Hardy, Roach, & Witmer, 1976).

Infested and non-infested sites shared some characteristics and differed in others (Figure 2 presents distribution of statistical means). Infested catchments had greater MATs and MAPs than non-infested catchments (6.1 °C and 1003 mm vs 9.1 °C and 1186 mm), due to the location of non-infested sites at higher latitudes than infested ones (mean of 44.68 °F vs 40.94 °F). Infested catchments were larger in area and had lower K_{SAT} and AWC values than non-infested...
catchments (1912 km², 2.08 cm h⁻¹, 137.85 mm vs 1578 km², 3.5 cm h⁻¹, 159.61 mm). Non-infested sites had lower forest cover than infested ones (mean of 60.8% vs 71.8%), yet twice the evergreen forest cover as a percentage of total land cover than infested sites (12% vs 6%). Catchments within our sample were largely undeveloped, with only one site out of 81 recording that >10% of its land cover was characterized by medium intensity development, an NLCD classification of areas where 50–79% of total land surface is impervious. Across all sites, the mean value for the percentage of landcover of which more than half was composed of impervious surface was 1%. To determine gross forest cover loss at our sites, we utilized a Global Forest Change map (Hansen et al., 2013), a dataset of 30-m spatial resolution that captures forest loss, defined as stand-replacement or transition from forest to a non-forest state, for the period 2000–2018. We clipped this map to HUC12 boundaries and calculated the ratio of pixels that reported forest-cover loss to total pixels within each watershed from 2000 to 2016. Forest loss was low across our sample, with only one catchment recording >10% loss (at 13%); mean and median loss were 2% and 1%, respectively.

2.2 | Flow data characterization

We characterized daily streamflow data at all sites with three hydrologic indices: Richards–Baker Flashiness index (R-B) (Baker, Richards, Loftus, & Kramer, 2004), baseflow index (BFI), and the ratio of total
ET to total precipitation (ETP). We calculated R-B, a measure of the frequency of short-term changes in stream-flow patterns, following the methodology of Baker et al. (2004), which measures the ratio of discharge oscillations to total discharge as:

\[ R-B = \sum_{i=1}^{n} |q_i - q_{i-1}| / \sum_{i=1}^{n} q_i \]  

(1)

where \( q \) is the mean daily flow and \( i \) is the single time-step. The R-B is an indicator of hydrologic regime volatility and can be used to assess ecohydrological changes to discharge. Numerically, R-B exists within the range (0, 2), where higher values suggest greater flashiness or streamflow volatility.

We quantified changes to stream baseflow, the component of total discharge that is generally attributed to sloping groundwater flows as opposed to surface runoff contributions and is thus less sensitive to precipitation patterns, as the BFI (e.g., Eckhardt, 2008), or the ratio of total baseflow to total discharge. To separate baseflow volumes from the discharge record, we used the baseflow separation function of Nathan and McMahon (1990) as implemented in the EcoHydRology package (Fuka et al., 2018). Higher values of BFI indicate a lower contribution of surface runoff to total discharge and that the stream likely exhibits less regime volatility.

We estimated ETP for all sites as the difference between annual precipitation and annual streamflow (i.e. ET) divided by annual precipitation during the period 1976–2017 from CPC data (2018), under the assumption that changes to catchment storage are negligible compared to multi-decadal fluxes in precipitation, discharge, and ET. For this calculation, we studied the streamflow record for all sites to identify data gaps and accordingly omitted precipitation data for those dates missing a discharge measurement.

### 2.3 Statistical approaches to estimating Forest cover influence on discharge

#### 2.3.1 Multivariate regression of streamflow characteristics

All statistical and simulation analyses and data visualization were performed using RStudio version 1.1.456.

We first attempted to understand the influence of EH infestation on the previously described streamflow indicators via multivariate regression analysis. We parameterized three general multivariate linear regressions of the form:

\[ y = \beta_1 x_1 + \ldots + \beta_p x_p + \epsilon \]  

(2)

where the response variable \( y \) (i.e. R-B, BFI, and ETP) was predicted by catchment characteristics \( x_{ip} \) (i.e. \( K_{S\text{AT}}, \text{MAT}, \text{condition, area, etc.}, \) see Table 1).

