RESEARCH ARTICLE

Scenario Planning Management Actions to Restore Cold Water Stream Habitat: Comparing Mechanistic and Statistical Modeling Approaches

M. R. Fuller[1] | N. E. Detenbeck[2] | P. Leinenbach[3] | R. Labiosa[3] | D. Isaak[4]

[1]Oak Ridge Institute for Science and Education (ORISE) Postdoc, Atlantic Coastal Environmental Sciences Division, U.S. Environmental Protection Agency, Narragansett, Rhode Island, USA | [2]Atlantic Coastal Environmental Sciences Division, U.S. Environmental Protection Agency, Narragansett, Rhode Island, USA | [3]Region 10, U.S. Environmental Protection Agency, Seattle, WA, USA | [4]Rocky Mountain Research Station, U.S. Forest Service, Boise, Idaho, USA

Correspondence: N. E. Detenbeck (detenbeck.naomi@epa.gov)

Received: 22 January 2024 | **Revised:** 13 August 2024 | **Accepted:** 25 August 2024

Funding: This work was supported by the US EPA Office of Research and Development (ORD) via a Region 10 Regional Applied Research Effort project (interagency agreements with the U.S. Forest Service (DW12924479) and U.S. Department of Energy support (DW 92429801-5 and DW 92525701-1)). US EPA ORD contribution number is ORD-046319.

Keywords: cold water habitat | fish habitat | mechanistic model | river restoration | scenario planning | spatial stream network model | temperature total maximum daily loads | thermal regime

ABSTRACT

Under the United States Clean Water Act, states are required to periodically assess state waters to determine compliance with water quality criteria (including temperature) and then to develop total maximum daily loads (TMDLs) for impaired waters as necessary to bring them into compliance. We compared the performance of mechanistic stream temperature models (HeatSource, QUAL2K, and QUAL2Kw) applied to the mainstem of three TMDL watersheds (Middle Fork John Day, OR; Wind River, WA; South Fork Nooksack, WA) with that of spatial stream network (SSN) models applied to the full watersheds and used these to evaluate the potential effectiveness of restoration strategies. SSN models performed well with slightly lesser accuracy (RMSE = 0.47–0.87) for mainstem predictions than mechanistic models (RMSE = 0.4) but provided additional benefits to inform management, including information on spatial and temporal heterogeneity of restoration effectiveness throughout the watershed. Of the four scenarios considered (restoration of riparian zones to potential natural vegetation, channel narrowing, increasing flow by restricting irrigation withdrawals, and combined applications), riparian zone restoration was consistently the most effective in reducing temperatures at the outlet, mainstem, and throughout the watersheds. Predicted restoration effectiveness for thermal regimes varied significantly both within and among watersheds. A focus on water quality criteria exceedance only at the watershed outlet or along the mainstem reach can obscure knowledge of restoration potential for fish habitat in tributaries and headwaters, potential for creation of thermal refuge areas along the mainstem critical for maintaining migration corridors, and thermal regime heterogeneity across space and time.

1 | Introduction

Watershed-scale stream temperature models may be used to assess compliance with temperature criteria designed to protect cold water fisheries throughout a stream network. Under the United States Clean Water Act, states are required to periodically assess state waters to determine compliance with applicable water quality criteria (WQC) (including temperature criteria, where applicable) and then to develop total maximum daily loads (TMDLs) for impaired waters as necessary to bring them back into compliance. Nationally, states of Oregon and Washington have the most 303(d) listed stream segments for temperature impairments, and most of these stream segments provide habitat for temperature sensitive salmonids. Accordingly, state regulatory agencies have been developing temperature TMDL assessments on these stream segments over the past several decades.

Temperature TMDLs typically differ from those for most toxic pollutants as numerous natural factors affect the thermal regime of water bodies in addition to anthropogenic sources (Poole and Berman 2001). Thus, regulators rely heavily on models to predict thermal regimes in streams and rivers under different management scenarios and to determine actions needed to bring waters back into compliance. Mechanistic models such as QUAL2K, QUAL2Kw, and HeatSource are commonly applied to mainstem reaches to evaluate compliance during the hottest days of summer under different management strategies (Kennedy and Butcher 2012; Middle Fork IMW Working Group 2017; Pelletier 2002). QUAL2K is a one-dimensional river and stream water quality model with steady state hydraulics, nonuniform steady flow, and diel heat budget/water-quality kinetics. It requires user input of hourly data for individual river segments via an Excel interface with multiple spreadsheets coded in VBA. The heat balance in QUAL2K takes into account heat transfers from adjacent elements, loads, withdrawals, the atmosphere, and sediments. QUAL2Kw is similar and provides continuous simulation with time-varying boundary conditions for periods of up to 1 year with the option to use repeating diel conditions and with either steady or nonsteady flows. HeatSource is used to simulate stream hydrology and thermodynamics over a single day with defined reaches of 100 ft in length and requires 25–40 parameters for operation.

These mechanistic models require extensive parameterization, and thus are typically only applied to mainstem reaches over limited periods. However, restoration of fish populations throughout watersheds requires consideration of spatial and temporal variation of thermal regimes in different habitats used during different phases of fish life cycles (Steel et al. 2018). These varying requirements are reflected in the assigned designated uses for streams and rivers, for example, for spawning, rearing, and migration for different species, as well as core cold water habitat. Spatial stream network (SSN) modeling (Peterson and Ver Hoef 2010) is one method used to characterize thermal regimes across entire river networks (Detenbeck et al. 2016; Isaak et al. 2017) that captures essential fish habitats (beyond the mainstem or network outlet) for designated fish uses (Fuller et al. 2022). SSN temperature models are statistical models, typically with less than a dozen parameters representing different water or heat sources and modifiers, and which taken into account spatial autocorrelation throughout a stream network.

SSN temperature models have been used to explore fine spatial resolution of thermal regimes across river networks at broad spatial extents (e.g., western half of and New England regions of continental USA) (Detenbeck et al. 2016; Isaak et al. 2017). Predictions from these types of SSN temperature models have also been used to evaluate the risk for fish habitat given future shifts in climate and how well riparian shade restoration can mitigate future warming (Fuller et al. 2022). Additional watershed management options could be investigated with these types of SSN models and these results can be compared with results from mechanistic temperature models previously developed to support temperature TMDL efforts.

The objectives of our research are to (1) compare mechanistic with SSN stream models, (2) apply these models to different restoration scenarios, and (3) evaluate the predictions of the two models with respect to several water quality temperature criteria that are applicable to protect aquatic life uses and water bodies in Oregon and Washington. Our project is structured into two parts. In part one, we used published mechanistic and SSN temperature models to predict stream temperatures under different management action scenarios and then we compared the models' performance (mechanistic versus SSN) for the mainstem. In addition to the model performance comparison, we discuss how SSN models can be used alongside traditional mechanistic models used in TMDLs to better characterize stream network thermal regimes. In part two, we focus on just the SSN model predictions to describe the response of stream temperature under different management actions in two ways: (1) across different spatial extents (river outlet, mainstem reach, and entire network) and (2) to evaluate how protective each action may be for designated fish habitat uses as defined by state agencies across our study systems.

We focused on three watersheds for which temperature TMDLs recently had been developed: the Middle Fork John Day (MFJD) in the state of Oregon (OR; Pelletier 2002), and the Wind River (WR) and South Fork Nooksack River (SFNR), both in the state of Washington (WA; Kennedy and Butcher 2012; Middle Fork IMW Working Group 2017). SSN models have recently been developed for each of these systems as well to predict thermal regimes (Fuller et al. 2023a, 2023b) throughout these watersheds. These SSN models refined the original NorWeST models developed for larger units (Isaak et al. 2017) by improving the characterization of shade at various scales and by incorporating variables related to potential restoration actions. We compared the performance of SSN models with mechanistic models for the mainstem both under baseline conditions and for management scenarios including riparian restoration, channel narrowing, reduction in water withdrawals for irrigation, as well as combined management strategies. Information on the effectiveness of various management strategies is useful for TMDLs developed to meet WQC for stream temperature.

2 | Materials and Methods

2.1 | Study Areas

Our scenario planning was conducted in three river networks within the Pacific Northwest, USA (Figure 1). These networks are the MFJD River, SFNR, and the WR. All three study basins have TMDLs for temperature (Butcher et al. 2010; Kennedy and Butcher 2012; Pelletier 2002). These temperature TMDLs include either a water quality management plan or provide strategies for managing the temperature load to meet specific water quality standards effective under the Clean Water

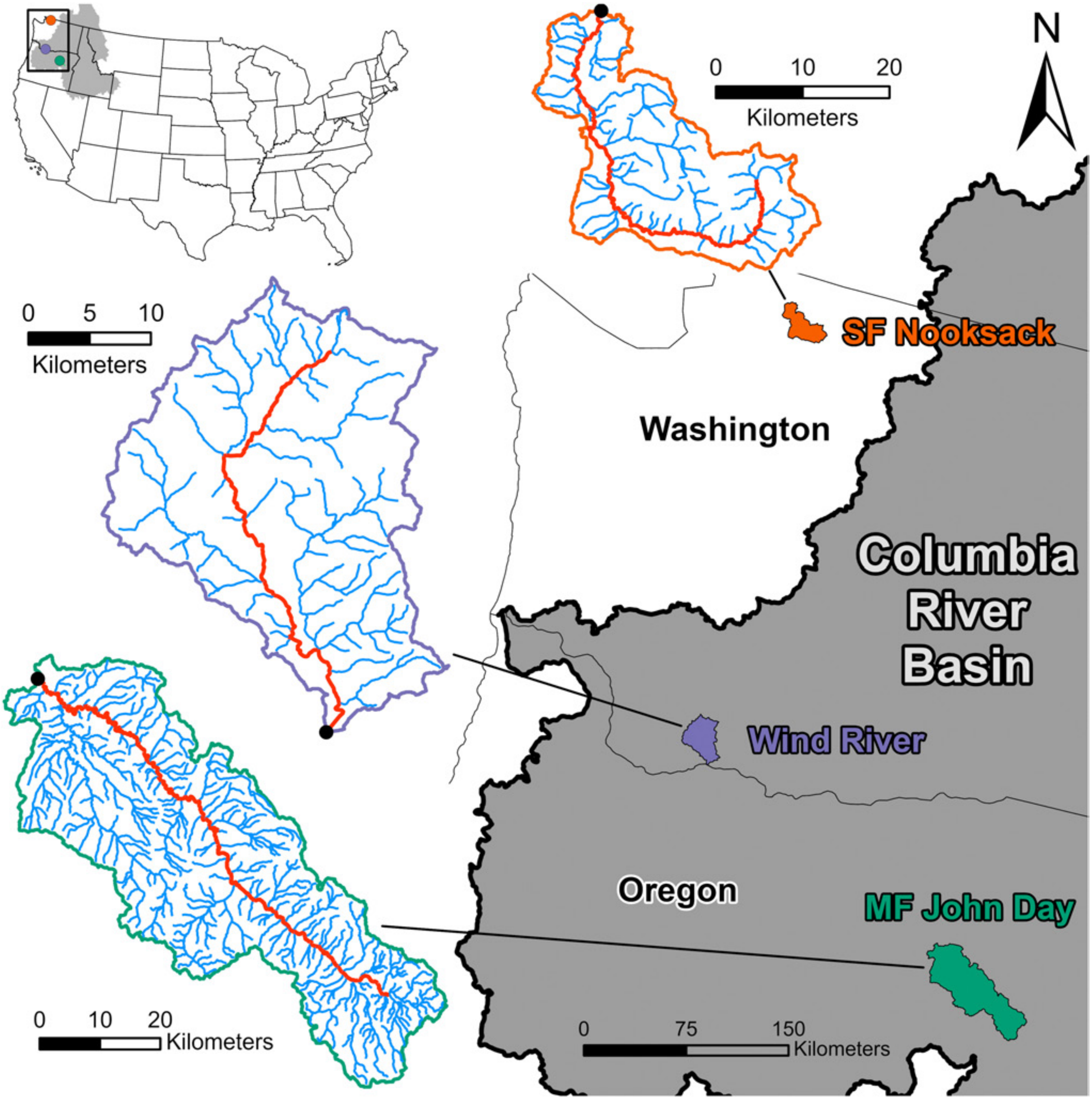

FIGURE 1 | Three study basins in the Pacific Northwest, USA used for scenario planning cold-water habitat restoration. River networks for each study area. Note the outlet (black dot) and distinguish between the mainstem reach (wider red reaches) and tributary reaches (narrower blue). The Columbia River Basin is provided for spatial context within the region and to the conterminous U.S. (CONUS) inset map. [Color figure can be viewed at wileyonlinelibrary.com]

Act for each state's waterbody as the target of each TMDL. These management activities were the basis of our modeling scenarios in both the mechanistic and SSN model predictions described below.