#### 2.3.2 Kolmogorov–Smirnov (K-S) two-sample test

We measured the distance between the empirical cumulative distribution functions (CDFs) of the R-B, BFI, and ETP indices across both non-infested and infested sites through the two-sample, non-parametric K-S test. We designed each test such that the null hypothesis (i.e. both CDFs are drawn from the same data generating process) stated that EH infestation does not change catchment discharge metrics. Rejection of the null hypothesis would indicate that EH infestation may be influencing flow partitioning and processes that create surface runoff. The test-statistic (Ds) representing the distance between two CDFs was evaluated at levels 0.05, 0.10, and 0.20, as a measure of sensitivity as in Hornberger and Spear (1981).

### TABLE 1 Hydroclimate and physical catchment characteristics used as covariates within the multivariate regression analysis

<table>
<thead>
<tr>
<th>Regressor</th>
<th>Description</th>
<th>Source</th>
</tr>
</thead>
<tbody>
<tr>
<td>Condition</td>
<td>Classification of site as ‘1’ if located both within EH and HWA range, and ‘2’ if located within EH range but outside HWA range</td>
<td>Little, 1971; USDA Forest Service, 2018</td>
</tr>
<tr>
<td>Area</td>
<td>Area of watershed within which gauge is located</td>
<td>USGS, 2018</td>
</tr>
<tr>
<td>Lat</td>
<td>Geographic latitude of gauge location</td>
<td>USGS, 2018</td>
</tr>
<tr>
<td>( K_{S\text{AT}} )</td>
<td>Saturated hydraulic conductivity, or the rate at which the saturated soil transmits water</td>
<td>Soil Survey Staff, 2018</td>
</tr>
<tr>
<td>AWC</td>
<td>Available water capacity, or the capacity of soil to hold water</td>
<td>Soil Survey Staff, 2018</td>
</tr>
<tr>
<td>MAT</td>
<td>Mean annual air temperature at site</td>
<td>CPC Global Daily Temperature, 2018</td>
</tr>
<tr>
<td>MAP</td>
<td>Average precipitation at site</td>
<td>CPC Unified Gauge-Based Analysis of Daily Precipitation over CONUS dataset, 2018; Chen et al., 2008</td>
</tr>
<tr>
<td>Percent forest</td>
<td>A summation of landcover classified as deciduous, evergreen, and mixed forest</td>
<td>Anderson et al., 1976; Homer et al., 2015</td>
</tr>
</tbody>
</table>
2.4 Hydrological model development

We analyzed flow partitioning at only those stream-gauge sites that met the criteria of exhibiting continuous daily streamflow records over the 20-year period from January 1, 1997 to December 31, 2017 (n = 45) using the Lumpsed VSA model, JoFlow (Archibald et al., 2014). JoFlow is a semi-physics-based hydrologic model that maintains a lumped daily water budget at the catchment scale while spatially distributing runoff processes. We simulated daily potential evapotranspiration (PET) and snowpack accumulation and melt dynamics following Archibald et al. (2014) and Walter et al. (2005), respectively. Within JoFlow, surface runoff is solved via a modified curve number approach, and vadose zone dynamics and ET are determined through the daily Thornthwaite–Mather soil water budget. We selected JoFlow for its low dimensionality and the inclusion of hydrologic processes relevant to our study (Georgakakos, Morris, & Walter, 2018; Knighton, Saia, et al., 2017b; Knighton, Steinschneider, & Walter, et al., 2017a).

Further, Buchanan et al. (2018) found that saturation excess (SE) overland flow prevails over Hortonian flow generation throughout the eastern U.S. and the coastal areas around the Great Lakes. As the dominant surface runoff generation process is homogenous and subject to VSA hydrology across our sample region, it is imperative to utilize a model that numerically describes SE runoff generation, such as JoFlow.

Forcing data supplied to the model consists of latitude (USGS, 2018), daily precipitation and daily minimum and maximum air temperatures (Chen et al., 2008). Unsaturation zone AWC and depth to confining layer at each site were determined from SSURGO (Soil Survey Staff, 2018). We assumed a standard albedo value of 0.23 for non-snow cover conditions. During snowpack accumulation and melt, we modelled surface albedo following Walter et al. (2005).

JoFlow hydrologic model parameters for all catchments (both infested and non-infested) were calibrated using the dynamically dimensioned search (DDS) algorithm (Toolson & Shoemaker, 2007).

We preserved the calibration parameter set for each catchment that maximized the predictive power of the model, as determined by the daily discharge Nash–Sutcliffe efficiency (NSE) coefficient (Nash & Sutcliffe, 1970), following 10,000 iterations of parameter value selection through Monte Carlo sampling within indicated boundaries (Table 2).

The \( S_{\text{min}} \) parameter describes the minimum available catchment storage on a given day, \( S_t \), through the relationship expressed in Equation 3, where \( \theta_t \) is the daily soil water content. \( S_t \) is in turn used to determine the surface runoff response, \( Q_t \), to precipitation, \( P \) (Equation 4).

\[
S_t = S_{\text{min}} + C1(AWC - \theta_t) \quad (3)
\]

\[
Q_t = \frac{P^2}{P - S_t} \quad (4)
\]

The percent impervious parameter allows for some portion of catchment precipitation to be converted directly into discharge independent of \( S_t \).

### Table 2: JoFlow model parameter values and ranges

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Boundary range</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Forest</td>
<td>(0, 1)</td>
<td>Fraction of forest cover within catchment</td>
</tr>
<tr>
<td>( T_p )</td>
<td>(0, 4)</td>
<td>Time to peak (h)</td>
</tr>
<tr>
<td>( l_a )</td>
<td>(0.05, 0.9)</td>
<td>Initial abstraction (mm)</td>
</tr>
<tr>
<td>( PET_{max} )</td>
<td>(2, 10)</td>
<td>Maximum daily PET (mm)</td>
</tr>
<tr>
<td>Percent impervious</td>
<td>(0, 100)</td>
<td>Fraction of impervious cover</td>
</tr>
<tr>
<td>( rec_{coef} )</td>
<td>(0.0001, 0.2)</td>
<td>Baseflow recession coefficient</td>
</tr>
<tr>
<td>( C1 )</td>
<td>(1, 5)</td>
<td>Parameter relating soil moisture to curve number</td>
</tr>
<tr>
<td>( S_{\text{min}} )</td>
<td>(40, 210)</td>
<td>Minimum daily curve number</td>
</tr>
</tbody>
</table>

The \( C1 \) parameter (Equation 3) describes the slope of the relationship between storage and soil water deficit (SWD) predicated on a modification of the curve number approach following Schneiderman et al. (2007). This parameter controls the sensitivity of surface runoff to catchment storage.

We compared CDFs of all calibrated model parameters as estimated for non-infested catchments against those estimated for infested catchments, relying again on the two-sample K-S test. This approach potentially provides further insight into the mechanisms of runoff generation. We note that selection of the sampling ranges of hydrologic parameters is a critical issue in sensitivity analysis, and hydrologic model parameter ranges for this research were selected following other studies utilizing JoFlow (Georgakakos et al., 2018; Knighton, Saia, et al., 2017b; Knighton, Steinschneider, & Walter, et al., 2017a). Further, in order to understand the sensitivity of this analysis to the parameters time-to-peak (\( T_p \), h), a measure of the time between the onset of a precipitation event and discharge peak associated with the event, and percent impervious cover (percent impervious, %), a measure of the proportion of groundcover impervious to water infiltration, we repeated this analysis twice: (a) by increasing the range of \( T_p \) to (0, 100) hours, allowing for longer catchment flow paths (see Figures S8 and S9), and (b) by removing percent impervious from the analysis, forcing runoff processes to occur solely as a saturation-excess process (e.g., Schneiderman et al., 2007) (see Figures S10 and S11).

3 Results

3.1 Relationship between Eastern hemlock condition and hydrologic indices

In general, watershed characteristics did not strongly predict variabilities in hydrologic indices (R-B, BFI, and ETP) (all \( r^2 \) values < 0.45; Table 3). While larger catchments tend to report lower flashiness (Baker et al., 2004), our data yielded only a weak trend in agreement (Figure 3). In addition, flashiness increased and baseflow, as a fraction of total flow, decreased at higher MAP levels. The proportion of precipitation removed by ET was greater in locations where soils had high
saturated hydraulic conductivity and where a higher fraction of total landcover was vegetated.