2.1.1 | MFJD River

The MFJD in Oregon State, USA (Figure 1) is in an arid region with a drainage area of 2051 km^2, river network length of 1581 km, and average discharge of 7.5 m^3/s. Basin elevation range is 666–2477 m with an overall mean of 1378 m, and mean annual precipitation range is 350–1000 mm. Most precipitation falls between November and March as snow. Mean annual air temperatures across the basin range between 4.4 and 9.1°C with an overall mean of 6.4°C. The MFJD TMDL for temperature identifies riparian shade, flow, and channel narrowing as important areas for restoration activities to focus on. The mainstem channel width ranges from 7 to 20 m with a mean of 15 m. Irrigated cropland from MODIS-based estimates covers an area of only 2.5 km^2, or 0.12% of watershed area, clustered in three small areas. Current riparian shade in this basin as determined using the "shade.xls" stream shade model (https://ecology.wa.gov/research-data/data-resources/models-spreadsheets/modeling-the-environment/models-tools-for-tmdls) averages 58.5% (range = 7.3%–97.6%). More generally, the MFJD basin is 54% forested with little area classified under the 2016 National Land Cover Database as urban (1%) (Yang et al. 2018). Grassland and shrubland were also well represented land covers with 26% and 11% coverage, respectively. Little (0.1%) cultivated agriculture is present in the MFJD.

2.1.2 | South Fork Nooksack River

The SFNR is a coastal watershed in northern Washington State, USA and has a drainage area of 482 km^2. The river network has a length of 317 km and mean annual discharge of 29 m^3/s. The basin elevation range is 11–3284 m with a mean of 1049 m. Mean annual precipitation ranges from 1125 to 3500 mm and the mean annual air temperature ranges from 1.4–10.2°C with a mean of 6.3°C (U.S. Geological Survey 2012). Base flow during the summer is supported by a small amount of melting ice but mainly groundwater inflow. During other seasons, flows are provided by a combination of rainfall and snowmelt. The temperature TMDL in the SFNR primarily aims to reduce thermal loading into the river (Kennedy and Butcher 2012). The SFNR is predominantly forested (77%) with some shrubland (8%), developed land (4%), and wetlands (2%). Current riparian shade in this basin (based on the "shade.xls" stream shade model) averages 66.9% (range = 7.4%–99.5%).

2.1.3 | Wind River

The WR flows into the Columbia River mainstem from Washington State, USA (Figure 1). The WR watershed covers 582 km^2 has a network length of 342 km and associated mean discharge of 34.24 m^3/s. The largest stream flows typically occur in response to rain-on-snow events, while low flows are maintained by a combination of late season snowmelt and areas of water retention or recharge. The study area elevation ranges from 19 to 3746 m with a mean elevation of 779 m. Mean annual precipitation range is 275–3250 mm and mean annual air temperature range is 1.5–11.6°C with a mean of 7.5°C (U.S. Geological Survey 2012). The temperature TMDL for the WR explores issues that include temperature impairment from riparian shade loss and summer flow reduction from instream withdrawals (Pelletier 2002). The WR watershed is predominantly forested (90%) with lesser coverage of shrubs (4%) and developed land (5%). Current riparian shade in this basin based on the "shade.xls" stream shade model averages 60.5% (range = 7.2–99.0).

2.2 | Mechanistic Models

2.2.1 | Model Description and Set Up

We ran available calibrated mechanistic models developed as part of past TMDL efforts within the three study basins (Butcher et al. 2010; Kennedy and Butcher 2012; Pelletier 2002). These mechanistic models included HeatSource (Boyd and Kasper 2003) in the MFJD, QUAL2Kw (Pelletier and Chapra 2008) in the SFNR and Qual2K (Chapra, Pelletier, and Tao 2008) in the WR. Our mechanistic model runs generated predictions for a subset of the management scenarios described in Section 2.4, which allowed us to directly compare their output to SSN model predictions. These calibrated mechanistic models were only built for the mainstem of each river network, which limits the spatial extent of our comparisons to those mainstem reaches. The HeatSource model can be used to predict hourly temperature response over multiple days, while the Qual2k and Qual2kw models report daily minimum, daily average, and daily maximum temperature response for a single day. The SSN models were fit to mean monthly temperatures (see Section 2.3) and therefore could not make predictions at a submonthly temporal scale like the mechanistic models. Therefore, HeatSource mechanistic model predictions for the MFJD were initially made at the daily time-step (mean daily temperature for each reach along the mainstem) for each day in August and then averaged across the month to generate a mean August temperature prediction. Predicted daily average temperatures for the SFNR and Wind R. basins using the Qual2kw and Qual2k models, respectively, were compared directly with predicted SSN average August values. Further details on mechanistic model inputs, boundary conditions, and manipulations to address management activity scenarios (described in Section 2.4) are provided in the Supporting Information.

2.2.2 | Comparing Mechanistic and SSN Models

Comparing the mechanistic and SSN model predictions demonstrates whether our SSN model manipulations were responding in a similar manner to the mechanistic models in each system. The mechanistic model predictions were generated for the calibration year of the mechanistic model in each basin (Table 1) and SSN model predictions were made to match those associated calibration years. Three error statistics (mean error, mean absolute error, and root mean square error [RMSE]) were calculated to determine how well the mechanistic and SSN model predictions diverged or matched each other along the mainstem reaches of each basin as well as with a small observed validation dataset.

In addition to comparing the two modeling approaches during the mechanistic model's calibration years, we also compared predictions between the two model types that were based on management activities proposed in each basin's temperature TMDL (Table 2). Changes in channel width were simulated by reducing estimates of channel width by 10%–50%, cessation of irrigation was simulated by changing estimated water losses from crop evapotranspiration (ET) to zero, and restoration of riparian zones was simulated by modifying riparian vegetation composition from current to potential vegetation conditions as attributed spatially by the USFS Landfire Environmental Site Potential (ESP) dataset (https://www.landfire.gov/), and assigning vegetation height and canopy cover target conditions for these ESP groups to conditions reported in TMDL documents (e.g., Appendix C in Butcher et al. 2010). Current vegetation height and canopy cover conditions were derived from two Landfire datasets: Existing Vegetation Height (EVH) and

TABLE 1 | Descriptions for hydrologic and climate scenarios (and their abbreviated IDs) for fish habitat use analyses.

Scenario (ID)	Description	MFJD	SFNR	WR
Calibration (Cal)	The year a given mechanistic model was calibrated	2002	2007	1999
Historical (Hist)	Historical period average covering 1990–2015	'90–'15	'90–'15	'90–'15
Dry year (Dry)	Driest year (from historical period) based on PDSI	2007	2015	2015
Wet year (Wet)	Wettest year (from historical period) based on PDSI	1993	1997	2011
Cool year (Cool)	Coolest water temperature year from historical period[a]	2006	2008	2000
Warm year (Warm)	Warmest water temperature year from historical period[a]	2001	2003	2004

Note: Years noted for each study basin indicate the period used to reflect each scenario type.
Abbreviation: PDSI = Palmer Drought Severity Index.
[a]Based on mean observation temperatures for sites that had records covering majority of historical period.

Existing Vegetation Cover (EVC), respectively. Stream shade throughout these basins was calculated for derived current and potential vegetation conditions using the "shade.xls" model available at ecology.wa.gov/Research-Data/Data-resources/Models-spreadsheets/Modeling-the-environment/Models-tools-for-TMDLs. Channel width scenarios were based on OR DEQ's estimates of the range of historic changes in channel widths in the MFJD (Butcher et al. 2010). For mechanistic models, irrigation reduction scenarios were modeled by removing points of diversion or ditch inputs and adjusting tributary flows to estimates of natural flow (Butcher et al. 2010). For the SSN models, we estimated potential crop irrigation withdrawals based on ET water loss from known irrigated crop land use coverages (Pervez and Brown 2010). This comparison provided a sensitivity assessment for how each model type responds to various modifications of model inputs and when generating predictions for management activities. These model sensitivities/differences allow managers to make more informed decisions (depending on which model types are being used) on activities to undertake when restoring or managing each of these systems with different conservation goals and budgets.

2.3 | SSN Models

We used SSN models (Peterson and Ver Hoef 2010) to build and quantify our scenario planning exercises and comparisons (Peterson, Cumming, and Carpenter 2003) for cold-water thermal habitat restoration. An SSN model is a type of linear mixed model that specifically addresses spatial autocorrelation associated with river network branching structure and connectivity. Addressing this form of network structure and connectivity is especially important for understanding response variables with strong relationships to flow (Erős and Lowe 2019), such as water temperature (Steel, Sowder, and Peterson 2016; Isaak et al. 2014).

2.3.1 | SSN Model Suites

The SSN temperature models we used are published sets of SSN models for each study area (Fuller et al. 2023a, 2023b; Tables S1 and S2). In each study area, a suite of SSN temperature models (hereafter "suite of models" or "model suite") was selected to generate the most accurate temperature predictions across each study network, while maintaining two management-related statistical covariates (riparian shade covariate and irrigated cropland water losses from ET covariate over the growing season [Table 2]) that could be manipulated for comparison of management outcomes for the three basins (i.e., MFJD, SFNR, and WR) during three different assessment periods (i.e., May, August, and September). Most analyses for TMDLs focus on conditions and mitigation during August when peak temperatures tend to occur as WQC are typically based on the 7 day average daily maximum (7DADM), but changes in thermal regime during spring and fall could also be important as different aquatic organism life cycle stages and processes may be affected during these periods (Steel et al. 2017).

Other covariates contained in the model suites varied among study areas, but included base flow recession coefficient, main channel slope and reach slope, latitude, mean annual precipitation, mean monthly air temperature, mean monthly discharge, proportion of catchment area covered by glaciers or lakes, reach elevation, and upstream catchment area (Fuller et al. 2023a; Tables S1 and S2). Interaction terms were also present between slope and covariates representing surface–subsurface exchange processes such as hydraulic conductivity of surficial sediments or base flow index.

2.3.2 | SSN Model Averaging

We used model averaging (Burnham and Anderson 2002) to generate management scenario predictions for comparison within each study network (Fuller et al. 2023a). There are benefits to using model averaging versus a single best model from a selection process. First, model averaging incorporates multiple working hypotheses from the different models being included in the averaging process (Grueber et al. 2011). Second, using model averaging reduces the uncertainty associated with selecting just one model from a model selection procedure like step-wise regression (Whittingham et al. 2006). Third, model averaging approaches provide more robust predictions that are more stable across datasets (Burnham and Anderson 2002).

We used Burnham and Anderson's (2002) model averaging methods and describe them as a two-step process for generating model averaged predictions for our study. The first step in the model averaging process calculates a model weight for

TABLE 2 | Definition of scenarios and comparison between SSN and mechanistic model predictions for each scenario in August.

Scenario Name	Scenario ID	Description	MFJD[a] (August 2002)	SFNR[b] (August 2007)	WR[c] (August 1999)
Baseline	A	Current estimate of temperature based on recent climate drivers (1990–2015)			
Riparian shade restoration	B	Used potential vegetation height and density across network for shade	ME = 0.578°C MAE = 0.702°C RMSE = 0.812°C	ME = 0.529°C MAE = 0.529°C RMSE = 0.568°C	ME = 0.272°C MAE = 0.365°C RMSE = 0.436°C
Channel narrowing (10%)	C1	Reduced the mainstem channel width by 10% (affects shade estimates)	Not compared	Not evaluated	Not evaluated
Channel narrowing (20%)	C2	Reduced the mainstem channel width by 20% (affects shade estimates)	Not compared	Not evaluated	Not evaluated
Channel narrowing (30%)	C3	Reduced the mainstem channel width by 30% (affects shade estimates)	Not compared	Not evaluated	Not evaluated
Channel narrowing (40%)	C4	Reduced the mainstem channel width by 40% (affects shade estimates)	Not compared	Not evaluated	Not evaluated
Channel narrowing (50%)	C5	Reduced the mainstem channel width by 50% (affects shade estimates)	ME = 0.041°C MAE = 0.136°C RMSE = 0.167°C	Not evaluated	Not evaluated
Crop water loss	D	Set estimated irrigated cropland water loss from evapotranspiration to zero	ME = 0.003°C MAE = 0.007°C RMSE = 0.010°C	ME = −0.025°C MAE = 0.048°C RMSE = 0.046°C	ME = −0.0002°C MAE = 0.0003°C RMSE = 0.0009°C
Cumulative effects	E	Combined scenarios in each study area. MFJD: B + C5 + D; SFNR and WR: B + D	ME = 0.877°C MAE = 1.076°C RMSE = 1.227°C	ME = 0.515°C MAE = 0.515°C RMSE = 0.550°C	ME = 0.270°C MAE = 0.364°C RMSE = 0.434°C

Note: The year matches the period for which the mechanistic model was calibrated. SSN predictions for this comparison were made for the corresponding calibration year to best match those of the mechanistic model in each basin. See Supporting Information for more details on this comparison.
Abbreviations: MAE = mean absolute error, ME = mean error, RMSE = root mean square error.
[a]Middle Fork John Day, HeatSource (Boyd and Kasper 2003).
[b]South Fork Nooksack River, Qual2kw (Pelletier and Chapra 2008).
[c]Wind River, Qual2k (Chapra, Pelletier, and Tao 2008).

each model in the suite of models based on the model's Aikake Information Criterion (AIC) value. These AIC weights (w_i) for a given model (i) are calculated as follows:

$$w_i = \frac{\exp\left(-\frac{1}{2}\Delta_i\right)}{\sum_{r=1}^{R} \exp\left(-\frac{1}{2}\Delta_r\right)}$$

where w_i is the AIC weight for the ith model relative to the model with the lowest AIC value in the suite of models, Δ_i is the difference in AIC values between model i and the model with the lowest AIC value in the suite of models, R is the total number of models in the suite of models, the numerator of the AIC weight equation $\left(\exp\left(-\frac{1}{2}\Delta_i\right)\right)$ is the likelihood of model i, and the denominator $\left(\sum_{r=1}^{R} \exp\left(-\frac{1}{2}\Delta_r\right)\right)$ breaks down as the sum of all model likelihood values in the model suite. These AIC weights (w_i) are used in step two of the model averaging process to generate predictions. A model average prediction ($\hat{\overline{\theta}}$) is calculated as follows:

$$\hat{\overline{\theta}} = \sum_{i=1}^{R} w_i \hat{\theta}_i$$

where $\hat{\theta}_i$ is the prediction from the ith model. Therefore, for a given prediction point, each model's AIC weight is multiplied times the prediction value from that model and then those products are summed for all models in the suite of models for a final AIC-weighted prediction estimate (a temperature) for that point ($\hat{\overline{\theta}}$). All calculations for SSN model predictions, model averaging, and summary statistics were conducted in R (R Core Team 2020) using the following packages: doParallel (Weston and Calaway 2014), foreach (Weston 2017), foreign (R Core Team et al. 2022), SSN (Ver Hoef et al. 2014), and tidyverse (Wickham et al. 2019). Data and R code for the model predictions are available online (doi: https://doi.org/10.23719/1529829).

2.4 | Management Scenarios

2.4.1 | Scenario Descriptions

Using both mechanistic and SSN models, we compared three individual management activities and one cumulative effects (combination of those three activities) scenario for four total alternative tests, with some exceptions noted below. The three individual scenarios included restoring riparian shade (B, hereafter "shade" scenario) across the basin, narrowing mainstem channel widths (C1-C5, hereafter "channel narrowing" scenario), and eliminating crop water loss via ET from the basin (D, hereafter "crop water loss" scenario) (Table 2). The cumulative effects scenario (E, hereafter "cumulative" scenario) employed all three of these activities in the MFJD, but only shade and crop water loss scenarios in the SFNR and WR since channel widening there is not primarily caused by anthropogenic disturbances. For the WR, model suites for the month of May did not retain a shade covariate, so a riparian shade restoration scenario was not run for May and consequently, no cumulative effects scenario was run in May. Cumulative effects scenarios for WR were generated for August and September as a shade covariate was retained in these model suites. For each alternative scenario, we compared temperature predictions to a baseline scenario that represents current conditions in these basins that was established by calibrated SSN models (Fuller et al. 2023a, 2023b).

Irrigated crop water loss scenario predictions for the MFJD and SFNR study areas were generated by taking the volume of water estimated to be needed to cover ET needs in irrigated croplands in each reach's upstream subcatchment and setting it to zero. In the WR, the crop water loss covariate was not included as a significant parameter in the model suites for any month. However, to generate a version of a crop water loss scenario in the WR, we took the volume of water that was lost via ET over irrigated croplands and added it back to the discharge volume. This discharge adjustment simulates the mechanism for eliminating irrigated crop water loss as the expectation is water would be retained in the stream if not extracted for irrigation.

Channel width was not directly included as a covariate in the models we used for predicting stream temperature. Instead, we modified how solar radiation reached the stream surface as a result of the channel narrowing. We used methods as described in Fuller et al. (2022) to estimate shade, which incorporates the shadow effects of topography, vegetation height/density, and channel aspect/width in its calculations of percent reach shade (shade.xls model). By narrowing the channel width in the shade estimate calculations, we could estimate how solar radiation to the stream would be reduced. We performed five levels of channel narrowing that included 10%, 20%, 30%, 40%, and 50% channel narrowing along the mainstem of the MFJD river. For cumulative effects scenarios, we used the 50% channel narrowing scenario for a potential maximum cumulative impact from combining all restoration activities in the basin.

In the MFJD, historic channel widening has been attributed to loss of riparian vegetation, stream straightening, reduction in larger woody debris, increased sediment loading, and decreased floodplain availability (Butcher et al. 2010). The reverse problem of channel incision, with disconnection from the floodplain, is also an issue in the MFJD and elsewhere, but was not addressed in the MFJD TMDL and was not assessed in this effort.

2.4.2 | Spatial Extents for Evaluation

The comparison to our baseline scenario provides a quantitative evaluation of the potential impact on stream temperature each management activity holds. These comparisons with the baseline were summarized at three spatial extents for the SSN models; (1) network (based upon the 1:100 k NHDPlus stream line network), (2) mainstem, and (3) outlet. At the network summary extent, we looked at the average network temperature change from the baseline for each alternative scenario, calculated as the average temperature from all our 1 km prediction reaches. At the network extent, we also generated difference maps (reach-by-reach temperature differences across the network) to see spatially which reaches diverged most from the baseline (Figures S1–S6). At the mainstem summary extent, we averaged the individual reach temperatures for only the mainstem of each study network (bold red lines in Figure 1 networks). We also calculated reach differences with the baseline scenario along the mainstem and plotted those mainstem temperature profile differences to see where along the mainstem the largest cooling effect was predicted. Finally, the outlet summary extent provides the temperature at the outlet reach from each study network (black dots in Figure 1 networks) for a point comparison with the baseline outlet temperature among scenarios.

2.5 | Designated Fish Habitat Use

To examine the spatial and data characteristics of the two models for determining the potential efficacy of management actions, the model outputs were compared with several WQC applicable to protect designated aquatic life uses in water bodies in Oregon and Washington. The SSN models were fit to observed mean monthly stream temperatures and the fitted models provide estimates of mean monthly stream temperatures for each reach within the study networks (data and R code available at https://doi.org/10.23719/1529829). We related predicted mean August stream temperatures to a 90% probability of WQC exceedances of 7DADM for designated fish uses by developing logistic regression models using the "glm()" function in the R "stats" package (R Core Team 2020). This allowed us to estimate changes in potential fish habitat use corresponding to different management scenarios.

We evaluated the thermal suitability of stream reaches for different fish species and uses in the study area under different climate/hydrologic regimes (Table 1). There were three fish habitat use classes used, which were mapped to reaches based on each state's designation. Numeric temperature WQC in streams of Oregon and Washington states were used as a proxy for thermal suitability. The fish uses evaluated in this study are a subset of designated fish and aquatic life uses in Oregon and Washington rules (see https://www.epa.gov/wqs-tech/water-quality-standards-regulations-oregon and https://www.epa.gov/wqs-tech/water-quality-standards-regulations-washington). We only evaluated numeric temperature WQC that are applicable year-round (including the summer months) and, for example, did not evaluate seasonal uses and their applicable numeric criteria. Therefore, our reference to fish uses in each state and their associated temperature numeric WQC are not exhaustive and only reflect the fish uses we selected for analysis due to their relevance to our study area and the warm summer months. The interpretation of these numeric temperature WQC made in this study is for research purposes only.

For each fish use criteria, for August, typically the warmest month of the year in the PNW, we evaluated the proportion of stream reaches (by length) currently designated for a particular use by each state that are likely to meet the criteria under different hydrologic and climate regimes (Table 1). We evaluated three relevant fish use categories (7DADM criteria for char spawning and rearing [12°C], core cold water habitat [16°C], and salmon and trout rearing and migration [18°C]) in the MFJD, two (char spawning and rearing [12°C] and core summer habitat [16°C]) in the SFNR, and one (core summer habitat [16°C]) in the WR. For each fish use criteria and each hydrologic/climate regime, we also evaluated how different restoration activities affected the proportion of designated stream reaches that meet the criteria. This full cross between fish use criteria, hydrologic/climate regimes, and restoration activities provides a variety of scenario outcomes to quantify and compare how different restoration activities affect designated fish use habitat availability in the summer for a range of climate and hydrologic conditions that could be representative of more common climate regimes in the future.

3 | Results

3.1 | Comparing Mechanistic and SSN Models for Mainstem Only

Mechanistic models and SSN models were compared along the mainstem reaches of each study basin (Figure 1). Error statistics calculated to explore divergence or convergence of mechanistic and SSN model predictions overall were ~1°C or below for all management scenarios (Table 2). The reduction in crop water loss scenario had the lowest error statistics between the two model types, followed by channel narrowing (MFJD only), riparian shade restoration, and then the combined management activities scenarios for each basin (Table 2). For the calibration year mechanistic model runs, SSN model predictions often presented more reach-to-reach temperature variability along the mainstem than mechanistic models (Figure 2).

When comparing how mechanistic and SSN models each responded to management scenarios, there were different patterns associated with each study basin. In the MFJD, HeatSource (mechanistic model used in MFJD TMDL) riparian shade restoration resulted in lower predicted temperatures along the mainstem than for the SSN model (Figure 3a), but comparable temperature predictions for the channel narrowing and crop water loss scenarios (Figure 3b,c). The combined management scenario accentuated the lower temperature predictions of HeatSource for shade restoration along the mainstem when compared with SSN predictions (Figure 3d). For the SFNR, management scenarios followed similar model comparison differences as in the MFJD (Figure 4). For the WR, Qual2k (mechanistic model used in WR TMDL) predictions overlapped with the SSN predictions, but SSN temperature predictions were more variable along the mainstem than Qual2k predictions among the scenarios used (Figure 5).

3.2 | SSN Management Action Predictions Across Spatial Extents

3.2.1 | MFJD Scenarios

In the MFJD, estimated SSN mean temperature differences from the baseline scenario for all summary extents follow a similar pattern across scenarios and months, though at varying degrees of cooling for each month, with the greatest shade restoration effects seen in May (Figure 6). The largest mean cooling effects were seen in the riparian shade restoration scenario. These riparian restoration shade cooling effects are most pronounced in