With the exception of the positive relationship between the percentage of total landcover classified as forest and the ETP ratio, our analysis uncovered tenuous relationships between forest cover, tree condition, and stream discharge characteristics. Our results are similar to the findings of Tan-Soo et al. (2016), which suggested weak relationships between forest cover and hydrologic characteristics. Although others have demonstrated that new and interesting hypotheses can be developed through statistical analysis of hydrologic datasets (e.g., Evaristo & McDonnell, 2017), a purely probabilistic framework approach is possibly too limiting to sufficiently identify a clear signal of change from within the somewhat large variance in the distributions of streamflow characteristics in our case study. Indeed, prior knowledge of the hydrologic functioning of catchments, as captured by physically based models, may be employed to explain some of the inherent variability between catchments.

3.2 | Differences in hydrologic indices between infested and non-infested catchments

We found that infested sites exhibited greater stream flashiness, lower baseflow volumes, and possibly higher surface runoff responses to precipitation than non-infested sites. CDFs of each of the three hydrologic indices across all catchments revealed statistically significant separation by EH condition for both the R-B index and BFI. While we expected infested sites to record lower ET volumes compared to non-infested sites, the distribution of ETP ratio data was not statistically significant (Figure 4).

3.3 | Lumped VSA model calibration, performance, and flow partitioning

JoFlow performed better with data from northeastern states (Figure 5), the region for which the model was designed (Archibald et al., 2014).
Runoff mechanisms from undeveloped areas include SE (i.e. overland runoff generated from the near-stream saturated extent of a catchment) and Hortonian flow (i.e. overland runoff generated when precipitation intensity exceeds soil infiltration capacity). The dominant mechanism is heterogeneously spatially distributed and often coupled across the U.S. (Buchanan et al., 2018). The northeastern U.S., characterized by dense vegetation, a humid climate, and higher soil $K_{\text{SAT}}$ values relative to the Midwest, exhibits runoff generation dominated by SE processes (Archibald et al., 2014; Buchanan et al., 2018; Walter et al., 2003). The 10th percentile, median, and 90th percentile NSE scores across model runs for all sites are 0.15, 0.52, and 0.64, respectively.

Although not statistically significant, we qualitatively observed separation between CDFs of the fraction of simulated surface runoff to total simulated stream discharge for non-infested and infested catchments. The distribution of this data suggested that overland flow was greater at the latter sites, as we expected based on the catchment-scale findings of Kim et al. (2017) (Figure 6, $D_s = 0.3$).

$P = 0.23$). A reduction in stomatal conductance following infestation may yield reduced ET (e.g., Domec et al., 2013), increased catchment storage, and increased water yield.
CDFs of calibrated values for the baseflow recession coefficient (rec_coef) separated significantly along the lines of HWA infestation (Figure 7). This parameter controls the release of catchment-stored water to stream discharge, where lower values indicate longer pathways through groundwater and likely represent streams with more sustained baseflows. Infested sites had higher baseflow recession coefficient values, suggesting shortened groundwater flow pathways.

Similarly, while a qualitative observation of CDFs of all other calibrated parameters suggested hydrological differences between infested and non-infested sites, K-S test results indicated that the data for these two groups may not necessarily be differently distributed (Figure 7). Median values for these parameters evinced slight differences between parameter set distributions (Table S1). These discrepancies were likely a result of our sample size; indeed, more data would generate greater confidence in our test statistic (Ds) results, and we encourage future modelling exercises on this theme to further expand the number of sites studied. While we cannot interpret the remaining parameters with as high a level of confidence as we did the rec_coef parameter, a comparison of the positions of CDFs for these other calibrations offers some insights about watershed processes that could possibly contribute to differences between infested and non-infested watersheds.

The Se_min values for non-infested sites had a higher median (Table S1) than infested sites, which is reflected in the CDF distributions for this parameter (Figure 7). This may capture differences such as canopy storage provided by EH needles (Guswa & Spence, 2012). This parameter represents the lower bound on catchment water storage and is indicative of increased runoff. Lower calibrated Se_min values may reflect dilation of the VSA as a hydrologic response of EH loss, as suggested by Ford and Vose (2007).

We observed higher values for the percent impervious parameter at infested sites compared to non-infested sites, but only near the median of each distribution. As with the Se_min distribution, this calibration result would generally reflect increased runoff and VSA dilation.

Assuming SE flow dominated hydrology across our region of study, higher median values of the C1 parameter for infested sites (Table S1) indicate that surface runoff potentials at these locations have a higher sensitivity to catchment storage than at non-infested sites.