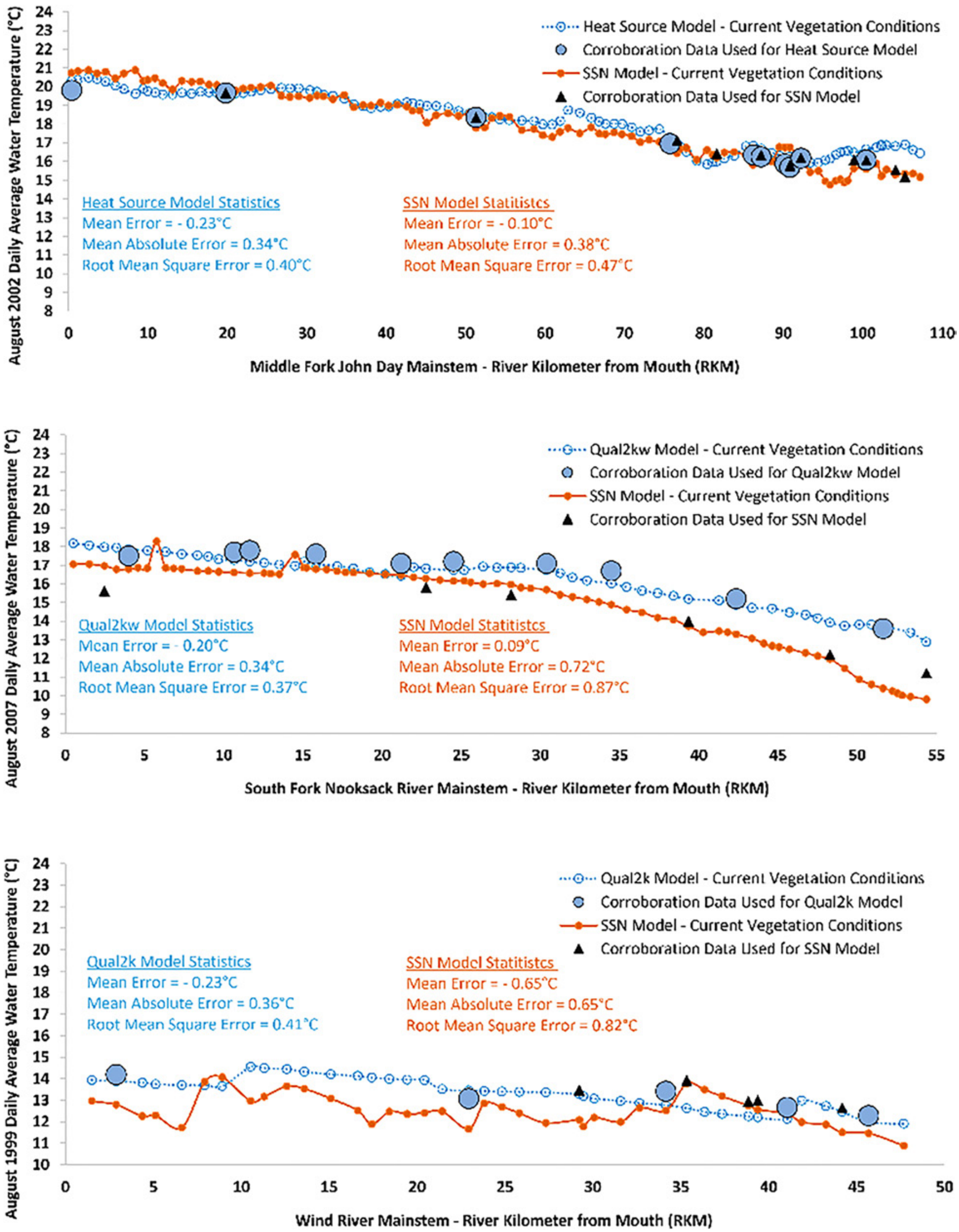

FIGURE 2 | Comparison between field measurements (blue circles and black triangles) with predicted mean August stream temperatures for "current" conditions during the mechanistic modeling calibration periods in each basin (Middle Fork John Day River, South Fork Nooksack River, and Wind River were August 2002 (averages shown), 8/2/2007, and 8/5/1999, respectively). [Color figure can be viewed at wileyonlinelibrary.com]

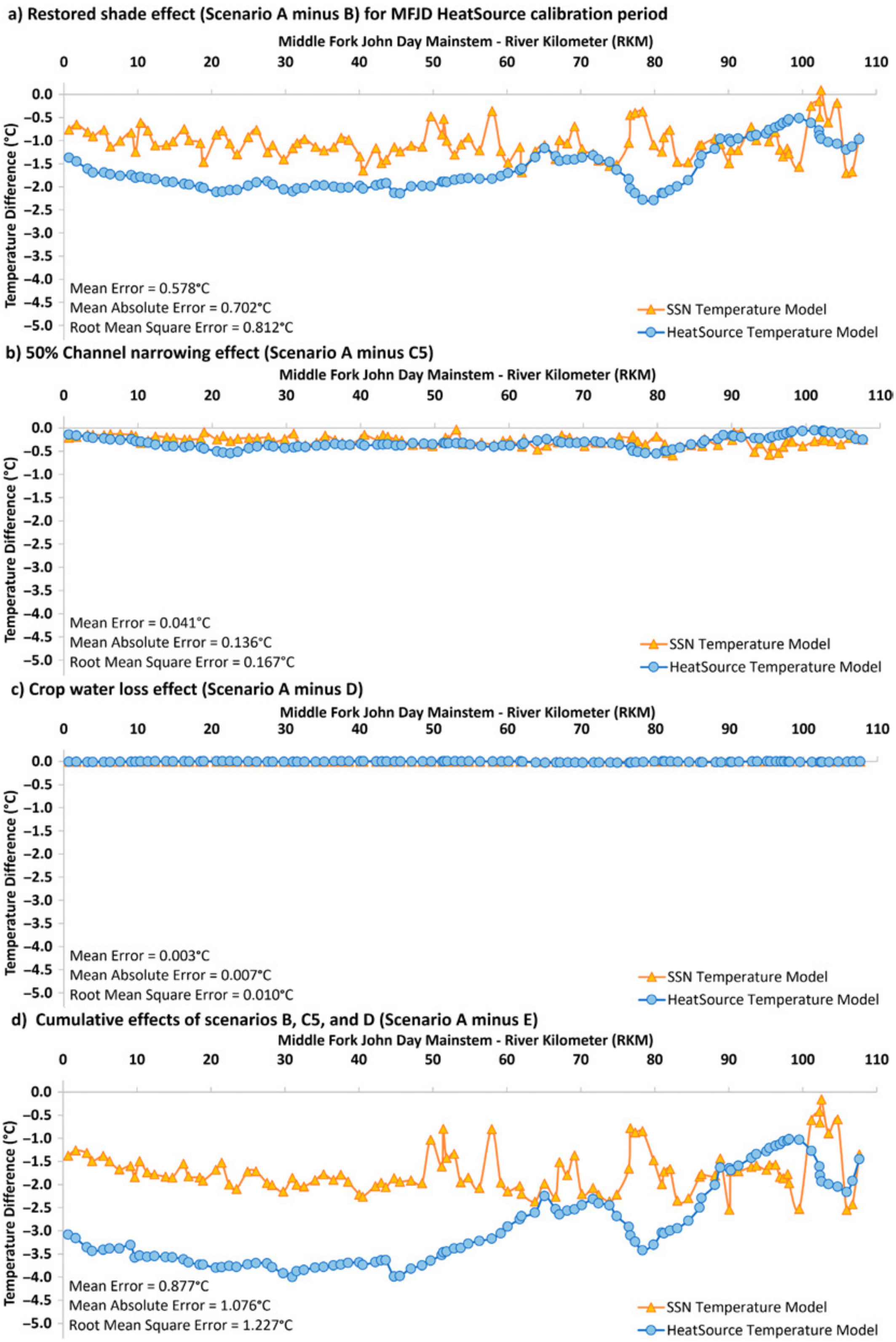

FIGURE 3 | Middle Fork John Day River mean August stream temperature response to management scenarios predicted using SSN and HeatSource mechanistic models. [Color figure can be viewed at wileyonlinelibrary.com]

May (average −3.4°C), followed by August (−1.5°C), and then September (−1.3°C). The cooling observed in the channel narrowing scenarios (−0.01 to −0.06°C) was much smaller than riparian shade restoration, but larger percent channel narrowing did generate slightly larger cooling effects than smaller percent narrowing scenarios. Eliminating crop water losses also had a small temperature effect across all scenarios and months (<−0.02°C; Figure 6).

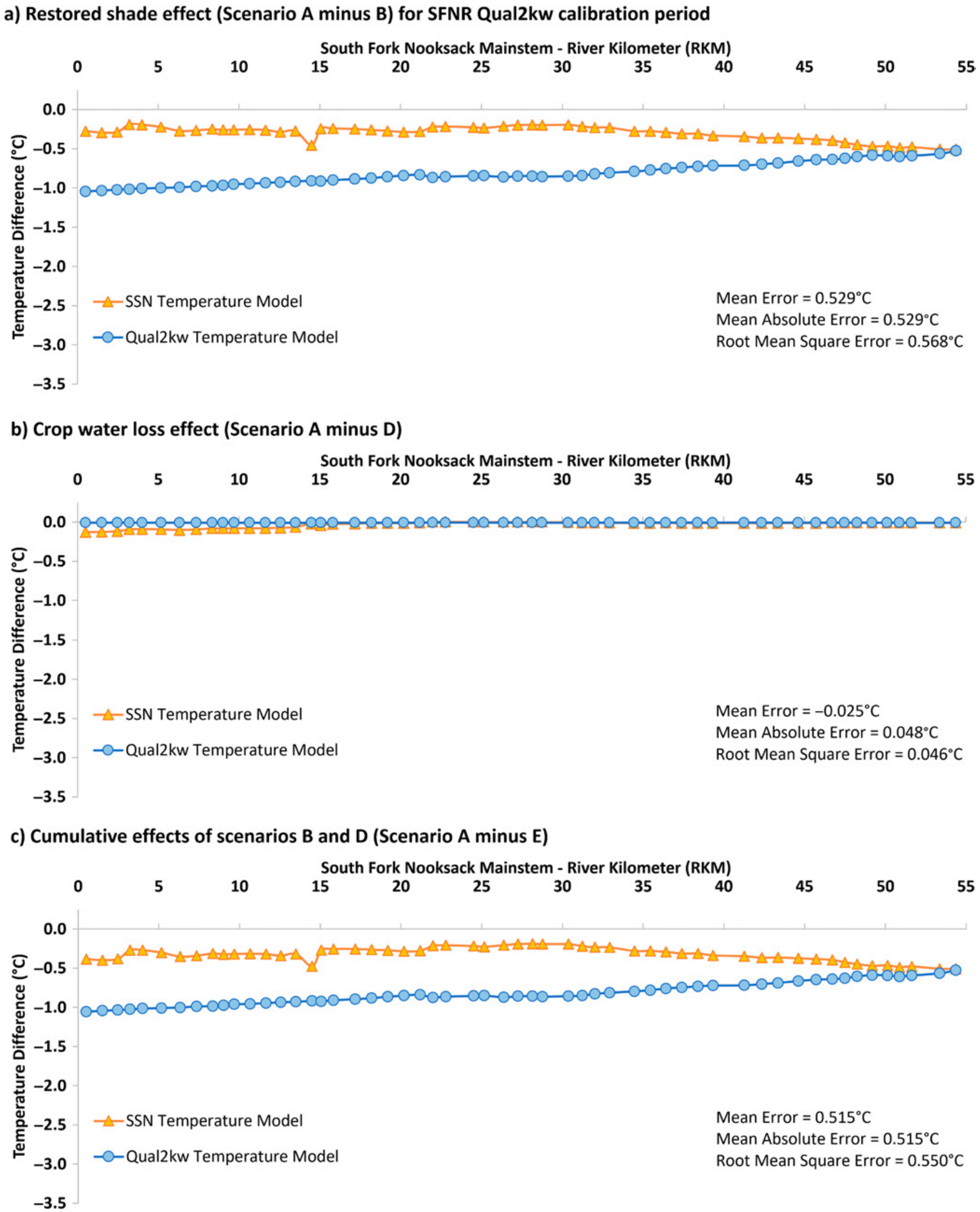

FIGURE 4 | South Fork Nooksack River August mean August stream temperature response to management scenarios predicted using SSN and Qual2kw mechanistic models. [Color figure can be viewed at wileyonlinelibrary.com]

Along the mainstem of the MFJD, management activities provided individual reach cooling of up to −5.7°C in May, −2.5°C in August, and −2°C in September (Figure 7). Mainstem channel narrowing scenarios had smaller cooling effects of up to −0.7°C at the reach-scale than riparian shade restoration. Elimination of crop water losses had little to no cooling effect

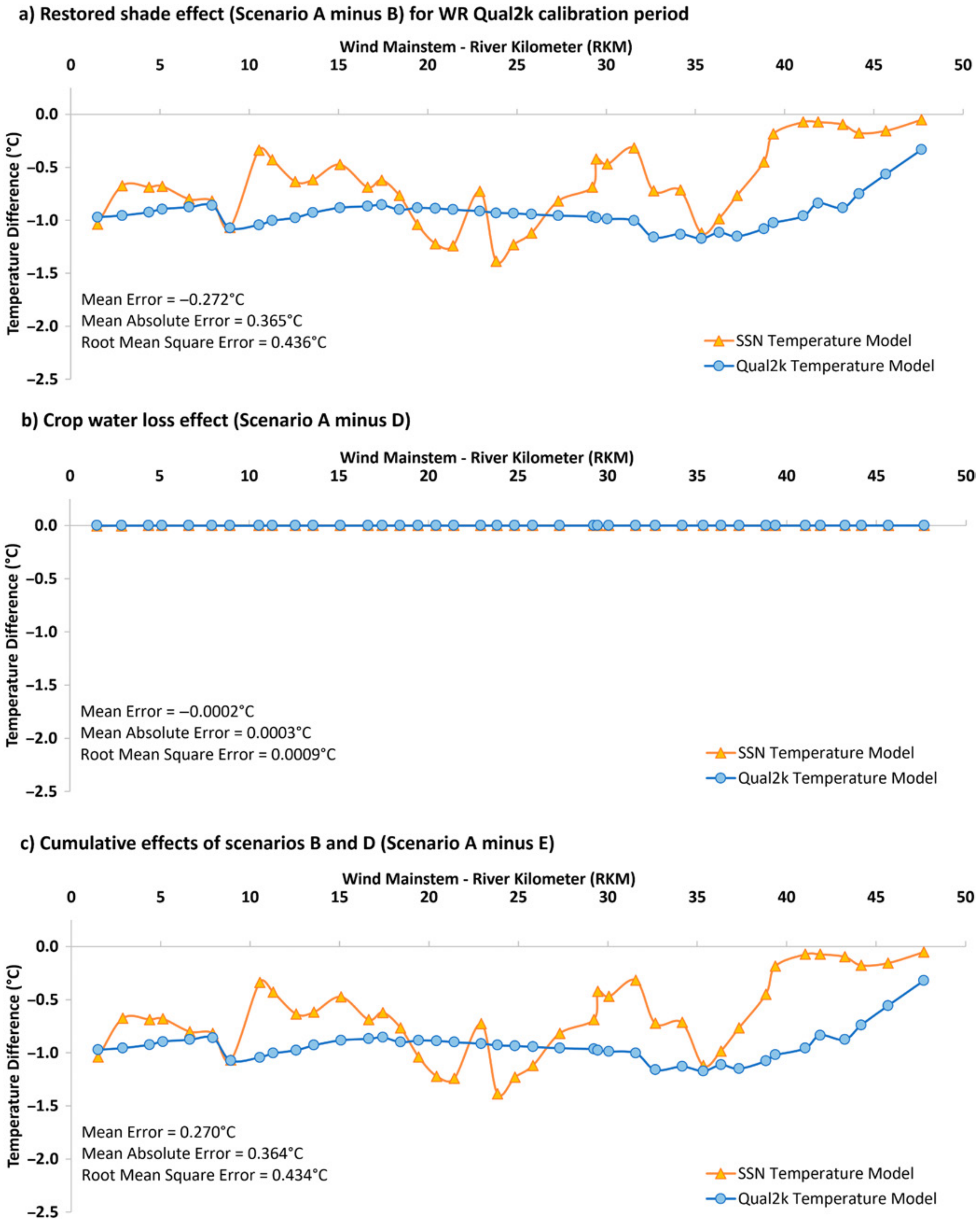

FIGURE 5 | Wind River August mean August stream temperature response to management scenarios predicted using SSN and the Qual2k mechanistic models. [Color figure can be viewed at wileyonlinelibrary.com]

along the mainstem (<−0.01°C) and in May resulted in an unexpected slight warming trend moving toward the outlet from the headwaters when compared with the baseline scenario. The cumulative scenario was the most effective scenario for cooling the mainstem reach on average (Figure 6), as well as reach-by-reach along the mainstem (Figure 7). The cooling effect from cumulative management activities was smaller toward the outlet (−1.3 and −0.4°C) than in the headwater sections of the mainstem (−2.1 and −1.6°C) for all scenarios in August and September. In May, however, the opposite was observed with cooling effects increasing toward the outlet (cumulative −2.9 to −3.5°C) except in the crop water loss scenario.

3.2.2 | SFNR Scenarios

In the SFNR, shade effects also dominated over crop water loss effects; channel widening was not evaluated (Table 2). The mean monthly temperature differences from the baseline scenario for all summary extents and months follow a similar pattern among scenarios except at the outlet in September (Figure 6). In September, the predictions for all restoration scenarios were warmer than the baseline at the outlet of the network (0.2–0.7°C) rather than cooler. For all other scenarios and months, riparian shade restoration provided larger mean cooling effects (−0.28 to −0.36°C) than eliminating crop water loss (−0.01 to −0.03°C). The cumulative scenario resulted in the largest mean cooling at the network and mainstem spatial extents in August.

Cooling effects along mainstem reaches of the SFNR were greater near the headwater reaches (−0.5°C), as compared with near the outlet (−0.3°C) except for the August crop water loss scenario (Figure 7). Specifically, the August riparian shade restoration and cumulative scenarios both reached nearly 0.6°C cooler near the headwaters with smaller cooling effects downstream near the outlet of 0.3 or 0.4°C (shade or cumulative scenarios, respectively). Large temperature differences were noted at river kilometer 14 (−0.7°C) in both months in the riparian shade restoration and cumulative scenarios. The cumulative scenario was almost identical to the shade scenario between river kilometers 14 and 60 (Figure 7).

3.2.3 | WR Scenarios

Due to the lack of a shade covariate in the May model suite, there was no cumulative effects scenario to generate as only the

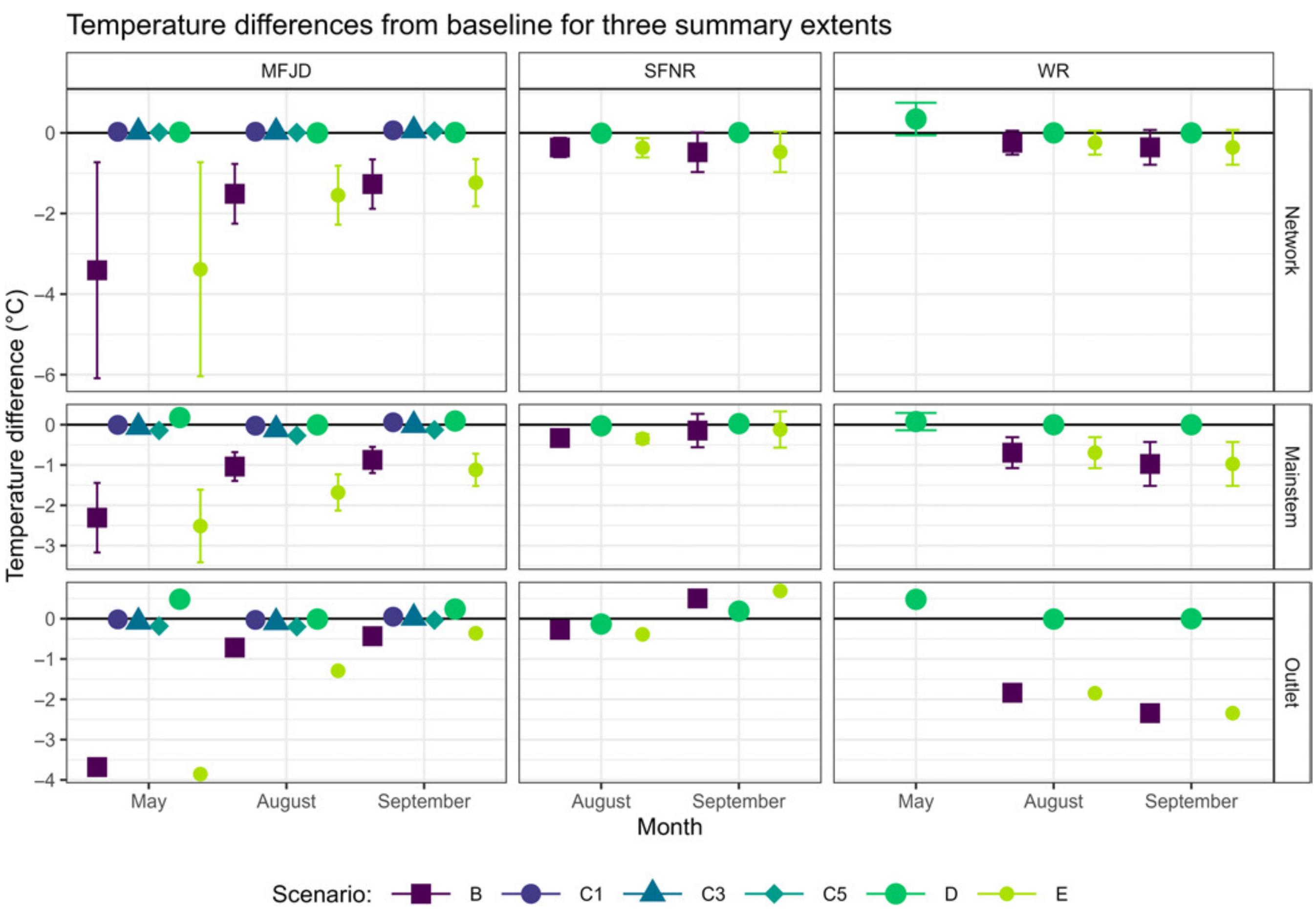

FIGURE 6 | Summary extent (entire network, mainstem reaches only, and outlet reach) temperature differences from baseline for management activity scenarios in each month for (a) Middle Fork John Day River (MFJD), (b) South Fork Nooksack River (SFNR), and (c) Wind River (WR). Network and mainstem summary extents have error bars (+/−1 standard deviation) associated with their mean values while the outlet does not since outlet points represent a single reach temperature extracted for its summary extent in each scenario. Scenario IDs as in Table 2. [Color figure can be viewed at wileyonlinelibrary.com]

crop water loss scenario was present. In May, the crop water loss scenario resulted in slightly warmer (0.3°C) or similar stream temperatures (<−0.01°C) when compared with the baseline at any summary spatial extent, and only a small cooling impact in August and September (Figure 6). The August and September shade scenarios cooled streams at all summary spatial extents (−0.24 to −2.3°C) and were coolest at the outlet (− 2.3°C). Shade restoration resulted in slightly larger cooling effects in September than in August at all summary extents.

Along the mainstem of the WR during the month of May, removing the crop water loss had an extended cooling effect between river kilometers 20 and 35 (−0.03 to −0.32°C, Figure 7). There was one mainstem reach at around kilometer 7 that also cooled in the crop water loss scenario, but otherwise the predicted stream temperatures were warmer than the baseline scenario (Figure 7). Removing crop water loss in August and September had very limited cooling effects along the mainstem (<0.01°C). Shade scenario mainstem profile differences with the baseline in August and September presented little to no cooling (≈0.2°C) in the headwater reaches (river kilometer 40–44) with more cooling (up to 2°C) occurring toward the outlet. Differences in the cooling effect were present along the mainstem, and between months, that varied by up to ±1°C. Cumulative scenarios in the August and September months mirrored the shade scenarios.

3.3 | Management Practice Effects on Fish Habitat

In the MFJD, the proportion of reach length predicted to meet assigned designated uses is only 10% under current conditions for char spawning and rearing habitats, ~30% for salmon and trout rearing and migration and highest for core cold water habitat (~50%) across all hydrologic and climate regimes (Figure 8). The proportion of reach length designated as char spawning and rearing habitat meeting the required 7DADM of 12°C is most affected by hydrologic and climate regime variability, varying by over an order of magnitude between the wet and warm temperature years. The thermal suitability of char spawning and rearing designated reaches is also most strongly impacted by management regimes, with percentages more than doubling over a range of hydrologic regime conditions if riparian shading is implemented either alone or in combination with other management practices investigated. However, predicted management effectiveness diminishes in dry years and is extremely low in the warm temperature year. Improvements in percentages of those reaches classified as core cold water habitat and salmon and trout rearing/migration habitat predicted to achieve thermal requirements under management strategies involving restoration of riparian shade either alone or in combination were relatively consistent, increasing by ~20% across different hydrologic regimes.