Finally, while not statistically significant as determined by the K-S test, a higher median of the PETmax parameter at non-infested catchments implies that discharge characteristics are better explained when allowing for higher ET demand, in tenuous agreement with our multivariate regression analysis results.

To further assess the influence of HWA infestation on hydrologic data, we conducted a time-series analysis by calibrating JoFlow to two paired three-year periods at a subset of catchments, pre- and post-infestation (data not shown). The results were broadly inconclusive, with only the distribution of C1 separating significantly (paired samples t-test P < 0.05), reporting a median increase in the post-infestation calibration. This indicates that surface runoff potentials exhibit greater sensitivity to catchment storage after infestation, relative to before infestation.

4 | DISCUSSION

4.1 | Regional influence of Eastern hemlock infestation on catchment hydrology

In concert, CDFs of optimal parameter calibrations across our dataset provide evidence that sites located within HWA range exhibit shorter catchment flow-paths, lower catchment storage at median values, and possibly VSA dilation, suggesting more prominent surface runoff responses to precipitation. These results support trends in our flow
indices showing that catchments within HWA range exhibited greater stream flashiness and lower baseflows than non-infested catchments. However, EH infestation was spatially correlated with higher MAP and lower soil K\textsubscript{SAT} and AWC (Figure 2), potentially confounding our conclusions about the hydrological importance of EH condition. Further, the poleward movement of the HWA yields a latitudinal gradient of EH infestation (Figure 1), possibly suggesting climate and local weather as explanations for the differences in hydrologic pathways between the two groupings.

The hydrologic simulation approach allowed us to control for varied soil properties and climate to an extent, and still yielded evidence of the hydrologic effects of HWA infestation. As determined by the K-S test, differences in baseflow recession emerged as the most significant control on hydrological variation between infested and non-infested catchments (Figure 7). Considering that EH generally populates the riparian zone, this result agrees with previous research results that vegetation growing adjacent to streams relies considerably on groundwater flows (Dunford & Fletcher, 1947).

We observed that the relationship between infestation and runoff was not related to changes in catchment ET. Both the ETP ratio (Figure 4) and the simulated PETmax parameter (Figure 7) did not establish ET as a statistically significant or physically prominent driver of observed bifurcations in hydrologic trends across infested and non-infested sites. That the distributions of ETP ratios between our two categories of watersheds were similar, and that mean precipitation was greater at infested sites, compared to non-infested sites, suggests that ET is greater at infested sites. This result is interesting as previous studies have identified decreased ET as one of the dominant results of EH infestation (Figure 1), and that ET is greater at infested sites. This result indicates that while the reduction in forest ET as a result of EH decline may not necessarily transform precipitation partitioning to the extent of increasing runoff (as in Brantley et al., 2014), reduced canopy interception and VSA dilation could possibly drive increased overland flow. This is further supported by the higher Se\textsubscript{rain} parameter distributions at non-infested sites (Figure 7), which appreciably alter the canopy storage effects. The BFI distribution (Figure 4) and the modelled rec_coef outputs (Figure 7) also illuminate differences in catchment storage dynamics along EH condition, with non-infested sites feeding a greater proportion of their total stream discharge through sloping groundwater input.

Collectively, our analyses identified variance in components of infiltration-runoff partitioning between non-infested and infested sites, such as baseflow, VSA hydrology, and flow responses to throughfall, as the primary contributors to hydrological differences observed (e.g., Ford & Vose, 2007; Stadler et al., 2005). We note that while JoFlow resolves some critical climatological and edaphic factors, we cannot account for ecological succession effects, which invariably modulate such partitioning over time. Guswa and Spence (2012), for example, found that while EH canopies intercepted a greater volume of precipitation relative to deciduous species, the latter more than accounted for these differences with higher transpiration during the growing season.

### 4.2 Broader implications of Eastern hemlock decline

Sudden defoliation of any dominant tree species may initiate a cascade of ecosystem transformations (e.g., Ellison et al., 2005), and studies examining the hydrologic consequences of the loss of other foundation species broadly align with our research in their results.