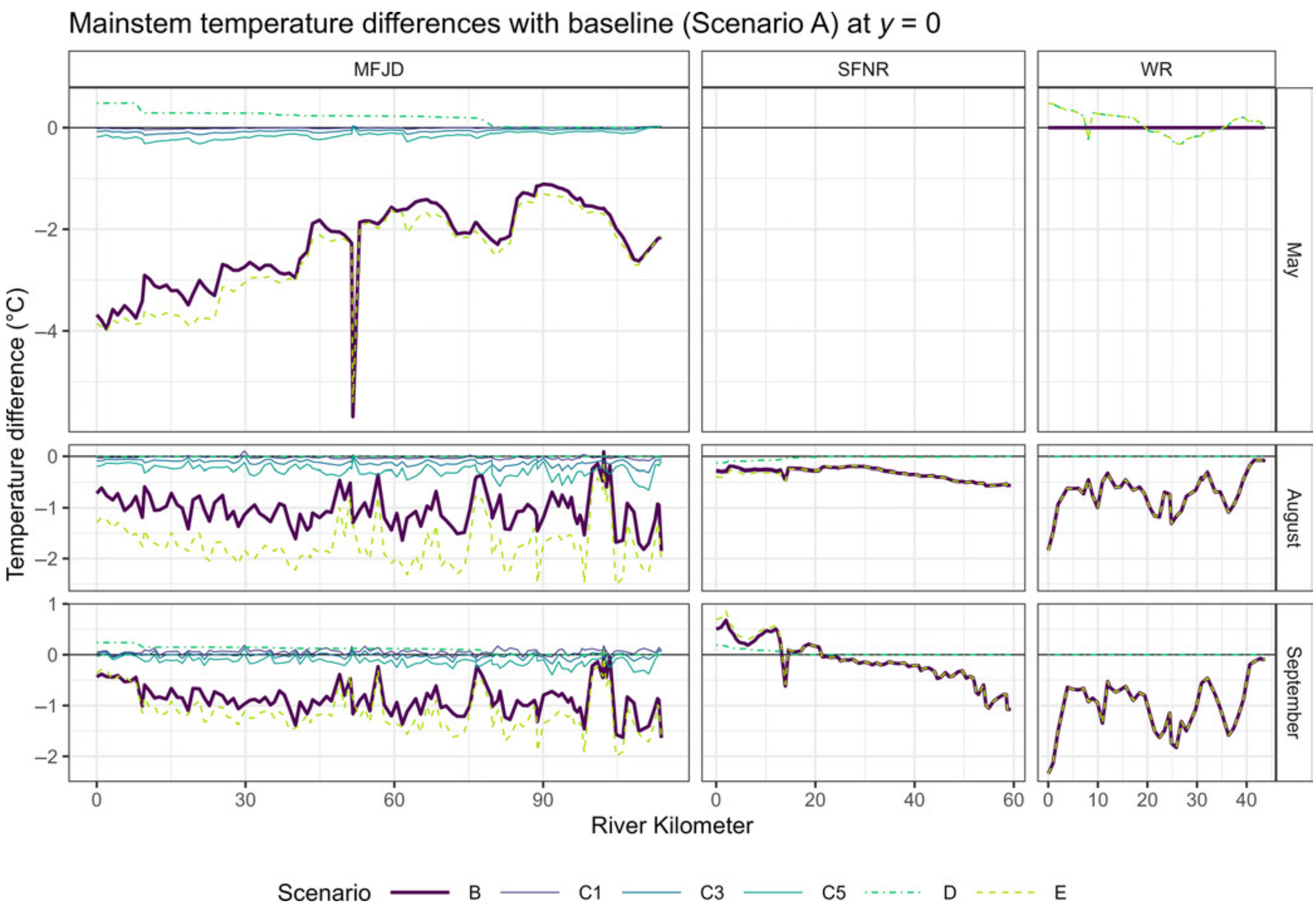

FIGURE 7 | Mainstem profiles for temperature differences between baseline (zero on axis) and management activity scenarios for each month for (a) Middle Fork John Day River (MFJD), (b) South Fork Nooksack River (SFNR), and (c) Wind River (WR). River kilometer 0 is the outlet and scenario IDs are as in Table 2. [Color figure can be viewed at wileyonlinelibrary.com]

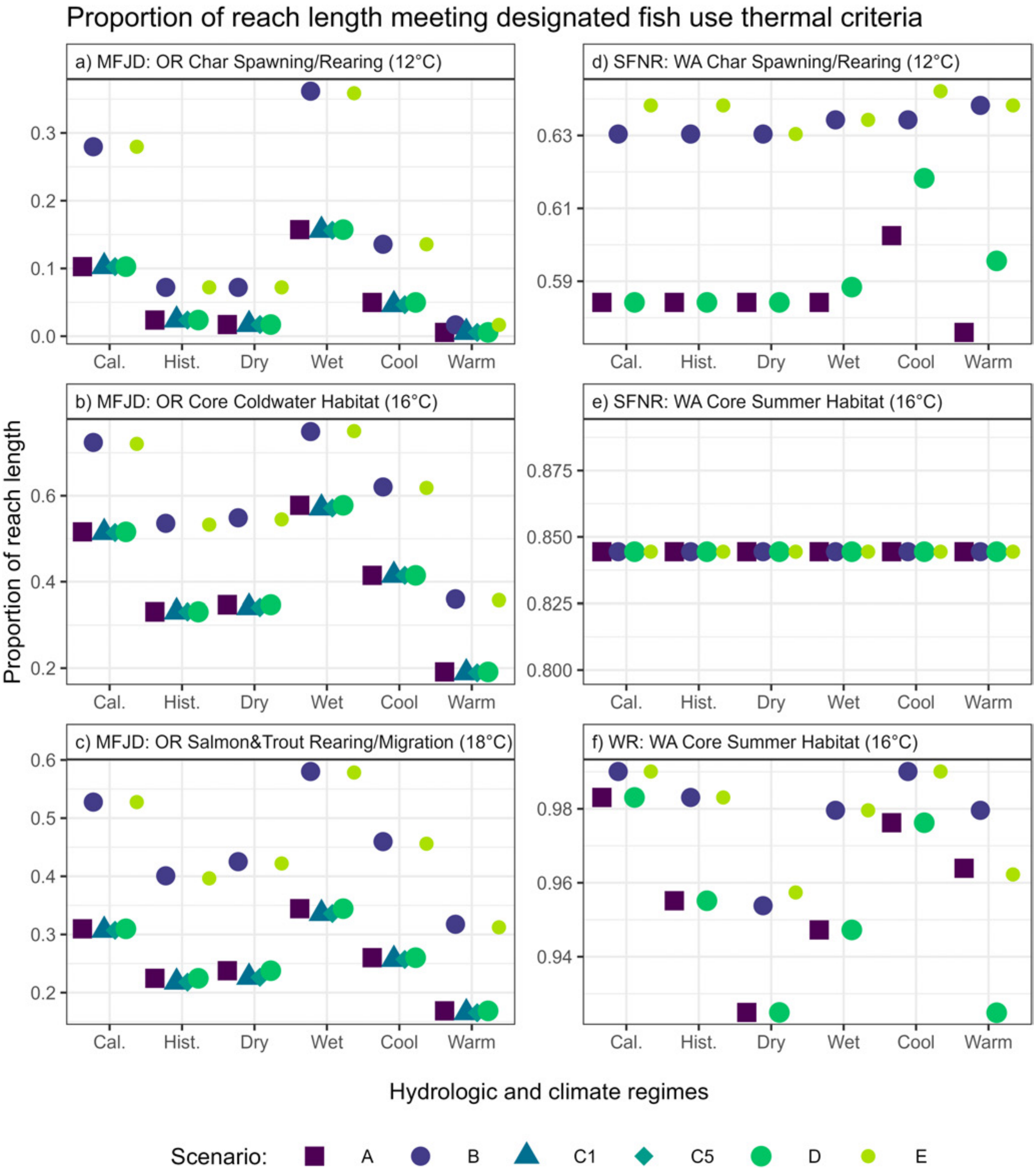

FIGURE 8 | Fraction of potential August fish habitat projected to meet 7DADM water quality criteria to protect the designated use for specified fish under variable management improvement scenarios and hydrologic regimes. (a) fraction of OR Char Spawning/Rearing Habitat meeting 12°C for estimated 7DADM in Middle Fork John Day (MFJD), (b) fraction of OR core cold water habitat meeting 16°C for estimated 7DADM in MFJD, (c) fraction of OR salmon and trout rearing/migration meeting 18°C for estimated 7DADM in MFJD, (d) fraction of WA char spawning/rearing habitat meeting 12°C for estimated 7DADM in South Fork Nooksack River (SFNR), (e) fraction of WA core summer habitat meeting 16°C for estimated 7DADM in SFNR, and (f) fraction of WA core summer habitat in WR basin. See Table 2 for definition of management practice scenarios. S801 = 2002 mechanistic model calibration year, S802 = 1990–2015, S803 = 2007 dry year, S804 = 1993 wet year, S805 = 2006 low water temperature year, and S806 = 2001 high water temperature year. [Color figure can be viewed at wileyonlinelibrary.com]

For the SFNR, the projected habitat meeting char spawning and rearing habitat thermal requirements under baseline conditions was 58%–60% (Figure 8). The SSN model predicted the proportion of habitat meeting these requirements would be increased by about 9% across all hydrologic regimes when management strategies involving riparian restoration were implemented, with slightly lesser improvement (from higher baselines) in a cool temperature year and slightly greater improvement (from lower baselines) in a warm temperature year. Somewhat paradoxically, reducing water use appeared to affect char habitat to a lesser extent and only for the cool and warm water temperature years. Proportion of habitat projected to meet core summer habitat requirements was ~84% under baseline conditions (concentrated in tributaries) and was not predicted to improve further within mainstem reaches under any of the planned management strategies implemented (Figure 8). In a dry year, implementation of management practices is projected to reduce fragmentation of suitable char spawning/rearing habitat in the upstream reaches (Figure S8a,b).

In the WR basin, percentages of stream kilometers predicted to meet designated uses under baseline conditions were higher than 92% for core summer habitat (Figure 8). Thus, percentage of reaches predicted to meet designated uses for streams within the WR basin was less sensitive to variation in hydrologic regime than that reported for the MFJD streams. Increases in the projected attainment of the core summer habitat were primarily attributed to riparian zone restoration either alone or in combination. In a dry year, combined restoration strategies helped restore connection of core summer habitat throughout one of the lower tributaries (Figure S9).

4 | Discussion

4.1 | Comparison of Mechanistic Versus SSN Model Predictions

Our models showed accuracy comparable with that reported in the literature for mechanistic versus statistical (SSN) model performance in predicting stream/river temperatures. For the mainstem comparisons, the mechanistic models used for TMDLs had slightly better accuracy, with RMSE of 0.4°C for prediction of August daily average water temperatures for the year modeled, as compared with SSN model performance (RMSE = 0.47–0.87°C; Figure 2). Across the full network, SSN model errors were somewhat higher (RMSE = 1.0–1.4°C) for mean August temperatures (assessed against observed data) as compared with mechanistic model results for the mainstem (Fuller et al. 2023). Historically, when developing predictive temperature models for streams, there has been a tradeoff between accuracy of model predictions and practical spatial extent of model coverage. Mechanistically-based heat budget models such as SNTEMP (Krause et al. 2004) can predict stream temperature within a few tenths of a degree but are difficult to apply across large regions. Using SSN models, Isaak et al. (2010) have reported r^2 values of greater than 0.90 for model predictions of mean or maximum stream temperatures in the Pacific Northwest, with RMSE of approximately 1°C. Siegel and Volk (2019) were able to achieve similar model performances (averaged RMSE for each watershed ranged from 0.85 to 1.54°C for daily mean temperature and 23% less for monthly means) with a simpler GAM modelling approach without accounting for spatial autocorrelation but focusing on key interactions between spatial and time-varying parameters (day of year or air temperature).

Although there are several reviews comparing the accuracy of statistical or mechanistic stream temperature model predictions across studies (Benyahya et al. 2007; Brennan 2015; Detenbeck et al. 2016; Holthuijzen 2017), with different models applied to different systems, this is only the second study we know of that compares the predictions of process-based mechanistic and empirical SSN models within the same systems using comparable datasets (Lee et al. 2020). In our case, although the SSN models provided predictions for the full watershed stream networks, our direct comparisons had to be limited to the mainstem of these systems due to the more restricted mechanistic model boundaries. Lee et al. (2020) used the process-based DHSVM-RBM model to demonstrate the implications of climate change for streamflow and water temperature in two watersheds. Both mechanistic and SSN models projected generally similar future spatial patterns of maximum water temperature in the Snoqualmie watershed (WA) and Siletz watershed (OR), USA, with reductions in number of cool reaches particularly in lower reaches of the watershed. However, the process-based model projected higher spatial heterogeneity in water temperature as compared with the NorWeST SSN model due to the spatially explicit simulation of streamflow and because the model was calibrated with spatially continuous remotely sensed water temperature data.

The added spatial complexity in our SSN models yielded slightly more heterogeneous temperature profiles along the mainstem of modeled systems as compared to mechanistic models (Figure 2), which was accentuated even more in longitudinal trends predicted under different management scenarios (Figures 3–5). Our SSN models incorporated greater spatial heterogeneity in parameters related to hydrologic processes such as reach and month/year-specific discharge and time of travel estimates, hot spring distribution, irrigated crop water loss, and surface water extraction use for public supply as compared with the traditional NorWeST SSN models. The latter use watershed yield from reference stations to account for interannual variation in flow and a baseflow index to capture spatial variation in groundwater inputs. Our SSN models also incorporated terms capturing interactions between variables related to groundwater inputs and slope, and between discharge and air temperature.

For the modeled mainstems for our systems, the SSN and mechanistic models produced similar temperature response trends to the various stream shade and flow scenarios, that is, comparison statistics were mostly below 1°C. The SSN models and two of the mechanistic models produced a linear temperature response to stream shade changes, but the mechanistic models were slightly more responsive to stream shade changes in management scenarios than was observed in the SSN models (Figures 3–5). Greater responsiveness of the mechanistic models to riparian zone restoration could be due to the integrative nature of shade estimates in the SSN models, which incorporate effects of upstream changes in shade aggregated by distance or travel time. In contrast, if only applied to the mainstem, the mechanistic models cannot fully describe spatial variation in effects of upstream restoration throughout the network.

4.2 | Restoration Target Scales: Outlet Versus Mainstem Versus Full Watershed

Conclusions about the projected success of various management scenarios in likely meeting temperature WQC differ substantially depending on the scale of the management target and, to a lesser extent, the modelling approach applied. For mechanistic models evaluating riparian shade restoration, the largest predicted temperature reduction was at the watershed outlet (decrease ranging from 1.0 to 3.0°C), and this impact decreased moving upstream from the watershed outlets. This could be the result of imposing management strategies only along the mainstem for the mechanistic models, while keeping boundary conditions constant. In contrast, the SSN models showed improvements to tributaries throughout the watershed, yielding more variable improvements along the mainstem as compared with the mechanistic modeling results. Projected decreases of 3°C at the outlet of MFJD for the combined August management scenario for the HeatSource modeled year (2002) would have been insufficient to meet even the least restrictive cold water use criterion for this watershed (OR salmon and trout rearing and migration) with 90% probability. The latter requires a 7DADM of 18°C, which corresponds to a mean August daily average of 13.6°C. The MFJD SSN model predicted far less temperature improvement at the outlet (decrease of 1.0°C). For the SFNR, the initial mean August conditions were cooler at the outlet (18°C based on Qual2K) as compared with SSN model projections, however projected improvements for the combined scenario were lower (1.0°C). For the WR, the initial mean August conditions at the outlet were even cooler for the modeled year (14°C) based on Qual2K as compared with SSN predictions, and both Qual2K and the SSN model projected only a 1°C decrease in temperature for the combined August scenario. Thus, the combined August scenario for the WR meets requirements for the majority of criteria but modeled improvements would not be sufficient to meet the most stringent target (the lowest criteria 7DADM of 12°C/mean August temperature of 10.4°C).

When the full networks are considered, the SSN model projections show a wider range of responses to the combined management scenarios than projected by the mechanistic models, however, spatial patterns differ among watersheds. As described above, suitable fish habitat areas increased in all fish use categories although improvements in coldest categories were small. Variability is greatest for the MFJD combined scenario (1–4°C reductions), with greatest improvements from riparian restoration clustered in selected tributaries (Figure S2), as compared with 50% channel narrowing. The WR showed an intermediate variability in responsiveness (up to 2.0°C), with greatest improvements focused along the mainstem and higher order tributaries (Figure S6). The SFNR showed the least variability in responsiveness (up to 1.2°C), with greatest improvements focused in headwater tributaries (Figure S4).

4.3 | Comparison of Management Scenario Effectiveness in Protecting Fish Habitat

Our results showed restoration to potential natural vegetation yielded the greatest potential improvements in habitat suitable for char spawning and rearing (≤12°C), salmon and trout rearing and spawning, and OR core cold water or WA summer core habitat (≤16°C) fish aquatic life uses across the basins, with relatively minor improvements attributable to either channel narrowing of 10%–50% or reduced water use for the studied basins. Using either empirical or mechanistic (HeatSource) models for tributaries to the Snake River in Washington, others have predicted large gains in the percentage of stream miles meeting the ≤16°C criterion resulting from increased stream shade conditions following restoration of riparian vegetation, with relatively smaller projected improvements associated with channel width reduction (Justice et al. 2017; Seixas et al. 2018). White et al. (2018) reported greatest effect of riparian restoration in Snake River tributaries narrower than 40 m, suggesting that channel widths were narrow enough that modifications would have significantly affected channel shading (Seixas et al. 2018). In the lower Columbia River Basin, riparian zone restoration was projected to have greatest impact on thermal regimes for relatively small streams with bankfull widths less than 5 m (Fuller et al. 2022). In our study, MFJD mainstem shade was projected to increase by only 6%–7% for the 50% reduction in channel width scenario (with no change in riparian vegetation) and the riparian restoration alone scenario resulted in an average shade gain of 24% (Table S4). However, shade gain increased by an average of 40% with the combined scenario of 50% reduction in channel width and restoration of riparian vegetation, indicating that the effects of channel narrowing and riparian restoration were synergistic for the mainstem of MFJD. Finally, our simulation of effects of channel narrowing only evaluated potential effects via riparian shading and did not consider potential ancillary effects associated with water depth or current velocity, both of which would likely reduce stream temperatures further. Depending on channel morphology, in some cases, channel narrowing might have negative impacts if channel complexity were reduced with fewer side channels and spring channels, which tend to dampen and delay diurnal temperature fluctuations (Arrigoni et al. 2008).

Contrary to conditions often reported within literature, we found little predicted effects of restoring flows by eliminating withdrawals for irrigation or other uses on either temperature profiles in the mainstems or on cold water habitat availability. However, the predicted changes in flow from water withdrawals were extremely low, ranging from 0.01% (WR) to 2.65% (MFJD; Table S6). Our simulations of reduced irrigation flows via both the SSN and mechanistic models were relatively simplistic, considering only effects related to thermal inertia via flow volume. In reality, the predicted effect of changes in irrigation practices can differ widely depending on the source of irrigation withdrawals—groundwater versus surface water (Essaid and Caldwell 2017), and the contribution of lateral flows or subsurface tile drainage versus surface runoff to return flows (Alger, Lane, and Neilson 2021; Bjorneberg 2015). On a broader scale, irrigation can affect the overall heat balance of the upland watershed, with nighttime warming exceeding daytime cooling (Chen and Jeong 2018). The study of such effects (e.g., nighttime warming vs. cooling) was not within the scope of our analysis.

4.4 | Evidence of Restoration Effectiveness in Study Watersheds as Compared to Model Predictions

In only two of the TMDL watersheds have recent riparian plantings been targeted to have the greatest impact (Middle Fork IMW Working Group 2017). For example, in the WR watershed, around the turn of the 21st century riparian restoration was targeted to specific reaches (Connolly 2001; Connolly and Jezorek 2008; Connolly et al. 1998) and temperature monitoring to track improvements has taken place since 1996. Various restoration activities have been focused on the Trout Creek subbasin, which was logged in 1948 with subsequent removal of large woody debris in the 1970's resulting in increased flow velocity, which led to channel downcutting. Most restoration activities implemented in 2003–2008, with the exception of reduction of bankfull width to depth ratios from over 60 to less than 25, have long-term targets (https://srp.rco.wa.gov/project/140/11558): "Objectives: (1) Restore riparian conifers along Upper Trout, Crater, Compass and Layout Creek to eight trees/acre >31″ in diameter (200 years). (2) Increase shade >80% (60 years). (3) Increase bank stability >80% (2 years). (4) Reduce bank full width to depth ratios <25 (2 years). (5) Increase in-stream large woody debris >100 pieces/river mile (1 year). Project actions: Riparian areas will be thinned and under-planted with native conifers." In the MFJD, most riparian restoration has taken place within the last 5–10 years, while model projections of vegetation growth and shade evolution suggest that 50 years are needed to see significant differences in shade needed to improve thermal regimes (Middle Fork IMW Working Group 2017). A detailed HeatSource model applied to the Oxbow Tailings Restoration project demonstrated that a narrowing of the channel resulted in decreased temperatures due to a decrease in stream surface area. HeatSource modeling was required to tease out the effect of surface area due to interannual variability in weather before and after restoration activities (Hall, Chiu, and Selker 2020).

Led by the Nooksack Tribe in the SFNR, a 5 year cycle has been established for assessing effectiveness of restoration activities starting in 2020, but it is too soon for results to be available. The SFNR is dominated by forests, both managed and preserved. Suggested strategies for forest-dominated landscapes such as this, with fewer opportunities to reduce anthropogenic water use, include strategic creation of forest gaps to increase snow storage thus delaying peak snowmelt and groundwater inputs, and an increased rotation frequency for tree harvesting to increase soil moisture by reducing ET (Dickerson-Lange et al. 2022).

4.5 | Limitations to Management Scenario Simulations

Our simulations of the effect of reducing anthropogenic water use on thermal regimes were restricted to those factors identified in the existing TMDLs as contributors to temperature exceedances. We were also limited by the quality and resolution of data available on water use. USGS compiles data on water use at the county scale, which then must be disaggregated down to finer resolution based on population estimates. Estimates of irrigated cropland were indirect, relying on relatively coarse remote sensing data for irrigated crop locations and ET estimates. We did not have access to information on the location of return flows. A more complete analysis of reducing irrigation impacts on thermal regimes would require more spatially explicit data, information on the nature of withdrawals (groundwater vs. surface water), the potential influence of subsurface drainage on return flow temperatures, and the nature of return flows (lateral flows vs. runoff). Simulation of irrigation impacts was limited to crop water losses and did not consider other factors affecting irrigation impacts, which could require linkage with more sophisticated hydrology and/or landscape energy balance models.

In the TMDL watersheds we examined, anthropogenic water use was relatively low compared with total discharge in streams. In other systems, the impact of interbasin transfers could have significant intrabasin effects on thermal regimes even absent significant water use, depending on timing and location of withdrawals and return flows (Liu et al. 2022).

In watersheds dominated by forests, there is a need for simulation of other activities to help restore more natural hydrologic regime timing and baseflow, for example, forest management, wetland restoration, and green infrastructure (Dickerson-Lange et al. 2022). Incorporation of these effects into SSN models may require sequential application of models or literature mining to simulate monthly changes in flow regimes.

Our simulation of channel narrowing scenarios could be improved by distinguishing between historic changes expected for constrained and unconstrained channels and considering the importance of floodplain reconnection for incised channels. The latter restoration activity will increase thermal heterogeneity at scales finer than that projected by SSN models. More detailed process-level models and field experiments may be needed for simulating effects of finer scale restoration activities (reach to planform scale) that can increase local thermal heterogeneity.

Our riparian management scenarios did not consider time required for mature forest development given different starting conditions. This could be built into future simulations given information from studies such as White et al. (2017) that provide information on expected shade evolution over time. More information is needed on recovery of shade following different forest management and restoration practices.

5 | Conclusions

SSN models performed well with slightly lesser accuracy as judged by discrete instream temperature monitoring than mechanistic models but provided additional benefits to inform management, including information on spatial and temporal heterogeneity of restoration effectiveness. In theory, the mechanistic models commonly applied in temperature TMDLs could yield information on responses covering a larger fraction of the watershed but effort required to parameterize these across a full network is not realistic. Application of SSN models via scenario simulations can provide critical insight for restoration planning and implementation especially when paired with mainstem mechanistic models. Predicted restoration effectiveness can vary significantly both within and among watersheds. Analysis of the effectiveness of restoration activities in progress will need to take into account the effect of interannual variability in weather on thermal regimes under baseline conditions as well as effectiveness of management actions on fish habitat under dry versus wet years. Interaction between climate variability and management scenario effectiveness could be important in reducing population bottlenecks (Kanno et al. 2015). A focus on WQC exceedance only at the watershed outlet or along the mainstem reach can obscure knowledge of restoration potential for fish habitat in tributaries and headwaters, potential for creation of thermal refuge areas along mainstem critical for maintaining migration corridors, and thermal regime heterogeneity across landscapes in space and time.

Acknowledgments

Funding for this research was provided by the US EPA Office of Research and Development (ORD) via a Region 10 Regional Applied Research Effort project (interagency agreements with the U.S. Forest Service (DW12924479) and U.S. Department of Energy support (DW 92429801-5 and DW 92525701-1)). US EPA ORD contribution number is ORD-046319. We thank Chris Knightes, Gunnar Johnson, and John Palmer for comments on early drafts.

Data Availability Statement

The SSN model fitting and prediction data used in this study come from open source datasets provided by federal and nonprofit institutions (see data listed/linked to in Fuller et al. 2023b). The R scripts (R version 4: R Core Team 2020) used for selecting the final suites of models used in this study can be found here: https://doi.org/10.23719/1524755 (Fuller et al. 2023b). The temperature predictions (and associated R code and input datasets for predictions) for each scenario in this paper is available via data.gov at: https://doi.org/10.23719/1529829.

References

Alger, S., B. Lane, and B. Neilson. 2021. "Irrigation Influences on Summer Stream Temperature Variability." *Authorea*, January 05. https://doi.org/10.22541/au.160982186.65568005/v1.

Arrigoni, A. S., G. C. Poole, L. A. K. Mertes, S. J. O'Daniel, W. W. Woessner, and S. Thomas. 2008. "Buffered, Lagged, or Cooled? Disentangling Hyporheic Influences on Temperature Cycles in Stream Channels." *Water Resources Research* 44, no. 9: 637. http://digitalcommons.unl.edu/natrespapers/637.

Benyahya, L., D. Caissie, A. St-Hilaire, T. B. M. J. Ouarda, and B. Bobée. 2007. "A Review of Statistical Water Temperature Models." *Canadian Water Resources Journal* 32, no. 3: 179–192. https://doi.org/10.4296/cwrj3203179.

Bjorneberg, D. J. 2015. "Water Temperature in Irrigation Return Flow From the Upper SnakeRock Watershed." *Agricultural Water Management* 158: 209–212.

Boyd, M., and B. Kasper. 2003. "Analytical Methods for Dynamic Open Channel Heat and Mass Transfer: Methodology for the HeatSource Model." Version 7.0., 204. http://www.deq.state.or.us/wq/TMDLs/tools.htm.