In wetland environments, for example Scots pine (Pinus sylvestris) and downy birch (Betula pubescens) serve a foundational role, and have been observed to draw down the water table through a combination of precipitation interception and transpiration, balancing the water relations of Sphagnum, which promote submergence (e.g., Frankl & Schmeidl, 2000; Rutter, 1963). Among other cases where foundation tree species are threatened by invasive insect outbreaks, bark beetle-related mortality of pines has been linked to increased surface runoff, baseflows, and annual catchment water yield (Bearup, Maxwell, Clow, & Mccray, 2014; Behlalmy, 1974; Hubbard, Rhoades, Elder, & Negron, 2013; Potts, 1984). The decline of white
Another important point of support for the thesis that HWA outbreak will initiate marked changes in catchment flows, at least in certain seasons, is that ecological succession is expected to shift forest demography in favour of deciduous vegetation. Previous studies have found that the former forest cover type more effectively tempers the conversion of extreme precipitation events to high streamflow, owing in part to its considerably higher rainfall interception relative to hardwood vegetation, as discussed earlier (Ford et al., 2011; Helvey, 1967; Kelly, Mcguire, Miniat, & Vose, 2016). As a foundation species, EH exerts considerable influence in maintaining a stable local microclimate and regulating water fluxes, a set of unique contributions that cannot be approximated by other conifers or ecological successors (e.g., Daley et al., 2007; Ford & Vose, 2007). Field examinations have revealed that frequently the pioneer species following EH mortality in northeastern plots is the early successional black birch (Betula lenta), whereas in the southern plots it is red maple (Acer rubrum), both of which species differ from EH in water use patterns and precipitation interception (Brantley et al., 2013; Daley et al., 2007; Ford et al., 2011; Orwig & Foster, 1998). Research predicts this transition will reduce warm season flows to potentially unsustainable volumes and increase winter discharge (Brantley et al., 2013; Daley et al., 2007). In the Appalachian region, the encroachment of rosebay rhododendron (Rhododendron maximum) following EH decline is expected to moderate peakflow increases attributable to extreme precipitation (Brantley et al., 2014). However, this evergreen understory shrub is less common in higher latitude forests, suggesting that cool season streamflow response to storm events may be more pronounced in the Northeast once EH is replaced by deciduous tree species (Knighton, Coneelly, & Walter, 2019a).

It is further important to consider climatic trends occurring in parallel to changes in forest structure. The frequency and intensity of extreme precipitation and flooding events in the northeastern U.S. have been increasing (DeGaetano, 2009; Kunkel, 2013; Armstrong, Collins, & Snyder, 2014; Frei, Kunkel, & Matonse, 2015; Huang, Winter, Osterberg, Horton, & Beckage, 2017), driven by a phase shift of the Atlantic Multidecadal Oscillation and concomitant increases in sea-surface temperatures and the magnitude of tropical cyclones in the region (Huang, Winter, & Osterberg, 2018). Frei et al. (2015) found that both extreme streamflow and extreme precipitation events, a product of synoptic scale meteorological processes, are increasing in frequency in the warm season. This joint-trend diverged through the year, however, with extreme precipitation events consistently larger in the warm season and extreme streamflow events exhibiting greater intensity in the cold season (November–May). Further, recent regional increases in precipitation, driven primarily by summer and fall season meteorology, could be attributed to few extreme precipitation events resulting from the increased magnitude of tropical cyclones (Huang et al., 2017, 2018).

The effects of diminished ET and canopy interception (Ford & Vose, 2007; Lustenhouwer et al., 2012) in the cold season following EH loss are likely to intersect with this meteorological transition and further enhance flooding in the growing season (e.g., Knighton, Coneelly, & Walter, 2019a). The loss of the insulating properties of EH...
canopy may also contribute to hydrologic regime instability in the form of snowpack melt (Lustenhouwer et al., 2012).

EH decline can also potentially complicate human use of water resources. For example, surveys of HWA progression through the Catskill Mountains find substantial EH decline concentrated in a few areas (e.g., Hanavan, Pontius, & Hallett, 2015). The Hudson Valley region, which is prone to riverine flooding, supplies drinking water to residents of New York City (e.g., Saleh, Ramaswamy, Georgas, Blumberg, & Pullen, 2016). Extreme precipitation-related changes to streamflow volume, timing, and quality can in turn challenge potable water access. Also, streams draining EH stands have been found to host considerably higher trophic diversity of both macroinvertebrates and freshwater fish than those in purely hardwood forests (e.g., Ross et al., 2003; Snyder et al., 2002). Beyond forest managers concerned about invasive species and the extirpation of a foundation species, EH health, insofar as it impacts recreational and drinking water use, may be of interest to a diverse range of stakeholders.

### 4.3 Statistical analysis in hydrology

In this study, we considered three methodological approaches for evaluation of our research questions: (1) a parametric statistical test, (2) a non-parametric statistical test, and (3) comparison of calibrated hydrologic model parameter values. The parametric statistical test did not identify a convincing influence of EH infestation on several catchment discharge metrics. In contrast, the non-parametric (K-S test) and the hydrologic modelling-based approaches both suggest a clear influence of HWA infestation on catchment hydrology.

The multivariate regressions helped control for climate (e.g., MAP, MAT) and land surface variables (e.g., K\textsubscript{SAT}) that would have some physical influence on catchment hydrology. Yet, our results yielded the insight that regression analysis is perhaps greatly limited by forcing linear relationships between regressors and the response variable, and thus neglecting to account for ecohydrological relationships between explanatory variables and catchment discharge and increasing the risk of multicollinearity. JoFlow avoided these pitfalls through a mechanistic framework.

Further, inherent structural differences in the analyses may favour the precision of simulation over statistical modelling results in the context of our study. While inputs provided to JoFlow include a dynamic record of hydrologic and meteorological data, with most static parameters such as percent impervious calibrated to values that best align with observed hydro-physical trends (i.e. water content, albedo, and depth to confining layer), inputs to our regression analysis are either all static point data or long-term average values. For example, interrelated catchment landcover characteristics such as percent forested cover, K\textsubscript{SAT} and AWC, have likely changed over time, confounding our effort to explore a clear relationship between HWA status and stream flashiness. For this reason, future statistically based investigations that utilize dynamic regression models integrating time-series data may yield a more promising methodology.

We found that mechanistic hydrologic models used in concert with probabilistic analysis provided a clear path towards circumventing some of the challenges of purely probabilistic analysis. Physically based models offer a convenient framework to leverage knowledge of well-described physical processes (e.g., mass balance, energy balance of the forest canopy, and snowmelt dynamics) and extract additional information contained in the time-order of observations, without over-emphasizing results that may be attributed to autocorrelated observations.

### 5 CONCLUSIONS

We investigated the relationship between EH condition and various ecohydrologic factors that determine stream discharge characteristics. HWA presence has expanded throughout much of coastal eastern and mid-Atlantic U.S., posing a veritable challenge to the continued health of EH stands and possibly transforming the hydrologic and biogeochemical paradigms they modulate. We employed both statistical and simulation modelling analyses for increased methodological rigour and to evaluate similarities and differences in the outcomes of both approaches. While our regional-scale multiple regression analysis did not identify infestation as a statistically significant predictor of hydrologic regime, a non-parametric test suggested that catchments within HWA range exhibit higher flashiness and lower baseflows. We share indeterminacy of statistical analysis outcomes with other meta-analyses examining the relationship between land cover characteristics and hydrologic regime, potentially indicating that a purely statistical methodology may not capture dynamic physical ecohydrologic relationships. A semi-physically based hydrologic simulation model further elucidated that sites within HWA range exhibit shorter flow pathways and likely lower baseflows, as well as lower catchment storage and higher sensitivity of storage to SWD. Through a coarse, regional scale analysis, we find that HWA presence modifies a number of ecohydrological characteristics and precipitation partitioning between groundwater flows and surface runoff, and that a semi-physically based simulation illuminates trends not captured by a simple regression analysis.

### ACKNOWLEDGMENTS

This research was supported by USDA Forest Service Project Number 16-CR-11092000-001 to the NYS Hemlock Initiative.

### DATA AVAILABILITY STATEMENT

Data sharing is not applicable to this article as no new data were created or analyzed in this study.

### ORCID

Kanishka Singh  https://orcid.org/0000-0001-8445-9545

### REFERENCES


**SUPPORTING INFORMATION**

Additional supporting information may be found online in the Supporting Information section at the end of this article.

**How to cite this article:** Singh K, Knighton J, Whitmore M, Walter MT, Lassoie JP. Simulation and statistical modelling approaches to investigate hydrologic regime transformations following Eastern hemlock decline. *Hydrological Processes*. 2020;1–15. [https://doi.org/10.1002/hyp.13666](https://doi.org/10.1002/hyp.13666)