Brennan, L. 2015. "Stream Temperature Modeling: A Modeling Comparison for Resource Managers and Climate Change Analysis." In *Environmental & Water Resources Engineering Masters Projects*, 72. Amherst, MA: Univ. of Massachusetts.

Burnham, K. P., and D. R. Anderson. 2002. *Model Selection and Multimodel Inference: A Practical Information-Theoretic Approach*, edited by K. P. Burnham and D. R. Anderson, 2nd ed. New York: Springer New York.

Butcher, D., J. Crown, K. Brannan, K. Kishida, and S. Hubler. 2010. *John Day River Basin Total Maximum Daily Load (TMDL) and Water Quality Management Plan (WQMP)*. Portland, OR: Oregon Department of Environmental Quality. https://www.oregon.gov/deq/FilterDocs/jdTMDLwqmp.pdf.

Chapra, S., G. Pelletier, and H. Tao. 2008. "QUAL2K: A Modeling Framework for Simulating River and Stream Water Quality." Version 2.11: Documentation and Users Manual.

Chen, X., and S.-J. Jeong. 2018. "Irrigation Enhances Local Warming With Greater Nocturnal Warming Effects Than Daytime Cooling Effects." *Environmental Research Letters* 13: 024005.

Connolly, P. J. 2001. "Wind River Watershed Restoration: 1999 Annual Report." *Office of Scientific and Technical Information (OSTI)*. https://doi.org/10.2172/790163.

Connolly, P. J., and I. G. Jezorek. 2008. "Wind River Watershed Restoration 2004–2005 Annual Report." *Office of Scientific and Technical Information (OSTI)*. https://doi.org/10.2172/962643.

Connolly, P., S. James, K. Wieman, et al. 1998. "Wind River Watershed Restoration Project, Segments II-V." Project No. 1998-01900, 235, (BPA Report DOE/BP-00004973-1).

Detenbeck, N. E., A. C. Morrison, R. W. Abele, and D. A. Kopp. 2016. "Spatial Statistical Network Models for Stream and River Temperature in New England, USA." *Water Resources Research* 52, no. 8: 6018–6040. https://doi.org/10.1002/2015WR018349.

Dickerson-Lange, S., O. Grah, J. Jay, and R. Mitchell. 2022. "Modeling the Effects of Forest Management on August Streamflow: South Fork Nooksack River Pilot Research Study." A Report Prepared for The Nooksack Indian Tribe Natural Resources Department.

Erős, T., and W. H. Lowe. 2019. "The Landscape Ecology of Rivers: From Patch-Based to Spatial Network Analyses." *Current Landscape Ecology Reports* 4, no. 4: 103–112. https://doi.org/10.1007/s40823-019-00044-6.

Essaid, H. I., and R. R. Caldwell. 2017. "Evaluating the Impact of Irrigation on Surface Water—Groundwater Interaction and Stream Temperature in an Agricultural Watershed." *Science of the Total Environment* 599–600: 581–596.

Fuller, M. R., P. Leinenbach, N. E. Detenbeck, R. Labiosa, and D. J. Isaak. 2022. "Riparian Vegetation Shade Restoration and Loss Effects on Recent and Future Stream Temperatures." *Restoration Ecology* 30, no. 7: 1–17. https://doi.org/10.1111/rec.13626.

Fuller, M. R., N. E. Detenbeck, P. Leinenbach, R. Labiosa, and D. J. Isaak. 2023a. "Spatial and Temporal Variability in Stream Thermal Regime Drivers for Three River Networks During the Summer Growing Season." *Journal of the American Water Resources Association* 60: 57–78. https://doi.org/10.1111/1752-1688.13158.

Fuller, M. R., N. E. Detenbeck, P. Leinenbach, R. Labiosa, and D. J. Isaak. 2023b. "Spatial and Temporal Variability in Stream Thermal Regime Drivers for Three River Networks During the Summer Growing Season [Dataset, Code, Metadata]." Data.gov. https://doi.org/10.23719/1524755.

Grueber, C. E., S. Nakagawa, R. J. Laws, and I. G. Jamieson. 2011. "Multimodel Inference in Ecology and Evolution: Challenges and Solutions." *Journal of Evolutionary Biology* 24, no. 4: 699–711. https://doi.org/10.1111/j.1420-9101.2010.02210.x.

Hall, A., Y.-C. Chiu, and J. S. Selker. 2020. "Coupling High-Resolution Monitoring and Modelling to Verify Restoration-Based Temperature Improvements." *River Research and Applications* 36: 1430–1441. https://doi.org/10.1002/rra.3668.

Holthuijzen, M. F. 2017. "A Comparison of Five Statistical Methods for Predicting Stream Temperature Across Stream Networks." All

Graduate Theses and Dissertations, Utah State University. https://digitalcommons.usu.edu/etd/6535.

Isaak, D. J., C. H. Luce, B. E. Rieman, et al. 2010. "Effects of Climate Change and Wildfire on Stream Temperatures and Salmonid Thermal Habitat in a Mountain River Network." *Ecological Applications* 20: 1350–1371. https://doi.org/10.1890/09-0822.1.

Isaak, D. J., E. E. Peterson, J. M. Ver Hoef, et al. 2014. "Applications of Spatial Statistical Network Models to Stream Data." *WIREs Water* 1, no. 3: 277–294. https://doi.org/10.1002/wat2.1023.

Isaak, D. J., S. J. Wenger, E. E. Peterson, et al. 2017. "The NorWeST Summer Stream Temperature Model and Scenarios for the Western U.S.: A Crowd-Sourced Database and New Geospatial Tools Foster a User Community and Predict Broad Climate Warming of Rivers and Streams." *Water Resources Research* 53, no. 11: 9181–9205. https://doi.org/10.1002/2017WR020969.

Justice, C., S. M. White, D. A. McCullough, D. S. Graves, and M. R. Blanchard. 2017. "Can Stream and Riparian Restoration Offset Climate Change Impacts to Salmon Populations?" *Journal of Environmental Management* 188: 212e227.

Kanno, Y., B. H. Letcher, N. P. Hitt, D. A. Boughton, J. E. B. Wofford, and E. F. Zipkin. 2015. "Seasonal Weather Patterns Drive Population Vital Rates and Persistence in a Stream Fish." *Global Change Biology* 21: 1856–1870. https://doi.org/10.1111/gcb.12837.

Kennedy, J. T., and J. Butcher. 2012. "South Fork Nooksack River Temperature Total Maximum Daily Load Water Quality Study Design (Quality Assurance Project Plan)." Olympia, WA. https://apps.ecology.wa.gov/publications/SummaryPages/1203126.html.

Krause, C. W., B. Lockard, T. J. Newcomb, D. Kibler, V. Lohani, and D. J. Orth. 2004. "Predicting Influences of Urban Development on Thermal Habitat in a Warm Water Stream." *Journal of the American Water Resources Association* 40: 1645–1648. https://doi.org/10.1111/j.1752-1688.2004.tb01612.x.

Lee, S.-Y., A. H. Fullerton, N. Sun, and C. E. Torgersen. 2020. "Projecting Spatiotemporally Explicit Effects of Climate Change on Stream Temperature: A Model Comparison and Implications for Coldwater Fishes." *Journal of Hydrology* 588: 125066. https://doi.org/10.1016/j.jhydrol.2020.125066.

Liu, N., G. R. Dobbs, P. V. Caldwell, et al. 2022. "Inter-Basin Transfers Extend the Benefits of Water From Forests to Population Centers Across the Conterminous U.S." *Water Resources Research* 58, no. 5: e2021WR031537. https://doi.org/10.1029/2021WR031537.

Middle Fork IMW Working Group. 2017. "Middle Fork John Day River intensively monitored watershed final summary Report".

Pelletier, G. 2002. "Wind River Watershed Temperature Total Maximum Daily Load." Olympia, WA. https://apps.ecology.wa.gov/publications/documents/0203010.pdf.

Pelletier, G., and S. Chapra. 2008. "QUAL2Kw: A Modeling Framework for Simulating River and Stream Water Quality." Version 5.1: Documentation and Users Manual.

Pervez, M. S., and J. F. Brown. 2010. "Mapping Irrigated Lands at 250-m Scale by Merging MODIS Data and National Agricultural Statistics." *Remote Sensing* 2, no. 10: 2388–2412. https://doi.org/10.3390/rs2102388.

Peterson, E. E., and J. M. Ver Hoef. 2010. "A Mixed-Model Moving-Average Approach to Geostatistical Modeling in Stream Networks." *Ecology* 91, no. 3: 644–651. https://doi.org/10.1890/08-1668.1.

Peterson, G. D., G. S. Cumming, and S. R. Carpenter. 2003. "Scenario Planning: A Tool for Conservation in an Uncertain World." *Conservation Biology* 17, no. 2: 358–366. https://doi.org/10.1046/j.1523-1739.2003.01491.x.

Poole, G., and C. Berman. 2001. "An Ecological Perspective on in-Stream Temperature: Natural Heat Dynamics and Mechanisms of Human-Caused Thermal Degradation." *Environmental Management* 27: 787–802. https://doi.org/10.1007/s002670010188.

R Core Team, R. Bivand, V. J. Carey, et al. 2022. "Foreign." https://cran.r-project.org/web/packages/foreign/foreign.pdf.

R Core Team. 2020. "R: A Language and Environment for Statistical Computing." In *R Foundation for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing. https://doi.org/10.1007/978-3-540-74686-7.

Seixas, G. B., T. J. Beechie, C. Fogel, and P. M. Kiffney. 2018. "Historical and Future Stream Temperature Change Predicted by a Lidar-Based Assessment of Riparian Condition and Channel Width." *Journal of the American Water Resources Association (JAWRA)* 54, no. 4: 974–991. https://doi.org/10.1111/1752-1688.12655.

Siegel, J. E., and C. J. Volk. 2019. "Accurate Spatiotemporal Predictions of Daily Stream Temperature From Statistical Models Accounting for Interactions Between Climate and Landscape." *PeerJ* 7: e7892.

Steel, A. E., C. Sowder, and E. E. Peterson. 2016. "Spatial and Temporal Variation of Water Temperature Regimes on the Snoqualmie River Network." *Journal of the American Water Resources Association* 52, no. 3: 769–787. https://doi.org/10.1111/1752-1688.12423.

Steel, E. A., T. J. Beechie, C. E. Torgersen, and A. H. Fullerton. 2017. "Envisioning, Quantifying, and Managing Thermal Regimes on River Networks." *Bioscience* 67, no. 6: 506–522. https://doi.org/10.1093/biosci/bix047.

Steel, E. A., A. Marsha, A. H. Fullerton, et al. 2018. "Thermal Landscapes in a Changing Climate: Biological Implications of Water Temperature Patterns in an Extreme Year." *Canadian Journal of Fisheries and Aquatic Sciences* 76, no. 10: 1740–1756. https://doi.org/10.1139/cjfas-2018-0244.

U.S. Geological Survey. 2012. "National Hydrography Dataset Plus—NHDPlus Version 2.1 for Hydrologic Unit (HU) 17." Washington, DC. http://www.horizon-systems.com/NHDPlus/index.php.

Ver Hoef, J., E. Peterson, D. Clifford, and R. Shah. 2014. "SSN: An R Package for Spatial Statistical Modeling on Stream Networks." *Journal of Statistical Software* 56, no. 3: 1–45. https://doi.org/10.18637/jss.v056.i03.

Weston, S. 2017. "foreach: Provides foreach Looping Construct for R." R Package Version 1.4.4. http://cran.r-project.org/package=foreach.

Weston, S., and R. Calaway. 2014. "Getting Started With doParallel and foreach." R Package Version, 1.0.8.

White, S. M., C. Justice, D. A. Kelsey, D. A. McCullough, and T. Smith. 2017. "Legacies of Stream Channel Modification Revealed Using General Land Office Surveys, With Implications for Water Temperature and Aquatic Life." *Elementa: Science of the Anthropocene* 5: 3. https://doi.org/10.1525/elementa.192.

Whittingham, M. J., P. A. Stephens, R. B. Bradbury, and R. P. Freckleton. 2006. "Why Do We Still Use Stepwise Modelling in Ecology and Behaviour?" *Journal of Animal Ecology* 75, no. 5: 1182–1189. https://doi.org/10.1111/j.1365-2656.2006.01141.x.

Wickham, H., M. Averick, J. Bryan, et al. 2019. "Welcome to the Tidyverse." *Journal of Open Source Software* 4, no. 43: 1686. https://doi.org/10.21105/joss.01686.

Yang, L., S. Jin, P. Danielson, et al. 2018. "A New Generation of the United States National Land Cover Database: Requirements, Research Priorities, Design, and Implementation Strategies." *ISPRS Journal of Photogrammetry and Remote Sensing* 146: 108–123. https://doi.org/10.1016/j.isprsjprs.2018.09.006.

Supporting Information

Additional supporting information can be found online in the Supporting Information section.